Far-Infrared Spectroscopy on Proteins: Bringing together Experiments, Simulations, and Theory

Von der Fakultät Mathematik und Physik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Christian Ulrich Stehle

aus Stuttgart

Hauptberichter:	**Prof. Dr. Martin Dressel**
Mitberichter:	**Prof. Dr. Jörg Wrachtrup**
Tag der mündlichen Prüfung:	**18. Juni 2012**

1. Physikalisches Institut der Universität Stuttgart

2012

Table of Contents

Kurzfassung

Diese Doktorarbeit befasst sich mit Experimenten, Simulationen und theoretischen Fragestellungen zur Ferninfrarotspektroskopie an Proteinen. In diesem Frequenzbereich werden für die Funktionalität von Proteinen wichtige kollektive Bewegungen erwartet. Trotzdem fehlten bisher reproduzierbare Ferninfrarotspektren unter definierten Bedingungen, sowie Messungen zur Abhängigkeit bzgl. der wichtigen Parameter Temperatur und Hydrierung. Gründe hierfür sind vor allem experimentelle Herausforderungen im Ferninfrarotbereich, die es zu überwinden gilt wie z. B. das Fehlen von starken Lichtquellen zusammen mit dem Bedarf für sehr sensitive, in unserem Fall heliumgekühlte Detektoren. Außerdem müssen die biologischen Proben so präpariert werden, dass deren natürlicher Zustand und damit die Funktionalität erhalten bleibt, eine signifikante Absorption gewährleistet ist, aber trotzdem noch eine ausreichende Transmission für die Messung vorhanden ist.

Für den experimentellen Teil, dem Schwerpunkt dieser Arbeit, wurden insgesamt fünf Messaufbauten konzipiert, die es erlauben Proteine unter einen Infrarotmikroskop zu messen, welches als Quelle ein Fourier-Transform-Spektrometer nutzt. Mit dem wesentlichen Aufbau, der sogenennten „Atmosphärenzelle“, sind Messungen im Frequenzbereich von ca. 40 cm^{-1} bis 690 cm^{-1} möglich, was ungefähr 1 bis 20 THz entspricht. Damit reicht der Messbereich über den Ferninfrarot- in den Terahertz-Frequenzbereich hinein. Der Messaufbau erlaubt es außerdem die beiden wichtigen Parameter Temperatur und Hydrierung zu variieren. Diese beeinflussen die Proteindynamik, so dass Auswirkungen auf das Ferninfrarotspektrum von Proteinen zu erwarten sind, was sich auch in den Messergebnissen bestätigte. Um temperatur- und feuchtigkeitsabhängige Messungen zu realisieren wurde eine eigene Präparationsmethode entwickelt, die es erlaubt freitragende Proteinfilme herzustellen. Somit konnte auf das sonst übliche Trägermaterial verzichtet werden, welches insbesondere die externe Hydrierung der Filme erschwert bzw. behindert. Zusätzlich bieten die freitragenden Filme die Möglichkeit die Proteinmenge auf ein optimales Signal-Rausch-Verhältnis für die technisch schwierigen Messungen im Ferninfrarotbereich einzustellen. Grundlage dieser Probenpräparation ist die Herstellung eines Films mittels Lyophilisation auf Polypropylen, welches einen ähnlichen thermischen Ausdehnungskoeffizienten besitzt wie der Proteinfilm selbst. Somit werden bei den für die Herstellung notwendigen Temperaturänderungen Risse im Film verhindert. Das hydrophobe Trägermaterial erleichtert außerdem eine Ablösung des Proteinfilms nach

Abschluss des Lyophilisationsprozesses, so dass anschließend freitragende Filme in die Atmosphärenzelle für die eigentliche Messung übertragen werden können. Mittels eines Peltier-Elements und eines PT-100 Temperatursensors kann die Temperatur in der Atmosphärenzelle automatisiert auf 0,1 K genau eingestellt und gehalten werden. Der für Proteine physiologisch relevante Temperaturbereich von 4°C bis 55°C wird vollständig abgedeckt. Die externe Hydrierung der Filme erfolgt über gesättigte Salzlösungen. Über diesen stellt sich eine so konstante relative Luftfeuchtigkeit ein, dass diese Methode auch für die Kalibrierung von Feutigkeitsensoren verwendet wird. Mittels verschiedener Salzlösungen kann die relative Luftfeuchtigkeit und damit die Hydrierung des Proteinfilms in einem Bereich von 33% bis 97% relativer Luftfeuchtigkeit mit einer Stabilität von besser als 1% relativer Luftfeuchtigkeit variiert werden. Über die sogenannte Sorption-Isotherm-Gleichung kann die Hydrierung des Proteins quantitativ beschrieben werden. Unter Ausnutzung der Verhältnisse zwischen unseren bekannten Parametern der Probenpräparation und dem Messbereich unter dem Mikroskop konnten wir sogar erstmals direkt auf die Anzahl der Wassermoleküle und deren Zu- und Abnahme bei Veränderung der relativen Luftfeuchtigkeit schließen. Eine solche externe Hydrierung reicht aus eine annähernd natürliche Umgebung eines Proteins in wässriger Lösung zu simulieren. Beweis dafür ist das Einsetzen der enzymatischen Aktivität verschiedener Proteine bereits bei sehr geringer externer Hydrierung, die die für die Funktionalität notwendige Proteindynamik auslösen. Somit kann das Protein nahezu in seinem natürlichen Zustand gemessen werden und die starke Absorption von Wasser im Ferninfrarotbereich umgangen werden. Das Dreikomponentensystem aus Protein, biologischen Wasser und Wasser bei einer Messung in wässriger Lösung kann mit unserem Messaufbau auf die ersten beiden Komponenten reduziert werden, was die Interpretation der Ergebnisse erleichtert. Biologisches Wasser werden Wassermoleküle genannt, deren Eigenschaften durch die Bindung oder die Nähe zum Protein gegenüber freien Wasser abweichen. Dies betrifft z. B. translatorische und rotatorische Freiheitsgrade einzelner Wassermoleküle, sowie das Wasserstoffbrückennetzwerk und somit auch die spektroskopischen Eigenschaften dieser Moleküle im Ferninfrarotbereich.

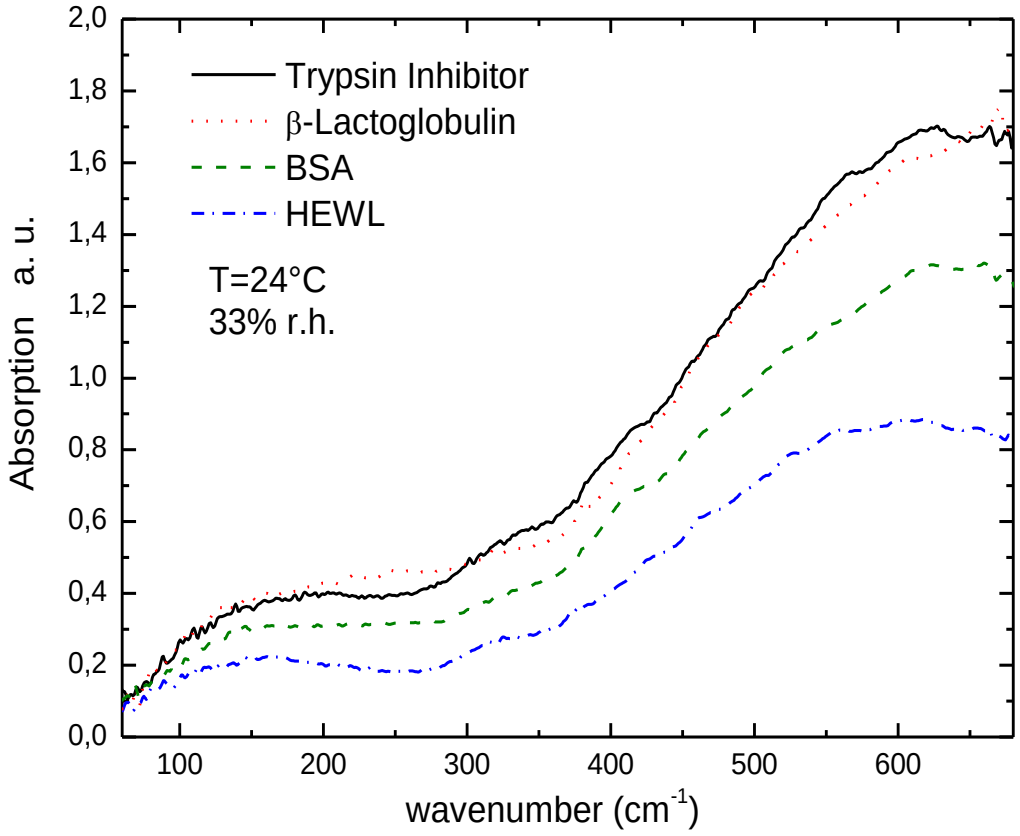

Abb. 1: Ferninfrarotabsorptionspektren der vier gemessenen Proteine bei einer Temperatur von 24°C und 33% relativer Luftfeuchtigkeit. Im Gegensatz zum Terahertz-Frequenzbereich lassen sich die vier Proteine anhand mehrerer Absorptionsbanden deutlich voneinander unterscheiden.

Gemessen wurden vier Proteine Bovine Serum Albumin (BSA), Hen Egg White Lysozyme (HEWL), β-Lactoglobulin und Trypsin Inhibitor. Teilweise sind dies die ersten publizierten Ferninfrarotmessungen an diesen Proteinen, jedoch sind es für alle Proteine die ersten Messungen unter kontrollierten Bedingungen bezüglich Temperatur und Hydrierung. Außerdem wurde die Temperatur- und Feuchtigkeitsabhängigkeit der Proteinspektren im Ferninfrarotbereich untersucht. Wie in Abbildung 1 zu sehen ist, lassen sich die vier Proteine anhand ihrer Ferninfrarotspektren unterscheiden und die mittels eines Fits identifizierten Absorptionsbanden können für Vergleiche zu Simulationsergebnissen herangezogen werden. In der Analyse wurden Hinweise darauf gefunden, dass im Ferninfrarot sowohl allgemeine als auch proteinspezifische Absorptionsbanden auftreten. In den angrenzenden Frequenzbereichen des mittleren Infrarot und des Terahertzbereichs ist hingegen eine Proteinunterscheidung anhand der Spektren kaum bzw. nicht möglich. Im mittleren Infrarot

werden die in allen Proteinen vorkommenden kovalenten Bindungen gemessen. Für eine Unterscheidung zwischen Proteinen müssen kleinste Abweichungen in Frequenz, Höhe und Form der Absorptionsbanden herangezogen werden, die durch die sekundäre und tertiäre Struktur und das dynamische Verhalten von Proteinen verursacht werden. Im Terahertzbereich zeigen alle Proteine einen fast linear ansteigenden Absorptionsverlauf. Bei Kenntnis des Ferninfrarotspektrums von Proteinen ist dies als tieffrequente Flanke einer breiten Absorptionsbande zwischen 100 cm^{-1} und 200 cm^{-1} erklärbar. Eine Unterscheidung zwischen Proteinen ist aufgrund der fehlenden Absorptionsbanden im Terahertzbereich bisher nicht möglich. Zusätzlich zu den eigentlichen Messungen wurden weitreichende Untersuchungen zur Reproduzierbarkeit der Messergebnisse vorgenommen indem Spektren unterschiedlicher Proben aus verschiedenen und gleichen Präparationsläufen, sowie verschiedene Probenmesspunkte auf einer Probe miteinander verglichen wurden.

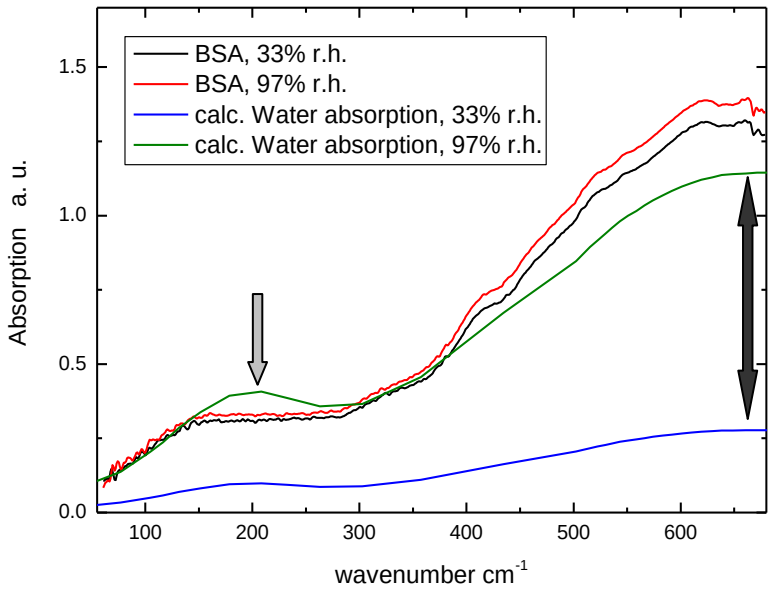

Abb. 2: Vergleich der feuchtigkeitsabhängigen Veränderung des Ferninfrarotspektrums von Bovine Serum Albumin (BSA) mit einer kalkulierten Absorption des im Messbereich vorhandenen Wassers bei Annahme einer Absorption von flüssigem Wasser. Bei 97% relativer Luftfeuchtigkeit würde allein die Absorption der Wassermoleküle die gemessene Absorption des Protein-Wasser-Komplexes übersteigen. Außerdem müsste die relative Änderung der Absorption mindestens eine Größenordnung größer als beobachtet sein. Beide Argumente lassen auf eine deutlich geringere Ferninfrarotabsorption von biologischem Wasser schließen, welches durch das Protein beeinflusst wird und somit andere (dynamische) Eigenschaften gegenüber flüssigem Wasser zeigt.

Eine Variation der externen Hydrierung der Proteine hat nur einen geringen Effekt auf das Ferninfrarotspektrum der Proteine und zeigt, wie in Abbildung 2 zu sehen ist, nur eine kleine, aber reproduzierbare frequenzunabhängige Zunahme der Absorption mit steigender Hydrierung. Dies ist auf den ersten Blick überraschend, da Wasser im Ferninfrarot stark absorbierend ist und eine große zusätzliche Anzahl Wassermoleküle hinzukommt und die Absorption stark zunehmen müsste, wie der Vergleich mit einer kalkulierten Wasserabsorption in Abbildung 2 zeigt. Da jedoch biologisches Wasser hinzukommt, welches durch das Protein in seinen Eigenschaften beeinflusst wird, konnte aus dem quantitativen Vergleich geschlossen werden, dass diese Wassermoleküle eine deutlich geringere Ferninfrarotabsorption zeigen als flüssiges Wasser. Ein Vergleich zum Ferninfrarotspektrum von Eis legt nahe, dass die eingeschränkte Dynamik der biologischen Wassermoleküle eher eisähnlich ist. Somit unterstützen unsere Messungen das sogenannte „Eisberg“-Model für biologisches Wasser. Somit konnten wir ein bisher auf Einzelmolekülebene beschriebenes Modell basierend auf Simulationen und Strukturanalysen mit der makroskopischen Größe Absorption verknüpfen, die eine Mittelung über das dynamische Verhalten von vielen Molekülen darstellt.

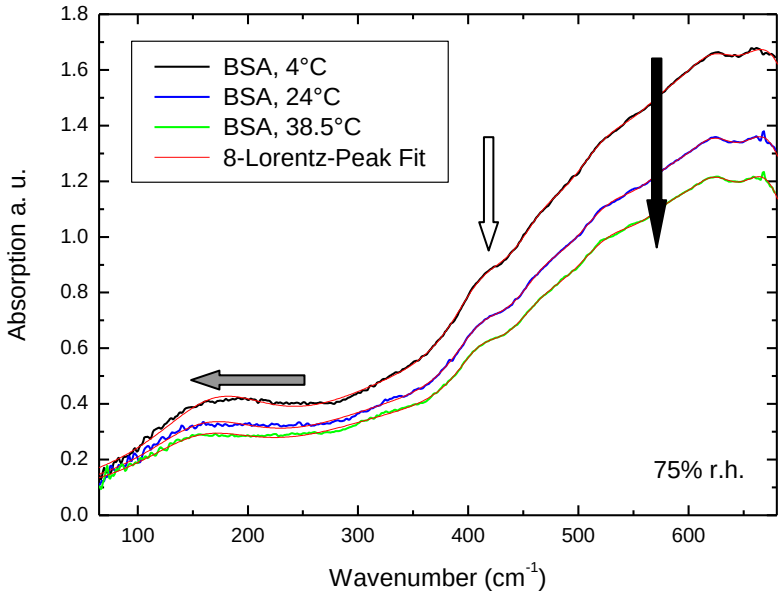

Abb. 3: Ferninfrarotspektrum von Bovine Serum Albumin (BSA) unter einer Variation der Temperatur. Als größten Effekt konnte eine frequenzunabhängige Abnahme der Absorption mit steigender Temperatur ausgemacht werden (schwarzer Pfeil). Für eine Diskussion der Messergebnisse wurde der gesamte Temperatureffekt in drei unabhängige Temperatureffekte aufgeteilt. Somit konnten zudem zwei frequenzabhängige Effekte identifiziert werden. Die roten Graphen stellen jeweils Lorentz-Fits der Proteinspektren dar.

Temperaturänderungen hingegen haben einen deutlich größeren Einfluss auf die Ferninfrarotspektren als eine Änderung der Hydrierung. Zur Diskussion wurde der Temperatureffekt wie in Abbildung 3 angezeigt in drei verschiedene Effekte aufgeteilt. Die frequenzunabhängige starke Abnahme der Absorption mit steigender Temperatur wird von frequenzabhängigen Effekten überlagert, für deren Ursache wir im unteren Frequenzbereich eine Erklärung nahelegen, die nicht für den frequenzabhängigen Effekt bei höheren Frequenzen gelten kann und daher separat diskutiert wird.

Für den frequenzabhängigen Temperatureffekt im unteren Frequenzbereich (grauer Pfeil in Abbildung 3) konnte für alle Proteine außer β-Lactoglobulin eine Verschiebung zu tieferen Frequenzen mit steigender Temperatur beoachtet werden. In diesem Frequenzbereich sind höhere besetzte Zustände zu berücksichtigen, da Raumtemperatur von 24°C mit einer Frequenz von 206, 5 cm^{-1} korrespondiert. Mit steigender Temperatur nimmt die Besetzung höherer Zustände zu und somit werden auch Übergänge zwischen höheren Zuständen häufiger, was zu einer Verschiebung der Absorption zu niedrigeren Frequenzen führt. Für das abweichende Verhalten von β-Lactoglobulin, welches als einziges Protein zwei Absorptionsbanden in diesem Frequenzbereich zeigt, könnte möglicherweise der höhere β-Sheet-Anteil verantwortlich sein. Im höheren Frequenzbereich (weißer Pfeil in Abbildung 3) zeigen alle Proteine Frequenzbereiche mit starker und mit geringer Temperaturabhängigkeit. Letztere überlappen mit für alle gemessenen Proteine gefundenen Absorptionsbanden, so dass ein Zusammenhang mit dem allen gemeinsamen Grundgerüst eines jeden Proteins nahe liegt.

Der frequenzunabhängige Temperatureffekt zeigt einen deutlichen Absorptionsrückgang mit steigender Temperatur für alle Proteine. Mögliche Erklärungen hierfür haben wir in unserem Model of Energy Related Correlation Effects in Proteins (MERCEP) zusammengefasst, welches zur Klärung spezifischer Fragestellungen mit molekulardynamischen Simulationen untermauert wurde. Wir glauben, dass die steigende Temperatur Korrelationen innerhalb des Proteins reduziert, welche die für die Infrarotabsorption relevanten Dipolmomente sowie kollektive Bewegungen betrifft. Insbesondere induzierte Dipolmomente zwischen nichtgebundenen Atomen in einem typischen Van-der-Waals-Abstand zueinander werden durch eine Erhöhung des mittleren Abstands geringer. Zusätzlich nimmt die Korrelation innerhalb der kollektiven Proteinbewegungen ab, da diese durch Van-der-Waals-Bindungen und Wasserstoffbrücken gekoppelt sind. Beide Bindungsarten werden durch die Temperaturänderung beeinflusst. Erste durch die erwähnte Abstandsänderung, für letztere konnten wir in Simulationen zeigen, dass zumindest die Lebensdauer der Wasserstoffbrücken bei einem Temperaturanstieg abnimmt. Die Anzahl der gebildeten Wasserstoffbrücken nimmt

sogar zu, dies weist aber in Summe auf eine höhere, temperaturinduzierte Proteindynamik hin. Wasserstoffbrücken werden kürzer ausgebildet und aufgrund der höheren Dynamik stehen in Summe andere und somit mehr Bildungsmöglichkeiten von Wasserstoffbrücken zur Verfügung. Insgesamt ist die Bewegung der Proteinatome untereinander weniger korreliert und die kollektiven Proteinbewegungen mit einer Absorption im Ferninfrarotfrequenzbereich werden geringer. Diese Reduktion der Anzahl der kollektiven Proteinbewegungen zusammen mit dem Rückgang der für die Infrarotabsorption relevanten Dipolmomente führt zu einem generellen Absorptionsrückgang mit steigender Temperatur in den Ferninfrarotspektren der Proteine.

Daneben ist in dieser Arbeit ein Review über die bisher publizierten Ferninfrarotspektren von Proteinen enthalten, sowie Reviews über relevante Messungen im Terahertz- und mittleren Infrarotfrequenzbereich. Außerdem wurden die Ergebnisse der Messungen und Simulationen in eine umfangreiche Betrachtung verschiedener theoretischer Themen eingebettet, die im Kontext dieser Arbeit behandelt wurden. So haben wir beispielsweise den in der Literatur bestehenden Dissenz bzgl. der Existenz eines Bosonpeaks in Proteinspektren aufgegriffen. Proteine zeigen zwar eine breite Absorptionsbande in einem ähnlichen Frequenzbereich wie der Bosonpeak in Gläsern, jedoch sprechen die Temperatur- und Feuchtigkeitsabhängigkeit, sowie die Existenz bei Temperaturen oberhalb des dynamischen Übergangs gegen die gleiche physikalische Ursache dieser Absorptionsbande. Der dynamische Übergang in Proteinen wird gerne auch als Glasübergang bezeichnet, jedoch zeigt der Verlauf der Wärmekapazität einen eindeutig anderen Verlauf als bei Gläsern, so dass auch dieser Begriff für Proteine nicht verwandt werden sollte. Außerdem ist in dieser Arbeit ein Weg beschrieben Ferninfrarot-spektren von Proteinen auf Basis von molekulardynamischen Simulationen zu berechnen. Dazu wurde auf Kubos quantenmechanische Formulierung des Dissipation-Fluktuation-Theorems und die Dipolmoment-Korrelationsfunktion zurückgegriffen. Wir hoffen damit korrespondierenden Simulationen zu unseren hier präsentierten Ferninfrarotmessungen an Proteinen den Weg zu bereiten.

1. Introduction

What is so interesting about far-infrared properties of proteins? A legitimate question after having spent some years of reseach in this field. The most condensed answer is importance, interest and challenge. Importance, since proteins are an essential component for all living organisms and the far-infrared is one of the most promising frequency ranges to gain information about functionally relevant motions which will lead to potent applications if their real nature can be once revealed. Interest, because the here treated topic is at the edge of physics, chemistry and biology requireing also sound mathematical understanding for the presented theoretical derivations, all in all fitting very well to my three different subjects of study which were finished prior and meanwhile this PhD thesis. There are no limits for interest and especially a topic with touchpoints to many other different areas of research is capable to satisfy a scientist's thirst for knowledge. But this unconfined field is also a challenge, directly passing over to the third reason, since much more research will be necessary for a complete understanding of proteins concerning all three aspects of this thesis experiments, simulations, and theory. For each of them specific challenges have been met. The experimental challenge was the developement of an own setup for the sample arrangement meeting particular requirements and enabling reproducible measurements in the far-infrared range, one of the crucial obstacles to overcome when probing biological matter. Unavailable powerful light sources and therefore the necessity of sensitive, usually helium cooled, detectors, representing an arrangement which requires additionally an affiliated sample preparation which considers carefully adequate absorption as well as sufficient transmission, to mention at least some of the experimental reasons for the lack of thorough protein research in this frequency range in the past. Due to the complexity of proteins, simulations are a powerful tool to get further insight into microscopic consequences, for instance if external parameters are changed as applied in our experiments. Here the challenge was to get accompanying simulations to our prior theoretically derived hypotheses and then finally to match them with the interpretation of our experimental results. Another simulations concerning topic leads directly over to the theoretical challenges in this field. Simulations which model the far-infrared absorption are not available since even less complex water is hardly matching the actual measurements properly up to now. Even the way to calculate such far-infrared spectra is probably just known to a very small group of people. Hence a lot of basic theoretical work on how the requirements of an experimental physicist can be matched

with the capabilities of simulation researchers had to be done, representing a challenge in the combined field of experiments, simulations and theory. Since protein research can be approached from physics, chemistry, biology, and medicine either on an experimental or on a theoretical way some confusions, misconceptions and disagreements are present in the current scientific discussion often caused by a focussation each time on the own well known field of research. To get familiar with adjacent research areas has been accordingly the principal theoretical challenge but afterwards combination of this knowledge allowed us to derive several conclusions clarifying some of the above mentioned controversies. Therefore the upcoming theory chapter includes already such “theoretical results” and provides also the basis for the discussion of our experiments presented in chapter 4 to which a setup and sample preparation details concerning section was placed in front.

2. The Theory of Far-Infrared Spectroscopy on Proteins and Affiliated Topics

All Proteins are made of the same basic amino acid elements, but the developing three dimensional structure is rich in varieties. An overview over structure, different types of bonds and the role of proteins for living organisms is given at the beginning of this chapter. The timescale of protein dynamics spans many orders of magnitude from femtoseconds to seconds and these dynamics are essential for the functionality. Infrared spectroscopy is one possibility to investigate motions theoretically predicted to be related to functionality and these kinds of motions will be in focus of this particular section. But the interpretation of the gained experimental data is aggravated by the complexity of the sample, namely the protein macromolecule. Especially molecular dynamics simulation is a very promising technique to model proteins and is widely used for calculations of the structure factor in X-ray experiments, meanwhile with a well conformity. Unfortunately this method still has to overcome some obstacles for calculating infrared spectra even of water, far less complicated than the demanded protein-water systems. Only one close to the experimental data calculated water spectrum was published in 1996 [Sou96]. To date a second group claimed by personal communication to have modelled a far-infrared spectrum of liquid water properly but no spectrum was published so far. As it can be assumed that just few MD Simulation groups are capable of calculating protein infrared spectra, the theory is presented here in a comprehensive way. This theoretical chapter will be closed by a discussion of the wide field of protein hydration and in particular the procedure of calculating the water content of proteins out of external macroscopic values, as it is used in our experiments.
Besides presenting here the theoretical basis for our measurements we derive also “theoretical results” as already mentioned above to explain the subtitle of this thesis. As such we understand our own extension of several models and formulas as well as new conclusions that we draw from a synopsis of theoretical discussions. Some more results, already included in the theoretical chapter, will lead to new findings if their derived arguments are joint later with the experimental data. That there is also a large theoretical outcome within this experimentally focused thesis, is the reason for such a comprehensive presentation of theory as given here.

2.1. Proteins – Structure, Bond Types, and Folding

Even made out of just 20 different natural occurring amino acids the variety of proteins is manifold as well as the various functions in living organisms. The importance of protein research can be understood by recapturing essential processes where proteins are indispensable. Proteins are involved in the storage and transport of energy, electrons, and small molecules [Nie96]. The most famous example is hemoglobin for the O_2 and CO transport. Almost all enzymes are proteins [Jac06] catalyzing processes in living organisms. Usually specific chemical reactions would take place at a much slower rate but enzymes can accelerate these reactions by orders of magnitude typically around 10^8 to 10^{12} with a record of about 10^{20} up to date [Jac06]. Exemplary serine proteases like trypsin trigger the hydrolysis, basically the breaking, of peptide bonds to digest proteins into smaller peptides and amino acids. A good summary of protein functions is found in [Les04] covering, besides the two already named tasks, mainly protein functions in tissue structures and also as antibodies. An example for the first group of proteins, emphasizing the structure relevance, is collagen, which is the most abundant protein in mammals. The latter group is one important reason why protein research is pushed forward in Life Science companies. Another class of proteins acts as hormones. For instance insulin maintains the blood sugar at appropriate levels and thisprotein was one of the first being produced by pharmaceutical industry on a larger scale for Diabetes patients.

The essential role of proteins for understanding living organisms becomes apparent not only by the abundance of different types and functions but also by the fact that about 50% of the dry mass of a living cell consists of proteins [Tus07]. To understand protein related diseases and curing them, is another driving force for protein research. Best known are the protein aggregation diseases like Alzheimer or the Parkinson disease. Senile plaques, which are dense insoluble extracellular protein deposits, indicate the first and so called Levy bodies the latter, which are inclusions in brain cells containing protein aggregates [Les04]. That proteins can also be the transmitting agent of diseases became apparent even to nonscientists in 2001 as a serious epidemic spread out from the United Kingdom called Bovine Spongiform Encephalopathy (BSE) or the so called “mad cow” disease. This was associated with the occurrence of a new form of the Creutzfeldt-Jakob disease in humans giving rise to a transferability of the disease from cattle to humans. People got aware that prions (proteinaceous infectious particles) can cross specie barriers since it is assumed that BSE originates from the Scrapie disease in sheep. A lot of preventation action and thorough testing

is ongoing since that time including focused investigations on proteins. At the moment it is believed that a particular protein named PrP, which role is not completely revealed yet, exists in two forms differing in their conformation as seen in fig. 2.1. The infectious version can form aggregates which can block functions in the nervous system. It can also catalyze natural proteins to the infectious form leading to an outspread of the disease or an infection of other individuals [Les04].

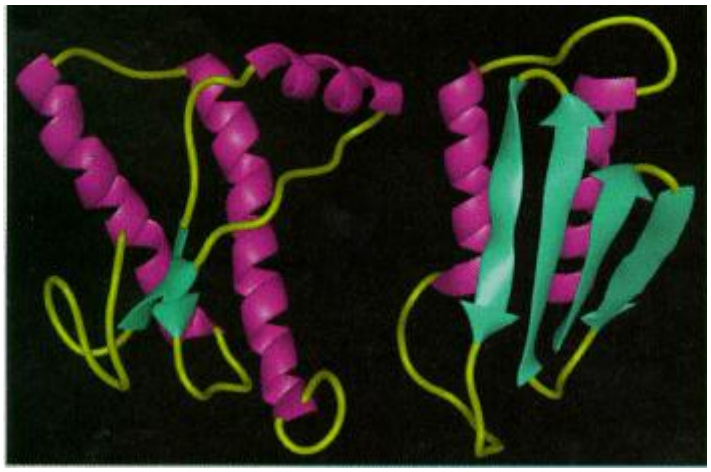

Fig. 2.1: The natural and infectious form of PrP causing the disease Bovine Spongform Encephalopathy (BSE) [Wik08].

Even beyond understanding proteins and their caused diseases a long term goal of research in this field is protein design. Especially pharmaceutical industry seeks for adverse effect free medication through the highly specific functionality of proteins, making protein research a promising field for many applications in the future.

2.1.1. The Structure of Proteins

As we have seen for prion diseases the structure of proteins is crucial for their functionality, or in this case, for their disfunctionality. The structure-function hypothesis claims that the present protein structure is the result of a million years ongoing evolutional selection process which suits perfectly the certain biological function in that particular organism. Understanding the structure of a protein will lead to a further understanding of the functionality. This also called Anfinsen Dogma [Anf73], that the three dimensional structure of a protein is determined solely by its amino acid sequence, is widely acknowledged in a

nonstrict form due to easy experiments. Loosing structure results in loosing functionality, restoring the structure restores protein functionality [Kar06]. Just in the last two sentences the two major open questions in protein research are summarized as visualized in fig. 2.2. How can the protein structure be predicted out of the primary sequence alone? And how can the protein structure and its dynamics be related to the functionality? With this thesis we will contribute to answering the latter.

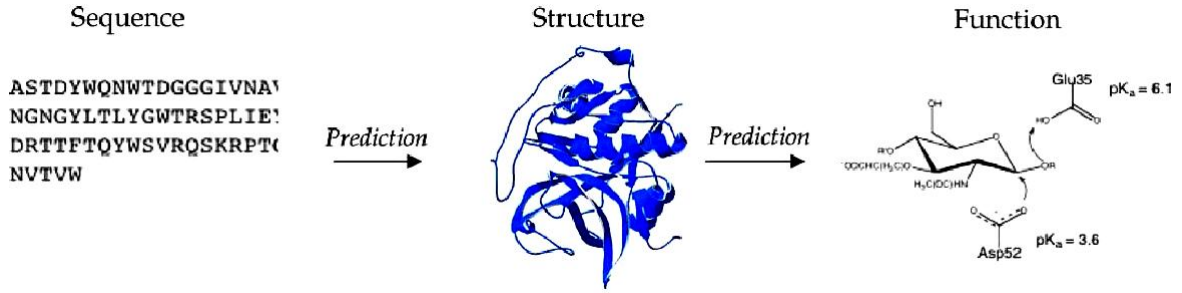

Fig. 2.2: The two major open questions in protein research: How can the protein structure be predicted by the knowledge of the primary sequence? How can the functionality be predicted by an analysis of the protein structure and dynamics [Col06]?

Starting with the amino acid chain sequence, the different forms of protein organization are referred to the primary up to the quarterny structure which will be discussed in detail in this section, regarding topics needed in later chapters.

2.1.1.1. The Primary Structure of Proteins

Encoded in the DNA (or RNA for some species) the basic sequence of proteins is translated into a messenger RNA and transcripted by the transfer RNA tripletts (codons) into an amino acid sequence at the ribosomes of the rough endoplasmatic reticulum. The amino acids are linked by a peptide bond which is formed between the carboxyl group and the amino group of two amino acids under formation of one water molecule (fig. 2.3). The sequence of the 20 natural occurring α-L-amino acids is called the primary structure of a protein which is known for more than 100,000 proteins [Ber03]. This large number is owed to the development of the automatized Edman degradation allowing a fast analysis of every protein primary structure.

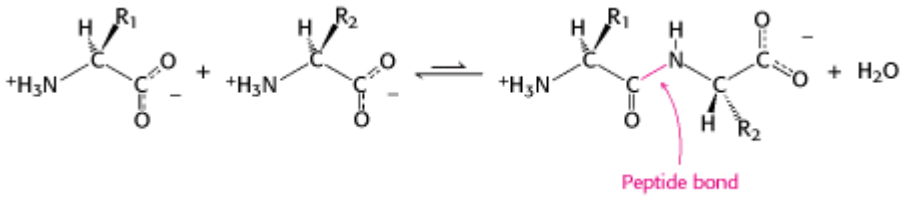

Fig. 2.3: Formation of a peptide bond between two amino acids and an additional water molecule [Ber03].

The non polar peptide bond is rather unable to rotate which arises from a slight double bond character. As it can be seen from fig. 2.4 the length of the peptide bond is typically 1.325 Å, confirmed by X-ray structure analysis and quantum mechanical calculations. This is much closer to the double bond length C=N of about 1.25 Å than to the single bond C-N length of 1.445 Å [Dau97]. This results in a planar structure of six different atoms from the left (black) C_α atom via C_β, O, N, and H to the second C_α atom of the next amino acid. In contradiction to that the two single bonds around the C_α atom are able to rotate freely, restricted only by steric considerations (fig. 2.4). This is the major reason for the many different possible protein conformations of which many are explored in the so called rough energy surface at physiological temperatures (see chapter 2.2.2).

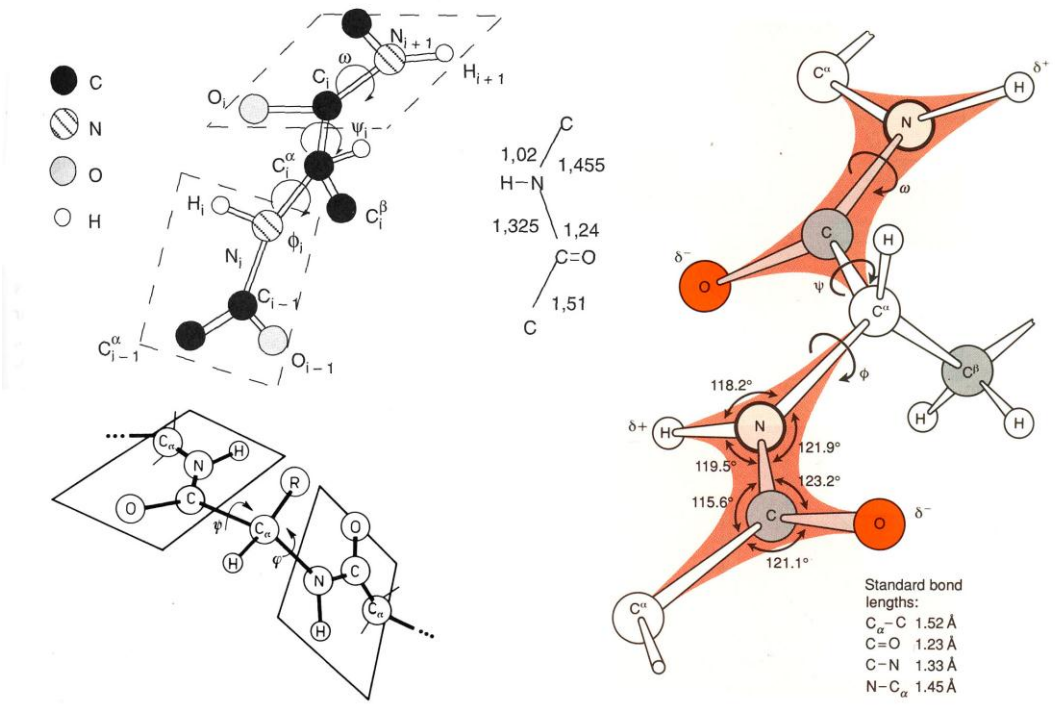

Fig. 2.4: Bond lengths and angles of the peptide bond. The resulting planar structure is indicated [Ber03, Dar93, Twa94, Dau97].

All amino acids just differ in side groups which can be classified in different ways. Nonpolar, polar and charged side groups are found as well as side groups that can consist of a simple H atom up to aromatic rings or even containing sulfur atoms. For this thesis the properties concerning the interaction with water molecules is decisive which are well presented for every single kind of side group in [Ser06].

2.1.1.2. The Secondary Structure of Proteins

Internal hydrogen bonds are the major component for the formation of two groups of secondary structure elements in proteins, namely helical and sheet type structures. The peptide chain which is not organized in one of these two forms is called random coil. Here a lot of the possible protein conformations occur due to the large degree of freedom especially of the two single bonds around the C_α atom of every amino acid, which are able to rotate as discussed in the previous section. Also the interaction with surrounding water molecules is stronger for the random coil compared to the denser packed helices and sheets.
In the α-helix secondary structure every amino acid C=O group forms a hydrogen bond with the amino acid group N-H usually situated four residues ahead as seen in fig. 2.5 [Ber03]. Essentially all occurring α-helices in proteins are right-handed, which is slightly energetically more favorable than the left-handed version because of steric reasons between the side groups and the backbone. In right-handed α-helices all amino acid side groups point outwards with 3.6 amino acids per turn. Amino acids spaced three or four apart are very close to each other as discussed below. In some proteins another helical type, a so called 3_{10} helix, often occurs at the C-terminal end of an α-helix.

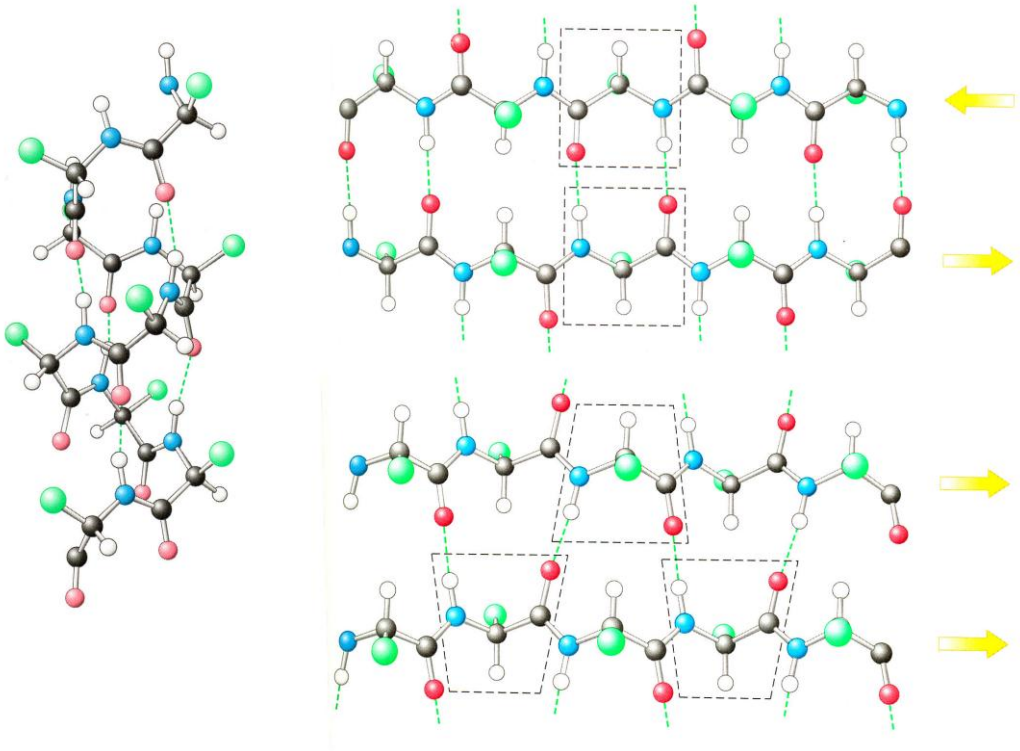

Fig. 2.5: Secondary structure elements of proteins. Left the α-helix, upper right the parallel β-sheet and lower right the anti-parallel β-sheet structure. Carbon atoms are black, nitrogen atoms blue, oxygen atoms red and hydrogen atoms white [Ber03].

Of the many varieties of β-sheets the common antiparallel and parallel β-stranded forms are discussed. Fig. 2.5 visualizes the stabilization by hydrogen bonds between two different strands which are an arrangement of almost fully extended amino acid chains. The side groups are usually in an alternating conformation. All β-sheets are more or less planar structures which exhibit just a slight twist over the complete length. In contrast to the α-helix structure, where different amino acids are as close as 1.5 Å, the distance in β-sheets is much larger, up to 3.5 Å [Ber03]. In the antiparallel β-sheet one amino acid is bound to another one in the adjacent strand by two hydrogen bonds at the C=O and N-H sites. In the parallel form one amino acid is bound to two amino acids by one hydrogen bond each as seen in Fig. 2.5.

2.1.1.3. The Tertiary Structure of Proteins

The spatial arrangement of secondary structure elements is called the tertiary structure of proteins. The reasons for a further packing will be discussed in the folding chapter 2.1.3. The tertiary structure of a protein is stabilized by a combination of van der Waals, hydrogen and

disulfide bonds [Tus07]. The first type is a much weaker interaction but due to the dense protein packing and the many possible atom pairs this contribution is significant. The tertiary structure is first to change by denaturation which can be defined in several ways. Here we will use the definition of a significant structural change that unables the protein to function. Typical thermal denaturation energies of many proteins are not sufficient to break up secondary structure elements. Hence the tertiary structure must be relevant for the protein functionality since this is the only arrangement that is altered as the functionality changes or disappeares. Theory predicts that tertiary structure motions and binding energies are related to the far-infrared frequency range.

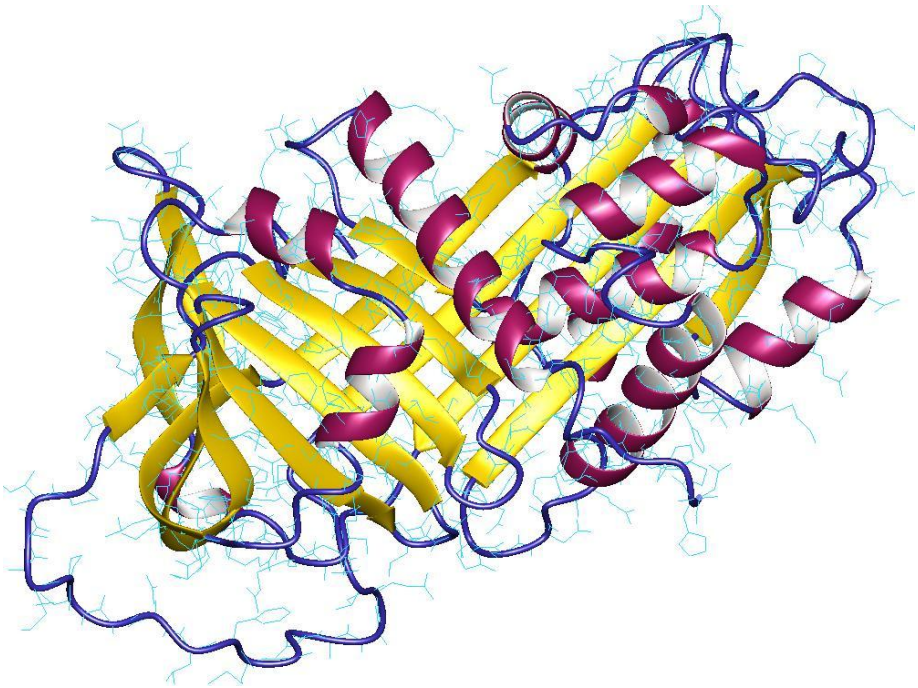

Fig. 2.6: The protein A1-antitrypsin as an example for tertiary structure including α-helices in red, β-sheets in yellow and random coil structures in blue [PDB08].

2.1.1.4. The Quarternary Structure of Proteins

If more than one peptide chain forms a superstructure which sometimes also includes non organic groups, this is called the quarternary structure of a protein. For humans the most important example is hemoglobin (fig. 2.7) consisting of four peptide chains and additional iron containing heme groups.

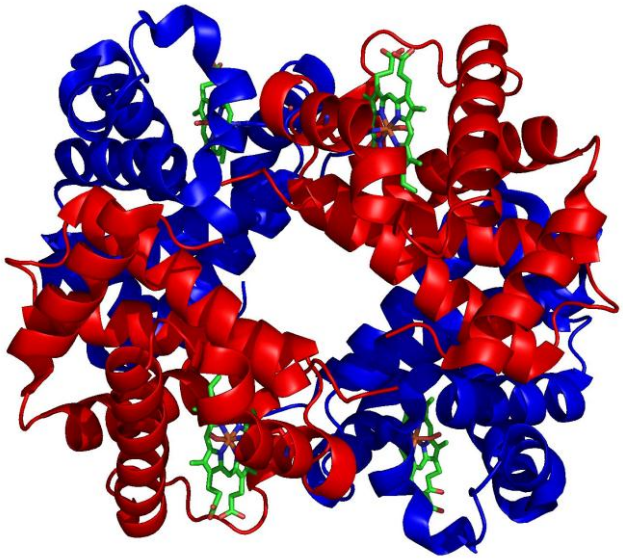

Fig.2.7: The quarternary structure of human hemoglobin consisting of four peptide chains (red and blue) and additional iron containing heme groups (green) [PDB08].

2.1.2. Bond Types in Proteins

The different types of bonds occuring in proteins and their corresponding motions can be observed in different infrared frequency regions. As a basis for the later discussion an overview over the respective properties, relations and energy scales is presented here. Since bonds concerning anorganic groups to be found sometimes at functional binding sites of proteins do not occur in our measured proteins and disulfide bonds are just found in special side groups, the focus of this section will be on the relevant covalent and hydrogen bonds as well as the van der Waals interactions. The latter two are important for our later derived MERCEP model used for the interpretation of the temperature dependent far-infrared measurements in chapter 4.6.

2.1.2.1. Covalent Bonds in Proteins

In covalent bonds the electron shells overlap and an electron sharing mechanism between the atom pair is found. Single covalent bonds are usually free to rotate if they do not have a double bond character, like the peptide bond through isomerisation as an example in section 2.1.1.1. Pure double bonds are not able to rotate. Atomic pair motions like stretching vibrations, bending and rotational modes along the bond can be observed in the infrared spectra. The first type is typical for the mid-infrared frequency range. For instance in N-methylacetamide, a model compound for peptide bonded molecules like proteins, the lowest covalent stretching mode is well above 600 cm^{-1} [Mir96] which indicates that these types of motion do not play a significant role in the far-infrared frequency range and if, that they may contribute just to the very upper part of the here presented measurements. It is more typical to observe bending and rotational motions of covalently bound atom pairs in the far-infrared. Single covalent bond lengths [Eng91] in proteins are in the range of 1 Å (O-H bond) to 1.5 Å (C-C bond) and double bond lengths are slightly shorter than their single bond analogue. The bond energy [Tus07] ranges from 290 kJ/mol (C-N bond) to 720 kJ/mol (C=O bond) and is well above the thermal energy of 2.5 kJ/mol at room temperature and the used infrared radiation within a range of 0.48 kJ/mol at 40 cm^{-1} to 8.25 kJ/mol at 690 cm^{-1} [Wil63].
It is important to mention the interplay between covalent bonds and hydrogen bonds like it occurs for N-H and C=O bonds for instance. If there is a hydrogen bond formed, broken or altered by temperature, conformational or random effects, the covalent bond is influenced and therefore sensitive to such small changes, too. A formation or a stronger hydrogen bond leads to a weakening and larger average distance between the covalently bound atoms due to electron shell relocations.

2.1.2.2. Hydrogen Bonds in Proteins

Proposed for the first time in 1912, the basic work on hydrogen bonds was done by Linus Pauling [Pau67]. By comparing water to other substances like H_2Te or H_2S the theoretical value for the melting point would be about -100°C and for the boiling point about -80°C. The large difference arises from the hydrogen bond network which is in water obviously a significant energy contribution.

In proteins hydrogen bonds play an important role in the formation of the secondary structure elements and are hence involved in protein folding and stability. The conclusion from energy measurements on proteins is that every hydrogen bond contributes around 4.2 kJ/mol to the protein stability [Mye96]. Structural relevant buried polar groups inside a protein are to 90 % bound via hydrogen bonds as *ab initio* calculations on several biomolecules show [Mei05]. Hydrogen bonds involving side chains are a key element in tertiary structure stabilization [Bor94]. Even in some enzymatic reactions hydrogen bonds are involved, emphasizing their role for protein functionality [Der95].

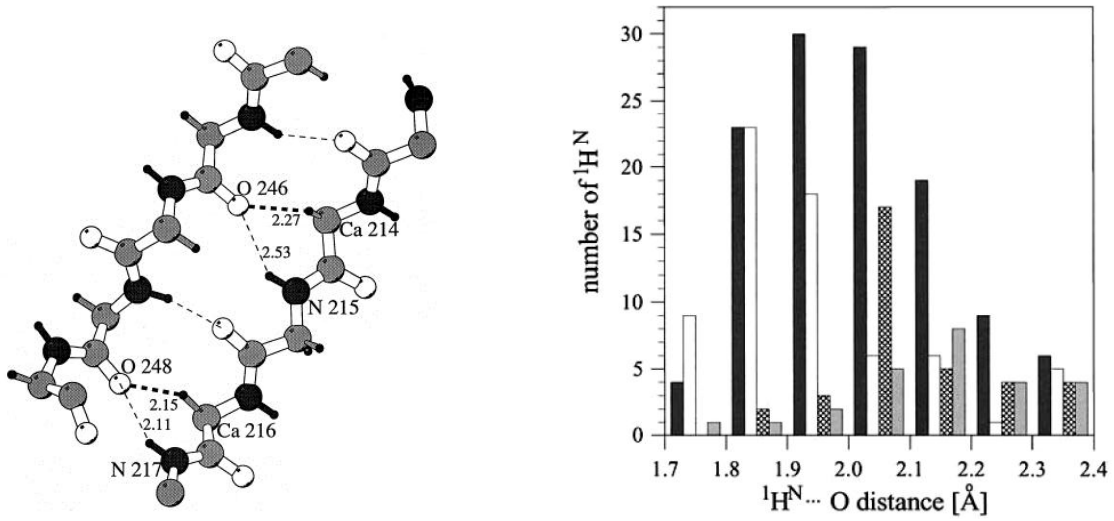

Fig. 2.8: Parallel β-sheet structure with typical atom to atom hydrogen bond lengths on the left [Der95]. On the right, it is shown that the average hydrogen bonds length in parallel β-sheets is about 1.7 Å to 1.9 Å, in α-helices 1.8 Å to 2.1 Å and in antiparallel β-sheets 2.1 Å to 2.3 Å – ordered with increasing average hydrogen bond length [Cie01a].

A common criterion for a hydrogen bond refers to a non-covalent bond average distance smaller than 3.5 Å between an electronegative acceptor atom and a hydrogen atom obeying also some restrictions on the angle between the hydrogen bond and the covalently bound polar group, the latter acting as a donor within the three atom system X-H⋯Y [Kim05, Dau97]. In figure 2.8 typical atom to atom hydrogen bond lengths of a β-sheet structure are shown. Depending on the type of secondary structure elements, different hydrogen bond lengths and hence energies are found as it is seen in fig. 2.8 on the right, which are relevant for the later described secondary structure content determination derived from covalent bond motions by mid-infrared spectroscopy. Typical hydrogen bond lengths depending on the atom type are found for instance in [Tee91, Hey06]. In the densely packed protein interior a distinction of

bound atom pairs is not always easy from experimental data as well as from simulation results. There is no sharp transition from hydrogen bonds to van der Waals type bonds dealt with in the following section. For instance carbon-carbon interactions are clearly of van der Waals type since the majority of next neighbors is found between 3.5 Å and 4 Å distance in radial distribution functions. Just a few of them are found closer than 3.5 Å. Usually N⋯O distances are at maximum 2.8 Å and hence clearly hydrogen bond dominated. But C⋯O distances exhibit an intermediate character of the two described before with the majority at 3.5 Å to 3.75 Å distance which is just a little bit closer than the C⋯C distance and well around the upper limit of the hydrogen bond definition [Der95].

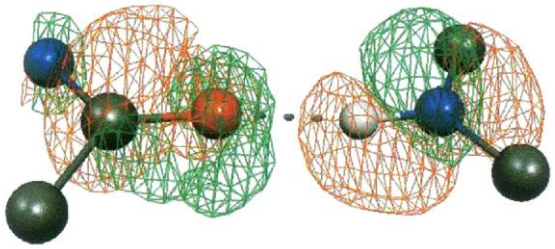

Fig. 2.9: The electron density is changed along a hydrogen bond. The red mesh around the oxygen atom (red) indicates an increased electron density. The green mesh around the hydrogen atom (white) visualizes the decreased electron density upon the polarization of the hydrogen bond. The electron density around the nitrogen atoms (blue) is affected depending on the covalently bound atom type [Tse07].

As it can be seen from fig. 2.9 the electron density around the hydrogen bond atoms is shifted even affecting the neighboring covalently bound atoms, leading to an enhanced dipole moment of the hydrogen bond system [Tse07]. When taken into account the van der Waals radii, where the repulsive term dominates the van der Waals interaction, the hydrogen bond between H⋯O is 0.3 Å less than the mere summation of 1.2 Å (carbon bound hydrogen) and 1.5 Å (oxygen) [Der95]. Hence upon a hydrogen bond breaking in proteins at elevated temperatures we conclude from this fact that we can expect a non-linear distance increase, larger than the sole thermal expansion, of such a weak hydrogen bond. The hydrogen bond breaks and the repulsive part of the van der Waals interaction will increase the distance at least those 0.3 Å between the former bound atom pair leading to local non-linear expansion. Also the energy potential for the covalently bound A-H group is altered under the formation of a hydrogen bond as seen in fig. 2.10 [Jef91]. The typical hydrogen bond energy is in the range of 8.4 kJ to 62.8 kJ per mol [Sch94] whereas in proteins the hydrogen bond energy is

more in the range of 2 kJ to 21 kJ per mol [Des91, Ber03]. This is a value just slightly above the thermal energy at room temperature which is about 2.5 kJ. So hydrogen bonds will be affected by temperature and hence are likely to play a role in a denaturation event.

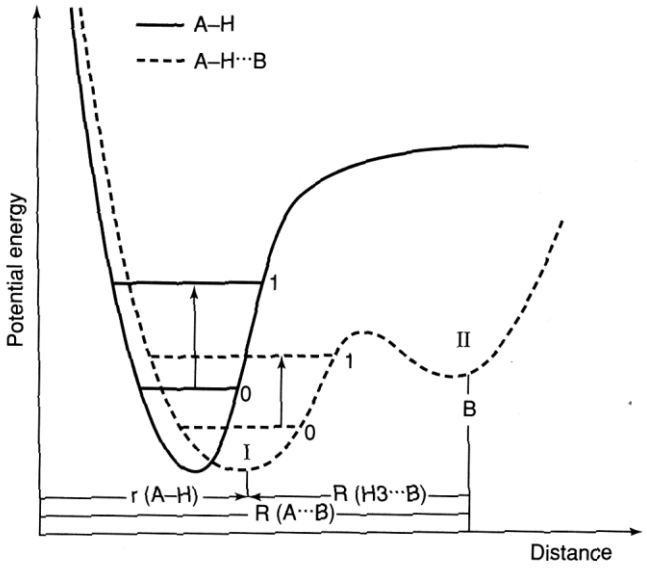

Fig. 2.10: On the formation of a hydrogen bond the potential energy curve flattens and develops a second minimum whose depth increases with the strength of the hydrogen bond. The A-H stretching vibrational energy levels get closer and the energy barrier between these two wells is very sensitive to the environment [Jef91].

Proteins form on average 1.1 hydrogen bonds per residue which sums up to a sizeable number in large proteins as measured here. The backbone hydrogen bond C=O···H-N is the most common with 68 % followed by C=O···side chain and N-H···side chain hydrogen bonds with 11 % and 10 % respectively. Hydrogen bonds between two side chains occur with 11 %, too. The average hydrogen bond length is 3.04 Å. To a close approximation the number of hydrogen bonds scales with the secondary structure content. All these results have been gained in a statistical manner from Stickle et al. by analyzing information of 42 different proteins gained by many different X-ray experiments. [Sti92]. Concerning the hydration shell of proteins, the C=O group has the option to form to two or more hydrogen bonds in contrast to the N-H bond. Therefore 70 % of the water-protein hydrogen bonds involve oxygen atoms of the protein rather than nitrogen atoms [Dem97].

Hydrogen bond stretching and bending motions can be observed in the far-infrared frequency region [Zun96]. Raman spectra on hydrated lysozyme and myoglobin were attributed to hydrogen bond vibrations in the range from 60 cm^{-1} to 210 cm^{-1} [Eav03]. Combining several

publications [Hal63, Joe74] all agree to find hydrogen stretching motions below 250 cm^{-1} or at least below 600 cm^{-1} [Jef91] down to about 50 cm^{-1}. Hydrogen bending motions are assumed to occur below 200 cm^{-1} and 50 cm^{-1} respectively. Hence it can be concluded that at least the hydrogen stretching modes will contribute to our measured frequency range. As Joesten and Schaad [Joe74] pointed out, two hints on the observation of hydrogen bonds can be given concerning our measurements, since an assignment of hydrogen bond absorption features is difficult. First the polarizability change upon the formation of hydrogen bonds is small and so the corresponding Raman features should be weak. Second very broad and asymmetric absorption bands are expected in the far-infrared region due to the strong influence of the surrounding on the weakly coupled hydrogen bonds.

2.1.2.3. Van der Waals Bonds in Proteins

Real gases show deviations from the ideal gas formulas especially on increased pressure and at lower temperature where the interaction between molecules is increased. Going back to the publication from van der Waals in 1873 [Waa73] this was explained by an interplay of a strong short ranged repulsive and a weak but long range attractive force [Kra83]. The two new introduced constants *b* and *a* belong to the above described forces in the van der Waals equation (2.1).

$$\left(P + \frac{an^2}{V^2}\right)(V - bn) = nRT \tag{2.1}$$

Substances where no ionic or covalent bonds between molecules are present, like helium for instance, exhibit some attractive force between the molecules which arises from fluctuations in permanent and induced dipoles. This force is very weak which becomes obvious for instance in enhanced evaporation even by small temperature increases in these substances out of the liquid state [Tip98]. The attracting long ranged part of the van der Waals interaction has been often neglected due to the focus on the strong short range repulsive part, important for volume calculations and structural concerns [Cha83].
If in general van der Waals interactions are relevant, the main contribution comes from the dispersion interaction part arising from dipole induced dipole interactions and hence we will use the term van der Waals interactions in the following sections assuming this dispersive part

as the most important one, without an exact knowledge of the share each different subtype of interaction contributes.

Usually van der Waals interactions are defined as a sum of Keesom, Debye and London contributions [Rot96]. The first one refers to dipole – dipole interactions, the second to dipole – non polar interactions and the last one to non polar – non polar interactions, based on the work from London in 1930 [Lon30]. The last interaction is also named dispersion force because London found a link between his second order perturbation theory, concerning this energy, and optical dispersion [Buc80].

In general van der Waals interactions arise from fluctuations in the electron density which can induce a dipole in an adjacent atom or molecule. Therefore the van der Waals interactions increase with the size of the electron shell and depend on the polarizabilities of the influenced atoms or molecules. But many aspects about van der Waals interactions remain unclear even today [Gla01]. Their spacial distribution and exact distance dependence within dense matter are just two of the open questions. The existence of a molecule specificity and a frequency dependence of the interaction is also still under discussion.

Hamaker and Lifshitz [Ham37, Dzy61] described the van der Waals interaction between macroscopic bodies in their modern form. The classical way to estimate the van der Waals interaction energy between two colloidal bodies is given in the first paper by summing pairwise each contribution. Here the so called Hamaker constant, depending on the polarizabilities and the density of atoms in both bodies as well as on the polarizability of the medium in between, describes the free energy of this interaction. This Hamaker constant can be derived from quantum mechanical calculations, too. Here the origin of the van der Waals interaction comes from the classic electrodynamical theory, where a zero-point radiation is included. By using the retarded Green's function the homogenous solution of the Maxwell's equations corresponds to a random fluctuating radiation with a Lorentz invariant spectrum. This random fluctuating radiation drives polarizable particles into random oscillations. The radiation which is emitted from two particles is hence correlated to the initial random radiation. This gives a nonvanishing average for the classical Lorentz force on each particle named van der Waals force. At elevated temperatures just the spectrum of the fluctuating random radiation changes [Boy75]. A more modern description is given for instance by Laberge [Lab98] derived from a multipole approach. A list of the different interactions and their distance dependence is given in table 2.1. Laberge pointed out that the dispersion forces can be of long range spreading out even more than 10 Å. For energy calculations the additivity of interactions of a three atom system is relevant. To complete the distance

dependence of table 2.1, Coulomb interactions are found to be non-additive, electrostatic interactions additive, induced dipole interactions non-additive and dispersion interactions non-additive [Buc80, Elr94, Kar06].

Interaction	**Spatial geometry dependence**	**Energy distance dependence**
Electrostatic		
Charge-Charge	no	1/r
Charge-Dipole	yes	1/r
Charge-Dipole (long range)	yes	$1/r^2$
Dipole-Dipole	yes	$1/r^3$
Charge-Quadrupole	yes	$1/r^3$
Dipole-Quadrupole	yes	$1/r^4$
Quadrupole-Quadrupole	yes	$1/r^5$
Induced Polarization	yes	$1/r^4$
Dispersion	no	$1/r^6$
Exchange Repulsion	no	$1/r^{12}$
H-Bonding	yes	~1.2-3.5 Å

Table 2.1: Overview over non-covalent interactions in proteins including their geometry and distance dependence [Lab98].

We will have a closer look on the dispersion forces arising from London interactions for the arguments given above. Under the assumption of two harmonic oscillators the energy potential caused by the attractive dispersion forces can be calculated from the oscillation frequencies ν and the polarizabilities α to the well known r^{-6} distance dependence [Dau97]:

$$E_P = -3h\frac{\nu_1\nu_2\alpha_1\alpha_2}{(2\nu_1+\nu_2)r^6} \quad . \tag{2.2}$$

The attractive dispersion force and the repulsive part, which is often described by a r^{-12} distance dependence, lead to the Lennard-Jones potential description of the van der Waals potential energy function as visualized in fig. 2.11. The point of equal attraction and repulsion gives a distance optimum for the atom pair. At larger distances the attractive part is the major component decreasing with r^{-6} towards larger distances.

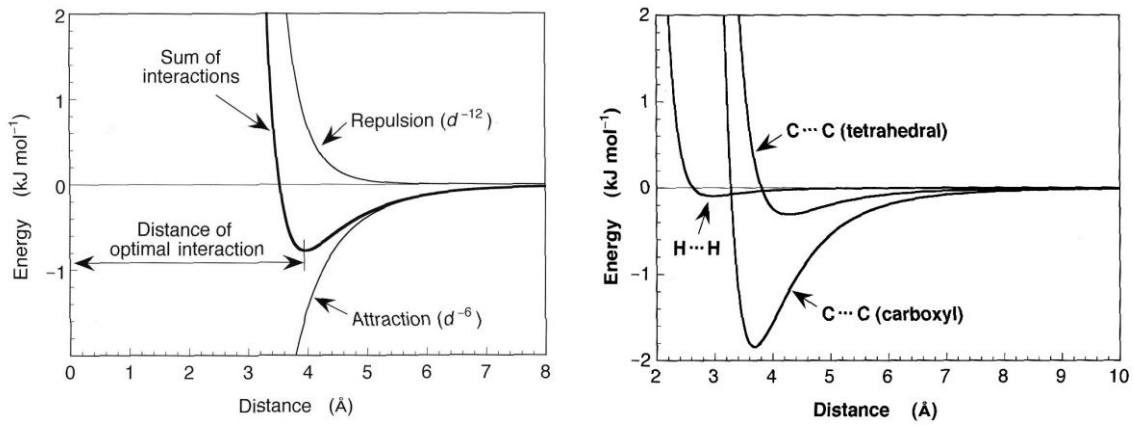

Fig. 2.11: Left: Van der Waals potential for two carbon atoms. The repulsive term originates mainly from the quantum mechanical Pauli exclusion principle and the attractive term from the induced dipole-dipole interactions [Nöl06].
Right: Examples of van der Waals potentials for two hydrogen atoms, two carbon atoms in a tetrahedral arrangement and two carboxyl atoms for their distance dependence and their calculated energy [Fer85]. The potential minimum as well as their long-range attraction differ significantly between these types of atom pairs [Nöl06].

As fig. 2.11 shows, the potential minimum as well as the long range attraction are different depending on the type of atom pairs. Depending on the literature [Nöl06, Tus07] the values of this pair distance range all between 2.6 Å and 4.4 Å. The relevant energy scale for biological samples is even more scattered, starting from very small values of 0.1 kJ/mol up to 2 kJ/mol [Nöl06]. A little bit larger values are proposed in the books from Berg and Tuszynski [Ber03, Tus07] which are in the, in our oppinion more likely, range of 2 kJ/mol to 4.5 kJ/mol per atom pair. The van der Waals energy is hence about 100 times lower than for covalent bonds and 2 to 10 times lower than for hydrogen bonds, almost reaching the lower edge energy values of weak hydrogen bonds. But due to the huge number of possible pairs (much larger than the statistical 1.1 hydrogen bonds per residue in proteins as cited above) this interaction can become substantial and hence an important factor in densely packed systems like proteins. Also it is very interesting to mention that the van der Waals bond energy in the range from 0.4 kJ/mol to 4.2 kJ/mol is around the thermal energy of 2.5 kJ/mol at room temperature and therefore strongly influenced by this parameter. We will come back to this in the discussion of the temperature dependent measurements and in the section about our new model in chapter 4.6.

For several reasons the importance of van der Waals interactions in biological matter can be emphasized. For protein stabilization van der Waals interactions are indispensable [Cre85].

Collagen chains are held together by van der Waals forces and some hydrogen bonds. As a debatable example for secondary structure stabilization the results from modeling the forces of an α-helix in cytochrome revealed that the van der Waals energy is the major component of the total energy. It is two to three times larger than the energy contribution arising from hydrogen bonds due to their large number. In general the energy solely due to van der Waals interactions was found to be around 84 kJ/mol to 126 kJ/mol protein, which is a very large value since here now many van der Waals interactions per atom are summed up and not only one pair interaction as above. The conclusion of this modeling is that the van der Waals interaction determines the local barriers for forming or non-forming α-helices, since hydrogen bonds are energetically profitable in both helical and non-helical regions. Therefore the van der Waals interaction has to determine the formation as the authors suggest [Kil04]. For the stabilization of protein tertiary structures many publications indicate that the packing of apolar side chains is here the driving force and all side chains interact by van der Waals forces [Kel88, Kel89, Tus07].

The dense packing of proteins is anyway related to the van der Waals interactions. On one hand van der Waals interactions are necessary to reach such a dense packing and on the other hand the dense packing enhances the distance dependent van der Waals interactions and make them a relevant contribution in proteins [Lim91, Pap96, Ber03]. In a simulation of lysozyme adsorption on charged surfaces, the van der Waals interaction was found to play a significant role. From the total energy at a distance of 3 Å between protein and surface, the van der Waals interaction contributes about 20 % of the total energy. At 5 Å this is still about 15 % of the total energy. For weakly charged proteins the computed results reveal that van der Waals attraction is the predominant contributor to the adsorption on the surface. This result shows qualitatively that even if electrostatics are the major component, van der Waals interactions with their fast spacial decay cannot be neglected even at comparable large distances [Rot93]. Later the results have been verified in general by corresponding experiments [Rot95]. Also it was concluded that in matter with an overall dipole moment less than 1 D it can be assumed that the London dispersion forces dominate the infrared absorption relevant total dipole moment [Rot96]. Even for calculations concerning the long range interactions between two bovine pancreatic trypsin inhibitor proteins in electrolyte solution, van der Waals forces have been analyzed and included into the molecular dynamics simulation [Son04]. Van der Waals interactions are essential for a real description of water structure and dynamics as the outcome of *ab initio* molecular dynamics simulations revealed. A comparison to neutron data indicated that at distances where dispersion forces rule, from about 3 Å to 5 Å distance, just models

including van der Waals interactions are able to describe the experimental data [Lin03]. In biological systems, with dominating hydrophobic groups, van der Waals interactions are assumed to play an even more crucial role than for the highly polar and hydrogen bond forming liquid water.

To get an estimate for the relevant frequency range of van der Waals interaction contributions, a study on the van der Waals Trimer Ar_2HCl [Elr91a,Elr91b] will be quickly discussed here. Intermolecular vibrations of this molecules have been found by THz spectroscopy and compared to calculations. Bands exist at 37 cm^{-1} as in-plane bend, at 40 cm^{-1} as Σ-bend and at 45 cm^{-1} as out-of-plane bend motions with a calculated frequency deviation of less than 4 cm^{-1}. This demonstrates the occurrence of intramolecular vibrations at that frequency range in systems having significant van der Waals interaction. Another indication comes from simulations of the dielectric properties of methyl cyanide which is a highly polar fluid [Edw84a, Edw84b]. The comparison of the calculation with the actual spectrum shows mainly reorientation effects around 70 cm^{-1} but induced dipole effects are essential to match the spectrum derived from the simulation with the experimental one, in particular in this low frequency range. In general, motions of weak bonds, especially from larger molecules are found in the far-infrared or THz frequency range, whereas motions of strong bonds are found in the mid-infrared.

Photons in the lower half of our frequency range have enough energy to excite van der Waals bond motions or even to break them. The van der Waals bond energy of an atom pair is around 4 kJ/mol as stated above. Taking the energy values of the lower and upper edge from Tuszynski [Tus07] and calculating the attributed frequency at which this energy is provided, we get for the energy range of 2.1 kJ/mol to 18.8 kJ/mol the frequencies of 42 cm^{-1} and 376 cm^{-1} with the attributed temperatures (so that k_BT provides the energy) of -220°C and 270°C respectively. Van der Waals interactions, together with interactions mediated directly over an ensemble of covalent bonds, will lead to the concept of collective motions, which means a correlated motion of several atoms in a concerted way. In summary it is very likely to observe temperature sensitive effects especially upon van der Waals interactions in our far-infrared frequency range.

2.1.3. Protein Folding

Starting from the stretched out amino acid chain assembled at the ribosoms the protein starts to fold under suitable conditions which is a process not completely understood up to now [Fra94]. Those suitable conditions include the right temperature range, the type of solvent and the pH value. As we have seen in chapter 2.1.1, the protein structure is relevant for the functionality and misfolded proteins can cause diseases. Concerning our measurements, the knowledge about the folding process will lead us to the concept of the energy landscape which is the basis for protein dynamics in the next section. Also the later presented comparison of proteins to glass formers evolves from the global, but not completely defined, minimum reached at the end of the folding process justifying this section about the folding process.

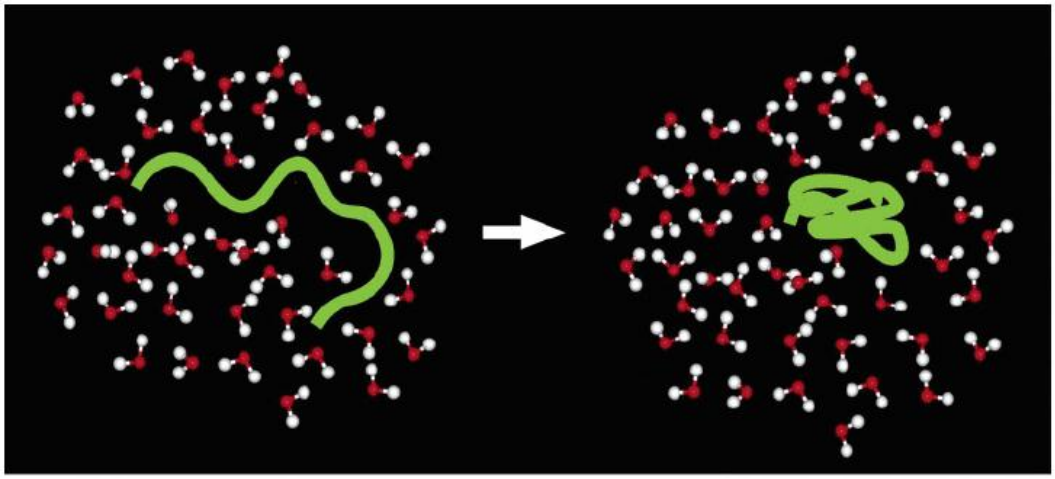

Fig. 2.12: The hydrophobic effect as a major driving force for protein folding in aqueous solutions [Pal04].

One of the major "driving forces" of protein folding is the hydrophobic effect. Assuming folding in an aqueous surrounding like cell plasma or blood, at least some side chains of amino acids are hydrophobic and are energetically in a lower state by forming an almost water free protein interior. It is not always possible to get all hydrophobic groups to the inside due to steric considerations. A hydrophobic side chain could be adjacent to a hydrophilic one leading to a large friction for this part of the protein during the folding process. But in general more hydrophilic side groups are situated at the water covered protein surface but still hydrophobic ones are found there for this reason, too. The amount of hydrophobic nonpolar

amino acid groups at a protein surface is on average hence a small majority of 57 % [Jan99]. But in lysozyme, as one of the measured proteins, a summation derives a ratio of 59 % polar (N and O groups) and 41 % non polar surface areas for instance [Lee71]. A visualization of the hydrophobic effect, caused by the interaction of protein side chains and the surrounding water molecules, which triggers a folding of the peptide chain is presented in fig. 2.12.
Fig. 2.13 shows the side group classification of myoglobin, the "hydrogen atom" of protein science. Many charged and polar hydrophilic (blue) amino acids are found at the surface (left picture) among a large number of hydrophilic (yellow) and other (white) amino acids. The cut through the folded protein (right picture) reveals that in the interior the majority of amino acids is hydrophilic. But also hydrogen bond formation, van der Waals interactions, steric hindrance and water molecule dynamics influence the protein folding process in addition to the hydrophobic effect.

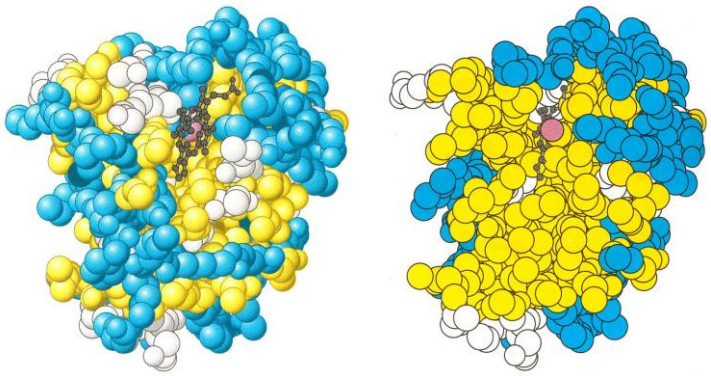

Fig. 2.13: Amino acids at the myoglobin surface (left) and in the protein interior (right). Charged and polar hydrophilic amino acids (blue) are found mainly on the outside, whereas hydrophobic (yellow) and other (white) amino acids are found mainly at the interior but also on the water covered protein surface [Ber03].

A visualization of the dynamics of a small side chain folding is picturized in fig. 2.14. Several energy barriers on the way to the final conformation have to be surmounted. After a first nucleation some intermediate, free state has to be formed again prior a different ordering is able to be formed. Backbone fluctuations together with parts frozen in their folded or misfolded states indicate the randomness of this process and anharmonic nature of motions. Even here with just a few connection points in a side chain, the large number of different possible conformations and intermediate states becomes evident that have to be formed on the way to the proper folding.

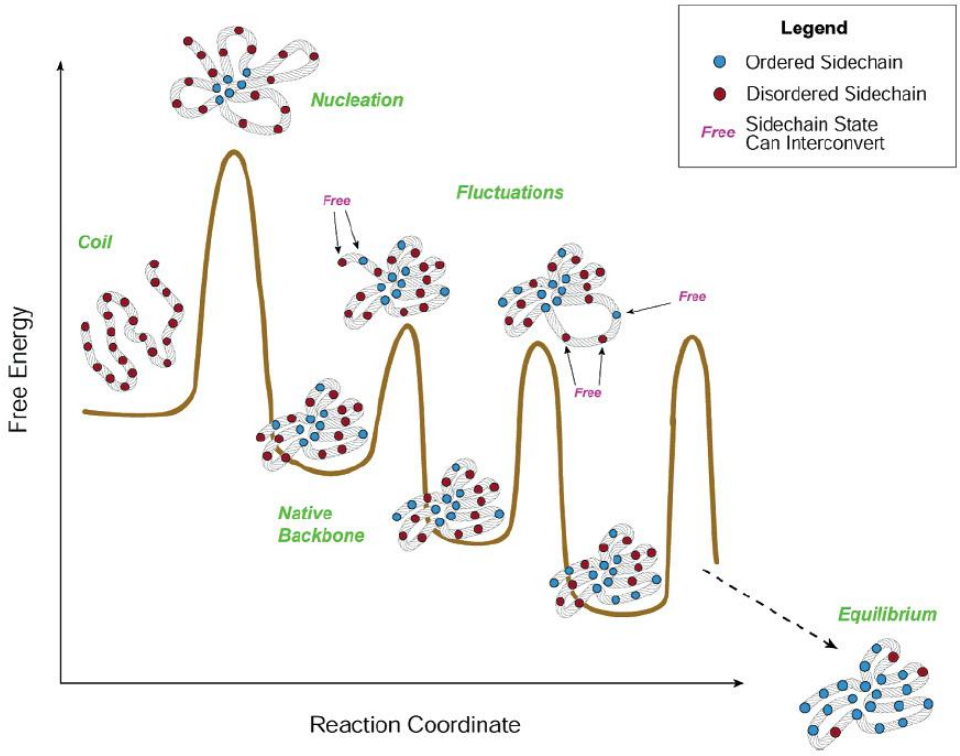

Fig. 2.14: Side chain organization in a folding process. Barriers heights and intermediate conformations are just chosen for the visualization of the random character. The connection points indicate the ordered or disordered conformation state, respectively [Shakhnovich2006].

This side chain organization leads us to the so called Levinthal paradoxon which is a strong argument against a random walk during protein folding [Nie04, Pet87]. Going back to the example given by Jaenicke and Kell [Jae84, Pet87] even at a theoretically assumed very fast rate of exploration of 10^{15} s^{-1}, the time of the universe, estimated to be 10^{17} s, will be not enough to discover all possible 10^{80} different conformations of an average 200 amino acid containing protein. Since in nature even the slow folding proteins get their conformation within several seconds, there must be a principal direction the folding process is heading for to achive such a many orders of magnitude faster folding result. A directional folding is nowadays described by a so called folding funnel as seen in fig. 2.15 and fig. 2.16 which was investigated and verified by model calculations [Leo92]. The form is given by the energy distribution in a hyperspace based on all possible conformations. In figure 2.15 two possible conformational degrees of freedom are depict, in fig. 2.16 just one degree of freedom, representing certain cuts through the hyperspace.

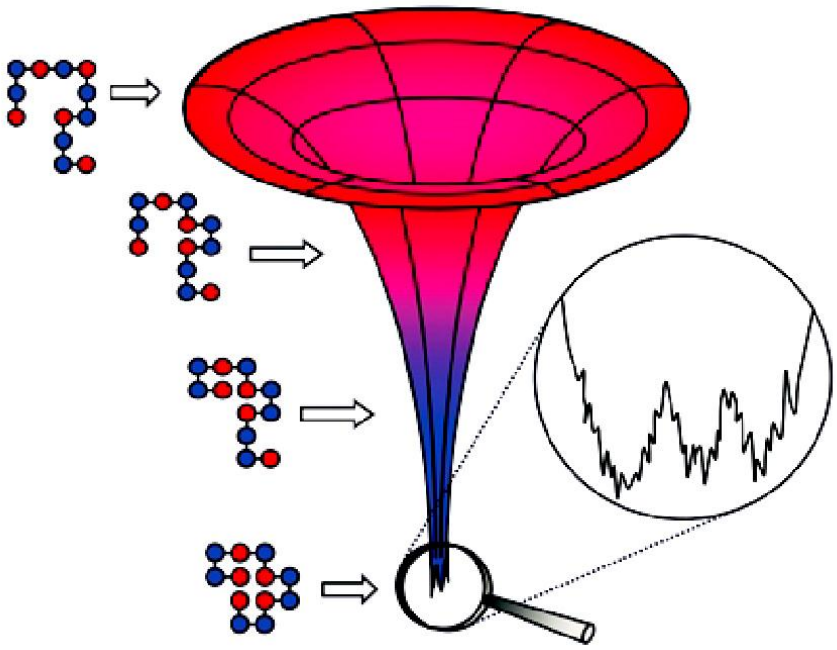

Fig. 2.15: The folding funnel of a peptide chain. Not all possible minima are tested but a directional folding downhill sets in after the initial (random) conformational state. The tip is magnified emphasizing the presence of several energy minima. A unique lowest energy state is missing and the thermal energy is usually high enough to provide adequate energy for testing many different conformations above the very tip of the folding funnel [Nie04].

The initial state of the peptide chain is random as well as the folding into the first conformational substate. Then a random testing of several minima sets in but all downhill in energy towards the folding funnel tip which leads to a "directional random walk". Some states similar to a denaturized state are formed on the way to the natural form which are often named molten globule states as in fig. 2.16. The over all shape is already reached, but functionality lacks. Further folding leads to the energy level at the physiological temperature, regarding the influence of the natural surrounding. In a folded protein no unique state of lowest free energy exists. The lowest state is highly degenerate in slightly different conformations called conformational substates. These substates differ just in details and may exhibit different functional rates for instance. So natural proteins are all folded in the same general manner but differ in the many conformational substates they are in at the moment and which they can leave by transitions to other conformational substates. If the temperature is lowered the glass (fig. 2.16), but better called dynamical transition (see chapter 2.3.1.1) is reached. Here the degrees of freedom get much less upon cooling and the dynamical behavior is significantly restricted. The lowest tips are reached by further cooling and the proteins will

be “trapped” in the discrete minima distributed in a statistical manner. The lack of a unique minimum is one similarity of proteins to glass formers.

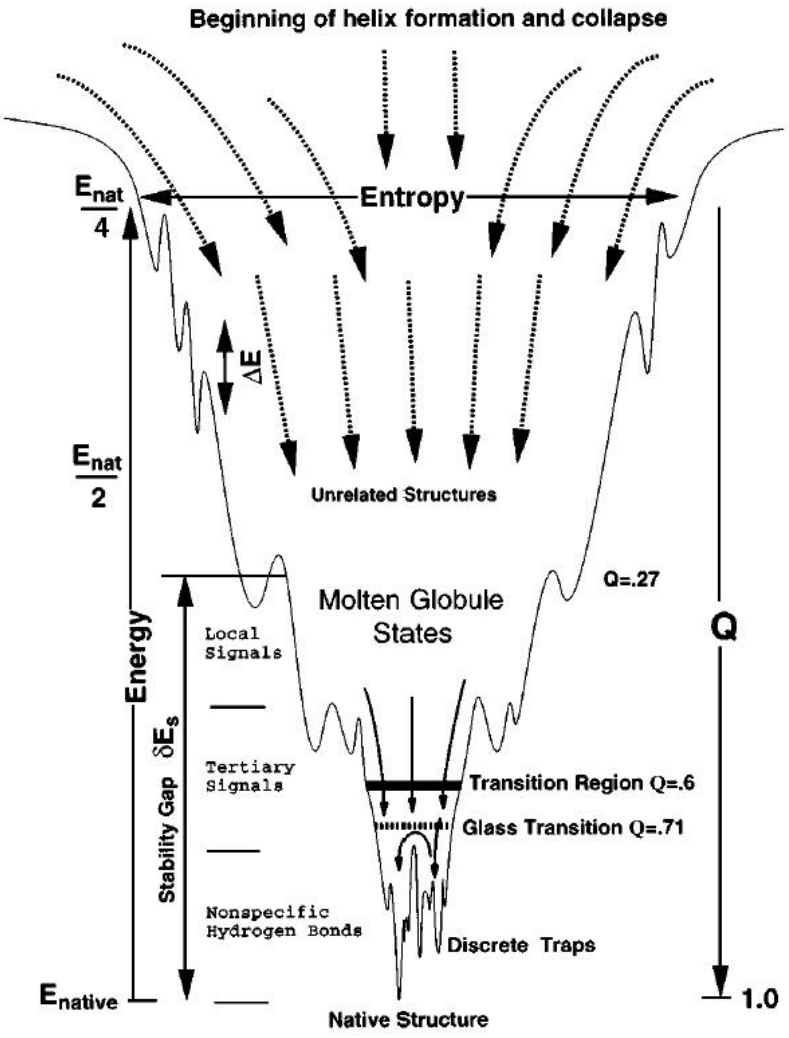

Fig. 2.16: Cut through a schematic folding funnel. The misfolded or molten globule state is an intermediate or denaturized state in which the protein can be trapped under several conditions. Above the discrete traps at the very tip of the folding funnel, the glass transition and the natural energy levels are indicated, which define possible conformations. The parameter *Q* gives the percentage estimate of the native state conformation [Onu97].

After folding, the efficient packing of a protein is comparable to a dense liquid [Nie96]. This is caused mainly by hydrogen bonding and van der Waals forces especially from the many apolar parts and side chains of the protein interior [Har94]. On the other hand the dense packing makes van der Waals interactions favorable. If they are disrupted, we assume here a temperature increase for instance, the protein becomes destabilized [San89]. The delicate optimal conformation of the protein might be fundamentally disturbed and this small change

can hence lead to a much larger "non-linear" effect. This could explain retrospectively the influence of the weak but multiple van der Waals interactions on protein stability and structure as discussed in the former section. In their analysis the authors separated this van der Waals interaction based "packing" effect from the well known hydrophobic effects and found the first one to be almost twice as large in their energy contribution than the latter, namely in the range of 3.4 kJ/mol to 12.1 kJ/mol compared to 0.4 kJ/mol to 6.7 kJ/mol [San89]. In the folded protein the hydrophobic effect is smaller than the van der Waals interactions which are, in contradiction to the first one, temperature sensitive which will be used later in the temperature dependent measurement discussion chapter 4.6.

2.2. Protein Dynamics

In the sections so far the protein had been treated more as a rigid molecule of which we have analyzed the structure. Motions and related frequency ranges have been mainly concerned with the stretching and bending of distinctive bonds. From the dynamical process of protein folding, which ends in no unique energy minimum, we want to derive now the dynamics of the folded protein which is important for the functionality. To understand proteins one has to analyze the motions of many thousands of protein atoms and of some hundreds of attached water molecules together, to get a complete picture which is necessary for a fundamental insight into protein functionality. First we will demonstrate the way that led to the dynamics-functionality association. Then we will look at the huge range of different timescales and amplitudes of all motions occurring in proteins. Out from that the functional relevant motions in proteins will be identified together with their properties, for instance their attributed frequencies and amplitudes, their harmonic/anharmonic character or in general equilibrium fluctuations/non-equilibrium transitions. The concepts of the energy landscape and hierarchy describing folded protein dynamics will be introduced before the influence of temperature and of the solvent on this dynamical behavior will be in focus. Since water plays a crucial role for proteins the properties of "different types" of water and their dynamics will be discussed, too.

2.2.1. The Relation between Functionality and Protein Dynamics

2.2.1.1. Historical Review

Perutz and Kendrew [Per60, Ken60] were the first that got an atomic resolution of protein structures in 1960 on hemoglobin and myoglobin respectively. "The remarkable detail and intrinsic beauty of the molecular structures led to the misconception that proteins are rigid molecules and most attempts to understand protein function have been based on rigid models" [Nie96]. The key which fits into a lock was the visualization of enzyme functionality. This changed slowly in the 1980ies as it came out more and more that protein dynamics plays a crucial role for functionality. The first hint on the importance of protein dynamics for functionality came from Perutz and Mathews [Per66] who found that there is no open path in myoglobin and hemoglobin for gas molecules to reach or leave the binding site – at least on physiological time scales. Otherwise a side chain has to swing out and to allow access which would include protein motions on a large scale. That some transitions in myoglobin require a dynamical protein and do not occur in rigid myoglobin was proven by Case and Karplus in1979 [Cas79]. And that groups of atoms in proteins can move in a concerted way from one minimum to another was found by molecular dynamics simulation from Mao et al. on ferrocytochrome c [Mao82].

The picture of the energy landscape with many different conformations a protein travels through came up around the same time by Bryngelson and Wolynes [Bry87a]. Along this knowledge, different techniques for investigations of these motions were demanded because X-ray crystallography, most used at that time, averages over all these different conformations and just the average value of motion, without detailed frequency or amplitude composition, can be gained. Here the progress in (inelastic) neutron scattering experiments, far-infrared and THz spectroscopy, and the femtosecond pump-probe experiments are mentioned. Also computer simulations on proteins to model this complex dynamical behavior became more important since that time. Yet the relation between protein dynamics and their functionality is known but is still a far from being understood topic.

2.2.1.2. Functionality Related Experiments

Today the importance of protein motions for the functionality is best studied on myoglobin (fig. 2.17). All possible paths for an entering or leaving of oxygen in a rigid structure are related to energy barriers of more than 420 kJ. This would take billions of years, in contrast to the very much faster physiological time scales. Calculated energies of 35.6 kJ, related to many small conformational changes, can decrease the relevant energy barrier down to 20.9 kJ which leads to time scales close to the observed physiological ones [Kar86]. This underlines the connectivity between motions on different timescales, amplitudes and locations within a protein (see chapter 2.2.1.3). Small local conformations, able to overcome the low energy barriers separating the next substate, prepare a way for a global transition within the protein. In addition to the large number of moving atoms in a protein, this connection adds even more complexity to protein dynamics.

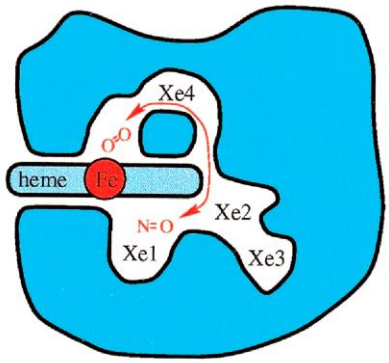

Fig. 2.17: Schematical drawing as a cut through the overall structure of sperm whale myoglobin. The heme group binding site is situated in a protein cavity. The conformation of the two parts, opening and closing the protein cavity, influence the protein functionality [Fra01].

Conformational changes are important to the protein function and their accessibility is critical as it is also seen from mutagenesis and temperature manipulating experiments. In lysozyme for instance so called hinge bending motions are involved by clamping down polysaccharides in bacteria cell walls when breaking these down. Lysozyme is described as a protein with two lobes covering the active site in between, whose relative motions have been analyzed by a very early model assuming two rigid simplified lobes after the functional relevance was discovered [McC76]. The lysozyme and myoglobin examples indicate among others that in particularly motions involving many atoms seem to be relevant for protein functionality. That so called collective motions are important to the protein function, has been also stated by

Huber and Bennett and was confirmed by simulation results on lysozyme for instance [Hub83, Lev85]. The first molecular dynamics simulation on proteins to support experimental findings has been done several years earlier by the group of McCammon and Karplus on bovine pancreatic trypsin inhibitor in 1977 [McC77]. These early simulation results already showed the comparability and usefulness of the gained information when comparing them to real observables. Concerted motions of some atoms have been found and a connection to low frequency modes has been already derived at that time.

2.2.1.3. Timescales and Amplitudes of Protein Dynamics

Table 2.2 shows the timescales and amplitudes of protein motions based on the work of McCammon in 1984 [McC84]. The timescale spawns 15 orders of magnitude. Even if more recent results indicate that the here presented overview might be about one order of magnitude too low for some motions, a first indication is given in which spectroscopic relevant frequency range and respectively energy scales the different protein motions can be expected. Vibrations of bound atom pairs are found in the mid-infrared with main features between 1500 cm^{-1} and 3500 cm^{-1} covering the timescales around 100 THz, corresponding to 10^{-14} s per period as seen in table 2.2. The above discussed motions of globular regions of a protein are found in the THz range, which we reach on the lower frequency side of our measurements. Due to the damping of the solvent and due to the dense internal packing of proteins, the motions of surface side chains and the torsional liberation of buried groups occur on even longer timescales situated in the upper GHz frequency range. Hinge bending motions are referred to be very low in frequency but McCammon stated already in the same paper that most experiments up to that date indicate a timescale around 20 ps which would be in the very upper part of the given range in table 2.2 [McC84]. (Local) denaturation, reaching the timescale of seconds, is the slowest dynamical process in proteins. As it is known from medicine a denaturation can also occur on a timescale spawning several days, even extending the protein dynamics some orders of magnitude on this frequency side.

Motion	Spatial extent (nm)	Amplitude (nm)	Logarithm of characteristic time per second
Relative vibrations of bonded atoms	0.2 – 0.5	0.001 – 0.01	-14 to -13
Elastic vibration of globular region	1 - 2	0.005 – 0.05	-12 to -11
Rotation of sidechains at surface	0.5 - 1	0.5 - 1	-11 to -10
Torsional liberation of buried groups	0.5 - 1	0.05	-11 to -9
Relative motion of different globular regions	1 - 2	0.1 – 0.5	-11 to -7
Rotation of medium sized sidechains in interior	0.5	0.5	-4 to 0
Local denaturation	0.5 - 1	0.5 - 1	-5 to 1

Table 2.2: Timescales, spacial extent and amplitudes given for internal protein motions. Here just the order of magnitude was estimated based on the information available at that time [McC84].

As an argument for maybe too low frequencies that have been assumed, two molecular dynamics simulations of bovine pancreatic trypsin inhibitor are discussed. They concluded that atomic fluctuations take place on subpicosecond time scale as local fluctuations and as more collective motions in the range of 1 ps or longer corresponding to about 3 cm^{-1} up to 30 cm^{-1} [Swa82]. Large fluctuations within the protein have always been due to these collective types of motion [Ich91a, Swa82]. This indicates that the interface from local to collective type of motions is found in the far-infrared or THz frequency range. But atomic fluctuations are already found around 200 cm^{-1} in this early simulation [Swa82] which is too low in frequency, since these motions are proven to be found in the mid-infrared. So in general, the results seem to be a little bit too low in frequency. But we anyway do not expect a sharp transition frequency at which special protein motions are found and hence it is very likely to see the onset of discovering collective motions in our probed frequency range. More recent reviews on THz spectroscopy and corresponding simulations covering the frequency range of about 1 cm^{-1} to 100 cm^{-1} expect breathing modes of secondary structures, correlated motions of some amino acid groups and tertiary structure fluctuations in this frequency range [Lei06]. All of them important for the protein functionality supporting our choice of measurements on proteins in the far-infrared frequency range.

Many studies found that in particular functionally relevant ligand binding and catalysis involved protein regions often exhibit large amplitude displacements, emphasizing the importance of these kinds of motion [Art79, Hub83, Fra79]. But if one group of atoms moves other protein groups or solvent molecules have to give way, therefore motions in proteins often occur in a correlated manner. This underlines the analogy of protein motions to viscous liquids which is often stated. The dense protein packing causes that on timescales shorter than 0.5 ps, small amplitude motions occur similar to simple liquid motions [McC84]. This arises from a temporary trapping between other protein parts or solvent molecules. There is a "rattling in a temporary cage" motion assumed for proteins. The frequent internal and external solvent collisions induce also random like motions as in liquids. On a longer timescale protein motions behave more solid like. This can be understood by the fact that most of the atoms in a protein have defined average positions corresponding to the native protein state. There is a "restoring force" pointing backwards to the average position after a local motion caused a disturbance. All together overdamped as well as underdamped motions occur in proteins. Large amplitude motions are assumed to be in the THz frequency range or lower, like the hinge-bending in lysozyme or the rotation of side chains at the surface (see table 2.2). Both are subject to damping from the surrounding solvent and frictional forces could be sometimes even dominant. Whereas covalent bonds and small amplitude motions move freely and just small friction effects are involved.

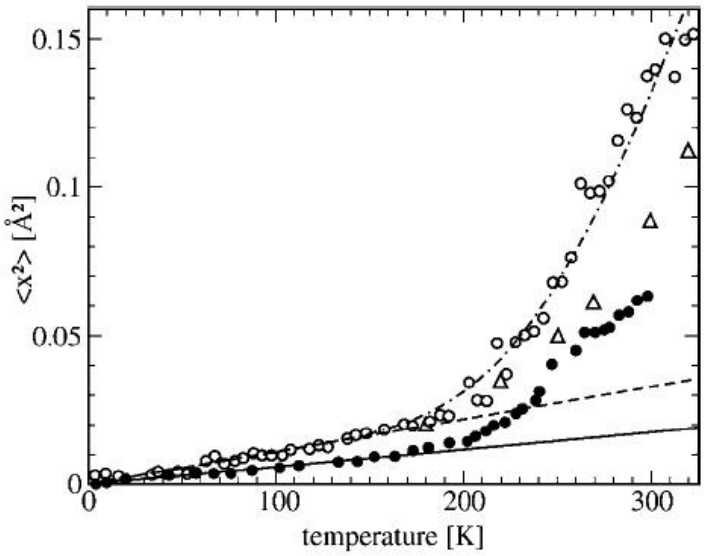

Fig. 2.18: Mean square displacements from neutron scattering (white) on deoxy myoglobin with 80 ps time resolution (circles) and 12 ps time resolution (triangles) from [Dos98]. Mössbauer results are indicated by black circles [Par03].

An overall amplitude of all motions is given by the mean square displacement. Mössbauer spectroscopy and neutron scattering experiments indicate that proteins exhibit a very large value compared to other solids. Fig. 2.18 shows mean square displacements derived from both techniques on deoxy myoglobin states. The later discussed dynamical transition is seen around 200 K. Here we would like to emphasize that molecular dynamics simulations found that just a few anharmonic modes are responsible for the major contribution to this large mean square displacement [Gar92, Ama93, Hay94, Hay95, Kit98, Vli99]. Smith et al. [Smi04] derived from a molecular dynamics simulation of lysozyme at 300 K that less than 1% of the total dynamics contribute more than 90% to the total mean square displacement. This is in line with the observed few kinds of motion in proteins expected to show large amplitudes (table 2.2). Combined together, this indicates that functional relevant motions seem to be related to large amplitudinal, very likely collective, motions situated in the low frequency range.

2.2.1.4. Collective Motions and their Frequency Range

Throughout this thesis we refer to collective motions as a concerted dynamical behaviour of at least five, usually much more, atoms of a protein. Such groups of atoms will range from small sidegroups up to large regions or parts of the whole protein, which are in motion relative to each other. We want to clearly separate collective motions from the typically two or three atoms involving vibrations or rotations along direct bonds. Roughly defined, a group of atoms reveals collective motions if correlated movement of atoms further away than their next neighbours exist. Every study based on small "model" molecules for proteins can therefore not reveal any intramolecular collective type of motion. Since such small molecules are often used for normal mode analysis, information about collective motions are often completely missing which, we want to emphasize here, is not caused by the method itself.
The connection of collective motions to the far-infrared and THz frequency range is based on the joint publication of Brooks, Karplus and Pettitt [Bro88] which stated that even above 100 cm^{-1} collective motions of several atoms like aromatic ring motions occur. Motions in the range of 20 cm^{-1} to 100 cm^{-1} are collective motions of many atoms together usually correlated to large amplitudes. Torsional vibrations of secondary structure elements are assumed in this frequency range, with the helical structures more likely to be situated in the upper half and the sheet structures in the lower half. Hinge bending motions are suggested to be in the THz range precisely between 3 cm^{-1} and 5 cm^{-1}.
At the same time experimental results of neutron scattering in the low frequency range below 150 cm^{-1} did not show a distinction between proteins containing different secondary structure elements. This was shown by Smith et al. on bovine pancreatic trypsin inhibitor [Smi86], myoglobin [Dos89, Cus90] and further proteins as found in [Cus88 and references therein]. The conclusion is that the type of secondary structure elements does not contribute a lot to inelastic neutron scattering and presumably in general to the low frequency range, but so does their spatial arrangement and the relative motion of the secondary structure elements. Since mainly this tertiary structure information is required to unravel the protein functionality process this is an encouraging result for experiments in the far-infrared frequency range.
Molecular dynamics simulation results based on the paper from Go et al. [Go83] found modes in the range from 100 cm^{-1} to 200 cm^{-1} which are of collective nature involving several neighboring atoms or residues. Below 100 cm^{-1} the motions become more and more global involving already large parts of the protein. These motions contribute a large part to the

overall net displacement and are basically of anharmonic character linking up to the next section.

2.2.1.5. Basic Types of Protein Motion

We will extend the conception originally given by Doster and Settles [Dos98] that motions in a protein can be compared to people in a tram. In a tram two major types of motions occur. The first types are small amplitude, harmonic fluctuations caused by the motions of a single person reacting on his moving surrounding. But from time to time large amplitude, anharmonic disturbances, representing the second type, occur which drive the system out of equilibrium and a relaxation to the initial or a new equilibrium state sets in. If a person is surrounded by a bunch of friends the energy arising from a disturbance is faster dissipated in a larger system of people than alone. Hence in this case the equilibrium state is reached on a faster timescale. In a protein these connections are very likely to be predominately the hydrogen bonds and energy is mediated and distributed quickly through this network. Also the dissipation-fluctuation theorem which connects the equilibrium fluctuations with the energy dissipation can be explained by this illustration. Motions from small fluctuations, which always occur in the system, contain as well as the reaction to a large disturbance and the way how dissipation of energy is carried out, information about the whole system. They are connected to each other and from the analysis of just one of these types, the network motions are completely defined and the reaction or occurrence of the other type can be predicted.

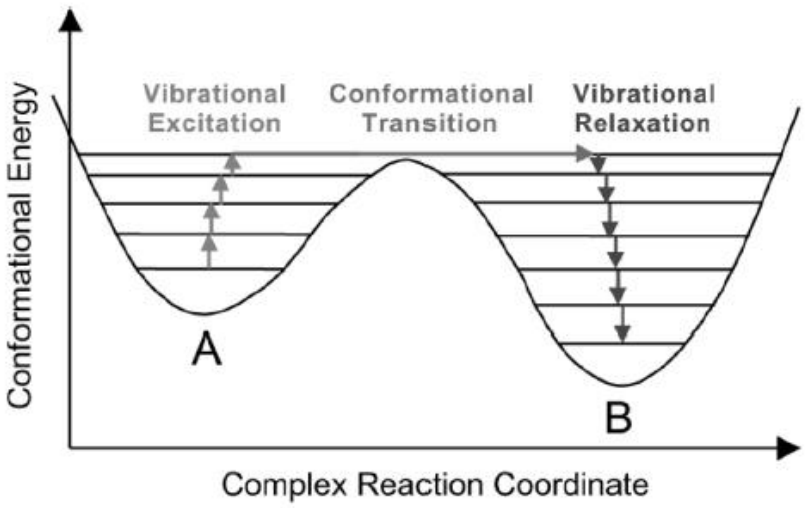

Fig. 2.19: By equilibrium fluctuations slightly different conformations or excited vibrational levels can be reached. The energy is provided by thermal fluctuations and by anharmonic coupling to other states or to the solvent. From some excited states a transition is possibly followed by a relaxation behavior to the new ground state. The anharmonic character of this transition is displayed by the nonsymmetric energy barrier between the two different states. Tunneling effects are not included in this model [Xie02].

Such two types of motions are assumed to occur in proteins, too [Fra88]. Which are namely equilibrium fluctuations and functionally important motions. The first describes an approximated harmonic and the second a relaxation type process. For instance after the photodissociation of myoglobin and carbon monoxide, the new conformational state of the protein after this event is slightly different from the one before. This state is reached by several functionally important motions outspreading like waves within the protein. This event was described by Ansari et al. [Ans85] as a protein quake due to the similarity to an earthquake leading to a new equilibrium state by several follow up processes caused by the first event. The two different kinds of motion, which can cause transitions between different conformational states, can be described in the following way. The equilibrium fluctuations arise from permanent fluctuations within the energy, entropy and volume caused by the interaction with the surrounding, namely the solvent. The solvent can be seen as a thermal bath to which these protein motions are coupled to [Cooper1976]. Nonequilibrium relaxations occur after a perturbation like light quantum absorption, the above described ligand dissociation or even after a transition to a new conformational state. Afterwards a relaxation from this nonequilibrium state to an equilibrium state takes place as seen in fig. 2.19. Both kinds of motion differ in their harmonic and anharmonic behavior respectively. If linear

response of the system is assumed, both types of motion are connected via the dissipation-fluctuation theorem, described in a more general way using correlation functions by Kubo [Kub57, Kub66], which is the basis for the below presented calculation of protein absorption using dipole moment correlation functions. The latter are derived from normal mode analysis or molecular dynamics simulation and hence the details will be discussed in the belonging chapter. Encapsulated, the theoretical models attribute the functional relevant motions to anharmonic, large amplitude motions. A collective character of motions is very likely in such a densely packed and highly connected state which proteins reveal. Methods analyzing the excited state after a major external disturbance as well as methods analyzing the harmonic fluctuations of the system are both capable to gain results which contain information about the complete protein-water system.

2.2.2. The Energy Landscape

As emphasized in the protein folding section (2.1.3), the folding process does not end in a unique energy minima or a well defined state. Instead there exist many slightly different possible conformations, which are all roughly folded in the same manner and shape. This section here will be about the "context" in which both types of transitions between different conformations occur that have been described in the prior section. The influence of temperature and the solvent will be discussed and the most recent version, "the protein seashape model", will be introduced and extended.

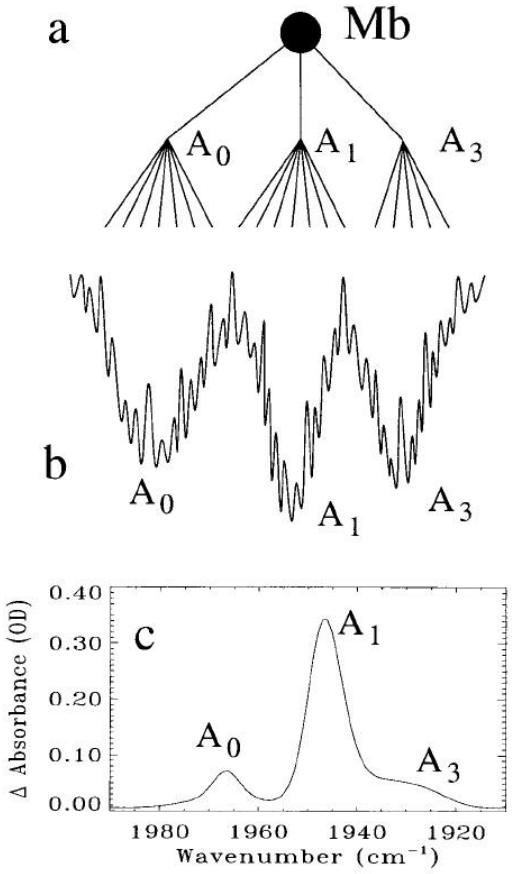

Fig. 2.20: The hierarchical ordering of several tiers is presented in the upper picture. Each tier splits into several different states, here A_0, A_1 and A_2, which were identified in the mid-infrared spectra of the carbon monoxide stretching vibrations. Each substate consists of several further minima and hence another tier in the hierarchical energy landscape [Fra01].

Mirsky and Pauling [Mir36] were the first that proposed the lack of an "uniquely defined configuration" in denaturized proteins. They suggested that upon denaturation a change of these configurations is taken place. The modern energy landscape concept has its roots in the statistical mechanics of glasses and phase transitions based on the work of Bryngelson and Wolynes [e.g. see Bry87a, Wol08]. On the other side Frauenfelder established his model of conformational substates [Fra79, Fra88] which are arranged in a hierarchical way, derived from experimental observations on myoglobin research [Fra98, Fra01, Fra02b]. He was able to identify at least four hierarchical tiers in myoglobin and related one of them to three mid-infrared C=O stretching vibrations and to functional aspects like the carbon monoxid binding as demonstrated in fig. 2.20 [Fra01]. Both approaches were brought together in joint papers of Frauenfelder and Wolynes [Fra91, Fra94, Fra99]. An example for the modern protein energy landscape arising from the macroscopic folding funnel is picturized in fig. 2.21

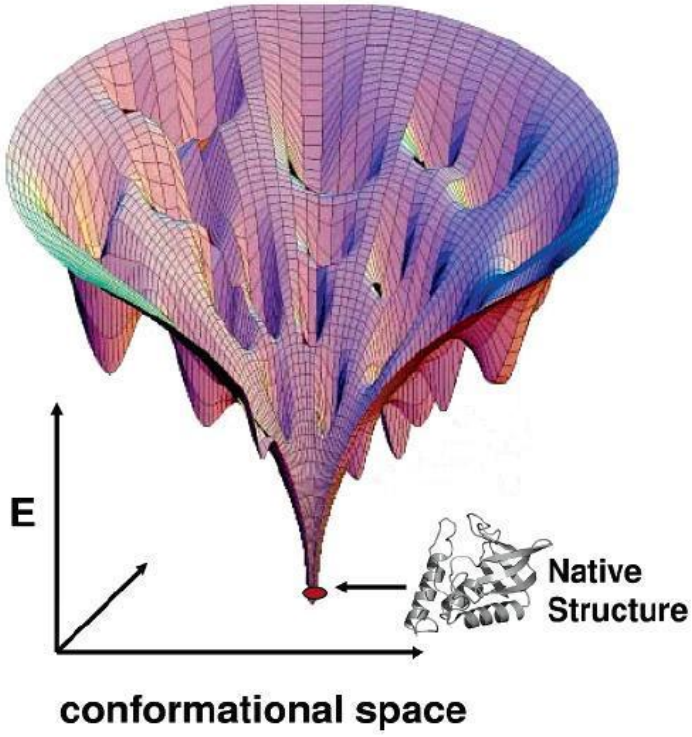

Fig. 2.21: The energy landscape of a protein in a two dimensional conformational subspace. The native state is at physiological temperatures well above the rough tip on a certain energy level. On this level several local energy minima represent possible intermediate states. The protein can be "trapped" for some time in such a local minimum before a large transition to another minima occurs. Each of these wells has no distinct minimum but several local minima again, which can be identified with the hierarchy that Frauenfelder added to the energy landscape concept of Wolynes [Hil06].

Temperature is the main factor which defines the access to other conformational states and hence the hierarchical level. Proteins move rapidly from one substate to another, changing their structure by breaking and reforming non-covalent bonds [Fra79]. When the temperature is lowered the protein is trapped in one of the tiers just able to switch between the different subconformations at the global minimum of this particular energy well. The barrier to another global conformational substate is now invincible. This trapping, but still in a non-defined minimum, is a similarity to a glassy state.

Temperature dependent experiments are one possibility to explore and analyze the energy landscape. A kink in the total and further derived conformational mean square displacement term of X-ray measurements on myoglobin, were one of the first indications of a transition around 200 K [Fra79]. Above this temperature, a strong temperature dependent increase in the

conformational mean square displacement is observed, whereas below the transition the conformational mean square displacement remains more or less constant, indicating a trapped state of the protein. The large increase suggests that the energy barriers, separating many different conformational substates, are overcome at a certain temperature and therefore a large increase of the conformational mean square displacement is observed. These results lead to the conclusion that the access to many conformational substates is non-linearly increased around 200 K, which is referred to the glass or better dynamical transition in proteins. There the thermal energy is able to "fill" the protein folding funnel up to a height where the funnel broadens a lot (fig. 2.16). Many new substates can be reached now and the possible transitions increase to a large extend. Simulations for instance on hydrated carboxy myoglobin revealed that below the dynamical transition the dynamics are mainly of harmonic nature in the trapped tiers [Ste96]. Above the dynamical transition temperature, the non-equilibrium transitions to other wells increase the anharmonicity which governs the dynamics at physiological temperatures.

Direct information about the energy landscape is gained by fs pump-probe experiments as some groups, including the group of Zewail, perform on biomolecules [Hil94, Zhu94, Ham95, Ham98, Edl02, Pal02, Zho02, Pal04]. Other possibilities to discover the energy landscape are hole burning and photon echo experiments [Lee97, Sch01, Pon06]. Their results indicate both similarities and differences to glasses as well. The time laws seem to be different in proteins compared to glasses and even compared to an unfolded protein. But the solvent around the protein seems to play an important role which underlines the differences to glass formers as well as the importance of the solvent in a protein system. The first molecular dynamics simulation which tried to probe the effect of the solvent on the collective protein motions revealed an unexpected result [Hay93]. The energy landscape of the bovine pancreatic trypsin inhibitor protein is significantly altered by the solvent compared to a simulation of the same protein in a free space. The protein is "slaved" to the dynamics of the solvent as Iben et al. first described the influence of water molecules on the glass/dynamical transition in carbonmonoxy-myoglobin [Ibe89]. Frauenfelder generalized this effetc on the whole protein dynamics and folding procedure [Fra01, Fra06]. Fenimore et al. [Fen05] stated that a well defined energy landscape is not accounting for the surrounding solvent. In his recent review on the dynamical transition of proteins Doster [Dos08] concluded that the concept of a well defined energy landscape is hence misleading. The solvent acts as a plasticizer which modulates the energy barriers and he suggested therefore the name protein seascape. We add here that not only the solvent but also the anharmonic coupling between many different

protein motions will lead to permanent modulations of the protein energy levels and hence we support a designation “seascape” to emphasize this aspect.

2.2.3. The Properties of Water

Proteins and their surrounding water molecules build together a functional relevant system which cannot be treated separately. Typical values of protein concentration in their natural environment are for cytoplasm 170 mg/ml to 350 mg/ml which corresponds to about 5 *h* to 2.5 *h*. The *h* value is the mass ratio of gram water per gram protein. This value belongs in this particular case to about 7 to 9 hydration shells around each protein and intermolecular protein distances varying between 50 Å to 40 Å [Mid96]. The most intriguing argument for the relevance of water molecules around a protein comes from hydration dependent enzyme activity measurements as presented in fig. 2.22 [Row80]. Lysozyme, which is also studied in our measurements, has no detectable enzymatic function below a hydration value of 0.2 *h* [Rup83]. At that hydration it seems to start that internal hydrogen bonds are broken and randomly replaced by protein-water hydrogen bonds faciliated by the energy provided from the surrounding water molecules which are necessary to cause fluctuations to finally perform enzymatic activity. Another reason is that proteins consist of ordered regions, namely α-helices and β-sheets, connected by disordered random coils where the main motions of the larger parts and collective motions are taken place. Water molecules interacting with these random coil sections can have a large effect on the dynamics of these parts and further on the whole protein. As it can be seen from fig. 2.22 above 0.2 *h* the activity increases with increasing hydration. For a full activation of globular proteins a typical threshold of 0.4 *h* is required which is less than a fully coverage of the protein surface [Biz02].

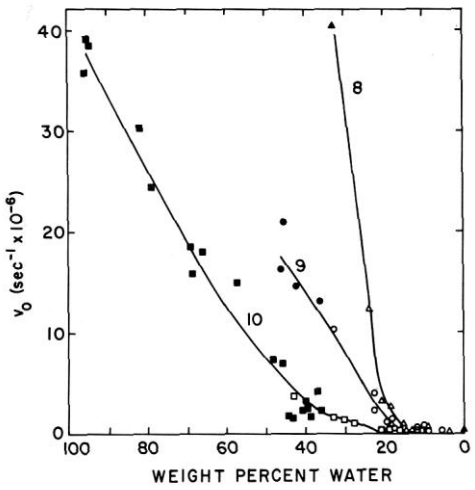

Fig. 2.22: Enzymatic activity of lysozyme as a function of the mass water content *h* at three different pH values. Open symbols belong to externally hydrated measurements and closed symbols to solvent directly added to protein powder measurements. All three graphs show no enzymatic activity below 0.2 *h* [Row80].

We will follow the discrimination from Denisov and Halle and distinguish between three categories of water [Den96]. Strongly bound internal water which almost belongs to the protein structure itself, water molecules interacting with the protein surface, which we call protein bound water or better biological water (since also non-bound water molecules in the vicinity of the protein surface have different properties) and further departed liquid bulk water which is always present in the natural surrounding as stated above for the low protein concentration in cytoplasms.

Structural water molecules are observed in X-ray protein structure analyses for instance. Another hint comes from the fact that protein crystals contain typically 40 % to 50 % water [Row80]. This very large number is maybe due to high humidity during the protein storage time. Fact is that a complete drying does not remove all water molecules, so there seems to be a type of very strongly bound water molecules within the protein structure. The actual number of these water molecules will be discussed in the following section. Afterwards the dynamic and spectroscopic properties of water and ice will be presented as a basis for the discussion of our results on biological water derived from our hydration dependent measurements in chapter 4.5.

As discussed at the end of the energy landscape section, water has the ability of promotion conformational changes, which was proven by Sundaralingam and Sekharudu from protein structure analysis and from the influence of water molecules on the secondary structure

elements [Sun89]. So there must be some water molecules which influence the protein and some water molecules which are influenced by the presence of the protein, too. These two groups do not have to be the same. A definition of these water molecules will hence be very difficult and this problem is far from being unraveled. Even a distinction between bulk water molecules and the biological water molecules is not always easy concerning the experimental techniques. For the above enzymatic activity experiments it was therefore very important that measurements involving neat organic solvents also found a loss of functionality [Mat02]. This indicates in addition that some water properties are essential for functionality and that water cannot be replaced easily by other solvents. Sometimes a second effect occurring in the protein-water system makes the interpretation of the results difficult. For the here given example Griebenow and Kilbanov [Gri95] found in their work on aqueous-organic mixtures that despite the fact that dehydration increases the tendency for a protein to denaturize and hence a loss of functionality, on the other hand the secondary structure elements are even stabilized by a lesser conformational mobility. Often just a synopsis of information gained by different techniques allows access to this complex protein-water interaction and to entangle the different contributions.

The different methods of revealing the protein-water interactions are quickly summarized by replacing the type of solvent, hydration and dehydration experiments as our approach, and cryogenic experiments freezing out the motions of the solvent and the protein. All three types and their findings point into the direction of a linkage between protein-water interactions with the protein functionality. There is a large exchange between biological water and bulk water impeding a distinction between these two types. But this exchange seems to be important for the protein motions and seem to act as mediator for conformational changes the protein undergoes [Mat02]. In our experimental procedure we have the advantage of measuring just one type of water molecules which is biological water. Due to the two compound protein-water system, we can straight forward conclude on biological water properties without using assumptions on bulk water properties, as most of the three component system measurements have done it so far. In this thesis we will stick as much as possible to a “definition” of biological water molecules exhibiting any property that is different from bulk water molecules. This seems to be more general than using the definition of water molecules directly influencing the protein dynamics.

2.2.3.1. Strongly Bound Internal Protein Water

The identification of strongly bound water molecules is mainly based on X-ray structure analyses experiments on proteins. Regarding to the large number of water molecules, exceeding often one hundred, the refinement of protein structure decreased the found number of internal water molecules significantly over the years. So many older protein structures overestimate the number of strongly bound water molecules due to indirect conclusions on hydrogen atoms on which the results are predominately based. As an example, the estimations of the number of water molecules within the structure of myoglobin decreased from 106 water molecules in 1968 to 40 just one decade later [Row80]. Spin relaxation rates gained by special nuclear magnetic resonance experiments performed on lysozyme, myoglobin, trypsin and serum albumin characterize the strongly bound water molecules by enhanced residence times from about 10^{-3} s to 10^{-9} s. The exact number of buried water molecules within the protein was not possible to be derived but the experiments indicate at least for BSA 30 to 60 water molecules per protein [Den96]. Recent molecular dynamics simulations typically assume just a handful of internal water molecules per protein which are strongly bound. In real protein crystals this number can differ because additional water molecules might fill cavities within and between the proteins leading to larger number of water molecules. But even at low hydration the number of water molecules attached to the protein surface becomes much larger than the internal ones as it will be seen in the section about the Sorption Isotherm Equation (see chapter 2.4). Therefore it can be safely assumed that for our measurements the internal strongly bound water molecules are already present even at our lowest hydration. A contribution to the number of water molecules changed by the external hydration can be excluded.

2.2.3.2. Liquid Water and Ice

Liquid water maintains the tetragonal ice structure on a statistical basis as seen in fig. 2.23. Two molecules act as a donor and two as an acceptor of the hydrogen bonds. Recent experimental data for the internal distances in liquid water derived for O-O 2.82 Å, for O-H 1.85 Å, and for H-H 2.40 Å [Tu99]. There exists no perfect tetragonal network in liquid water. The coordination number is about 4.5 indicating some “extra water molecules” within the structure disturbing the pure tetrahedral arrangement. It is assumed that due to thermal

fluctuations and dynamical reordering the ideal tetrahedron is disturbed by a permanent formation and breakage of hydrogen bonds.

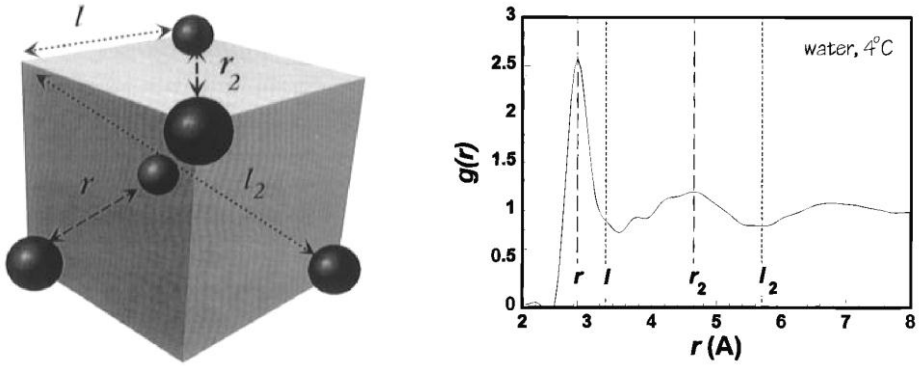

Fig. 2.23: Left: The tetrahedral structure of water visualized by highlighting the oxygen atoms of several water molecules. At 4°C, the lengths in the graph are given to r = 2.85 Å, l = 3.29 Å, r_2 = 4.65 Å, l_2 = 5.7 Å [Agm96].
Right: The radial distribution function of liquid water at 4°C going back to the work of Narten and Levy [Nar71]. The distances for the tetrahedral are indicated [Agm96].

The radial distribution function of liquid water, based on the work from Narten and Levy [Nar71], is picturized in fig. 2.23. Ice usually reveals similar peaks in the radial distribution function except for the peak in liquid water around 3.7 Å which is not occurring there [Dau97]. In ice I*h*, which is the normal hexagonal form based on the tetrahedral structure, the distances at 170 K are a little bit shorter. X-ray studies on ice derived for the hydrogen atom distance a range from1.72 Å to 1.75 Å and for the oxygen to oxygen distance a value of about 2.76 Å, whose average for the next neighbor distance is hence 0.09 Å shorter than for liquid water at 4°C [Isaacs1999, Agm96]. But this has to be taken with caution since there are at least nine different types of ice that can be formed regarding mainly on pressure and the way of cooling. An overview over the different forms of ice is given in [Book Vinogradov]. There the general differences of ice and liquid water are described via the five to six orders of magnitude lesser reorientational and translational motions in ice, which are for liquid water at 0°C about 10^{11} to 10^{12} per second, reaching the ps range. This is supported by X-ray studies which found at the ice to water transition the onset of ps motions [Eds83]. A correlation length in liquid water was found to be around 8 Å, but neutron scattering, directly sensitive to the hydrogen atoms, comes to the conclusion of about 5 Å. The rotational relaxation time of the motions mentioned above, was derived to be around 10 ps from dielectric and NMR

measurements [Eds83]. These timescales lead us to the spectroscopic properties of water and ice, which are important for the later discussion.

The far-infrared spectrum of liquid water has been measured by several groups [Dra66, Afs77, Afs78, Zel95, Ber96b, Bru05]. Two far-infrared absorption peaks of liquid water have been observed around 685 cm^{-1} and in the range from 170 cm^{-1} to 195 cm^{-1} depending on the publication. The temperature dependence, except for the two most recent ones which are discussed later [Bru05, Agm96], shows an increase in absorption with increasing temperature, which changes to the other direction around 600 cm^{-1} just shortly before the upper absorption peak. In general the temperature dependence was found to be most prominent between the two absorption peaks as seen in fig. 2.24. In this study a third liquid water absorption peak was necessary at 395 cm^{-1} to fit the absorption spectrum. This peak was widely discussed and is not accepted by everyone since also two asymmetric absorption peaks are possible to fit the spectrum properly. Also another absorption peak, assumed to be in the range of 50 cm^{-1} to 60 cm^{-1}, has not been confirmed by all groups.

Hindered translational motions and liberational motions of water are assumed in the frequency range of 50 cm^{-1} to 700 cm^{-1} from molecular dynamics simulations [Bop98, Sut98]. The upper absorption peak is widely understood to arise from the latter motion. Here hydrogen bonds are broken and restored in connection with rotational and translational motions of water molecules. This liberational band of liquid water around 685 cm^{-1} is hence "very sensitive to the degree of hydrogen bonding" [Ven01] as a study of the influence of micelles on water molecules revealed. Weaker hydrogen bonding, caused by the hindrance of the optimal configuration within micelles, has been found to cause a much lower liberational band frequency.

The origin of the lower absorption peak is still under discussion [Zas07, Sha05, Vij04, Nan00]. The question is whether (σ-stretch of) hydrogen bonds cause this peak alone, or rotational motions of water molecules, often connected to IR inactive translational motions within the water network, contribute, too. If O···O (antisymmetric) stretch vibrations [Fra76] are assumed to be the origin, the more recent temperature dependent measurements can be explained. The lower band red shifts, widens and decreases in intensity with increasing temperature [Agm96, Bru05], very similar to the general behavior of our temperature dependent protein results. Also oxygen vibrations appear in the low frequency range (30-300 cm^{-1}) often named intermolecular "restricted translational" motions. But these motions might be just normal modes of an oxygen tetrahedron when hydrogen atoms are ignored in a zoomed out model. This would fit also to the occuring intensities of these bands in both, IR

and Raman, measurements. The importance of the temperature dependent infrared absorption of the lower liquid water peak for this thesis arises from several publications indicating an involvement of dipole induced dipole interactions – derived mainly from simulations that demand an inclusion of this property to reproduce qualitatively such a peak in this frequency range [Nar92, Gui91, Mad86, Bur98, Nie93, Nan00].

Connections between the two prominent absorption peaks in water and additionally to the lower relaxation process in the GHz range are indicated from molecular dynamics simulation results. Two types of fluctuations have been found in simulations. A fast component in the range of 10 fs to 100 fs, related to the liberational motions of water molecules attributed to the 685 cm^{-1} absorption, and a slower one in the range of 1 ps to 10 ps related to structural changes within the water network. It is assumed for the latter that, in a correlated manner, water molecules undergo transitions from a stable to an unstable configuration together with the surrounding water molecules, leading to huge potential energy changes on a picosecond time scale. The large fluctuations of the fast component cause small geometrical changes which could trigger the slower process. This underlines again the connection of the dynamics of water molecules to protein dynamics, which could be the substitute for the lower frequency part in a coupled protein-water system [Tan87, Ohm88].

The refractive index was found to be rather flat in the far-infrared with a steeper increase towards lower frequencies than 200 cm^{-1}. At the lower frequency range down to the GHz range Vij et al. [Vij04] performed temperature dependent measurements showing an increasing absorption with increasing temperature as another group found, too [Ron97]. Here the measurements in the neighboring frequency range indicate a contrary temperature dependence to the more recent results in the far-infrared as discussed above.

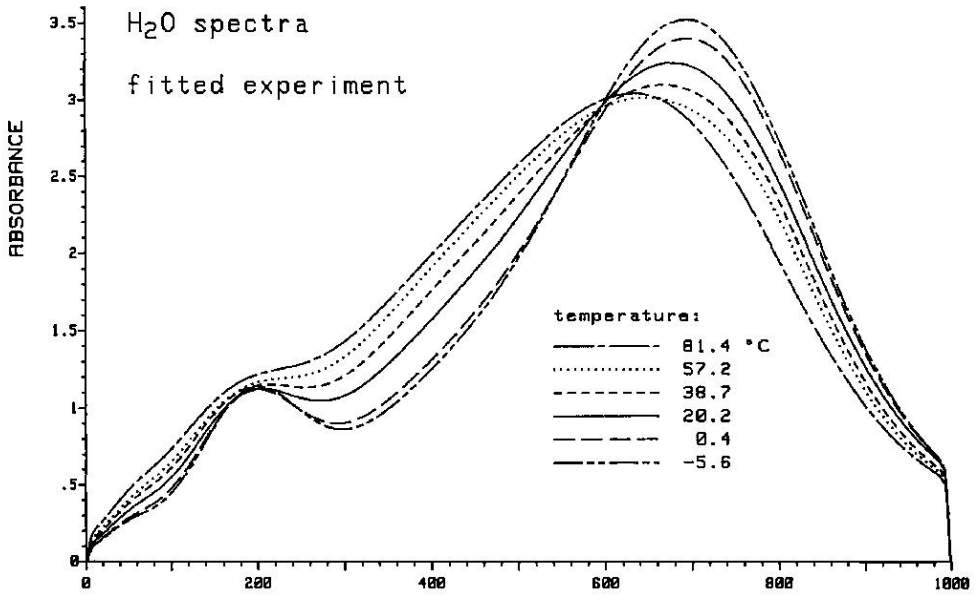

Fig. 2.24: Liquid water absorption spectra derived from a fitting procedure of the experimental data. The temperature dependence reverses around the upper absorption peak which is, compared to the lower one, much more influence by this parameter [Zel95].

Temperature dependent Raman scattering experiments even including pressure variations have been performed by Walrafen et al. [Wal86, Wal89, Wal96] which are shown in fig. 2.25. Here clearly a peak in the range of 50 cm^{-1} to 60 cm^{-1} is found. The temperature dependence changes already at 128 cm^{-1} and above the intensity sinks with increasing temperature in contradiction to the infrared absorption. At 180 cm^{-1} hydrogen bond stretching was assumed by the authors and hence the intensity decrease was explained by the breakage of hydrogen bonds with increasing temperature.

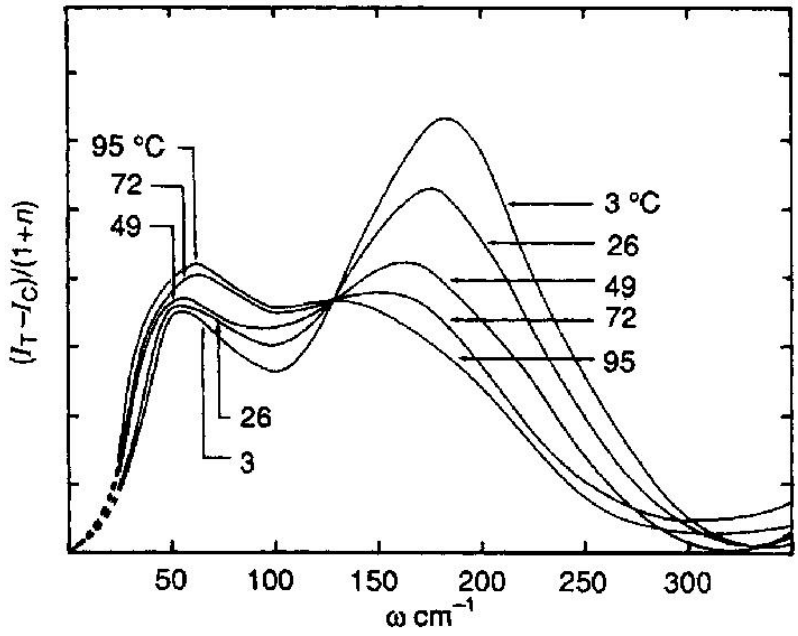

Fig. 2.25: Temperature dependent low frequency Raman spectra of liquid water going back on the work from Walrafen et al. [Wal89]. The temperature dependence changes its direction at 128 cm^{-1}. Above this point the intensity decreases with increasing temperature [Nie93].

Because we will claim an "ice-like" behavior of the biological water around proteins it is important to look on the spectroscopic properties of different states of water. First of all water vapor in the THz exhibits a minimum of nine sharp and strong absorption lines in the range from 0.2 THz to 1.45 THz [Ext89]. This is significantly different from the broad liquid absorption which shows an almost linear increasing absorption in the same range (fig. 2.24). It is assumed that translational motions and rotational motions of the water molecules are responsible for the change from the sharp gaseous water vapor lines to the broad liquid water spectrum [Kin96]. The absence of such sharp lines in our later presented spectra proofs the well withdrawn water vapor absorption by our background measurements. For a direct comparison the complex dielectric constants of liquid water and ice are shown in fig. 2.26 and fig. 2.27 respectively. First indications to their underlying molecular mechanisms have been made by the group of Gaiduk [Gai06, Zas07]. Other studies on the far-infrared absorption of ice can be found in [Ber64, Ber67, War84].

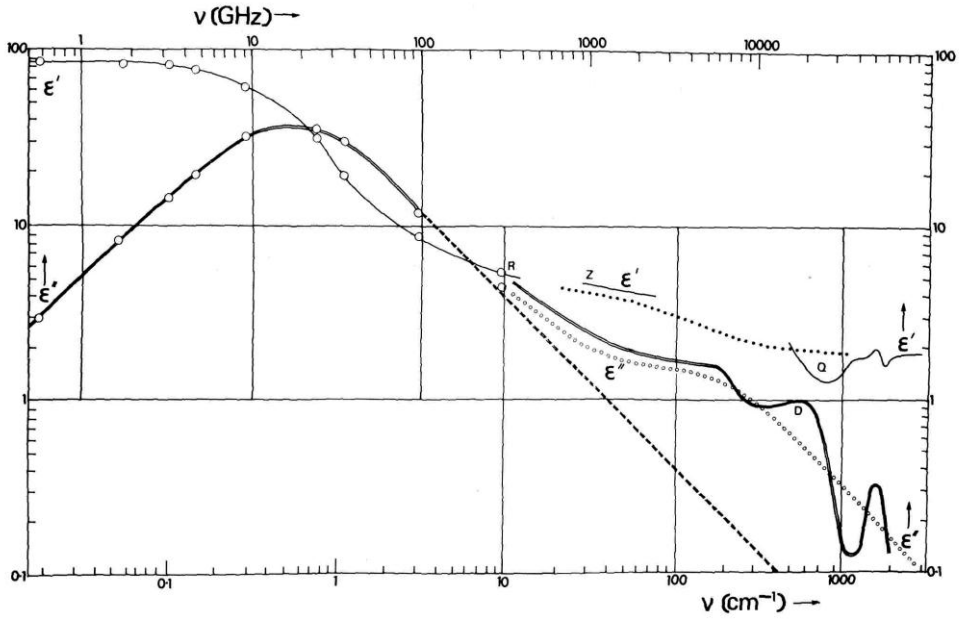

Fig. 2.26: The complex dielectric constants of liquid water. The dashed and the dotted lines represent fits for two relaxation processes assumed to occur in liquid water [Has73].

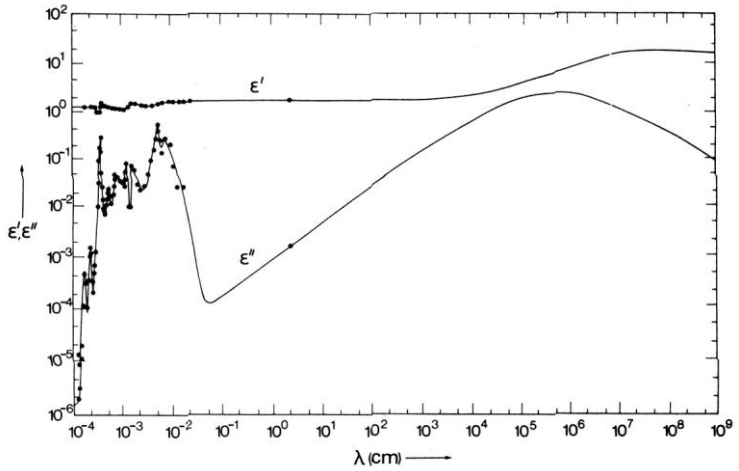

Fig. 2.27: The complex dielectric constants of ice. The dots are the measured values and the lines indicate the most likely fit [Has73].

Comparing the ε'' values of both states of water it becomes evident that in ice ε'' is in the order of 10^{-2} to 10^{-3} between 1 cm^{-1} and 800 cm^{-1}, increasing with frequency without any absorption peak. In liquid water ε'' exhibits two absorption peaks in the same range but is at least two orders of magnitude larger, which means around 10^{0} to 10^{1}. Since ε'' scales directly to the absorption coefficient, under additional consideration of *n*, as seen from formula 2.3, at least a much larger absorption of liquid water is expected compared to ice.

$$\alpha = \frac{2k\omega}{c} = \frac{\omega\mu_1\varepsilon''}{nc} \tag{2.3}$$

The refractive index of liquid water was found to be rather flat in the far-infrared with a steeper increase at lower frequencies than 200 cm^{-1} [Ber96b] as it can be also theoretically assumed, since sharp absorption peaks are missing here, which would be the origin of large changes in the refractive index. Due to the complete lack of absorption peaks in ice in this frequency range, here a flat refractive index, slightly increasing with frequency, can be safely assumed. Another strong argument for a small contribution of the refractive index is, that the refractive index changes in maximum by a factor of two or three but not as much as two orders of magnitude as ε'' in this frequency range. The here presented large decrease in far-infrared absorption for ice compared to liquid water is the basis for the first spectroscopic hint on "ice-like" properties of biological water, which we found in our hydration dependent measurements presented in chapter 4.5.

2.2.3.3. Protein Bound or Biological Water

The first hints on water molecules with different properties were derived from experiments with polymers which seem to interrupt locally the natural structure formation in the liquid water phase [Row80]. This was referred to "non-freezing water" which has been proved by several different experimental techniques and samples [Row80]. Also around proteins such a type of water was found. Crystallization of protein bound water can be suppressed far below the freezing point as for instance caliometry and IR spectroscopy experiments as well as molecular dynamics simulations revealed [Sar94, Dos86, Pac98]. Several other publications stated that hydration water in the vicinity of the protein surface has amorphous character similar to supercooled water properties [Gre95, Dos90, Gre94, Gal96, Biz00, Biz02]. But also not directly bound water molecules have different properties compared to bulk water molecules as fig. 2.28 indicates.

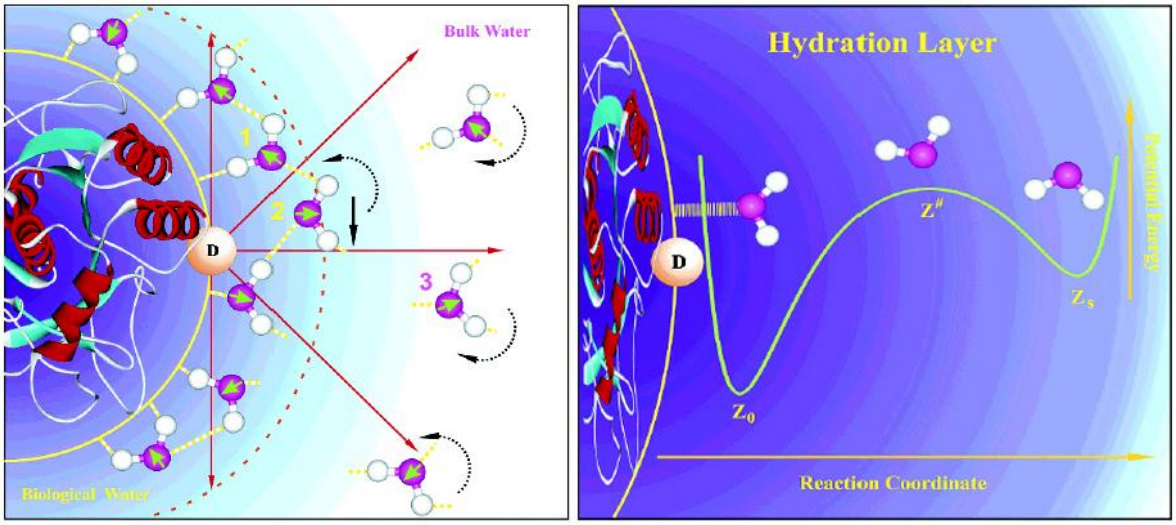

Fig. 2.28: Illustration of a protein and its hydration shell. There are two subtypes of biological water. Protein bound water (1), "quasi-free" water (2) and bulk water (3) molecules. The energy potential determining an exchange from bound to bulk water is illustrated on the right [Pal02].

One certainly different property is the diffusion constant of biological water compared to bulk water. Some results are at first glance in contradiction to each other but, if presented in a comprehensive way like here this is not the case. We will present here a microscopical picture of the biological water in the vicinity of the protein surface and derive some arguments for the following discussion of our hydration and temperature dependent measurements. Experimental results using neutron scattering on C-phycocyanin as well as dielectric permittivity and NMR measurements on hemoglobin at physiological temperatures found at least a three times lower translational diffusion coefficient of water molecules at the protein surface than for bulk water, and also in addition rotational hindering [Ste93a, Bel96]. A combined method of quasi elastic neutron scattering and molecular dynamics simulations on a prototype amino acid found further two subtypes of biological water [Rus04]. At the innermost layer water molecules with suppressed rotational behavior and with a translation diffusion behavior, like supercooled water at -10°C. The second type, within the second or third hydration layer, reveals normal bulk like rotational behavior but still lesser translational diffusion values than bulk water, but much faster than water molecules in the first hydration layer. Very early an outspread of a "protein effect" beyond the first hydration shell was suggested from X-ray and early molecular dynamics simulations on lysozyme [Hag78]. The authors claim that effects up to 6 Å or 7 Å distance from the protein surface can be assumed as supported by other groups which found effects up to 10 Å, all suggesting more than one single hydration shell around proteins [Bro89, Kom93]. Lower in energy, the biological water

molecules are capable to mediate interactions with the protein to a larger extend than just the first hydration shell.

Detailed simulations concerned the interaction behavior of water molecules with special surface groups. The simulations on lysozyme derived an increasing diffusion constant for water molecules of an ordering from attached to apolar, charged and polar sidegroups to bulk water . For water in the vicinity of apolar groups a disruption of the tetrahedral structure of water was suggested for the cause which seems to hinder the usual translational diffusion and hence decreases the mobility. For water around charged groups the strong protein-solvent interaction seems to form a rigid cage restricting the motion. Hence the solvation shell water exchange with the bulk seems to be slowed down. Water around polar groups has similar behavior like bulk water but with about half of the diffusion constant. Maybe the extend of near neighbor interactions is therefore the important difference [Bro89]. But two other simulations found no correlation of the solvent diffusion coefficient to the type of atoms at the surface which they hydrated [Kom93, Smi04]. Some common sense seems to be now that the size of the first hydration layer depends on the types of surface groups. The first hydration shell was found in the radial distribution function at 2.8 Å around polar and charged groups, stabilized by protein-water hydrogen bonds. Around non-polar groups the first water molecules are found further away around 3.4 Å stabilized by water-water hydrogen bonds. So microscopically a different behavior due to the type of bonding can be assumed, whereas the whole hydration shell will not differ significantly on the ratio of surface group types. But more a dependence on the surface topology is expected as we will see in a later chapter where this information becomes relevant.

The general spacial distribution of diffusion constants was in focus of another molecular dynamics simulation [Mak98]. The diffusion rate parallel to the surface is even higher than in bulk whereas perpendicular to the surface a much slower one is observed. The latter can be explained by restriction of the space by the protein in this direction. The increase in the parallel component would be then just due to energy conservation leading to a faster diffusion. This picture is not completely in line with an assumption of a rough protein surface that should lead to an decreased diffusion in the parallel direction, too. But the distance dependence of this effect in parallel direction would be again in line with a rough protein surface leading to an effect observed as far as 15 Å apart from the surface. No large difference between myoglobin and two types of DNA has been found suggesting basic physical principles being responsible for this diffusion behavior also supporting our results derived

from the hydration dependent measurements, which do not differ a lot with the type of protein.

This on the first sight contradictory result above, of an even faster diffusion constant for stronger bound water molecules, we will explain by matching it with some findings on liquid water behavior. There a distribution of water molecules is found regarding the number of hydrogen bonds, which can range usually from none up to five to get the average, also called coordination number, of 4.5 hydrogen bonds per water molecule as mentioned in the section before [Sci92]. Comparing directly the two largest groups with four and five hydrogen bonds on time average, revealed a faster diffusion of the latter group in a molecular dynamics study on liquid water, even if these water molecules might be interpreted to be stronger bound. The extra neighboring water molecule is assumed to offer more possible network arrangements and so decreasing the energy barrier for a breaking of a hydrogen bond [Sci92]. On this way, using some kind of "friction" within the system, the thermal energy can be capable to break a hydrogen bond even if k_BT is smaller than the sole hydrogen bond energy itself, which is an interesting result which we will also use for our later introduced new MERCEP model throughout chapter 4.6.4. In the end this can lead to an increased "defect" propagation expressed in an enhanced diffusion constant for stronger bound molecules, which is understood here in a way of an increase in numer of hydrogen bond formation, but as we will elucidate later by our molecular dynamics simulations maybe the lifetime of every hydrogen bond is decreased and this quantity has hence to be considered, too. Such a dynamical behavior is also sometimes called a "cooperative water mechanism" which distributes energy very quick among the neighboring molecules by wandering of local defects in the hydrogen network of liquid water [Ohm88, Tan87, Sci92]. The same mechanism could be possible close to proteins for instance for hydrogen bonds at the protein-water interface. But the long residence times of these water molecules do not support this indicating the above mentioned consideration of additional information for a comprehensive description of the dynamics. Residence times, defined as the average time before translation of one molecular diameter is observed, are with 7 ps shorter in bulk water than the surface water residence times of 10 ps to 50 ps [Den96]. We conclude from this discussion that bulk water seems to be more irregular formed than biological water, which would support our later discussed "ice-like" picture. Nevertheless the friction of the protein can lead to local disturbances enhancing or inducing a "defect propagation", which may lead to some regions or directions of increased diffusion for biological water molecules in the vicinity of proteins.

2.3. Analysis of Protein Spectra and Corresponding Models

Protein spectra are very complex as well as their originating sample are - at least for their underlying motions causing the absorption. Broadening of absorption peaks plays a crucial role since the protein-water system with its energy seascape provides a locally different environment for every motion and many thousands of different states are probed and averaged in an absorption spectrum like ours.

The analysis of the far-infrared spectra can be approached from either methods proved to be valid in the mid-infrared range, mainly a direct treatment of protein motions, or from methods often used in the GHz range, where on this timescale the here presented model of relaxation governs the protein behavior. The relaxation model will be closely connected to the glass or dynamical transition debate, which will be summarized here. The typical successful mid-infrared method is normal mode analysis which will be elucidated in the second section. Some normal mode based histograms of far-infrared protein spectra will be presented later. Due to the less complicated pllied force field and the smaller calculation effort compared to other methods, they provide the majority of indications for a peak assignment in proteins so far. But we will emphasize the restriction of the classic version to harmonic potentials, preventing a complete description of far-infrared protein spectra. Because in our frequency range at least the onset of anharmonicity is observed, associated often with the underlying collective protein motions. Therefore we will afterwards present, from our point of view, the method of choice, which is molecular dynamics simulations together with a discussion of other simulation techniques including anharmonic protein dynamics in the same section. We will present later molecular dynamics simulation results in particular for an analysis of hydrogen bond dynamics which we made in cooperation especially for accompanying our MERCEP model. Since all methods discussed up to that point will just gain the density of states but not the infrared absorption, several possibilities for a calculation are derived from the fluctuation-dissipation theorem, but also some approximations useful for a direct application are presented. This part is included here because the theoretical framework for the infrared spectra calculations are barely known as discussions on recent conferences revealed. This section should encourage simulations directly corresponding to the in chapter 4 presented measurements and summarizes the theoretical work on calculating far-infrared protein

absorption that has been done during this project besides the experimental achievements and results.

2.3.1. Relaxation Mechanisms in Proteins

There is a widely discussed comparison of proteins to glass formers based on both theoretical models and experimental results. The over all question is, if the similarities are caused by the same underlying physical properties in both types of materials. This would be important for a proper description of proteins and for a further insight into protein dynamics and hence the mechanism of protein functionality. Relaxation indicates anharmonic dynamical behavior and is a strong argument for the treatment of protein dynamics by simulations including anharmonicity, as an introduction to the following sections. Therefore we want to present here the theoretical description of glasses and their corresponding features like the boson peak, since we are touching with the lower frequency part in our measurements the region, where glass features might be observed.

2.3.1.1. The Dynamical Transition in Proteins

By several experimental techniques an enhanced increase in protein dynamics is found around 200 K as in fig. 2.18 already anticipated, representing an example for neutron scattering as well as Mössbauer experiments.
A transition around 180 K was found in inelastic neutron scattering experiments on hydrated myoglobin [Dos89]. Below that temperature myoglobin exhibits the properties of a harmonic solid which means that the mean square displacement is proportional to the temperature. Above, such a description is not valid anymore for the very much enhanced motion within the system that is observed. Access to new degrees of freedom is proposed to be the origin. The dynamical transition temperature correlates with the glass-like transition observed in the hydration shell of the same protein some years before by the same group [Dos86]. On the other hand no structural change of the crystal structure of myoglobin was found down to 80 K and this can be excluded as the origin. From these findings a coupling between fast local and slower collective motions was proposed, which would be also a property found in glass-like systems.

In X-ray experiments such a transition was found, too [Til92]. The anisotropic increase of the Debye-Waller factor with increasing temperature suggests the origin of intramolecular motion [Pet84] rather than static atomic or lattice disorder, since the latter two should be, in first approximation, temperature independent. Mössbauer experiments found the same transition at a slightly higher temperature at 210 K [Par81]. The mean square displacement contributions from Mössbauer results on myoglobin have been calculated [Ach02]. The authors found that about 25 % of the total sum are caused by motions below 8 cm^{-1} (1 meV), 40% from modes between 8 cm^{-1} and 24 cm^{-1} (1 meV and 3 meV), and 35 % from modes between 24 cm^{-1} and 80 cm^{-1} (3 meV and 10 meV), since modes with energies above 80 cm-1 do not contribute a lot to the total mean square displacement. This is in line with other publications which found that few low frequency modes are responsible for the majority of the total mean square displacement as stated already. Mid-infrared studies [Dos86, Ibe89] and several molecular dynamics simulations [Kar90, Smi90, Ste93b, Ste96] support these findings.
Claims that this transition can be also observed in the for us relevant THz range [Mar07] are debatable. As fig. 2.29 shows, the imaginary part of ε shows a slightly linear increasing behavior at all measured THz frequencies from 6.7 cm^{-1} to 66.7 cm^{-1} (0.2 THz to 2 THz) with a slightly larger increase setting in around 200 K. The largest increase is seen between 240 K and 250 K, well above the transition temperature found by other methods in the range from 180 K to 210 K. But this increase culminates already 10 K higher at 250 K and sinks down for the measurement at 260 K for unknown and not further discussed reasons. Also the room temperature data is completely different again as the inset shows and this inconsistency, starting just 10 K above the unusual transition around 240 K, makes these results questionable.

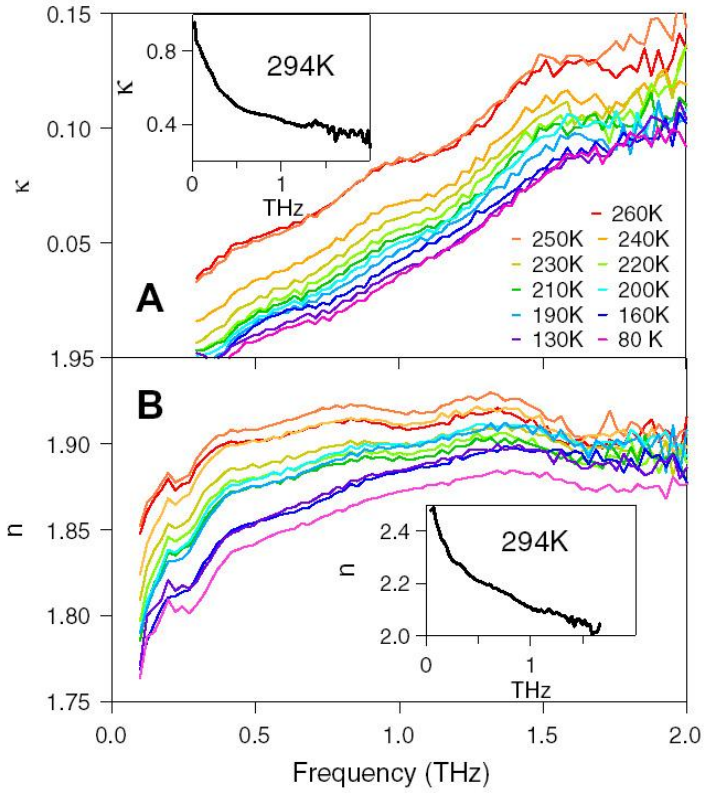

Fig. 2.29: The real and imaginary part of cytochrome C solutions measured at different temperatures. A slight increase in the imaginary part is found around 200 K and a strong increase between 240 K and 250 K. Afterwards a decrease is observed with a complete different slope at room temperature as the insets shows. The real part reveals three jumps and decreases also at 260 K. The room temperature data is again completely different [Mar07].

In addition to the transition found in dynamical properties, a transition in caliometric and thermal expansion properties characterizes a glass transition. Fig. 2.30 visualizes the typical behavior of the enthalpy *h* and the volume *V* of a liquid either forming a glass upon a fast cooling rate or a crystal. In glasses the thermal expansion coefficient

$$\alpha = \left(\frac{\partial \ln V}{\partial T} \right)_p \tag{2.4}$$

and the isobaric heat capacity

$$c_p = \left(\frac{\partial h}{\partial T} \right)_p \tag{2.5}$$

change abruptly but continuously at the glass transition temperature T_g. Calorimetric studies of hydrated lysozyme, hemoglobin and myoglobin revealed the absence of a clear glass transition in proteins [Sar94]. But the other indicator seems to be present in proteins. A transition in the unit cell parameters in the region of 200 K to 230 K was found in a temperature dependent X-ray study on crystal lysozyme [Kur95]. A determination of the Young's modulus and the linear thermal expansion coefficient on lysozyme, BSA and myoglobin indicate both a glass transition around 240K and 160K respectively [Mor85].

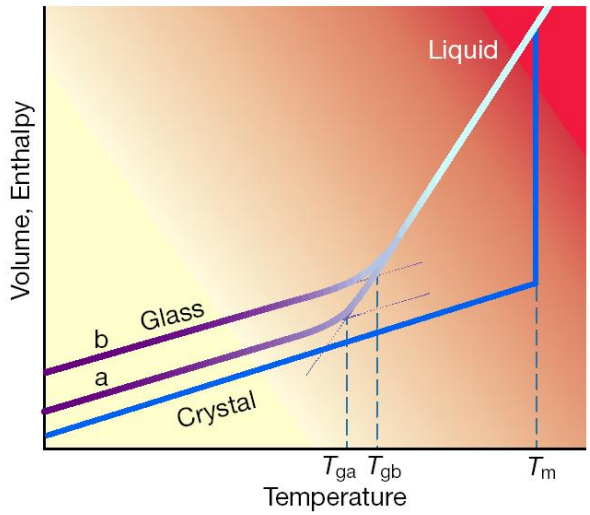

Fig. 2.30: Enthalpy and volume changes during the formation of a glass or a crystal from a liquid. Depending on the cooling rate two different glass transition temperatures (T_{ga} and T_{gb}) can be found and T_m indicates the melting temperature [Deb01].

There are also many indications, for instance on CO escape functions in myoglobin [Dos98], that the onset of protein functionality is situated right in the same temperature range underlining a correlation of this dynamical transition to functional relevant (anharmonic, collective) motions. Also the hydration of the protein is important for the transition. An absent or much less prominent transition in proteins hydrated with less than 0.2 *h* has been found by neutron scattering and NMR experiments [Fer93, Roh05, Zan99], which supports again a relation of this transition to functional relevant anharmonic motions associated with the far-infrared and THz frequency range.

2.3.1.2. Description of the Relaxation Process

Relaxation behavior of a system due to one underlying mechanism is described by the Debye relaxation. The Debye equation [Sca63] describes an exponential response of the complex dielectric constant to an external field by

$$\hat{\varepsilon}(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega) = \varepsilon_\infty + \frac{\varepsilon_S - \varepsilon_\infty}{1 + i\omega\tau} \tag{2.6}$$

where τ is the relaxation time. ε_s is the static permittivity, which is the low frequency limit of the permittivity and ε_∞ the high frequency limit respectively.
The Cole-Cole plot compares the real and imaginary part of polarizabilities or dielectric constants, which is used to visualize the underlying relaxation behavior weather one (Debye-) or several relaxation times are present in a system. Cole and Cole eliminated the frequency dependence and derived a semicircle equation:

$$\left(\varepsilon' - \varepsilon_\infty\right)^2 - \left(\varepsilon_S - \varepsilon_\infty\right)\left(\varepsilon' - \varepsilon_\infty\right) + \varepsilon''^2 = 0 \quad , \quad \varepsilon' > \varepsilon_\infty \quad , \quad \varepsilon'' > 0 \tag{2.7}$$

with

$$\text{Radius:} \quad \frac{1}{2}\left(\varepsilon_S - \varepsilon_\infty\right) \quad \text{and} \quad \text{Center:} \quad \left[\frac{1}{2}\left(\varepsilon_S + \varepsilon_\infty\right), 0\right]. \tag{2.8}$$

If the system has more than one relaxation time, the general formula for the complex dielectric constant becomes

$$\hat{\varepsilon}(\omega) = \varepsilon'(\omega) - i\varepsilon''(\omega) = \varepsilon_\infty + \frac{\varepsilon_S - \varepsilon_\infty}{(1 + i\omega\tau)^\beta} \quad , \quad 0 < \beta < 1 \tag{2.9}$$

which has the form of a skewed arc. Therefore a Cole-Cole representation gives a “one glimpse visualization” if a single Debye relaxation time or a distribution of relaxation times is describing a system properly. Usually protein relaxation cannot be described by a simple Debye relaxation after a perturbation and the Cole-Cole plot becomes a skewed arc as it was found for instance from Knab et al. [Kna06]. To model such a relaxation process in proteins, a

so called stretched exponential has to be used with an additional factor β somewhere between zero and unity, which is also applied for moddeling glassy systems.
The corresponding relaxation functions *C(t,T)* [Ang95, Bay02], which depend on the time *t* and the temperature *T* are in the Debye case

$$C(t,T) = \exp\left(-\frac{t}{\tau}\right) \tag{2.10}$$

and in the case of several relaxation times

$$C(t,T) = \sum_{i=1}^{n} \exp\left(-\frac{t}{\tau_i}\right). \tag{2.11}$$

The latter formula can be approximated by the Kohlrausch-William-Watts relaxation function conveniently depending again on one single relaxation time with the above introduced stretched exponential parameter β.

$$C(t,T) = \exp\left(-\left(\frac{t}{\tau}\right)^{\beta}\right) \tag{2.12}$$

A comparison of the frequency behavior of these two relaxation types is seen in fig. 2.31.

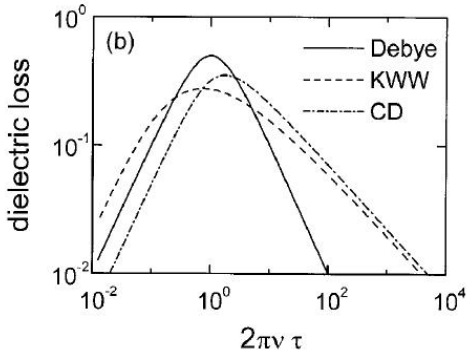

Fig. 2.31: Comparison of the Debye, Kohlrausch-William-Watts and Cole-Davies relaxation. Especially the higher frequency component is altered [Lun00].

The inverse relaxation time, occurring in all formulas above, is the so called rate coefficient *κ(T)*. In simple relaxation systems this coefficient can be described by an Arrhenius law

$$\kappa(T) = A\exp\left(-E/k_B T\right) \tag{2.13}$$

which describes the probability to escape a potential well with depth E per unit time [Sne05]. A is a temperature independent parameter. In glasses often two empirical relations are used to get a more realistic description of the rate coefficient. Both are recently used to describe the rate coefficient derived from mean square displacements of Mössbauer and neutron scattering in myoglobin [Fen02, Fen04, Fra06]. One is the Ferry relation

$$\kappa(T) = A_F \exp\left(-\left(T_F/T\right)^2\right) \tag{2.14}$$

and the other the Vogel-Fulcher-Tammann relation [Fra91]

$$\kappa(T) = A_{VFT} \exp\left[-B/(T-T_0)\right] \tag{2.15}$$

using an additional temperature independent parameter B. The Vogel temperature T_0 at which the rate coefficient diverges, has the physical meaning of an extrapolated temperature where infinite viscosity is reached during the cooling process. But this is usually not observed since the glass transition sets in before at a higher temperature [Bay02]. T_0 falls almost together with the Kauzmann temperature T_K. The Kauzmann temperature arises from an extrapolation of the entropy of supercooled liquids, which could then reach values lower than the crystalline state, known as the Kauzmann paradoxon. Therefore a transition at the Kauzmann temperature has to be proposed to avoid this unphysical behavior [Ang95, Sti01]. Interesting about that coincidence is that T_K is a thermodynamical property and T_0 has its origin in a dynamical treatment of the glass transition. So thermodynamics and kinetics are connected in glass formers described via the Adam and Gibbs theory [Bay02]. All these descriptions originally used for glasses or spin glasses are also able to model the dynamical transition in proteins and hence a relation is often assumed.

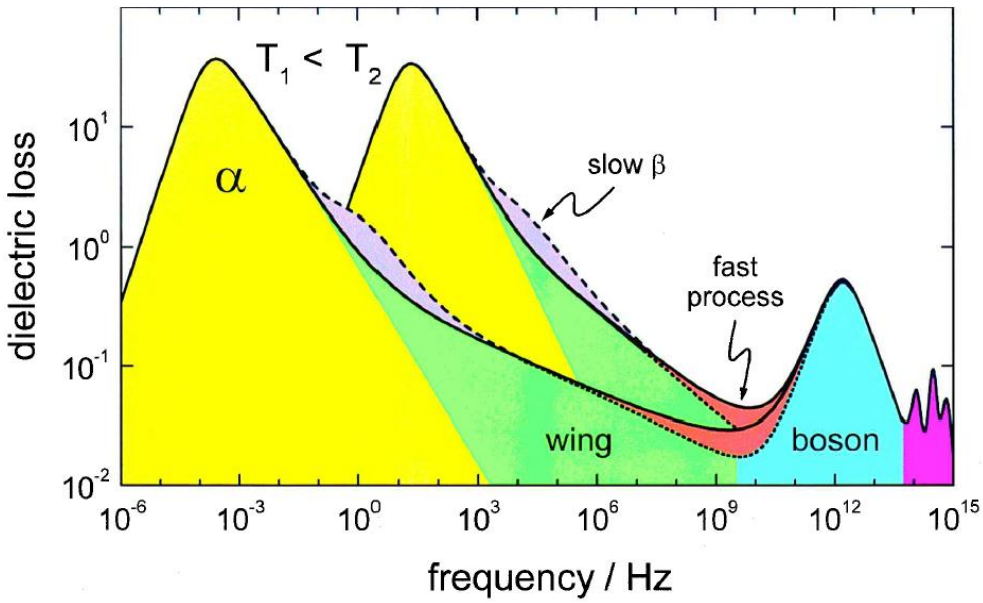

Fig. 2.32: The typical dielectric loss of glass formers over a broad frequency range: The α-relaxation (yellow), a possible slow β-relaxation above the excess wing (green), a broad minimum with the sometimes occurring fast β-process (orange), the boson peak (blue) and the infrared bands (violet). Everything below the temperature independent boson peak is shifted towards lower frequencies upon cooling, usually several orders of magnitude [Lun00].

Over the whole frequency range several relaxation regimes occur in glasses as well as in proteins. As fig. 2.32 shows, there exists a broad α-relaxation in the low frequency range shifting orders of magnitude towards lower frequencies upon cooling. In many glass formers a slow β-relaxation is observed above the high frequency side excess wing. Both processes are strongly temperature dependent as well. Approaching the for us more relevant region, the almost temperature independent boson peak is situated in THz and far-infrared range. This peak will be therefore discussed separately in the next chapter.

In between sometimes a fast β-relaxation is occurring, too. This fast process is often connected to the relaxation at much lower frequencies. Such a connection has been found for water molecules in the protein hydration shell which could trigger, with their motions in the far-infrared, the slower protein relaxation behavior. For that reason temperature dependent measurements of proteins may be able to probe properties lower in frequency. This connection is for instance described by the mode-coupling theory also used for proteins, which is controversially discussed for two decades now [Dos89, Ibe89, Dos90, Goe92, Dos98, Goe99, Lun00]. This theory also proposes the cage effect and a “rattling in a cage” motion as already mentioned in the timescale of protein motions section 2.2.1.3. Motion is just possible if the next neighbor moves, too. Therefore the neighbors are seen as a cage and

based on several experimental results this gives a microscopic insight into the liquid to glass transition. This theory is able to explain the strong increase in the observed mean square displacement properly around the dynamical transition temperature [Dos98]. Even without the questionable theoretical description of the experimental findings, additionally a sound understanding of the biological water properties seems to be essential to answer this open question.

Hence this will be here our example for glassy relaxation behavior, starting already the biological water discussion, to which we will contribute with our hydration dependent measurements. That proteins need a glassy surrounding to function, experience in protein storage simply proofs. There is a long known preserving effect of sugar molecules protecting proteins against denaturing effects from elevated temperatures or dehydration. This is used in food and pharmaceutical industry despite a lack of comprehensive understanding. Three explanations are possible. Most likely sugars may provide a glassy environment and hence conserve the protein's conformational flexibility. One argument is that the effectiveness of the preservation scales usually with the glass transition temperatures of different types of sugars. The other approach is that hydrogen bonds between protein and sugars stabilize the protein structure, when water molecules and their hydrogen bonds are removed by temperature or dehydration effects. Trapping of important water molecules within the built protein-sugar complexes might be also another possibility for an explanation of protein conservation effect [Lop00].

Molecular dynamics simulations investigated the relaxation behavior of protein-water systems. The closer water is to the protein surface, the more its dynamics differ from a single exponential relaxation in the survival time correlation function [Roc98]. At 14 Å the exponent becomes β=0.9 indicating hence a close to a single exponential relaxation behavior. But at 6 Å and 4 Å the exponent becomes 0.6 and 0.5 respectively. The later type with 4 Å is close to the first hydration shell of a protein and indicates a clearly different behavior from liquid water [Gar93, Abs96, Roc98, Biz02]. For a validation, bulk water in these studies is to be found around β=1, hence biological water in the vicinity of proteins deviates significantly in its properties from bulk water and reveals a "glass-like" behavior.

2.3.1.3. The Boson Peak in Proteins

Usually disordered materials behave similar to external excitation regardless to their chemical composition. This is related to a pileup of vibrational modes in the low frequency range around 30 cm^{-1} known as the boson peak [Tar06]. There are several publications using many different techniques which found a "boson peak like feature" in proteins under certain parameters which we want to discuss here. In supercooled liquids and glasses the boson peak is almost temperature independent as already mentioned above. In proteins this seems to be different and hydration was found to influence the boson peak in this two component protein-water system, too.

Cusack and Doster [Cus90, Dos90] derived both a low temperature peak in neutron scattering experiments on myoglobin at about 25 cm^{-1}. This peak in myoglobin was reproduced by temperature dependent molecular dynamics simulations by Smith and Steinbach et al. [Smi90, Ste91] but slightly lower in frequency, between 10 cm^{-1} and 20 cm^{-1}, than experimentally observed. This was attributed to deviations in treating the interactions of more than three molecules or in general to deviations within the force field compared to the experimental arrangement [Tar00]. But in general these low frequency modes can be still attributed to intramolecular motions. In several other neutron scattering and Mössbauer experiments such a low frequency peak has been confirmed [Die97, Ley99, Fit99, Ach02]. The temperature and hydration dependence of the boson peak is puzzling of which the influence of temperature is discussed first. In general many boson peaks are found above the so called glass transition in proteins. In glasses they just occur below T_g. Even up to physiological relevant temperatures a boson peak (or at last a distinguishable shoulder) in proteins was found with neutron scattering on myoglobin and pig liver esterase [Ley99, Kur05]. Such a peak was gained also by optical heterodyne detected Raman induced Kerr effect spectroscopy (fig. 4.10) on several peptides and proteins in solution [Gir03], by very early Raman experiments as picturized in fig. 4.9 [Bro72] and, for us most important, by synchrotron far-infrared transmission measurements (fig. 2.33) revealing a low frequency peak in low hydrated lysozyme at 19 cm^{-1}, even up to a temperature which is also probed in our experiments [Moe92].

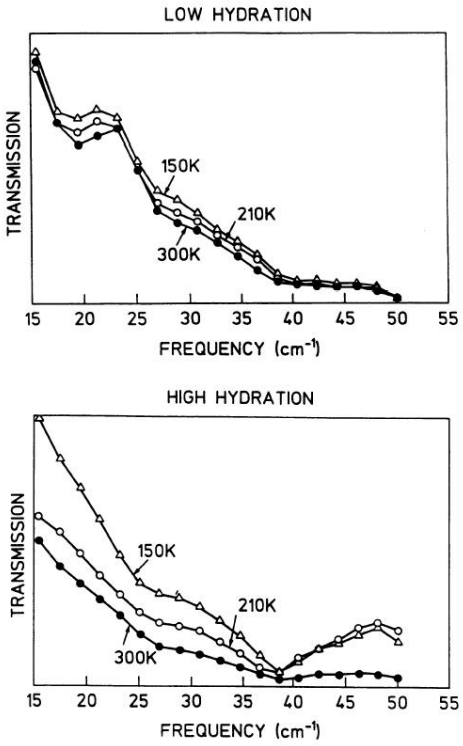

Fig. 2.33: Far-infrared transmission spectra of low and high hydrated lysozyme at different temperatures. Two features can be found around 40 cm^{-1} and at 19 cm^{-1} and 25 cm^{-1} respectively. Just the one at 19 cm^{-1} in dry lysozyme is assumed to be the boson peak [Moe92].

Corresponding molecular dynamics simulations and neutron scattering experiments have found a boson peak arising from the hydration shell around proteins which stoke up the controversy about the origin of the boson peak in proteins again [Pac98, Pac99, Biz00, Biz02]. Raman and neutron scattering experiments found a broad peak around 15 cm^{-1} and 30 cm^{-1} respectively either with decreasing temperature or hydration [Ley99]. That the boson peak gets more pronounced at lower hydration was also found from Diehl et al. [Die97] and in the experiment presented in fig. 2.33 [Moe92]. By neutron scattering on pig liver esterase Kurkal et al. [Kur05] confirmed the same dependence. The boson peak with a broad maximum around 25 cm^{-1} is clearly seen at 120 K in the structure factor and was found to be hydration dependent. The intensity decreased and the peak shifts towards higher frequencies upon increasing hydration. Hayward and Smith supported these findings because they found in experiments and in corresponding molecular dynamics simulations that a dynamical transition takes place even in a dry protein [Hay02].

So the origin of the “boson” or better low frequency peak in proteins remains unclear but in the following discussion we will demonstrate that there are significant differences in the nature of both peaks. So besides the above explanations involving the hydration shell water

molecules, collective protein motions have been at least suggested as a cause for the low frequency peak in dry or low hydrated proteins [Die97, Kur06].

2.3.1.4. Protein-Glass Comparison

Our conclusion from the above discussion is that we will name the transition in proteins around 200 K better a dynamical transition and not a glass transition. Fact is that there exist some properties which are similar to glasses. The lack of a defined energy minimum and some kind of frustration occurring within the system upon cooling leads to relaxation behavior of the dynamics. The "boson" peak seems to be unspecific concerning protein structure and folding as it is assumed for such a general property of disordered materials [Ley99]. Glass models are well suitable to describe the relaxation peaks and the rate coefficients as mentioned above, but we see clear arguments for a completely different physics underneath these observations. If all results are summarized in their entirety, our conclusion is that the name glass transition, used in the majority of all (even recent) publications, makes no sense. This theoretical result will be summarize by a synopsis of all arguments.

One major characteristics of glass forming systems is a defined glass transition temperature which is observed by a sharp decrease in the heat capacity on cooling down, when the systems is frozen in a non-equilibrium state. On warming up this becomes an even sharper increase as soon as an equilibrium state is reached again. The origin of this behavior is believed to be a crossing of the experimental and glass time scales, when the cooling/heating rate becomes faster than the relaxation time, which reaches typical values of some minutes at the crossing. Proteins lack this kind of sharp transition completely in the above described form. Due to the lack of heat capacity indicators there seems to be more an increase of motional freedom in hydrated proteins in the range of 160 K to 230 K than a real glass transition [Gre94]. This is supported by the regarded publication of Lee and Wand [Lee01] which stated that thermal excitation would be enough to explain the increase in motion that is observed around 200 K; even without a glass transition model.

Anyway strange about this increase is the timescale of relevant motions in the ps range, as shown in figure 2.34. Relaxation of glasses occurs typically in the order of seconds leaving many orders of magnitude difference between these kind of motions and this theoretical description of proteins. This could be either explained again by mode coupling theory or that

the fast motions trigger the much slower relaxation process which is therefore observed on such a different time scale. Molecular dynamics simulation results [Bec03] indicate that side chains of the protein surface, or at least close to the surface, are most affected at the transition temperature, namely protein atoms interacting strongly with the solvent. This underlines the importance of the solvent without knowledge of the exact role. By simulations of a "cold" solvent and a "warmer" protein, the dynamics of the protein is caged proofing also the important role of the solvent for protein dynamics. This close connection to the surrounding water molecules [Dos98, Set92] would be in contradiction to glass formers which show glassy behavior as an intrinsic property.

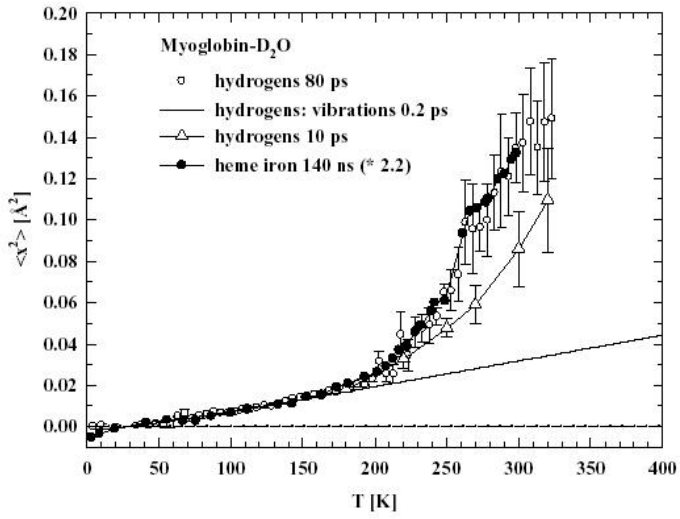

Fig. 2.34: Timescales of non-polar hydrogen atoms in myoglobin derived from neutron scattering and (rescaled) Mössbauer spectroscopy on the iron atom. Increasing motions in the ps and ns range for the hydrogen atom and iron atom displacements respectively, have been found, in contradiction to motions, many orders of magnitude away, in the time scale of seconds in glasses [Dos98].

We completely disagree with publications using the occurrence of a boson peak in proteins as an argument for a glass comparison [Gre94] and not for an argument of an indication of a just disordered state. Here a lack of knowledge about the above discussed temperature independence of boson peaks in glasses per definition has likely to be assumed. To our best knowledge this argument has never been discussed regarding to proteins and hence we would like to emphasize it here. First of all, a large hydration and temperature dependence of the peak position as well as on the intensity was observed; in clear contradiction to glasses. Secondly, the boson peak was unusually observed in many experiments above the "glass"

transition, which we summarize to the final result that this low frequency peak in proteins is by far not the boson peak as observed in glass formers, but just a broad peak in the same frequency range maybe arising from collective motions which show an intrinsic anharmonic or relaxational behavior of a wide energy distribution after an excitation.
Combined with the lack of a sharp transition in the heat capacity, we conclude that just the terminology of a dynamical transition can be used for proteins which we strongly recommend. Glass related tools to analyze protein properties might fit the experimental data, but we believe in a completely different physics behind this similarity based on the above argumentation.

2.3.2. Normal Mode Analysis and other quasi-harmonic Methods

For a harmonic or quasi-harmonic analysis of the motions within proteins, there have to be two approaches treated separately which are closely related to each other. First of all normal mode analysis uses in the classical form a purely harmonic approximation of the energy potential. Secondly, the principal component analysis which uses data derived from simulations or calculations which could include anharmonicity. Then these results are splitted up again into harmonic motions approximating the real dynamical behavior to gain the main (principal) underlying motions with their corresponding frequencies. This step is similar to the overall methodology for a normal mode analysis and is hence presented together in this section.
A molecule consisting of *N* atoms has *3N-6* internal degrees of freedom which becomes already for small proteins a very large number. The reason for a splitting of motions into normal modes is to derive completely decoupled ones. Therefore the molecule is seen as an assembly of harmonic oscillators under the influence of small displacements around their equilibrium position. Every atom is described in its own coordinate system whose origin is at the minimum of the harmonic potential well. For the kinetic energy in Cartesian coordinates the following relation holds

$$E = \frac{1}{2}\sum_{i=1}^{N} m_i \left(\dot{x}_i + \dot{y}_i + \dot{z}_i\right)^2 \tag{2.16}$$

Expansion into a series, truncated after the quadratic term leads to

$$\begin{aligned} E = E_0 &+ \sum_{i=1}^{N}\left[\left(\frac{\partial V}{\partial x_i}\right)_0 x_i + \left(\frac{\partial V}{\partial y_i}\right)_0 y_i + \left(\frac{\partial V}{\partial z_i}\right)_0 z_i\right] \\ &+ \frac{1}{2}\sum_{i=1}^{N}\left[\left(\frac{\partial^2 V}{\partial x_i^2}\right)_0 x_i^2 + \left(\frac{\partial^2 V}{\partial y_i^2}\right)_0 y_i^2 + \left(\frac{\partial^2 V}{\partial z_i^2}\right)_0 z_i^2\right] \\ &+ \frac{1}{2}\sum_{i=1}^{N}\left[\left(\frac{\partial^2 V}{\partial x_i y_i}\right)_0 x_i y_i + \left(\frac{\partial^2 V}{\partial x_i z_i}\right)_0 x_i z_i + \left(\frac{\partial^2 V}{\partial y_i z_i}\right)_0 y_i z_i\right] \end{aligned} \tag{2.17}$$

and a subsequent coordinate transformation to normal coordinates Q_i simplifies this equation significantly. E_0 can be chosen as zero and the first derivative at the origin vanishes completely, since this is exactly the energy minimum of a harmonic potential. The cross variable terms get entangled and vanish, too. In Cartesian coordinates one complex motion is described by multiple contributions from up to all *3N-6* different motions via the latter part, whereas now the dynamics is described by fewer uncoupled motions. The Hamiltonian gets the simple form of

$$\hat{H}_{vib} = \sum_{i=1}^{3N-6}\left(\frac{-\hbar}{2}\frac{\partial^2}{\partial Q_i^2} + \frac{1}{2}\lambda_i Q_i^2\right) \tag{2.18}$$

in which λ_i is the wavelength of mode *i*. The wavelength is again related to the force constant κ and the effective mass μ by

$$\lambda_i = 2\pi c\left(\frac{\kappa_i}{\mu_i}\right)^{-\frac{1}{2}} . \tag{2.19}$$

Solving the Schrödinger equation with the above Hamiltonian derives the sought for eigenvalues representing the frequency distribution of the modes, called also a histogram. The first normal mode analysis of a protein was performed on glucagon in 1982 [Tas82]. The first complete mass weighted normal mode calculation was done on bovine pancreatic trypsin inhibitor by Brooks and Karplus one year later [Bro83].

Important results like the contribution of collective protein motions to the far-infrared frequency range have been deduced from normal mode analyses. But the question about the damping of this motions and the role of the solvent have been controversially discussed and were not completely answered by this purely harmonic method [Lev85, Bro85, Kit91]. In

normal mode studies focused on bovine pancreatic trypsin inhibitor the results indicated that collective motions have a global character even in this small protein and are found in the low frequency regime below 120 cm^{-1}. Modes above 200 cm^{-1} are more localized and in between, motions involving some atoms are to be found [Go83, Nis87]. Even up to day measurements are compared to normal mode analysis, despite the lack of anharmonicity inclusion, due to the very quick and convenient calculation [Mar02, Whi03]. One example is given in fig. 4.8, where the THz slope of protein absorption is presented in combination of a corresponding histogram.

Principal component analysis entangles the complex protein dynamics in a similar way as described above. Such a quasi harmonic analysis is derived for instance from molecular dynamics simulations which include anharmonicities at the basis of the calculation. The eigenvectors of the atomic fluctuation matrix (given in formula 2.20) are the quasiharmonic modes which can be treated in the same way as normal modes.

$$A_{i,j} = \sqrt{m_i m_j}\left\langle \left(\vec{r}_i(t) - \vec{r}_i^{\,m}\right)\left(\vec{r}_j(t) - \vec{r}_j^{\,m}\right)\right\rangle \tag{2.20}$$

This mass weighted matrix is used for diagonalization and includes the mean position r^m of the atom i and the actual position r_i. The average is taken over different trajectorial time frames [Smi04]. Anharmonic motions within a single minimum will lead to the same results as normal mode analysis but if multiple minima are present, which can be assumed for proteins, the result will differ [Bro95]. For models not including a hydration shell, Teeter and Case [Tee90] proofed on crambin that the “cumulative” frequency distribution derived from both normal mode analysis and from principal component analysis is comparable. This points out the importance of hydration shell including experiments and simulations, because the anharmonicity seems to arise from the water molecules which trigger or damp the protein motions, giving them a more anharmonic character. But treated not in a cumulative way, there is still a considerable difference in the derived fluctuations gained by these two methods. But they agree qualitatively about the low frequency modes and the large contribution to the over all fluctuations [Hay95].

To encourage the next section about anharmonicity including methods, we will emphasize the difference between normal mode analysis and principal component analysis, which take place in a harmonic approximation from the beginning or at the end respectively. Fig. 2.35 shows the calculated mean square displacement of myoglobin separated into the components which are gained by normal mode analysis and principal component analysis.

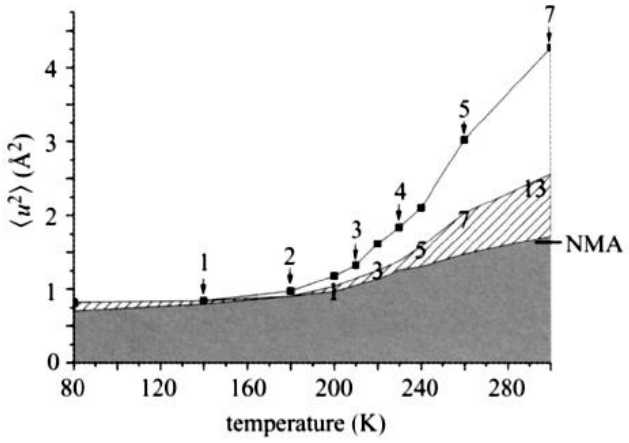

Fig. 2.35: Mean square displacement of myoglobin splitted into different principal component contributions. Dark grey are the harmonic modes whose derived value at 300 K fits well with the result from normal mode analysis using the same model system and potential function. The hatched area indicates the contribution from quasi harmonical modes and the open area the contribution from multiminima modes [Smi04].

It becomes obvious from fig. 2.35 that at physiological temperatures, as applied in our measurements, a large contribution from (few) anharmonic, collective modes to the overall motion is expected. Normal mode analysis will fail to describe the dynamics properly in our temperature range. Hence the methods on which the discussed more suitable principal component analysis can be applied to, will be presented in the following chapter.

2.4.3. Molecular Dynamics Simulations and Related Methods

As we have seen in the preceding section, regular normal mode analysis is insufficient to model the real low frequency spectrum of large biomolecules due to the highly anharmonic contribution that is expected in this range [Roi97]. Different types of simulations can be used instead, regarding to the size of the sample (including maybe a large hydration shell), the relevant timescale and the wanted properties. An overview is given in fig. 2.36.

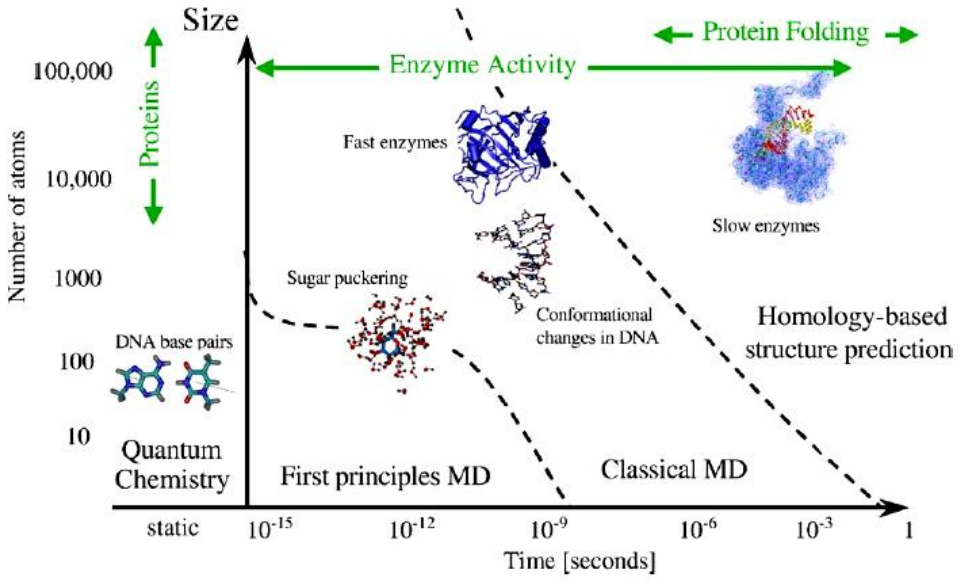

Fig. 2.36: Methods used to model processes in biomolecules under consideration of sample size and timescales [Col06].

For very short timescales and smaller systems, quantum mechanical based methods like Car-Parrinello quantum mechanical molecular dynamics simulations or quantum mechanical *ab initio* calculations can be used. Simulations using molecular mechanics are typically molecular dynamics (MD) and Monte Carlo simulations to model larger systems as proteins together with their hydration shell. The latter method has the advantage of just using the energy potential without an evaluation of a force compared to MD simulations. On the other hand with its randomized steps it is much more inefficient in exploring the many conformational substates and the large energy landscape. Before explaining the methodology of MD simulations we would like to mention an approximation, which can be used in every method, especially for gaining information about the far-infrared relevant collective protein motions. Proteins can be seen as partially rigid molecules. In a first approximation rigid secondary structure elements are connected by loops or random coils where motions relative to each other become possible. This concept of coarse graining or subspace approximation reduces the calculation effort and can be directly focused on these types of motion which are seeked for.

The process for a typical MD simulation is presented in fig. 2.37.

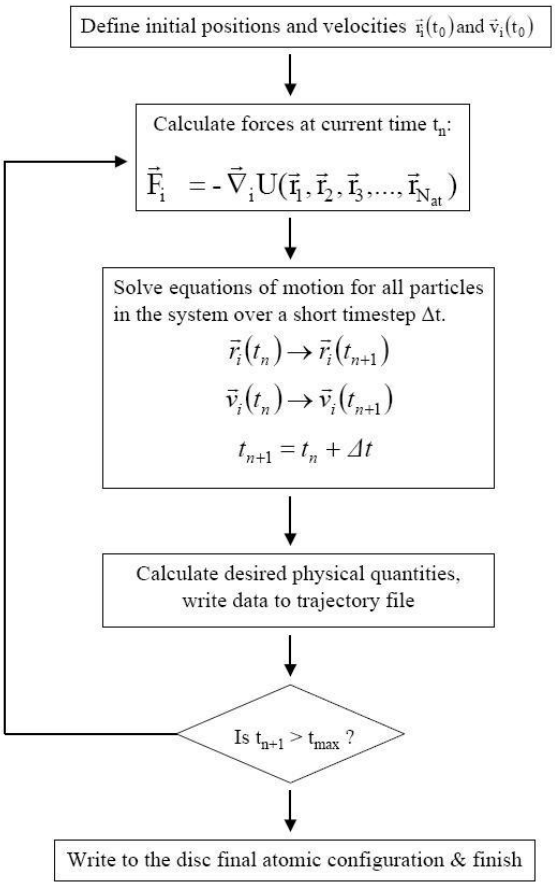

Fig. 2.37: Outline of typical molecular dynamics simulation steps.

The initial atomic positions are usually known from protein structure analysis. The corresponding velocities are derived by a Boltzmann distribution of energies at that particular temperature. The forces which apply on every atom are calculated by using the underlying force field. Then, from this starting set on, the equations of motion are solved and the simulation runs under the influence of the derived forces for the wanted time step. The end positions and velocities are saved and used as an initial state for the next step. Usually a certain equilibration time at the beginning of the simulation is not used in the later analysis, to prevent unphysical non-equilibrium fluctuations caused by the initial set of positions and velocities.

Important for close to reality simulations are the applied force fields for proteins. A typical force field has the following form.

$$V = \frac{1}{2}\sum_{bonds} K_b\left(b-b_0\right)^2 + \frac{1}{2}\sum_{\substack{bond\\angles}} K_\theta\left(\theta-\theta_0\right)^2 + \sum_{\substack{dihedral\\angles}} K_\phi\left[1+\cos\left(n\phi-\delta\right)\right] + \sum_{\substack{nonbonded\\pairs}}\left[4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r}\right)^{12}-\left(\frac{\sigma_{ij}}{r}\right)^{6}\right]+\frac{q_i q_j}{\varepsilon r}\right] \tag{2.21}$$

The first sum describes the contribution for every covalent bond, the second for the corresponding deformation angles and the third sum refers to the deformation energy due to rotation, all describing the contribution from covalently bound atom pairs. Non-bound pairs are accounted through the Lennard-Jones potential, with the first term for the core repulsive contribution and the second term for the attractive dispersion force. Electrostatic interactions are described by the last term using partial charges for instance [Nie96, Bro93].

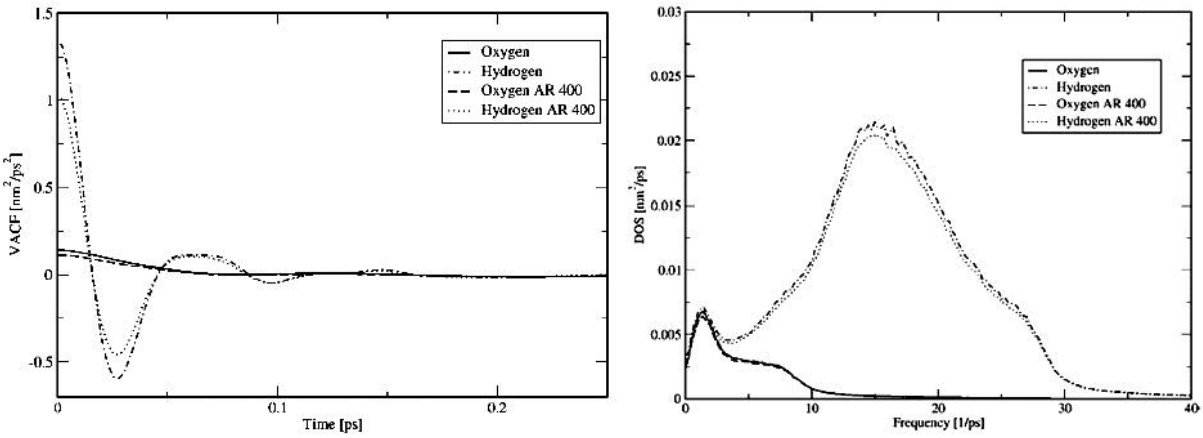

Fig. 2.38: Velocity autocorrelation functions of water splitted up into hydrogen and oxygen atom contributions derived by two different methods. The Fourier transform gives the density of states. Comparison to the absorption spectrum of water indicates especially deviations for the lower absorption peak, for which a proper modeling of induced dipoles seem to play a crucial role [Rog03].

For a further analysis of the MD simulation results usually (auto-)correlation functions are used. As seen in figure 2.38 the Fourier transform of the velocity autocorrelation function derives the power spectrum or better called frequency dependent density of states. To model the lower absorption peak in water, a full treatment of induced dipoles (see chapter 2.3.6) seems to be necessary and explains the deviations from the experimental data set in this particular simulation.

2.3.4. The Fluctuation-Dissipation Theorem

As a basis for the following chapter, focusing on the calculation of infrared intensities, the Fluctuation-Dissipation Theorem will be presented, going originally back on the work of Callen, Welton and Greene [Cal51, Gre51, Cal52, Gre52]. Written in a short and more modern form [Sca89] the Fluctuation-Dissipation Theorem is given by

$$C_m(\omega) = 2k_B T \frac{\alpha''(\omega)}{\omega}. \tag{2.22}$$

It relates the spectral density of multipolar fluctuations *C* with the imaginary part of the multipolar polarizability α. The latter was derived using the Kramers-Kronig relations [Kro26, Kra27] originally connecting the real and imaginary parts of complex dipole polarizabilities. This formula applies also for higher multipole approaches as it is used here. Prerequisites for this short version of the theorem is a linear system, which gives the possibility to get rid of the integrals, present in the original formula, by comparing the integrands on both sides as it is presented here. The importance of this theorem becomes evident when it is pointed out that the left side of the equation is determined by statistically distributed equilibrium fluctuations and the right side by non-statistical dissipative dielectric properties, which were connected for the first time here.
Kubo's important contribution was the formulation of the fluctuation-dissipation theorem via correlation functions and that he gave a more general form which holds also for quantum mechanical systems and not only for classical systems. We will use this general correlation function formalism to calculate the infrared intensities and line shapes in the following section. This theorem describes how the admittance for an external disturbance is related to the square of a Fourier component of a physical quantity which fluctuates in time in a thermal equilibrium. If linear response is assumed, the admittance for the disturbance is reduced to the calculation of the time-fluctuations of the equilibrium [Kub57]. The best explanation of the meaning and possible applications of this formula was given by Kubo himself which we want to cite here [Kub66]:

> "This theorem states a general relationship between the response of a given system to an external disturbance and the internal fluctuation of the system in the absence of the disturbance. Such a response is characterized by a response function or equivalently by

> an admittance, or an impedance. The internal fluctuation is characterized by a correlation function of relevant physical quantities of the system fluctuating in thermal equilibrium, or equivalently by their fluctuation spectra. The fluctuation-dissipation theorem can thus be used in two ways: it can predict the characteristics of the fluctuation or the noise intrinsic to the system from the known characteristics of the admittance or the impedance, or it can be used as the basic formula to derive the admittance from the analysis of thermal fluctuations of the system. The Nyquist theorem is a classical example of the first category [Nyq28], whereas, perhaps, Onsager's proof of the symmetry of kinetic coefficients is the oldest example of the second [Ons31a, Ons31b]."

We have to embed the Kubo formulas in a more general context to emphasize the prerequisites and restrictions of this formula. The Langevin equation describes Brownian motions under the influence of a "white" noise. There is no memory effect included which means that the next distortion does not depend on the former state. This can be hence described by a Markov process which is "memory free". Also some kind of memory function must be introduced, since the response of the systems is not instantaneously. Based on the work of Kubo and Mori, a generalized Langevin equation was introduced [Eva90] or the Liouville equation can be used, both leading to a memory effect containing non-Markovian processes. The latter approach, as an alternative possibility for deriving the Kubo formulation of the fluctuation-dissipation theorem, can be found for instance in [Mei64, Eva90] but this way will be not presented here, since it is completely equivalent to the less laborious approach by using the generalized Langevin equation. Along within this framework, it can be derived that the breakdown of Debye's relaxation theory above the microwave range is due to this absence of a memory effect. The imaginary part of the complex polarizability in this theory is given by

$$\alpha''(\omega) = \frac{\omega\tau\alpha'(0)}{1+\omega^2\tau^2} \tag{2.23}$$

with the maximum at the position $\omega\tau = 1$. A distribution of relaxation times gives a broader absorption peak with a larger high frequency contribution as discussed in the relaxation section. But at the high frequency limit, indicated by very short relaxation times τ, the absorption goes towards zero, unable to describe absorption peaks from the far-infrared upwards [Eva90]. This is the reason for the proper description of far-infrared absorption on

the basis of the fluctuation-dissipation theorem and not by extending relaxation models applied in the frequency range below.
There are several possibilities to express the Kubo master equations [Che63, Zwa65]. We will give here the most general one emphasizing the role of the correlation function. Before, the connection to the much earlier derived Wiener-Khintchine-Theorem will be shown, too [Wie30, Khi34, Wax54]. Wiener and Khintchine found for a stochastic (e.g. Brownian, again Markovian memory free) process that the correlation function *C* and the spectra density *S* form a Fourier transform pair [Sca89, McQ76].

$$C(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{i\omega\tau} d\omega \tag{2.24}$$

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} C(\tau) e^{-i\omega\tau} d\tau \tag{2.25}$$

This property, that the Fourier transform of a correlation function can be interpreted as a spectral density or the corresponding Fourier transform quantity, is the origin of the later used relation of the dipole moment correlation function to the infrared absorption spectrum.
The fluctuation-dissipation theorem can be given in a general form via correlation functions of two physical quantities *A* and *B* describing the complex admittance or response of the system.

$$\chi'_{BA}(\omega) = \frac{\omega}{2iE_\beta(\omega)} \int_{-\infty}^{+\infty} \left\langle \{A(0)B(t)\} \right\rangle e^{-i\omega t} dt \tag{2.26}$$

$$\chi''_{BA}(\omega) = \frac{\omega}{2E_\beta(\omega)} \int_{-\infty}^{+\infty} \left\langle \{A(0)B(t)\} \right\rangle e^{-i\omega t} dt \tag{2.27}$$

Translated to the response of the above introduced multipolar polarizations, we gain for the dipole correlation function $c_m(t)$ of the dipole moment *m* [Sca89, Eva90]

$$c_m(t) = \langle m(0)m(t) \rangle \tag{2.28}$$

whose limes towards infinity is zero. The result for the complex multipolar polarization is given by

$$\alpha(\omega) = \frac{1}{k_B T}\left[c_m(0) - i\omega \int_o^\infty c_m(t) e^{-i\omega t} dt \right] \tag{2.29}$$

which is a generalization of the Kirkwood-Fröhlich relation (here as an example for a sphere with radius *a*, since the formula is shape dependent), which is given for the frequency independent polarizability α_S and its relation to the static permittivity by

$$\alpha_s = \frac{\langle M^2 \rangle_0}{3k_B T} = 4\pi\varepsilon_0 \frac{a^3(\varepsilon_s - 1)}{(\varepsilon_s + 2)}. \tag{2.30}$$

Formula 2.27 is the general basis for the derivation of the infrared absorption calculation, concerning the intensity and the line shape presented in the following section.
Concluding remarks for this section concern the underlying linear response theory and direct applications. The theorem is just an approximation because it assumes linearity which just holds for small applied fields. It is assumed that the system is weakly driven out of equilibrium and the equilibrium fluctuations are caused by the internal density of the system [All95]. Our introduction about underlying processes can be used for direct applications. For instance molecular dynamics simulations on a simplified protein model concluded that at picosecond time scales some kind of memory function is found within the protein dynamics, since the description by a Langevin model failed. At time scales longer than some hundred picoseconds a Markovian model without a memory function was able to model the dynamics properly [Gru94]. This indicates a qualitatively change of the dynamics within proteins around the ps time scale in which we measure. And it is another example for the breakdown of the Debye theory in this frequency range as described above.
A nice visualization of the Fluctuation-Dissipation theorem (even in the Langevin form) is seen in fig. 2.39.

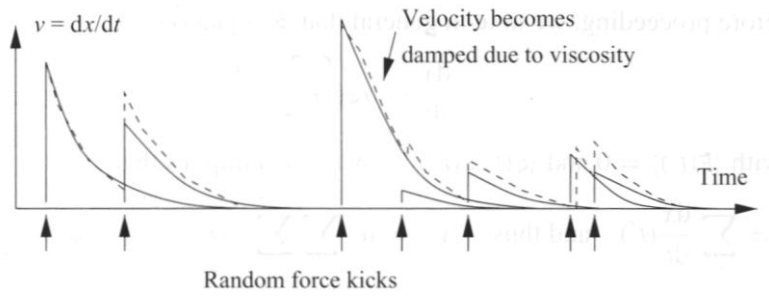

Fig. 2.39: Random forces and viscosity connected by Newtonian equations can visualize the Fluctuation-Dissipation theorem. To make sure that the average velocity of the total system is equal to the energy k_BT distributed over all degrees of freedom, a relation between the random force kicks and the damping due to viscosity must exist [Sne05].

Even in such a simplified form, it follows that for single particle diffusion the ratio of mass times diffusion constant D and the energy k_BT gives the timescale τ of dissipation.

$$\tau = \frac{mD}{k_BT} \tag{2.31}$$

This can be seen as a memory timescale after that a molecule has forgotten its former excitation. On longer timescales, the molecule properties can be then derived by simple first order Langevin dynamics towards which the dissipation-fluctuation theorem degenerates into under these conditions. Using this equation for a single albumin protein molecule (assumed diffusion of 10 μm per second) in water, a typical timescale is derived of about 1.7 ps [Sne05] confirming the relevant ps timescale, found also in the above example using more complex molecular dynamics simulations.

2.3.5. The Calculation of Infrared Spectra and Intensities

Due to the lack of any published calculated far-infrared protein spectrum including a dipole moment treatment, we want to present here in a comprehensive way the different possibilities to derive infrared intensities from any calculation or simulation of protein dynamics. Three reasons can be assumed for the complete absence of calculated far-infrared spectra of proteins. First of all the theoretical approach is not straight forward and a publication for a standard approach - a “manual” to do this - is missing. Secondly, there are many experimental

obstacles to overcome. Even the calculations for the far-infrared absorption of simple water are just recently close to the experimental spectrum. A more complex protein molecule with more than a single type of each covalent and hydrogen bonds seems to be too challenging. Also simplifications based on the symmetry of the water molecule and concerning just two sorts of atoms with their charge disproportion cannot be used for proteins. Thirdly, the lack of reproducible measurements of proteins in this frequency range under defined conditions made an optimization procedure of the eventually gained simulation results impossible, which is a prerequisite for progress in the field of simulation. These former missing measurements will be now presented in chapter 4. Regarding the first reason, we will pave the way for such calculations as much as it is possible to include this in a thesis focused on experimental physics, to encourage a calculation of protein spectra corresponding to our measurements.
So far, usually an easy approximation has been done by relating the spectral density of modes to the absorption spectra. Of course it can be assumed that the absorption at some frequency is larger if more possible motions contribute there. But neither a statement about the dipole moment change, essential for an infrared activity, can be derived on this way nor the real contribution itself. Some motions with a large fluctuation dipole moment will lead to a strong absorption, whereas others will contribute less. For a modeling and comparison to an experimental spectrum, the intensity as well as the absorption band shape has to be included as presented below, otherwise just the peak positions can be safely assigned.
The first approach for an IR intensity formula goes back to Wilson, Wells and Thorndike using dipole moment derivatives [Wil46, Tho47, Wil55]. Several versions of this formula exist of which we will give here the one for dilute gases, most suitable for applications [Wat61, Ary66].

$$A_i = \frac{N\pi}{3c^2}\left(\frac{\partial\mu}{\partial Q_i}\right)^2 \tag{2.32}$$

with A_i as the infrared band intensity, N as the Avogadro number, μ as the dipole moment and Q_i as the normal coordinates. Some authors suggest an additional prefactor [Dav63, Tai97]. In a solid or liquid phase this becomes

$$A_i(\nu) = \frac{g_i N\pi}{3c^2 n(\nu)}\left(\frac{\nu_i}{\omega_i}\right)\left(\frac{\partial\mu}{\partial Q_i}\right)^2\left(\frac{E_i^2(\nu)}{E^2(\nu)}\right) \tag{2.33}$$

of which E_i is the internal effective electrical field caused by the external field E, g_i is the degree of degeneracy of the vibration and a relation to the refractive index n is now necessarily included. The first bracket is the above mentioned prefactor, using the ratio of the absorption frequency to the harmonic angular frequency [Per82, DeL92]. For a direct application the last formula is usually unsuitable because the internal field is no property which can be measured or simulated properly. For instance infrared intensities of the protein model system N-methylacetamide, using *ab initio* calculations of the dipole moment derivative, have been gained by applying therefore the gaseous state formula 2.32. A good agreement with the amide I and amide II band in the experimental spectra has been found by this method [Che85, Ban92].

A more exact formula for the calculation of infrared and Raman intensities is based on the extensive work of Gordon [Gor63, Gor64a, Gor64b, Gor65, Gor68a, Gor68b, Gor71]. The formula can be used either for a quantum mechanical calculation or in the classical limit, depending on the underlying correlation functions. Different versions and normalizations of this formula exist. We will present here our favorite one as a combined version of several publications, which can be seen as an extension of the version from Souaille and Smith that led to good experimental results for water [Kra80, Mad86, Nie93, Sou96, Dut09].

$$\alpha(\omega) = \frac{4\pi^2\omega}{3\hbar cVn(\omega)D(\omega)}\left(1-\exp\left(-\beta\hbar\omega\right)\right)C(\omega) \tag{2.34}$$

$$C(\omega) = \int_{-\infty}^{+\infty} \frac{1}{2\pi}\exp\left(-i\omega t\right)\left\langle M(0)\cdot M(t)\right\rangle dt \tag{2.35}$$

Here $\alpha(\omega)$ is the absorption coefficient, $\beta = 1/k_BT$, c the speed of light, V the sample volume, $n(\omega)$ the refractive index and $D(\omega)$ an internal field correction. The term in front of the time correlation function $C(\omega)$ is a population factor. It was stated that above 400 cm^{-1} the exponential term can be neglected for measurements at physiological or colder temperatures which becomes evident in the explanation of the balance condition discussed below [Nie93]. The lengthy complete derivation of this formula, based on the dissipation-fluctuation theorem and (directly for the infrared absorption) on Fermi's golden rule of perturbation theory, can be found in the Gordon and McQuarrie publications cited above. We will focus more on the explanations necessary for an application of this formula to derive (far-)infrared spectra. Formula 2.34 is just valid under two prerequisites. First of all the scattering and the incident infrared wave vectors have to be negligible small. Secondly, the frequency dependent

difference between the external and internal Maxwell electric field should be so small so that $D(\omega)$ can be assumed to be unity over the complete frequency range. But this is a restriction especially for calculating polar liquids like water [Bra80]. Otherwise some additional experiments or calculations to derive $D(\omega)$ are necessary. The refractive index is also an often unknown quantity and hence set to unity. But for the broad protein absorption peaks just minor frequency dependent changes of n are expected. Hence a rescaling of the absolute absorption to any known experimental value should be a proper way to model the complete experimental far-infrared spectra of proteins.

$C(\omega)$ can be interpreted as the quantum mechanical line shape function. But if the dipole moment is calculated classically, as it is usually done in MD simulations, then the Fourier transformation is the classical approximation of the band shape. Just if a quantum-mechanical calculation for the dipole moment motion is used the true spectrum is obtained. Via the detailed balance correction the formula can be symmetrized (which is important for mathematical reasons as demonstrated below) and then quantum effects are minimized for calculations in the classical limit [Bra80].

The quantum mechanical correlation function $C(\omega)$ fulfills the balance condition [Kra80]

$$C(-\omega) = \exp(-\beta\hbar\omega)\, C(\omega) \quad . \tag{2.36}$$

The meaning of this condition is that the absorption of a photon, described by $C(\omega)$, and the emission, corresponding to $C(-\omega)$, are connected by the Boltzmann probabilities of these states [for further details see Ber81a, Ber81b, Ber83a]. The classical limit $C_{cl}(\omega)$ (which is $\hbar \rightarrow 0$ and means infinitesimal small distances between energy levels) has hence the different symmetry of

$$C_{cl}(-\omega) = C_{cl}(\omega) \quad . \tag{2.37}$$

To regain the original relation, the classical time correlation function is identified with the real part of the quantum mechanical counterpart, leading to the following transformation [Sou96]:

$$C_{cl}(\omega) \rightarrow \frac{2}{1+\exp(-\beta\hbar\omega)} C_{cl}(\omega) \, . \tag{2.38}$$

For the absorption calculated by the classic dipole moment correlation function, as it is usually derived from molecular dynamics simulations, we finally get for our extended version

$$\alpha(\omega) = \frac{4\pi^2\omega}{3\hbar cVn(\omega)D(\omega)} \tanh\left(\frac{\beta\hbar\omega}{2}\right) C_{cl}(\omega) \tag{2.39}$$

in which two parameters (*n*,*D*) might be set to unity over the whole frequency range for a practical handling as described above. This is the so called harmonic approximation, which is derived by replacing the Kubo-transformed quantum correlation function by a classical one [Che08]. As suitable for infrared spectroscopy, absorption caused by rotational and vibrational motions is modeled (which is most relevant for our frequency range) but (to mention it just for completeness) no absorption arising from translational motions, due to the invariance of the total dipole moment *M* under translation [Kra80].
Now we can focus our attention on the calculation of the determining dipole moment correlation function by MD simulations. The total dipole moment *M* has two contributions arising from the permanent *p* and induced *d* dipoles.

$$\vec{M} = \vec{p} + \vec{d} \tag{2.40}$$

The total dipole correlation function [Sou96]

$$\left\langle \vec{M}(0) \cdot \vec{M}(t) \right\rangle \tag{2.41}$$

for calculating $C_{cl}(\omega)$ above, is a sum of three terms. The permanent dipole correlation function is given by

$$C_p(t) = \sum_{i,j} \left\langle \vec{p}_i(0) \cdot \vec{p}_j(t) \right\rangle \tag{2.42}$$

with the permanent dipole moment *p* related to the charge disproportion at distance *r* via

$$\vec{p} = \Delta q \cdot \vec{r} \ . \tag{2.43}$$

The induced dipole correlation function

$$C_d(t) = \sum_{i,j} \left\langle \vec{d}_i(0) \cdot \vec{d}_j(t) \right\rangle \tag{2.44}$$

is calculated from the sum of the induced dipole momenta μ by

$$\vec{d} = \sum_i \vec{\mu}_i \quad . \tag{2.45}$$

The third interference term of the correlation function is given by

$$C_{pd}(t) = \sum_{i,j} \left[\left\langle \vec{p}_i(0) \cdot \vec{d}_j(t) \right\rangle + \left\langle \vec{d}_i(0) \cdot \vec{p}_j(t) \right\rangle \right]. \tag{2.46}$$

This leads us to the challenging calculation of the induced dipole moment in large systems. A general formula is given for instance by [Man06]

$$\vec{\mu}_i = \alpha_i \left(\vec{E}_i + \sum_{j \neq i} \hat{T}_{ij} \vec{\mu}_j \right). \tag{2.47}$$

α is the polarizability of the atom *i* which is here assumed to be isotropic. The first term describes the electric field caused by permanent charges or polar groups at the position of atom *i*. One possible derivation is

$$\vec{E}_i = \sum_j \left(\frac{\left(4\nu_{ij}^3 - 3\nu_{ij}^4\right) q_j \vec{r}_{ij}}{r_{ij}^3} \right) . \tag{2.48}$$

v is here a parameter which avoids unphysical induced dipole momenta at small distances [Sou96]. The second term is the electric field arising from induced dipoles. The tensor *T* can be chosen in different ways and is usually gained by an iterative process. A non-iterative method is found in [Gui91] for instance. Also an Ewald Sum approximation can be used to cut down the calculation time. Therefore on the outside of a cutoff sphere an infinite dielectric medium is assumed with a corresponding dielectric constant. In general all tensors *T* have to be calculated in every step of the simulation by calculating μ_i in the n^{th} step and using it as μ_j in the $(n+1)^{st}$ step. This is usually done until μ_i shows a convergence. Such an induced dipole

moment inclusion requires a large additional calculation effort. But especially in the far-infrared region this is necessary and is emphasized by many authors [Tse87, Pal98], in particular for water or water containing systems like proteins as discussed in section 2.2.3.2 with the references given there. The lower absorption peak of water arises from the induced dipole momenta and is completely absent in the spectrum calculated just from the permanent dipoles. The contribution from the induced dipoles, as derived from MD simulations, was even larger than the one from the permanent ones in the low frequency range around the peak at 200 cm^{-1} [Sou96]. This is also an important result for proteins since the hydration shell has to be included for a completely functional protein and because our later presented new theoretical model of the temperature dependent protein absorption is also based on induced dipole momenta.

Another application of the absorption calculation via dipole moment correlation functions, derived from the dissipation-fluctuation theorem, can be used to analyze the absorption band shape. Even the coupling between mid-infrared stretching modes of hydrogen bonded molecules in inert solvents can be analyzed as done from Robertson [Bra80]. This would be a nice application for the secondary structure analysis and for joining mid- and far-infrared measurements. The weak low frequency hydrogen bond stretching mode of a X-H···Y substance will be under the influence of fast stochastic Brownian motions of the surrounding solvent. This leads to a strongly fluctuating amplitude of the hydrogen bond stretching mode, which the associated covalent bond will follow adiabatically. From gas phase experiments it is known that the hydrogen bond stretching mode and the covalent stretching mode are strongly coupled. As a proof for that we only need to mention any of the above cited publications which stated an observation of secondary structure influences on covalent bonds in the mid-infrared. The randomness in the fluctuations of the hydrogen bond will lead to broadening of the covalent absorption band shape. The reason for this is the loss of phase coherence of the covalent stretching mode under maintenance of the amplitude. The excited state after an absorption will hence relax much slower to the ground state than the phase relaxation which governs the band shape. This phase relaxation was found the take place on a subpicosecond time scale and hence in our relevant frequency range. Information about the coupling between these modes, the hydrogen bond stretching frequencies and the relaxation process can be gained using the above derived method. The here presented model fulfilled also the tests on validity on absorption band shape changes upon deuteration, solvent exchange and temperature variations [Bra80] which is an encouraging result for further calculations hopefully also on larger systems like proteins.

2.3.6. Advanced MD Simulations: Polarizable Force Fields and the Inclusion of Induced Dipoles

We have emphasized the importance of induced dipole momenta in the preceding section. An inclusion of polarizabilities in the underlying models is therefore inevitable. But even in 2007 it was stated from Palmo [Pal07] that current standard force fields just provide a fixed charge disproportion and do not include polarizabilities. At least atomic polarizabilities have been implemented in modern force fields but it is obvious that polarization is more a group property than an atomic one. Since polarization including simulations are in general not widely used and still have to overcome some obstacles, we want to refer here to some introductory publications in order of our recommendation without giving all details [Yu05, Adc06, Ma00].

In polarizability including simulations on N-methylacetamide and other small amides the induced dipole moment is responsible for up to the half of the total dipole moment [Din95]. The induced dipole momenta were found not to scale simply with the polarizability of the molecule. In solvation free energy calculations, a coupling between the permanent and induced dipoles has been found to contribute significantly, a result which could have been also theoretically derived from the permanent and induced dipoles connecting cross-term in our above presented calculation. This interaction was largest in the protein model system N-methyacetamide compared to the other amides. This range of contribution is supported by simulations of several protein model systems and proteins itself [Dan97, Kam01, Cie01b, Kam02, Har05]. The origin of the induced dipole moment may be due to shape anisotropies. Large polarization arises from the wide inhomogeneous surface of proteins due to side chain reorientation. The protein interior is assumed to contribute little [Wei07]. This is supported by joint experimental and simulation studies focusing on the total dipole moment in biomolecules and their origin [Tak96, Erm98, Tak99, Tak00].

Older estimations of the dipole moment and the different contributions are mainly based on the Kirkwood-Shumaker effect. The main statement is that dipole moment fluctuations or contributions from sidgroups, associated with partially charged or polar atoms, could be the origin of a considerable large total dipole moment reaching even the main contribution [Kir52]. Due to the large number of charged and highly polar groups in biomolecules a “counterion atmosphere” is attracted which reveals a very high polarizability. This should lead to a significant induced dipole moment in biomolecules likely exceeding the static contribution [Ort68, Ort70, Ant89]. This effect is discussed even decades later because dipole

momenta and polarizabilities are frequency dependent and are hence an origin of errors which must be assumed in some publications [Por96, Por97]. But in general, a large permanent and induced dipole moment can be both expected in proteins. The O=C-N-H peptide group has a large dipole moment of 4 D to 5 D in the condensed protein medium, whereas liquid water in the bulk phase has about 3 D supporting the significance dipole momenta in proteins. The protein core is densely packed with a fraction of 0.74, which is comparable to the face-centered cubic lattice of hard spheres, and the α-helix has its peptide dipoles almost perfectly aligned [Sim03]. This is all leading to an enhancement of induced dipole momenta in proteins. Our best argument for a large contribution to the total dipole moment in proteins, arising from induced dipole momenta which we will use later in our new theory, comes from well known and proven water properties. The dipole moment of gas phase water is 1.85 D, whereas in liquid it is about 2.9 ± 0.6 D. This increase is just due to polarization effects caused by a denser packing of the same kind of molecules having a permanent dipole moment and a large polarizability. This polarization and packing effect is assumed be sizable in proteins, too [Yu05].

So in summary it can be stated that in densely packed proteins the contribution of induced dipole momenta is too large to be ignored and has to be certainly taken into account for analyzing protein far-infrared absorption spectra. We will introduce in chapter 4 a new model regarding such an inclusion and relate the induced dipole momenta directly to temperature effects observed in our measurements which has never been done in such a form for proteins before.

2.4. Quantitive Analysis of Protein Hydration: The Sorption Isotherm Equation

The Sorption Isotherm Equation allows one to derive the mass water content of a molecule from the external relative humidity which is identified with the so called water activity. Together with some properties of our measurements we are able to quantify our hydration dependence giving us also the possibility for extensive comparison possibilities to results derived with other techniques.

We will use here the extended model of the B.E.T. (Brunauer, Emmett, Teller) approach using three quantities to describe the sorption isotherm of a substance, going back to the work of

Bone et al. [Bon77, Gas77, Ede80]. The needed quantities are known from theoretical calculations and several experimental methods [Tow95, Car80] for three of our measured proteins bovine serum albumin, lysozyme and β-lactoglobulin. This theory has been even recently proven again to be valid and was several times reviewed [Smi02, Gre95, Bon96, Sha98, Zha04]. In food industry this is a well established theoretical and experimental method mainly using there a five parameter based theory from Watt and D'Arcy [Arc70, Igl84, Igl86] which are unfortunately not known for our measured proteins. Since the sorption isotherm shape is originally an experimental result, a more parameter including system just leads to a better theoretical description. But the graph from our here used approach from Bone et al. is already almost perfectly describing the experimental data sets as it can be seen in fig. 2.40 on the left for our measured protein BSA [Gas77] and more recently in fig. 2.41 on also measured β-lactoglobulin. It fits also very well to the results from Kuntz and Kauzmann which were pioneers in the field of protein hydration [Kun74]. The sigmoid shape of the sorption isotherm is typical for glassy polymers [Wor00], which can be seen as an additional similarity for the discussion in the last section of our protein-glass comparison.

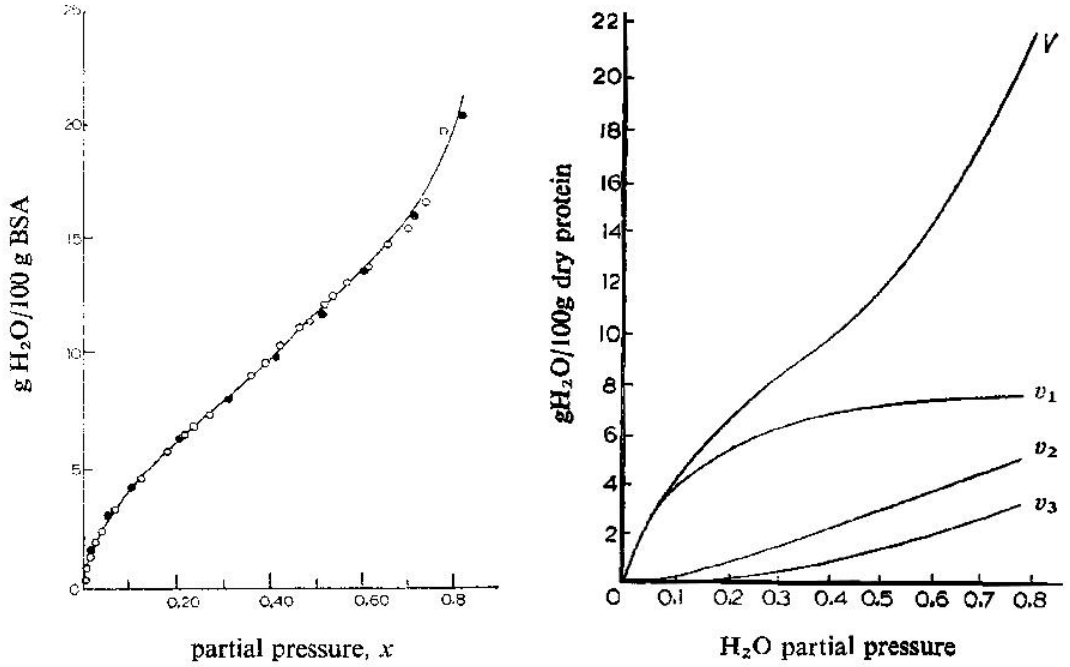

Fig. 2.40: Left: Comparison of the theoretical derived sorption isotherm of BSA to measurements provided from two different groups (open and filled spots) [Gas77].
Right: The sorption isotherm for BSA. Here the filling of the first three hydration shells is included for a visualization of the mathematical derivation of the Sorption Isotherm Equation [Bon77].

We will give a quickly summarized mathematical derivation of the Sorption Isotherm Equation [Gas77, Bon77] for two reasons. First of all, the original publications lack of a detailed description and therefore a short derivation is presented here by using colors, quickly indicating the important steps without the need of further extensive explanations. And we will also derive some additional formulas on the way for our own version of calculation, concerning mainly different hydration shells. Secondly, we found a typo in the final Sorption Isotherm Equation from 1977, which seems to cause confusion up to day. Upon derivation of the Sorption Isotherm we will prove the correct formula.
Starting under the assumption that water is adsorbed on identical surface adsorption sites of a molecule in the following way. There is a first hydration layer which has its own adsorption behavior expressed in the factor *ab*. The factor is written in this form only for mathematical reasons. Subsequent hydration layers exist of different adsorption properties to the first one, but still different to the bulk phase. Here an infinite number of hydration shells is assumed which contribute to the adsorption less and less upon distance to the surface, depending on the preceding hydration shell by the factor *b*.

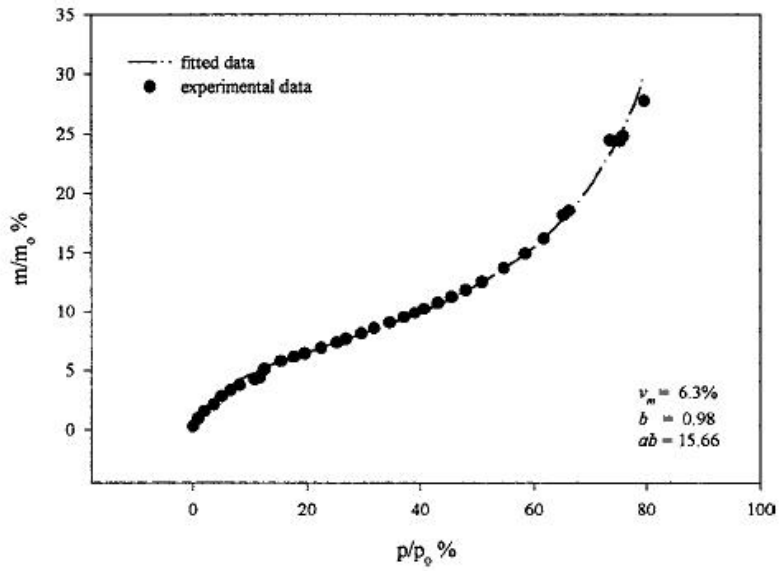

Fig. 2.41: The Sorption Isotherm of β-lactoglobulin. Comparison of the theoretical prediction to experimental results using the here presented formula [Lop00].

This leads us to a simplified Grand Partition function for one sorption site being expressed in a form close to an infinite geometrical series, which we extend here in an identity approach to

$$f(x) = a + abx + a(bx)^2 + a(bx)^3 + ... + 1 - a \quad . \tag{2.49}$$

The parameter x is the activity of the sorbed gas which is identified, without loss of generality, with the relative partial pressure and hence the relative humidity. Written down in this form now, the first part is an infinite geometrical series

$$f(x) = a\sum_{k=0}^{\infty}(bx)^k + 1 - a \tag{2.50}$$

which can be expressed in a much shorter form, like it is found in textbooks, leading to the following result for the whole expression:

$$f(x) = \frac{1 + (ab - b)x}{1 - bx} \quad . \tag{2.51}$$

It was derived that the total adsorption ν_{tot}, which is given in quantities of the mass water content h, is connected to the monolayer sorption capacity of the first shell ν_m by

$$\nu_{tot} = \nu_m x \frac{\partial}{\partial x} \ln f(x) = \nu_m \frac{x \cdot f'(x)}{f(x)} \quad . \tag{2.52}$$

The denominator is already derived in formula 2.51 and the nominator is given then as

$$x \cdot f'(x) = \frac{abx}{(1 - bx)^2} \quad , \tag{2.53}$$

finally leading to the “error free” Sorption Isotherm Equation:

$$\nu_{tot} = \frac{\nu_m abx}{(1 - bx)(1 + (ab - b)x)} \tag{2.54}$$

For a calculation of the water content ν_n of the n^{th} hydration shell, which we will need later, we will provide the not published ansatz and a short indirect proof of the formula. The ansatz starts with a splitting of the total absorption and continuous then with an expansion to infinite series for the amount of water contained in all hydration shells above the seeked for n^{th}

hydration shell as seen in formula 2.55. We want to emphasize that here a transformation of the variables from *n* to *n-1* was made before, to start the summation with 0 for which the limes are known.

$$\nu_{tot} = \sum_{k=0}^{n-1} \nu_k + \nu_n + \sum_{k=n+1}^{\infty} \nu_k = \sum_{k=0}^{n-1} \nu_k + \nu_n + \left(\sum_{k=0}^{\infty} \nu_k - \sum_{k=0}^{n} \nu_k \right) \tag{2.55}$$

Back-transformation of *n-1* to *n* gives the result, which can be indirectly proven by substituting the resulting formula 2.56 below and the Sorption Isotherm equation formula 2.54 in equation 2.55 above.

$$\nu_n = (1-bx)(bx)^{n-1} \nu_{tot} \tag{2.56}$$

The water content of the first three hydration shells is presented in fig. 2.40 on the right in comparison to the total adsorption of the Sorption Isotherm. They all rise towards their limes $\nu_{n,max}$ which can be easily calculated by setting the relative humidity *x* to 100 %. Then the filling ratio for each shell can be derived as we will calculate it in the discussion section of the hydration dependent measurements.

What is referred to be a hydration shell is an open question since experimental proofs lack due to the above argumentation of the unknown outspread of biological water around proteins. But this brings us back to one of the first assumptions for the derivation of identical adsorption sites. This is definitely not the case for a protein and the needed parameters can be seen as averaged values of different sites to model the experimental isotherm shape. A modern interpretation of the several hydration shells is the hydration process of different sites like charged, polar and apolar groups on the protein surface as well as their influence on the following water layers in the vicinity of these sites.

The sigmoidal shape of the isotherms indicates two transition points in the hydration of proteins. One very likely interpretation among others [Gre95, Row80, Tow95, Smi02] is that up to the first transition mainly charged and highly polar groups are hydrated and the increase in water content is steep. Afterwards backbone and less polar groups are hydrated but the increase of water content is smaller than before. At the last transition point weak polar and even apolar sites are hydrated finishing the complete surface coverage. In addition multilayers are assumed to form and explain the rapid increase in water content again. Another interpretation of the “plateau-like” mid part is the formation of multilayers which is discussed

and repudiated by Rupley et al. for instance [Row80]. We also see that this might explain the slower water sorption up to that point but an explanation of the steep increase at higher hydrations lacks in this model, so that we prefer the first definition of a hydration shell to be surface group dependent. This makes sense to us since we assume a parallel formation of multilayers around charged and polar surface groups to the formation of the first hydration shell around apolar ones. So the theoretical model including several hydration shells in the sorption isotherm agrees not only with the experimental but also with the microscopic picture of protein hydration.

Different phases in protein hydration are supported also by several other techniques. Caliometric heat capacity measurements of lysozyme support an additional transition point around 0.27 *h* [Yan79], between the two transitions found in the sorption isotherms, corresponding to 0.05 *h* and 0.38 *h*. A comparison of heat capaticity, enzymatic activity, mid-infrared spectroscopy and diamagnetic susceptibility in a paper also focused on lysozyme confirmed changes at 0.05/0.07 *h*, 0.2/0.27 *h* and 0.38 *h* [Car80]. The questionable one in the middle, which will play an important role in the later discussion, is just observed by the first method. Enzymatic activity and diamagnetic susceptibility agree on a transition lower than 0.2 *h*. Three different optical measurements are mainly flat without a clear sign of a transition. We summarize that there are more indications for no transition in the hydration of lysozyme at 0.27 *h*.

A hysteresis is discussed for sorption and desorption without a clear tendency as some groups observed and some groups did not observed this in their experiments [She83, Row80, Bry87b, Sha98]. If there is one, the deviation has been in most cases rather small within the percent range [Bol89, Gre95]. The temperature dependence of the sorption isotherm has also been tested many times with the result of a sinking water content of some percent with an temperature increase of 20 K [Hno61, Mor85, Igl84, Igl86, Sha98, Wor00]. But since we observed just a small hydration dependence in our measurements this influence on the temperature dependent measurements can be safely neglected.

3. Experimental Section

In this chapter we want to present the technical details of our measurements and developed setups; starting with the description of the used spectrometer with the attached infrared microscope and the employed far-infrared detector. Five different setups have been developed for protein measurements of which especially the atmosphere cell setup and the corresponding free standing protein film sample preparation are described in detail. For the three liquid cell setups interesting results and their discussion are just quickly summarized already within this experimental section, reserving the main chapter of "results and discussion" for the much more relevant atmosphere cell measurements.

We have seen in the preceding chapter that temperature and the surrounding water molecules are the most important influencing parameters for protein dynamics and hence the spectra that we measure. For a relation to the functionality it is important to investigate proteins in a state as close as possible to the natural one. For the relevant proteins such a state is usually formed in an aqueous surrounding within the physiological temperature range. But the strong far-infrared liquid water absorption complicates these measurements, in addition to low intensity sources and the large temperature caused background in this frequency range. To develope a liquid cell for protein measurements using FTIR or THz time domain spectrometer is unsuccessfully attempted for more than three decades now and we also did just partially succeed in constructing such a liquid cell.

In the THz range Xu, Plaxco and Allen [Xu06a, Xu06b, Xu06c] managed to measure proteins in solution by applying a free electron laser as a strong source to overcome the dominant liquid water absorption. Within a very small frequency range (81 cm^{-1} to 96 cm^{-1}) this can be also achieved by using a special developed far-infrared laser [Ber05]. This advance is just related to the usage of stronger sources but not to an advanced development of a liquid cell which allows measurements without the need for large facilities or (to small frequency ranges restricted) laser systems.

Three liquid cell setups and an atmosphere cell with exchangable windows have been constructed during this thesis covering the whole frequency range from the THz to the mid-infrared as shown in fig. 3.1.

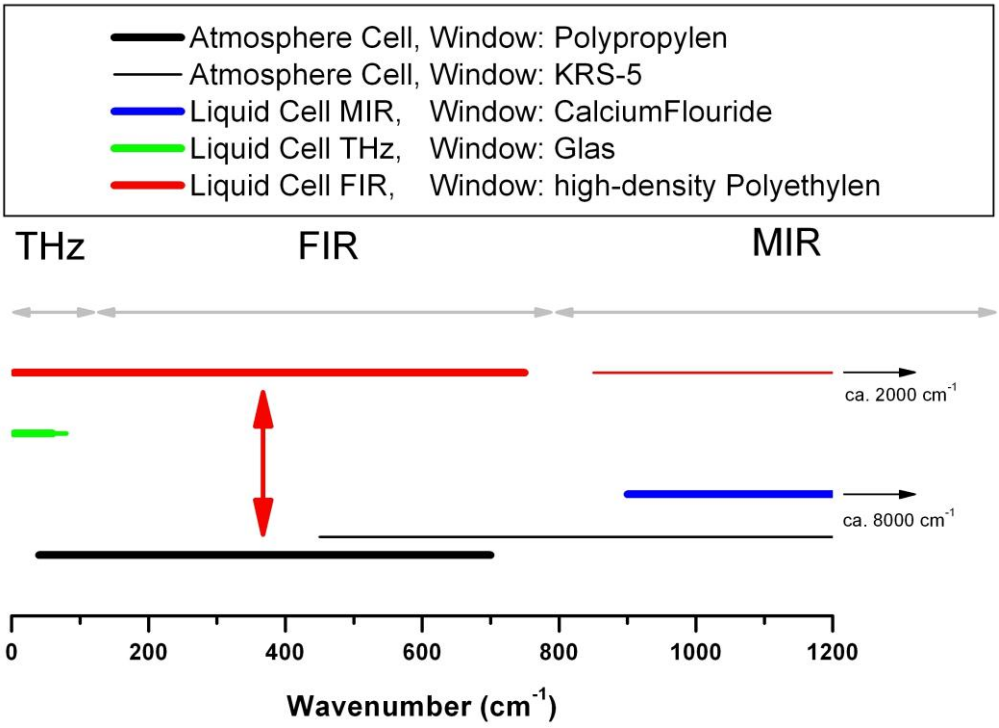

Fig. 3.1: Overview over the developed setups and their corresponding frequency ranges. The small lines indicate prototype setups or possible extensions of developed setups which have been constructed for an application of different window materials.

For the THz and far-infrared range it was achieved to penetrate the very thin solution layer completely but the differences between water and protein solutions were too small to be resolved using these liquid cell setups. The mid-infrared liquid cell was a development to support our cooperation partners in Palestine. Here it was possible to distinguish between water and protein solution mainly due to the available high intensity sources. For the above reasons the three setups are presented just in a very short manner.

Another idea to probe proteins in a close to the natural state emerged to be anyway the much better approach. The fourth developed setup measures proteins under external hydration. This leads to an only two type containing system of protein and biological water since just water adsorbed to the protein surface is measured and no liquid (bulk) water is present. In the liquid cell setups there is a three type system probed consisting of protein, biological water and liquid (bulk) water. In absorption measurements just an indirect conclusion on the contribution of each of these three compounds is possible. Whereas in the atmosphere cell just two contributions are found and, most important, the biological water amount can be varied and hence the contribution can be directly identified. For the atmosphere cell development we

focussed on the far-infrared region due to the importance of collective motions as discussed in the preceding chapter and due the lack of measurements in this frequency range.
To handle temperature changes is challenging in both setup types. For liquid cells a proper sealing under temperature variation has to be guarantied and in the atmosphere cell the occurrence of cracks within the protein film has to be prevented. The technical solutions will be presented in the following sections, too.

3.1. The Fourier-Transform Spectrometer

For all measurements presented here a Bruker IFS66 v/s FTIR spectrometer has been used (Fig. 3.2).

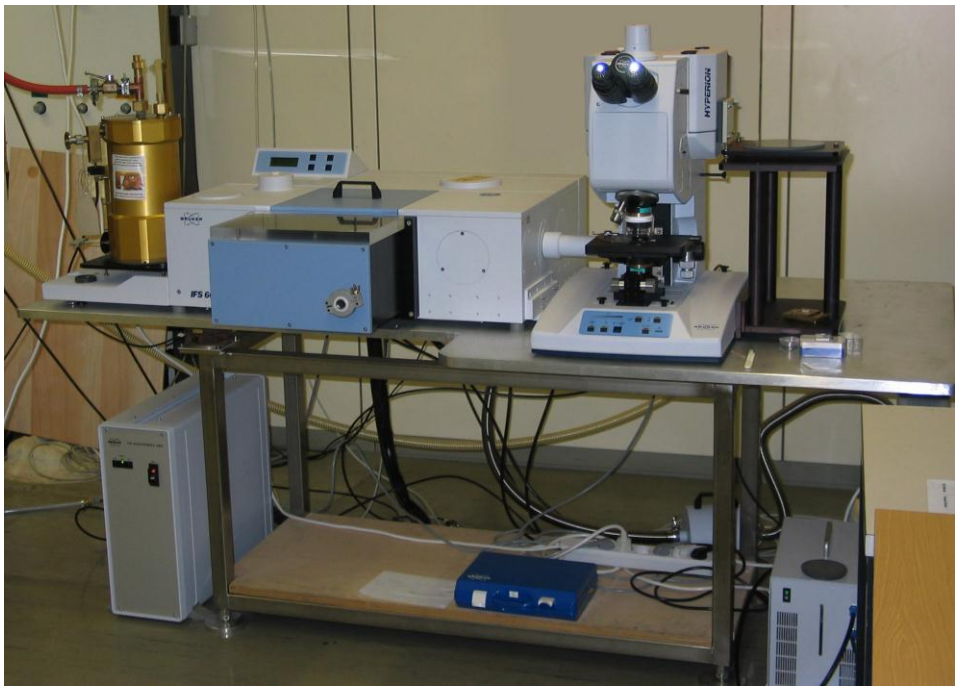

Fig. 3.2: Bruker IFS66 v/s FTIR spectrometer with the attached Hyperion infrared microscope and the helium cooled far-infrared bolometer.

The far-infrared liquid cell measurements were performed in the front compartment of the spectrometer whereas all atmosphere cell measurements used the infrared microscope. The spectrometer is based on a Michelson interferometer with one fixed and one moving mirror that combine the beams again behind the interferometer regarding to the path difference of the two beams. The measured interferogram, which is intensity versus mirror position, is then

Fourier transformed to the transmission spectrum. A schematic drawing of the internal construction is given in fig. 3.3.

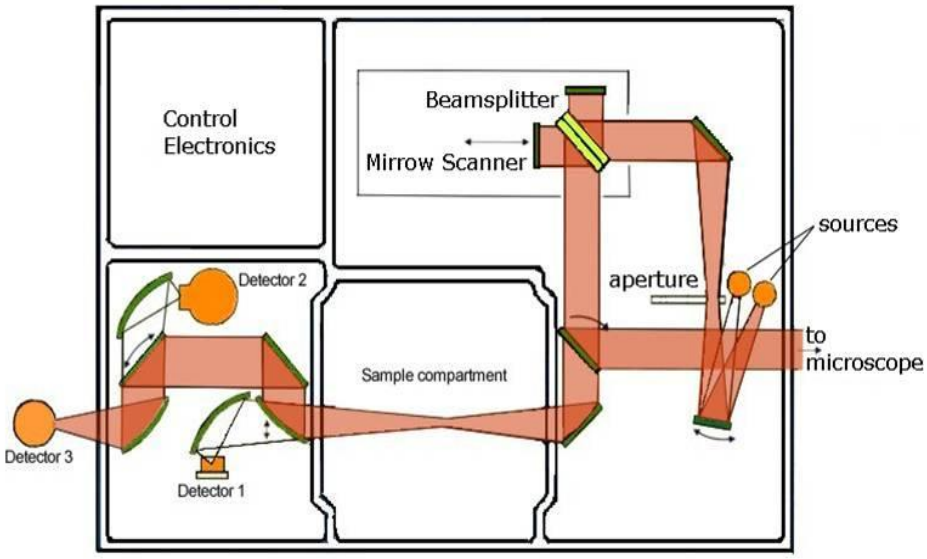

Fig. 3.3: Schematic drawing of the internal path of light and the important components. For far-infrared investigations a Mylar beamsplitter is used and a helium cooled bolometer is attached at detector position 3 for measurements on the far-infrared liquid cell situated in the sample compartment. The atmosphere cell measurements have been performed under the infrared microscope.

By applying different light sources, beamsplitters and detectors the spectrometer covers the whole range down from 30000 cm^{-1} to about 10 cm^{-1} with highest possible resolution of 0.25 cm^{-1}. For our far-infrared measurements a mercury arc lamb, a 6 μm Mylar beamsplitter and a bolometer have been used leading to a frequency range of about 40 cm^{-1} to 690 cm^{-1}. No sharp features have been found in proteins and the actual resolution in our measurements is 2 cm^{-1} which was necessary to cut down scanning time. On the other hand we decided to average over many scans to reach on this way a high stability during each of the many hours enduring measurements. The absorption $A(\nu)$ was calculated via

$$A(\nu) = -\ln\left(\frac{I_S(\nu)}{I_B(\nu)}\right) \tag{3.1}$$

from the ratio of the sample and background intensities $I_S(\nu)$ and $I_B(\nu)$. Each intensity was averaged over 250 scans each. For preventing water vapor absorption the spectrometer can be either evacuated or dry nitrogen purged.

For the atmosphere cell measurements the infrared Hyperion microscope was used (Fig. 3.4).

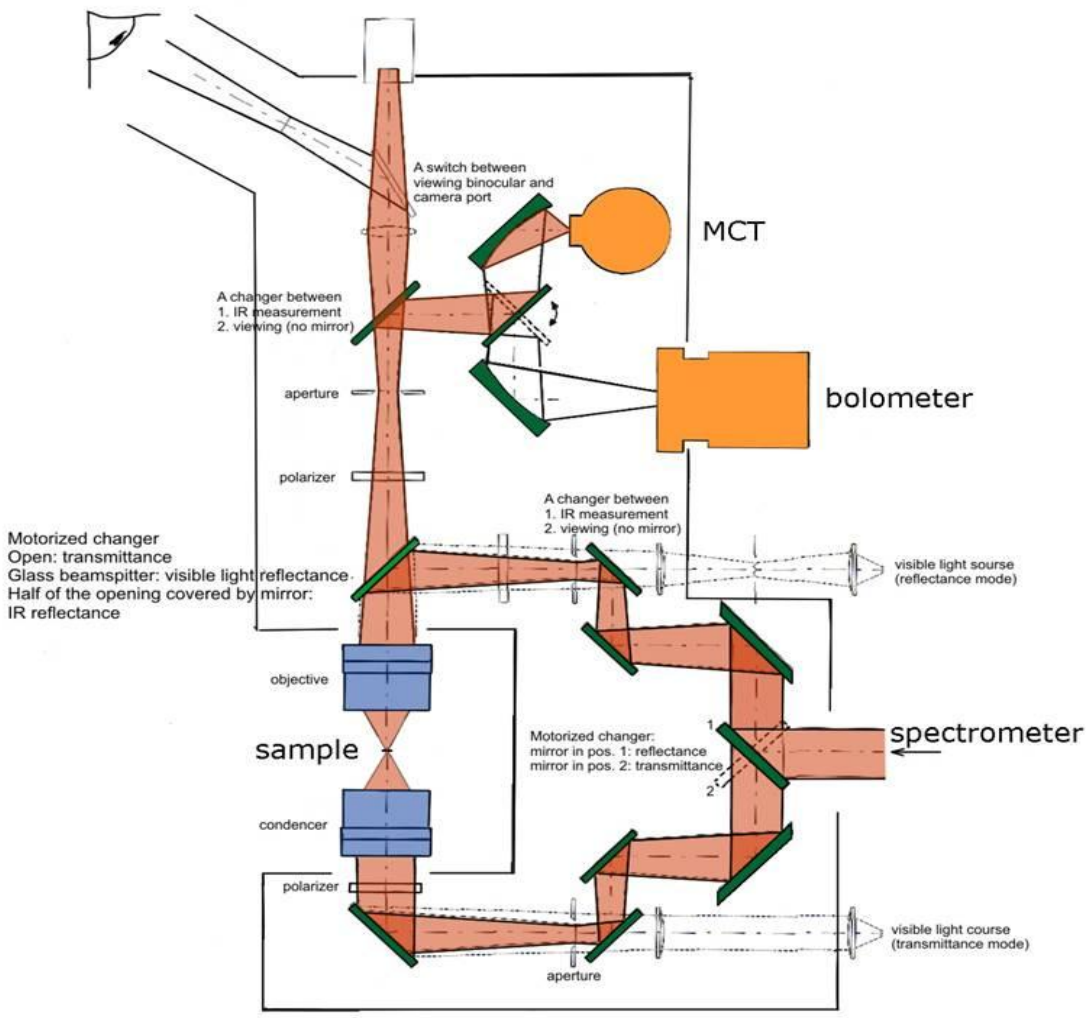

Fig. 3.4: Schematic drawing of the Bruker Hyperion microscope. The beam from the FTIR spectrometer is used as a light source. For transmission measurements the sample is in a confocal arrangement which means that the sample is in the focus of both the condenser below and of the objective above. The light is focused and collected via a Cassegrain mirror arrangement due to the lack of suitable infrared lenses.

The beam from the FTIR spectrometer is used as a light source. Due to the lack of suitable lenses in the infrared the focusing is achieved by a Cassegrain mirror arrangement as known from telescopes. The whole microscope is purged with dry nitrogen to prevent the water vapor absorption for the same reason as mentioned above. But in addition we covered the complete

arrangement, including the microscope and the atmosphere cell, by another large nitrogen purged cover to avoid even slight fluctuations caused by water vapor, therefore further increasing the accuracy which was in particular important to enable the undoubtedly observation of the small hydration effect on the far-infrared protein spectra. By a self made polypropylene window at the spectrometer output we extended the frequency range of the microscope down to about 40 cm^{-1}. The weak intensity at the lower edge gives then a reliable frequency range of 60 cm^{-1} to 690 cm^{-1} by using the microscope, but the monotonic decrease of protein absorption with decreasing frequency can be easily followed down to 40 cm^{-1}. The spot size can be varied by several apertures down to 20 µm whereas we used the largest possible spot size of 250 µm. The magnification of the microscope is 15 and the focal length is 24 mm. We attached a video camera on top of the microscope to get to exactly the same measuring spot again after a background measurement. This significantly increased the precision of our measurements concerning the relative changes caused by the two external parameters temperature and hydration.

Fig. 3.5: The helium cooled bolometer which is used for far-infrared measurements.

We used a helium cooled bolometer as a far-infrared detector (fig. 3.5). The temperature dependence of the silicon based semiconductor resistance of the sensitive element is capable

to detect very small changes of the far-infrared radiation with a large dynamical range. Two filters for the frequency range of 30 cm^{-1} to 700 cm^{-1} and for 10 cm^{-1} to 100 cm^{-1} can be chosen. But the microscope restricts the lower edge anyway to the above mentioned 40 cm^{-1} for our measurements.

3.2. Developed Setups and Techniques

To measure proteins, the experimental conditions as well as the sample preparation have to be taken into account. This is the reason for special setups that we developed for different frequency regions. For liquid cell measurements the sample preparation is a simple protein solution but therefore the sealing of the cell, especially under temperature variation, becomes a difficult task. Also the solution thickness is crucial due to the high absorption. To work around the large liquid water absorption an atmosphere cell using external hydration was constructed leading in addition to the advantage of a distinguishable two compound system as described above. But to gain a quick and complete external hydration a development of a new sample preparation technique was necessary. Thin protein films are also used in the adjacent THz frequency range but to our best knowledge a film without a supporting substrate underneath was never achieved so far. The free standing pressed pellet technique, which is also widely used in the THz, is not able to be hydrated externally and often polymers are added as a matrix in which the proteins are embedded in. The occurrence of cracks during the film preparation and during temperature dependent measurements have been the major obstacles to overcome in achieving free standing protein samples. In section 3.2.2 we will give a comprehensive description of the developed sample preparation process. The important results have been gained by the atmosphere cell setup and hence a detailed description of this setup is given in the last section. Prior to this section the liquid cell setups are presented. The results and the discussion of the far-infrared liquid cell measurements are already included in this chapter as mentioned above.

3.2.1. The Liquid Cell Setups and Important Results

Three liquid cell setups have been developed covering the frequency ranges of the THz, far-infrared and mid-infrared frequency region. The latter was a development for our Palestinian partners which uses carved calcium fluoride windows opaque in the mid-infrared. A variation of the temperature becomes possible by a special sealing we want to describe here.
For an exact layer thickness the inner part of the cell consists of a flat and a carved window which has to be polished down to just some microns accuracy to gain a penetrable, very thin solution layer. The complete setup of the mid-infrared and the first far-infrared liquid cell are similar to the developed THz setup which is seen in fig. 3.6, both using the "carved window" approach for realization.

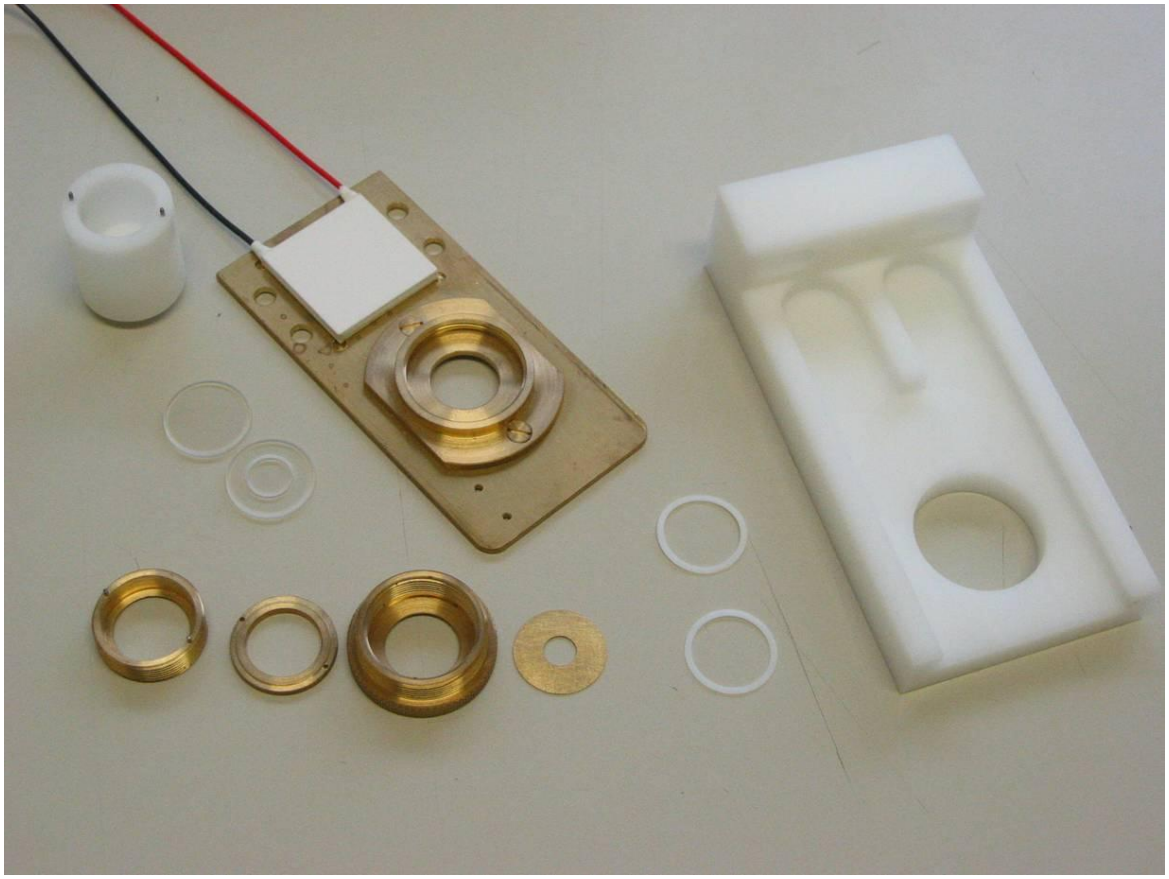

Fig. 3.6: The THz liquid cell setup with an inclusion of temperature control and variation. The many different screwable compartments are necessary for applying uniform pressure on the very breakable windows.

A screwable arrangement of several compartments is necessary to apply a uniform pressure to the very breakable windows. The ditch around the inner part of the carved window (fig. 3.7) is very important for two different reasons. First of all for a complete and "bubble free" filling of the thin layer, a larger amount of solution has to be filled into and this overflow has to be collected somewhere. Otherwise the overflow would prevent a direct contact of the two windows for a complete sealing. The second reason is that experience showed that capillary forces lead to a steady drying of the internal layer. The vacuum or dry nitrogen purged spectrometer surrounding extracts water molecules at the edge of the liquid cell, which are replenished from the inside leading at worst to a complete drying. By this ditch the capillary forces are interrupted and a better sealing of the cell is achieved compared to windows without this additional ditch around the actual solution vessel.

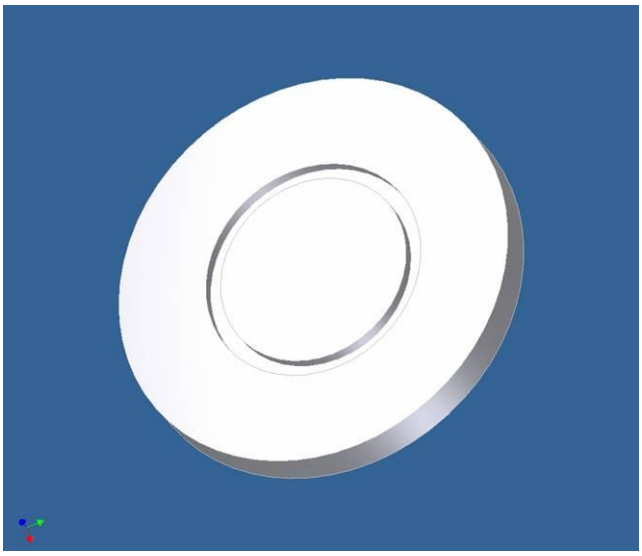

Fig. 3.7: Carved window for the inner part of the THz or mid-infrared liquid cell. Either glass or calcium fluoride are polished down on the inner part several microns to gain a fixed layer thickness. The ditch collects the protein solution overflow as well as it interrupts the capillary forces which otherwise lead to an unwanted steady drying process of the enclosed inner volume.

Even by using glass windows in the THz liquid cell setup, a complete closure of the cell is not possible just by pressing two windows together. Several oils with different viscosity have been tested to reach a well sealing. That the oil layer becomes not too thick and disturbs the defined layer thickness, a mixture of alcohol and oil proofed to be most suitable with a final recipe of 50 % polyethylene glycol 400 and 50 % propanol. The alcohol evaporates very quick and leads afterwards to an almost monolayer type of oil sealing for the cell, even

capable to store the liquid part during small temperature variations in the physiological relevant range as successful tests with the mid-infrared liquid cell demonstrated.
The THz liquid cell measurements, using backward wave oscillators as light sources, were not able to detect a significant difference between pure water and protein solution in the first version of the setup. A planned further improvement became unnecessary due to the meanwhile published free electron laser gained measurements in exactly the same year of the development [Xu06a, Xu06b]. Our setup was not able to compete with the high intensity and the reproducibility of the absolute values reached in these experiments for the same frequency range.

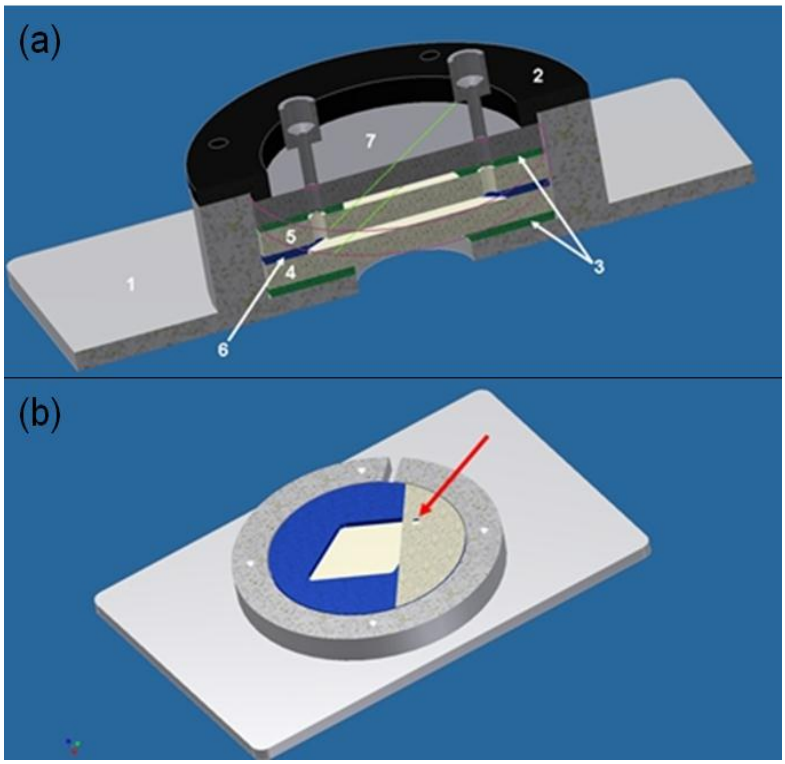

Fig. 3.8: a) Cut through the far-infrared liquid cell No. 2: The base plate for fitting into the spectrometer holder (1), the outer ring with four screws for uniform pressure (2), viton rings to prevent breaking of the windows including a passage for the liquid (3), flat high density polyethylene window (4), drilled high density polyethylene window with drillings for a solution pathway (5), self made diamond shape teflon spacer (6), covering plate with syringe adapters for filling (7).
b) The diamond shape teflon spacer and the drilling for a solution pathway through the upper compounds are indicated with the red arrow. This allows instantaneous solution exchange without an opening of the cell.

The first approach for the far-infrared liquid cell setup was to use carved windows again for the inner part. Possible window materials for this frequency range are diamond, silicon and several polymers. The first one is too expensive and the second one has a very large refractive index leading to a large loss of intensity due to reflection. In the far-infrared region, where high intensity light sources lack, this made the use of silicon as window material impractical. The polishing of the third option to use polymer windows emerged to be a difficult task. The softness even of high density polyethylene lead to more smearing than polishing of the material. Also some stress was induced into the material which relaxed after some days leading to a disappearance of the about 10 µm to 20 µm depression. Tempering of the polymer material improved but not completely overcame this situation. So a second far-infrared liquid cell was constructed using spacers between two flat polished high density polyethylene windows (fig.3.8). The difficulties related to this approach could be finally better handled. Starting out from a mid-infrared liquid cell, an adoption for the far-infrared region with self made windows, spacers and additional parts was developed.

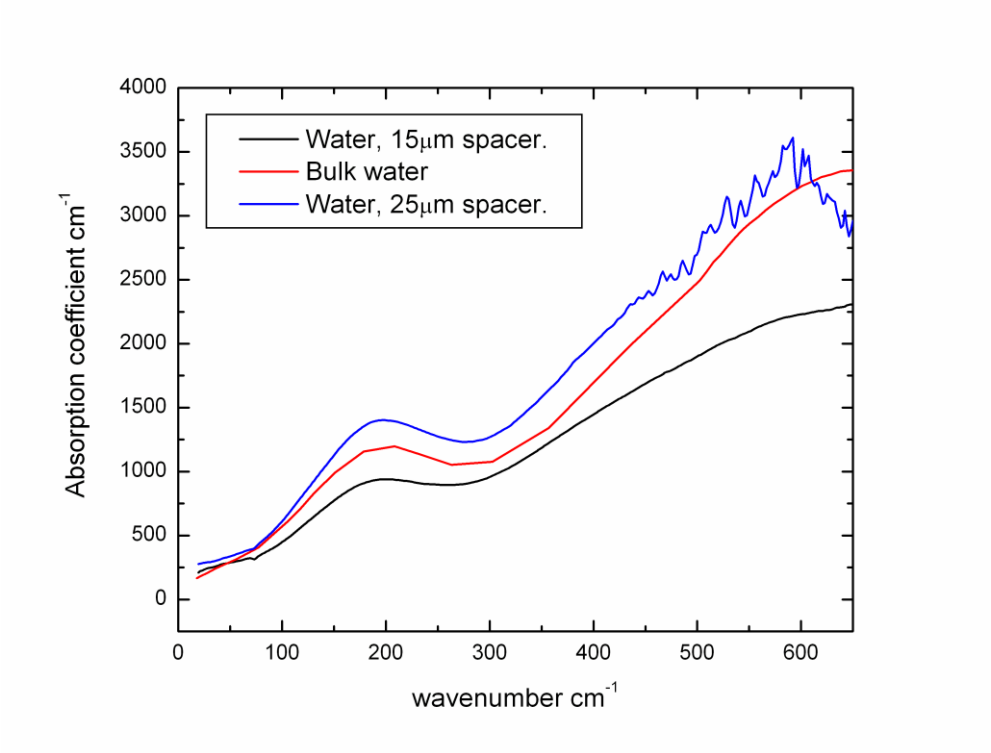

Fig. 3.9: Comparison of the measured pure water spectra with a 15 µm and 25 µm spacer to the values published in [Pal98]. There is a good agreement with the overall shape and the low frequency absorption peak position. In the upper frequency range the increasing absorption causes a larger deviation from the published data.

Even with defined teflon spacers the control of the water layer thickness was difficult as well as the sealing of the cell with an oil film, both leading to deviations of the absolute absorption value. Also pure water had to be degassed in vacuum to prevent additional scattering from microscopic gas inclusions. As a test measurement, the absorption of pure water was compared to the standard values from Palik [Pal98] which reproduced the shape and peak position in the lower frequency range very well. In the upper frequency range the increasing absorption leads to larger error bars and explains the deviation to the standard values. With an inclusion of the water layer thickness even the absolute values become close but have not been reproduced for all measurements that well as seen in fig. 3.9.

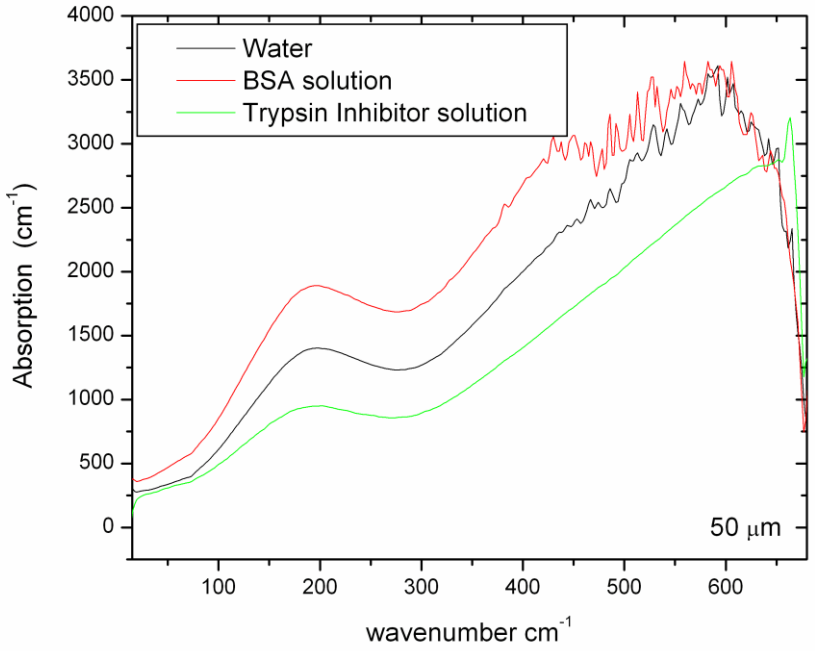

Fig. 3.10: Pure water spectrum compared to BSA and trypsin inhibitor protein solutions using a 50 µm spacer. The absolute absorption value has to be taken with care but the shape indicates no difference between water and protein solution in line with measurements from the THz frequency range.

The measurements on protein solutions revealed no significant differences to pure water as indicated in fig. 3.10. The measurements have been made with protein solutions as dense as possible. This means that the solution was optimized to a solution which did not show any precipitation which would then cause additional and unwanted scattering. A comparison to results in the THz, demonstrated in fig 3.11, which were gained up to 124 cm^{-1}, could be

reproduced in that manner that there was no difference to be found in the absorption shape between water and protein solution [Xu06a, Xu06b]. The observed small absorption decrease around 5% of the protein solution compared to pure water could not be confirmed from our measurements since the absolute value deviations made this conclusion impossible.

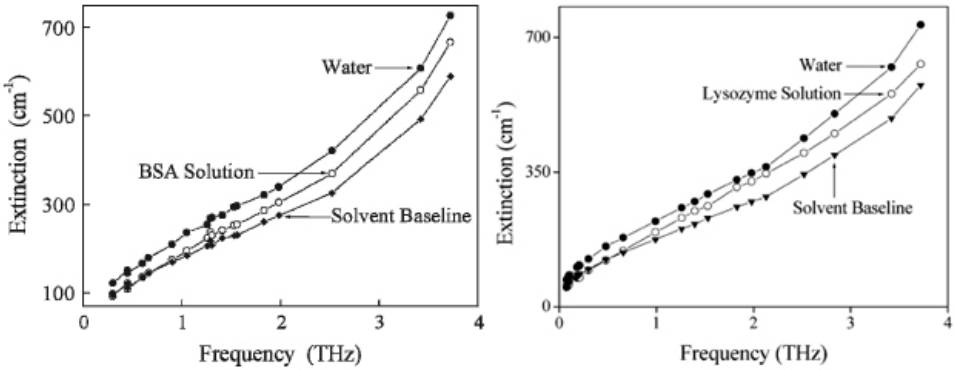

Fig. 3.11: THz spectra of BSA and lysozyme protein solutions compared to buffer. The lower graph is gained by calculation of the liquid water absorption arising from the replaced volume. No assumption about a hydration shell or biological water was made. The most important finding is that protein solution absorbs less than pure water. This is the first protein measurement in aqueous solution for the THz frequency range proofing a difference in the absolute absorption value [Xu06a, Xu06b].

Besides the experimental breakthrough of measuring protein solutions by using a free electron laser as a high intensity THz light source there have been some conclusions derived for the protein absorption. The protein solution absorbs less than pure water (or buffer) which means that the protein and biological water system absorbs less than the water volume which is displaced. The sole protein spectrum was calculated by withdrawing the amount of pure buffer solution spectrum that is assumed to be displaced by the protein. The non-zero difference is then attributed to the protein and hydration shell absorption as seen in fig. 3.12 which is a more precise terminology than originally stated in the publication just refer to protein absorption. The contribution from the biological water is not defined due to the indirect calculation method. But the results are in line with the completely featureless THz spectra of externally hydrated proteins underlining the appropriateness of an external hydration technique to gain spectra of proteins in a state close to the natural one. Another advantage is the direct probe of biological water which is the motivation for the following two sections describing the new sample preparation technique and the developed atmosphere cell.

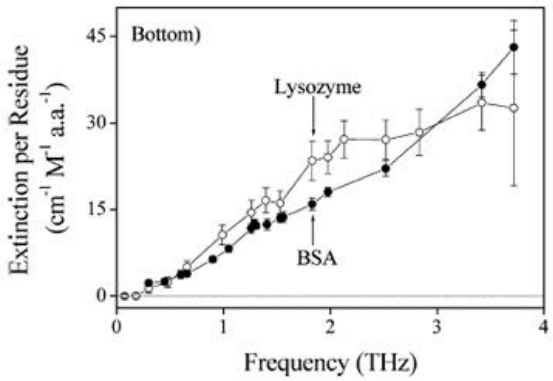

Fig. 3.12: Calculated BSA and lysozyme THz absorption spectra normalized to the absorption per residue to account for the different molar weight and volume of the two proteins. The spectra were indirectly gained by withdrawing the absorption of the buffer solution whose volume was replaced by the protein. More precise, than stated in the publication, this is the absorption of the protein plus the biological water shell. The authors claim that the differences in absorption are statistically significant [Xu06b].

3.2.2. Free Standing Protein Film Sample Preparation

External hydration of proteins and the variation of this parameter lead to the above discussed advantage of a direct relation to the absorption of biological water. The crucial point for the development of such a setup is the sample preparation that demands a quick and complete hydration of the final protein sample. Due to low intensity sources and the hydrated protein absorption, thin layers are employed from groups in the THz region where the absorption is even significantly less than in the far-infrared. Those films are usually prepared on a THz opaque substrate which completely hinders the hydration from below. This substrate causes often cracks in the protein layer on top due to a different thermal expansion coefficient of the two materials. Surface interactions and possible condensation at the substrate at high relative humidity are also problematic. The latter demands a precise temperature controlled substrate and in general slow temperature changes due to the larger heat capacity of the substrate. Occurring temperature differences would lead instantaneous to condensation effects at high relative humidity. Without a supporting substrate, to our best knowledge just pressed pellets, unsuitably for this application, have been used for protein measurements so far. Therefore we present here the details of our developed sample preparation technique for free standing

protein films which is the basis of the construction of the atmosphere cell described in the following section.

Besides the condensation and worse external hydration, a substrate underneath the protein film causes usually cracks during the sample preparation and during temperature dependent measurements. The first consequence might be the main reason for the lack of protein samples of appropriate size which can be used as free standing protein films. We identified two reasons which are the major obstacles to overcome on the way to free standing protein films. Firstly, the different thermal expansion of the two materials which are connected to each other at the surface. And secondly, exactly this connection of the protein film and the substrate which has to be prevented as much as possible. Or examined from another point of view, one obstacle is that the film is prepared directly on the opaque material and not on a suitable sample preparation substrate, where the sample is then afterwards transferred to the transparent substrate if necessary. The latter was our approach and a supporting substrate turned out to be unnecessary after further optimization of the process, finally leading to free standing protein films.

As a preparation substrate we tested several polymers whose thermal expansion coefficient is similar to typical protein ones. A preventation of the connection between the final protein film and the substrate was best achieved by large hydrophobicity directly at the first sample preparation step starting with a protein solution. From several experiments polypropylene finally turned out to fulfil best these requirements and was used throughout the sample preparation of the here presented measurements.

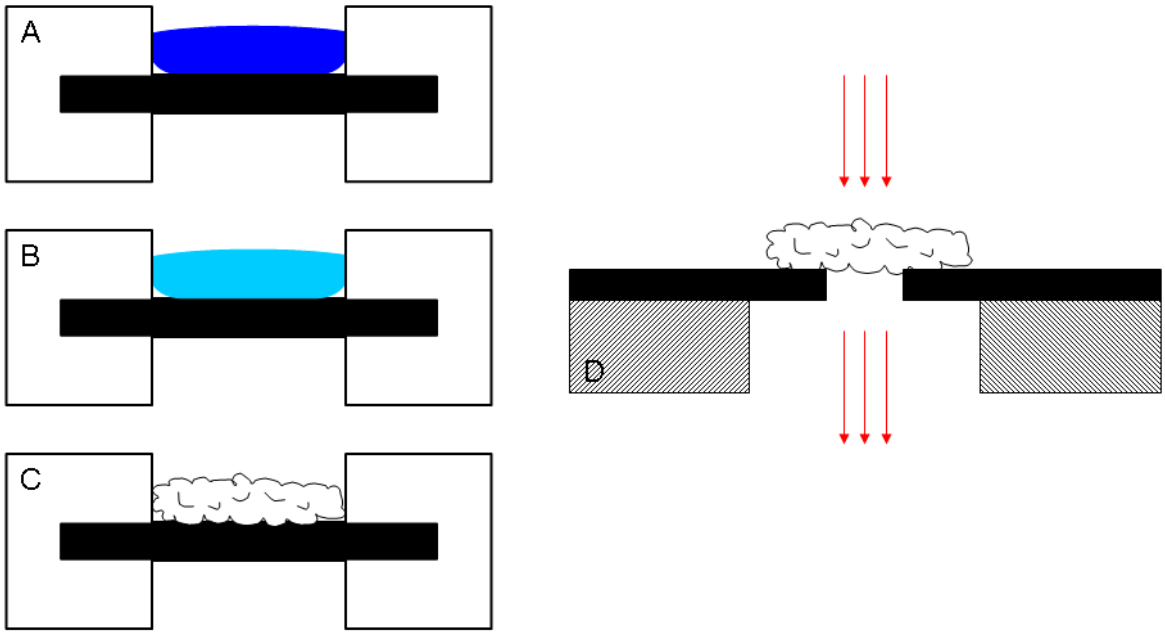

Fig. 3.13: The free standing protein film sample preparation:
A: Protein solution on polypropylene substrate
B: Cooling down to about -80°C depending on the type of protein
C: Lyophilization leads to a cotton wool like protein film
D: Transfer to the actual sample holder for experiments
The black color indicates polypropylene, the white color simplified plastic and rubber holders and hatched parts the actual metallic sample holders used for experiments.

The whole procedure for gaining free standing protein films is seen in fig. 3.13. Protein solution optimized to the later required optical density is filled into the sample preparation holders. These are made of two screwable plastic parts in which two rubber rings press down a disc of polypropylene foil. This arrangement was chosen to ensure water tight sample holders even under large temperature changes as they occur during the process. An entire cleaning of all parts after every preparation process is hence possible and the substrate is always cut out of new foil to avoid any contamination which is especially important for biological matter. For the protein solution purchased lyophilized powder from Sigma was used without any further purification and dissolved in deionised and micro filtered water. The freezing process has to be quick enough to prevent crystallization but also slow enough to avoid cracks upon cooling. By lowering the pressure (~1 mbar) under the triple point, sublimation sets in and leads to a drying process leaving a cotton wool like protein film behind. Therefore the concentration and the protein solution thickness are important parameters influencing the final result. Careful transfer on the actual sample holders with apertures lead to samples which are large enough for hundreds of possible measuring spots

under the microscope. Since a temperature variation is also taken place during the extensive measurements, the protein film is not directly fixed on the metal sample holders which provide well thermal coupling to the Peltier element. The protein film is attached again to a polypropylene foil in which apertures of up to 8 mm diameter have been embossed in. Then this set is finally fixed on the metal holders buffering the different thermal expansion of the employed materials. This construction successfully prevented cracks within the protein films even under extensive temperature and hydration dependent measurements.

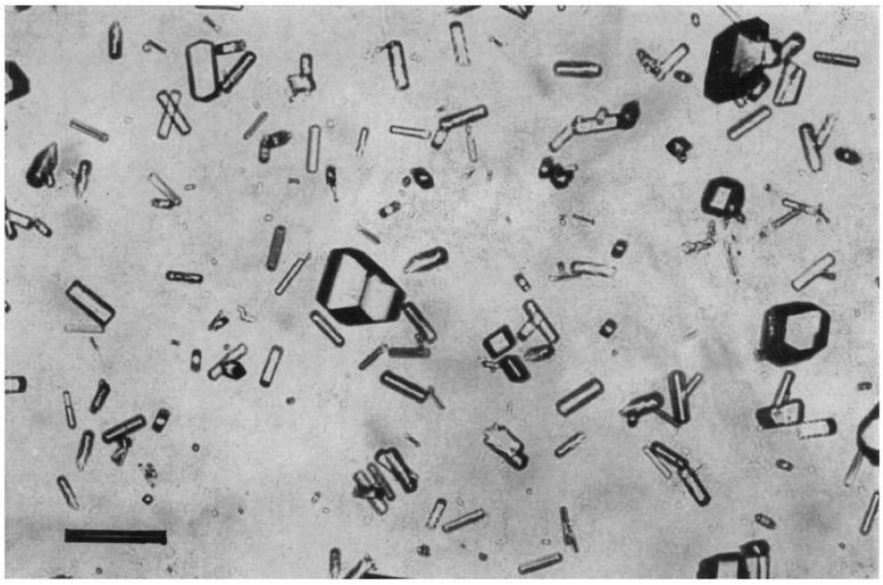

Fig. 3.14: Lysozyme crystals as grown from protein solution. The bar on the lower left has a length of 0.02 mm [Fin98].

Regarding to the protein, the lyophilization process parameters have to be optimized for each type. This concerns mainly the freezing temperature and the timescale of both cooling and drying. Sometimes crystallization can be induced into the system leading to larger polycrystalline parts interfering with the wavelength of our far-infrared radiation. Here either a much larger crystal or a much smaller polycrystalline sample than the far-infrared wavelength in the order of 10 μm has to be produced to prevent scattering. We managed to reach the latter since just lysozyme is an easy to crystallize protein as seen in fig. 3.14.

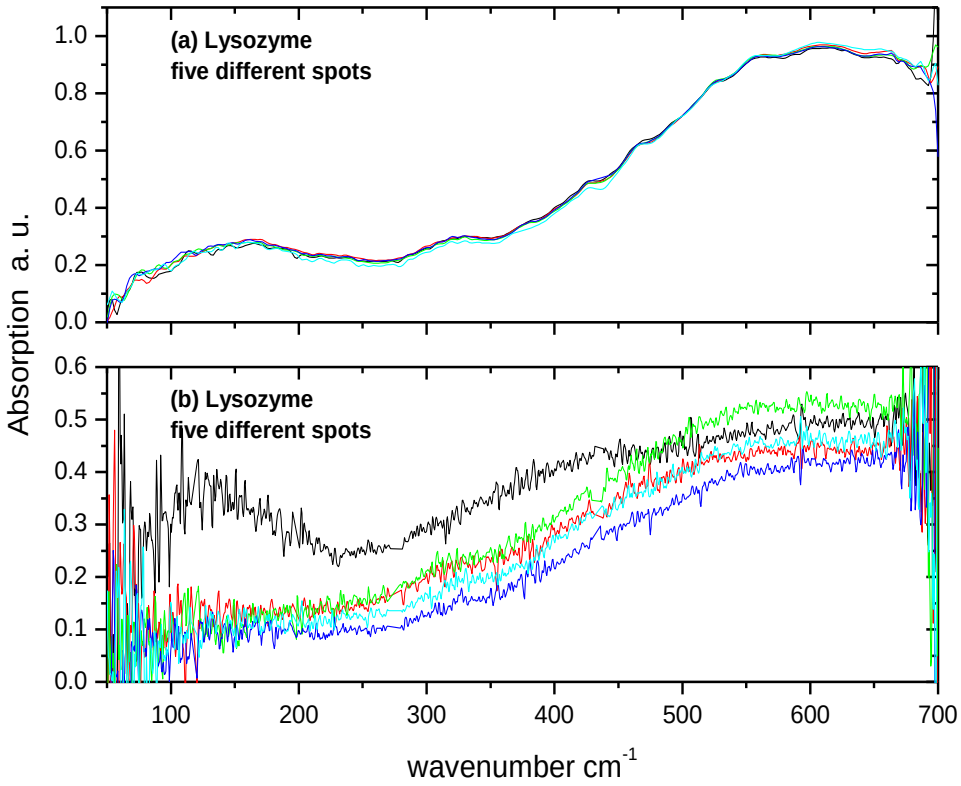

Fig. 3.15: Spectroscopic analysis of the optimization of the sample preparation process. The spots measured of the upper sample (a) have been produced by liquid nitrogen cooling and the self constructed lyophilizator. The lower spectra of a sample produced by slower refrigerator cooling (b) reveal for one spot an additional low frequency peak around 150 cm^{-1} which is not due to absorption but caused from scattering of a large lysozyme crystal of presumably 60 µm size. The fluctuations in the lower part arise from a lack of hydration control which was not applied in this quick sample quality check.

To visualize this optimization we want to present the spectra of lysozyme in fig. 3.15 whose sample preparation was most difficult. We were able to estimate from the additional scattering peak that some large lysozyme crystals are present in sample (b). This explains the dependence of the spectra to the measured spot which is not observed in the uniform sample without such large crystals due to precisely controlled parameters and a faster cooling during the preparation process. This was our motivation for cooling the samples with liquid nitrogen and developing our own lyophilizator.

The cooling process is crucial for the size of crystals. Our protein concentrations are high and provide a basis for crystallization which means that an as short as possible delay up to the

complete freezing of the sample is important to prevent crystallization. Also slow cooling supports a crystallization process and fast freezing can stop the process immediately. But on the other hand too fast cooling leads to cracks in the frozen solution and later in the protein film finally destroying the sample. Optimization and experience showed that the best way is a manual procedure by a precooled copper block for a quick freezing at the beginning, followed by a slower cooling rate to about -80°C. At temperatures below, the risk of cracks increases again significantly. Therefore we controlled the sample temperature during the complete cooling process instead of cooling the sample in a freezer as it was used in earlier attempts. For the cooling process we made the experience that a copper holder with tips placed close to the middle of each holder leads to the best results (fig. 3.16). Otherwise some volcano shape protein films were produced because freezing started mainly from the outside, after the holders got cold enough, causing a pile up of frozen solution around the center. For achieving a more uniform and flat sample, the cooling has to be balanced which we reached by much smaller tips than the diameter of the sample finally resulting into an almost sudden freezing of the entire protein solution.

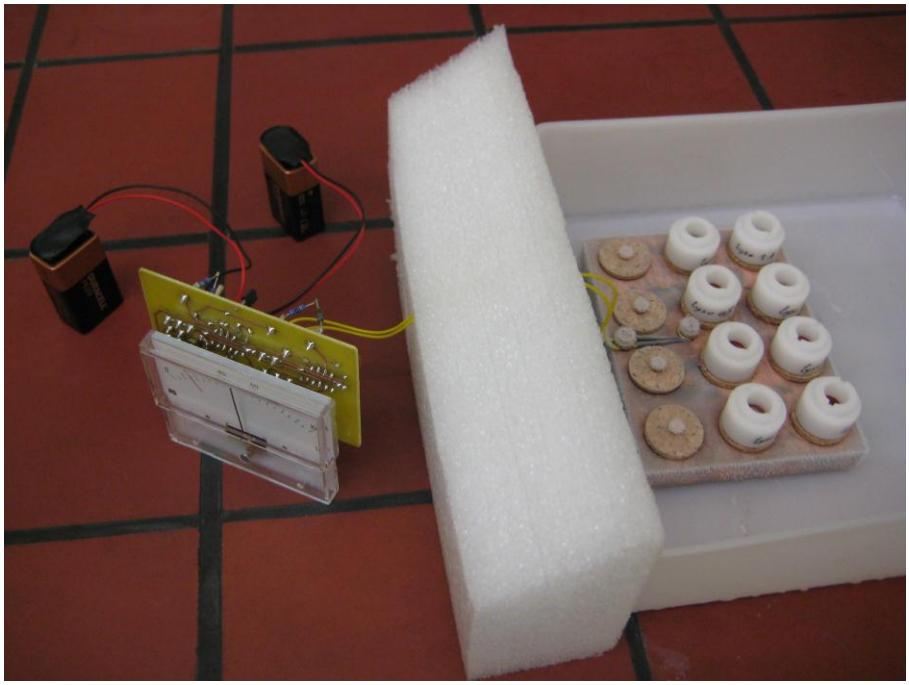

Fig. 3.16: The white sample preparation holders on top of the final version of the copper block with insulating cork discs and small tips. A vacuum tight thermometer attached to a PT-100 temperature sensor is seen on the left also important for the temperature control in the following lyophilzation process. The cooling rate was manually varied by liquid nitrogen.

The usage of a standard lyophilizator in the following step lead to unsatisfactory results because neither the temperature nor the applied low pressure could be controlled or varied. Therefore a self made lyophilizator (fig.3.17) was constructed on basis of a large desiccator. The temperature was controlled by using another precooled copper block and a self made vacuum tight thermometer with its own power supply. With the individual control of the parameters this step of the sample preparation procedure was finally cut down to several hours from almost one day before.

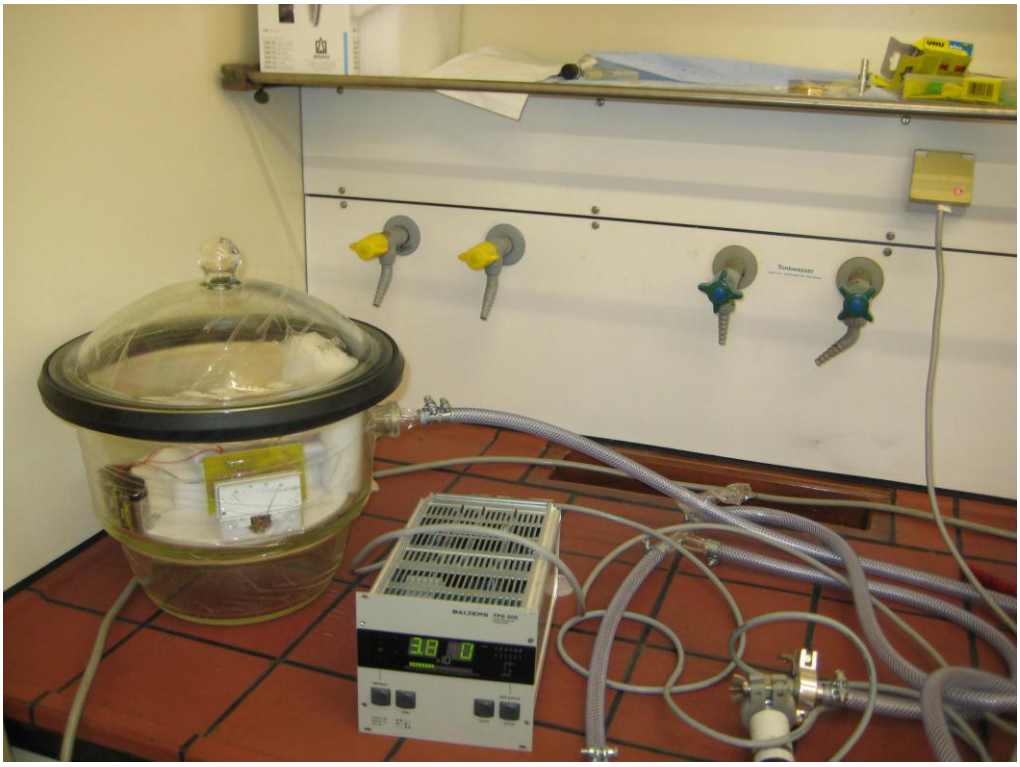

Fig. 3.17: The self made lyophilizator based on a large desiccator. The built up vacuum tight thermometer was used to control the warming up procedure of the samples.

After being removed from the sample preparation holders and being transferred to the actual sample holders used for the measurements, the protein films could be stored for several days in the refrigerator without changing their far-infrared spectra as test measurements reveal. These free standing protein film preparation was the basis for the development of the atmosphere cell setup described in the upcoming section.

3.2.3. The Atmosphere Cell

There are several indications based on experiments investigating different proteins that even low external hydration (0.1 h to 0.4 h) leads to protein functionality and can be used to study proteins in a close to natural state [Poo83a, Poo83b, Poo84, Sch85, Zak88a, Zak88b, Row80]. For us this opened a path to work around the strong liquid water absorption and to get direct access to the absorption of biological water as already discussed throughout this chapter.

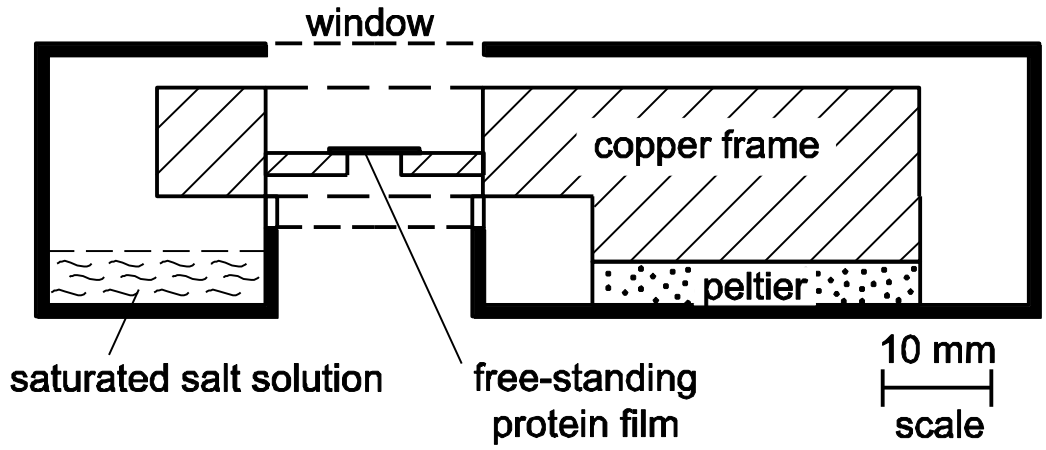

Fig. 3.18: Schematic, simplified overview over the atmosphere cell construction restricted to the most relevant parts.

The prerequisites for an atmosphere cell construction are the free standing protein film as described in section 3.2.2, a temperature and humidity control as well as a restriction to very tiny proportions to fit under the microscope in the confocal setting. A schematic, simplified overview picturing the main relevant parts is given in fig. 3.18 and a detailed technical drawing of the inner parts of the atmosphere cell is pictured in fig. 3.19.

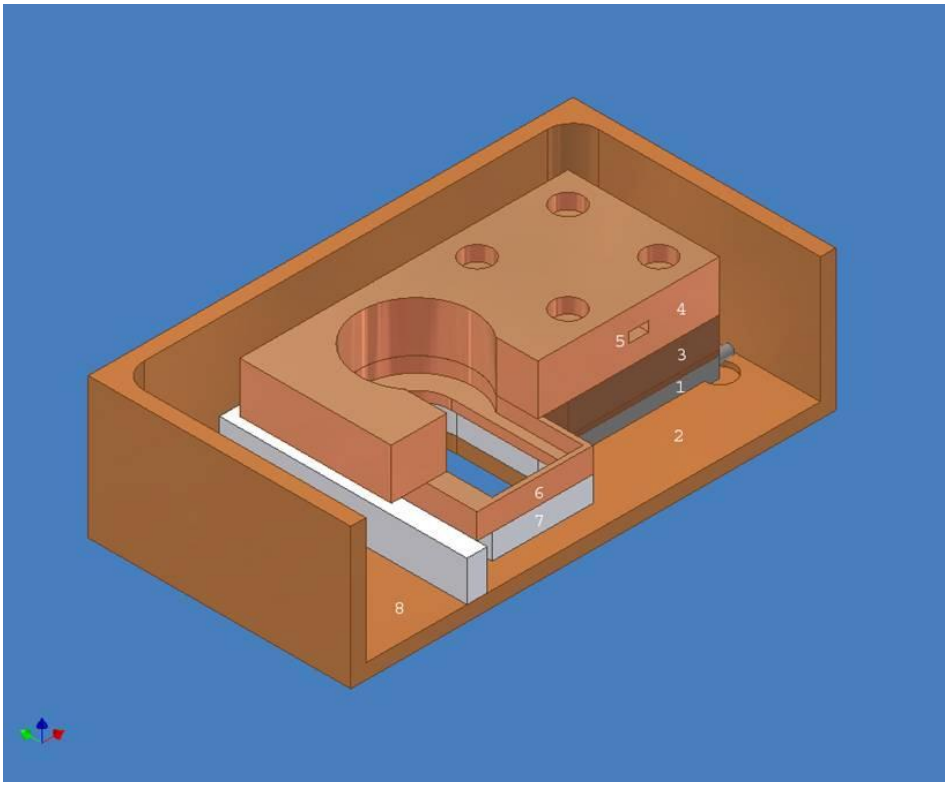

Fig. 3.19: The inner parts of the atmosphere cell. The Peltier device (1), the cell itself made of copper, copper parts for thermal coupling (3,4,6), the embedded PT-100 temperature sensor, insulation for thermal decoupling (7), the compartment for the salt solution (8).

The temperature is controlled via a Peltier device and a PT-100 temperature sensor connected to a developed temperature controller calibrated to this particular cell. A water cooling in the base plate increases the efficiency of the Peltier device covering the physiological relevant temperature range of 4°C to 55°C. In addition it stabilizes the temperature control to better than 0.1 K. For external hydration a standard calibration technique for humidity sensors is applied [Ber80, Web95, Ben98]. In a closed volume the air above saturated salt solutions lead to a very precise and stable relative humidity. By changing the saturated salt solutions different relative humidity values can be reached as presented in table 3.1.

	Temperature (°C)									
	5	10	15	20	25	30	35	40	45	50
Salt	**Relative Humidity (%)**									
$MgCl_2$	34	34	34	33	33	33	32	32	31	30
$Mg(NO_3)_2$	58	57	56	55	53	52	50	49	46	-
NaCl	76	76	76	76	75	75	75	75	75	75
KCl	88	88	87	86	85	85	84	82	81	80
K_2SO_4	98	98	97	97	97	96	96	96	96	96

Table 3.1: Relative humidity over saturated salt solutions and their temperature dependence. Based on the work of O'Brien [Bri48].

A humidity sensor proofed for this setup that after about one hour the humidity fluctuations became less than the sensor sensitivity of about 1 % relative humidity. But anyway a sensor is always less reliable than its calibration technique and in the later versions the humidity sensor was removed to gain some additional space within the cell for some refinements of the arrangement. The protein sample and a gap to perform the background measurement are situated in part (6) of fig. 3.19, so that the cell does not have to be opened in between.

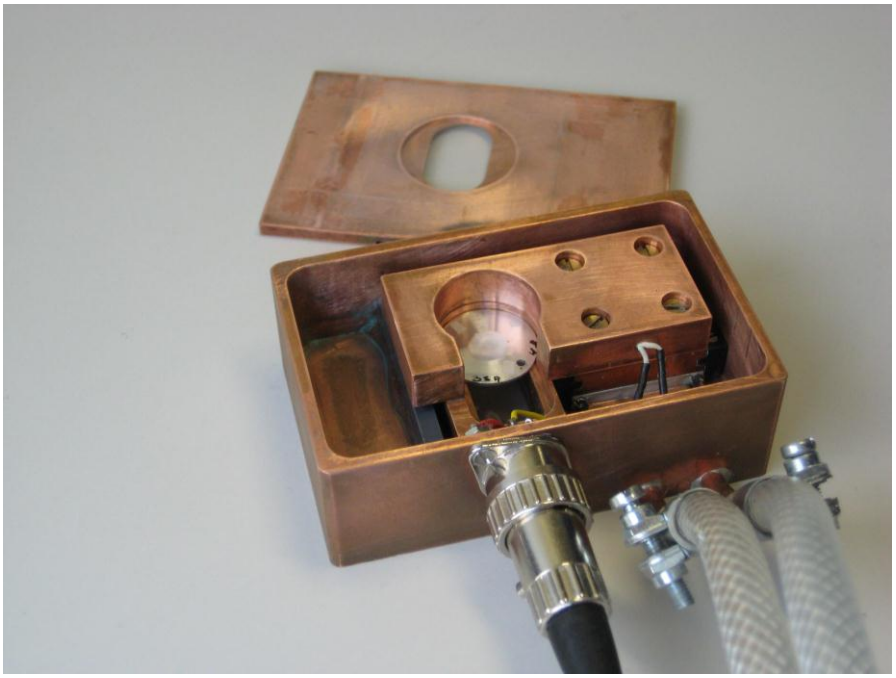

Fig. 3.20: The atmosphere cell with the free standing protein film placed in the center. The saturated salt solution is filled to the vessel on the left, the connections are for the temperature controller and the water cooling. The size of the cell is about 30/85/25 mm (l/w/h) to fit under the infrared microscope. The screwing arrangement is built in for a better cleaning and maintenance including a possible exchange of the sensors and devices.

As it can be seen from fig. 3.20 the cell is sealed with two 13 μm mylar foil windows suitable for our frequency range. To gain the best accuracy and stability of the measurements we found out that humidity fluctuations have to be prevented as much as possible even in the very short open (non-evacuated and not dry nitrogen purged) path of the beam. The part above the cell is still partially purged with dry nitrogen from the microscope but the part below the cell of about 1 cm caused unnecessary noise during the nine to sixteen hours enduring measurement procedure for every single parameter dependence. Therefore we packed the whole microscope under another completely nitrogen purged cover which lead to the accuracy of the measurements as presented in the first part of the next chapter.

4. Far-Infrared Measurements on Proteins and Accompanying Molecular Dynamics Simulations

At the beginning of the main results and discussion chapter we summarize all discoveries up to day in the relevant far-infrared frequency range and the important ones from the neighboring mid-infrared and THz region in an own review about spectroscopy on proteins, regarding also findings from other techniques before presenting our own results. Extensive measurements on Bovine Serum Albumin, Hen Egg White Lysozyme, Trypsin Inhibitor and β-Lactoglobulin have been performed. The protein samples are introduced already with their corresponding spectra. Some of them are the first far-infrared measurements of these proteins, but all measurements are the first under the control of both temperature and hydration including an improved resolution, high accuracy and a proven reproducibility, which is most important for measurements on biological matter. Therefore a separate part on the reproducibility is included prior the discussion of further results. The basic frequency components of the far-infrared protein spectra have been analyzed and will be discussed afterwards. Little is known about the origin of the far-infrared absorption but the desired experimental improvement has been made yet and a description for the calculation of infrared absorption protein spectra has been given in the theory chapter of this thesis. So the here presented results can be seen as an encouragement to restart again the theoretical interpretation by molecular dynamics simulations of the broad protein absorption bands in the far-infrared. The hidden information in this part of the electromagnetic spectrum is very promising and revelation would be a large step ahead on the way of understanding proteins, which would lead to important applications in many fields.

In addition the so far unknown influence of the two important parameters temperature and hydration on the far-infrared protein spectra has been extensively investigated and constitute the main part of this result and discussion chapter.

From the hydration dependence it was possible to derive the far-infrared properties of biological water. Our spectroscopic results support the “iceberg” model which is originally based on a dynamical picture of this type of water molecules indicating a relation to macroscopic detectable properties.

Three temperature dependent effects were identified from our measurements. Two frequency dependent and one large frequency independent effect is discussed. The explanation of the latter led to our new developed MERCEP model. Corresponding molecular dynamics simulations support this new model in some aspects and initiated some refinements of the MERCEP model. The model relates spectroscopic properties with a microscopic insight into the protein. Since all these results just exhibit their conclusion if presented in a comprehensive way, theory, simulations and the measurements concerning this effect are presented together in the last section of this chapter.

4.1. Review of Spectroscopy on Proteins

The aim of this section is the comprehensive presentation of the prior results in the far-infrared and the neighboring frequency ranges to identify the basis of experimental information this thesis started from. This should emphasize to what extend the different aspects have been pushed forward with the outcome of this thesis. This section will be split up into the standard mid-infrared, the rarely studied far-infrared and the recently fashionable THz range for studying biological matter. At the end some important results from other techniques are presented.

4.1.1. Mid-Infrared Spectroscopy on Proteins

The first measurements on proteins revealed some strong absorption bands in the mid-infrared (700 cm^{-1} to about 4000 cm^{-1}), which were not differing between various proteins within the accuracy of that time. Since in this frequency range mainly local vibrations of atom pairs are detected, a difference between the same elements containing proteins was anyway not expected. Elliott and Ambrose [Ell50] proposed from that a correlation of the, later named amide I and II bands, to the peptide structure and later found indications for the shift of the peak positions by the influence of the secondary structure arrangement [The79]. Normal mode studies from Miyazawa et al. proofed that the amide I, II and III are modes located within a peptide group [Miy58]. Since they occurred in all proteins, nine prominent absorption bands, mainly in the mid-infrared, were then named amide A and B and amide I to VII, starting from the highest frequency down to the far-infrared. Few authors use a different numbering of

amide I to IX which we want to mention. This does not make any sense due to historical reasons, which became later the official labeling. Another argument for the historical naming is that the amide I band is the strongest absorption band of proteins and hence it makes sense to give this one the first number and not the third. We will use here the original labeling throughout the thesis.

The identification of absorption bands in proteins have been mainly derived from the model compound N-methylacetamide for the peptide bond [Miy58]. The hints on the influence of the secondary structure came from normal mode analysis and spectroscopy on polyglycine ($-CH_2-CO-NH-)_n$ which is one of the simplest synthetic polyamino acids [Miy61, Gup68]. Polyglycine I has linear zigzag chains in an antiparallel β-sheet form in which neighboring chains are hydrogen bonded via N-H⋯O=C. N-H⋯O hydrogen bonds seems to play a role in both types of chains running in the same or opposite direction whereas C-H⋯O is very weak in chains of antiparallel direction, like in polyglycine I, and is used for the distinction of direction. Polyglycine II is in a helical form to model α-helical proteins but the hydrogen bond formation is different. Hydrogen bonds of the form C-H⋯O are found at just one third of all possible positions but these model systems proofed to be very useful for a protein band assignment in the mid-infrared. An overview over the underlying vibrations of all amide bands, except amide IV, including an infrared active dipole moment change, is given in fig. 4.1. Here also the peak positions in different secondary structures based on the two types of polyglycine are presented, too.

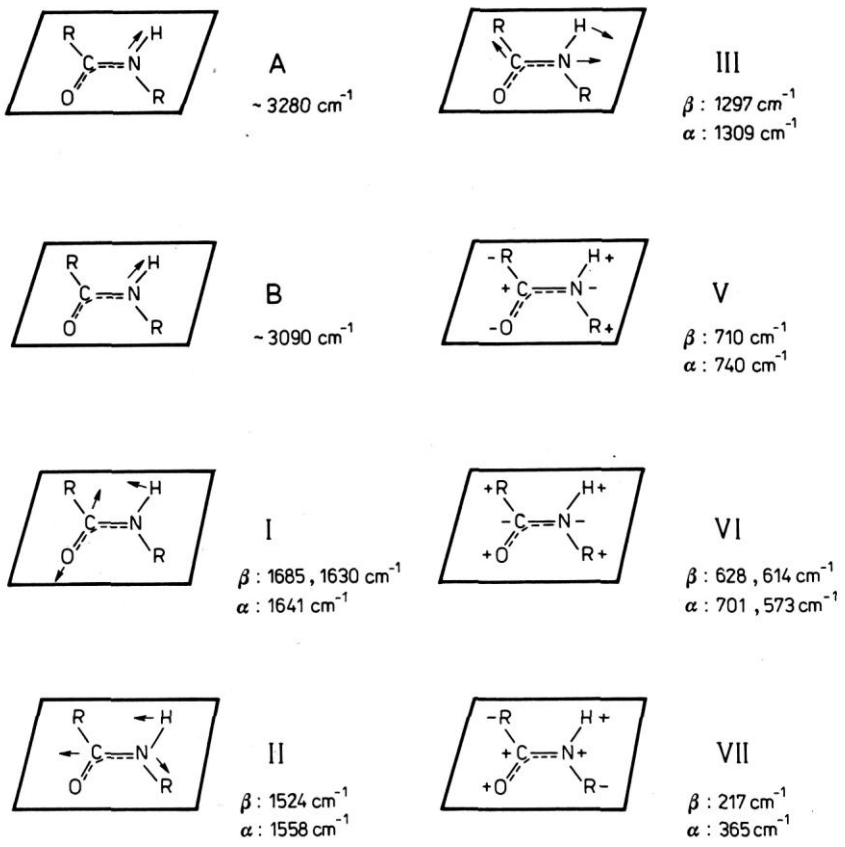

Fig. 4.1: The underlying vibrations of the amide absorption bands derived from polyglycine in two different conformations as labeled. Arrows indicate the atom displacements during the vibration. The signs + and – indicate movements of atoms above and below the plane respectively [Twa94].

Krimm and Bandekar [Kri86] and Bandekar [Ban92] described the conformational sensitivity of amide bands in a modern way by hydrogen bonding and dipole coupling. This coupling leads to a band splitting depending on the orientation and distance of the interacting dipoles and hence on the secondary arrangement of the polypeptide chain. The amide I band is mainly used for infrared secondary structure analysis. The amide II band is not so sensitive to the secondary elements, but the amide III band is again. The latter is very weak in the mid-infrared but, due to different selection rules, this is the band of choice for Raman studies and for a comparison of both methods [Sur88, Sur93]. Integrated intensities over the amide sub-bands (fig. 4.2) lead to secondary structure element content estimations as it is used nowadays in software packages for the chemical and pharmaceutical industry. Exact amide I peak positions for all possible secondary structure elements can be found for instance in [Arr93].

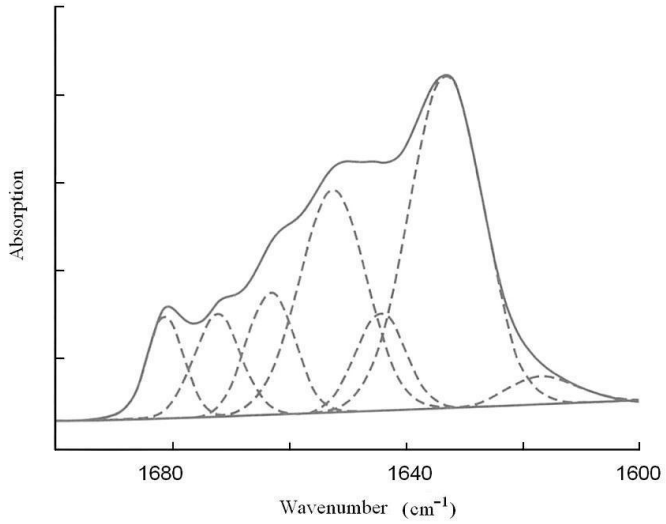

Fig. 4.2: Mid-infrared spectrum of the amide I region of ribonnuclease S. The deconvolution into several sub-bands underneath the amide absorption peak indicates the contribution from α-helices, β-sheets and random coil secondary structure elements [Hol98].

The possible applications for mid-infrared spectroscopy on proteins are diverse. This frequency region is well studied by many groups and by industry, since mid-infrared spectroscopy became a standard method. Some important examples of what can be observed, by using sophisticated methods and an application of the extensive knowledge about mid-infrared spectroscopy on proteins, are quickly presented here to close this section. For instance rebinding dynamics of CO after photolysis in myoglobin can be observed through mid-infrared bands in the so called temperature derivative spectroscopy [Ber90]. Temperature and hydration dependent shifts can be sensitively detected and connected to conformational changes [Pre93, Don95, Grd05a, Grd05b], often reaching an outstanding accuracy of about 10^{-5} which enables to detect very small absorption changes [Suk05]. Due to experimental difficulties such accuracy cannot be reached up to day in the far-infrared frequency range which is in focus of the next section.

4.1.2. Far-Infrared Spectroscopy on Proteins

Low light source intensities and an energy scale around or even below k_BT make the far-infrared region an experimental challenging frequency range. The large absorption of the often water containing proteins and sample sizes in the order of the magnitude of the

wavelength causing etalon effects, are in addition experimental difficulties. Hence little work on biological samples has been done in the far-infrared region, despite the theoretical expectations and simulation results to find here functional relevant collective motions and direct information about the hydrogen bond network as described before. On the first glance our measured region from 40 cm^{-1} to 690 cm^{-1} seems to be just a small part of the electromagnetic spectrum. But in terms of octaves, which means a doubling of the frequency, this region spans almost five octaves. This is about the same range as everything above the far-infrared up to the visible going through all over the widely used mid- and near-infrared regions. If we use the term far-infrared we mean in particular frequencies from 100 cm^{-1} up to 700 cm^{-1}. The lower edge is the typical very upper limit for most of the standard THz setups which is rarely reached. So our measurements have a well overlap with the THz range which we will use for comparison of our measurements to some publications of this region.
We give here a historical review on all far-infrared measurements on proteins and most biomolecules to our best knowledge, since they are anyway of little number. To emphasize the novelty of our measurements we will always outline the controlled experimental parameters, focused on temperature and hydration, due to their importance for detecting the natural protein dynamics as it has been discussed already. Also different sample preparations are proven to influence the results [Bro72] and can cause etalon effects as narrow peaks, which unfortunately overlay most of the measurements from the group of Genzel [Mar00], the main contributors to far-infrared spectroscopy on biomolecules for decades.

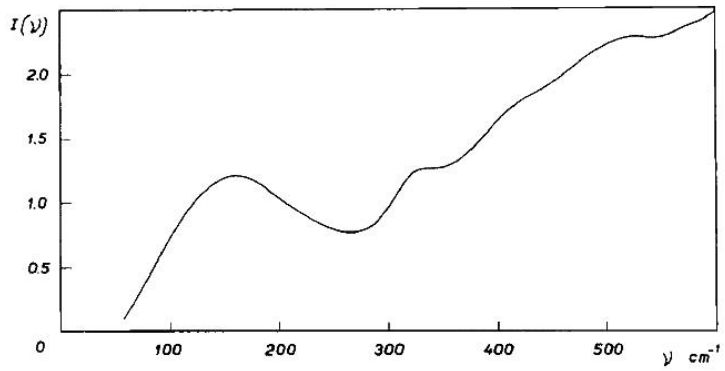

Fig. 4.3: The absorption spectrum of hen egg white lysozyme without any control of temperature and hydration in a polyethylene pellet [Buo71].

Starting in 1971 Buontempo et al. investigated several globular proteins prepared as films or pellets at ambient condition without special caution taken as far as the temperature and hydration is concerned [Buo71]. All globular proteins exhibit a broad absorption band

between 100 cm^{-1} and 200 cm^{-1}. In a detailed study of lysozyme (fig. 4.3) which represents maybe the best protein measurement for decades, an indication of several peaks above has been found, too. After that myoglobin was measured in a Vaseline paste at not further controlled room temperature and after heat denaturation; a detailed analysis of the absorption peaks was included [Chi73]. Shotts and Sievers measured pressed pellets of proteins at liquid helium temperature [Sho74].

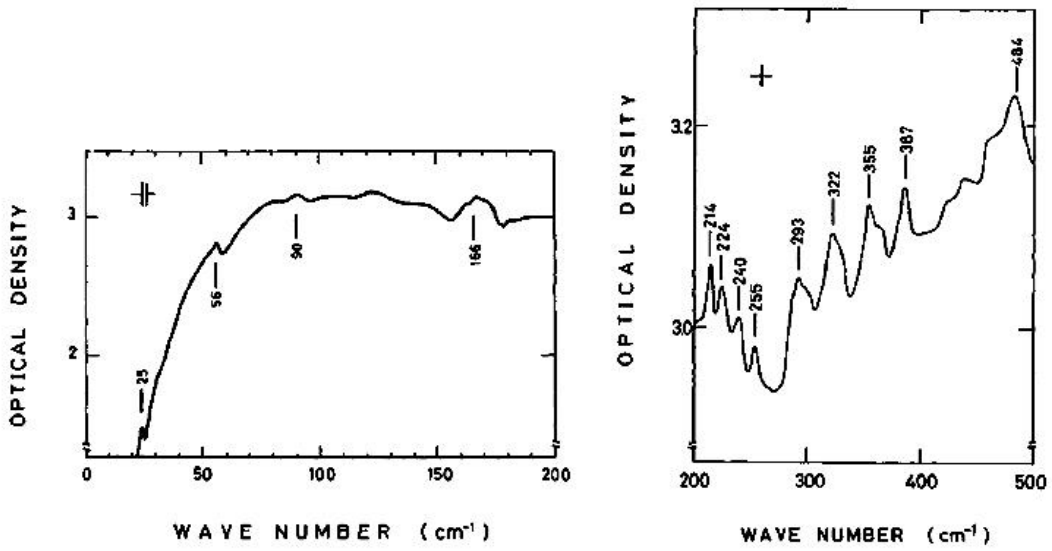

Fig. 4.4: Absorption spectra of crystalline layers of lysozyme. The narrow peaks found on the right typical in several publications up to the 1990ers, have not been reproduced in measurements on large biomolecules and are likely to be due to etalon effects caused by the sample preparation [Ata79].

The measurements of Ataka and Tanaka on crystalline lysozyme layers at room and liquid nitrogen temperature were performed in a very dry state due to the vacuum of the grating spectrometer [Ata79]. The group of Genzel conducted room-temperature measurements of protein films in a similar condition of an evacuated or with dried air flushed shroud [The79]. Other infrared and Raman experiments on biomolecules have been performed, too [Gen76, Ban83, Wit86, Pow87, Pow91]. Several narrow infrared absorption peaks have been found (similar to figure 4.4) and even compared to normal mode calculations which are later stated to be all due to etalon effects [Mar00]. Except for the low frequency measurements of Ataka and Tanaka as seen on the left in fig. 4.4 in the decade before and in the last decade, no small peaks in this frequency range on proteins or larger biomolecules have been found again. But encouraged by the renewed interest in this field, the old results were republished – together

with additional unpublished data from that time [She07]. Amazingly, more than a third of a century elapsed since the first experiments in this field, but a "detailed explanation is still lack and leaving a big room for further investigations". The best spectra of three proteins from this group, without a lot of etalon effects, are seen in fig. 4.5.

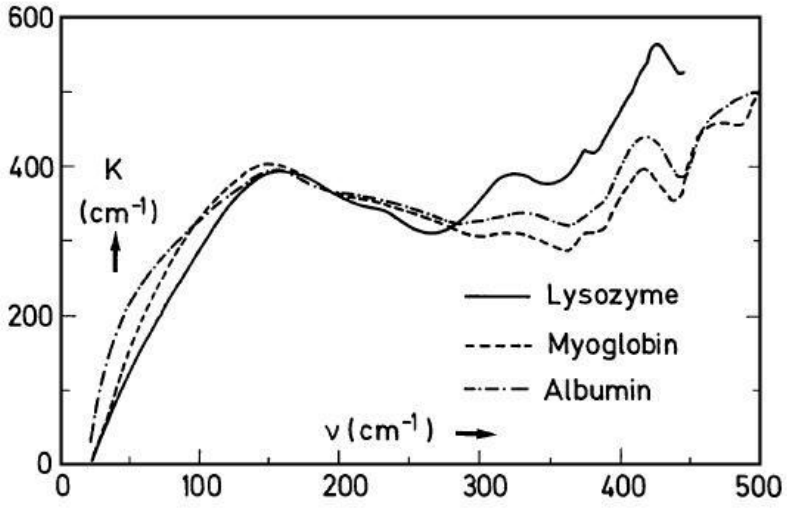

Fig. 4.5: Far-infrared spectra of lysozyme, myoglobin and albumin from cast films of about 70 µm thickness at room temperature without a temperature or hydration control [She07].

A spectrum of a film of bacteriorhodopsin was published without further discussion, probably because the paper is focused on lifetime of vibrational excitations [Xie02]. The film was stored in 90 % relative humidity atmosphere and measured at room temperature. In the far-infrared amino acids and blood serum have been measured [Mat03a]. A broadband absorption spectrum of cytochrome c in solution reaching down into the far-infrared has been published by Berthomieu et al. [Ber06]. We are not aware of any published measurements on proteins in the frequency range of 150 cm^{-1} to 500 cm^{-1} under controlled temperature and humidity so far, although these conditions strongly influence the protein dynamics and the spectra.

To explain the far-infrared absorption peaks of proteins several approaches indicate functional relevant collective type of motions causing the absorption in this range [Bah97, Bah98] reaching down to the THz range as the results for bovine pancreatic trypsin inhibitor and human lysozyme reveal [Roi95, Hay95]. Of the several protein absorption features in our measurements, so far the existence and the origin has been just discussed for the three amide bands discussed below.

In the relevant frequency range the best peak assignment uses an (hypothetical) isolated planar C=O···N-H group, which represents the peptide bond and refers to motions mainly

derived from the normal mode analysis of N-methylacetamide, which is found in [Arr93]. Many of these amide bond assignments go back to the studies of Krimm and Bandekar [Kri86], Miyazawa and others as summarized in [Par83] for slightly different peak positions depending on the used model system. Here the amide IV found at 625 cm^{-1} is caused by C=O bend (40 %), C-C (methyl-) stretch (30 %) and C-N-C deformation. The amide VI around 600 cm^{-1} arises from C=O out of plane bend and C-N torsion. Amide VII, found at about 200 cm^{-1}, originates from N-H out of plane bend, C-N torsion and C=O out of plane bend. Also non-amide modes at 436 cm^{-1} have been found due to C-C-N deformation (60 %), C=O in plane bend (10 %) and C-N stretch (~8 %) and at 289 cm^{-1} due to C-N-C deformation (70 %), C=O in-plane bend (20%) and C-C-N deformation (10 %). All frequencies are rounded on N-methylacetamide values and the percentage is approximated for a potential energy distribution but which depends strongly on the selected force field. Our measurements will be later compared to these model compounds absorption in the peak assignment section. This overview underlines the lack of reliable measurements on protein in the far-infrared in general and especially under defined conditions concerning sample preparation, temperature and hydration.

4.1.3. THz Spectroscopy on Proteins

The mainly used technique of THz time-domain spectroscopy covers typically the range from 5 cm^{-1} up to 70 cm^{-1}. Just with free electron lasers or sometimes synchrotron radiation the intensity can be increased and the frequency range can be extended up to 120 cm^{-1} [Mat03b]. Therefore we refer here to the THz range covering the frequency region from 3 cm^{-1} to 100 cm^{-1}, which is roughly 0.1 THz to 3 THz. Simulations and theoretical approaches agree upon the delocalized, collective type of motions that are expected here, whereas the far-infrared region seems to be some intermediate range and indications are hence ambiguous [War74, Jaa98, Tam00, Tam01]. Due to this collective character the THz frequency range is assumed to be sensitive to conformation, structure and the environment of the biomolecule [Wal02]. But the following review about experimental THz results does not indicate such a straight forward connection which should result in (not observed) large absorption deviations between different proteins.

One of the first measurements bridging the THz gap for biomolecules [Moe92] used synchrotron radiation in the frequency range from 15 cm^{-1} to 45 cm^{-1} to investigate lysozyme

at different hydrations (0.1 h and 0.3 h) and in water, in the region where the lowest underdamped vibrations are expected [Cus88]. The hydration dependence was performed because at low hydrations of lysozyme, the tertiary structure and some activity remains but the picosecond dynamics was found to be reduced [Eds83]. In summary a hydration dependent and almost temperature independent absorption was found. Myoglobin and the protein components biotin and retinal were measured as pressed pellets at room temperature as well as cryogenic temperatures [Plu03, Wal00]. The different retinal isomers led to different THz spectra which were interpreted to be due to the different isomeric structures. Such a statement, that THz measurements are able to detect tertiary structure changes, has been also made by the group of Markelz who claimed to be the first ones proving this, based on their mutant to wild type comparing measurements as seen in fig. 4.6 on the left hand side [Whi03, Mar02, Mar03, Hei08]. The mutant is assumed to have a reduced conformational flexibility and shows also a much lower THz absorption supporting this assumption. Later a different THz absorption was also found for the oxidized and reduced state of cytochrome C. Both interpretations just use differences in the erratic quantity of absolute absorption since no (reproducible) absorption peaks have ever been found in the THz range for a distinction.

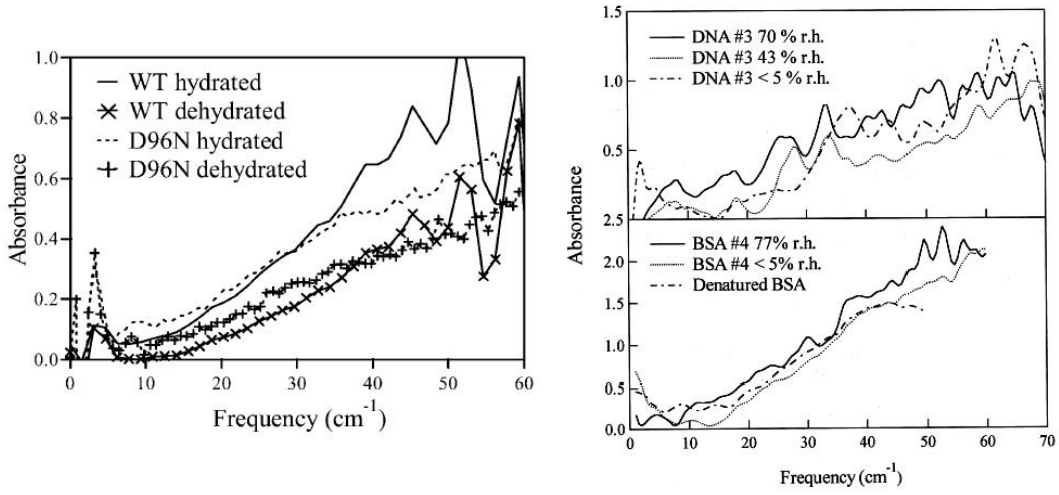

Fig. 4.6: Wild type and D96N mutant bacteriorhodopsin thin film absorption spectra. From the difference in the absolute absorption value the authors claim to present here the first THz measurement which is able to detect tertiary structure changes. Two spectra are given for each mutant in a dehydrated state below 5 % external relative humidity and in a hydrated state at 80 % relative humidity. The 30 % relative humidity spectra are left out due to no significant difference indicating no or just a very weak hydration dependence [Whi03]. A weak hydration dependence is also supported by the measurements on the right side showing hydration dependent absorption spectra of DNA and BSA as well as denaturized BSA [Mar00].

Cytochrome C was investigated together with polypeptides as lyophilized powder at room temperature, showing featureless absorption behavior [Yam02, Yam05]. Hydration dependent room-temperature data of BSA, collagen and cytochrome C were published by the group of Markelz in which the contrary findings of a weak or a strong hydration effect (fig. 4.6 on the right compared to fig. 4.37) seem to depend on the employed analyzing method [Mar00, Mar02, Mar03, Che05]. Also thin films of lysozyme were measured at controlled room temperature and under variation of the surrounding humidity [Kna06, Kna07a, Kna07b]. Fig. 4.7 shows hydrated and dehydrated myoglobin spectra and the corresponding hydration dependence that has been analyzed by another group. They found a strong hydration dependence but used the same data treatment as the Markelz group [Zha04, Zha06].

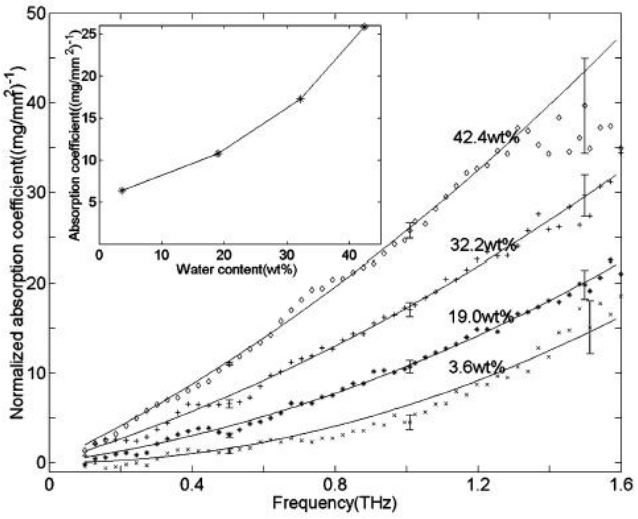

Fig. 4.7: Hydration dependent THz absorption of myoglobin. The line is a quadratic fit through the data points. Oscillations are due to etalon effects and the actual spectra can be assumed to be smooth and free of identifiable features. The inset shows the hydration dependence of the absorption coefficient at 1 THz [Zha04].

A real experimental breakthrough for the THz have been the publications from Xu, Plaxco and Allen on measurements of pure liquid water and of BSA and lysozyme in aqueous solution, already presented in figures 3.11 and 3.12 [Xu06a, Xu06b, Xu06c]. The strong absorbing liquid water has been surmounted by using a free electron laser as a THz radiation source for the frequency range from 10 cm^{-1} to 124 cm^{-1}. A distinction between liquid water absorption and protein absorption was difficult and just derived by withdrawing liquid water absorption values. The existence of unknown absorption due to biological water was treated in a general approximate way. But this was the only possibility to derive the contribution of the protein to the spectrum in such a three component system. To complete this review with measurements of other important biomolecules, the THz absorption of DNA and RNA compounds has been measured from Fischer et al. [Fis02, Fis05].
When taking out a lot of narrow peaks above the linear increase, usually caused by etalon effects [Mar08, Bal08], in general all measurements on proteins agree in a featureless increasing absorption with increasing frequency, making a distinction between different proteins very unlikely in the THz range as it is supported by model calculations [Kes00]. Here the advantage of far-infrared measurements on proteins becomes most evident. Even measurements on completely different small inorganic substances like liquid benzene, cyclohexane or tetrachloride show a similar absorption to proteins without recognizable features in this frequency range [Ped92].

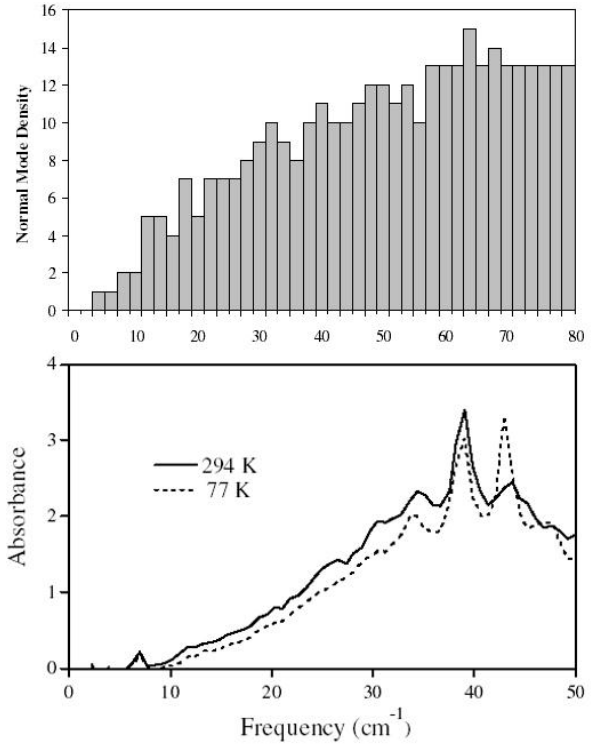

Fig. 4.8: Comparison of a normal mode density histogram and the THz absorption spectrum of hen egg white lysozyme. The linear increase in absorption with frequency is modeled correctly but a different frequency axis is used for both graphs which has to be considered because this is not obvious at the first sight. The decrease above 40 cm^{-1} is certainly due to an experimental uncertainty in the upper frequency range. Narrow peaks are caused by etalon effects overlaying the protein spectrum [Mar02].

Painter et al. [Pai82] pointed out that bigger molecules should have a contribution lower in frequency than small ones situated below 100 cm^{-1} based on a very simple model. This has not been confirmed, since a continuously increasing frequency spectrum and no single low frequency dominating mode is found in the THz [Cho86]. More sophisticated intends to explain the THz absorption of proteins have been made by normal mode analysis results which have been compared to measurements as seen for instance in fig. 4.8. The normal mode density calculation, unfortunately without any dipole inclusion, shows an almost linear increase, too. But this increase goes up to 60 cm^{-1} before reaching a plateau. The experimental absorption spectrum starts to decrease again from about 40 cm^{-1} on to the upper frequency cutoff at 50 cm^{-1}; most likely due to an experimental uncertainty. Both hen egg white lysozyme and myoglobin spectra are very similar and show the same behavior. The normal mode distribution is assumed to cause a higher normal mode density for myoglobin than for lysozyme but the actual calculated and experimental absorption spectra do not differ

significantly between these two proteins underlining the unreliability of absolute absorption values calculated on this way. Therefore the authors concluded that the gained featureless THz spectra are not sufficient to distinct between different proteins. Etalon and scattering effects have been confessed being present in these fine experiments. But nevertheless this publication has been one of the first comparisons of experimental and (normal mode) calculated spectra in this frequency range [Mar02].

4.1.4. Results from other Techniques and Frequency Ranges

Many different experimental techniques are used to study biological matter. Here we will not give a complete overview but present important findings which are necessary for the later discussion of our measurements.

As a first approximation Raman scattering yields complementary information to infrared absorption in small molecules. In large molecules without any symmetry like proteins usually the same vibrational bands are seen. But the intensity of Raman scattering is related to polarizability changes caused by nuclear motions and not to dipole moment changes as for infrared spectroscopy and hence different band intensities can be expected. An advantage of Raman compared to infrared is that water losses are weak and do not hinder the protein absorption in that manner as water does it in many infrared regions. Therefore early Raman studies have been already able to analyze proteins in solution for which Edsall has been the pioneer. Nevertheless Raman bands are usually weak and therefore hard to study. A nice technique called resonance Raman scattering can overcome this problem. If the questionable Raman band is coupled to an electronic transition, a signal enhancement is observed. This is done for instance in investigations on chromophores and active sites in proteins, where an additional electronic absorption is excited, too.

One of the first low frequency Raman measurements on proteins has been done on α-chymotrypsin which confirmed low frequency modes in proteins [Bro72]. Large differences are seen by the used method of sample preparation which effect the low frequency modes significantly (fig. 4.9). This is an important finding we have already used and will use in our argumentation and is one reason for our extensive reproducibility checks. A peak at 29 cm^{-1} was observed at the same frequency where the boson peak of glasses is to be found. It remains unclear if this peak is always present at room temperature as seen here. Low frequency

Raman spectroscopy in the range from 1 cm^{-1} to 250 cm^{-1} on lysozyme crystals and DNA confirmed such a peak at room temperature in this range [Ura98].

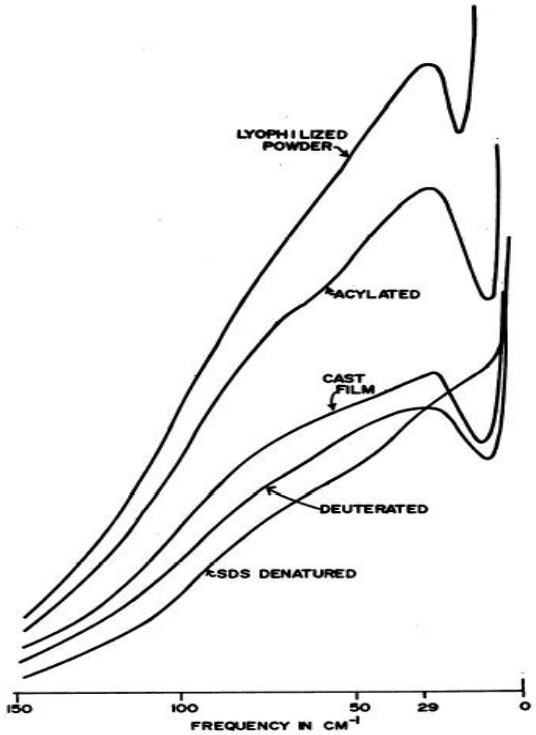

Fig. 4.9: Low frequency Raman spectra of α-chymotrypsin prepared in different ways and in two denaturized states. A band at 29 cm^{-1} is clearly seen which is at the same frequency where the boson peak in neutron scattering for instance occurs, but usually at lower temperatures. Some people claim this to be an argument for a spectroscopic observable and at room temperature occurring boson peak in proteins [Bro72].

OHD-RIKES – optical heterodyne detected Raman induced Kerr effect spectroscopy (details of this technique in [Lot95]) - is claimed by the authors that "this technique has proven to be superior to infrared and Raman spectroscopy at low frequencies (<400 cm^{-1}) and its high signal to noise ratio allows a detailed analysis of the spectra" [Gir03]. But in contradiction, the authors state in the same paper that their effective resolution is 9 cm^{-1} and that "the validity of the data above ~300 cm^{-1} should be regarded with a certain degree of suspicion" and large differences concerning the liberational peak of water around 600 cm^{-1} to the far-infrared are seen to these and Raman results on liquid water spectra. Nevertheless we want to present the spectra of α- and β-lactoglobulin, pepsin and lysozyme in aqueous solution, because they are clearly different from the liquid water spectrum as seen in fig. 4.10. This is in contradiction to the far-infrared spectra of proteins in solution [Xu06a, Xu06b] where

almost the same shape of absorption is gained as for the (slightly more absorbing) pure water. Also the far-infrared measurements just reveal a linear increase up the highest measured frequency of 124 cm^{-1} and did not peak in the whole range from 10 cm^{-1} on. Selection rules for Raman and infrared typically involve symmetry which should play no significant role in complex systems like proteins. From a theoretical point of view these different findings for proteins are puzzling and more or less the same peaks for protein absorption spectra are expected to occur. If the controversial water absorption peak at 60 cm^{-1} is neglected, the Raman and far-infrared peak positions on pure water are the same with a slight shift of about 10 cm^{-1} of the Raman water peak to lower frequencies. For the pure protein spectra, all proteins exhibit a significant broad peak in the OHD-RIKES spectra around 80 cm^{-1} with slight changes in the slope towards higher frequencies, derived after a "hand-scaled" subtraction of water. That the peak at 80 cm^{-1} is not seen in inelastic neutron scattering is explained by the authors that (torsional) backbone motions and not hydrogen atoms involving motions are the origin of this peak as it is suggested by Simonson and Perahia [Sim96], because inelastic neutron scattering is proportional to the density of vibrational modes involving hydrogen atoms. But Raman scattering is related to the density of states, including all types of atoms, multiplied by the square of the polarizability tensor derivative.
The imaginary part of the OHD-RIKES spectrum R(ω) is related to the low frequency depolarized Raman spectrum $R_{DRS}(\omega)$ via the Bose thermal occupation factor [Gir03] by

$$\operatorname{Im} R(\omega) = R_{DRS}(\omega)\left[1-\exp\left(-\hbar\omega / k_B T\right)\right]. \tag{4.1}$$

A comparison on this way to spontaneous Raman scattering measurements on lysozyme from Genzel et al. [Gen76] found that the results have been reproduced with these measurements. So despite the expected similarity, the peak positions in proteins seem to be different derived from these methods compared to far-infrared protein spectra which has to be regarded in our discussion. The origin of this is very questionable and the answer might contain important information for the understanding of protein dynamics in this frequency range.

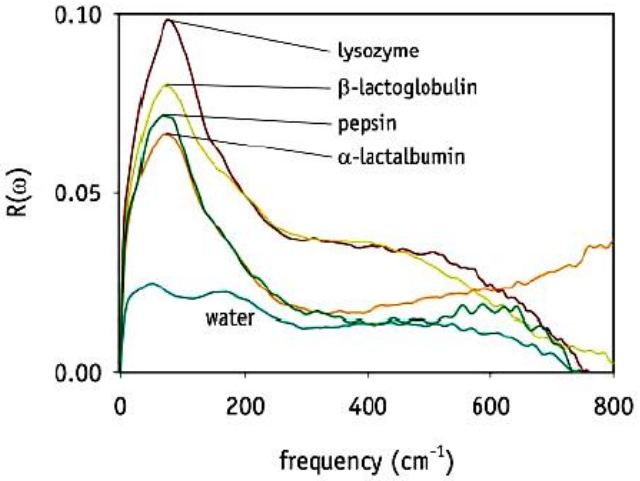

Fig. 4.10: Raw OHD-RIKES spectra of α- and β-lactoglobulin, pepsin and lysozyme in aqueous solution compared to liquid water which is scaled to the proportion of the water content [Gir03].

X-ray crystallography is based on the scattering on the electron shell of protein crystals which are not possible to be produced for all proteins and which is very complicated to gain large size crystals. Sometimes crystallization is just reached by introducing stabilization atoms or molecules which also alter the protein structure to an unknown amount. Static and dynamical disorder cannot be distinguished properly and hydrogen atoms are hard to be resolved. Averaging the experimental results over a long time causes an inability to discriminate between real motions and static disorder. Just by analyzing the temperature dependence of the Debye-Waller factor it is possible to distinguish between these two contributions; but just under the assumption of a temperature independent static disorder. The assumed isotropic Debye-Waller factor is therefore given by

$$B = 8\pi^2 \left\langle x^2 \right\rangle = 8\pi^2 \left(\left\langle x^2 \right\rangle_{ld} + \left\langle x^2 \right\rangle_{d} \right) \tag{4.2}$$

where the first term is the lattice or static disorder and the second term disorder due to motions [Kur95].

Neutron scattering is in contrary to X-ray mainly hydrogen atom sensitive because scattering off these nuclei is ten times larger than with other nuclei including deuterium [Dos90, Cha03, Dos05]. Water and protein-internal motions are separated by proceeding the experiment once with H_2O and then with D_2O. Complementary to X-ray experiments, the protein dynamic behavior is derived indirectly from the hydrogen atom motions. The experimentally gained

quantity is the incoherent dynamic structure factor $S(\boldsymbol{q},\omega)$, $\hbar\boldsymbol{q}$ and $\hbar\omega$ are respectively the momentum and the energy transfer between the system and the incident neutron [Dos89]. In contrast to neutron scattering experiments far-infrared and THz measurements on proteins contain additional information about the nature of protein motions and the hydration water [Moe92]. Sample size and setup depending data collection times restrict inelastic scattering experiments [Dos08].

Nuclear magnetic resonance uses the splitting of energy levels in an external magnetic field due to the spin of the nuclei. Most important for biological samples are the nuclei ^{1}H, ^{13}C, and ^{15}N, which split into two energy levels due to spin quantum number ½. Information about the local environment comes from slight different resonances, called chemical shift, which is a very complex quantity in large molecules like proteins [Nie96]. One of the first NMR studies analyzed hydrated and temperature dependent BSA, in which also a very nice introduction to NMR on biomolecules is given, too [Ful68].

One reason for the discrepancy in the small and large number of derived biological water molecules between NMR and neutron diffraction experiments is explained by Gu and Schoenborn, that the average residual time of bound water molecules are too short to be detected by NMR experiments [Gu95]. This complicates the discussion about the real outspread and number of the biological water shell molecules around proteins in the later presented discussion.

Mössbauer spectroscopy is mainly used for iron containing proteins like hemoglobin due to the fact that most often ^{57}Fe is used. The emitted gamma quant is used for absorption spectroscopy with an extremely small energy width. But since just the same types of nuclei can absorb the radiation, the above named limitation is given. Also dynamic information about the motions around the absorbing atom can be gained from Mössbauer experiments by an analysis of the line width. Since just the dynamics of the (iron) atoms is observed, the general information about protein dynamics is missing but the advantage is the direct probe of conformational changes at the active center of this protein and hence likely the functional important conformational changes. This is the origin of the knowledge about the described functional relevant collective motions and their timescales reaching into the far-infrared frequency range.

4.2. The Protein Samples

Four different proteins have been extensively studied for this thesis. They differ in their size and their secondary structure content and can be seen as some cross section of different kinds of proteins. But due to the time consuming measurements the number is unfortunately yet too small for a direct relation to different protein properties in a statistical manner. On the other hand the observed differences in the far-infrared spectra led to a clear distinction between all proteins which is very promising to reach such a relation between spectra and protein structure one time. For instance the small differences in mid-infrared amide absorption bands, analyzed for just twenty different proteins, led to an implementation for industrial applications to derive the secondary content of unknown proteins. As an introduction to our measurements we therefore present our four measured proteins together with their properties and their corresponding far-infrared spectra.

4.2.1. Bovine Serum Albumin

The largest measured protein is Bovine Serum Albumin which is the most abundant protein in blood plasma of cattle. The functionality is related to the transport of substances and to the regulation of osmotic blood pressure. It is widely studied due to the similarity to human serum albumin whose primary sequence is about ¾ analog and the secondary/tertiary structure is even more similar as fig. 4.11 shows.

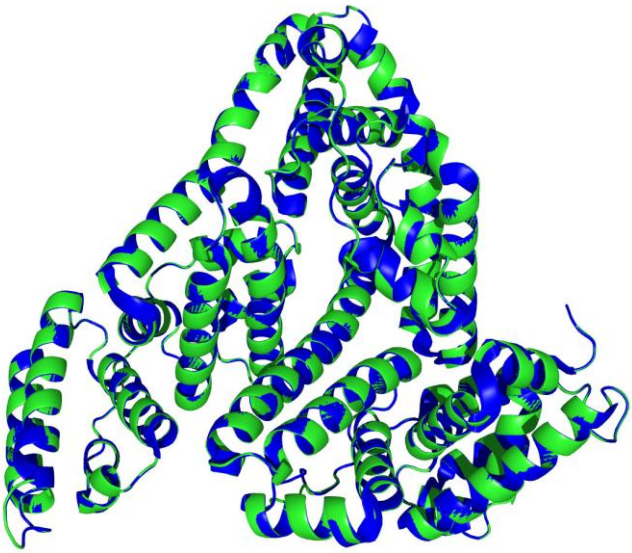

Fig. 4.11: Structure comparison of bovine serum albumin (green) and human serum albumin (blue) [PDB08].

Solely α-helices are found in this protein. 583 amino acids belong to the actual protein without precursors summing up to a molecular weight of 66.4 kDa.
Purchased lyophilisized powder from Sigma (A2153), was used without further purification. The measurements presented here were gained on free standing protein films derived from 120 µm drops of a $7.53 \cdot 10^{-8}$ mol/ml solution. The corresponding far-infrared spectrum is presented in fig. 4.12 together with a five spot comparison at 24°C and 75 % relative humidity.

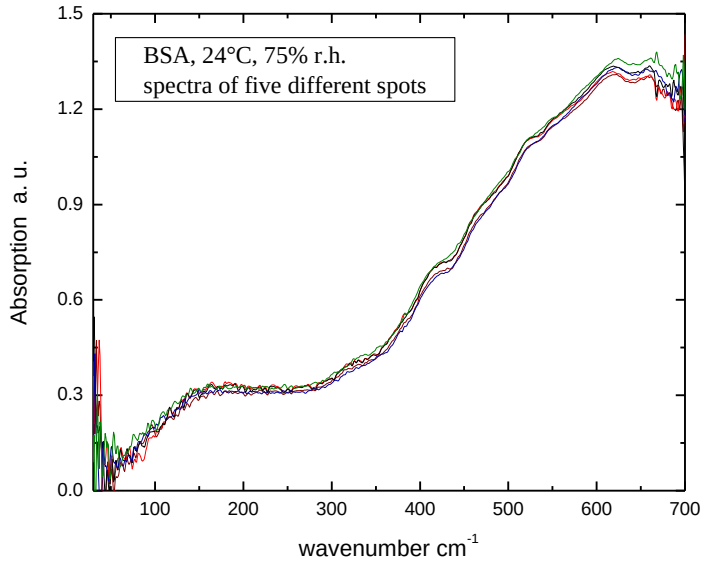

Fig. 4.12: Far-infrared spectra of bovine serum albumin measured at five different spots of the same protein film at 24°C and 75 % relative humidity.

There is a controversial debate about a proposed change in the secondary structure content of BSA based on circular dichroism measurements of BSA solutions at pH 7 [Tak89, Tak93]. Above 30°C a decrease in the α-helical content sets in, getting stronger above 40°C. A comforting finding is that below 45°C this change was completely reversible and that other authors found in solutions at pH 8 no structural change in BSA up to 42°C. The whole discussion is summarized in the above cited two publications. From our measurements we have no indications for such a transition.

4.2.2. Trypsin Inhibitor

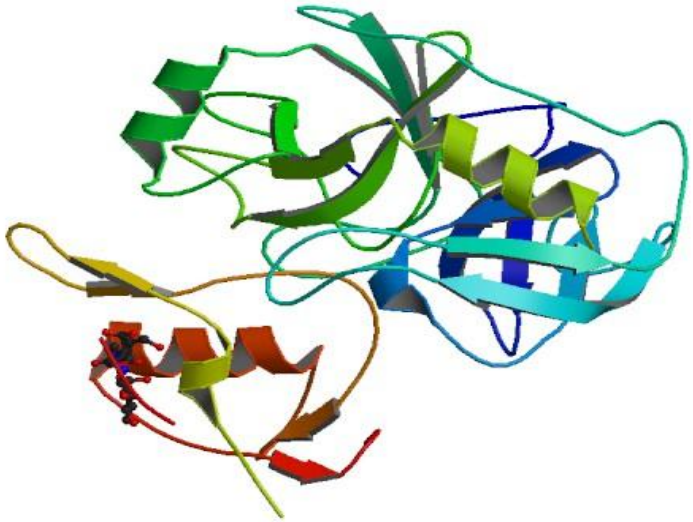

Fig. 4.13: The structure of turkey egg white trypsin inhibitor [PDB08].

Consisting of 186 amino acids and with a molecular weight of 28.0 kDa trypsin inhibitor is significantly smaller than the prior discussed BSA. As the name indicates the function of this protein is the inhibition of the hydrolytic (digesting into smaller compounds) activity of the protein trypsin. Trypsin inhibitor is also called ovomucoid and is the main cause of allergic reactions to egg white [You02]. Helical and sheet structures are both present. As a basis for the sample preparation we used here 120 µm drops of $1.79 \cdot 10^{-7}$ mol/ml solution of trypsin inhibitor from hen egg white (Sigma; T9253) whereas fig. 4.13 pictures trypsin inhibitor from turkey egg white. The far-infrared spectra of five different spots are presented in fig. 4.14.

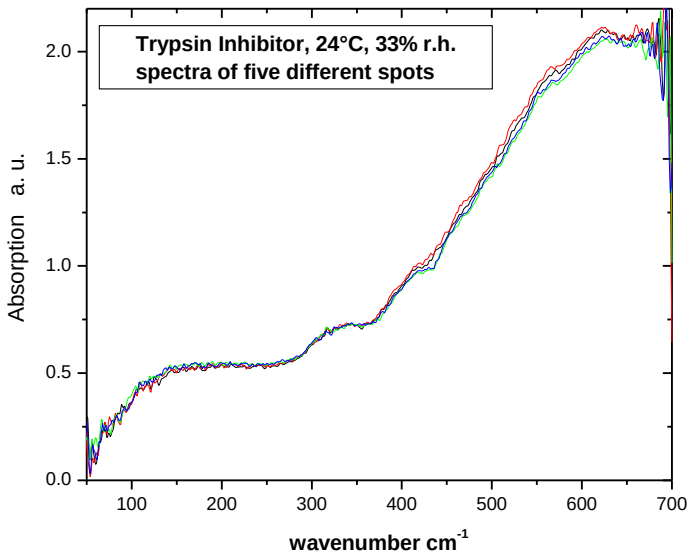

Fig. 4.14: Spot comparison for a trypsin inhibitor sample measured at 24°C and 33 % relative humidity.

4.2.3. β-Lactoglobulin

β-lactoglobulin occurs quite numerous in (bovine) milk. The exact function seems to be still unclear but some relation to the binding of ligands was confirmed. It easily forms gels by aggregation and is therefore often used as a thickener in food industry [Hoe02]. The 162 amino acids sum up to a molecular weight of 18.3 kDa.

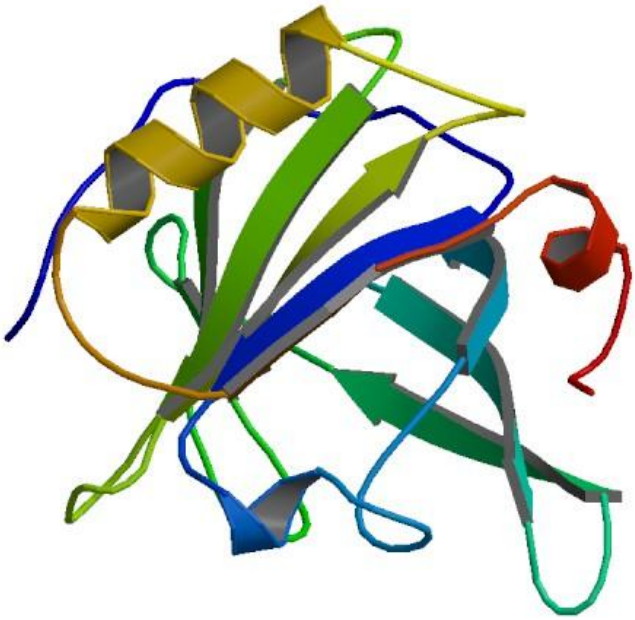

Fig. 4.15: The structure of β-Lactoglobulin [PDB08].

There is a large amount of random coil and the largest amount of β-sheet formation per mass causing β-lactoglobulin to stand out of the other proteins in the measured group (fig. 4.15). This might be a possible explanation for the most different far-infrared protein spectrum and the contrary low frequency temperature dependence. A spot comparison at 24°C and 53 % relative humidity is presented in fig. 4.16. For this protein a 120 µm drop of a $2.73 \cdot 10^{-7}$ mol/ml solution was the starting value for the measured β-lactoglobulin films. The initial solution was made from lyophilisized powder from Sigma (L0130) which contains β-lactoglobulin from bovine milk.

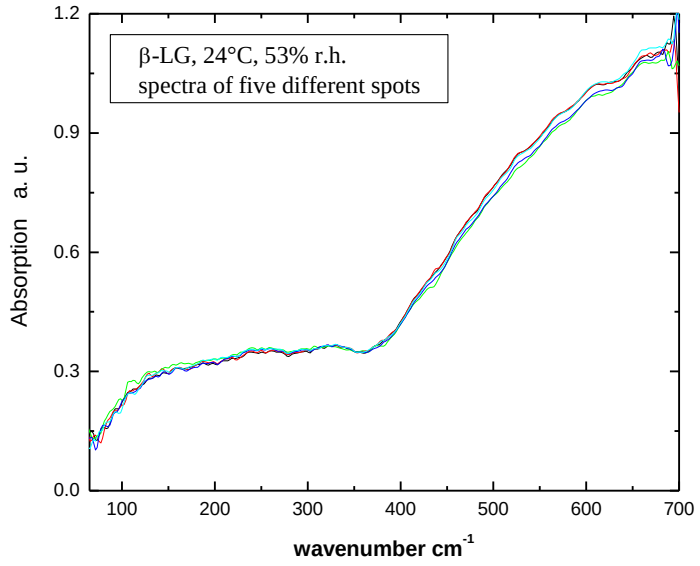

Fig. 4.16: 5 different spots of a β-lactoglobulin sample measured at 24°C and 53 % relative humidity.

4.2.4. Hen Egg White Lysozyme

The smallest of the four measured proteins is hen egg white lysozyme with 129 residues and a weight of 14.3 kDa presented in fig. 4.17. Here the α-helical content dominates with 41 % over the β-sheet content of about 10 %. Lysozyme proteins occur quite frequently and their function is mainly the break down of bacteria cell walls involving a hinge bending motion as already described in the theory chapter. The functionality was already derived along the original discovery of the protein from Fleming in 1922 [Eds83].
Human and (here studied) hen egg white lysozyme differ in their main crystal forms of tetragonal and triclinic structure respectively, and in their internal water content as known from X-ray experiments [Mou76]. These early findings are based on the easy crystallization property of lysozyme as already mentioned in chapter 3.2.2. By using lyophilisized powder from Fluka (62970) the samples have been made from 120 μm drops of $1.74 \cdot 10^{-7}$ mol/ml solution leading to the absorption spectra plotted in fig. 4.18.

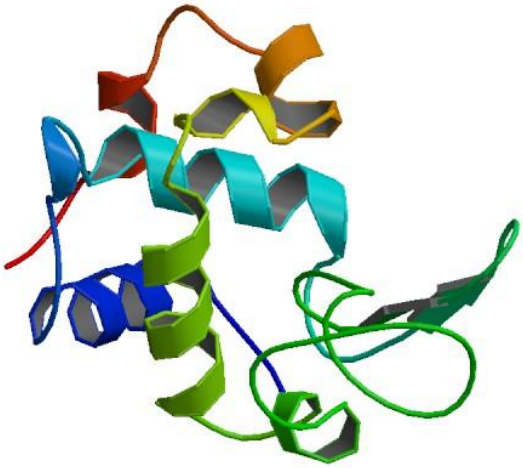

Fig. 4.17: The structure of hen egg white lysozyme [PDB08].

X-ray structure analysis revealed different possible structures for lysozyme, sometimes assumed to be temperature or hydration dependent which would then affect also our temperature and hydration range [Ber83b, Kod90, Bon96]. But under variation of both parameters we have no indication in our far-infrared spectra for such a structural transition. Our results provide strong arguments against such transitions which we would like to discuss right here in this part. From a theoretical point of view and from calculations on proteins, collective motions should be highly affected by any (tertiary) structural change. Therefore we should see for lysozyme and maybe for BSA an effect which should be absent in the other two measured proteins. But there is nothing to be observed under the variation of both parameters which are assumed to alter the structure leading to our above stated conclusion.

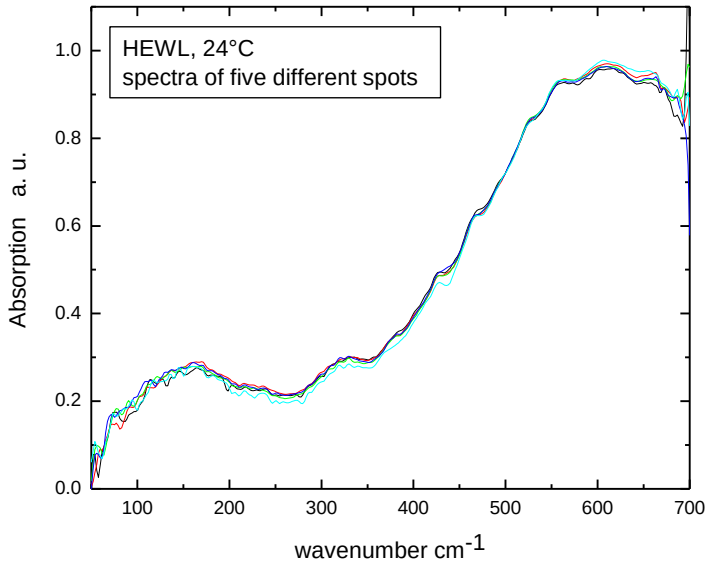

Fig. 4.18: Spot comparison for hen egg white lysozyme at 24°C. The absence of defined humidity control in this particular set of measurements causes the small fluctuations present mainly in the lower frequency range.

4.3. Reproducibility Tests

Biological samples have a natural variation which restricts reproducibility. Reasons are impurities or functional disabilities combined with a large sensitivity to the sample preparation. Therefore all presented hydration and temperature dependent spectra have been measured at least two, but usually three times, for all possible 60 combinations of hydration and temperature settings. Hence just a much smaller number of spectra can be presented here in this thesis, focused on those which are necessary for a comprehensive discussion and to emphasize the relevant property. Therefore fig. 4.19 compares different samples of which sample 1B and 2B have been made within the same process and 3B in a different process. For a better comparability have been the spectra normalized at 350 cm^{-1} to withdraw different layer thicknesses.

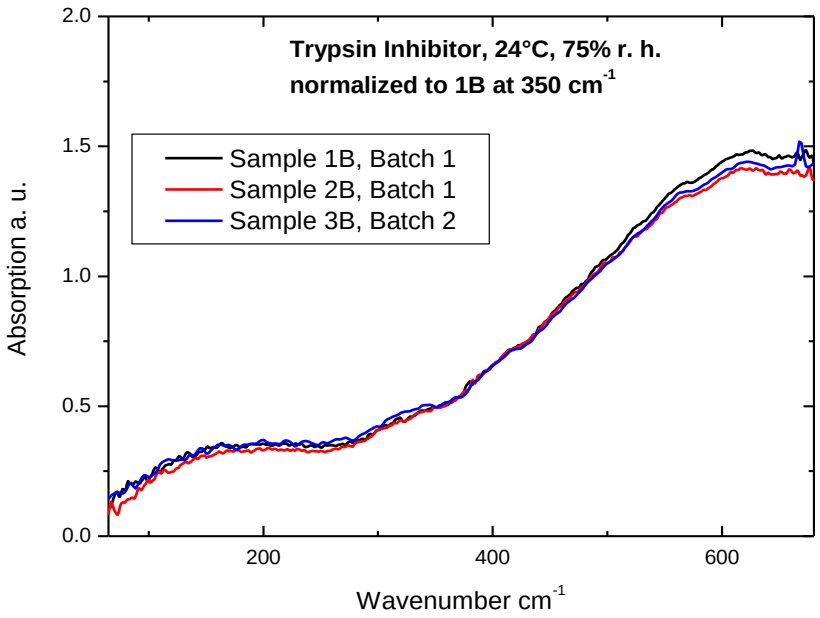

Fig. 4.19: Comparison of three samples of Trypsin Inhibitor. 1B and 2B made within the same sample preparation process; 3B in a different one. A normalization at 350 cm^{-1} to sample 1B has been applied to withdraw minor thickness or density differences.

All spectra show the same overall slope and reveal the same absorption bands. Slight deviations are present in particular concerning absolute values at the more absorbing high frequency range. An important result is that the spectra of samples made within different sample preparation processes do not exhibit larger deviation than samples made within the same process. So the sample preparation causes less deviation than the differences within the samples itself. Since different spots on one sample exhibit also just little variation, as it will be discussed later in this section, the origin for that seem to be impurities and more likely local differences in each sample holder during the freezing or following lyophilization process. The difference of the absolute absorption values is around 1.5 % in the low frequency range and around 2.5 % in the high frequency range. But even the small absorption features present in the spectra are reproducible and can be referred to the protein spectrum.

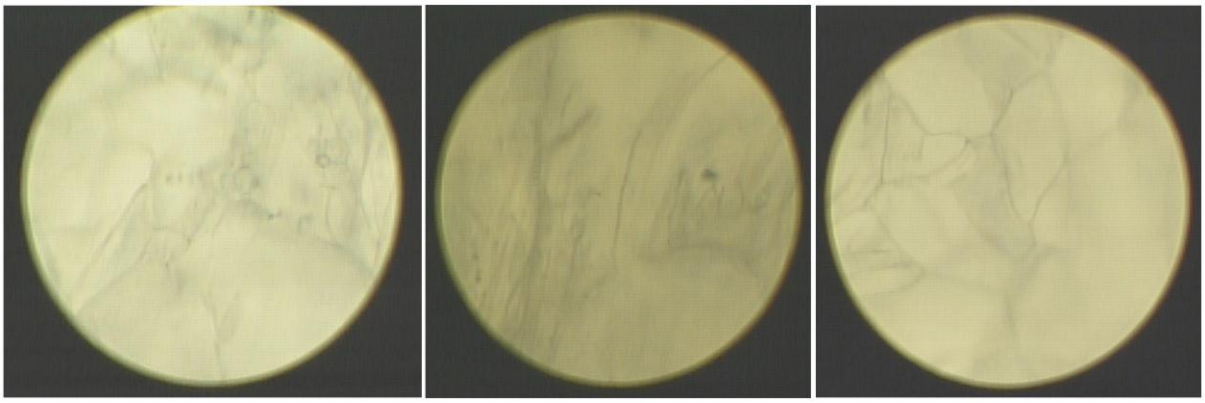

Fig. 4.20: Surface picture of the samples 1B(S2), 2B(S3) und 3B(S4) from fig. 4.9. The diameter of the sample spot is 250μm.

Fig. 4.20 shows the corresponding surface pictures of the sample spots from fig. 4.19 in transmission. The cotton wool like consistence of the samples is displayed in the surface structure. Protein filaments are the main features which are actually used by us to find again the same spot in this 3-dimensional arrangement after the background measurement. A crystallization was successfully hindered with our improved sample preparation at least for the relevant size for this frequency range (around 15 μm to 150 μm wavelength) based on two reasons. Firstly, to a large extend crystals above 50 μm size could be excluded by optical investigations under the microscope in transmission and reflection mode. Secondly, the spectra of highly scattering samples, caused by crystallization, differ significantly as we described it in the prior chapter during our sample preparation improvement procedure. Such a slope was never seen again after the improvement of sample preparation which was used for all presented results in this thesis.

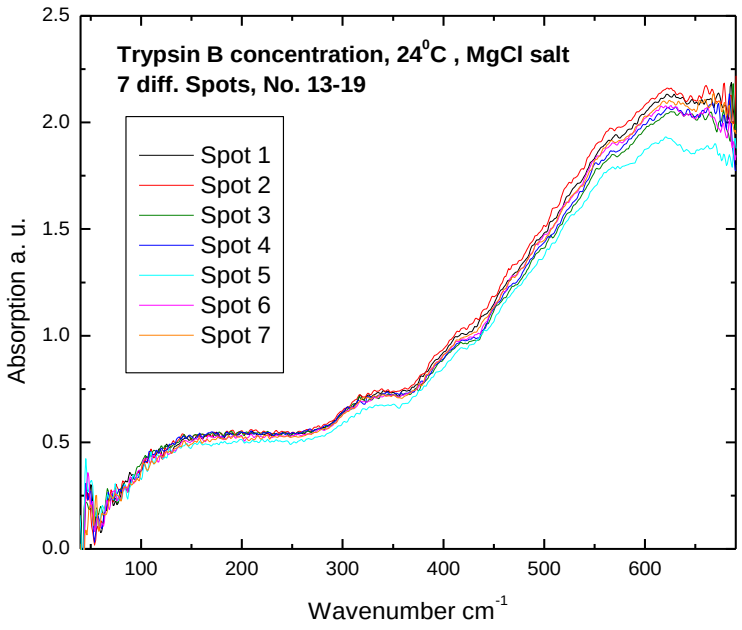

Fig. 4.21: Comparison of seven different spots on one sample to emphasize the reached sample homogeneity and reproducibility.

For our later presented temperature and hydration dependent measurements just the deviations during the measurement of one single spot are relevant, since relative changes under the influence of these parameters are studied. They will be demonstrated to be much smaller than the above presented absolute value differences. As fig. 4.21 shows one sample reveals slight variations in thickness or density. This has been overcome later by using the same spot for a complete measurement run allowing a further reduction of deviations and hence the accuracy concerning relative changes under the variation of temperature and hydration.

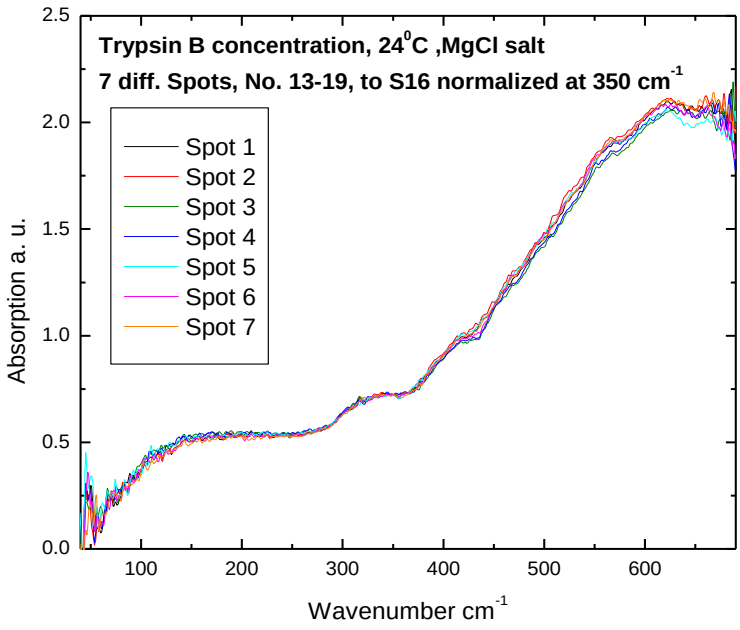

Fig. 4.22: Normalized comparison of seven different spots on one sample to emphasize the reached sample homogeneity and reproducibility; normalized to the absorption of sample spot No. 4 at 350 cm^{-1}.

The seven different measurement spots show all the same absorption bands as it can be better seen in the normalized fig. 4.22. To ensure a sample homogeneity just samples which had at least four sample spots in common, but no different spot spectra, were used for actual measurements in which just one spot was used throughout the whole procedure.

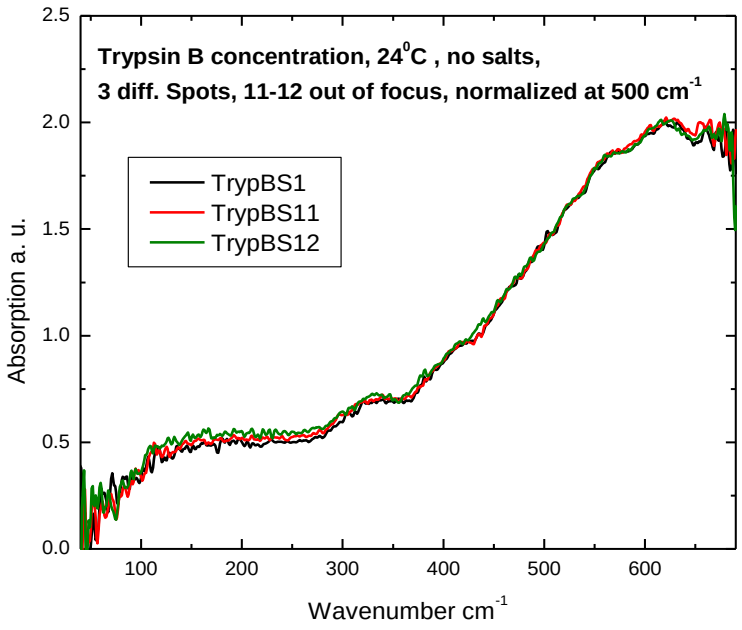

Fig.: 4.23: Spectra with different focusing on the sample surface, normalized on the absorption at 500 cm^{-1} for graph S1. No control of hydration was applied which explains the present additional noise.

That even a very bad focusing on the sample has little influence on the spectra can be derived from fig. 4.23. In general the same slope and the same absorption bands are found. Since focusing is difficult within a 3-dimensional protein web, this test demonstrates that minor focus changes do not influence the spectrum. Here extreme out of focus spectra have been taken which caused just little consequences to prove the independence from this experimental setting.

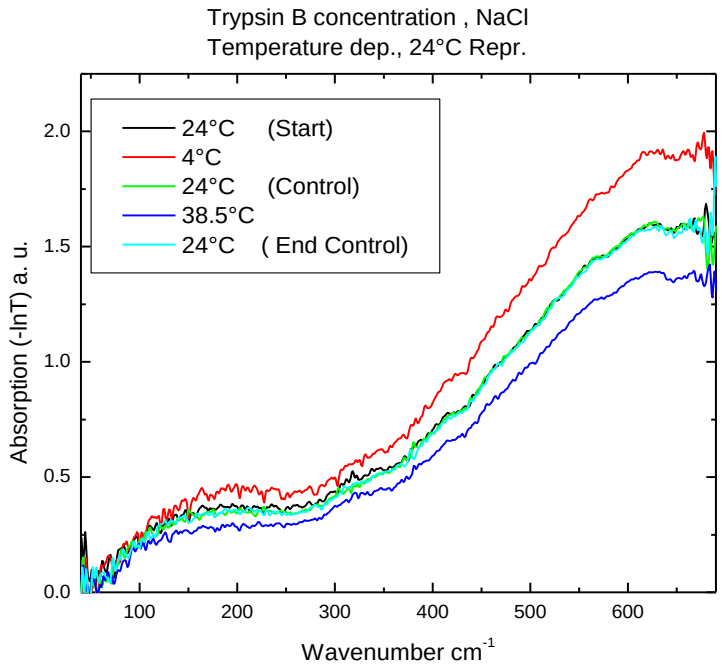

Fig.: 4.24: Typical temperature dependent measurement (in this case on trypsin inhibitor) with a three times retest of the 24°C absorption spectrum. The 24°C spectra have been taken at the beginning, after coming down from 38.5°C and at the end after the 4°C measurement. A good reproducibility of all features is demonstrated.

The temperature dependent measurements have been tested directly on their reproducibility and can be used to estimate the error bars. As fig. 4.24 shows the 24°C spectrum has been taken at the beginning, in between and at the end of the roughly ten hours enduring measurement. At the beginning the fluctuations within the cell seems to be larger than for the following measurements, since the control and end control graphs match better. Maximum care to get back to the exact same spot has been taken. It can be clearly seen that the temperature effect is much larger than the deviations of the three test measurements. All features are as much pronounced but the start measurement has slight higher absolute absorption values below 400 cm^{-1}. The error bars for the temperature dependent measurements can be hence estimated to a maximum of 2 %. Actually the error bars in the presented measurements are smaller because of the installation of our additional covering around the whole setup and spectrometer, which was unfortunately applied after all these reproducibility tests presented here. In summary, the observed absorption change caused by temperature is around 10 % for all proteins based on the 24°C graphs and is therefore well above our error bars.

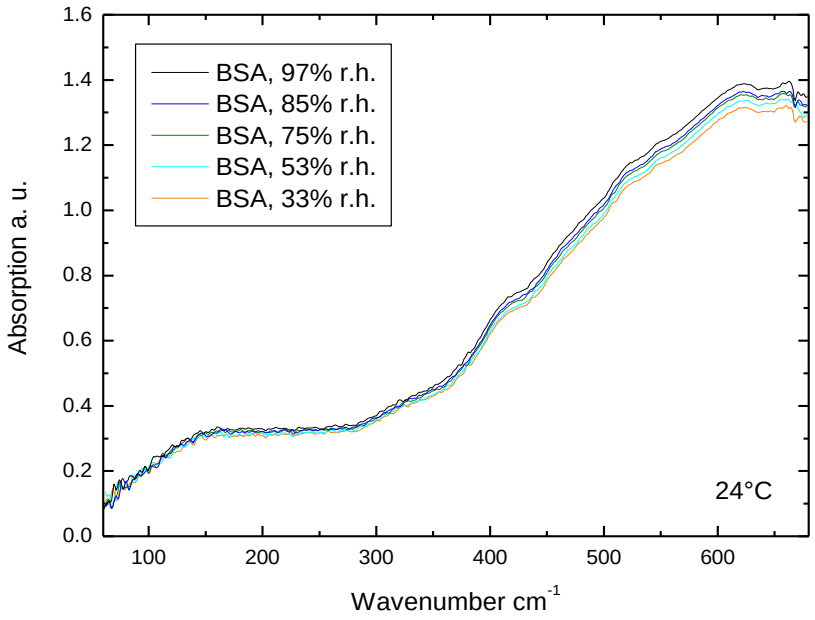

Fig.: 4.25: Typical hydration dependent measurement. Since all absorption spectra are very close together, especially at the lower frequency side, but still ordered in the right way, this proves our very small error bars.

As it can be seen from fig. 4.25 all hydration dependent spectra are very close together. Even at the lower frequencies, where differences are small, the graphs are ordered along increasing absorption with increasing hydration. The error bars are hence smaller than the differences between two adjacent graphs because otherwise the order would have been altered at least for some measurements. But all of them showed the same exact order with increasing hydration. We therefore conclude that the error bars are smaller than 1 % for these kinds of measurements. This is very surprising because the spectrometer and the atmosphere cell have to be held stable for a minimum of at least 14 hours. But for the hydration dependence the atmosphere cell seems to be driven out of equilibrium much less than upon temperature variations. The error bars become significantly less even if the cell was opened completely to exchange the salt solution which is not necessary during the temperature variation measurements.

4.4. Basic Frequency Assignment

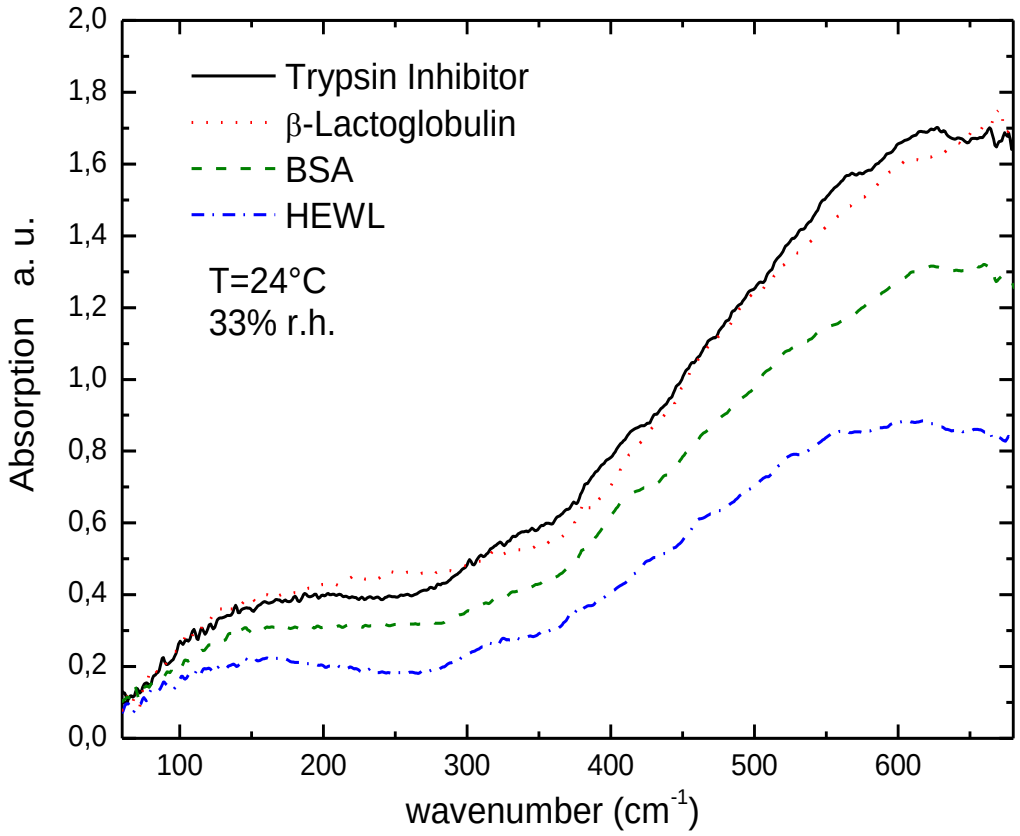

Fig. 4.26: The far-infrared spectra of all measured four proteins at 24°C and 33 % relative humidity. In contradiction to the THz range the spectra are clearly different to each other and exhibit several absorption bands.

We observed from our measurements that the four proteins can be clearly distinguished in the far-infrared region as demonstrated in fig. 4.26. There are several absorption peaks and shoulders present which we want to analyze further in this section. Normal mode analyses even based on molecular mechanics fail in the low frequency range as Joti et al. [Jot04] showed in their comparison to molecular dynamics simulation results. But unfortunately at the moment mainly normal mode analysis results, usually not derived for proteins but for smaller model compounds, are the only hint to date which motions might cause the measured protein spectra. With our here given basic frequency assignment we would like to encourage corresponding calculations of the infrared absorption spectra for which we provide the basis for a comparison with the here presented peak assignment.

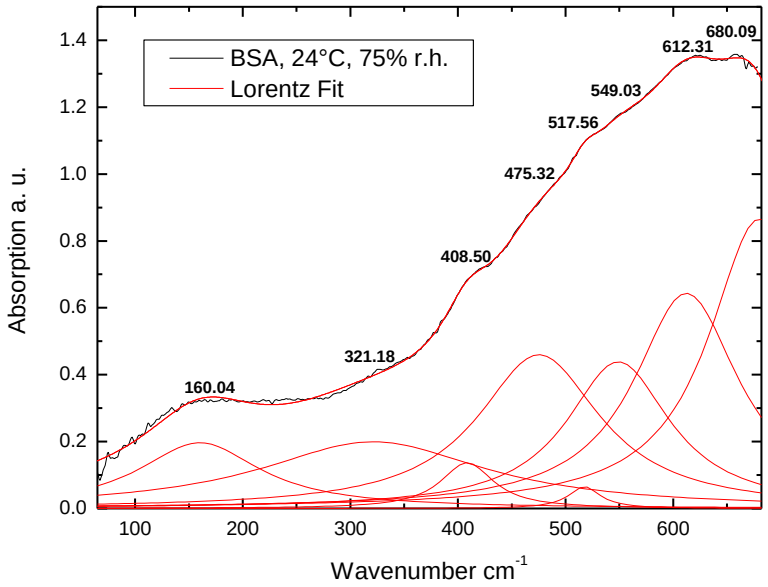

Fig. 4.27: Basic frequency assignment on the far-infrared spectrum of BSA at 24°C and 75 % relative humidity. The eight Lorentzian peaks and the combined fitted graph are shown in red color.

We decided to fit the protein spectra with several Lorentzian peaks without the inclusion of any physical model underneath for an assignment of the major contributors to the observed absorption features.

$$y = y_0 + \frac{2A}{\pi}\frac{w}{\left(x - x_c\right)^2 + w^2} \tag{4.3}$$

A is the area, w the width and x_c the center frequency of the peak. The offset y_0 was set to zero. To get reliable conclusions the fitting was done by two different scientists starting from the pure absorption spectra and the gained final results have been afterwards compared. The number of peaks necessary for the fitting was chosen to be as small as possible. By calculating the deviation of the fitted graph to the actual measured one, the actual number of peaks was reached if an additional peak did not lead to a large improvement anymore. There was always a direct consensus about the number of peaks from the two independent scientists. The results from the separate fitting on the same graphs turned out to differ in the contributions (e.g. the area A underneath every peak as seen in formula 4.3) but confirmed the peak positions with better than ±1 cm^{-1}. Therefore we restrict ourselves to the latter reliable

quantity in this discussion and left out the interpretation of the area value given by the fitting procedure, too.

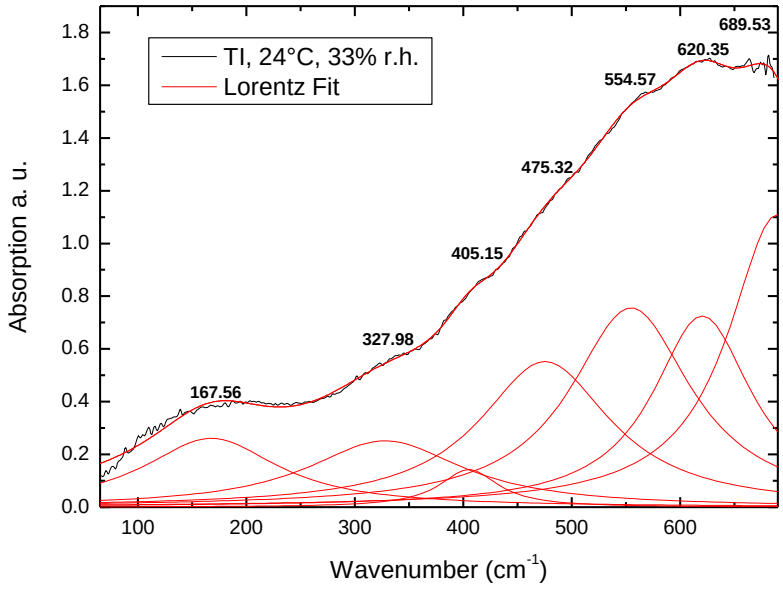

Fig. 4.28: Peak assignment of trypsin inhibitor measured at 24°C and 33 % relative humidity fitted with seven Lorentzian peaks.

We take the chance to present here a random sample for the many peak assignments derived at different hydration values for each protein. As we will see later just a small (almost frequency independent) hydration dependence has been found for all proteins which is confirmed by this peak assignment. For the below presented comparison we choose of course the same (lowest) hydration for all proteins. The effect of temperature on the peak positions will be discussed in the corresponding chapter 4.6.

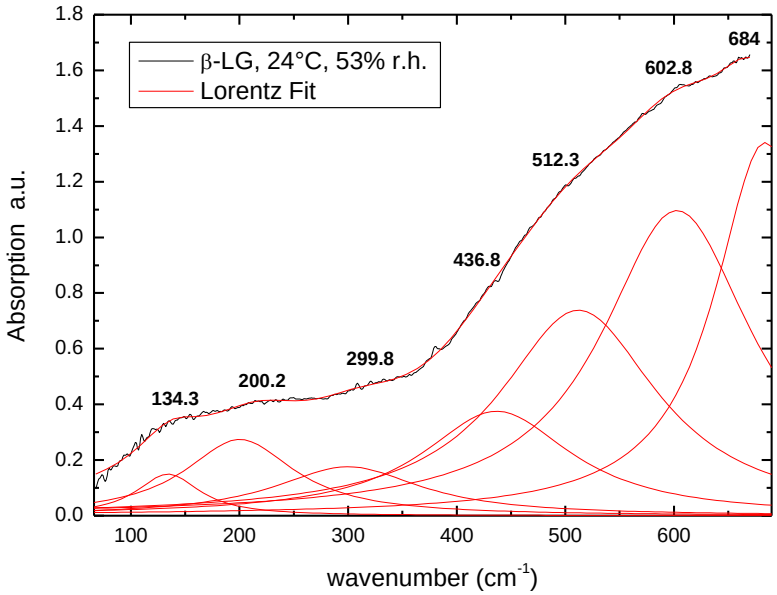

Fig. 4.29: Lorentz peak analysis of the far-infrared spectrum of β-lactoglobulin at 24°C and 53 % relative humidity. Here seven peaks were used to fit the spectrum properly.

The basic frequency analysis of the far-infrared spectrum of BSA is seen in fig. 4.27. The experimental parameters are 24°C and 75 % relative humidity. Eight Lorentz peaks fitted the spectrum properly and the peak positions are given in the graph. For all proteins except β-lactoglobulin (fig. 4.29) we fitted the broad low frequency absorption band on purpose with only one Lorentzian as we will see in chapter 4.6.2. discussing the temperature dependence of this particular peak. Therefore the deviations of the fitted graph to the actual measured spectrum is slightly larger in this region than for the part above. The spectrum of β-lactoglobulin (fig. 4.29) has more a shoulder than a low frequency peak which is most pronounced in the spectrum of lysozyme in fig. 4.30. Spectra of trypsin inhibitor (fig. 4.28) and β-lactoglobulin have been able to fit properly by just seven Lorentz peaks but which are very different to each other. The trypsin inhibitor spectrum is in general much closer to the spectra of BSA and lysozyme but interestingly for the latter two additional peaks occur in the upper frequency region.

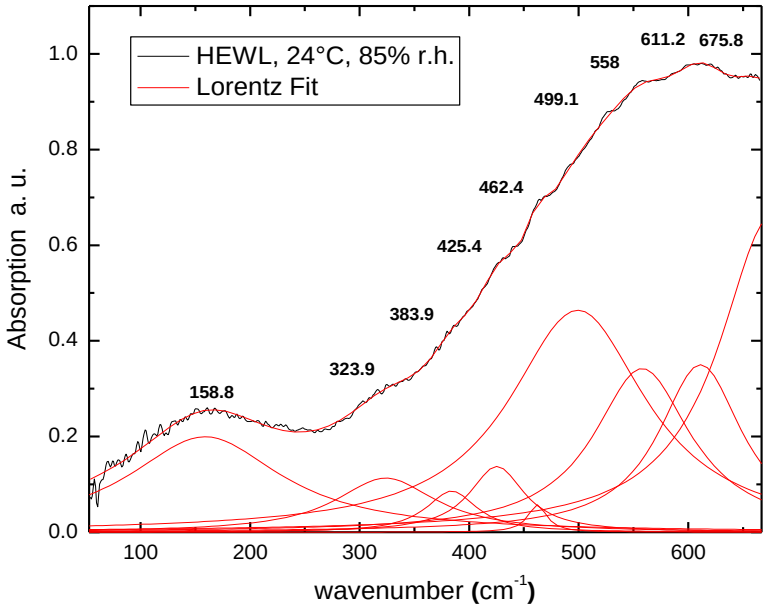

Fig. 4.30: Basic frequency assignment using nine Lorentz peaks for a hen egg white lysozyme spectrum at 24°C and 85 % relative humidity.

Fig. 4.31 shows the protein peak comparison at 24°C and for the lowest hydration of 33 % relative humidity. The highest frequency peak above 675 cm^{-1} can be attributed to a well known amide absorption band but it remains unclear if this is the amide IV or amide VI band, because both are assumed to occur in a broad frequency range between 630 cm^{-1} to 750 cm^{-1} which could even alter the ordering. It is more likely to see the amide VI since it is on average lower in frequency and no other broad peak below is present. This would mean that bending motions are the main cause of this absorption feature if the normal mode analysis from N-methylacetamide motions of fig. 4.1 is applied. The broad low frequency absorption peak below 200 cm^{-1} is usually referred to be the amide VII band whose origin is widely debated. The assumed motions have to be taken with caution since they are mainly based on normal mode analysis as well as small model compounds. Both restrictions are not capable to derive anharmonic collective motions assumed in this frequency range for many different reasons as discussed in the theory chapter before.

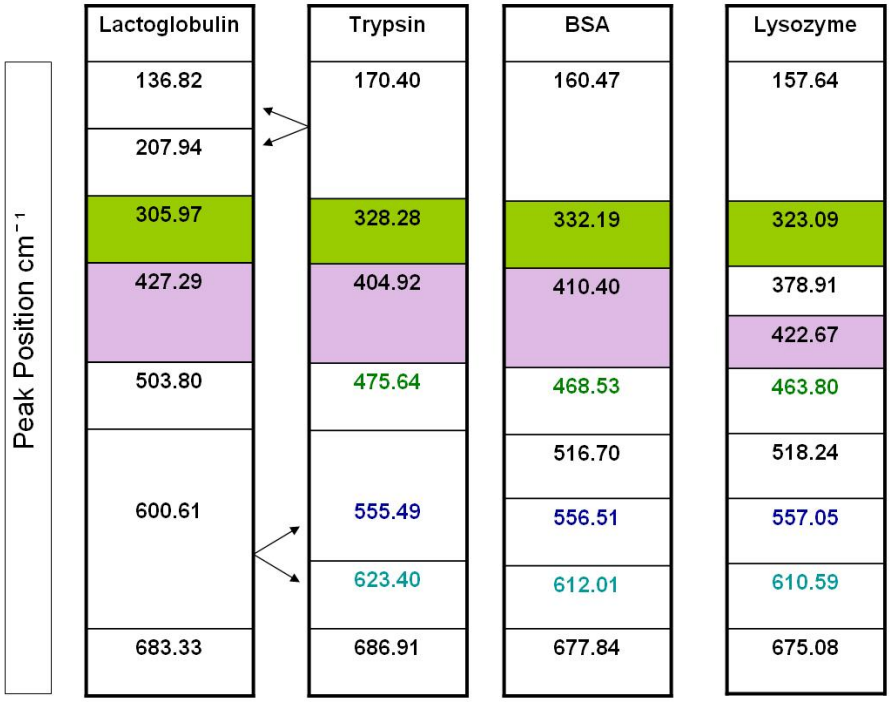

Fig. 4.31: Protein comparison of the peaks derived from the fitting of the far-infrared spectra at 24°C and 33 % relative humidity. A possible splitting of bands is indicated as well as the bands which seem to be in common. β-lactoglobulin stands out of the group due to a much more different absorption spectrum compared to the other three. Hence more bands which are in common for the others have been found. The most prominent bands at the lowest and highest frequency are not highlighted since they were observed before as one of the labeled amide bands in proteins.

The comparison in fig. 4.31 indicates that some absorption bands seem to be caused by general motions in proteins. From the peak positions the two absorption features around 300 cm^{-1} and 400 cm^{-1} (highlighted with colors in fig. 4.31) seem to be common in all proteins of course with individual deviations as it is observed for all amide bands, too. Both peaks are none of the former known two far-infrared amide bands in proteins and are assigned here for the first time. If β-lactoglobulin is left out then three more bands can be assumed to be caused from similar motions. Four proteins are not enough for a statistical significant analysis but we suggest the large β-sheet content to be the origin of the different spectra and hence the different absorption bands. As the above presented structures reveal this is the most striking divergence, since the size of the proteins can be excluded as a cause and the tertiary structure is anyway different for all proteins. A possible splitting of bands is indicated by arrows in fig. 4.31.

For a direct comparison to our results, at least low frequency peaks in lysozyme from OHD-RIKES (optical heterodyne detected Raman induced Kerr effect spectroscopy) have been determined by multiple peak fits to 164 cm^{-1}, 106 cm^{-1} and 73 cm^{-1} [Gir02]. The first one is very close to our fitted peak at 158 cm^{-1} in lysozyme. Here the expected conformity of infrared and Raman results occurs due to the lack of symmetry in proteins.
One step ahead to derive finally the origin of especially these low frequency modes is the connection to protein functionality. Experiments on heme proteins showed low frequency modes which indicate a strong coupling to the reaction with ligands [Ros02, Gro02]. Photo echo experiments found a relation of low frequency, 67cm^{-1} to 306 cm^{-1} , modes to antigen-antibody coupling [Jim02]. Also the internal coupling of modes has been studied in a molecular dynamics simulation on myoglobin near a hypothetical temperature of 0 K, but was split up into normal modes due to a valid harmonic assumption at this very low temperature. One particular mode was chosen for a detailed analysis occurring at 827 cm^{-1} because this is the frequency of a functional relevant vibrational mode of the heme group excited by photodissoziation. A strong coupling to three modes at 363 cm^{-1} and 464 cm^{-1} and 1654 cm^{-1} was found [Mor00]; the first two ones are interestingly close to our frequencies present in all proteins. The other several thousand protein modes were almost unaffected. But this underlines again the collective character and the (anharmonic) coupling of protein motions in the low frequency range. Besides direct calculations of modes this seem to be another promising way to find the underlying motions of these peaks even if no exact assignment has been derived yet. Measurements on cytochrome c and cytchrome c oxidase found peaks at 530 cm^{-1} and 550 cm^{-1} were an involvement of C=O in plane vibrations was supposed [Dör08]. These peaks are close to our gained BSA and lysozyme features, but since we do not observe them for all measured proteins we cannot support this assumption for a peak assignment from our results.
At the end of this discussion concerning basic frequency assignment results, we would like to speculate what could be the origin of the absorption features by using the up to today available publications based on model compounds and normal mode analysis. The latter method will provide certainly important hints but an exact peak frequency comparison makes no sense for the reasons already given before. In small model compounds motions involving more than three atoms are usually not observed making in addition a comparison to collective protein motions difficult. But most of the work has been done on N-methylacetamide as described in chapter 4.1. More recent calculations on trans-N-methylacetamide suggest that additional peaks arise from hydrogen bond motions whereas the prominent two amide bands

are mainly caused by deformation of covalently bonded atom pairs or triples [Mir96]. A peak around 450 cm^{-1} was found to be caused mainly from C-C-N deformation and a peak around 300 cm^{-1} from C-N-C deformation motions both close to peak frequencies observed in our protein spectra. Due to the small model compounds the intermolecular motions have been calculated, leading then to additional contributions to bands below 200 cm^{-1} involving hydrogen bonds [Her01, Her03]. The results can be seen as a hint for such hydrogen bond contributions in proteins but due to the completely different hydrogen bond structure we hesitate to apply a direct comparison to these calculated frequencies. This underlines the lack of exact protein calculations (not based on much smaller model systems) which we hope to promote with the here presented analysis of our experiments.

4.5. Hydration Dependence

With our developed atmosphere cell setup we are able to change the amount of water molecules situated at the protein surface by external hydration. Due to a precise background measurement the absorption from water molecules in the gas phase is almost completely withdrawn. As measurements with the first prototype of the cell demonstrated, water vapor anyway just causes small absorption peaks in the far-infrared that can be clearly distinguished from the broad protein absorption as visualized in fig. 4.32.

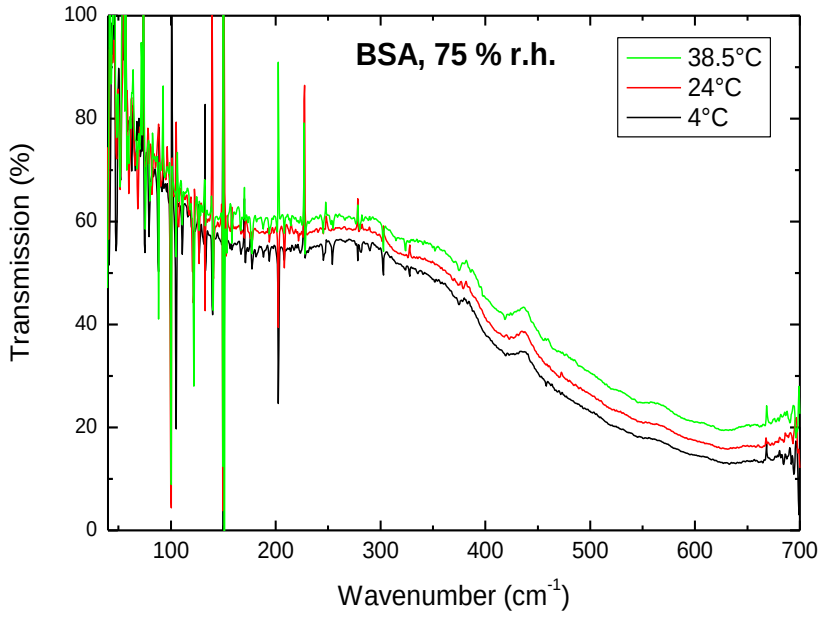

Fig. 4.32: Temperature dependence measurement of BSA at 75 % r.h. with the first prototype of the atmosphere cell to visualize the very sharp water vapor absorption peaks. These peaks are later completely removed by the background measurement due to a much higher stability of the system reached with the most improved version of the cell used throughout the presented results in this thesis.

Optimization of the complete experimental setup made it finally possible to derive spectra where these small absorption peaks are completely absent. This gives us the advantage to present here spectra without any additional data treatment, which we want to emphasize prior to the comparison of our results to THz time domain spectroscopy publications in chapter 4.5.2.

The small “noise” on top of the protein spectra is mainly the remainder of small hydration fluctuations during the about fourteen hours enduring measurement to get one single

hydration dependence. But all discussed temperature and hydration dependent effects are much larger than our error bars as described in the reproducibility chapter 4.3. Since the water vapor absorption is withdrawn by the background measurement at exactly the same hydration, the differences in the hydration dependent absorption are solely due to the influence of biological water molecules attached to the protein sample.

For all measured proteins just a small effect of hydration on the far-infrared absorption spectra has been observed. This result seems to be quite uninteresting on a first glance but yields important information about the far-infrared properties of biological water. The absorption caused by this kind of water is much less than for liquid bulk water which is the main outcome of our hydration dependent measurements. By a comparison to ice we can even connect these results to microscopic dynamical behavior and support therefore the proposed "iceberg" model of biological water which will be discussed in chapter 4.4.3.3. We will present first our experimental findings and the results using the sorption isotherm formalism, including the application of our derived formulas in chapter 2.4. Then the following discussion is split into a general part about the hydration dependence and into a part concerning biological water which is probed with our experiments.

4.5.1. Spectroscopic and Calculated Results of the Hydration Dependence

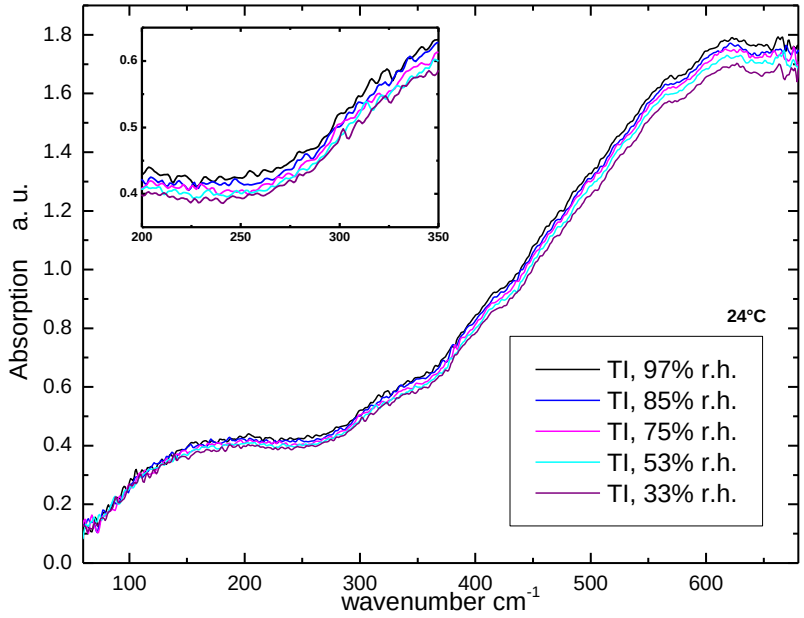

Fig. 4.33: Hydration dependent measurements on trypsin inhibitor at 24°C. The insets shows the well ordered graphs with increasing hydration even in the lower frequency range, where the differences in absorption are smaller than for the higher frequencies.

Fig. 4.33 shows the hydration dependent measurements of trypsin inhibitor. Regarding the absolute values the increase in total absorption with increasing hydration gets larger at higher frequencies. Detailed analysis reveals that the percentage increase is similar over the complete far-infrared range. No frequency dependence nor a shift of bands could be observed leading to a small but general absorption shift under the influence of hydration. As all hydration dependent graphs show, the same small general effect was found for every protein even for β-lactoglobulin as demonstrated in fig. 4.34. So despite the different absorption bands discussed in the prior section and despite some individual protein temperature dependent effects discussed later, no hints concerning size, secondary or tertiary structure can be gained from the hydration dependent measurements. Hence the influence of additional water molecules seems to be connected to a general protein property.

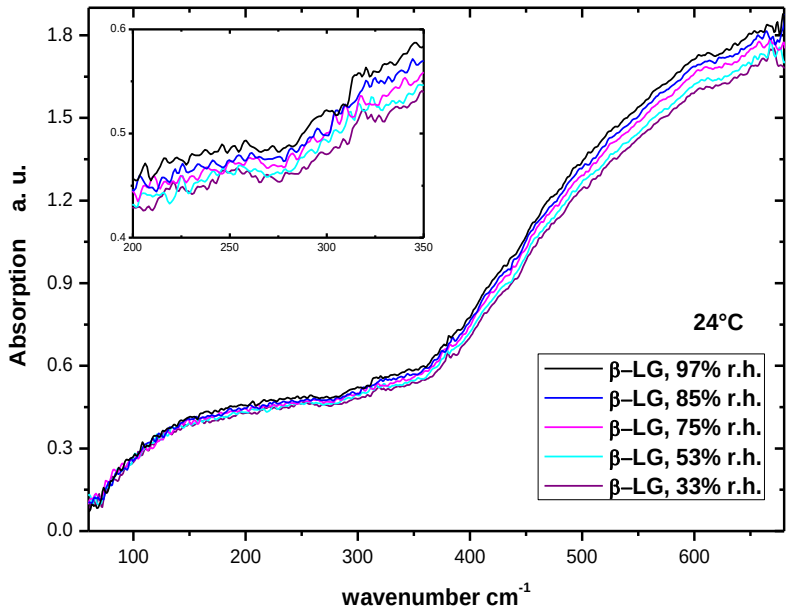

Fig. 4.34: Far-infrared spectra of β-lactoglobulin under the influence of external hydration at 24°C.

For three of our measured proteins the parameter values for applying the sorption isotherm are known and listed in table 4.1.

Protein	ν_m	**ab**	**b**
BSA	7.87	9.63	0.81
HEWL	8.15	10.5	0.82
β-LG	6.3	15.66	0.98

Table 4.1: Sorption isotherm parameters for bovine serum albumin, hen egg white lysozyme and β-lactoglobulin [Gas77, Lop00].

Via the sorption isotherm formula 2.54 the *h* value, which is gram water per gram protein, can be calculated from the relative humidity (r.h.). For BSA we get for instance from the different saturated salt solutions:

- 0.0873 *h* (33 % r.h., magnesium chloride $MgCl_2$)
- 0.1240 *h* (53 % r.h., magnesium nitrate $Mg(NO_3)_2$)
- 0.1902 *h* (75 % r.h., sodium chloride NaCl)
- 0.2434 *h* (85 % r.h., potassium chloride KCl)
- 0.3590 *h* (97 % r.h., potassium sulphate K_2SO_2)

The filling of the hydration shells can be calculated, too. We will present here the calculation of BSA to derive an important argument used in the following section. By applying our formulas derived in chapter 2.4, the filling of the first four hydration shells is given in table 4.2.

h	**rel. Hum.**	**1st shell**	**2nd shell**	**3rd shell**	**4th shell**
0.0873	**33 %**	82.95 %	27.35 %	9.03 %	2.98 %
0.1240	**53 %**	91.81 %	48.62 %	25.78 %	13.65 %
0.1902	**75 %**	96.81 %	72.55 %	54.44 %	40.80 %
0.2434	**85 %**	98.33 %	83.51 %	71.02 %	60.33 %
0.3590	**97 %**	99.79 %	96.72 %	93.86 %	90.98 %

Table 4.2: Hydration shell filling of BSA at the applied relative humidity and the corresponding *h* value as derived from the sorption isotherm formalism. The first hydration shell is already well filled at our lowest relative humidity.

For a comparison to the results gained by other techniques it is important to derive the number of attached water molecules n_w on each protein which is calculated via

$$n_w = h \cdot \frac{m_p}{m_w} \tag{4.4}$$

by using the protein molecule mass m_p and the water molecule mass m_w. The results for BSA can be found in table 4.3, for lysozyme in table 4.5 and for β-lactoglobulin in table 4.6.

Rel. Hum.	**33 %**	**53 %**	**75 %**	**85 %**	**97 %**
h	0.0873	0.1240	0.1902	0.2434	0.3590
$\mathbf{n_w}$	322	457	701	897	1324

Table 4.3: Calculated values for BSA derived from the quantitative analysis via the sorption isotherm and the sample preparation parameters.

The calculated number of attached water molecules per protein fits very well with other published results. At 0.1 *h* for BSA about 400 water molecules per protein have been reported [She83]. From X-Ray studies on crystals of human and tortoise egg-white lysozyme about 50 to 140 water molecules per protein have been found which is conform with our low hydration values [Bla83, Ful86, Hag78, Mou76]. Simulations and experimental data on myoglobin indicate about 350 to 390 water molecules per protein at sufficient hydration values and around 0.39 *h* respectively [Ste93b, Dos86]. With 18 kDa [Nie92] myoglobin is of comparable size to lysozyme and β-lactoglobulin, hence about three times smaller than the above presented BSA results. Around 0.39 *h* they have 310 and 379 water molecules attached in our calculation, being very close to the given number for myoglobin. All these examples further support the reliability of our above calculation besides the before mentioned direct proofs of the sorption isotherm formalism.

4.5.2. Discussion of the Hydration Dependence

To our best knowledge so far there exist no publications about hydration dependent far-infrared measurements on proteins and for the THz frequency region they are rare. Since we have an overlap to THz spectroscopy experiments we want to compare our results at least to the few published data of this adjacent frequency range. In addition there seems to be a general disagreement in the THz-community about the hydration dependence of protein THz absorption spectra for three reasons. Firstly, there are groups which unfortunately lack of basic knowledge to perform these measurements properly. We would like to mention here the work from the group of Globus which claimed for instance polarization effects in, per se isotropic, liquid protein solutions. For such reasons these publications are not used in the following discussion but are, of course, known to us. Secondly, the reached accuracy of the measurements in the THz region leaves still enough room for interpretation. Therefore no

convincing arguments for the intrinsic hydration dependence of protein THz absorption exist so far. Thirdly, even within experimental results published from the one particular group, as already presented in fig. 4.6, there are inconsistencies in the conclusions about the hydration dependent measurements.

For the discussion we will focus on the large hydration dependence found from Zhang et al. (fig. 4.7) and the more recent publications from Markelz [Che05, Kna07a] which are, on first sight, in clear contradiction to our small far-infrared hydration dependence shown in fig. 4.35 for lysozyme.

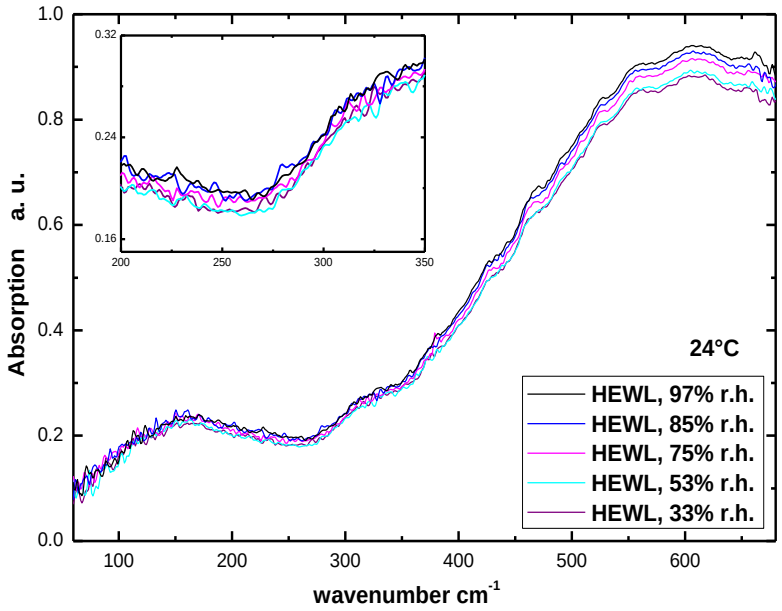

Fig. 4.35: Hydration dependent measurements on hen egg white lysozyme at 24°C.

If analyzed more in detail it becomes evident that the absorption spectra gained by these different methods cannot be directly compared since different experimental conditions were used and, most important, an additional data treatment was applied for all THz publications showing a large hydration dependence. To embed our results in the framework of other publications we have to discuss these differences and we will see that there are many indications that actually our results are in line with the experiments of the THz frequency range.

Since we are using a free standing protein film, which is completely surrounded by the external hydration, our sample hydration should be more complete than for the protein films on substrates as it is used for the THz measurements to which we compare our results. The substrate blocks the protein hydration completely from one side in this case from below. Since

we observe the smaller hydration effect with a two times more efficient hydration, this difference in the setups cannot be the origin of the deviation. This is a clear indication that our results reveal the actual, setup independent, hydration dependence. For a better comparability to the THz publications we present our hydration dependent results for several frequencies versus the hydration *h* value (fig.4.36).

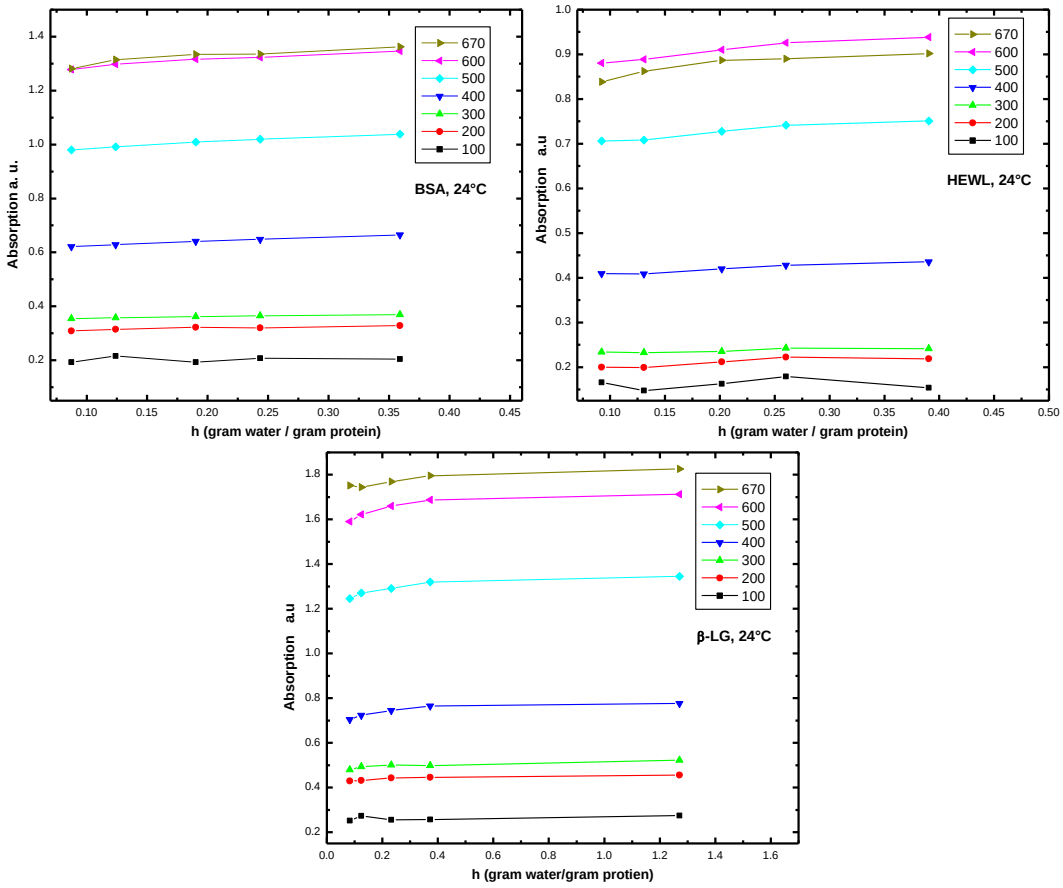

Fig. 4.36: Hydration dependence at several frequencies versus the *h* value of BSA, hen egg white lysozyme and β-lactoglobulin.

We want to focus on our lysozyme measurements in the range from 40 cm^{-1} to 690 cm^{-1} which is the extension of the THz time domain measurements in the frequency range of 7 cm^{-1} to 63 cm^{-1} by the group of Markelz [Kna06]. The corresponding hydration dependence to fig. 4.36 is pictured in fig. 4.37 for a direct comparison.

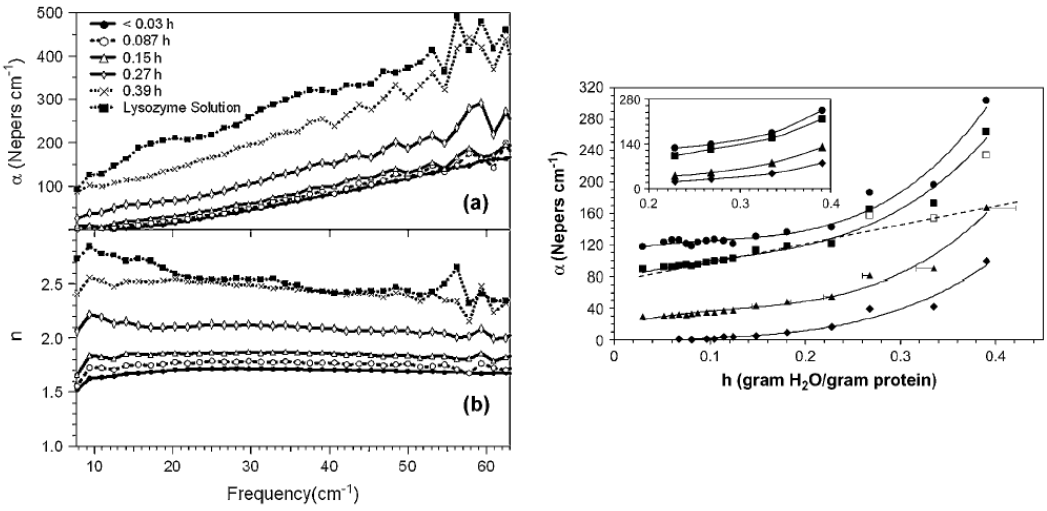

Fig. 4.37: Hydration dependence of the absorption and the refractive index derived from THz time domain spectroscopy on lysozyme at several hydration values as seen on the left. An enormous increase in hydration of more than 200 % is observed just between 0.23 *h* and 0.39 *h* as visualized by plotting the absorption against the hydration aspresented on the right. The inset shows the measurement of another film which reveals also a strong hydration dependence but clearly less than 200 %. This demonstrates the dependence on the used sample in these experiments and that one order of magnitude larger error bars (minimum of 10 %) than for our hydration dependent measurements must be assumed [Kna06].

The comparability is difficult since several differences in gaining the absorption of hydrated lysozyme are present and an overlap of the data is just given for the fact that we can confirm the more or less linear absorption increase without any absorption feature in the THz range of lysozyme. But our hydration dependence gets clearly smaller in absolute absorption values towards lower frequencies. We want to emphasize also the one order of magnitude better precision of our measurements compared to this publication as explained in the subtitle of fig. 4.37.

First of all the absorption coefficient (formula 4.5) of the Markelz group is calculated, or better to say divided, by a factor which uses the measured phase Φ and the further derived refractive index n respectively. These two values are included in the factor F and the transmission t from which the absorption coefficient is calculated. A geometry factor correction G depending on the used setup is occurring, too. The used calculation procedure [Kna06] is similar to the one used from the group of Zhang and Durbin [Zha04, Zha06].

$$\alpha(n,\phi,G) = -\frac{2}{d}\ln\left(\frac{|t|}{FG}\right) \tag{4.5}$$

In fact this is the more precise way to gain the real absorption coefficient than just measuring the transmission but the exact phase measurement is then the crucial experimental challenge. Unfortunately an error propagation is not given in all publications how deviations of the phase measurements affect the final absorption after the extensive recalculation procedure. The presented calculated refractive index shows by itself a clear transition at 0.27 *h* after being almost constant below (fig. 4.37). The same transition occurs in the recalculated absorption coefficient based on the former value. So the spectra of both constants exhibit a slight increase with increasing hydration at values smaller than 0.03 *h*, at 0.087 *h*, and 0.15 *h*. Afterwards both coefficients increase enormously. Since it is not possible with our method to measure the phase and hence *n*, the published results could have been just compared with the uncalculated or pure absorption coefficients. But the extensive non-linear recalculation procedure makes it impossible to derive the original value for a direct comparison. Since the results up to 0.15 *h* and interestingly all former publications of the same group without using this recalculation [Mar00, Hei08, Che03, Whi03] agree very well with our results, this indicates that the deviation to the absorption coefficient in these publications is very likely caused by the sudden increase of the refractive index. This leads then to a calculated increase in the absorption coefficient not present in our pure experimental data.

In the presented graphs unfortunately important hydration dependent measurements are left out by the authors giving a misleading picture of a large hydration dependent THz absorption, solely due to the selection of published measurements. Just in the text statements are found concerning a small or absent hydration effect. “The 80 % r.h. and 30 % r.h. data overlay for all samples measured (80 % r.h. shown in the figure (fig. 4.6 in this thesis, where just one of two possible hydration measurements is presented)), whereas when the cell was purged with dry nitrogen gas, the net absorbance decreased in a uniform way” [Whi03]. In summary, just complete drying of the proteins below 5 % r.h. lead to a significant hydration caused absorption difference. Since we measure well above 30 % r.h. our findings are in line with the results from the Markelz group up to this publication even if the for us important graphs were unfortunately not included in the published figures.

Another possible explanation for the different published THz hydration dependence of the absorption coefficient is that the protein film was dissolved in a buffer solution prior drying. Typically these buffers do not exhibit a THz absorption which is true, but the authors unfortunately decided from this fact not to include these molecules in their background measurement. But due to the ionic or polar character of the buffer molecules [She00] additional water is adsorbed out of the relative humidity. The absorption (and also the

refractive index) will increase much more than for the protein sorption itself, likely leading to an enhanced water collection and hence increased absorption caused by supplementary liquid (or buffer attached) water molecules. Therefore the missing buffer in the background could play an additional crucial role by such an effect even with a sole transparent property in the THz.

The last difficulty for a straight forward comparison of the results to the cited publications comes from the large temperature fluctuations of ±1°C of the used dew point generator. In a later paper [Kna07a] the temperature fluctuations are attributed as the primary source of uncertainty during a complete hydration measurement. There will be an increasing uncertainty in humidity with rising relative humidity right where the deviations to our results occur. Highly saturated atmospheres tend to condensate on every part within the sample cell which is slightly colder. Fluctuations of ±1°C can easily lead to dew or drops on windows (or substrates, if no free standing film is used) enhancing the "absorption" or even introducing additional scattering, which is also then mistakenly attributed to an increased sample absorption. In the early versions of our atmosphere cell we struggled with the same problems when changing the temperature at high humidity. This was overcome in the latest version, as used for all here presented results, due to the at this point finally possible new arrangement within the cell by using the developed free standing protein film technique. Of course our absolute humidity may be less accurate in a range within ±1 % r.h., even if the saturated salt solution method has intrinsic calibration standard, but our humidity fluctuations within the cell and the chance for an unwanted condensation will be much less due to our one order of magnitude more accurate temperature control (±0.1 K). This reduction of humidity fluctuations may be important especially in the upper hydration range where the deviations to our neighboring frequency range measurements occur.

For instance 100 % relative humidity at 20°C corresponds to just 80 % relative humidity at 24°C [Web95]. The difference with temperature fluctuations increase significant with increasing relative humidity. At 20°C a temperature fluctuation of ± 1°C causes at 10 % r.h. just about 0.6 % change, at 50 % r.h. already 3.2 % change and at 90 % r.h. about 5.7 % change in the absolute value of relative humidity. Therefore a precise temperature control is essential even if condensation problems are ruled out which might occur in addition to these deviations and which could trigger in such a case much larger differences to the actual relative humidity.

But besides all these experimental setup issues, the observed transition in the hydration dependent THz measurements around 0.27 in fig. 4.37 is supported by dielectric relaxation

experiments on adsorbed lysozyme molecules from the 1970ers [Har72]. In the GHz range a transition was found in the imaginary part of the dielectric constant around 0.3 *h* in lysozyme at 25°C. Bulk water has typically a dielectric relaxation peak at 19 GHz [Gas81]. Harvey and Hoekstra did not found a relaxation for 0.34 *h* hydrated lysozyme there and concluded that protein bound water is different from bulk water. At 0.3 *h* the authors suggested a completation of a single monolayer around lysozyme which causes the transition in the dielectric constants. One likely origin of the dielectric relaxation in this case is the rotational relaxation behavior of permanent or induced dipoles as they can be found for polar side chains of proteins or for protein bound water molecules. Very recent dielectric measurements on 14 different proteins under different hydration levels, also including lysozyme, in the frequency range of 1 MHz to 20 GHz did not confirm a special hydration transition and are therefore an additional argument against the above claimed transition in the THz region [Bon08].
Also a histogram of corresponding normal mode calculations of the same group [Kna07a] to model the THz lysozyme measurements did not reproduce such a large increase of more than 200 % just between 0.23 *h* and 0.39 *h* as seen in fig. 4.38 versus fig. 4.37, too. This might be due to the fact that this method is not sufficient to model the THz absorption properly as the authors themselves state. But interestingly the calculations fit much more to our results of a smaller, almost linear absorption increase and an absent transition at 0.27 *h*.

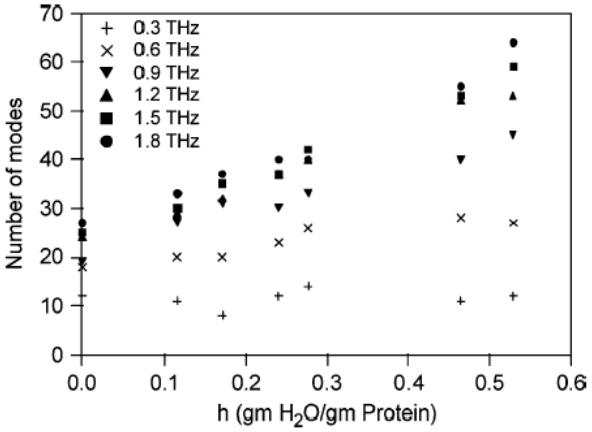

Fig. 4.38: Normal mode based histogram of the hydration dependence for hen egg white lysozyme. If identified with the absorption, the number of modes at certain frequencies do not double between 0.23 *h* and 0.39 *h*. Also the corresponding experimentally observed transition at 0.27 *h* (shown in fig. 4.27) is completely absent. The results are much more in line with our results of a small and almost linear hydration dependence concerning the protein absorption [Kna07a].

As in table 4.2 presented, the filling of the first shell is already larger than 80 % at our lowest hydration and there is also no indication of a transition from these values. The protein can be seen as to be in a well hydrated and functional state throughout the measurements where just the amount of biological water around the protein is increased. Also other THz measurements on myoglobin do not support any transition in this range [Zha04]. At the end of the theory chapter we summarized already that there are also more indications from non-spectroscopic experiments for an absent transition at 0.27 *h* in the hydration dependence of lysozyme, which the authors of the above cited publications claim to observe in the THz range. So we conclude that our experimental results clearly support a small and continuous absorption increase without any transition. This is in line with the majority of results derived by other techniques and in common with the theoretical expectations for completely functional proteins at this hydration level (fig. 2.22). Summarizing the above discussion we finally claim that our findings are valid for both, the far-infrared and the THz region, and that for these frequency ranges our measurements revealed the intrinsic hydration dependence of protein absorption.

4.5.3. Discussion of Biological Water Properties

The sorption isotherm provides even the possibility for a further quantitative analysis of our hydration dependent results. To derive biological water properties, it is important to know the number of proteins and water molecules measured within the sample spot. Since we are not able to measure exactly the thickness and the density of our fluffy free standing film, we bypassed this obstacle through calculating the number of proteins in the spot without using these unknown parameters but using ratio functions.
Theoretically the density ρ of our sample can be calculated via

$$\rho = \frac{m_{sample}}{A_{sample} \cdot t} \tag{4.6}$$

with m_{sample} as the known total mass of protein that we dissolved in water at the beginning of the preparation, A_{sample} is the area of the sample calculated from the known diameter of 9 mm and t is the unknown thickness of the sample.
The mass of protein in our sample spot is then given by

$$m_{spot} = \rho \cdot t \cdot A_{spot} \tag{4.7}$$

A_{spot} can be calculated by using the microscope spot diameter of 250 µm.
By bringing together (4.6) and (4.7), the mass of protein in the spot can be calculated without the unknown thickness t and density ρ of the sample with

$$m_{spot} = m_{sample} \cdot \frac{A_{spot}}{A_{sample}} \tag{4.8}$$

For the sample preparation of BSA and β-lactoglobulin 120 µl of a 5 mg/ml protein solution (for our here presented B concentrations) have been taken which leads to a total of 0.6 mg protein in each sample. For lysozyme half of the concentration namely 2.5 mg/ml was used. The number and hence the mass of proteins can be easily assumed to be constant in the measured spot. Even if the protein film soaks more water, the free-standing film gets heavier and lowers a little bit, but the light is still shined through the same amount of protein. Thermal expansion will be not present in these hydration dependent measurements at a defined temperature. Also a thickness increase and density changes would not alter the number of proteins within the spot. What will change significantly is the number (and therefore the mass) of water molecules.
The h value calculated for BSA increases from 8.73 % (33 % relative humidity) to 35.90 % (97 % r. h.) during the hydration measurement which is, with a protein mass of 462.9593 ng, a water mass of 40.41 ng and 166.20 ng respectively. The increase in water molecules is hence larger than a factor of four during the experiment. Since this increase arises just from a part of the total mass, the ratio of water mass based on the total mass increases from 8.03 % to 26.41 % which is still a factor of more than three.
To verify our conclusion that protein bound water absorbs much less and behaves different to liquid water in the far-infrared, we will derive several arguments for that in the following two subsections which also contain important conclusions about the general origin of the protein spectra.

4.5.3.1. Refutation of the Hypothesis of Water Dominated Far-Infrared Protein Spectra

Due to the large absorption coefficient of liquid water in the far-infrared it is often assumed that the protein spectra are governed by the absorption of water molecules around a protein. This would lead to almost similar far-infrared spectra for all proteins which is not the case as already presented in chapter 4.4. But we want to test this hypothesis because we will gain a nice comparison to liquid water absorption which we will use for a distinction to biological water below.

Since we know the mass of water within the measured spot, we are able to calculate the sole absorption from water without protein absorption. We assume, on basis of the hypothesis, liquid water behavior and theoretically concentrate the water in a layer of the length *l*. The volume of the water layer is given by $V=m/\rho$, with m as the water mass and the density ρ which is 997.048 g/l at 25°C. For the lowest hydration of BSA we have a water mass of 40.41 ng and this calculates to a volume of $40.4296 \cdot 10^{-9}$ ml. For the highest hydration with 166.20 ng water this becomes $166.6692 \cdot 10^{-9}$ ml. This volume is theoretically concentrated in a layer $l=V/A$ of a 250 µm measuring spot. The hypothetical water layer becomes hence 823.6315 nm and 3395.3837 nm for the lowest and highest hydration respectively.

With the far-infrared absorption coefficient taken from [Pal98] the following calculated water absorption spectra (fig. 4.39) are gained for all five measured hydrations. From the given far-infrared water absorption coefficient $\alpha(\nu)$ [Pal98] the molar absorption coefficient $\varepsilon(\nu)$ needed here is calculated by

$$\varepsilon(\nu) = \frac{\alpha(\nu)}{c_{molar}} \tag{4.9}$$

with c_{molar} as the molar concentration, which is 55.5084 mol/l based on a mol mass of 18.01528 g for water.

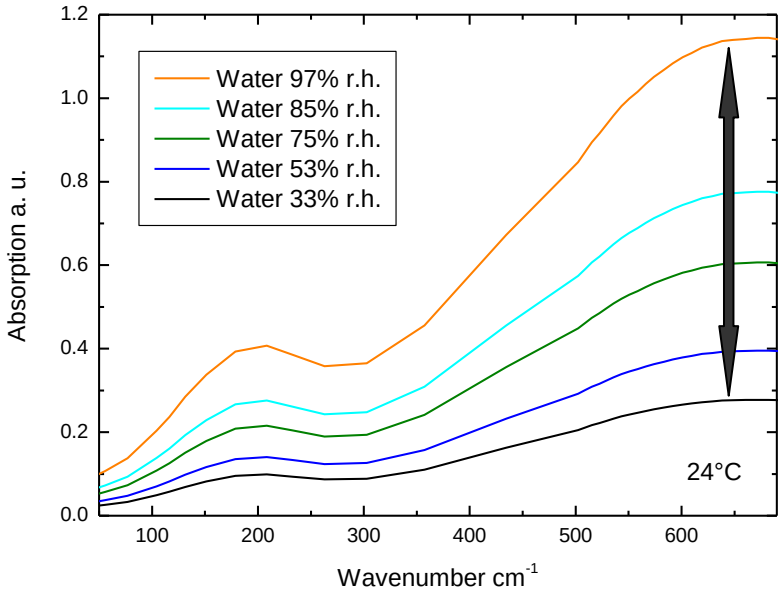

Fig. 4.39: Hypothetical pure liquid water absorption of the water molecules attached to BSA at the measured five different relative humidity values. For a calculation the liquid water absorption coefficient from [Pal98] has been used. The absorption increases more than a factor of four during the hydration.

It is obvious from fig. 4.39 that the protein spectrum cannot be governed by water absorption which is at least comparable to the absorption coefficient of liquid water in the far-infrared. We will give two arguments – an absolute one (highlighted with a grey arrow in fig. 4.40 to fig. 4.42) and a striking relative one (highlighted with a black arrow in fig. 4.39, fig. 4.40 and fig. 4.42) – with an argumentation first from the high and then from the low hydration point of view.

The absorption of the thin cumulated water layer at the highest hydration will be already stronger than that of the measured total spectrum in the lower half of the measured frequency range and hence the absorption of the protein bound water molecules must be less than liquid water (fig.4.40). The observed absorption change during the hydration from low to high water content should be in that case dramatically. Hence the water molecules cannot be responsible for a large absorption percentage of the total spectrum if similar absorption behavior is assumed as for liquid water.

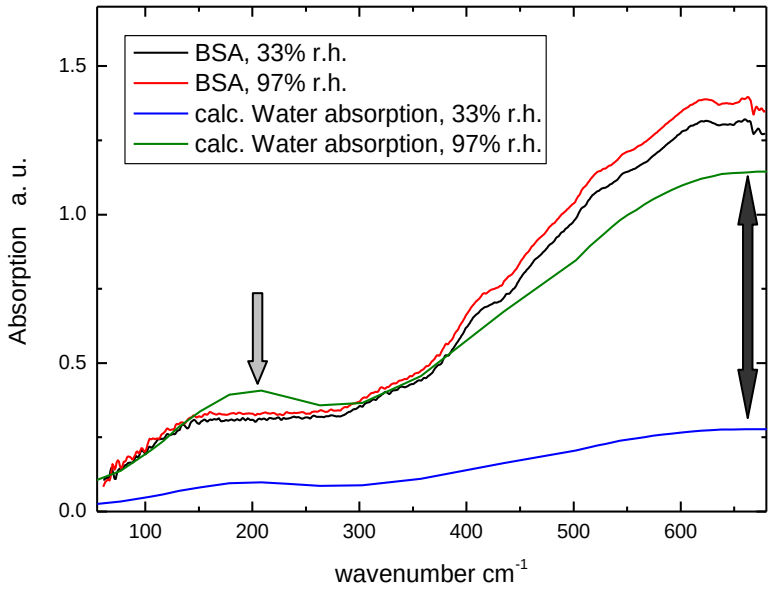

Fig. 4.40: Comparison of BSA spectra and the calculated liquid water absorption at the lowest and highest hydration. At the highest hydration the liquid water absorption exceeds already the protein spectra in the low frequency region. The relative change in absorption differs at least one order of magnitude. Both are strong arguments for a much lesser far-infrared absorption of biological water compared to liquid water.

Even if a small ratio of water to protein absorption is assumed, to e.g. 20 % of the total absorption at the lowest hydration, it can be easily seen from the graph that the increase in absorption must be still significantly stronger during the hydration than the small absorption increase that is actually experimentally observed. So the outreach of the absolute values due to the water absorption and the not observed large, relative increase in absorption of more than a factor of three, as seen in fig. 4.40, are strong arguments that protein bound water has a significant less absorption coefficient than liquid water and does not govern at all the measured protein spectrum.

Frequency	**100 cm^{-1}**	**200 cm^{-1}**	**300 cm^{-1}**	**400 cm^{-1}**	**500 cm^{-1}**	**600 cm^{-1}**
Absorption Increase	6.16 %	6.39 %	4.60 %	6.95 %	6.04 %	5.41 %

Table 4.4: Absorption increase of BSA in per cent at several frequencies from the lowest to the highest hydration.

Another argument is the calculated much too low water content of the protein sample with an assumed liquid water absorption behavior. To cause an overall absorption increase of about 6 % (table 4.4), the mass water content is calculated to about 1.2 % at the lowest hydration. This rough estimation is done via $w*4 + p = 1.06$, with w as the mass water content and $p = 1 - w$ as the protein content. The actual value is 8.03 % at the lowest hydration which is more than a factor of six apart from 1.2 % and the first one is a well proven value. At higher hydration values this becomes much more than an order of magnitude difference. So this discrepancy is far beyond our error bars and shows the striking different behavior of protein bound water to liquid water.

The measurements and calculations for lysozyme and β-lactoglobulin support this conclusion as well. The derived parameters for lysozyme are presented in table 4.5.

Rel. Hum.	**33 %**	**53 %**	**75 %**	**85 %**	**97 %**
h	0.0923	0.1309	0.2018	0.2602	0.3905
$\mathbf{n_w}$	71	103	159	206	310
Mass / ng	21.3679	30.3788	46.7172	60.2194	90.3905
Vol / 10^{-9} ml	21.4312	30.3987	46.8555	60.3977	90.6581
Length / µm	0.4366	0.6189	0.9545	1.2304	1.8469

Table 4.5: Calculated values for lysozyme at different hydration values.

For lysozyme the sole absolute water absorption is not above the hydrated protein absorption like it is seen for BSA and as we will see it for β-Lactoglobulin. Therefore in fig. 4.41 we just show the spectrum at the highest hydration to emphasize this point. Nevertheless the water content of lysozyme increases also more than a factor of three during the hydration, which is due to a very similar sorption isotherm for BSA and lysozyme (table 4.1). So the same argumentation as above, for the not observed but expected large relative absorption change, holds with the same significance for lysozyme.

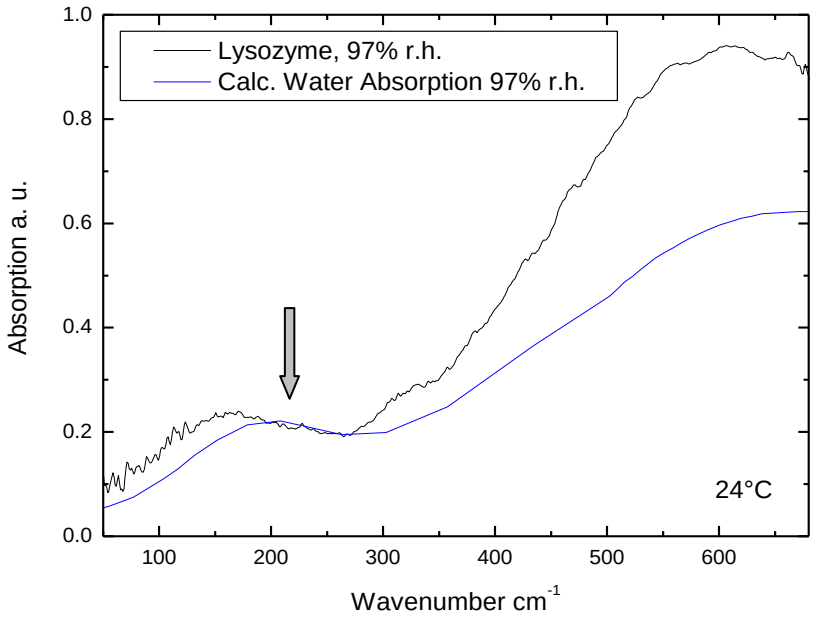

Fig. 4.41: Lysozyme absorption spectrum compared to liquid water absorption of the sample water molecules at the highest hydration.

The calculated values for β-lactoglobulin are presented in table 4.6. Due to the strong water sorption the parameters of the sorption isotherm differ a lot from these of BSA and lysozyme. The isotherm curve, as can be seen from fig. 2.41, diverges already to very large values at the highest hydration of 97 % relative humidity. The mass of attached water molecules exceeds the mass of β-lactoglobulin. Therefore a large increase of the water content between the two highest hydrations is present in the calculation.

Rel. Hum.	**33 %**	**53 %**	**75 %**	**85 %**	**97 %**
h	0.0823	0.1239	0.2325	0.3726	1.2712
n_w	84	126	236	379	1292
Mass / ng	38.1154	57.3653	107.6334	172.4847	588.4999
Vol / 10^{-9} ml	38.2283	57.5351	107.9520	172.9954	590.2424
Length / µm	0.7788	1.1721	2.1992	3.5243	12.0244

Table 4.6: Parameters derived for β-lactoglobulin.

As the β-lactoglobulin spectrum in fig. 4.42 shows, the discrepancy between expected liquid water absorption and the measured absorption change is enormous. Since a large contribution

comes from the divergence of the β-lactoglobulin sorption isotherm at highest hydration, the calculated factors might be too large. But even with a linear and not diverging extension, the effect would be at least as pronounced as for the other proteins. That we observe the largest absorption increase for β-lactoglobulin is perfectly in line with the larger water activity compared to the other two proteins as seen in table 4.1 which proofs again the Sorption Isotherm parameters and the derived calculation.

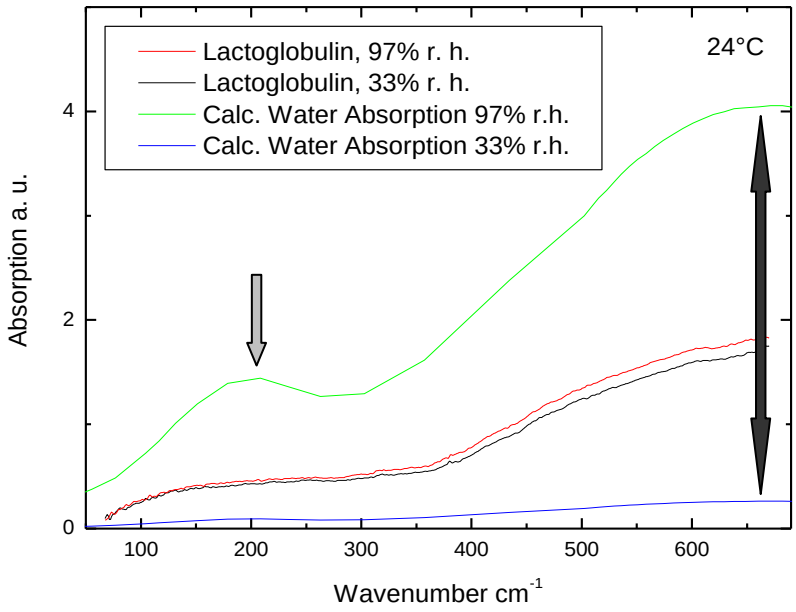

Fig. 4.42: The spectra of β-lactoglobulin at the lowest and highest external hydration compared to the liquid water absorption of the corresponding water molecules. The latter shows an increase of more than 1400 % whereas the hydrated protein absorption increases about 9 %.

The calculated absolute liquid water absorption exceeds the protein absorption more than three times. So this is our most intriguing absolute value argument for a much lesser absorption coefficient of protein bound water. The expected increase calculated from the lowest hydration is more than a factor of 14 for the complete range. This underlines the experimentally observed one order of magnitude smaller effect than it is expected for the increase in water content leading to this knowledgeable conclusion.

4.5.3.2. Estimation of the Biological Water Absorption Coefficient

We would like to give at least an estimation of the far-infrared absorption coefficient of biological water. Therefore we assume that biological water, as a part of the protein in a more general sense, absorbs just as much per mass as the protein itself without the analytical withdrawn water. This cuts down the strong absorption coefficient from liquid water which was hypothetically assumed in the prior section and can be seen as a lower edge estimation for which the observed effect still significantly holds.

The water content of the overall mass in BSA is at the lowest hydration 8.03 %. The number of water molecules is increased by a factor of more than four which becomes rescaled to the total mass a factor of more than three. Since the number of protein molecules stays constant in the spot and if an absorption change due to hydration is neglected for the protein absorption itself, 92 % of the overall absorption will be not altered. So we neither assume that the absorption ability is lowered due to a damping of motions nor we assume that with further hydration the motions and hence the absorption ability of the protein alone increases. But the water mass content of 8 % increases mass weighted by a factor of three which leads to 118 % absorption of the complete system at the highest hydration based on the lowest hydration absorption.

The experimentally observed average absorption increase of about 6 % from the lowest to the highest hydration (table 4.4) is still much smaller than the expected increase of 18 %. So even with the assumption of a comparable absorption from protein bound water to the protein itself, and hence a much smaller absorption coefficient than for liquid water, the observed change is still much smaller than the expected value. Hence a significant lower absorption of protein bound water compared to liquid water and still less than the protein absorption itself has been demonstrated.

The measured general absorption increase of lysozyme is about 6.5 % which compares to an expected 27 % increase using the same assumption as above, both based on the lowest hydration. Also for this protein the effect is still significant even with this low absorption estimation. As expected, the hydration dependence in β-lactoglobulin is more pronounced as in the two other proteins with an average absorption increase of 9 %. Based on the lowest hydration and with an assumed protein like absorption of the biological water molecules, still an increase of 116 % is expected. This shows that even for this low absorption behavior of biological water molecules the expected effect is still more than an order of magnitude above the measured contribution.

We can summarize here that the far-infrared absorption coefficient of biological water is about an order of magnitude smaller than for liquid water. From a purely mass weighted approach we see that the effect is still present. This leads to the estimation that the absorption coefficient of biological water is in the order of mass weighted water free protein absorption and maybe even below.

4.5.3.3. The "Iceberg" Model for Biological Water as a Visualization for Dynamical Behavior and Density Considerations

As presented in chapter 2.2.3 there is another type of water which has much lower far-infrared absorption than liquid water, which is ice. This is also for instance supported from recent spectra of cold buffer solution in an ice like state [Che07]. This opens up a possibility for a connection of our experimental data to microscopic properties concerning biological water which are suggested by several theoretical and experimental approaches.
As already discussed in chapter 2.2.3 there exists no sharp transition from biological water to bulk water around a protein and hence various quantities probed by different techniques will conclude a different outspread of the hydration shell aggravating a discussion about biological water. In the mid-infrared the interpretation of the hydration dependence is complicated. The difference upon hydration were on one hand attributed to conformational changes and on the other hand used to gain information about biological water properties [Pre93]. If no significant conformational change above a sufficient hydration for functionality is assumed, then the mid-infrared results indicate a stronger hydrogen bonding for biological water than for bulk water [Dos86]. Estimations derived an order of magnitude decreased mobility of biological water compared to bulk water [Pet92, Dem97]. In general mid-infrared results indicate that biological water is assumed to be in an intermediate state between ice and bulk water [Dos86] which our far-infrared absorption measurements support.
On the other frequency side of our measured spectra, the results are not so uniform. Recent MD simulations suggest that THz absorption of the hydration shell is distance dependent [Ebb08] leading hence to an expectation of subtly differentiated findings. Bizzarri et al. [Biz02] proposed from simulations a rather rigid biological water structure around proteins. The same is assumed in general for biomolecules [Ros05] and for carbohydrates [Lee05]; the latter are used in a small frequency range to derive indirectly biological water properties similar to the situation around proteins. The simulation suggests that the tetrahedral

arrangement of water is disturbed, more stable hydrogen bonds are formed and the breaking of hydrogen bonds is retarded. In the small range from 2.1 THz to 2.8 THz (70 cm^{-1} to 93 cm^{-1}) it was found that hydration water within about 6 Å to 7 Å distance, meaning about two hydration layers around disaccharides, have a 2 % to 4 % enhanced absorption compared to bulk water. For monosaccharides the same was observed but just up to about 3 to 4 Å [Hey08]. Interestingly the corresponding simulation revealed, in contradiction to the THz measurement of the same group, a decreased absorption of the hydration shell water molecules. To explain this inconsistency the neglection of polarization effects in the used force field models was addressed. As one argument the suppressed hydrogen bond network absorption band between 150 cm^{-1} and 200 cm^{-1} was given as it is found in Car-Parinello simulations including this property [Sha05, Sha07a, Sha07b]. But the simulation is in line with our findings of a decreased absorption of biological water in this frequency range and in line with theoretical considerations. Measurements of the same group on a five helix bundle compound [Ebb07, Ebb08] showed decreased and enhanced THz absorption of biological water depending on the pH value. From a theoretical point of view the enhanced THz absorption of biological water was assumed to arise from correlated in-phase motions of the stronger bound water molecules compared to the completely dephased liquid water molecules. The more efficient hydrogen bond network was meant to cause correlated motions and hence leading actually to an enhanced absorption [Heu06, Lei06]. In general this is an attractive way to look at these motions, but in which frequency range this should happen we want to address as an open question before including such an argument for far-infrared or THz protein absorption. We agree that collective motions of proteins are expected to occur in the THz range which are typically spread out over significant areas involving many atoms. But the frequency range of this possible effect, concerning now biological water, needs to be further clarified. Especially the involved force constants of this weakly bound network, the timescales for the presence of hydrogen bonds and the outspread of correlation would be the important, up to now undiscussed, factors to get an estimation of the relevant frequency range. From our point of view, the existence of internal collective protein motions in the THz range cannot be straight forward transferred to smaller, hydrogen bonded, biological water molecule networks. Therefore, from the short range order of biological water networks, and even within ice formations, we state from the known experimental spectroscopic data, that it is most likely to expect a decreased THz absorption for the biological water network compared to liquid water.

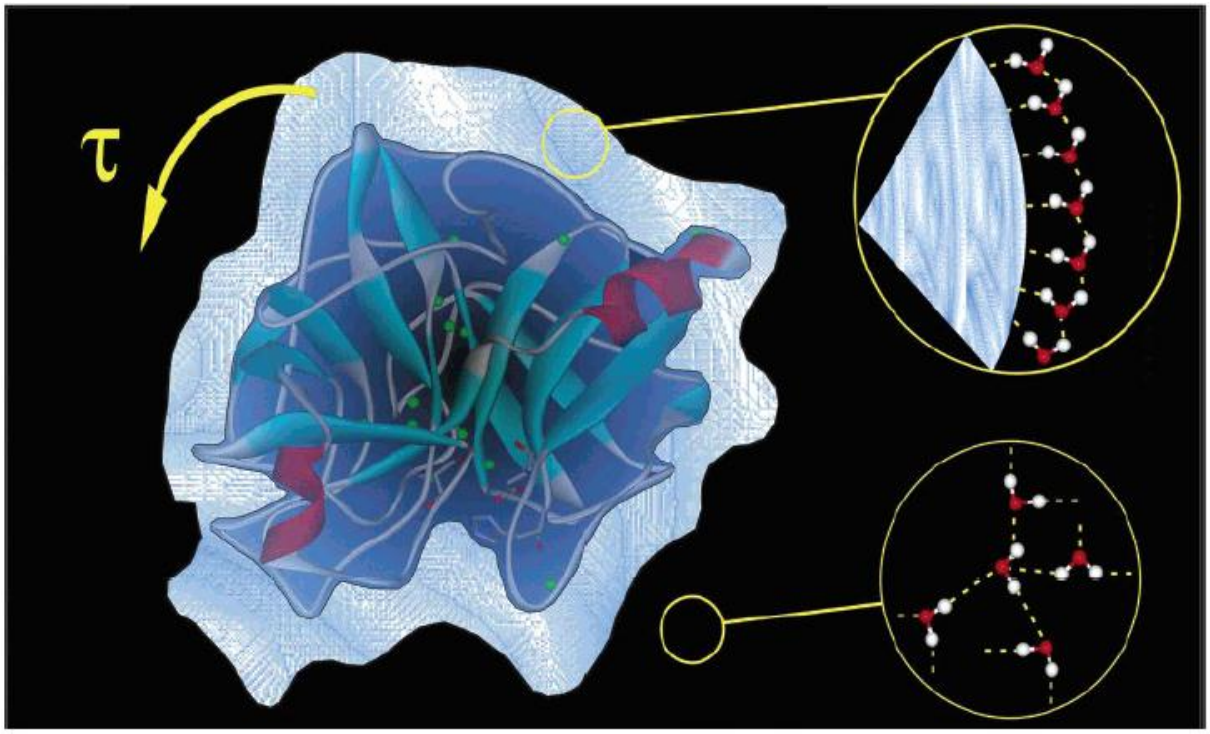

Fig. 4.43: Water molecules at the protein surface are bound via hydrogen bonds and are able to follow the orientational relaxation in the "iceberg" model. Outside the typical tetrahedral arrangement is formed by the water molecules [Pal04].

After the discussion of biological water absorption and the comparison to the absorption of ice, there are many publications regarding the dynamical behavior of biological water molecules, connecting the discussion from macroscopic observations, which concern a general property like absorption, down to a microscopic level. One model to describe the properties of biological water is the so called iceberg model, pictured in fig. 4.43, which will be supported by our hydration dependent measurements. The water molecules at the protein surface are bound via hydrogen bonds and are able to follow the orientational relaxation τ of the protein as indicated. Further outside the liquid water molecules adapt their typical tetrahedral arrangement. The iceberg model for biological water, whose roots can be followed many decades backwards [Hal04], visualizes two more properties of the following discussion. Firstly the retarded dynamics of biological water molecules around biomolecules and secondly differences in density compared to bulk water.

The dynamics of biological water molecules can be analyzed by focusing on different parameters. Dynamical behavior is related to the orientation of water molecules, since the rearrangement of the local tetrahedral structure is one of the important motions within liquid water. The water molecules around polar and charged sites are more structured and exhibit a

long range order compared to the non-polar methyl sites [Kov97]. We want to mention already here that this cited MD Simulation on α-helical polypeptide SP-C in water is in contradiction to more recent simulations claiming a relation rather to the surface topology than to the chemical type of surface groups as discussed in the following section. But besides this further distinction more ordered water molecules, and hence less dynamical biological water molecules, are indicated by MD simulations concerning these order parameters.
Other parameters to look at are the residence or correlation times of water molecules at possible hydrogen bond binding sites. As found with NMR in a titration method [Ful86], lysozyme perturbs water motions within 2.6 water layers. Several correlation times have been found for different types of water molecules. Bound water molecules to polar groups of the protein showed 10^{-9} s and other water molecules in the vicinity of the protein revealed 10^{-11} s correlation times – an order of magnitude longer than for bulk water with 10^{-12} s. Again a differentiation by surface groups was observed as it will be further discussed below.
Combined studies on BPTI with NMR and X-Ray [Ott91] concluded on the other hand that residence times of X-Ray observable bound water molecules on the surface are not longer than residence times of other ones, in contradiction to the NMR and two-dimensional Nuclear Overhauser Enhancement Spectroscopy findings which indicated different dynamical behavior [Ott89]. For the above mentioned studies on sugars in solution, the retardation of the hydrogen bond water dynamics predicted by MD simulation were confirmed with depolarized Rayleigh Scattering [Pao07]. It was found in this MD simulation that the retardation spreads out to 6 Å which is about two hydration layers around these sugars. The hydrogen bond lifetime gained by fitting the hydrogen bond correlation function was 1.5 ps for bulk water and around 2 ps for water molecules in the first hydration layer of different sugars as a second MD study concluded [Hey08]. To come back to simulations directly on biomolecules, MD simulations on plastocyanin at different hydrations, focused on residence times for biological water, lead to the conclusion that the dynamics of the hydrogen network of biological water is slowed [Biz95]. This would be also in line with the minimum amount of water molecules for an activation of the functionality of many proteins.
Viscosity studies are another approach to derive indirectly the dynamical behavior of biological water molecules. Usually the interaction of protein and solvent motions are studied. Simulations and experiments are able to relate the viscosity to certain motions and their corresponding frequencies [Vit00]. Two motions on different time scales seem to be involved in the mediation of motions from the solvent to the protein. One of about 10^{13} s^{-1}, which is not directly related to the viscosity of the solvent but performs rapid collisions with the protein,

and a much slower one [Dos83, Set92] depending mainly on the viscosity of the solvent and influences transitions among different protein substates [Fra02b].

But most MD simulation studies directly probe the dynamical behavior of biological water molecules. In a MD simulation on ubiquitin [Abs96] it was found that translational motions within the first hydration shell are retarded by a factor of 3.3 compared to bulk water which is in line with the experimentally found retardation by NMR of a factor of four in BPTI [Ott91]. Other molecular dynamics simulations calculated a sublinear development of the mean square displacement of biological water molecules in time, of cause depending on the distance from the protein surface [Biz96]. This was confirmed by neutron scattering together with the assumption of anomalous diffusion as an origin of this sublinear increase, since a retardation of the water molecules was observed for the low frequency modes of water molecules interacting with the protein [Set96]. Biological water molecules reach the diffusive state very fast as well, but there is a slowing down of the translational and rotational motion, calculated from the translational velocity and rotational autocorrelation function [Roc98] indicating a different nature of this processes [Biz96]. The Fourier transform giving the frequency distribution of the translational process shows two neutron scattering peaks around 10 THz (333.5 cm^{-1}) and 50 THz (1667 cm^{-1}) close to the ones in liquid water which exhibits a much larger peak at 10 THz (333.5 cm^{-1}) and a shoulder at 40 THz (1334 cm^{-1}) [Bal96]. So a protein has a retardational influence on the translational and rotational motions of water molecules close to the protein surface. An additional disorder caused by the protein's structure and dynamics is suggested. The protein transitions through many conformational substates, also on the here relevant picosecond timescale, lead to a steady changing local environment for biological water molecules and hence a distortion of the diffusive water process for water molecules in the vicinity of the protein surface. Still at distances more than one hydration layer apart, effects of the protein's presence have been found in this simulation [Roc98]. For water molecules within 4 Å and 6 Å a fast relaxation process of about 0.4 ps and a slower one of 13 ps and 40 ps respectively was found. The fast relaxation is suggested to arise from cage formations and the vibrations therein. Relation to supercooled water effects, as Gallo et al. described the biological water, can be made [Gal96].

In general all these results at least agree in indicating a decreased mobility of biological water molecules in respect to the water molecules further away which supports our experimental results and their interpretation. Controversially discussed are occasional findings in simulations showing in addition an enhanced mobility at large distances (about 10 Å to 15 Å) around biomolecules and sugars, for instance from Levitt et al. around bovine pancreatic

trypsin inhibitor [Lev88] and as discussed in the article from Teeter with references therein [Tee91]. This would fit to the above presented observation from the group of Havenith who claim firstly, a very large outspread of the hydration shell, and secondly, an enhanced absorption of biological water compared to bulk water. As a possible explanation a mismatch of the biological water structure with the pure bulk water at some kind of interface at these distances was proposed. But no direct experimental indication for the existence of some kind of interface has been found yet. Such an interface could also likely be introduced as an artifact in simulations caused by boundary conditions or the used long range effect treatment [Tee91].
The second microscopic property the iceberg model describes are differences in density compared to the surrounding bulk water. Already Richards [Ric77] assumed density variations, namely a lower density close to hydrophobic groups on the surface and, around hydrophilic ones, a larger density. Since polar and non-polar groups are usually split close to equal parts on a protein surface (see chapter 2) the effect is of local nature and presumably over all very small.
Stronger bound water molecules would be more “ice-like” and hence this would usually indicate a decreased density of biological water around a protein compared to the density of liquid water [Fra76]. But as stated by Perkins [Per86] after a comparison of several own and external experimental data, a (protein) bound water molecule occupies about 20 % less of the volume a liquid water molecule does, resulting in a larger density of these kind of water molecules compared to bulk water. Perkins agreed that in ordinary ice the density is lower than that of liquid water and of course much lower than for protein bound water. But if an average density of nine different forms of ice is taken, of which some are just stable under high pressure as we would like to add, the density of blended ice is again higher than for bulk water and interestingly close to the value of protein bound water. So higher density ice-like formations are in general possible and anyway a locally different arrangement around the protein surface must be assumed. Perkins summarized that if a water molecule is bound to the protein surface via hydrogen bonds its volume decreases due to “electrostriction”. This leads, by over all summation, to a “reduced total protein volume” [Per86] and this can be derived to protein bound water of higher density compared bulk water.
Later a MD simulation on trypsin inhibitor [Ger95] found that apolar protein surface atoms and their corresponding water atoms are less tightly packed, compared to atoms in the core and compared to bulk water. Charged atoms of proteins, including their corresponding water molecules, are on the opposite tighter packed at the surface, leading to a large variety of packing density on a protein surface underlining the statement of the publications before.

Hints from the experimental side came from the analysis of X-ray data concerning the protein-water interface of 22 different protein crystal structures [Ger96]. Water molecules at the protein surface are characterized by forming mainly one or two hydrogen bonds with polar groups of the protein. But also van der Waals contacts to non-polar protein groups can be found at the protein surface which is also of significance due to the presence of polar and non-polar groups in an equal manner. The formation of hydrogen bonds with other water molecules is found for most of the above water molecules, too. An average volume increase of a protein upon hydration is hypothetically derived to about 7 % per protein atom at the surface. This effective volume increase is caused by less efficient packing and voids occurring at the interface where the protein disturbs the typical tetrahedral water network. Also a correlation to the surface structure was found named by the authors "structural hydrophobicity". No indication of a correlation to the chemical structure has been seen. Especially in grooves within the protein surface the hindrance of the water network seem to be the origin of this volume increase. Concerning biological water, molecules close to the protein surface exhibit, in line with the discussion of the prior publication, a larger density than bulk water in this study. The volume occupancy is still larger than the van der Waals volume of a single water molecule but clearly smaller than that of bulk water. Here a volume decrease up to 20 % was found based on the bulk water volume. This compensates in general the above mentioned volume increase of the protein atoms at the surface. In general at the protein – water interface the very dense protein packing and the open hydrogen bond water network structure are in competition to each other. This leads to a protein surface structure influenced by the local arrangement where on average the protein is less packed and the biological water molecules are denser packed at the interface.

One of the earliest simulations directly predicting an increased density of biological water compared to bulk water around proteins was done on bovine pancreatic trypsin inhibitor, which showed an increased density of water molecules at a 2.5 Å distance from polar groups on the surface and a slightly decreased density within 3 - 4.5 Å distance [Lev88]. Experimentally this was confirmed by neutron scattering and X-ray experiments on lysozyme and two other proteins [Sve98]. An about 10 % denser hydration shell compared to liquid water was found and the authors proposed this to be a general property around proteins, which is supported qualitatively and quantitatively within the same percental range from other authors [Ste01, Nan00, Biz02]. Several publications from the group of Smith focused on biological water around biomolecules. An increased density enhancement of 15 % compared to bulk water has been found for the hydration shell water molecules within 3 Å distance from

a lysozyme surface in a simulation. The major effect was attributed to be of topological nature and will be discussed below. One third, namely 5 % originate from the perturbation of the water structure from the bulk. This splits up into two almost equal effects. One is due to the increase in the coordination number and the other from a contraction of the O···O distance by 0.8 % in average which corresponds to a 2.4 % volume reduction [Mer02, Smi04]. Despite earlier papers just an over all 5 % enhanced density was presented in a more recent publication based again on a MD simulation of lysozyme [Mer05]. Within the first hydration layer, where increased density is observed, the number of formed hydrogen bonds per water molecule is also increased which hence seems to be correlated with the increased density.
In first order, for spectroscopic observations such an increased biological water density would lead to a purely general increase in the absolute value of absorption without a significant effect on the absorption band shape. On the other hand our experiments show about one order of magnitude decreased FIR absorption due to the altered structure of biological water compared to bulk water. But it is very likely, as the major outcome of the above discussion, that the observed decrease is just partially compensated by a much smaller enhanced absorption effect, too small to be quantified from our experiments, which is caused by the above discussed larger density of biological water molecules close to the protein surface.
We would like to summarize the above discussion about the iceberg model. Due to the anomaly of water an analogy to the iceberg model for the increased density of biological water at the protein surface is not completely given. This is just valid if all kinds of ice formations are averaged which is again related to an increased density and hence be described by the model. Since this effect is proposed to be small, due to a summation of locally different higher as well as lesser densities of biological water, the over all picture is anyway not strongly affected. Besides this imprecision the iceberg model for biological water fits very well with the spectroscopic properties of both, the far-infrared and the THz frequency range, as well as for the restricted dynamical behavior. Therefore the iceberg model is at the moment the best way to look at microscopic properties of biological water molecules. That even its extension to macroscopic properties of biological water holds, we demonstrated above within the discussion part of our spectroscopic results.

4.5.3.4. Biological Water Properties Found by Simulations and Their Consequences for FIR Absorption

There are two more molecular properties, namely topological and dipole moment related arguments, which we want to discuss here in the context of our measured hydration dependent far-infrared absorption. Several suggestions derived from simulations can be supported by our findings trying to make even another step forward into the direction of joining simulation and experiments. This section provides also an outlook what properties might be on the agenda for further research in this field as a closure for the discussion of our hydration dependent results. The dynamical properties of biological water molecules and density considerations, which are both related to the iceberg model, have been discussed in the previous section. Thereby we were able to get directly a connection of our experimental results to these microscopic properties of biological water molecules. Several times it was already mentioned in this section if in those studies a correlation to surface groups was found or denied. At the moment there are strong indications that the surface topography of a protein affects the surrounding biological water molecules more than the chemical composition. Especially experimental observations of grooves at the protein surface dragged the attention to a possible topological influence. Blake et al. were one of the first realizing that water molecules close to a groove or a concave surface are more localized, namely better resolved in X-ray experiments than water molecules close to other surface structures [Bla83]. Harpaz, Gerstein and Cothia found for proteins and their biological water molecules no indication for a correlation of the individual atom volume to the chemical composition, but to the surface structure which they named "structural hydrophobicity" [Har94, Ger96]. Grooves at the protein surface in particularly disturb the water network which seem to alter the occupied volume and hence the density of biological water molecules. In general at the protein – water interface the very dense protein packing and the open hydrogen bond water network structure are in competition to each other. This leads to a "protein surface structure influenced" local arrangement where on average the protein is less packed and the water molecules are denser packed at the interface.
Simulations from the group of Smith provided further insight into a possible microscopic arrangement [Mer02, Smi04, Mer05]. Mainly in depressions of the protein surface the water molecules align along the electrostatic field generated from the protein and form places with higher density of water molecules. In these depressions the randomizing and disorientating effect from neighboring water molecules was found to be less present allowing a more parallel alignment which hence causes a denser packing of these kind of water molecules.

This geometrical effect is responsible for two third of the increase in density within the 3 Å hydration shell around lysozyme in this simulation [Mer02]. In a following publication the size of this effect was reduced, but still it was stated that this topographic effect seems to be the most important influence on the increased density of biological water molecules [Mer05]. If this increased density of biological water molecules is either caused by the chemical surface composition or more likely by topological reasons will be an interesting "follow up" question. This open question might be able to be answered also from the experimental side by hydration dependent far-infrared measurements after this small absorption increasing effect can be clearly separated from the one order of magnitude larger absorption reducing effect, to give an outlook to prospective applications of the here developed technique.

The last topic of the hydration dependent measurement discussion concerns a possible molecular origin of the observed decreased absorption of biological water. This question is far from being answered but we would like to present at least our favorite idea as an open end to this section. This idea will be anyway picked up again in the following temperature dependent measurement discussion where possible origins of the much larger temperature caused absorption changes are discussed.

Infrared light couples to the dipole moment of molecules which consists of two contributions which are permanent dipole moment and induced dipole moment. The latter is closely entangled with the polarizability of molecules. Changes in dipole moment would lead mainly to a frequency independent absorption variation as it is observed for our hydration dependence and for the below discussed major temperature dependence effect. There are several indications that the decreased far-infrared absorption of biological water compared to bulk water is caused by a reduced dipole moment of the first type. Gascoyne and Pethig [Gas81, Pet92] presented early calculations of the effective dipole moment of biological water with the result of 0.79 D, compared to 1.84 D in the gaseous state and 2.5 D in the liquid state of water. This is explained by rotational hindrance and decreased correlated reorientations with the surrounding molecules, maybe caused by additional hydrogen bonds compared to liquid water. Several other publications support this hypothesis in a general way [Jol95, Vil00, Des04a, Des04b]. Even more interesting are the statements of a (already above discussed) recent publication about THz measurements on myoglobin concerning this issue, too. Biological water is assumed to have a decreased absorption compared to liquid water due to protein-induced hindrance of rotational and vibrational degrees of freedom within the biological water network .Therefore the increased absorption upon hydration is explained with an increased protein's polarizability [Zha06]. The latter is different to our observations

since we found just a slight absorption increase without using a recalculation procedure as already discussed above. But the authors see “substantial evidence” from the two publications [Gas81, Pet92] that biological water has a reduced THz absorption, naming for the first time a possible relevant frequency range. From our hydration dependent measurements we derived about an order of magnitude lesser far-infrared and THz absorption for biological water compared to bulk water. Our measurements clearly support this idea that at least some contribution is caused by the reduced effective dipole moment. But if the reduction to about one third of the bulk water value is right, we have to claim that an additional effect will be necessary for a complete explanation of our experimental observations. One possibility, which we can derive in our below presented MERCEP model, up to now just for the temperature caused effects, might be the contribution of the induced dipole moment and corresponding correlation effects. So far this, from our point of view indispensable, refinement was not included in the above publications by a treatment of the sole effective dipole moment. An inclusion may lead in consequence finally to an additional further reduction of the far-infrared and THz absorption of biological water compared to bulk water. In this case, which we can for biological water just utter as a hypothesis at the moment, dipole moment based correlation effects would then provide the theoretical basis for our derived biological water properties and also for the microscopic origin of our observed absorption behavior, too.

4.6. Temperature Dependent Far-Infrared Measurements on Proteins

After the discussion of the hydration dependence results affecting mainly water molecules around proteins, we would like to discuss our temperature dependent measurements. Our observations will be presented along three experimental sections, describing two frequency dependent shifts and one major frequency independent shift within the protein absorption spectra. In this temperature dependence chapter we will gain mainly insight into internal protein properties even if the derived model and several conclusions could be extended to the properties of the hydration shell, too. Since we have been in need to develop our own model for explaining some of these new experimental results, due to the lack of theories concerning this topic, we hesitated to extend our conclusions to the hydration shell yet, because the observed effect of hydration, discussed in the prior section, is much less expressed than the

effect caused by temperature variation demonstrated in this section. But from a theoretical point of view an extension of the model can be done without any loss of generality. The reason for a restriction of our new model solely to internal protein motions and their corresponding far-infrared absorption at the moment is, that to this point our model is additionally substantiated by both experimental and MD simulation results. We will clearly separate indications without this additional support. At all extension possibilities, just an outlook will be given where the hydration shell molecules can be included in the near future. Because collective protein motions is a far from being understood field of research we decided, before leaving this reliable basis of bringing together experiments, simulations, and theory, first to establish our new model and our resulting comprehensive ansatz in approaching this topic.

4.6.1. Spectroscopic Results and the Analysis Procedure of the Temperature Dependence

In this section we will present our temperature dependence results and the corresponding analyzing procedure in a comprehensive way for one exemplary protein, prior to the detailed discussion concerning all measured proteins in the following sections.
All measured proteins revealed significant decreasing far-infrared absorption with increasing temperature which is the major effect on the protein spectra caused by a variation of this parameter. In fig. 4.44 this decrease is indicated by a black arrow. With a total average absorption change from 38.5°C to 4°C of about 35.7 % for BSA, this effect is much more than an order of magnitude larger than the maximal possible error bar estimation as described in section 4.3.

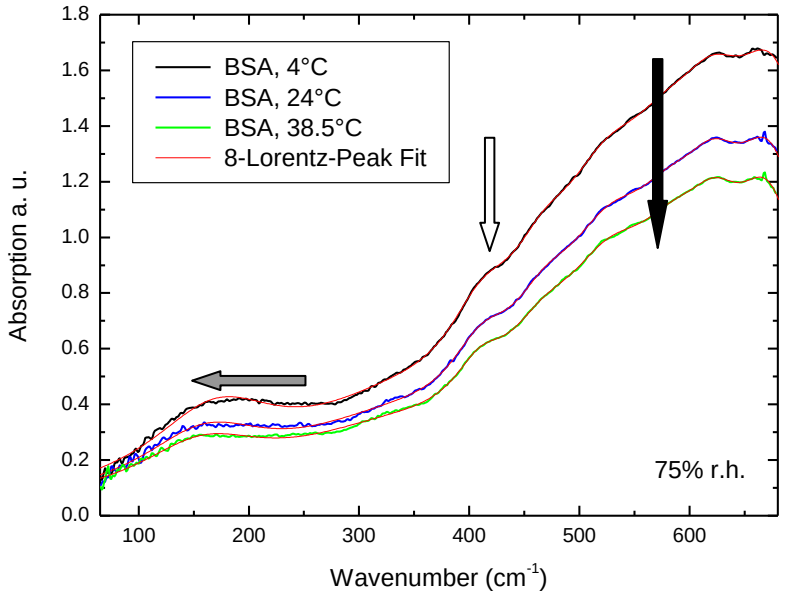

Fig. 4.44: Far-infrared spectrum of BSA under the variation of temperature. Most significant is the strong absorption decrease with increasing temperature as indicated by the black arrow. The complete temperature caused absorption change of all measured proteins is split into three different effects for the discussion as indicated by the arrows. The grey arrow labels the low frequency temperature shift and the white arrow temperature caused frequency dependent variations. The red lines represent again Lorentz fits as described in chapter 4.4.

By a closer analysis of fig. 4.44, as a representative for all four proteins, a shift of the lowest broad absorption peak becomes evident. Here we will use the Lorentz fits of chapter 4.3 again to analyze the frequency dependent shift of this peak. This analysis is the reason for a single peak fit of this broad feature in the peak assignment section to follow these temperature caused changes. Exception will be β-lactoglobulin where we have to discuss both low frequency absorption features of the spectrum.

We analyzed all protein spectra if there are other frequency dependent shifts underneath the large absorption decrease, to answer the question whether the decrease is caused by a superior frequency shift or really caused by a sole decrease of far-infrared and THz absorption. Also the large absorption decrease will be further analyzed whether a linear or non-linear dependence is observed.

Wavenumber in cm^{-1}	Absorption change 24°C to 38.5°C in %	Absorption change 4°C to 24°C in %	Absorption change 4°C to 38.5°C in %
200	15.4	25.1	44.3
300	11.0	20.5	33.7
400	13.2	22.8	39.0
500	12.2	22.4	37.4
600	11.6	22.6	36.7

Table 4.7: Relative absorption change for BSA under the variation of temperature.

A first hint on the existence of a frequency dependent shift is given in the last column of table 4.7 where the total absorption change for several frequencies is presented. For the large absorption decrease a comparison of the other two columns of the same table underlines that there is no overall linear absorption change by temperature but two temperature regions which are stronger and less affected respectively. Since we chose the upper temperature as the physiological body temperature of the corresponding specie, a normalization of the absorption change caused by 1°C will be presented in fig. 4.45 to visualize these two effects.

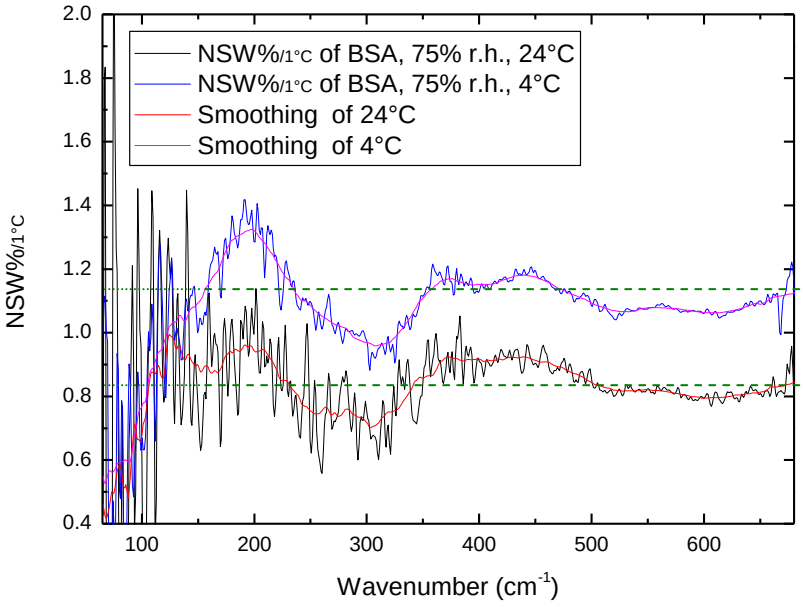

Fig. 4.45: Normalized spectral weight as a visualization for the absorption change caused by 1°C temperature change. Smoothing was gained by averaging over 40 points. A straight horizontal line (as indicated by the dashed green lines) would represent a complete frequency independent absorption decrease with increasing temperature. An overlap of both graphs in the absolute value would demonstrate a linear decrease with temperature which is not the case.

For normalization the following calculation procedure was used:

$$NSW\%_{/1°C}[\nu] = \frac{A_1[T_1,\nu] - A_2[T_2,\nu]}{A_2[T_2,\nu]} \cdot \frac{100}{T_2 - T_1} \tag{4.10}$$

with $NSW\%_{/1°C}$ as the frequency dependent, normalized to 1°C, spectral weight change in percentage and A as the absolute absorption value for every single measurement point at temperature T and frequency ν. For T_2 always the upper temperature measurement was chosen, which is for the here presented BSA a temperature of 38.5°C. As a guide to the eye a smoothing by averaging over 40 points is included in fig. 4.45, which allows a quick identification of the frequency dependent effects. On the very left of fig. 4.45 the low frequency shift is seen which is discussed separately in its own section 4.6.2. Several regions where frequency dependent shifts occur, mainly between 200 cm^{-1} and 500 cm^{-1}, can be identified which will be discussed in section 4.6.3 together with the different temperature sensitive regions represented by the mismatch of the absolute values of both graphs in fig. 4.45. The frequency dependent effects can be seen as small modulations of the much larger

general temperature effect, which are on top of a straight line as indicated in fig. 4.45 by the dashed green lines, to be discussed in section 4.6.4.
To emphasize it prior to the following discussion: Our split into these three discussion sections in this thesis is purely phenomenologically driven by our experimental results. Especially sections 4.6.3 and 4.6.4 will be described finally by the same microscopic origin on which our MERCEP model is based. But the explanation by this model of the second frequency dependent effect, discussed in section 4.6.3, is not as much accompanied by additional arguments as the much larger general temperature effect. Therefore we decided not to include a complete application and hence elucidation by this model to this effect yet, the reason for the separate discussion. We will present instead our indications we have for this experimentally observed frequency dependent shift without claiming to explain the observed effect within the protein spectra in detail. The same argumentation for leaving out the hydration shell in the new model at the moment, as mentioned already above, will be given here: We hesitated to mix just indications with the substantial argumentation for the most pronounced temperature effect, to get a clear split regarding the dependability of our new model. Due to this reason and to reduce complexity within the complete temperature dependence discussion, this form of presentation, based on a split into two separate sections even for effects likely caused by the same microscopic origin, has been chosen.

4.6.2. Discussion of the Temperature Effect on the Low Frequency Modes of the Protein Spectra

Besides the phenomenologically observed effect on the broad low frequency absorption feature of all measured proteins, the low frequency side of the spectra is in particular interesting to be analyzed separately for several reasons. First of all THz measurements on proteins are not able to distinguish between different origins of absorption changes. These could be either caused by a real increase/decrease of absorption or by a frequency dependent shift of the broad absorption feature above 100 cm^{-1} which is far out of the measuring range of usually employed THz setups and of which just the, in both cases similar, affected slope of this peak is measured. Hence we would like to present and discuss the temperature caused effects on this, in our measurements completely covered, broad absorption peak in detail. Furthermore there is an important change within spectroscopic properties around 200 cm^{-1} making a separate discussion of the low frequency part of our measured spectra advantageous.

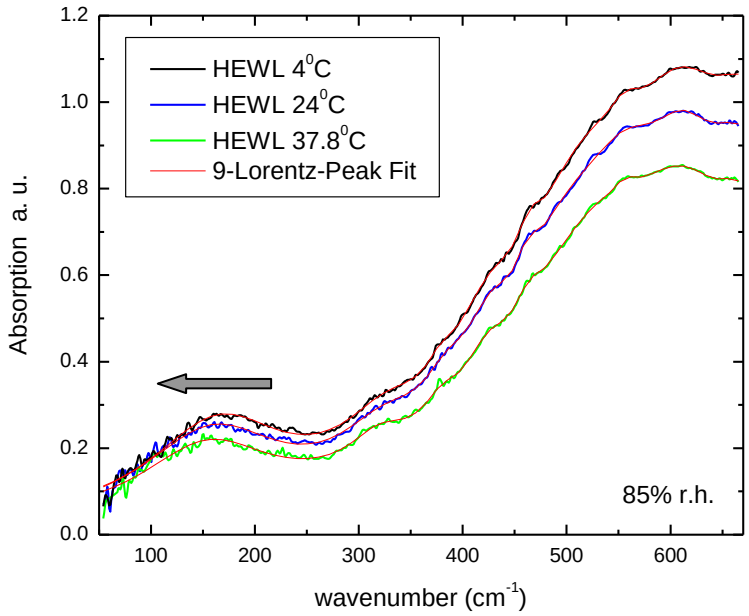

Fig. 4.46: Temperature dependent spectra of hen egg white lysozyme. The broad absorption feature in the very low frequency range exhibits a clear redshift of the peak position with increasing temperature. The red line represents the fit by nine Lorentzian peaks as described in chapter 4.3.

Due to the very pronounced low frequency absorption peak, the frequency shift caused by temperature variation is most intriguing in the spectra of lysozyme as presented in fig. 4.46. The peak position of the feature redshifts towards lower frequencies upon increasing temperature.

In our fitting procedure, using several Lorentzian peaks for the complete spectrum, the broad low frequency absorption feature was modeled by just one single peak to analyze this temperature caused frequency dependent shift within the protein spectra. The corresponding peak position for lysozyme is plotted in fig. 4.47. The peak around 160 cm^{-1} shifts about 10 cm^{-1} down in frequency by an overall temperature increase of 33.8°C.

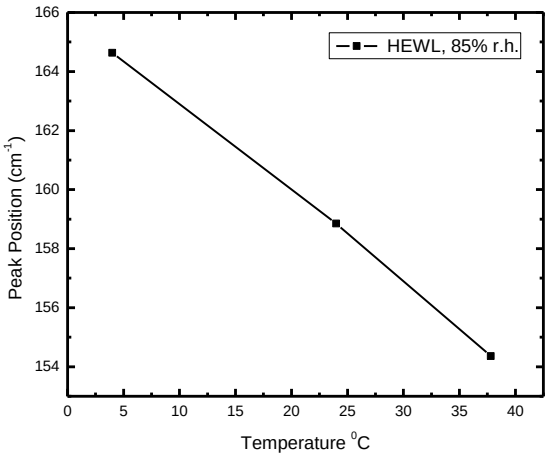

Fig. 4.47: Shift of the peak position of the low frequency absorption feature in hen egg white lysozyme by temperature variation.

The same analysis of the spectra of BSA, which were presented already in fig. 4.44, reveals an even larger shift of more than 12 cm^{-1} towards lower frequencies by a temperature increase of 34.5°C as seen in fig. 4.48. BSA seems to be more affected by temperature changes in the first range between 4°C and 24°C, which fig. 4.45 already indicated for the complete measured frequency range. This will be discussed for all proteins in the following section. But in general the deviation from a linear behavior is small considering solely the low frequency peak.

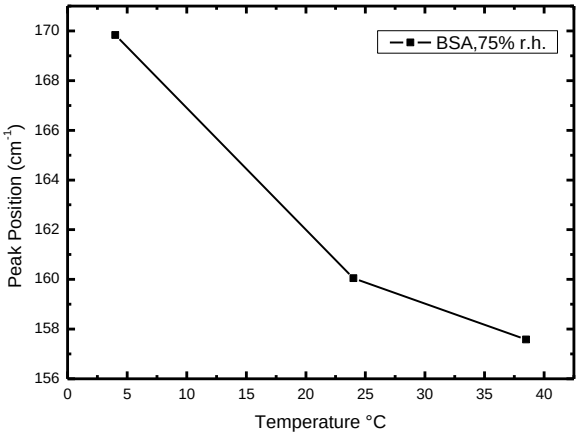

Fig. 4.48: Center peak position of the low frequency absorption feature of BSA plotted versus the temperature. The kink indicates that the low frequency peak is stronger affected by temperature changes between 4°C and 24°C than between 24°C and 38.5°C which is also observed over the whole frequency range.

A decrease of about 12 cm^{-1} in peak position is also observed for trypsin inhibitor as pictured in fig. 4.49, whose low frequency peak is situated around 170 cm^{-1}. Here the temperature difference is again 33.8°C, due to the lower body temperature of poultry, which leads to a slightly faster decrease than for the two proteins shown above.

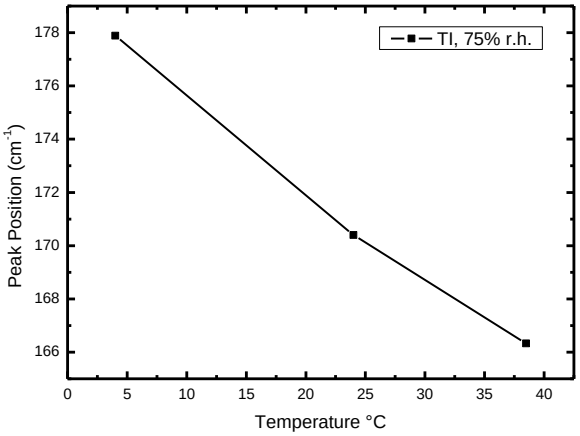

Fig. 4.49: Influence of temperature on the low frequency peak position of trypsin inhibitor as derived by a Lorentz peak fitting procedure.

Fig. 4.50 shows the temperature dependence of the far-infrared spectra of β-lactoglobulin. As already discussed in chapter 4.3 this spectrum was not able to be fit by just one single low frequency Lorentzian peak. Especially towards higher temperature values two shoulders in the absorption graph can be identified. Therefore both low frequency peaks are included in the discussion of the temperature caused effects on the very low frequency side of the protein spectra.

The corresponding peak positions are presented in fig. 4.51. First of all the temperature caused frequency shift of both absorption peaks is contrary to the direction which is observed for the other three proteins above. Both peaks shift towards higher frequencies which eliminates the fitting procedure by two peaks as the cause of this contrasting behavior. Also the very large shift of almost 30 cm^{-1} over the complete temperature variation of 34.5°C is a strong argument for a protein related effect. The increase for this peak is about three times larger than the observed decrease in frequency for the other proteins.

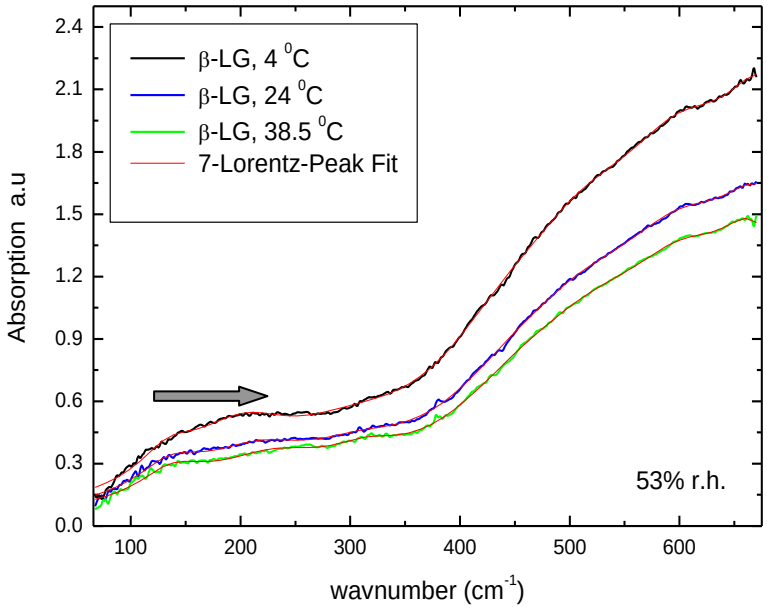

Fig. 4.50: Temperature dependent far-infrared measurements of β-lactoglobulin. The spectra were fitted properly by seven Lorentzian peaks over the whole frequency range. But the low frequency absorption differs from the other three measured proteins and had to be fitted by two peaks which are both included for a discussion of the low frequency effects upon temperature variation.

Temperature dependent THz measurements on purine and adenine [She03] as well as on polycrystalline saccharides [Upa04] revealed also a shift towards lower frequencies with increasing temperature as it is observed for the majority of our measured proteins. These are much smaller molecules which show, in contrary to large proteins, distinct THz absorption peaks whose shifts can be completely followed in the measured frequency range of these groups. The peak frequencies shifted some wavenumbers over a much larger temperature range of almost 300 K. Hence the underlying protein motions in this frequency range seem to be much more affected by temperature than the motions of these small molecules. Another reason might be the chosen temperature range. In both studies it was stated that the temperature effect becomes larger at physiological temperatures due to a similar argumentation for the cause of this frequency shift as we will give below. Therefore the temperature range of the measurements, performed just up to 295 K, might be an additional reason for observing a less pronounced frequency shift for these small molecules than for proteins. THz measurements on amino acid homopolymers using synchrotron radiation also revealed, in the same large temperature range as in the studies above, a similar shift of

absorption bands towards lower frequency [Xie99]. Again this shift is much smaller than in proteins regarding the complete, much larger temperature range.

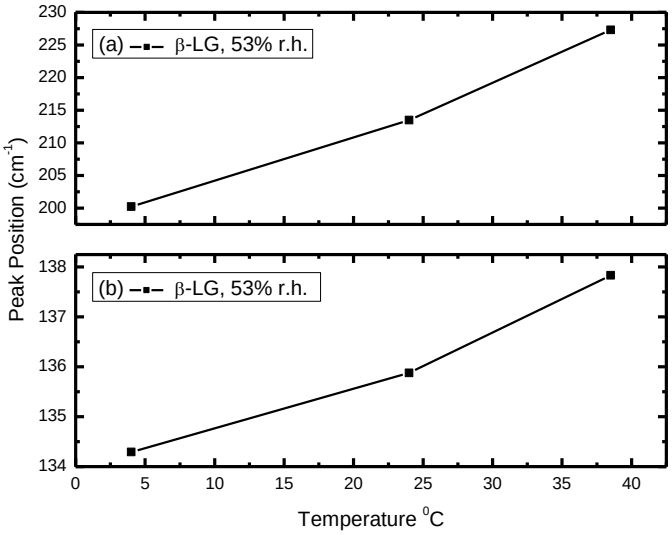

Fig. 4.51: Temperature dependence of the two low frequency absorption features in β-lactoglobulin. It is obvious that they exhibit a contrary behavior to the three proteins analyzed above. Since both peaks exhibit the same dependence, the different fitting procedure using two peaks is most probably not the cause and a diverse type of protein related origin of this effect is very likely.

For an explanation of the temperature caused redshift of the spectra we would like to emphasize the measurement of excited states in the here discussed frequency range, since for most of the published far-infrared and THz measurements on proteins such considerations are completely absent in the analysis and discussion.

As discussed already in the theory chapter, the corresponding frequency of radiation to the thermal energy at 24°C is situated in the far-infrared range, namely at 206.5 cm^{-1}. This means that in the lower far-infrared and in the complete THz region mainly transitions between excited states are probed since thermal energy is capable to populate higher levels of vibrational modes. Due to thermal fluctuations this effect will level off slowly above the frequency belonging to the average thermal energy.

The population of molecular vibration levels is described by a Boltzmann relation. If a harmonic oscillator is assumed, the population of the n^{th} level of a molecular vibration is given by

$$P_n = \frac{\exp\left(\frac{-nhc\nu}{k_B T}\right)}{\sum_{n=0}^{\infty} \exp\left(\frac{-nhc\nu}{k_B T}\right)} \tag{4.11}$$

where ν is the vibrational frequency in wavenumbers, h the Planck constant, k_B the Boltzman constant, c the speed of light and T the absolute temperature in Kelvin. We want to point out that this formula is different to the one given in the already above cited publication, for which we claim to be published incorrectly [Xie99, Pet79].

Using formula 4.11, we calculated the population of excited states for molecular vibrations at certain relevant frequencies which is presented for the first five excited levels in fig. 4.52.

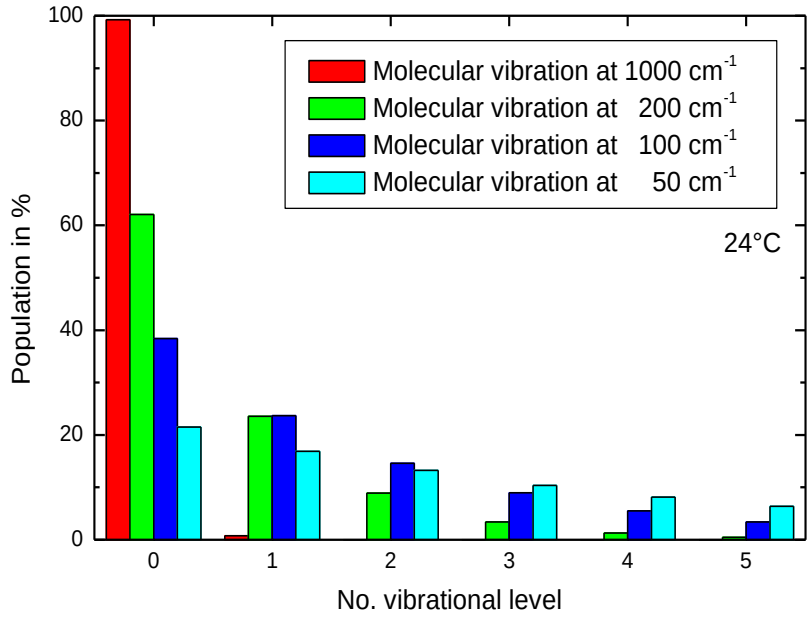

Fig. 4.52: Filling ratio of vibration levels at 24°C for typical MIR, FIR and THz molecular vibration frequencies.

Fig. 4.52 visualizes that for the MIR region above 1000 cm^{-1} the ground level is populated to more than 99 % and hence MIR radiation probes mainly transitions from this level to the first excited state. In the far-infrared, radiation at 200 cm^{-1} and 100 cm^{-1} probe about 40 % and 60 % excited states respectively. In the THz region transitions from the ground state reach a low share of about 20 % and consideration of excited states becomes indispensable. When assuming anharmonic vibrational potentials, which are more likely in this frequency region, the excited state populations will be even larger than in the harmonic case. The reason is that

including anharmonicity the energy difference between two vibrational levels decreases especially for higher vibrational states (fig. 4.54), actually leading to a larger filling of these levels than indicated from a harmonic calculation. So the importance of excited states is even underestimated in the above presented example.

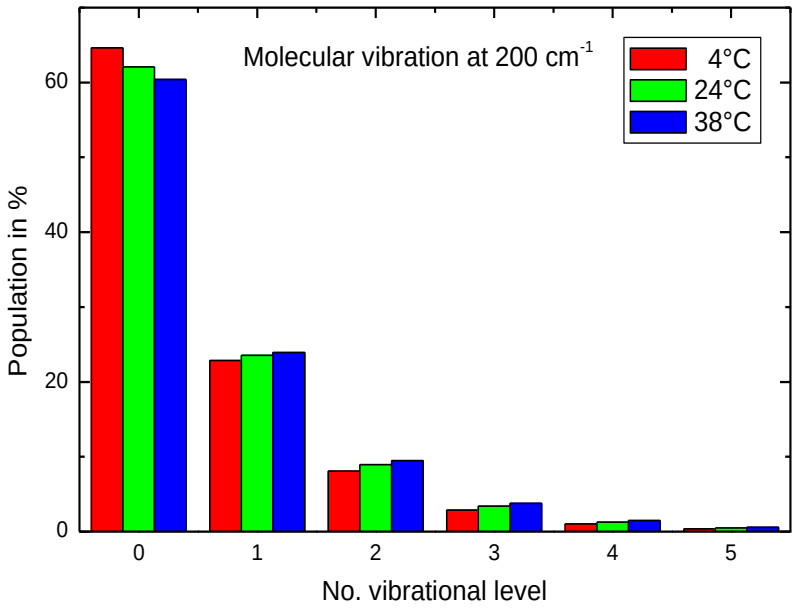

Fig. 4.53: Temperature caused changes in the population of molecular vibrational levels for the used temperatures in the measurements.

Since we observe the shift in peak frequency by a temperature variation, we want to demonstrate the change within the population of vibrational levels for these certain temperatures. For the thermal energy which is provided in average, the corresponding frequencies are for 4°C 193 cm^{-1}, belonging to 2.3 kJ/mol, and for 38°C 216 cm^{-1}, belonging to 2.6 kJ/mol, respectively. The consequences of our temperature variation up to the first five excited vibrational levels are picturized in fig. 4.53. The ground state is depopulated whereas mainly the population of the first up to the fourth excited vibrational level increases. Concerning spectroscopy, in the harmonic model this change within the distribution of the population would not affect the absorption frequency of these molecular vibrations. But in the realistic anharmonic case the energy gaps between two levels get smaller and the lower absorbed frequency of the radiation leads to a redshift of absorption features. Fig. 4.54 demonstrates typical energy levels in an anharmonic model with their decreasing gaps towards higher vibrational states. Therefore, in rising the temperature more molecules make a transition caused by IR radiation from *n* to *n+1* which occurs at lower frequency than a

transition from *m* to *m+1*, if *m* is smaller than *n* or, in the most extreme case, the ground state. The measured absorption overall redshifts with a characteristic softening of this effect towards higher frequencies.

That just probabilities are shifted aggravates a quantitative analysis of this effect. Even if the population shift would be derived, the energy potential must be also known to include anharmonicity effects on the population distribution. Further the potential is needed to quantify the energy gap reduction between every vibrational level. Additionally many different vibrational potentials in proteins must be assumed further increasing complexity. Simulations with their included force fields may be capable of modelling such a frequency shift if the prior encouraged far-infrared protein spectra calculation is applied. Such ideas can be seen as an outlook for an extended analysis of spectroscopic results after much more about protein dynamics and the underlying energy potential is understood. So far, an explanation by population levels elucidates qualitatively the observed redshift in the peak of the low frequency absorption features, at least for three of our four measured proteins. Also the observed softening of the redshift towards 300 cm^{-1} fits very nicely to this explanation which will be picked up again in the following section concerning the frequency dependent effects caused by temperature on the upper frequency region of our protein spectra.

An open and therefore very interesting question is the contrary behavior of β-lactoglobulin for which we have no explanation up to now. We can just exclude that protein size plays a crucial role for this low frequency shift since BSA, as the largest proteins in the group, is even among the three similar behaving proteins just average. The larger β-sheet content is the major difference of β-lactoglobulin to the other measured proteins of which some also contain a small amount this type of secondary structure, but we cannot identify a sound relation at the moment which would explain this contrasting behavior.

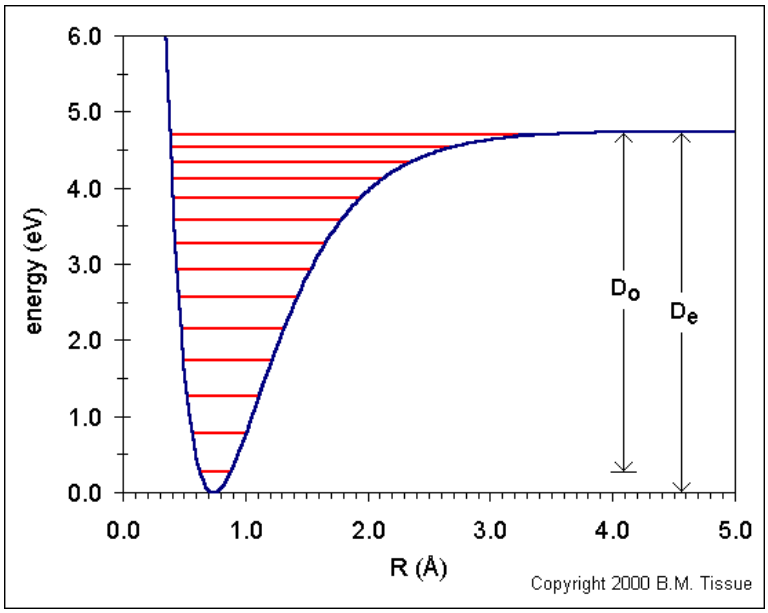

Fig. 4.54: Typical anharmonic potential with corresponding energy levels. The energy gaps get smaller towards excited levels which should visualize the explanation for the observed redshift of low frequency absorption peaks in proteins.

Excited states also lead to another, presumably minor, effect which we would like to address here for completeness of in particular low frequency absorption affecting effects. Since many molecules are in a highly excited state, a stimulated emission can take place, too. This event gets more likely at elevated temperatures since more molecules are in an excited state and the probability for a stimulated emission rises. This should reduce the measured absorption especially in the lower part of the spectra with a softening towards higher frequencies. But so far our analysis does not reveal any indication for a special or additional low frequency absorption reduction effect which is different from the large absorption decrease observed over the complete measured frequency range. If there is such an additional effect underneath this large decrease its influence on the protein spectra will be at least one order of magnitude smaller and definitely below the above given error bars of our measurements. Or, to mention all possible explanations, there is an unlikely, elaborately balanced corresponding high frequency absorption decrease which covers the significant low frequency characteristic of this stimulated emission effect in our analysis.

In general, probing excited states in the THz and lower far-infrared frequency range has been so far rarely considered for protein measurements despite the importance for any derived spectroscopic properties. Even if the low frequency shift caused by temperature was not able to be explained for all measured proteins we wanted to point out the typically expected behavior of the spectra as it is derived from basic physical properties. This should sensibilize

the discussion for taking into account excited states as a basis for a more detailed analysis of low frequency spectra. Also we would like to focus the attention on observed abnormal effects, whose explanation in particular might lead to further insight into special protein related properties.

4.6.3. Discussion of Temperature Caused Frequency Dependent Shifts in Protein Spectra

The remaining observed frequency shifts within the protein spectra caused by temperature variation are just small effects compared to the large general absorption decrease. Therefore we will only discuss indications of possible origins of these frequency shifts but which can be likely included into our new model presented in chapter 4.6.4. The basic temperature influence, affecting the bond lengths and hence the volume increase, will be already presented here since we need some of the properties for a discussion of the, in this section addressed, frequency shifts. The major consequences of the temperature caused effects will be discussed in the derivation of the new MERCEP model to which this part will finally guide to.

Protein	Absorp. change per 1°C upper temperature range	Absorption change per 1°C full temperature range
BSA	0.84 %	1.07 %
β-Lactoglobulin	0.89 %	1.42 %
Lysozyme	1.21 %	0.81 %
Trypsin Inhibitor	1.36 %	1.18 %

Table 4.8: Relative absorption changes per 1°C derived as an average from the normalized spectral weight analysis. The first two proteins are more affected by temperature in the lower temperature range whereas the other two are affected more in the upper temperature range.

To identify frequency dependent effects within the temperature variation measurements the normalized spectral weight per 1°C was derived via formula 4.10 as seen for β-lactoglobulin in fig. 4.55. Despite the abnormal temperature behavior of the low frequency peak of this protein, the frequency dependent effects in the upper part of the spectra under temperature variation are similar to the other three measured proteins.

Concerning the large frequency independent absorption decrease this analysis reveals the kind of temperature dependence and will be hence discussed in this section. By our 24°C measurement the temperature range is split into an upper and lower section. Two groups of proteins can be identified as table 4.8 reveals. The far-infrared absorption of BSA and β-lactoglobulin is more affected by temperature in the lower temperature range from 4°C to 24°C, whereas lysozyme and trypsin inhibitor are more temperature sensitive in the upper temperature range from 24°C to 37.8°C. This indicates a non-linear dependence of the large temperature effect which seems to be somehow protein type related. That both groups are split by the originating specie, namely cattle and poultry, can be not further clarified if this is accidentally or visualizes for instance the different body temperatures and hence the functional temperature range of the proteins.

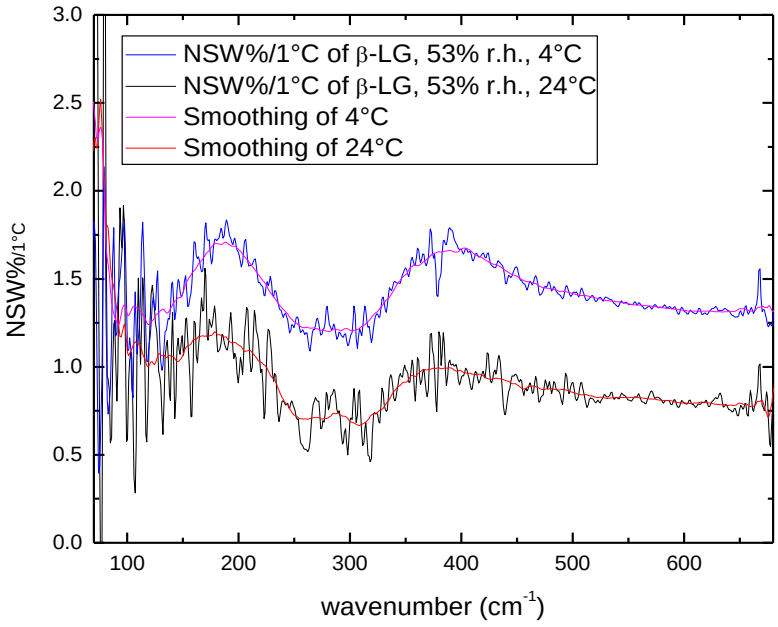

Fig. 4.55: Normalized spectral weight per 1°C as derived by formula 4.10 for β-lactoglobulin. The red line represents the average over 40 points.

Focusing again on frequency dependent effects, for all four measured proteins maxima around 200 cm^{-1} and 400 cm^{-1} separated by a minimum around 300 cm^{-1} have been found. A maximum in the normalized spectral weight, as derived in our procedure, means that at this particular frequency the absorption decrease is more pronounced than for other frequencies. The first maximum around 200 cm^{-1} might be caused by the observed redshift of the low frequency peak as discussed in the prior chapter. But including the analysis of β-lactoglobulin which shows the same peak, it is more likely to observe in general the absorption of very

temperature sensitive protein motions in this particular frequency range. The underlying motions seem to be common in all measured proteins giving reason for this similar behavior. Contrary type of motions seem to be situated in the frequency range around 300 cm^{-1}. For all proteins a minimum was derived as fig. 4.56 picturizes for lysozyme. Here the underlying protein motions seem to be less affected by temperature variations. This minimum clearly separates the above discussed low frequency temperature effect from the ones discussed in this section. Maybe spectral weight, at least for the three proteins exhibiting a redshift, is shifted towards lower frequencies causing still some part of this minimum, but this low frequency effect should level off somewhere in this region. But again, the same pronounced minimum in β-lactoglobulin indicates anyway another origin. As fig. 4.31 demonstrates there is a common absorption band in all proteins within the same region. Here, in first approximation, some unspecific protein motions seem to occur. This suggests some backbone related collective motions which fits well to the smaller temperature effect. Another argument is that strongly covalent bond structures are less affected by temperature changes than hydrogen bonds due to their different energy scale as already stated supporting our speculation.

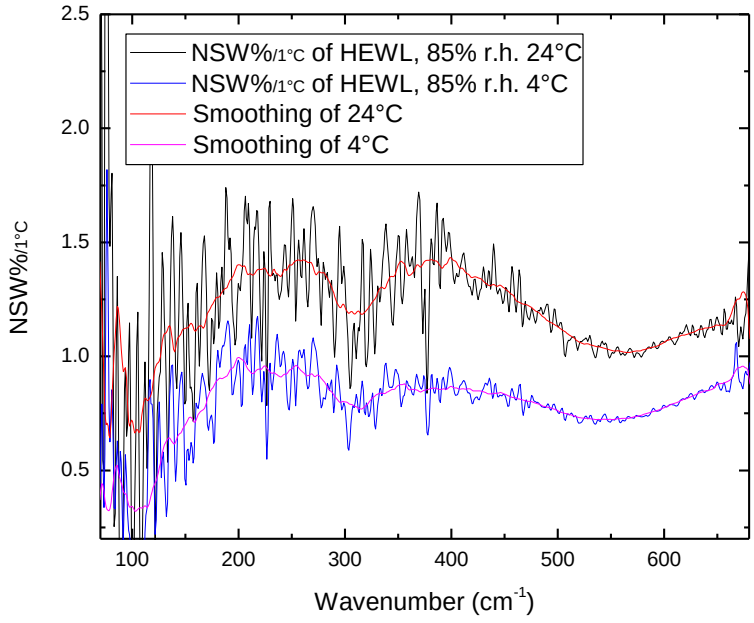

Fig. 4.56: Visualization of the temperature dependent effects by the normalized spectral weight of lysozyme. The average was taken over 40 points to get an indication for regions of different temperature affected absorption behavior.

Around 400 cm^{-1} strongly temperature affected motions seem to be present in all proteins, which the band assignment reveals, too. A larger decrease of absorption than average is indicated by the maximum in the normalized spectral weight. Only in trypsin inhibitor this peak is less pronounced as it can be seen fig. 4.57. In the upper temperature range the existence of a maximum for this protein is debatable. But above 500 cm^{-1} the temperature caused effect occurs to be under average indicating an additional shift of spectral weight within this frequency range.

By presenting the basic temperature caused effects on a molecular level, we will derive one possible explanation of these observed small frequency dependent shifts underneath the large general absorption decrease.

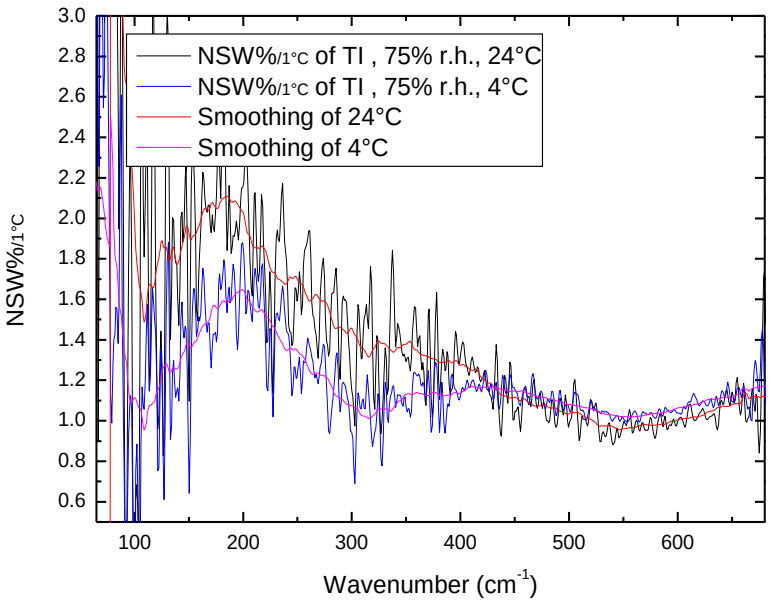

Fig. 4.57: Normalized spectral weight per 1°C for trypsin inhibitor. The same maxima and minima as for the other proteins are found except the maximum around 400 cm^{-1} for the upper temperature range. Just over the complete temperature range this maximum is present, too.

As already mentioned before in chapter 2.1.2 and 2.3.5. there is an interaction between covalent and hydrogen bonds. For instance, if a hydrogen bond length changes there is a relocation of the electron density of neighboring atoms which affects their covalent bond length and vice versa. This change will affect also the frequency of the underlying motions expressed in frequency dependent shifts within the absorption. Therefore we will discuss the effect of temperature on the here relevant covalent and hydrogen bonds in proteins, starting with the macroscopic thermal expansion.

Very different thermal expansion coefficients have been found for proteins which demand further specification of the experimental details. One of the first published linear thermal expansion coefficients of crystalline lysozyme was determined to about $1.35 \cdot 10^{-4}$/K between 298 K and 323 K [Bul73]. For tetragonal lysozyme crystals the thermal expansion coefficient was found to be $0.9*10^{-4}$/K [You94] and even a pH value dependence for lysozyme was observed with thermal expansion coefficients in this study within a range from $2*10^{-4}$/K to $3*10^{-4}$/K [Hie91].

Using X-ray on metmyoglobin different approaches of data analysis gave linear expansion coefficients of $0.5*10^{-4}$/K up to $1.60*10^{-4}$/K. The average was assumed to be $1.15*10^{-4}$/K [Fra87]. By observing the amide II peak frequency shift of myoglobin a linear expansion coefficient of $0.3*10^{-4}$/K was derived [Dem97] directly linking up to the expansion of bond length which can be observed by spectroscopic methods for instance.

For the volume increase of a protein crystal, the derived values from X-ray experiments on metmyoglobin and ribonuclease are all around 1 % per 100 K [Fra87, Til92, You94]. For our measured temperature difference we estimate therefore a maximum volume effect in the order of 0.4 %. Since the volume increase just affects our measurements in two of the three dimensions as described in chapter 3, we conclude that the volume effect of less measured protein in our spot is two orders of magnitude smaller than our observed effect and can be hence neglected in the following discussion. The underlying microscopic effect concerning the unit cell of the crystal is about twice the above values. Here the filling of empty subspace within the crystal might be the most likely explanation for the observation of much smaller macroscopic effect.

To get further down within protein structure, we finally reach the temperature effect on bonds. First of all the energy at physiological temperatures is not large enough to break covalent bonds but to break hydrogen bonds at least during the frequently occurring fluctuations. Hydrogen bonds are about one order of magnitude more temperature affected than covalent bonds as NMR studies on ubiquitin proved for instance [Cor02]. For all backbone hydrogen bonds an average linear thermal expansion coefficient of $1.7(\pm0.2)*10^{-4}$/K was derived, for the α-helical hydrogen bonds $1.9(\pm0.3)*10^{-4}$/K and for the β sheet hydrogen bonds $1.6(\pm0.1)*10^{-4}$/K were found which do not differ significantly. Estimating the volume expansion coefficient, by taking three times the linear one, this sums up to about $5*10^{-4}$/K. This is in good agreement with values for five other proteins which are all between $5*10^{-4}$/K and $6*10^{-4}$/K around 25°C [Hin94]. Rising temperature from 5°C to 65°C hydrogen bonds expand about 0.03 Å which corresponds well to the total volume expansion of the protein.

Therefore a non-covalently bond atom distance increase is responsible for the main thermal expansion in proteins. Studies on water came to the same conclusion that the intramolecular, hence covalent, bond distances are not altered even by temperature increases much larger than the increase used for our protein measurements [Sop00, Dou98, Ich91b]. Similar results on a stronger affection of hydrogen bonds than for covalent bonds were found for brushite which forms intermolecular hydrogen bonds as well as hydrogen bonds with water molecules [Sch04]. Modig et al. published *ab initio* calculations together with a comparison to measurements on water from 0°C to 80°C and concluded that the covalent bond with about 1.85 Å does not change within the complete temperature range supporting the arguments of mainly hydrogen bond caused thermal expansion. In their introduction they stated that "there is surprisingly little information available about the temperature-dependent hydrogen-bond geometry" in water up to that date, underlining that these issues have not been addressed frequently concerning even less complex water. Including the argumentation of excited states in the THz and far-infrared frequency range, there is an additional increase in bond length not due to thermal expansion but just due to the increased anharmonicity at elevated temperatures which increases the average hydrogen bond distance, too [Gho77].

Experimental data on deuterated portlandite, $Ca(OD)_2$, using time of flight neutron diffraction reveal that the hydrogen bond length increases with increasing temperature causing another effect. When the D···O hydrogen bond in the D···O-D arrangement gets weaker, the O-D bond may become shortened and thus strengthened [Xu07]. The contrary effect is observed if the increased motion of the A-H stretching vibration at elevated temperature weakens the hydrogen bond which is less stable [Jef97]. These effects have not been systematically investigated up to the date of the book in 1997. For spectroscopy this means that in the mid-infrared, which mainly probes the covalent bond motions, either a red- or blue-shift can occur depending on the interplay between the covalent and hydrogen bond lengths.

To give one of the many mid-infrared examples: If a quasiharmonic solid is assumed the normal modes frequencies shift down with increasing volume described by the Grüneisen-parameter which is typically close to unity. For at least some vibrational modes, e.g. the O-H stretching vibration in the mid-infrared, the authors claim that this relation is valid, too. The amide I and II are both affected by hydrogen bonds. But two different temperature dependencies have been found. For the amide I, which is dominated by C-O stretching vibration, the peak frequency is increased with rising temperature due to a softening of the C-O···H hydrogen bond. Amide II, as a combination of N-H stretch and C-N bending, behaves the opposite because the bending force constant decreases with a softening of the N-H···O

coupling [Dem97]. This example shows that frequency shifts in both directions are observed in temperature dependent mid-infrared measurements [Hey02, Man00, Dem97], so the vice versa effects may partially compensate each other demanding further studies just focussed on this particular topic to finally answer the microscopic effect on bond lengths in proteins caused by temperature for which we can just present arguments at the moment without knowing which effect outplays the other for each type of triplets of A-B···C occurring in proteins. Since we are likely to probe hydrogen bond antagonist of the A-B···C triple in the far-infrared, either red and blue shifts by temperature variations can occur but lacks the same quantitative explanation for complex proteins as in the mid-infrared.
For our temperature triggered frequency dependent shifts we want to point out our indications for temperature sensitive motions in proteins around 200 cm^{-1} and 400 cm^{-1} and the, less temperature affected and hence presumably backbone related, motions around 300 cm^{-1}, each first of all independently of the type, the size and the secondary structure content of the measured proteins. Some protein type related and less temperature sensitive, maybe starting to be more covalent bond attributed, motions were assumed above 500 cm^{-1}. The discussion of the microscopic temperature effects concluded both red and blue shifts depending on the interplay of covalent and hydrogen bonds. Within a complex protein at least this qualitatively describes the presence of frequency dependent shifts underneath our major frequency independent effect which will be in focus of the following section.

4.6.4. Discussion of the Frequency Independent Temperature Effect on Far-Infrared Protein Spectra

The major temperature caused effect on the far-infrared spectra of all measured proteins is a general absorption decrease with rising temperature as indicated by a black arrow in fig. 4.44 at the beginning of chapter 4.6 about our temperature dependent measurements. Along the discussion of temperature caused frequency shifts in the previous section, we already demonstrated in the analysis that this effect is of non-linear nature. We also found that the temperature sensitivity of the spectra does not depend on the temperature difference but on the absolute value of temperature itself and that sensitivity is maybe related to the originating specie of the protein, or more precise, to the corresponding body temperature and hence the physiological relevant temperature range.

In this section we will focus on the interpretation of the observed absorption decrease. Since there has been no model present up to date to explain these new far-infrared measurements under variation of controlled temperature and humidity, we were forced to develop our own model for a further analysis and interpretation of our experimental results gained under the influence of temperature variation. We will start this section with a theoretical approach considering possible origins of the observed temperature caused absorption change, deriving the main ideas for our later presented MERCEP model. These ideas will be used as a guideline throughout the following discussion of the accompanying MD simulations for the MERCEP model and as an orientation where this section will finally lead to. The simulations have been employed as a custom made testing of certain aspects of the MERCEP model and to substantiate our new model with additional arguments. Further insight into the temperature triggered microscopic processes has been gained and also some, on a first glance unexpected findings, were used for a refinement of the here presented latest version of the MERCEP model. All these microscopic effects will be discussed along the two types of bonds which are most likely to be affected by our experimentally applied temperature changes due to their energy scales.
Thereafter the following section's topic deals with the transfer from all these microscopic effects to the consequences for far-infrared absorption spectra as a macroscopic, entire protein property to finally match the theoretical approach and the MD simulation results with our far-infrared measurements within the new model. At last the influence of temperature on the two important parameters, collective protein motions and dipole momenta, will be generalized to a, in this case energy triggered, change of correlation within the protein, subsummized in our Model of Energy Related Correlation Effects in Proteins (MERCEP).

4.6.4.1 Discussion of Possible Origins of the Measured Absorption Decrease as a Guideline to the MERCEP Model

As seen in fig. 4.58, we observe a large frequency independent absorption decrease upon increasing temperature for Trypsin Inhibitor as well as for all other measured proteins. This effect is much larger than the two other prior analyzed temperature caused frequency dependent effects. We will use several conclusions, already derived in preceding sections, which will be therefore summarized here to gain a comprehensive discussion about the major frequency independent temperature effect on far-infrared protein spectra.

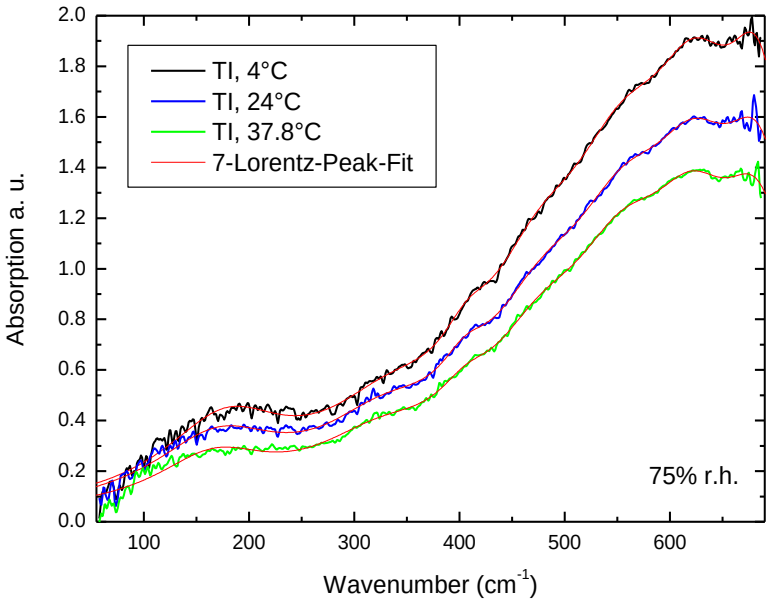

Fig. 4.58: Far-infrared spectra of Trypsin Inhibitor under temperature variation. As for all measured proteins, the major temperature caused effect is a frequency independent absorption decrease. The red lines represent the result of a fitting procedure applying seven Lorentzian peaks which were used for the analysis of the spectra.

One of the possible origins for a decreasing absorption might be a reduction of the far-infrared light absorbing water content within the protein sample caused by a temperature increase. In the hydration dependence chapter 4.5. we already discussed the influence of temperature on the surrounding relative humidity and hence on the amount of water molecules attached to the protein. Several arguments derived in that discussion can be used as evidence for an exclusion of water as the major origin of the here discussed temperature effect. First of all the relative humidity has been stabilized by a calibration standard method to the same value irrespective of the applied temperature. Therefore occurring deviations are expected to be very small as already stated in the experimental chapter 3 and the reproducibility section 4.3, hence too small to cause such a huge temperature triggered effect on the protein spectra. Secondly the altered amount of water molecules is about an order of magnitude larger during our hydration dependence measurements than deviations maybe occurring during the temperature variation. But the resulting hydration caused consequence for the far-infrared spectra is about two orders of magnitude smaller than it is observed for a temperature dependent measurement which is too far off to constitute the origin of such a large absorption decrease. Another argument can be derived from our conclusions about biological water properties. The absorption of these

molecules was found to be at least an order of magnitude smaller than for liquid water, estimated likely to be even smaller than the protein absorption itself, so that these molecules are in general not capable of causing such a large impact on the far-infrared absorption spectra.

In addition THz reflection spectroscopy on liquid water derived contrary temperature dependence to our protein spectra which we will use in our adjacent far-infrared frequency range as another argument for excluding water as the origin of our measured absorption decrease. It is very likely that extrapolation of the published THz measurements seen in fig. 4.59 indicates also an absorption increase with increasing temperature in the far-infrared. The temperature effect becomes even larger towards the upper far-infrared frequency side of the published data supporting our assumption. Under the hypothesis of a similar temperature dependence of biological water absorption as for liquid water, the opposite behavior under the influence of temperature variation excludes water as the provenance of the temperature effect. If biological water reveals the same absorption temperature dependence, this would just marginally antagonize our measured absorption decrease which we want at least to mention here once for considerations of a further refinement of the model in the future. But since we had to assume a temperature dependence of biological water, this argument should be just additionally taken to the other more substantial evidences that protein attached water molecules are not the origin of the absorption decrease with rising temperature.

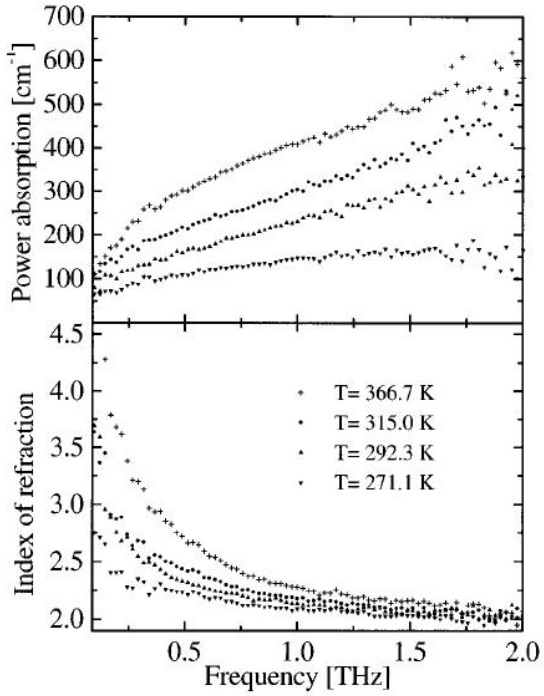

Fig. 4.59: The THz absorption of liquid water reveals an increasing absorption with increasing temperature [Ron97]. Since we observe an absorption decrease of our protein spectra upon a temperature elevation in the adjacent far-infrared frequency range this contrary temperature dependence represents one of several arguments for excluding protein attached water as the origin of our measured temperature effect.

A simple volume increase upon temperature elevation, leading to less dense and a smaller amount of protein within the measuring spot, can be excluded as well as the major origin of the observed decreased absorption. As stated in chapter 4.6.3. the volume change itself is about two orders of magnitude smaller than the measured absorption effect. But if the microscopic consequences of temperature caused volume and bond length changes could trigger an enhancement of relevant properties for the macroscopic observable far-infrared absorption then such a large effect might be explainable.

Several prerequisites for causing such a large effect, derived from the analysis of the measurements and from theoretical considerations, will be utilized as a guideline to the MERCEP model. For better storylining we will concentrate on the major temperature caused effect since the smaller frequency dependent effects have been in focus of the preceding sections: In the first place the temperature dependence measurements reveal a major frequency independent effect on the protein spectra. In such a simplified picture there are no indications for large scale shifts of spectral weight within the measured far-infrared frequency

range and the over all shape of the absorption graph remains under the variation of temperature. Therefore it is very likely that temperature affects mainly the amplitude of the far-infrared absorption. In the theory chapter we derived our extended version of the dipole moment correlation function to calculate far-infrared absorption spectra e.g. based on MD simulation results (see formula 2.39). The determining parameter for the absorption is the dipole moment which could be divided into a contribution from permanent and induced dipole momenta (formula 2.40, including also an additional cross correlation term). Breaking of bonds, internal distance increases between non-bonded atom pairs, sidegroups or secondary structure elements, and the loss of the collective nature of motions, likely assumed to be observed in the far-infrared frequency range, directly affect permanent and induced dipole momenta within the protein to which infrared light absorption is coupled to. Since either the number of motions is altered upon breakage of hydrogen bonds and the loss of collective motions mediated by hydrogen and van der Waals bonds, or just the strength of the induced dipole, with its the r^{-6} distance dependence of most important van der Waals bonds, is influenced by temperature variations, the sought for explanation for a frequency independent absorption change is at hand. These ideas of possible origins of the temperature caused consequences for far-infrared absorption will be the guideline to the MERCEP model which condenses all these effects in one generalized model regarding to the correlation within the protein.

As pointed out in sections 2.1.2.2. and 2.1.2.3 hydrogen and van der Waals bonds are likely to be influenced by our experimentally applied temperature change. Hence we will focus on the influence of temperature on these two bonds modeled by MD simulations before discussing the consequences for the far-infrared protein spectra.

4.6.4.2. Methodology of the MERCEP Model accompanying MD simulations

For a better understanding and validation of certain aspects of our MERCEP model, several hypotheses about temperature caused changes within the structural protein dynamics have been addressed and tested in MD simulations of β-lactamase. Since we were looking for general properties of protein dynamics the choice of the kind of protein was less important than existing experience concerning the sample's protein structure, surrounding water

molecules, reliable force fields and the sound knowledge about the model system which are all determining factors for the validity of the simulation results.

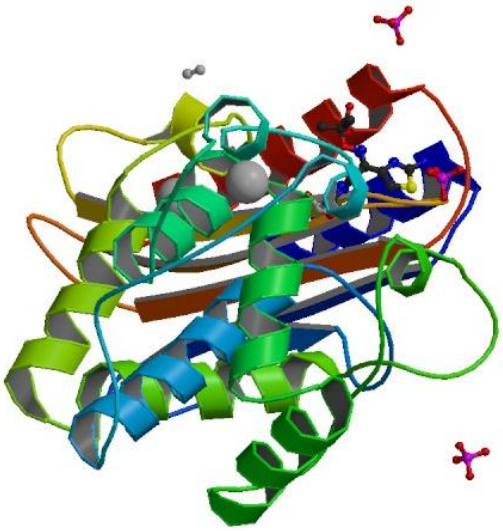

Fig. 4.60: All accompanying MD simulations of the MERCEP model have been performed on the enzyme TEM-1 β-lactamase, PDB code 1M40 [PDB08].

For the simulations the enzyme TEM-1 β-lactamase with PDB code 1M40 was used whose crystallographic structure corresponds to the mutant M182T [Min02, Wan02]. The protein consists of 263 amino acids with a total of 6,419 atoms, summing up to a weight of 29.7 kDa and a global charge of -7 at physiological pH. From the original crystallographic structure the analogue inhibitor, ions and water molecules have been removed before 1,383 hydrogen atoms, seven sodium counterions and 12,922 explicit water molecules were added by the LEaP module of AMBER. The system was immersed in a cubic box of TIP3 water solvent [Jor83] and a closeness factor of 1.0 Å was defined between the box and the protein surface. Using an AMBER 9 software package [Cas06] and an all-atoms AMBER force field "ff99" [Wan00], Particle Mesh Ewald molecular simulations have been performed in an isothermal-isobaric (NpT) ensemble. The whole system was coupled to a weak barostatic and thermostatic bath of 1.0 bar and the requested temperature respectively [Ber84]. To all C-H bonds the SHAKE algorithm [Gus77] of 1 fs time steps was applied.
By performing simulations at six different temperatures (270 K, 280 K, 290 K, 300 K, 310 K, 320 K) the influence of temperature on the structural protein dynamics was modeled. At every temperature three independent simulations of 4 ns length were carried out, starting from differing initial sets of Maxwell-Boltzmann distributed velocities. After an energy minimization for one thousand steps using the steepest descent and another thousand steps the

conjugated gradient algorithm, the system was heated up to the simulation temperature during three 20 ps periods by constraining the C^{α}-atoms of the protein backbone by a harmonic potential using force constants of 10.0, 1.0 and 0.1 kcal/mol respectively. Afterwards a 2 ns equilibration was applied until the root mean square deviation (RMSD) became constant. Every 0.5 ps the conformational space was sampled and for analyzing the ptraj module of AMBER was used.

One focus of the analysis concerns the influence of temperature on the hydrogen bonds of the system. Masks for potential hydrogen bond donors and hydrogen bond acceptors were defined, the percentage of occupancy and the lifetime was analyzed every 0.5 ps during the last 2 ns of each trajectory. The maximum number of interactions for a donor or an acceptor has been restricted to six and the cutoff distance and the cutoff angle for hydrogen bonds have been set to 3.5 Å and 120° respectively. Another focus of the analysis regards the atom pair distances defining the occurrence of induced dipole momenta in proteins which is one determining parameter for far-infrared absorption. Again 0.5 ps steps of the last 2 ns have been used to derive the radial distribution function (RDF) and root mean square deviations have been calculated upon the minimalized structure as a reference. For the RDF analysis a spacing of 0.002 Å between two bins has been considered up to a maximum range of 10 Å. For the protein volume calculation of each simulation step snapshot a Richard's rolling probe method [Ric77] and a Vox Volume Voxelator (^{3}V) algorithm [Vos06] were applied. The simulation results have been analyzed once for the backbone atoms C^{α}, C^{β}, N, and O only without the solvent, focusing on the internal protein distances, as well as for all protein atoms in a second analysis, whose discussion is the topic of the next section.

4.6.4.3. Microscopic Effects upon Temperature Variations: Discussion of the corresponding MD simulation results

The microscopic consequences of temperature variation upon hydrogen and van der Waals bonds are in focus of the here presented molecular dynamics simulations. The applied temperature change might be able to break or at least influence the two kinds of bonds due to their energy scales and these temperature caused effects we want to substantiate by corresponding simulations prior presenting our new MERCEP model. The published thermal expansion coefficients of hydrogen bonds and proteins agreed to be all in the range of 10^{-3} to 10^{-5} per K (see chapter 4.6.3). Hence the expected temperature effects, identifiable by shifts

within the radial distribution function, are small for our applied temperature elevation with an estimation in the range of a hundredth up to a tenth of an Angström reaching the error bars' edge of the present simulation results. Nevertheless the influence of temperature on the all atom distance has been modelled as seen in fig. 4.61. The observed peaks in the radial distribution function represent typical bond lengths within the protein. Upon rising temperature the peaks slightly lower and broaden due to a larger variation of bond lengths at elevated temperatures. If there the expected shift towards larger bond length is present in these simulations cannot be completely verified and is one reason for the below presented closer analysis of only one type of atom pairs to get clearer indications than by an all atom analysis.

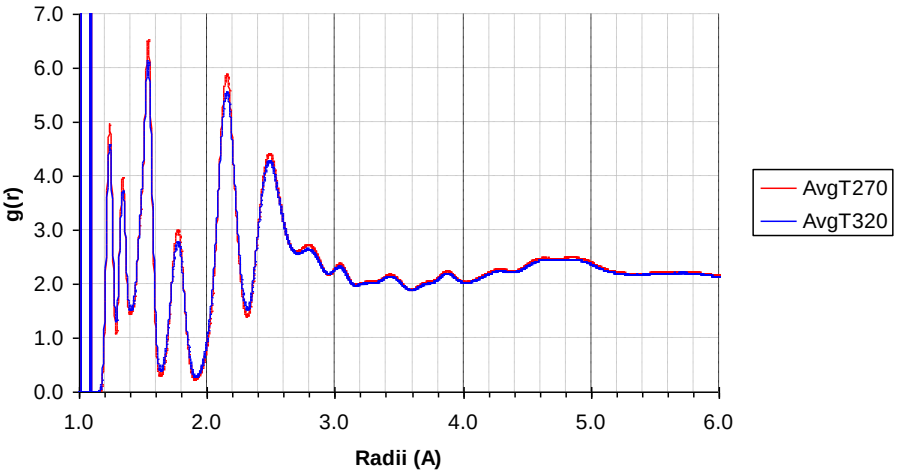

Fig. 4.61: Radial distribution function of all protein atoms derived from the last 2 ns of the simulation in 0.5 ps steps. An average over each of the three simulations for the temperatures of 270 K (red) and 320 K (blue) was taken to investigate the influence of temperature on the interatomic distances. The temperature effect of 50 K on the distances is small but leads to the expected lowering and broadening of the RDF peaks representing typical bond lengths in proteins.

For the further analysis of the MD simulations the O···H distance was chosen because between these atom pairs all three types of bonds covalent, hydrogen and van der Waals bonds occur in proteins, giving us the possibility for a focussed but comprehensive analysis of temperature effects on the structural dynamics of proteins. The radial distribution functions for the O···H distance of all 18 carried out molecular dynamics simulations are picturized in fig. 4.62.

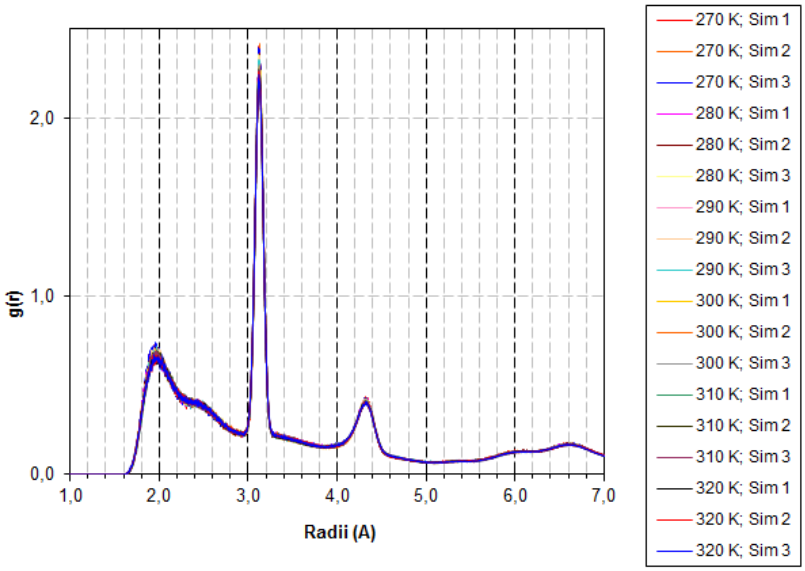

Fig. 4.62: Radial distribution functions for 18 MD simulations in the temperature range of 270 K to 320 K, based on the last 2 ns analyzed in 0.5 ps steps. Three peaks can be addressed which are predominately distances between a hydrogen atom and an oxygen atom respectively three, two or four residues ahead in the protein chain. The effect of temperature on each peak is scrutinized along the following three figures.

The analysis of just one atom pair leads to more pronounced and clearly identifiable peaks in the radial distribution function compared to the all atoms analysis. The distance between the hydrogen atom and the oxygen atom three residues ahead is closer than the distance between the atom pair two residues ahead which is intelligibly by recapturing the theory chapter about secondary structure formation 2.1.1.2. which also applies for this α-helix dominated but also a large amount of β-sheet containing protein.

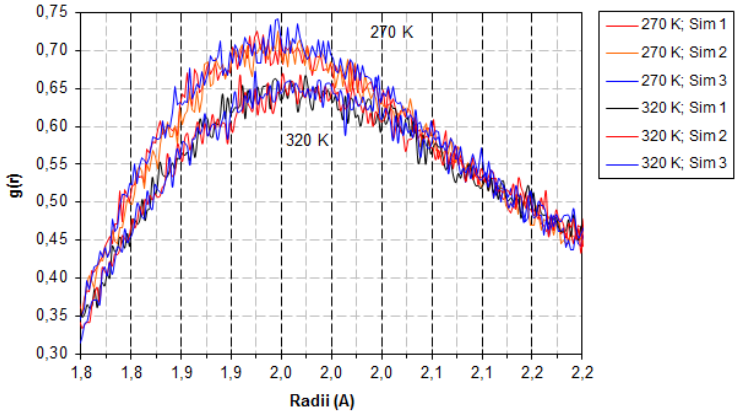

Fig. 4.63: Close up of the first peak in the radial distribution function (numbering from fig. 4.62) mainly referring to the $H_n \cdots O_{n-3}$ atom pair distance in β-lactamase. The upper three graphs represent the simulation results at 270 K and the lower three the results at 320 K. Upon temperature increase the left slope and the maximum are lowered indicating either a shift towards larger bond length and/or the breakage of shorter hydrogen bonds. There is a small indication that the maximum is slightly shifted towards a larger average bond length at elevated temperature.

The close up of the first peak in fig. 4.63, mainly attributed to the $H_n \cdots O_{n-3}$ distance, reveals an about 7% lowering of the peak upon the modelled temperature increase. The distance of 2 Å clearly indicates a hydrogen bond length but the shift of the maximum of about a few tenths of an Å is too small to hold as a sound proof for a temperature caused hydrogen bond length increase. At least the results indicate that a bond length increase is more likely than retention or a decrease. Especially the shift of the short distance slope on the left together with the contemporary match of the graphs at the longer distance side on the right are an allusion to a bond length increase and/or a loss of particularly shorter hydrogen bonds. As stated in chapter 2.1.2.2. a hydrogen bond breakage will lead to a non-linear distance increase of more than 0.3 Å caused by the repulsive van der Waals radii of the non-hydrogen bonded atom pair. But since we probe here an average distance the consequences of a breakage of just a few hydrogen bonds will be small. Therefore additional hydrogen bond properties like the total number of hydrogen bonds and occupancy ratios, based on an individual hydrogen bond analysis, will be separately discussed below after these radial distribution analysis to get further insight into the temperature caused effect on particular this type of bonds.

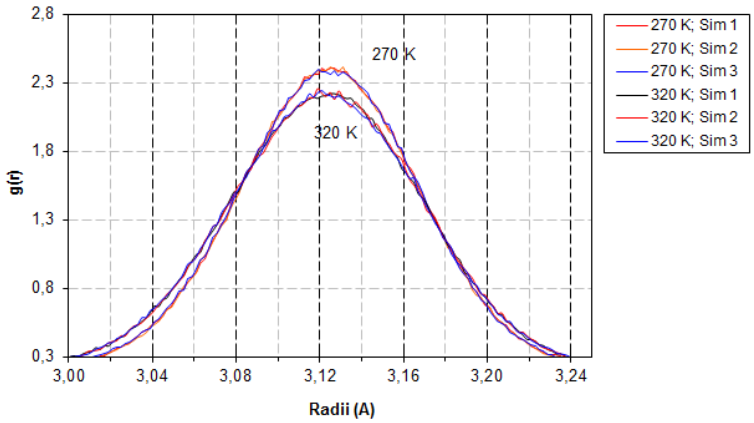

Fig. 4.64: Analysis of the second peak in the radial distribution function from fig. 6.42. Here the main contribution of the $H_n \cdots O_{n-2}$ atom pair distance is found with a maximum around 3.1 Å. The upper three graphs are the 270 K simulation results, the lower three the ones the 320 K. The maximum is lowered, maybe shifted towards shorter bond lengths and the width is increased on both sides indicating an extension within the bond length variation at elevated temperatures.

The second peak (fig. 4.64), indicating principally the $H_n \cdots O_{n-2}$ distance, gets about 8% lower and broadens upon the modelled temperature increase. The maximum is slightly shifted towards a shorter bond length, also the increase of shorter hydrogen bonds on the left slope is more pronounced than the increase of larger bond lengths to be found at the right slope pointing more in a direction of an average bond length decrease and/or again the breakage of some hydrogen bonds of this particular kind of atom pairs. Besides these indications, evidence for an increase in variation of bond lengths around 3.1 Å at elevated temperature is given by the simulation results.

Publications concerning the impact of temperature on hydrogen bonds derive an increase of the atom pair distance at least for water. The corresponding radial distribution function shows small shifts towards larger distances below 5 Å, most pronounced in the range between 3 Å and 5 Å, upon a temperature increase from 4°C over 25°C to 50°C [Nar71]. Temperature dependent X-ray experiments on crystal lysozyme [Kur95] revealed an increase of hydrogen bond distances but triggered by a much larger temperature increase from 110 K to 285 K which just reaches the edge of the protein's physiological relevant temperature range where the largest impact of temperature is assumed. Nitrogen atoms involving hydrogen bonds are found to be in average 0.1 Å longer and more temperature dependent than the oxygen hydrogen bonds. In general intermolecular contacts in a distance between 3 Å and 4.5 Å become less which concerns hydrogen bonds as well as atom pairs in a van der Waals distance. If the contacts get less also the induced dipole moment can be assumed to be reduced.

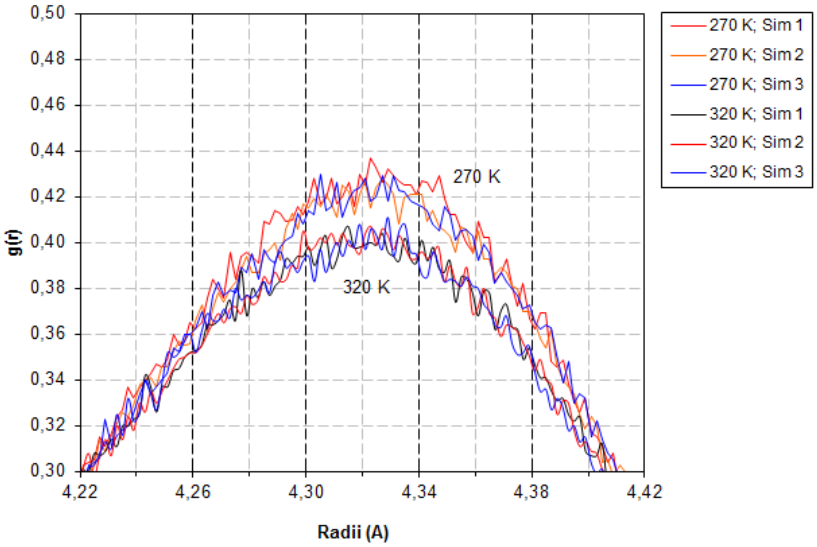

Fig. 4.65: Temperature effect on the third peak of the radial distribution function predominately arising from $H_n \cdots O_{n-4}$ atom pairs which are clearly in van der Waals bond distance. Three simulation results each depict the 270 K and 320 K simulations respectively. The maximum is lowered but almost remaining the atom pair distance upon the temperature increase. At shorter distances temperature has less impact on the van der Waals bonds than for the van der Waals bonds of larger distances at the slope of the right hand side of the figure. Indication for a decreased number of van der Waals bonds at larger distances than 4.26 Å is given.

Clearly in a van der Waals bond distance is the third peak of the radial distribution function picturized in fig. 4.65. Here the temperature increase causes again a lowering of about 5% but the position of the maximum remains around 4.32 Å. Starting at distances above 4.26 Å a decreasing number of predominantly $H_n \cdots O_{n-4}$ atom pairs can be identified, maybe one hint for a loss of van der Waals bonds within the protein upon rising temperature, an effect which we will pick up again in the following section.

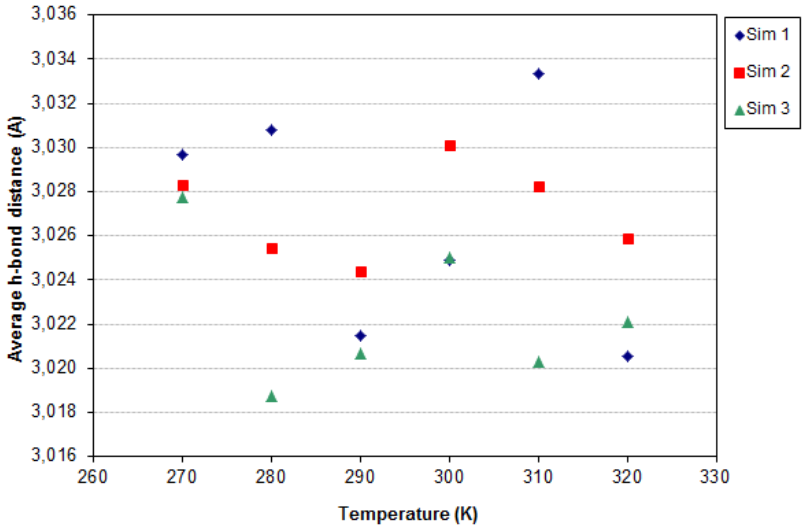

Fig. 4.66: Average hydrogen bond distance at different temperatures derived from an individual hydrogen bond analysis of three simulations each, giving no clear indication for a bond length change under temperature variation coinciding the outcome of the radial distribution function analysis.

Regarding the two hydrogen bonds concerning peaks of the radial distribution function our expectation about the lowering and the broadening of peaks have been met, indicating a shift of bond lengths and/or a reduction in number of hydrogen bonds upon a temperature increase. But puzzling have been indications for contrary shifts of the maxima and the slopes towards larger distances for the shorter hydrogen bonded atom pairs and the opposite for the larger hydrogen bonded atom pairs, provoking a further analysis which leads us to a more sophisticated temperature effect on hydrogen bonds than just a simple atom pair distance increase.

The molecular dynamics simulation results yield the possibility of an additional, individual hydrogen bond analysis which will lead to a better understanding of the temperature impact. As described in the prior section, for every O···H atom pair it was checked at each time step if the determining parameters for a hydrogen bond are fulfilled or not. But also this analysis (fig. 4.66), focussed on hydrogen bonded atom pairs and not only on atom pairs at typical hydrogen bond distances, derives no clear indication for a change of the average hydrogen bond length under temperature variation which is in line with the above ambiguous radial distribution function results.

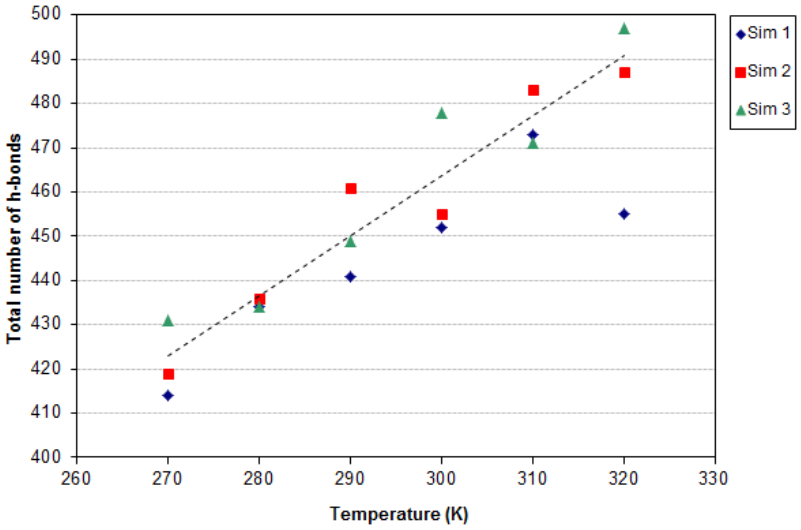

Fig. 4.67: Total number of hydrogen bonds which are formed at certain temperatures. For each temperature three different simulations have been analyzed. As a guide to the eye the line emphasizes the increasing number of hydrogen bonds with increasing temperature, representing an about 15% increase upon a 50 K temperature elevation based on the lowest simulated temperature of 270 K.

The temperature dependence of the total number of hydrogen bonds, as seen in fig. 4.67, reveals an increasing number of hydrogen bonds with increasing temperature. During an averaged time step at 320 K in total about 15% more hydrogen bonds are formed than at 270 K. Considered solely, this result might even indicate an increase in hydrogen bonding upon a temperature elevation but just together with fig. 4.68 the real nature of the impact of temperature on hydrogen bonds becomes apparent. Here the average occupancy, hence the average lifetime, of a hydrogen bond is analyzed, clearly resulting in a decrease of about 15% upon the same temperature increase. Hence during one averaged time step all O···H atom pairs are in total less hydrogen bonded at 320 K than at 270 K which coincides well with our hypothesis of a less bonded protein interior at elevated temperatures. The absolute number of formed hydrogen bonds does not allow such a conclusion by itself and unaccompanied interpretation indicates even the opposite on a first glance. But the above simulation results manifest that the lifetime of hydrogen bonds is the primarily parameter which is actually influenced by a temperature variation. The reduction of the temporal mean of all present hydrogen bonds exceeds the increase in number. Even more as we will elucidate below, under this perspective the increase in number of hydrogen bonds is a plausible consequence of the decreased lifetime of initially hydrogen bonded atom pairs at lower temperatures.

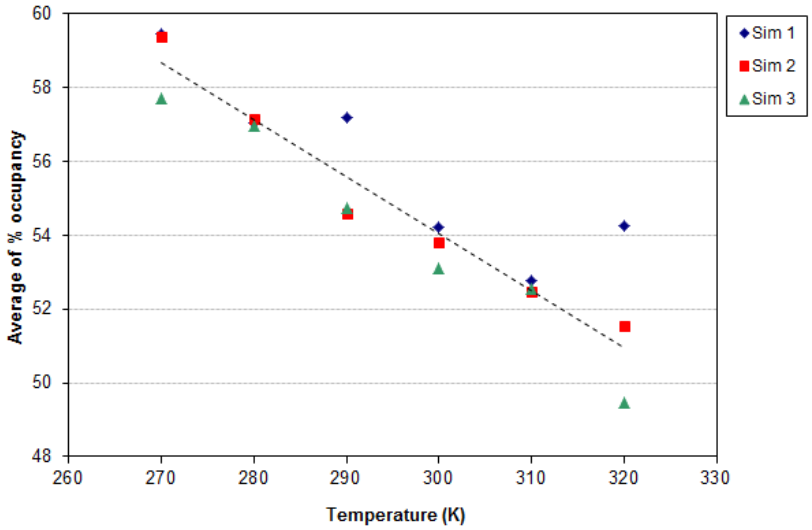

Fig. 4.68: Averaged occupancy of hydrogen bonds in per cent under the influence of temperature variation. The number of hydrogen bonds increases but the time scale for hydrogen bond formation is reduced. Together with fig. 4.67 this indicates an enhanced hydrogen bond formation and breaking at elevated temperatures.

Combined together we claim that the results characterize an enhancement of the structural dynamics within the protein. Triggered by elevated temperature, breakage of hydrogen bonds becomes more frequent but along, the formation of hydrogen bonds is also increased resulting in the temporal average to be the determining quantity to look at when probing a temperature caused effect on hydrogen bonds in proteins .This interpretation fits well with both, the increased number and the shorter lifetime of hydrogen bonds which are gained by the simulation. As the above findings suggest, temperature seems to affect not all hydrogen bonds but causes a change of the dynamical behavior of bond formation for some hydrogen bonds and further causes an increase in the number of hydrogen bonds which reveal such binding dynamics. The consequences on an average bond length are minor as the above results indicate and the temperature effect becomes just apparent by tracking the individual bond formation using each trajectory of the simulation.

A plot of the number of hydrogen bonds versus the occupancy, as seen in fig. 4.69, visualizes the intensified dynamics. As the most characteristic example the number of continuously linked atom pairs decreases, as the black arrow emphasizes, and more often hydrogen bonds of shorter lifetimes are formed with clear evidence at least for those hydrogen bonds within the dotted areas. Under temperature elevation one fifth of the strongest hydrogen bond formations at 270 K are at least temporarily broken, on the other hand especially short time hydrogen bonds of 10% to 30 % occupancy increase up to 50%. Encapsulated a temperature

increase causes not a complete separation of a former hydrogen bonded atom pair but reduces the duration of this particular hydrogen bond. During the extended non-bonded state of this explicit atom pair maybe additionally some alternative short term hydrogen bond formations can occur. But in total the structural dynamics within the hydrogen network of a protein is increased upon warming up at least in the here probed physiological relevant temperature range.

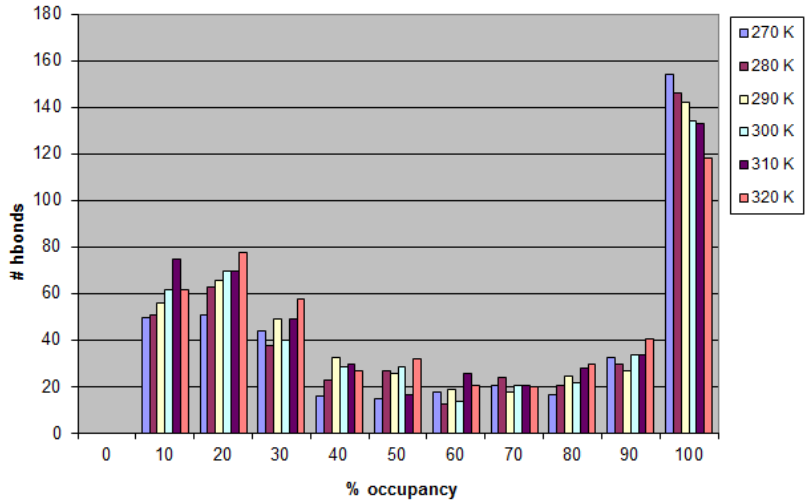

Fig. 4.69: Number of hydrogen bonds versus occupancy time for a 0.5 ps time step averaged over the last 2 ns. Here the results of one simulation set for the temperature range from 270 K to 320 K is presented. The total number of hydrogen bonds as well as hydrogen bonds with shorter occupancy times of 10 % to 30 % and of 80 % and of 90 % respectively are unambiguously increased by temperature elevation (dotted areas), whereas especially the continuously twinned hydrogen bonds (black arrow) start a temperature induced dynamical behavior of breakage and formation of hydrogen bonds resulting in the observed decrease for 100% occupancy.

We want to link up these simulation results and our interpretation with the theory chapter 2.2.3. where akin microscopic behavior was discussed. Both sections about biological water and liquid water offer possible touchpoints to the current discussion for which we will detail the latter.

In liquid water the perfect tetrahedron, usually formed in the icy state, is suppressed. The average coordination number, representing the next neighbors which are hydrogen bonded, is between four and five, hence larger than the coordination number of four in a perfect tetrahedron. In the case of liquid water, also upon a temperature elevation like in our measurements and the corresponding simulations, the coordination number and respectively the number of hydrogen bonds, is increased, finally not leading to a less but to an enhanced dynamical system. Upon the transition from ice to water, as an entail of some, thermally

triggered, initially broken hydrogen bonds within the system, some kind of friction for the adjacent tetrahedrons is caused by these “disturbing” water molecules offering an additional hydrogen bond site. The same kind of friction was involved in the discussion about publications which found indications for a spatially dependence of biological water diffusion constants.

Maybe a similar implication holds for proteins. Broken hydrogen and oxygen atom pairs offer at least temporarily new possibilities for hydrogen bond formations at higher temperatures. In the most extreme case a former hydrogen bonded, rigid structure becomes dynamic afterwards switching between different possible hydrogen bond arrangements. In the dense packed protein interior these new dynamical centers trigger chain reactions in their surrounding, affecting all kinds of bonds and their affiliated dynamics, from conformation determining torsional rotations of covalent bonds (e.g. of the backbone) up to steric hindrance of sidegroups, and in particular the hydrogen bond related stability of large scale secondary structures. If the structural dynamics is increased, the correlation, mediated by all these kinds of bonds, is reduced, leading to less collective type of motions and less induced dipoles. The latter is also determined by random fluctuations but correlated dynamical behavior outnumbers the random effect by far since extreme deviations from the mean become statistically more likely. The temperature caused effects on the number of collective motions, on the strength of permanent and induced dipoles, link up to the consequences for macroscopic properties, in particular the far-infrared absorption of proteins which is discussed in the following section.

4.6.4.4. Macroscopic Effects upon Temperature Variations: Consequences on the Far-Infrared Spectra of Proteins and the Model for Energy Related Correlation Effects in Proteins (MERCEP)

Two major hypotheses for the cause of the frequency independent temperature effect have been further investigated in the above accompanying molecular dynamics simulations: A broad band reduction of observed infrared active motions, which are mainly in the here relevant frequency range motions along single hydrogen bonds up to collective type of motions, by a detailed analysis of the temperature triggered hydrogen bond dynamics. And, not completely separated from the former, a reduction of the dipole moment by focussing on temperature caused consequences at typical van der Waals distances which are the main contribution to induced dipole momenta in proteins. Both hypotheses will be supported by additional arguments before deriving the MERCEP model below.
Besides our above deduced onset of an enhanced dynamical behavior of hydrogen bonds also theoretical assumptions support the temporary breakage of hydrogen bonds triggered by our experimentally applied temperature increase. Using the hydrogen bond energy range in proteins presented within theory section 2.1.2.2. from 2 kJ/mol to 21 kJ/mol [Des91, Ber02] and a typical average value of 4.2 kJ/mol [Mye96] our increase of provided thermal energy from 2.3 kJ/mol to 2.6 kJ/mol affects at least some hydrogen bonds of the lower energy scale. This coincides well with the simulation results where just the dynamical behavior of some hydrogen bonds is altered. Additionally we want to extend these theoretical considerations to the surrounding water molecules. We want to point out that these biological water molecules do not yield a large contribution to the far-infrared absorption as already discussed in the hydration dependence chapter but their influence on this internal protein dynamics concerning effect has been emphasized in chapter 2.2. In fig. 4.70 energies of hydrogen bonds between myoglobin and the surrounding water molecules are presented as derived from molecular dynamics simulations focussing on protein hydration. The majority of hydrogen bonds is situated in an energy range below 4.5 kcal/mol corresponding to 18.8 kJ/mol [Gu95]. About 18% of all hydrogen bonds, with an energy below 1.0 kcal/mol, are certainly affected by our temperature elevation indicating that also the dynamics within the biological water shell around a protein is increased, likely in a similar matter as internal protein hydrogen bonds as the above simulation results suggest.

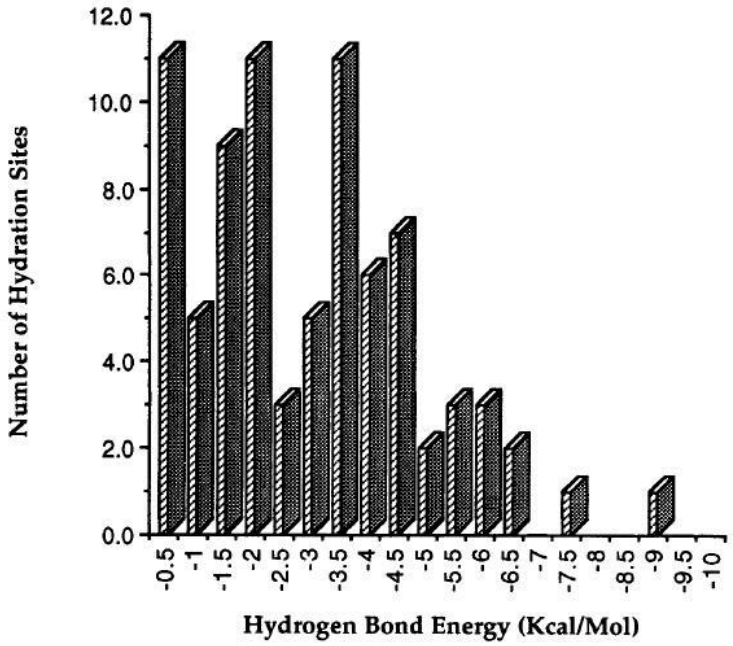

Fig. 4.70: Energy scale of hydrogen bonds formed between myoglobin and surrounding water molecules [Gu95]. 16 of 89 hydrogen bonds are in an energy range below 1kcal/mol which is likely affected by our temperature elevation corresponding to 0.55 kcal/mol and 0.62 kcal/mol respectively. In addition to the internal protein hydrogen bonds probed in our simulations this result indicates also a breakage of hydrogen bonds formed with surrounding water molecules likely to cause a similar increase in structural dynamics at the protein surface as well.

In the van der Waals bond distance there have been just indications but no sound proofs from the simulation results for a breakage or a distance increase by temperature elevation. To support our model we would like to refer to X-ray experiments at different temperatures on Ribonuclease A [Til92]. The number of contacts within a radius of 4.5 Å around a protein atom decreased almost 20 % upon a temperature increase from 98 K to 300 K. This temperature difference is about six times larger than applied in our experiments but it underlines the impact of temperature on atom pairs within a van der Waals distance. We want to concentrate here on the aspect of van der Waals bonded, hence correlated, atom pairs before discussing the induced dipole effect below. If the number of atoms within a 4.5 Å sphere around any protein atom is significantly reduced then even a smaller temperature increase as applied in our experiments is most probably capable to alter or break these kinds of bonds frequently. This is also supplemented by our statements in theory section 2.1.2.3, that this kind of bonds with an energy scale ranging from 0.4 kJ/mol to 4.2 kJ/mol is likely broken by the provided thermal energy ranging between 2.3 kJ/mol and 2.6 kJ/mol, representing in total a 13 % energy increase during our temperature dependent experiments.

Combined with the discussion about hydrogen bonds the first aspect of the MERCEP model can be now formulated in its final version. There is a broad band reduction in number of, in particular far-infrared light absorbing, motions upon temperature elevation. Our following elucidation will be structured from short to long range temperature effects as they apply for every randomly chosen atom within the protein. As the most close event a hydrogen bond to an adjacent atom might be broken at the higher temperature level resulting directly in a loss of the far-infrared absorption of this particular hydrogen bond vibration and rotation. Due to the very individual surrounding of each hydrogen bond their accompanying center frequencies in the spectra are not at all located at the same position but are scattered over a broader frequency range. The coupling to other atoms within the dense packed protein additionally broadens every single absorption peak. Encapsulated, the breakage of hydrogen bonds will lead more to a broad band than a located absorption decrease extending over the complete frequency range where hydrogen motions are observed. As we have already stated about hole burning experiments in proteins in the energy landscape section 2.2.2, anyway a huge number of states is probed by the proteins during one measurement due to the very short timescale of these motions. The loss of some possible states in some of the many thousand proteins within the sample will cause hence a more general lowering of the spectra because absorption can be treated as an average macroscopic property in our far-infrared frequency range by comparing the relevant timescales of our measurement to the occurring protein dynamics in chapter 2.2.1.3.

But the major effect on the hydrogen network within a protein is an increased dynamics of hydrogen bond breaking and formation, exemplary visualized by possible O···H atom pairs within an α-helix of β-lactamase in fig. 4.71. Just for one single oxygen atom three hydrogen atoms are in the vicinity of a formation of a hydrogen bond underlining the large number of atom pairs within a protein which are affected by a dynamical hydrogen bond arrangement, entailing in total a sizable temperature caused effect. Elevated temperatures favor the formation of just temporally existing hydrogen bonds with several of these offered hydrogen atoms to overcome the before discussed “friction” within the system. The dynamics of the neighboring atoms will permanently alter the conditions for these atoms resulting in continuously switching hydrogen bond arrangements. Also the momentary formation will again influence the hydrogen bond network of other neighboring atoms leading to an outspread of this temperature triggered effect causing an enhanced hydrogen bond dynamics in the whole protein. Important are the above simulation results which reveal that in total the temporal average of all hydrogen bonds is lower at elevated temperature. Hence the

correlation with neighboring atoms becomes less, weakening also collective motions whose concerted movement is partially mediated by hydrogen bonds and, on a larger scale, the stability of secondary structure units like α-helical structures as seen in fig. 4.71 for instance. All these collective motions are theoretically predicted to contribute to the far-infrared absorption. Again using the same arguments as above, their absorption will be broad and hence a loss of these motions will lead to a general absorption decrease.
We want to emphasize in an interjection that such conclusions from the well investigated mid-infrared measurements on proteins are difficult since in that frequency range hydrogen bonds can be just indirectly probed by their influence on the center frequency of covalent motions as stated in 2.1.2.2. Additionally aggravating an analysis of the temperature caused impact on mid-infrared spectra are contrary frequency shifts. The A-H stretching vibrations of an A-H···B hydrogen bonded triple are shifted towards higher frequencies but the A-H deformation modes are shifted towards lower frequencies upon hydrogen bond breakage, caused by the increased and decreased force constant of these motions respectively [Vin71]. In a complex protein system leading anyway to broad amide absorption bands which usually originate from several vibrations of many different, hydrogen bonded as well as non-hydrogen bonded, covalent bonds, an interpretation of such contrary temperature effects from mid-infrared measurements is hard. But the hydrogen mode absorption itself is situated in the far-infrared region where the influence of temperature on this kind of bonds can be hence probes directly [Vin71]. The same holds for the many atoms involving collective motions of which the mid-infrared range is almost unaffected, but far-infrared measurements are capable to directly investigate these motions and corresponding consequences caused by a temperature variation. These collective motions involve also van der Waals bonds which get less likely at larger atom pair distances which are assumed at elevated temperatures. This is the most outspread temperature effect which completes the above discussed reduction of hydrogen bond motions as well as collective motions and the consequences for their broad band far-infrared absorption.

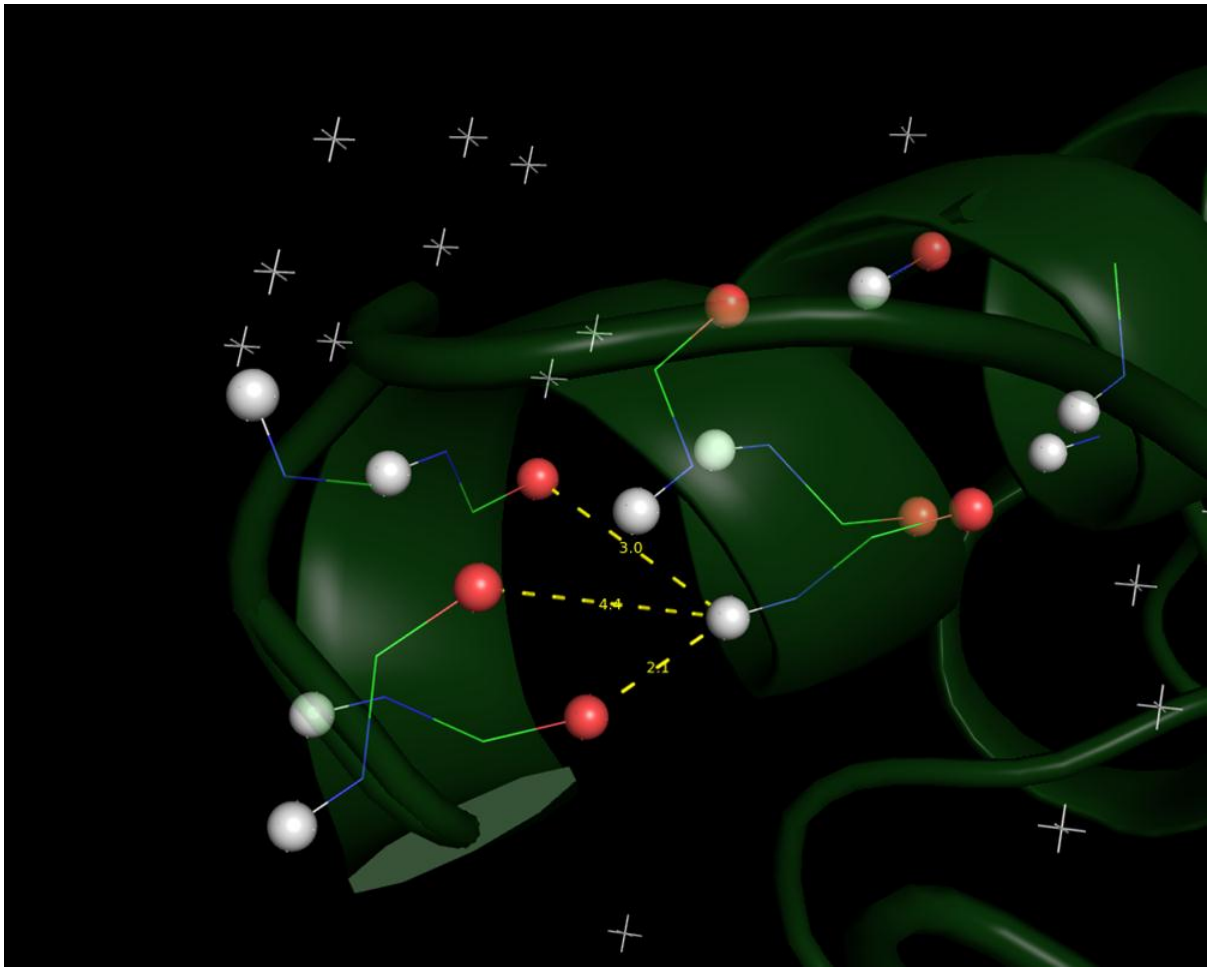

Fig. 4.71: Detail of an α-helical structure of β-lactamase (AA 194-251; PDB code 1M40). Highlighted are oxygen (white) and hydrogen atoms (red) as a random sample for the many different possibilities of hydrogen bond arrangements for every atom within the densely packed protein which is capable to form these kinds of bonds. In the MERCEP model temperature causes an enhancement of structural dynamics via a permanent formation and breakage of hydrogen bonds at elevated temperatures weakening correlated motions of this set of atoms.

The second aspect of the MERCEP model is a proposed decrease of dipole moment with increasing temperature. This effect cannot be completely separated from the above discussion constituting the main reason for uniting both aspects under the effect of correlation below. For instance, the breakage of hydrogen bonds directly reduces the permanent dipole moment and upon a distance increase of the former hydrogen bonded atom pair the induced dipole moment is altered, too (chapter 2.1.2.2). But one facet which is not included in the above reduction of motions is that the dipole moment determines the strength of the absorption of these motions also called the infrared activity. Therefore we want to focus mainly the contribution of induced dipole momenta caused by van der Waals interactions because in this case the dipole

moment is altered in a way which is not already included in the above discussion treating the breakage of bonds and the increased hydrogen bond dynamics.

We already emphasized the importance of an inclusion of induced dipole momenta in modeling far-infrared protein spectra or the dynamics of water particularly in sections 2.3.6 and 2.1.2.3. which we will use together with the references therein as a basis for this discussion. First of all fig. 4.72 demonstrates just for the highlighted nitrogen and carbon atoms the very large number of atom pairs within the densely packed protein which are in typical van der Waals bond distances. Hence very small changes will lead to a large effect due to the enormous number of affected atom pairs as already pointed out in section 2.1.2.3. Distinguishing the effect from the above discussion, we will use typical van der Waals bond lengths as a measure for distances between atom pairs where induced dipole momenta, caused by fluctuations within the electron density around atoms, still play a significant role. These fluctuations lead to a dipole induced dipole interaction between every atom pair. We will also use the r^{-6} relation of the van der Waals bond energy (formula 2.2) to estimate the loss of induced dipole moment between each atom pair which obeys the same distance dependence as the bond itself since it is the cause for the attractive force within the van der Waals equation. Molecular dynamics simulation also directly use the Lennard-Jones potential for modeling the induced dipole moment [e.g. Dan97]. We are aware that there is no clear dividing line between a van der Waals bond and a temporary occurring induced dipole effect of one atom on another especially in a dynamic system like a protein, but in our discussion we would like to separate the induced dipole effect on the absorption spectra from the above treated complete loss of some collective motions which are also stabilized by van der Waals bonds. The packing density of atoms, hence the distance between atoms, is the major component for the (induced) dipole moment as an already cited comparison of water in the gas to the liquid phase demonstrates [Yu05]. Temperature influences the internal protein distances as the extreme example of a protein denaturation, resulting in a partially entanglement of the protein, reveals. To distinguish the effect from the above discussion we hypothetically assume a (collective) motion which still remains in the same manner after all participating atoms just increased their distance. Infrared absorption scales with the change of dipole moment during the motion [Mir04] but under the above prerequisite it even scales with the dipole moment itself now alleviating the below discussion by our chosen separation into two certain aspects of the model. In chapter 4.6.3. the for this calculation relevant volume expansion coefficient, as a measure of the in all three dimensions increasing distance, was found to be between $5*10^{-4}$/K and $6*10^{-4}$/K around 25°C derived from measurements on six different proteins [Hin94,

Cor02]. For an estimation of the temperature effect on the induced dipole moment of one atom on another, we will use here the lower value of an expansion coefficient of $5*10^{-4}$/K for all atom pairs which represents of cause a simplification of a complex 3-dimensional expansion of a covalent, hydrogen and van der Waals bonded atom assembly. But we will just consider the expansion between non-covalent and hydrogen bonded atom pairs of which some will move less and some will move more further away in reality. Hence on average, the complete volume expansion coefficient for a protein will be a suitable guideline in particular for these kind of atom pairs to derive an estimation for the order of magnitude of this temperature effect because covalent bonded atom pairs, included in the over all expansion, can be safely assumed to exhibit no bond length increase as already stated above, increasing the actual thermal expansion coefficients of the atom pairs of interest beyond average. Our about 34 K temperature elevation causes hence a 1.7 % distance increase. Illustrative our temperature effect is estimated for a bond length of 4 Å which is well above typically applied cutoff ranges for hydrogen bonds at a distance of 3.5 Å and therefore in a typical van der Waals interaction distance (see section 2.1.2.2.). We want to mention here that van der Waals bonds of shorter distance also occur if non-hydrogen atoms are involved or steric restrictions prevent a hydrogen bond formation for instance. Applying formula 2.2 under the above assumptions that the induced dipole moment as the major cause for the attractive van der Waals interaction directly scales with the van der Waals bond energy and that all other parameters in the equation besides the distance remain unchanged, the induced dipole moment present at 277 K is reduced about 10.6 % by the 34 K temperature increase which is, as only one certain aspect of the model, already in the same order of magnitude as the observed absorption decrease. The majority of protein atoms and their motions will be mainly infrared active due to such an induced dipole moment effect, since for instance in average just about one hydrogen bond per residue exists in proteins (see chapter 2.1.2.3), whose number is hence at least one order of magnitude less than existing van der Waals atom pairs within a protein. Therefore the above scaling of the effect to the over all protein absorption can be justified. In addition, the permanent dipole moment of the protein is reduced because the simulation results indicate a decreased hydrogen bond formation at least in the, for absorption relevant, occupancy time. We claim in our MERCEP model that the dipole moment of hydrogen bond motions and collective motions are both reduced by a temperature increase. The observed broad far-infrared absorption features get hence less pronounced due to a weaker coupling to infrared light. Since a very large number, likely the majority, of motions are affected by such a dipole moment reduction the consequence of a broad band absorption decrease is in line

with our experimental results for whose explanation the model was evolved. Summarized we believe that a reduction of dipole moment upon a temperature increase, with their two contributions from permanent and induced dipoles, leads to a frequency independent absorption decrease as observed in the experiments. This concerns not only atom pairs or groups whose hydrogen or van der Waals bonds are broken, considered predominately above in the first aspect of the model, but still persisting motions of atom pairs and groups at elevated temperature which exhibit a decreased far-infrared absorption as a consequence of a less dipole moment change during an attributed motion.

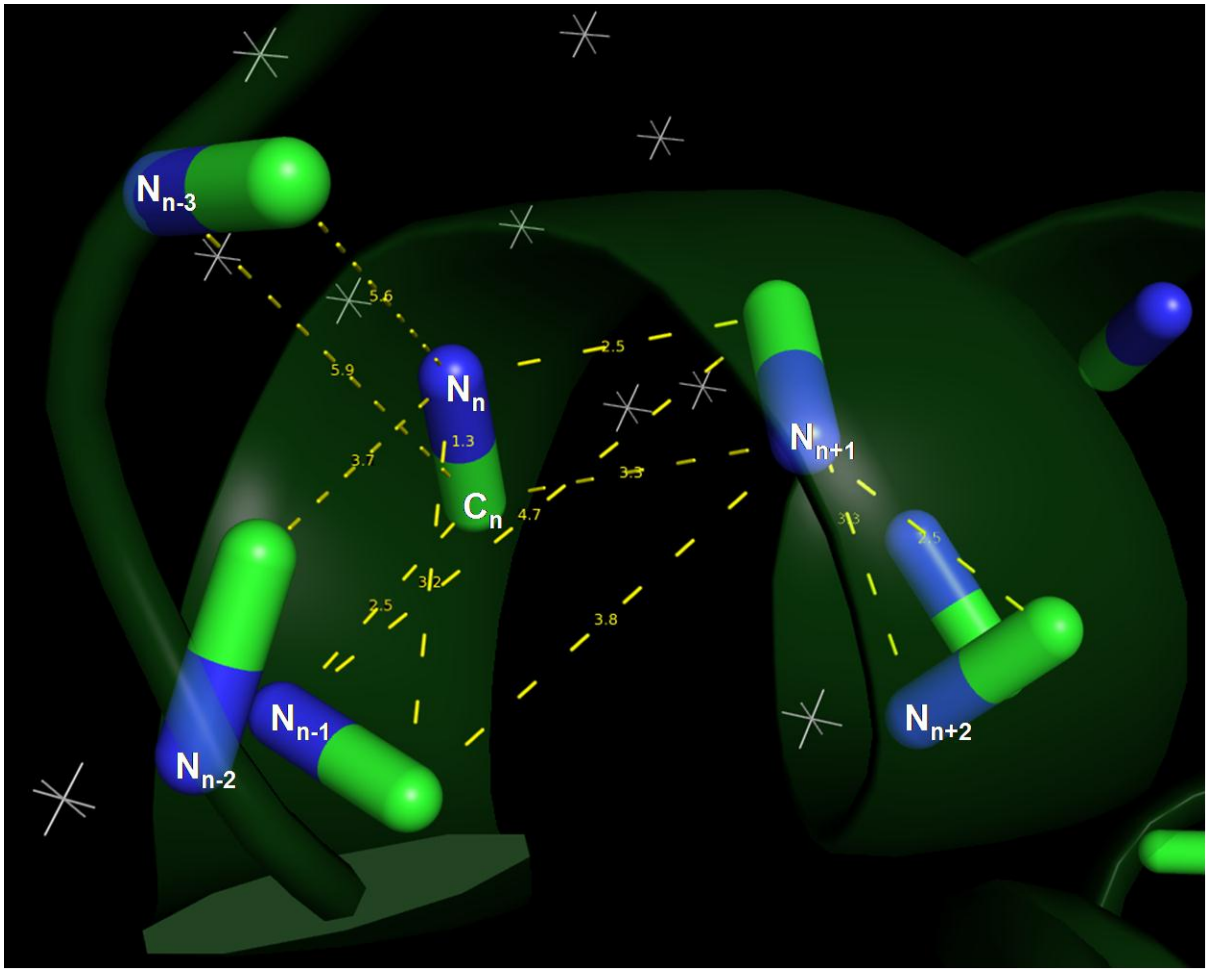

Fig. 4.72: Carbon and nitrogen atom distances within a 6 Å cutoff in an α-helix of β-lactamase (AA 194-251, PDB code 1M40). The given example emphasizes the abundance of non-covalent and non-hydrogen bonded atom pairs within a distance where van der Waals interaction are present. Since here just two sorts of atoms are highlighted and van der Waals interactions still influence atoms beyond the chosen typical cutoff range of 6 Å, which is there of course much less effective due to the r^{-6} distance dependence, the actual number of atom pairs is even much larger than suggested in this picture. In the MERCEP model we claim that temperature causes an increase of both, dynamics and distance within the system, reducing the correlation and

hence the (induced) dipole moment within every sector of the protein. Due to the huge number of affected atom pairs within the densely packed protein, even small temperature triggered effects add up to considerable consequences on macroscopic properties like the far-infrared protein absorption.

Since both aspects described above cannot be clearly separated from each other, especially concerning van der Waals bonds in the first and induced dipole interactions in the latter or the complete loss of collective motions versus just less dipole moment changes during collective motions respectively, it is useful to unite both aspects by a model which depends on an underlying joint, temperature influenced property. As we will see correlation offers such a superior framework to which both aspects can subsummized. The validity of the MERCEP model will be not affected by this generalization but we want to condense the above abundance of contributions to a catchy description without the necessity of elucidating several, maybe overlapping, aspects as above. By the way the model can be brought to a level where the correlation altering parameter must not be defined and hence not necessarily be the here applied temperature variation.

Interactions between protein atoms are involved in all above described facets of the temperature impact on the protein. Atom pairs or groups note the presence of each other as well as occurring variations which are caused by an extensive protein dynamics at physiological temperatures. Motions are often not random within proteins but strongly influenced by each other revealing a correlation on several length scales, starting from closely bonded atom pairs, over weakly affiliated atom pairs at van der Waals bond distances up to large secondary structure units performing functionally most relevant motions as our examples of lysozyme and myoglobin at the beginning of the theory chapter demonstrate. There we emphasized the importance of collective motions for protein functionality giving also a justification for our decision to perform experiments in particular in this far-infrared frequency range where these motions as the most intriguing example of correlated motions occur. The present correlation is enlarged beyond the covalent and hydrogen bond network by the dense packing of the protein in the folded state enabling additional van der Waals interactions, hence representing some kind of principal favorite energy levels in the continuously changing energy seascape of a protein where randomizing dynamics and inevitable resulting friction, again caused by the dense packing and even mediated by the same bonds as the above correlation, constitute the antagonist. Changing the provided energy for protein dynamics intervenes this balanced interplay determining the extend of correlation on all length scales. Hence the name Model for Energy Related Correlation Effects in Proteins

was chosen enclosing all above aspects which are necessary to explain our temperature dependent experimental results. Temperature is the most likely but not the only way to alter the correlation determining energy within the protein. The breakage of bonds by optical excitation can also provide additional energy to the system by triggering enhanced dynamical behavior and hence a loss of correlation within the protein, even a change of the surrounding solvent properties might be able to alter the energy seascape of the protein and hence the correlation, to provide another, maybe far-fetched, example as an outlook to which further derivations from the model might be employed sometime.

5. Summary

This thesis has been focussed on experiments, simulations as well as theoretical issues related to the far-infrared properties of proteins of which we will summarize here our most important results that have been derived in the preceding chapters.

Starting with the static structure of proteins we deduced that our experimentally applied temperature variation is able to affect hydrogen and van der Waals bonds already justifying the later discussion of the temperature dependent experiments along these two kinds of bonds. Their role for protein folding was emphasized and the evolving dense packing of proteins favours the development of a significant contribution of induced dipole momenta in proteins to which we came back in the derivation of the MERCEP model. At the passage from a static to the dynamical treatment of proteins we placed a section about protein folding as the first dynamical behavior of proteins in nature after being assembled at the ribosomes, also introducing the dynamics determining energy landscape of proteins with the lack of a defined minimum. For protein dynamics in the natural state we pointed out that the far-infrared frequency range is assumed to bear information about functional relevant motions in proteins namely hydrogen bond vibrations and rotations up to collective motions of several atoms together in a concerted way. Doster's conception of a tram as a model for protein dynamics was extended and even used for an illustrative description of the dissipation-fluctuation theorem. Afterwards we added anharmonic coupling as a further modulation factor to the most recent description of the energy potential which is the so called energy seascape of proteins.

One of the most influencing parameters on protein dynamics is the surrounding solvent usually indispensable for functionality and also indispensable for an understanding of proteins in their natural form. Discussing different types of water molecules in the protein-water system, focussing on biological water in the vicinity of the protein, we detailed especially the dynamical behavior and the affiliated explanations for enhanced diffusion, friction and ice like behavior, all of them reappearing in the sections about our hydration dependent measurements and the derivation of the MERCEP model.

In the following discussion about the boson peak and the glass transition, associated often with protein features, we derived for both an argumentation for resolving the dissent in the present scientific discussion. In fact, the sometimes observed low frequency peak in proteins occurs in a similar frequency range like the boson peak in glasses, but the deviating

temperature and hydration dependence as well as the presence of the peak above the dynamical transition are excluding that the same physics applies for both peaks. Hence the name boson peak is misleading for an assignment of this low frequency peak in proteins. In addition we recommended that the designation “glass transition” in proteins is better to be entitled a “dynamical transition” as the significantly differing heat capacity slope of proteins compared to glass formers proofs.

Since the way to calculate protein far-infrared absorption spectra is probably just known to a few scientists and a standard approach is missing up to date, we presented a thorough way to model far-infrared spectra based on dipole moment correlation functions by a derivation applying directly the dissipation-fluctuation theorem in Kubo’s quantum mechanical formulation. Therewith we hope to encourage accompanying calculations to our measured protein spectra in the future. At the end of the theory chapter we derived our own version for calculating the population of hydration shells based on the Sorption Isotherm Equation. For the latter we included also a correction of a confusion causing typo present in the original publication.

For the measurements five different setups for two types of samples, hydrated protein films and protein solutions, have been developed through several improvement steps also applying different window materials to suit certain frequency ranges. Covering with all setups the frequency range from the THz up to the mid-infrared, the far-infrared atmosphere cell was the predominant employed setup due to the importance of this particular sought-after frequency range and due to the most eminent gained results. In the thesis we additionally presented far-infrared measurements of protein solutions in the liquid cell which were in line with published measurements in the adjacent THz frequency range concerning the similar shape of absorption spectra as pure liquid water, but unfortunately we were not able to confirm the expected effect on the absolute value of absorption which has to be assumed to lie within our error bars for this particular setup.

Advantageous about our atmosphere cell setup is that external hydration of protein films reduces the three component system protein, biological water, and liquid water, present in the above described liquid cell measurements, to a two component system of protein and attached biological water. In connection with the knowledge of the hydration dependence of the spectra in which solely the amount of biological water is varied, this enables us to apply a different, conclusive analyzing procedure upon measurements of proteins in a close to natural state since the number of attached water molecules is adequate to simulate the typical aqueous surrounding. The major experimental breakthrough to gain these far-infrared measurements

on proteins has been the development of a free standing protein film preparation technique allowing us an improved design of the atmosphere cell which overcame many experimental obstacles and additionally permitting a faster and complete hydration of the protein film. The finally achieved reproducibility has been confirmed by several test measurements probing for instance the sample quality and the precision of the water vapor background removement. Reproducibility has been also a major topic in the review of spectroscopy on proteins which was given as overview of the up to date research in this field. Between the well known mid-infrared spectroscopy which probes primarily covalent bond vibrations and the unfortunately featureless but at the moment fashionable THz frequency range, we demonstrated that the far-infrared offers information about hydrogen bond and collective motions causing broad but distinctive absorption peaks in the protein spectra. For some proteins these spectra have been unknown before and for all measured proteins these have been the first measurements in the far-infrared, under defined conditions of both influencing parameters temperature and hydration as well as for their parameter dependence itself. By fitting the spectra with several Lorentzian peaks the center frequencies of all far-infrared features were assigned which can be used in future for a direct comparison to protein motion calculations. Also some absorption bands besides the known two amide bands, the latter are situated at both edges of the measured frequency range, have been found and assigned for the first time. The comparison between the four measured proteins β-Lactoglobulin from bovine milk, Bovine Serum Albumin, Hen Egg White Lysozyme, and Trypsin Inhibitor from hen egg white, revealed protein specific as well as presumably unspecific absorption bands within the measured frequency range.

With our atmosphere cell we are able to probe directly the influence of the attached water molecules on the far-infrared spectra of the protein-water system by varying the external humidity and hence solely the amount of water molecules. Our hydration dependent measurements revealed a small impact of this parameter on the protein far-infrared absorption. As we pointed out in the discussion, our results are in line with those measurements in the adjacent THz frequency range which were not additionally recalculated and hence disguising the absence of an effect. Maybe the expectation of a large effect by adding strongly absorbing water molecules to the protein-water system was misleading in those publications which overlooked that liquid but not necessarily biological water molecules are strongly absorbing in this frequency range and that the absence of a large effect is an even more intriguing result. By our developed formalism to calculate not only the attached water molecules at each protein but also the absolute number of water molecules in total within our

measured spot, we were able to deduce that biological water at least absorbs one order of magnitude less than liquid water in the far-infrared frequency range. Supported by a comparison to the far-infrared spectra changes of water upon a transition from the liquid to the icy state, our results can be seen as a sound spectroscopic argument for the iceberg model describing dynamical restricted biological water molecules, joining hence a microscopic model about single molecule dynamical behavior, based on simulations and structure resolving experiments, with our macroscopic property which averages over the dynamical behavior of an assembly of molecules.

The consequences of temperature variation on the measured protein spectra as the second independently changeable parameter within the atmosphere cell, can be split into three different effects. Underneath the major frequency independent absorption decrease, two frequency dependent temperature effects can be identified.

On the low frequency side of all measured protein spectra a broad absorption peak can be found which, except for β-lactoglobulin, shifts towards lower frequencies upon a temperature increase. We elucidated this shift by thermally triggered changes within the population of energy levels. If transitions between higher excited states at elevated temperatures are probed, revealing smaller energy differences than between lower states, a resulting shift of the absorption occurs towards lower frequencies as it is observed in the measurements. Probing excited states of proteins in the far-infrared and the THz frequency range has to our best knowledge never been mentioned before in any publication, even if room temperature corresponds to a frequency of 206.5 cm^{-1} and the importance of a consideration in the analysis around and below this frequency is evident. We want to emphasize this especially for THz measurements which are much more affected than our far-infrared experiments for which the levelling off above the temperature corresponding frequency is experimentally observed just as theoretically expected. For the different behavior of the β-lactoglobulin spectra, which also reveals in contradiction to the other proteins two distinguishable low frequency peaks, we can just suspect the larger β-sheet content to be the origin since for instance the size of the protein is just average among the other proteins. In the upper half of the spectra a different temperature effect can be observed, this time also for β-lactoglobulin, which is our best indicator for the presence of a separate temperature effect from the one above. Stronger and less temperature affected regions alternate and a relation to the protein backbone for the latter region can be assumed since this particular frequency range coincides with the position of a common absorption band in all measured proteins.

For the major frequency independent temperature effect on the protein spectra we observed non-linear temperature dependence. Split by the 24° measurement, in one temperature region the reduction of absorption was always more pronounced than in the other one. For our measured proteins these temperature regions were divided by the originating specie of distinctive natural body temperatures maybe indicating different functional temperature ranges of these proteins. Since there have been no models present to explain the large broadband absorption decrease upon a rise of temperature, we developed our own Model of Energy Related Correlation Effects in Proteins (MERCEP) which we substantiated by accompanying molecular dynamics simulations probing certain aspects of this model. We claim that temperature elevation causes a correlation decrease within the protein, reducing absorption strength relevant, predominately induced, dipole momenta as well as infrared active (collective) motions to be observed in this particular frequency region. For the first component there are indications for a slight distance increase especially between non-bonded atom pairs or groups which are in typical van der Waals bond distance, hence within the scope for vice versa inducing dipole momenta. Without necessarily altering the underlying dynamics of these compounded atoms, the absorption strength is reduced if a less change of dipole moment occurs during these motions. This distance increase has also considerable consequences on the second component, since in the far-infrared frequency range assumed collective motions are mediated by these van der Waals interactions as well as, also temperature sensitive, hydrogen bonds. The correlated motion of several atoms and their far-infrared absorption likely disappears if the connecting van der Waals or hydrogen bonds are at least temporarily broken. Especially for the temperature effect on hydrogen bonds the accompanying molecular dynamics simulations revealed an intriguing nature. Upon a temperature increase an enhanced dynamics of breakage and formation of hydrogen bonds sets in, even resulting in an increase in the total number of formed hydrogen bonds within a certain time step. But the for our measurements relevant temporal average of hydrogen bonding is reduced and an overall less stable protein arrangement is present at elevated temperatures. The enhanced dynamics reduces the above mentioned correlation between atom pairs and groups, since hydrogen bonds are elementary for the stabilization of small protein sections like sidegroups up to complete secondary structure elements like α-helices and β-sheets. In addition the far-infrared absorption of hydrogen bond vibrations and rotations itself can disappear upon a temperature triggered, temporary or permanent, breakage of hydrogen bonds. Since all the above mentioned absorption reduction effects are related to motions taking place in individual and hence differing surroundings, determined by the continuously

altering energy seascapes of proteins, our model predicts a broad absorption decrease in the same order of magnitude as it is observed in our temperature dependent measurements.
As the summary of the last result reveals, often just a synopsis of experiments, simulations, and theory is able to achieve progress towards revealing once the real, complex nature of proteins. By not meeting solely an experimental challenge during this thesis to perform just far-infrared measurements on proteins but also getting familiar with affiliated topics concerning simulations and theory, and by finally bringing together these three, usually separated, fields of research, it was possible to get further insight into many different, far-infrared frequency range related, protein properties, therefore hoping to encourage further, also interdisciplinary aspects concerning, research on proteins in this interesting and promising far-infrared frequency range.

6. Literature

[Abs96]: R. Abseher, H. Schreiber, O. Steinhauser, *The Influence of a Protein on Water Dynamics in Its Vicinity Investigated by Molecular Dynamics Simulation*, Proteins: Structure, Function, and Genetics **25**, 366 (1996)

[Ach02]: K. Achterhold, C. Keppler, A. Ostermann, U. van Bürck, W. Sturhahn, E. E. Alp, F. G. Parak, *Vibrational dynamics of myoglobin determined by the phonon-assisted Mössbauer effect*, Physical Review E **65**, 051916 (2002)

[Adc06]: S. A. Adcock, J. A. McCammon, *Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins*, Chem. Rev. **106**, 1589 (2006)

[Afs77]: M. N. Afsar, J. B. Hasted, *Measurements of the optical constants of liquid H20 and D20 between 6 and 450 cm*$^{-1}$, J. Opt. Soc. Am. **67** (7), 902 (1977)

[Afs78]: M. N. Afsar, J. B. Hasted, *Submillimeter Wave Measurement of Optical Constants of Water at various Temperatures,* Infrared Physics **18**, 835 (1978)

[Agm96]: N. Agmon, *Tetrahedral Displacement: The Molecular Mechanism behind the Debye Relaxation in Water*, J. Phys. Chem. **100**, 1072 (1996)

[All95]: P. B. Allen, *Linear Response Theory and Kubo Formulas,* Troisieme Cycle de Physique en Suisse Romande, (1995)

[Ama93]: A. Amadei, A. B.M. Linssen, H. J.C. Berendsen, *Essential Dynamics of Proteins*, Proteins: Structure, Function, and Genetics **17**, 412 (1993)

[Anf73]: C. B. Anfinsen, *Principles that Govern the Folding of Protein Chains*, Science **181**, 223 (1973)

[Ang95]: C. A. Angell, *Formation of Glass from Liquids and Biopolymers*, Science **267** (5206), 1924 (1995)

[Ans85]: A. Ansari, J. Berendzen, S. F. Bowne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, R. D. Young, *Protein states and proteinquakes*, Proc. Natl. Acad. Sci. USA **82**, 5000 (1985)

[Ant89]: J. Antosiewicz, D. Porschke, *The nature of protein dipole moments: experimental and calculated permanent dipole of alpha-chymotrypsin*, Biochemistry **28** (26), 10072 (1989)

[Arc70]: R. L. D'Arcy, I. C. Watt, *Analysis of Sorption Isotherms of non-Homogeneous Sorbents*, Trans. Faraday Soc. **66**, 1236 (1970)

[Arr93]: J. L. R. Arrondo, A. Muga, J. Castresana, F. M. Goni, *Quantitative Studies of the Structure of Proteins in Solution by Fourier-Transform Infrared Spectroscopy*, Prog. Biophys. molec. Biol. **59**, 23 (1993)

[Art79]: P. J. Artymiuk, C. C. F. Blake, D. E. P. Grace, S. J. Oatley, D. C. Phillips, M. J. E. Sternberg, *Crystallographic studies of the dynamic properties of lysozyme*, Nature **280**, 563 (1979)

[Ary66]: Y. B. Aryeh, *An application of infra-red spectra to the electronic structure theory of the water molecule*, Proc. Phys. Soc. **89**, 1059 (1966)

[Ata79]: M Ataka, S. Tanaka, *Far-Infrared Spectrum of Crystalline Lysozyme*, Biopolymers **18**, 507 (1979)

[Aus03]: R. H. Austin, A. Xie, L. van der Meer, M. Shinn, G. Neil, *Self-trapped states in proteins*, Nuclear Instruments and Methods in Physics Research A **507**, 561 (2003)

[Aus05]: R. H. Austin, A. Xie, L. van der Meer, B. Redlich, P. Lindgard, H. Frauenfelder, D. Fu, *Picosecond Thermometer in the Amide I Band of Myoglobin*, PRL **94**, 128101 (2005)

[Bah97]: I. Bahar, A. R. Atilgan, B. Erman, *Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential*, Folding & Design **2** (3), 173 (1997)

[Bah98]: I. Bahar, A. R. Atilgan, M. C. Demirel, B. Erman, *Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Motions in Relation to Function and Stability*, PRL **80** (12), 2733 (1998)

[Bal96]: U. Balucani, J. P. Brodholt, R. Vallauri, *Analysis of the velocity autocorrelation function of water*, J. Phys.: Condens. Matter **8**. 6139 (1996)

[Bal08]: R. Balu, H. Zhang, E. Zukowski, J. Y. Chen, A. G. Markelz, S. K. Gregurick, *Terahertz Spectroscopy of Bacteriorhodopsin and Rhodopsin: Similarities and Differences*, Biophysical Journal **94**, 3217 (2008)

[Ban83]: J. Bandekar, L. Genzel, F. Kremer, L. Santo, *The temperature-dependence of the far-infrared spectra of L-alanine*, Spectrochimica Acta **39A** (4), 357 (1983)

[Ban92]: J. Bandekar, *Amide modes and protein conformation*, Biochimica et Biophysica Acta. **1120**, 123 (1992)

[Bay02]: C. Baysal, A. R. Atilgan, *Relaxation Kinetics and the Glassiness of Proteins: The Case of Bovine Pancreatic Trypsin Inhibitor*, Biophysical Journal **83**, 699 (2002)

[Bec03]: T. Becker, S. Fischer, F. Noe, A. L. Tournier, G. M. Ullmann, J. C. Smith, *Protein Dynamics: Glass Transition and Mechanical Function*, Adv. In Solid State Physics **43**, 57 (2003)

[Bel96]: M. C. Bellisient-Funel, J. M. Zanotti, S. H. Chen, *Slow dynamics of water molecules on the surface of a globular protein*, Faraday Discussions **103**, 281 (1996)

[Ben98]: R. E. Bentley, *Handbook of Temperature Measurement: Temperature and humidity measurement*, Springer, Singapore, (1998)

[Ber64]: J. E. Bertie, E. Whalley, *Infrared Spectra of Ices Ih and Ic in the Range 4000 to 350 cm^{-1}*, J. Chem. Phys. **40** (6), 1637 (1964)

[Ber67]: J. E. Bertie, E. Whalley, *Optical Spectra of Orientationally Disordered Crystalls. II. Infrared Spectrum of Ice Ih and Ice Ic from 360 to 50 cm^{-1}*, J. Chem. Phys. **46** (4), 1271 (1967)

[Ber80]: M. A. Berliner, *Feuchtemessung*, Verl. Technik, Berlin, (1980)

[Ber81a]: P. H. Berens, K. R. Wilson, *Molecular dynamics and spectra. I. Diatomic rotation and vibration*, J. Chem. Phys. **74** (9), 4872 (1981)

[Ber81b]: P. H. Berens, S. R. White, K. R. Wilson, *Molecular dynamics and spectra. I. Diatomic Raman*, J. Chem. Phys. **75** (2), 515 (1981)

[Ber83a]: P. H. Berens, D. H. J. Mackay, G. M. White, K. R. Wilson, *Thermodynamics and quantum corrections from molecular dynamics for liquid water*, J. Chem. Phys. **79** (5), 2375 (1983)

[Ber83b]: J. Berthou, A. Lifchitz, P. Artymiuk, P. Jolles, *An X-ray study of the physiological-temperature form of hen egg-white lysozyme at 2 A resolution*, Proc. R. Soc. Lond. B **217**, 471 (1983)

[Ber84]: H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *Molecular dynamics with coupling to an external bath*, J. Chem. Phys. **81**, 3684 (1984)

[Ber90]: J. Berendzen, D. Braunstein, *Temperature-derivative spectroscopy: A tool for protein dynamics*, Proc. Natl. Acad. Sci. USA **87**, 1 (1990)

[Ber96b]: J. E. Bertie, Z. Lan, *Infrared Intensities of Liquids XX: The Intensity of the OH Stretching Band of Liquid Water Revisited, and the Best Current Values of the Optical Constants of $H_20(I)$ at 25 °C between 15,000 and 1 cm^{-1}*, Appl. Spectroscopy **50** (8), 1047 (1996)

[Ber03]: J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, Freeman and Company, New York, (2003)

[Ber05]: A. Bergner, U. Heugen, E. Bründermann, G. Schwaab, M. Havenith, D. R. Chamberlin, E. E. Haller, *New p-Ge THz laser spectrometer for the study of solutions: THz absorption spectroscopy of water*, Review of Scientific Instruments **76**, 063110 (2005)

[Ber06]: C. Berthomieu, L. Marboutin, F. Dupeyrat, P. Bouyer, *Electrochemically Induced FTIR Difference Spectroscopy in the Mid- to Far Infrared (200 µm) Domain: A New Setup for the Analysis of Metal–Ligand Interactions in Redox Proteins*, Biopolymers **82**, 363 (2006)

[Biz95]: A. R. Bizzarri, C. X. Wang, W. Z. Chen, S. Cannistraro, *Hydrogen bond analysis by MD simulation of copper plastocyanin at different hydration levels*, Chemical Physics **201**, 463 (1995)

[Biz96]: A. R. Bizzarri, S. Cannistraro, *Molecular dynamics simulation evidence of anomalous diffusion of protein hydration water*, Physical Review E **53** (4), 53 (1996)

[Biz00]: A. R. Bizzarri, A. Paciaroni, S. Cannistraro, *Glasslike dynamical behavior of the plastocyanin hydration water*, Physical Review E **62** (3), 3991 (2000)

[Biz02]: A. R. Bizzarri, S. Cannistraro, *Molecular Dynamics of Water at the Protein-Solvent Interface*, J. Phys. Chem. B **106**, 6617 (2002)

[Bla83]: C. C. F. Blake, W. C. A. Pulford, P. J. Artymiuk, *X-ray Studies of Water in Crystals of Lysozyme*, J. Mol. Biol. **167**, 693 (1983)

[Bol89]: B. A. Bolton, J. R. Scherer, *Raman spectra and water absorption of bovine serum albumin*, J. Phys. Chem. **93** (22), 7635 (1989)

[Bon77]: S. Bone, P. R. C. Gascoyne, R. Pethig, *Dielectric Properties of Hydrated Proteins at 9.9 GHz*, J. Chem. Soc., Faraday Trans. **1**, 1605 (1977)

[Bon96]: S. Bone, *Dielectric and gravimetric studies of water binding to Lysozyme*, Phys. Med. Biol. **41**, 1265 (1996)

[Bon08]: S. Bone, *Structural Flexibility in Hydrated Proteins*, J. Phys. Chem. B **112** (32), 10071 (2008)

[Bop98]: P. A. Bopp, A. A. Kornyshev, G. Sutmann, *Frequency and wave-vector dependent dielectric function of water: Collective modes and relaxation spectra*, J. Chem. Phys. **109** (5), 1939 (1998)

[Bor94]: D. Bordo, P. Argos, *The Role of Side-chain Hydrogen Bonds in the Formation and Stabilization of Secondary Structure in Soluble Proteins*, J. Mol. Biol. **243**, 504 (1994)

[Boy75]: T. H. Boyer, *Temperature dependence of Van der Waals forces in classical electrodynamics with classical electromagnetic zero-point radiation*, Physical Review A **11** (5), 1650 (1975)

[Bra80]: S. Bratos (Ed.), *Vibrational spectroscopy of molecular liquids and solids*, Plenum Press, New York, (1980)

[Bri48]: F. E. M. O'Brien, *The Control of Humidity by Saturated Salt Solutions*, J. Sci. Instrum. **25**, 73 (1948)

[Bro72]: K. G. Brown, S. C. Erfurth, E. W. Small, W. L. Peticolas, *Conformationally Dependent Low-Frequency Motions of Proteins by Laser Raman Spectroscopy*, Proc. Nat. Acad. Sci. USA **60**, No. 6, 1467 (1972)

[Bro83]: B. Brooks, M. Karplus, *Harmonic Dynamics of Proteins: Normal Modes and Fluctuations in Bovine Pancreatic Trypsin Inhibitor*, PNAS **80** (21), 6571 (1983)

[Bro85]: B. Brooks, M. Karplus, *Normal Modes for Specific Motions of Macromolecules: Application to the Hinge-Bending Mode of Lysozyme*, PNAS **82** (15), 4995 (1985)

[Bro88]: C. L. Brooks, M. Karplus, B. M. Pettitt; Proteins*: A theoretical perspective of dynamics, structure, and thermodynamics*; Advances in Chemical Physics **71**, (1988)

[Bro89]: C. L. Brooks, M. Karplus, *Solvent Effects on Protein Motion and Protein Effects on Solvent Motion Dynamics of the Active Site Region of Lysozyme*, J. Mol. Biol. **208**, 159 (1989)

[Bro93]: C. L. Brooks, D. A. Case, *Simulations of Peptide Conformational Dynamics and Thermodynamics*, Chem. Rev. **93**, 2487 (1993)

[Bro95]: B. R. Brooks, D. Janezic, M. Karplus, *Harmonic Analysis of Large Systems*, Journal of Computational Chemistry **16** (12), 1522 (1995)

[Bru05]: J. B. Brubach, A. Mermet, A. Filabozzi, A. Gerschel, P. Roy, *Signatures of the hydrogen bonding in the infrared bands of water*, J. Chem. Phys. **122**, 184509 (2005)

[Bry87a]: J. D. Bryngelson, P. G. Wolynes, *Spin glasses and the statistical mechanics of protein folding*, Proc. Nati. Acad. Sci. USA **84**, 7524 (1987)

[Bry83b]: W. P. Bryan, *Thermodynamic Models for Water-Protein Sorption Hysteresis*, Biopolymers **26**, 1705 (1983)

[Buc80]: A. D. Buckingham, S. Bratos (Ed.), *Vibrational spectroscopy of molecular liquids and solids*, Plenum Press, New York, (1980)

[Bul73]: H. B. Bull, K. Breese, *Temperature Dependence of Partial Volumes of Proteins*, Biopolymers **12**, 2351 (1973)

[Buo71]: U. Buontempo, G. Careri, *Far-Infrared Spectra of Some Globular Proteins*, Biopolymers 10, 2377 (1971)

[Bur98]: B. D. Bursulaya, H. J. Kim, *Spectroscopic and dielectric properties of liquid water: A molecular dynamics simulation study*, J. Chem. Phys. **109** (12), 4911 (1998)

[Cal51]: H. B. Callen, T. A. Welton, *Irreversibility and Generalized Noise*, Phys. Rev. **83**, 34 (1951)

[Cal52]: H. B. Callen, R. F. Greene, *On a Theorem of Irreversible Thermodynamics*, Phys. Rev. **86**, 702 (1952)

[Car80]: G. Careri, E. Gratton, P. H. Yang, J. A. Rupley, *Correlation of IR spectroscopic, heat capacity, diamagnetic susceptibility and enzymatic measurements on lysozyme powder*, Nature **284**, 572 (1980)

[Cas79]: D. A. Case, M. Karplus, *Dynamics of Ligand Binding to Heme Proteins*, J. Mol. Biol. **132**, 243 (1979)

[Cas06]: D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, P. A. Kollman, *AMBER 9*, University of California, San Francisco, (2006)

[Cha03]: T. Chatake, A. Ostermann, K. Kurihara, F. G. Parak, N. Niimura, *Hydration in Proteins Observed by High-Resolution Neutron Crystallography*, Proteins: Structure, Function, and Genetics **50**, 516 (2003)

[Che63]: G. V. Chester, *The theory of irreversible processes*, Rep. Prog. Phys. **26**, 411 (1963)

[Che08]: W. Chen, M. Sharma, R. Resta, G. Galli, R. Car, *Role of dipolar correlations in the infrared spectra of water and ice*, Physical Rev. B **77**, 245114 (2008)

[Cho86]: K. C. Chou, *Origin of Low-Frequency Motions in Biological Macromolecules*, Biophysical Chemistry **25**, 105 (1986)

[Cie01a]: T. Cierpicki, J. Otlewski, *Amide proton temperature coefficients as hydrogen bond indicators in Proteins*, *Journal of Biomolecular NMR* **21,** 249 (2001)

[Cie01b]: P. Cieplak, J. Caldwell, P. Kollman, *Molecular Mechanical Models for Organic and Biological Systems Going Beyond the Atom Centered Two Body Additive Approximation: Aqueous Solution Free Energies of Methanol and N-Methyl Acetamide, Nucleic Acid Base, and Amide Hydrogen Bonding and Chloroform/Water Partition Coefficients of the Nucleic Acid Bases*, Journal of Computational Chemistry **22** (10), 1048 (2001)

[Cha83]: D. Chandler, J. D. Weeks, H. C. Andersen, *Van der Waals Picture of Liquids, Solids, and Phase Transformations*, Science, New Series **220**, 787 (1983)

[Che85]: T. C. Cheam, S. Krimm, *Infrared Intensities of amide modes in N-methylamide and poly(glycine I) from ab initio calculations of dipole moment derivatives of N-methylamide*, J. Chem. Phys. **82** (4), 1631 (1985)

[Che03]: J. Y. Chen, J. Cerne, A. G. Markelz, *Far Infrared Sensing Of The Oxidation State Of Heme Proteins*, Arxiv cond. Mat. 0307557, (2003)

[Che05]: J. Y. Chen, J. R. Knab, J. Cerne, A. G. Markelz, *Large oxidation dependence observed in terahertz dielectric response for cytochrome c*, Phys. Rev. E **72**, 040901 (2005)

[Che07]: J. Y. Chen, J. R. Knab, S. Ye, Y. He, A. G. Markelz, *Terahertz dielectric assay of solution phase protein binding*, Applied Physics Letters **90**, 243901 (2007)

[Chi73]: Y. N. Chirgadze, A. M. Ovsepyan, *Observation of Small Conformational Changes in the Sperm-Whale Myoglobin by the Far-Infrared Spectra*, Biopolymers **12**, 637 (1973)

[Col06]: M. E. Colvin, C. J. Cramer, C.E. Dykstra, J. H. Jensen, S. Krimm, J. Rivail, A. J. Thakkar, M. Yanez, *Molecular quantum mechanics to biodynamics: Essential connections*, Journal of Molecular Structure: Theochem **764**, 1 (2006)

[Coo76]: A. Cooper, *Thermodynamic fluctuations in protein molecules*, Proc. Natl. Acad. Sci. USA **73** (8), 2740 (1976)

[Cor02]: F. Cordier, S. Grzesiek, *Temperature-dependence of Protein Hydrogen Bond Properties as Studied by High-resolution NMR*, J. Mol. Biol. **715**, 739 (2002)

[Cre85]: T. E. Creighton, *The Problem of How and Why Proteins Adopt Folded Conformations*, J. Phys. Chem. **89**, 2452 (1985)

[Cus88]: S. Cusack, J. Smith, J Finney, B. Tidor, M. Karplus, *Inelastic Neutron Scattering Analysis of Picosecond Internal Protein Dynamics Comparison of Harmonic Theory with Experiment*, J. Mol. Biol. **202**, 903 (1988)

[Cus90]: S. Cusack, W. Doster, *Temperature dependence of the low frequency dynamics of myoglobin Measurement of the vibrational frequency distribution by inelastic neutron scattering*, Biophysical Journal **58**, 243 (1990)

[Dan97]: L. X. Dang, T. M. Chang, *Molecular dynamics study of water clusters, liquid, and liquid–vapor interface of water with many-body potentials*, J. Chem. Phys. **106** (19), 8149 (1997)

[Dar93]: N. J. Darby, T. E. Creighton, *Protein Structure*, Oxford Univ. Press, Oxford, (1993)

[Dau97]: M. Daune, *Molekulare Biophysik*, Vieweg, Wiesbaden, (1997)

[Dav63]: M. Davies (Ed.), *Infra-red spectroscopy and molecular structure : an outline of the principles*, Elsevier, Amsterdam, (1963)

[Deb01]: P. G. Debenedetti, F. H. Stillinger, *Supercooled Liquids and the glass transition*, Nature **410**, 259 (2001)

[DeL92]: E. C. de Lara, A. M. Goulay, J. Soussen-Jacob, R. Kahn, *Effect of an electric field on a methane molecule*, Molecular Physics **76** (5), 1049 (1992)

[Dem97]: F. Demmel, W. Doster, W. Petry, A. Schulte, *Vibrational frequency shifts as a probe of hydrogen bonds: thermal expansion and glass transition of myoglobin in mixed solvents*, Eur Biophys J **26**, 327 (1997)

[Den96]: V. P. Denisov, B. Halle, *Protein hydration dynamics in aqueous solution*, Faraday Discuss. **103**, 227 (1996)

[Der95]: Z. S. Derewenda, L. Lee, U. Derewenda, *The Occurrence of C-H···O Hydrogen Bonds in Proteins*, J. Mol. Biol. **252**, 248 (1995)

[Des91]: G. R. Desiraju, *The C-H O Hydrogen Bond in Crystals: What is it?*, Acc. Chem. Res. **24**, 290 (1991)

[Des04a]: F. Despa, A. Fernandez, R. S. Berry, *Dielectric Modulation of Biological Water*, PRL **93**, 228104 (2004)

[Des04b]: F. Despa, A. Fernandez, R. S. Berry, *Publisher's Note: Dielectric Modulation of Biological Water*, PRL **93**, 269901 (2000)

[Die97]: M. Diehl, W. doster, W. Petry, H. Schober, *Water-Coupled Low-Frequency Modes of Myoglobin and Lysozyme Observed by Inelastic Neutron Scattering*, Biophysical Journal **73**, 2726 (1997)

[Din95]: Y. Ding, D. N. Bernardo, K. Krogh-Jespersen, R. M. Levy, *Solvation Free Energies of Small Amides and Amines from Molecular Dynamics/Free Energy Perturbation Simulations Using Pairwise Additive and Many-Body Polarizable Potentials*, J. Phys. Chem. **99** (29), 11575 (1995)

[Dör08]: S. Dörr, U. Schade, P. Hellwig, *Far infrared spectroscopy on hemoproteins: A model compound study from 1800–100 cm^{-1}*, Vibrational Spectroscopy **47**, 59 (2008)

[Don95]: A. Dong, S. J. Prestreski, S. D. Allison, J. F. Carpenter, *Infrared Spectroscopic Studies of Lyophilization- and Temperature-Induced Protein Aggregation*, Journal of Pharmaceutical Sciences **84** (4), 415 (1995)

[Dos83]: W. Doster, *Viscosity Scaling and Protein Dynamics*, Biophysical Chemistry **17**, 97 (1983)

[Dos86]: W. Doster, A. Bachleitner, R. Dunau, M. Hiebl, E. Lüscher, *Thermal Properties of Water in Myoglobin Crystalls and Solutions at subzero Temperatures*, Biophysical Journal **50**, 213 (1986)

[Dos89]: W. Doster, S. Cusack, W. Petry, *Dynamical transition of myoglobin revealed by inelastic neutron scattering*, Nature **337**, 754 (1989)

[Dos90]: W. Doster, S. Cusack, W. Petry, *Dynamic Instability of Liquidlike Motions in a Globular Protein Observed by Inelastic Neutron Scattering*, PRL **65** (8), 1080 (1990)

[Dos98]: W. Doster, M. Settles, *The Dynamical Transition in Proteins: The Role of Hydrogen Bonds*, Les Houches Lectures: Workshop on Hydration Processes in Biology: Theoretical and Experimental Approaches, IOS Press, (1998)

[Dos05]: W. Doster, M. Settles, *Protein–water displacement distributions*, Biochimica et Biophysica Acta **1749**, 173 (2005)

[Dos08]: W. Doster, *The dynamical transition of proteins, concepts and misconceptions*, Eur. Biophys. J. **37**, 591 (2008)

[Dou98]: R. C. Dougherty, *Temperature and pressure dependence of hydrogen bond strength: A perturbation molecular orbital approach*, J. Chem. Phys. **109** (17), 7372 (1998)

[Dra66]: D. A. Draegert, N. W. B. Stone, B. Curnuite, D. Williams, *Far-Infrared Spectrum of Liquid Water*, J. Opt. Soc. **56** (1), 64 (1966)

[Dut09]: P. Dutta, K. Tominaga, *Terahertz time-domain spectroscopic study of the low-frequency spectra of nitrobenzene in alkanes*, Journal of Molecular Liquids **147** (1), 45 (2009)

[Dzy61]: I. E. Dzyaloshinskii, E. M. Lifshitz, L. P. Pitaevskii, *The general theory of van der Waals forces*, Advances in Physics **10**,165 (1961)

[Eav03]: J. D. Eaves, C. J. Fecko, A. L. Stevens, P. Peng, A. Tokmakoff, *Polarization-selective femtosecond Raman Spectroscopy of low-Frequency motions in hydrated protein Films*, Chemical Physics Letters **376**, 20 (2003)

[Ebb07]: S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner, M. Havenith, *An extended dynamical hydration shell around proteins*, PNAS **104** (52), 20749 (2007)

[Ebb08]: S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, M. Gruebele, D. M. Leitner, M. Havenith, *Protein Sequence- and pH-Dependent Hydration Probed by Terahertz Spectroscopy*, JACS **130**, 2374 (2008)

[Ede80]: J. Eden, P. R. C. Gascoyne, R. Pethig, *Dielectric and Electrical Properties of Hydrated Bovine Serum Albumin*, J.C.S. Faraday I **76**, 426 (1980)

[Edl02]: J. Edler, P. Hamm, *Self-trapping of the amide I band in a peptide model crystal*, J. of Chem. Phys. **117** (5), 2415 (2002)

[Eds83]: J. T. Edsall, H. A. McKenzie, *Water and Proteins. II. The Location and Dynamics of Water in Protein Systems and its Relation to their Stability and Properties*, Adv. Biophys. **16**, 53 (1983)

[Edw84a]: D. M. F. Edwards, P. A. Madden, I. R. McDonald, *A computer simulation study of the dielectric properties of a model of methyl cyanide. I. The rigid dipole case*, Molecular Physics **51** (5), 1141 (1984)

[Edw84b]: D. M. F. Edwards, P. A. Madden, I. R. McDonald, *A computer simulation study of the dielectric properties of a model of methyl cyanide. II. The interference of permanent and induced dipoles*, Molecular Physics **51** (5), 1163 (1984)

[Ell50]: A. Elliot, E. J. Ambrose, *Structure of synthetic polypeptides*, Nature **165**, 921 (1950)

[Elr91a]: M. J. Elrod, D. W. Steyert, R. J. Saykally, *Tunable far infrared laser spectroscopy of a ternary van der Waals cluster Ar_2HCI: A sensitive probe of three-body forces*, J. Chem. Phys. **94** (1), 58 (1991)

[Elr91b]: M. J. Elrod, D. W. Steyert, R. J. Saykally, *An investigation of three-body effects in intermolecular forces. II. Far-infrared vibration-rotation-tunneling laser spectroscopy of Ar_2HCI*, J. Chem. Phys. **95** (5), 3182 (1991)

[Elr94]: M. J. Elrod, R. J. Saykally, Many-*Body Effects in Intermolecular Forces*, Chem. Rev. **1994**, 1975 (1994)

[Eng91]: R. A. Engh, R. Huber, *Accurate Bond and Angle Parameters for X-ray Protein Structure Refinement*, Acta Cryst. **A47**, 392 (1991)

[Erm98]: I. V. Ermolina, V. D. Fedotov, Y. D. Feldman, *Structure and dynamic behavior of protein molecules in solution*, Physica A **249**, 347 (1998)

[Eva90]: D. J. Evans, G. P. Morriss, *Statistical Mechanics of nonequilibrium liquids*, Acad. Press, London, (1990)

[Ext89]: M. van exter, C. Fattinger, D. Grischkowsky, *Terahertz time-domain spectroscopy of water vapour*, Optics Letters **14** (20), 1128 (1989)

[Fen02]: P. W. Fenimore, H. Frauenfelder, B. H. McMahon, F. G. Parak, *Slaving: Solvent fluctuations dominate protein dynamics and functions*, PNAS **99** (25), 16047 (2002)

[Fen04]: P. W. Fenimore, H. Frauenfelder, B. H. McMahon, R. D. Young, *Bulk-solvent and hydration-shell fluctuations, similar to α and β-fluctuations in glasses, control protein motions and functions*, PNAS **101** (40), 14408 (2004)

[Fen05]: P. W. Fenimore, H. Frauenfelder, B. H. McMahon, R. D. Young, *Proteins are paradigms of stochastic complexity*, Physica A **351**, 1 (2005)

[Fer85]: A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M M. Y. Waye, G. Winter, *Hydrogen Bonding and Biological Specificity analysed by Protein Engineering*, Nature **314**, 235 (1985)

[Fer93]: M. Ferrand, A. J. Dianoux, W. Petry, G. Zaccai, *Thermal motions and function of bacteriorhodopsin in purple membranes: Effects of temperature and hydration studied by neutron scattering*, Proc. Natl. Acad. Sci. USA **90**, 9668 (1993)

[Fin98]: S. Finet, F.Bonnete, J. Frouin, A. Tardieu, *Lysozyme crystal growth, as observed by small angle X-ray scattering, proceed without crystallization intermediates*, J. Eur Biophys, **27**, 263 (1998)

[Fis02]: B. M. Fischer, M. Walther, P. U. Jepsen, *Far-infrared vibrational modes of DNA components studied by terahertz time-domain spectroscopy*, Phys. Med. Biol. **47**, 3807 (2002)

[Fis05]: B. M. Fischer, M. Hoffmann, H. Helm, R. Wilk, F. Rutz, T. Kleine-Ostmann, M. Koch, P. U. Jepsen, *Terahertz time-domain spectroscopy and imaging of artificial RNA*, Optical Express **13** (14), 5205 (2005)

[Fit99]: J. Fitter, *The Temperature Dependence of Internal Molecular Motions in Hydrated and Dry α-Amylase: The Role of Hydration Water in the Dynamical Transition of Proteins*, Biophysical Journal **78**, 1034 (1999)

[Fra76]: F. Franks (Ed.), *Water: A comprehensive treatise*, Plenum Press, New York, (1976)

[Fra79]: H. Frauenfelder, G. A. Petsko, D. Tsernoglou, *Temperature-dependent X-Ray diffraction as a probe of protein structural dynamics*, Nature **280**, 558 (1979)

[Fra87]: H. Frauenfelder, H. Hartmann, M. Karplus, I. D. Kuntz, J. Kuriyan, F. Parak, G. A. Petsko, D. Ringe, R. F. Tilton, M. L. Connolly, N. Max, *Thermal Expansion of a Protein*, Biochemistry **26**, 254 (1987)

[Fra88]: H. Frauenfelder, F. Parak, R. D. Young, *Conformational Substates in Proteins*, Ann. Rev. Biophys. Biophys. Chem. **17**, 451 (1988)

[Fra91]: H. Frauenfelder; S. G. Sligar; P. G. Wolynes, *The Energy Landscapes and Motions of Proteins*, Science **254** (5038), 1598 (1991)

[Fra94]: H. Frauenfelder, P. G. Wolynes, *Biomolecules: Where the Physics of Complexity and Simplicity Meet*, Physics Today **47** (2), 58 (1994)

[Fra98]: H. Frauenfelder, B. McMahon, *Dynamics and function of proteins: The search for general concepts*, Proc. Natl. Acad. Sci. USA **95**, 4795 (1998)

[Fra99]: H. Frauenfelder, P. G. Wolynes, R. H. Austin, *Biological Physics*, Reviews of Modern Physics **71** (2), (1999)

[Fra01]: H. Frauenfelder, B. H. McMahon, R. H. Austin, K. Chu, J. T. Grove, *The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin*, PNAS, **98** (5), 2370 (2001)

[Fra02b]: H. Frauenfelder, P. W. Fenimore, B. H. McMahon, *Hydration, slaving and protein function*, Biophysical Chemistry **98**, 35 (2002)

[Fra06]: H. Frauenfelder, P. W. Fenimore, G. Chen, B. H. McMahon, *Protein folding is slaved to solvent motions*, PNAS **103** (42), 15469 (2006)

[Ful68]: M. E. Fuller, W. S. Brey, *Nuclear Magnetic Resonance Study of Water Sorbed on Serum Albumin*, Journal of Biological Chemistry **243** (2), 274 (1968)

[Ful86]: G. D. Fullerton, V. A. Ord, I. L. Cameron, *An evaluation of the hydration of lysozyme by an NMR titration method*, Biochimiea et Biophysica Acta **869**, 230 (1986)

[Gai06]: V. I. Gaiduk, D. S. F. Crothers, *Basic Molecular Mechanisms Underlying Complex Permittivity of Water and Ice*, Phys. Chem. A **110** (30), 9361 (2006)

[Gal96]: P. Gallo, F. Sciortino, P. Tartaglia, S. H. Chen, *Slow Dynamics of Water Molecules in Supercooled States*, PRL **76** (15), 2730 (1996)

[Gar92]: A. E. Garcia, *Large-Amplitude Nonlinear Motions in Proteins*, PRL **68** (17), 2696 (1992)

[Gar93]: A. E. Garcia, L. Stiller, *Computation of the Mean Residence Time of Water in the Hydration Shells of Biomolecules*, Journal of Computational Chemistry **14** (11), 1396 (1993)

[Gas77]: P. R. C. Gascoyne, R. Pethig, *Experimental and Theoretical Aspects of Hydration Isotherms for Biomolecules*, J. Chem. Soc., Faraday Trans. **1**, 171 (1977)

[Gas81]: P. R. C. Gascoyne, R. Pethig, *Effective Dipole Moment of Protein-bound Water*, J. Chem. Soc. Faraday Trans. **77**, 1733 (1981)

[Gen79]: L. Genzel, T. Theophanides (Ed.), *Infrared and Raman spectroscopy of biological molecules*, Reidel, Dordrecht, (1979)

[Gen76]: L. Genzel, F. Keilmann, T. P. Martin, G. Winterling, Y. Yacoby, H. Fröhlich, M. W. Makinen, *Low-Frequency Raman Spectra of Lysozyme*, Biopolymers **15**, 219 (1976)

[Ger95]: M. Gerstein, J. Tsai, M. Levitt, *The Volume of Atoms on the Protein Surface: Calculated from Simulation, using Voronoi Polyhedra*, J. Mol. Biol. **249**, 955 (1995)

[Ger96]: M Gerstein, C. Chothia, *Packing at the protein-water interface*, Proc. Natl. Acad. Sci. USA **93**, 10167 (1996)

[Gho77]: P. N. Ghosh, *Hydrogen bond distance and phase change in crystalline HCN*, J. Phys. C: Solid State Phys. **10**, 4421 (1977)

[Gir02]: G. Giraud, K. Wynne, *Time-Resolved Optical Kerr-Effect Spectroscopy of Low-Frequency Dynamics in Di-L-alanine, Poly-L-alanine, and Lysozyme in Solution*, JACS **124**, 12110 (2002)

[Gir03]: G. Giraud, J. Karolin, K. Wynne, *Low-Frequency Modes of Peptides and Globular Proteins in Solution Observed by Ultrafast OHD-RIKES Spectroscopy*, Biophysical Journal **85**, 1903 (2003)

[Gla01]: R. Glaser, *Biophysics*, Springer, Berlin, (2001)

[Go83]: N. Go, T. Noguti, T. Nishikawa, *Dynamics of a small globular protein in terms of low-frequency vibrational modes*, Proc. Natl. Acad. Sci. USA **80**, 3696 (1983)

[Gor63]: R. G. Gordon, *Molecular Motion and the Moment Analysis of Molecular Spectra in Condensed Phases. I. Dipole-Allowed Spectra*, Journal of Chemical Physics **39** (11), 2788 (1963)

[Gor64a]: R. G. Gordon, *Molecular Motion and the Moment Analysis of Molecular Spectra in Condensed Phases. II. The Rotational Raman Effect*, Journal of Chemical Physics **40** (7), 1973 (1964)

[Gor64b]: R. G. Gordon, *Molecular Motion and the Moment Analysis of Molecular Spectra in Condensed Phases. II. Infrared Spectra*, Journal of Chemical Physics **41** (6), 1819 (1964)

[Gor65]: R. G. Gordon, *Molecular Motion in Infrared and Raman Spectra*, Journal of Chemical Physics **43** (4), 1307 (1965)

[Gor68a]: R. G. Gordon, *Correlation Functions for Molecular Motion*, Adv. Mag. Resonance **3**, 1 (1968)

[Gor68b]: R. G. Gordon, W. Klemperer, J. I. Steinfeld, *Vibrational and Rotational Relaxation*, Annu. Rev. Phys. Chem. **19**, 215 (1968)

[Gor71]: R. G. Gordon, R. P. McGinnis, *Intermolecular Potentials and Infraed Spectra*, Journal of Chemical Physics **55** (10), 4898 (1971)

[Goe92]: W. Götze, L. Sjogren, *Relaxation processes in supercooled liquids*, Rep. Prog. Phys. **55**, 241 (1992)

[Goe99]: W. Götze, *Recent tests of the mode-coupling theory for glassy dynamics*, J. Phys. Condens. Matter **11**, 1 (1999)

[Grd05a]: J. Grdadolnik, Y. Marechal, *Hydrogen–Deuterium Exchange in Bovine Serum Albumin Protein Monitored by Fourier Transform Infrared Spectroscopy, Part I: Structural Studies*, Appl. Spectroscopy **59** (11), 1347 (2005)

[Grd05b]: J. Grdadolnik, Y. Marechal, *Hydrogen–Deuterium Exchange in Bovine Serum Albumin Protein Monitored by Fourier Transform Infrared Spectroscopy, Part II: Kinetic Studies*, Appl. Spectroscopy **59** (11), 1357 (2005)

[Gre51]: R. F. Greene, H. B. Callen, *On the Formalism of Thermodynamic Fluctuation Theory*, Phys. Rev. **83**, 1231 (1951)

[Gre52]: R. F. Greene, H. B. Callen, *On a Theorem of Irreversible Thermodynamics. II*, Phys. Rev. **88**, 1387 (1952)

[Gre94]: J. L. Green, J. Fan, C. A. Angell, *The Protein-Glass Analogy: Some Insights from Homopeptide Comparisons*, J. Phys. Chem. **98**, 13780 (1994)

[Gre95]: R. B. Gregory (Ed.), *Protein-solvent interactions*, Dekker, New York, (1995)

[Gri95]: K. Griebenow, A. M. Klibanov, *Lyophilization-induced reversible changes in the secondary structure of proteins*, PNAS **92**, 10969 (1995)

[Gro02]: M. L. Groot, M. H. Vos, I. Schlichting, F. van Mourik, M. Joffre, J. C. Lambry, J. L. Martin, *Coherent infrared emission from myoglobin crystals: An electric field measurement*, PNAS **99** (3), 1323 (2002)

[Gru94]: H. Grubmüller, P. Tavan, *Molecular dynamics of conformational substates for a simplified protein model*, J. Chem. Phys. **101**, 5047 (1994)

[Gu95]: W. Gu, B. P. Schoenborn, *Molecular Dynamics Simulation of Hydration in Myoglobin*, Proteins: Structure, Function, and Genetics **22**, 20 (1995)

[Gui91]: B. Guillot, *A molecular dynamics study of the far infrared spectrum of liquid water*, J. Chem. Phys. **95** (3), 1543 (1991)

[Gup68]: V. D. Gupta, S. Trevino, H. Boutin, *Vibration Spectra of Polyglycine*, J. Chem. Phys. 48 (7), 3008 (1968)

[Gus77]: W. F. van Gunsteren, H. J. C. Berendsen, *Algorithms for Macromolecular Dynamics and Constraint Dynamics*, Molecular Physics **34** (5), 1311 (1977)

[Hag78]: A. T. Hagler, J. Moult, *Computer simulation oft he solvent structure around biological macromolecules*, Nature **272**, 222 (1978)

[Hal63]: H. E. Hallam, M. Davies (Ed.), *Infra-red spectroscopy and molecular structure*, Elsevier, Amsterdam, (1963)

[Hal04]: B. Halle, *Protein hydration dynamics in solution: a critical survey*, Phil. Trans. R. Soc. Lond. B **359**, 1207 (2004)

[Ham37]: H. C. Hamaker, *The London – Van der Waals Attraction between spherical Particles*, Physica IV **10**, 1058 (1937)

[Ham95]: P. Hamm, M. Zurek, W. Mantele, M. Meyer, H. Scheer, W. Zinth, *Femtosecond Infrared Spectroscopy of Reaction Centers from Rhodabacter sphaeroides Between 1000 and 1800 cm^{-1}*, Proc. Natl. Acad. Sci. USA **92** (6), 1826 (1995)

[Ham98]: P. Hamm, M. Lim, R. M. Hochstrasser, *Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy*, J. Phys. Chem. B **102**, 6123 (1998)

[Har72]: S. C. Harvey, P. Hoekstra, *Dielectric Relaxation Spectra of Water Adsorbed on Lysozyme*, J. Physical Chemistry **76** (21), 2987 (1972)

[Har94]: Y. Harpaz, M. Gerstein, C. Clothia, *Volume changes on protein folding*, Structure **2** (7), 641 (1994)

[Har05]: E. Harder, B. Kim, R. A. Friesner, B. J. Berne, *Efficient Simulation Method for Polarizable Protein Force Fields: Application to the Simulation of BPTI in Liquid Water*, J. Chem. Theory Comput. **1** (1), 169 (2005)

[Has73]: J. B. Hasted, *Aqueous Dielectrics*, Chapman and Hall, London, (1973)

[Hay93]: S. Hayward, A. Kitao, F. Hirata, N. Go, *Effect of Solvent on Collective Motions in Globular Proteins*, J. Mol. Biol. **234**, 1207 (1993)

[Hay94]: S. Hayward, A. Kitao, N. Go, *Harmonic and anharmonic aspects in the dynamics of BPTI: A normal mode analysis and principal component analysis*, Protein Science **3**, 936 (1994)

[Hay95]: S. Hayward, N. Go, *Collective Variable Description of Native Protein Dynamics*, Annu. Rev. Phys. Chem. **46**, 223 (1995)

[Hay02]: J. A. Hayward, J. C. Smith, *Temperature Dependence of Protein Dynamics: Computer Simulation Analysis of Neutron Scattering Properties*, Biophysical Journal **82**, 1216 (2002)

[Hei08]: E. J. Heilweil, D. F. Plusquellic, S. L. Dexheimer (Ed.), *Pharmaceutical and Security Applications of Terahertz Spectroscopy in Terahertz Spectrocopy: Principles and Applications*, CRC Press, Boca Raton, (2008)

[Her01]: W. A. Herrebout, K. Clou, H. O. Desseyn, *Vibrational Spectroscopy of N-Methylacetamide Revisited*, J. Phys. Chem. A **105**, 4865 (2001)

[Her03]: W. A. Herrebout, K. Clou, H. O. Desseyn, N. Blaton, *Vibrational characterization of the peptide bond*, Spectrochimica Acta Part A **59**, 47 (2003)

[Heu06]: U. Heugen, G. Schwaab, E. Bründermann, M. Heyden, X. Yu, D. M. Leitner, M. Havenith, *Solute-induced retardation of water dynamics probed directly by terahertz spectroscopy*, PNAS **103** (33), 12301 (2006)

[Hey02]: C. D. Heyes, M. A. El-Sayedi, *The Role of the Native Lipids and Lattice Structure in Bacteriorhodopsin Protein Conformation and Stability as Studied*

by Temperature-dependent Fourier Transform-Infrared Spectroscopy, J. Biol. Chem. **277** (33), 29437 (2002)

[Hey06]: R. Heyrovska, *Dependence of the length of the hydrogen bond on the covalent and cationic radii of hydrogen, and additivity of bonding distances*, Chemical Physics Letters **432**, 348 (2006)

[Hey08]: M. Heyden, E. Bründermann, U. Heugen, G. Niehues, D. M. Leitner, M. Havenith, *Long-Range Influence of Carbohydrates on the Solvation Dynamics of Water: Answers from Terahertz Absorption Measurements and Molecular Modeling Simulations*, J. Am. Chem. Soc. **130** (17), 5773 (2008)

[Hie91]: M. Hiebl, R. Maksymiw, *Anomalous Temperature Dependence of the Thermal Expansion of Proteins*, Biopolymers **31**, 161 (1991)

[Hil94]: J. R. Hill, A. Tokmakoff, K. A. Peterson, B. Sauter, D. Zimdars, Dana D. Dlott, M. D. Fayer, *Vibrational Dynamics of Carbon Monoxide at the Active Site of Myoglobin: Picosecond Infrared Free-Electron Laser Pump-Probe Experiments*, J. Phys. Chem. **98** (43), 11213 (1994)

[Hil06]: V. J. Hilser, B. Garcia-Moreno, T. G. Oas, G. Kapp, S. T. Whitten, *A Statistical Thermodynamic Model of the Protein Ensemble*, Chem. Rev. **106**, 1545 (2006)

[Hin94]: H. J. Hinz, T. Vogl, R. Meyer, *An alternative interpretation of the heat capacity changes associated with protein unfolding*, Biophysical Chemistry **52**, 275 (1994)

[Hno61]: W. S. Hnojewyj, L. H. Reyerson, *Further studies on the sorption of H2O and D2O vapors by lysozyme and the deuterium hydrogen exchange effect*, J. Phys. Chem. **65** (10), 1694 (1961)

[Hoe02]: F. J. Hoedemaeker, R. W. Visschers, A. C. Alting, K. G. de Kruif, M. E. Kuil, J. P. Abrahams, *A novel pH-dependent dimerization motif in b-lactoglobulin from pig (Sus scrofa)*, Acta Cryst. D**58**, 480 (2002)

[Hol98]: K. E. V. Holde, W. C. Johnson, P. S. Ho, *Principles of Physical Biochemistry*, Prentice Hall, (1998)

[Hub83]: R. Huber, W. S. Bennett, *Functional Significance of Flexibility in Proteins*, Biopolymers **22**, 261 (1983)

[Ibe89]: I. E. T. Iben, D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach, A. H. Xie, R. D. Young, *Glassy Behavior of a Protein*, PRL **62** (16), 1916 (1989)

[Ich91a]: T. Ichiye, M. Karplus, *Collective Motions in Proteins: A Covariance Analysis of Atomic Fluctuations in Molecular Dynamics and Normal Mode Simulations*, Proteins: Structure, Function, and Genetics **11**, 205 (1991)

[Ich91b]: K. Ichikawa, Y. Kameda, T. Yamaguchi, H. Wakita, M. Misawa, *Neutron-diffraction investigation of the intramolecular structure of a water molecule in the liquid phase at high temperatures*, Molecular Physics **73** (1), 79 (1991)

[Igl84]: H. A. Iglesias, J. Chirife, *Technical note: Correlation of BET monolayer moisture content in foods with temperature*, Journal of Food Technology **19**, 503 (1984)

[Igl86]: H. A. Iglesias, J. Chirife, C. F. Fontan, *Temperature Dependence of Water Sorption Isotherms of Some Foods*, J. Food Science **51** (3), 551 (1986)

[Jaa98]: S. Jaaskelainen, C. S. Verma, R. E. Hubbard, P. Linko, L. S. D. Caves, *Conformational change in the activation of lipase: An analysis in terms of low-frequency normal modes*, Protein Sciences **7**, 1359 (1998)

[Jac06]: M. B. Jackson, *Molecular and Cellular Biophysics*, Cambridge University Press, Cambridge, (2006)

[Jae84]: R. Jaenicke, *Protein Folding and Protein Association*, Angew. Chemie 23 (6), 395 (1984)

[Jan99]: J. Janin, *Wet and dry interfaces: the role of solvent in protein–protein and protein–DNA recognition*, Structure **7** (12), R277 (1999)

[Jef91]: G. A. Jeffrey, W. Saenger, *Hydrogen bonding in biological structures*, *Springer*, Berlin, (1991)

[Jef97]: G. A. Jeffrey, *An introduction to hydrogen bonding*, Oxford University Press, Oxford, (1997)

[Jim02]: R. Jimenez, D. A. Case, F. E. Romesberg, *Flexibility of an Antibody Binding Site Measured with Photon Echo Spectroscopy*, J. Phys. Chem. B **106** (5), 1090 (2002)

[Joe74]: M. D. Joesten, L. J. Schaad, *Hydrogen Bonding*, Dekker, New York, (1974)

[Jol95]: P Jolles (Ed.), H. Jörnvall (Ed.), *Interface between chemistry and biochemistry*, Birkhäuser, Basel, (1995)

[Jor83]: W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *Comparison of simple potential functions for simulating liquid water*, J. Chem. Phys. **79**, 926 (1983)

[Jot04]: Y. Joti, A. Kitao, N. Go, *Molecular simulation study to examine the possibility of detecting collective motion in protein by inelastic neutron scattering*, Physica B **350**, 627 (2004)

[Kam01]: G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides*, J. Phys. Chem. B **105** (28), 6474 (2001)

[Kam02]: G. A. Kaminski, H. A. Stern, B. J. Berne, R. A. Friesner, Y. X. Cao, R. B. Murphy, R. Zhou, T. A. Halgren, *Development of a polarizable force field for proteins via ab initio quantum chemistry: First generation model and gas phase tests*, Journal of Computational Chemistry **23** (16), 1515 (2002)

[Kar86]: M. Karplus, J.A. McCammon, *Dynamics of Proteins*, Scientific American 254 (4), 42 (1986)

[Kar90]: M. Karplus, G. A. Petsko, *Molecular dynamics simulations in biology*, Nature **347**, 631 (1990)

[Kar06]: A. Karshikoff, *Non-covalent Interactions in Proteins*, Imperial College Press, London, (2006)

[Kel88]: J. T. Kellis Jr, K. Nyberg, D. Sali, A. R. Fersht, *Contribution of hydrophobic interactions to protein stability*, Nature **333**, 784 (1988)

[Kel89]: J. T. Kellis Jr, K. Nyberg, A. R. Fersht, *Energetics of Complementary Side-Chain Packing in a Protein Hydrophobic Core*, Biochemistry **28** (11), 4914 (1989)

[Ken60]: J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore, *Structure of Myoglobin*, Nature **185**, 422 (1960)

[Kes00]: O. Keskin, R. L. Jernigan, I. Bahar, *Proteins with Similar Architecture Exhibit Similar Large-Scale Dynamic Behavior*, Biophysical Journal **78**, 2093 (2000)

[Khi34]: A. Khintchine, *Korrelationstheorie der stationären stochastischen Prozesse*, Mathematische Annalen **109**, (1934)

[Kil04]: G. T. Kilosanidze, A. S. Kutsenko, N. G. Esipova, V. G. Tumanyan, *Analysis of forces that determine helix formation in α-proteins*, Protein Science **13**, 351 (2004)

[Kim05]: B. Kim, T. Young, E. Harder, R. A. Friesner, B. J. Berne, *Structure and Dynamics of the Solvation of Bovine Pancreatic Trypsin Inhibitor in Explicit*

Water: A Comparative Study of the Effects of Solvent and Protein Polarizability, *J. Phys. Chem. B* **109** (34), 16529 (2005)

[Kin96]: J. T. Kindt, C. A. Schmuttenmaer, *Far-Infrared Dielectric Properties of Polar Liquids Probed by Femtosecond Terahertz Pulse Spectroscopy*, J. Phys. Chem. **100**, 10373 (1996)

[Kir52]: J. G. Kirkwood, J. B. Shumaker, *The influence of dipole moment fluctuations on the dielectric increment of proteins in solution*, Proc. Natl. Acad. Sci. USA **38** (10), 855 (1952)

[Kit91]: A. Kitao, F. Hirata, N. Go, *The effects of solvent on the conformation and the collective motions of protein: normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum*, Chemical Physics **158**, 447 (1991)

[Kit98]: A. Kitao, S. Hayward, N. Go, *Energy Landscape of a Native Protein: Jumping-Among-Minima Model*, Proteins: Structure, Function, and Genetics **33**, 496 (1998)

[Kna06]: J. Knab, J. Y. Chen, A. G. Markelz, *Hydration Dependence of Conformational Dielectric Relaxation of Lysozyme*, Biophysical Journal **90**, 2576 (2006)

[Kna07a]: J. R. Knab, J. Y. Chen, Y. He, A. G. Markelz, *Terahertz Measurements of Protein Relaxational Dynamics*, Proceedings of the IEEE **95** (8), 1605 (2007)

[Kna07b]: J. R. Knab, Y. He, F. Lipps, J. Y. Chen, B. Moeller, S. Gregurick, A. G. Markelz, *Tertiary structural effects on protein picosecond dynamics: terahertz dielectric response*, Laser Science, OSA Technical Digest (CD) Optical Society of America, LThC2 (2007)

[Kod90]: R. Kodandapani, C. G. Suresh, M. Vijayan, *Crystal Structure of Low Humidity Tetragonal Lysozyme at 2.1-A Resolution*, J. Biol. Chem. **265** (27), 16126 (1990)

[Kom93]: Y. Komeiji, M. Uebayasi, J. Someya, I. Yamato, *A Molecular Dynamics Study of Solvent Behavior Around a Protein*, Proteins: Structure, Function, and Genetics **16**, 268 (1993)

[Kov97]: H. Kovacs, A. E. Mark, W. F. van Gunsteren, *Solvent Structure at a Hydrophobic Protein Surface*, Proteins: Structure, Function, and Genetics **27**, 395 (1997)

[Kra27]: H. A. Kramers, *La diffusion de la lumiere par les atomes*, Atti Cong. Intern. Fisica **2**, 545 (1927)

[Kra80]: J. van Kranendonk, *Intermolecular spectroscopy and dynamical properties of dense systems*, North-Holland, Amsterdam, (1980)

[Kra83]: J. van Kranendonk, *Solid Hydrogen*, Plenum Press, New York, (1983)

[Kri86]: S. Krimm, J. Bandekar, *Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins*. Advances in Protein Chemistry **38**, (1986)

[Kro26]: R. de L. Kronig, *On the theory of the dispersion of X-rays*, J. Opt. Soc. Am. **12**, 547 (1926)

[Kub57]: R. Kubo, *Statistical Mechanical Theory of Irreversible Processes I. General Theory and Simple Applications to Magnetic and Conduction Problems*, Journal of the Physical Society of Japan, **12** (6), 570 (1957)

[Kub66]: R. Kubo, *The fluctuation-dissipation theorem*, Rep. Prog. Phys. **29**, 255 (1966)

[Kun74]: I. D. Kuntz, W. Kauzmann, *Hydration of Proteins and Polypeptides*, Adv. Prot. Chem. **28**, 239 (1974)

[Kur95]: I. V. Kurinov, R. W. Harrison, *The Influence of Temperature on Lysozyme Crystals. Structure and Dynamics of Protein and Water*, Acta Cryst. D **51**, 98 (1995)

[Kur05]: V. Kurkal, R.M. Daniel, John L. Finney, M. Tehei, R.V. Dunn, Jeremy C. Smith, *Low frequency enzyme dynamics as a function of temperature and hydration: A neutron scattering study*, Chemical Physics **317**, 267 (2005)

[Kur06]: V. Kurkal-Siebert, J. C. Smith, *Low-Temperature Protein Dynamics: A Simulation Analysis of Interprotein Vibrations and the Boson Peak at 150 K*, JACS **128**, 2356 (2006)

[Lab98]: M. Laberge, *Intrinsic protein electric fields: basic non-covalent interactions and relationship to protein-induced Stark effects*, Biochimica et Biophysica Acta **1386**, 305 (1998)

[Lee71]: B. Lee, F. M. Richards, *The Interpretation of Protein Structures: Estimation of Static Accessibility*, J. Mol. Biol. **55**, 379 (1971)

[Lee97]: D. T. Leeson, D. A. Wiersma, K. Fritsch, J. Friedrich, *The Energy Landscape of Myoglobin: An Optical Study*, Phys. Chem. B **101** (33), 6331 (1997)

[Lee01]: A. L. Lee, J. Wand, *Microscopic origins of entropy, heat capacity and the glass transition in proteins*, Nature **411**, 501 (2001)

[Lee05]: S. L. Lee, P. G. Debenedetti, J. R. Errington, *A computational study of hydration, solution structure, and dynamics in dilute carbohydrate solutions*, J. Chem. Phys. **122**, 204511 (2005)

[Lei06]: D. M. Leitner, M. Havenith, M. Gruebele, *Biomolecule large-amplitude motion and solvation dynamics: modelling and probes from THz to X-rays*, International Reviews in Physical Chemistry **25** (4); 553 (2006)

[Leo92]: P. E. Leopold, M. Montal, J. N. Onuchic, *Protein folding funnels: A kinetic approach to the sequence-structure relationship*, Proc. Nati. Acad. Sci. USA **89**, 8721 (1992)

[Les04]: A. M. Lesk, *Protein Science*, Oxford University Press, Oxford, (2004)

[Lev85]: M. Levitt, C. Sander, P. S. Stern, *Protein Normal-mode Dynamics: Trypsin Inhibitor, Crambin, Ribonuclease and Lysozyme*, J. Mol. Biol. **181**, 423 (1985)

[Lev88]: M. Levitt, R. Sharon, *Accurate Simulation of Protein Dynamics in Solution*, PNAS **85** (20), 7557 (1988)

[Ley99]: H. Leyser, W. Doster, M. Diehl, *Far-Infrared Emission by Boson Peak Vibrations in a Globular Protein*, PRL **82** (14), 2987 (1999)

[Lim91]: W. A. Lim, R. T. Sauer, *The Role of Internal Packing Interactions in Determining the Structure and Stability of a Protein*, J. Mol. Biol. **219**, 359 (1991)

[Lin03]: S. T. Lin, M. Blanco, W. A. Goddard, *The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of Lennard-Jones fluids*, J. Chem Phys. **119** (22), 11792 (2003)

[Lon30]: F. London, *Some characteristics and uses of molecular force*, Zeitschrift für Physikalische Chemie **11**, 222 (1930)

[Lop00]: E. C. Lopez-Diez, S. Bone, *An investigation of the water-binding properties of protein and sugar systems*, Phys. Med. Biol. **45**, 3577 (2000)

[Lot95]: W. T. Lotshaw, D. McMorrow, N. Thantu, J. S. Melinger, R. Kitchenham, *Intermolecular Vibrational Coherence in Molecular Liquids*, Journal of Raman Spectroscopy **16**, 571 (1995)

[Lun00]: P. Lunkenheimer, U. Schneider, R. Brand, A. Loidl, *Glassy dynamics*, Contemporary Physics **41** (1), 15 (2000)

[Ma00]: B. Ma, J. H. Lii, N. L. Allinger, *Molecular Polarizabilities and Induced Dipole Moments in Molecular Mechanics*, Journal of Computational Chemistry **21** (10), 813 (2000)

[Mad86]: P. A. Madden, R. W. Impey, *On the Infrared and Raman Spectra of Water in the Region 5-250 cm^{-1}*, Chem. Phys. Letters **123** (6), 502 (1986)

[Mak98]: V. A. Makarov, M. Feig, B. K. Andrews, B. M. Pettitt, *Diffusion of Solvent around Biomolecular Solutes: A Molecular Dynamics Simulation Study*, Biophysical Journal **75**, 150 (1998)

[Man00]: E. S. Manas, Z. Getahun, W. W. Wright, W. F. DeGrado, J. M. Vanderkooi, *Infrared Spectra of Amide Groups in a-Helical Proteins: Evidence for Hydrogen Bonding between Helices and Water*, J. Am. Chem. Soc. **122** (41), 9883 (2000)

[Man06]: P. K. Mankoo, T. Keyes, *Induction model for molecular electrostatics: Application to the infrared spectroscopy of CO liquid*, J. Chem. Phys. **124**, 204503 (2006)

[Mao82]: B. Mao, M. R. Pear, J. A. McCammon, S. H. Northrup, *Molecular Dynamics of Ferrocytochrome c: Anharmonicity of Atomic Displacements*, Biopolymers **21**, 1979 (1982)

[Mar00]: A. G. Markelz, A. Roitberg, E. J. Heilweil, *Pulsed terahertz spectroscopy of DNA, bovine serum albumin and collagen between 0.1 and 2.0 THz*, Chemical Physics Letters **320**, 42 (2000)

[Mar02]: A. G. Markelz, S. Whitmire, J. Hillebrecht, R. Birge, *THz time domain spectroscopy of biomolecular conformational modes*, Phys. Med. Biol. **47**, 3797 (2002)

[Mar03]: A. Markelz, S. Whitmire, *Terahertz Applications to Biomolecular Sensing*, Int. J. High Speed Electronics **13** (4), 951 (2003)

[Mar07]: A. G. Markelz, J. R. Knab, J. Y. Chen, Y. He, *Protein dynamical transition in terahertz dielectric response*, Chem. Phys. Letters **442**, 413 (2007)

[Mar08]: A. G. Markelz, *Terahertz Dielectric Sensitivity to Biomolecular Structure and Function, IEEE Journal of selected topics in quantum Electronics **14** (1), 180 (2008)*

[Mat02]: C. Mattos, *Protein–water interactions in a dynamic world*, Trends in Biochemical Sciences **27** (4), 203 (2002)

[Mat03a]: A. Matei, M. Dressel, *Experimental Determination of the Far-Infrared Optical Properties of Biological Matter in Aqueous Solution*, Journal of Biological Physics **29**, 101, (2003)

[Mat03b]: Y. L. Mathis, B. Gasharova, D. Moss, *Terahertz Radiation at ANKA, the New Synchrotron Light Source in Karlsruhe*, Journal of Biological Physics **29**, 313 (2003)

[McC76]: J. A. McCammon, B. R. Gelin, M. Karplus, P. G. Wolynes, *The hinge-bending mode in lysozyme*, Nature **262**, 325 (1976)

[McC77]: J. A. McCammon, B. R. Gelin, M. Karplus, *Dynamics of folded proteins*, Nature **267**, 585 (1977)

[McC84]: J. A. McCammon, *Protein dynamics*, Rep. Prog. Phys. **47**, 1 (1984)

[McQ76]: D. A. McQuarrie, *Statistical Mechanics*, Harper & Row, New York, (1976)

[Mei64]: J. Meixner (Ed.), *Statistical mechanics of equilibrium and non-equilibrium*, North-Holland Publ., Amsterdam, (1965)

[Mei05]: L. Meinhold, J. C. Smith, *Pressure-dependent transition in Protein dynamics at about 4 kbar revealed by molecular dynamics simulation*, Physical Review E **72**, 061908 (2005)

[Mer02]: F. Merzel, J. C. Smith, *Is the first hydration shell of lysozyme of higher density than bulk water?*, PNAS **99** (8), 5378 (2002)

[Mer05]: F. Merzel, J. C. Smith, *High-Density Hydration Layer of Lysozymes: Molecular Dynamics Decomposition of Solution Scattering Data*, J. Chem. Inf. Model. **45** (6), 1593 (2005)

[Mey96]: J. K. Meyers, C. N. Pace, *Hydrogen Bonding stabilizes globular proteins*, Biophysical Journal **71**, 2033 (1996)

[Mid96]: H. D. Middendorf, *Neutron studies of the dynamics of biological water*, Physica B **226**, 113 (1996)

[Min02]: G. Minasov, X. J. Wang, B. K. Shoichet, *An ultrahigh resolution structure of TEM-1 beta-lactamase suggests a role for Glu166 as the general base in acylation*, Journal of the American Chemical Society **124** (19), 5333 (2002)

[Mir36]: A. E. Mirsky, L. Pauling, *On the structure of native, denatured, and coagulated proteins*, Proc. N. A. S. **22**, 439 (1936)

[Mir96]: N. G. Mirkin, S. Krimm, *Ab initio vibrational analysis of isotopic derivatives of aqueous hydrogen-bonded trans-N-methylacetamide*, Journal of Molecular Structure **377**, 219 (1996)

[Mir04]: N. G. Mirkin, S. Krimm, *Structural Dependence of NH Stretch Mode Frequency Shifts in Amide and Peptide Systems*, J. Phys. Chem. A **108**, 5438 (2004)

[Miy58]: T. Miyazawa, T. Shimanouchi, S. I. Mizushima, *Normal Vibrations of N-Methylacetamide*, J. Chem. Phys. **29** (3), 611 (1958)

[Miy61]: T. Miyazawa, *Internal Rotation and Low Frequency Spectra of Esters, Monosubstituted Amides and Polyglycine*, Bull Chem. Soc. Jpn. **34** (5), 691 (1961)

[Moe92]: K. D. Moeller, G. P. Williams, S. Steinhauser, C. Hirschmugl, J. C. Smith, *Hydration-dependent far-infrared absorption in lysozyme detected using synchrotron radiation*, Biophysical Journal **61**, 276 (1992)

[Mor85]: V. N. Morozov, S. G. Gevorkian, *Low-Temperature Glass Transition in Proteins*, Biopolymers **24**, 1785 (1985)

[Mor00]: K. Moritsugo, O. Miyashita, A. Kidera, *Vibrational Energy Transfer in a Protein Molecule*, Physical Review Letters **85** (18), (2000)

[Mou76]: J. Moult, A. Yonath, W. Traub, A. Smilansky, A. Podjarny, D. Rabinovich, A. Saya, *The Structure of Triclinic Lysozyme at 2.5 A Resolution*, J. Mol. Biol. **100**, 179 (1976)

[Nan00]: N. Nandi, K. Bhattacharyya, B. Bagchi, *Dielectric Relaxation and Solvation Dynamics of Water in Complex Chemical and Biological Systems*, Chem. Rev. **100**, 2013 (2000)

[Nar71]: A. H. Narten, H. A. Levy, *Liquid Water: Molecular Correlation Functions* from X-Ray Diffraction, J. Chem. Phys. **55** (5), 2263 (1971)

[Nar92]: M. Nardone, M. A. Ricci, P. Benassi, *Brillouin and Raman scattering from liquid water*, J. Mol. Struc. **270**, 287 (1992)

[Nie92]: G. U. Nienhaus, J. R. Mourant, H. Frauenfelder, *Spectroscopic evidence for conformational relaxation in myoglobin*, Proc. Natl. Acad. Sci. USA **89**, 2902 (1992)

[Nie93]: O. F. Nielsen, *Low-Frequency Spetroscopic Studies of Interactions in Liquids*, Annu. Rep. Prog. Chem. Sec. C Phys. Chem. **90**, 3 (1993)

[Nie96]: G. U. Nienhaus, R. D. Young, *Protein Dynamics*, Encyclopedia of Applied Physics **15**, 163 (1996)

[Nie04]: G. U. Nienhaus, *Physik der Proteine*, Physik Journal **3** (4), 37 (2004)

[Nis87]: T. Nishikawa, N. Go, *Normal Modes of Vibration in Bovine Pancreatic Trypsin Inhibitor and Its Mechanical Property*, Proteins: Structure, Function, and Genetics **2**, 308 (1987)

[Nöl06]: B. Nölting, *Methods in Modern Biophysics*, Springer, Berlin (2006)

[Nyq28]: H. Nyquist, *Thermal Agitation of Electric Charge in Conductors*, Phys. Rev. **32**, 110 (1928)

[Ohm88]: I. Ohmine, H. Tanaka, P. G. Wolynes, *Large Local Energy Fluctuations in Water. II. Cooperative Motions and Fluctuations*, J. Chem. Phys. **89** (9), 5852 (1998)

[Ons31a]: L. Onsager, *Reciprocal relations in irreversible processes I*, Physical Review **37**, 405 (1931)

[Ons31b]: L. Onsager, *Reciprocal relations in irreversible processes II*, Physical Review **38**, 2265 (1931)

[Onu97]: J. N. Onuchic, Z. Luthey-Schulten, P. G. Wolynes, *Theory of Protein Folding: The Energy Landscape Perspective*, Annu. Rev. Phys. Chem. **48**, 545 (1997)

[Ort68]: W. H. Orttung, *Anisotropy of proton fluctuations in proteins. Calculations for simple models*, J. Phys. Chem. **72** (12), 4066 (1968)

[Ort70]: W. H. Orttung, *Proton binding and dipole moment of hemoglobin. Refined calculations*, Biochemistry **9** (12), 2394 (1970)

[Ott89]: G. Otting, K. Wüthrich, *Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules*, J. Am. Chem. Soc. **111** (5), 1871 (1989)

[Ott91]: G. Otting, E. Liepinsh, K. Wüthrich, *Protein Hydration in Aqueous Solution*, Science **254** (5034), 974 (1991)

[Pac98]: A. Paciaroni, A. R. Bizzarri, S. Cannistraro, *Molecular-dynamics simulation evidences of a boson peak in protein hydration water*, Physical Review E **57** (6), R6277 (1998)

[Pac99]: A. Paciaroni, A. R. Bizzarri, S. Cannistraro, *Neutron scattering evidence of a boson peak in protein hydration water*, Physical Review E **60** (3), R2476 (1999)

[Pai82]: P. C. Painter, L. E. Mosher, C. Rhoads, *Low-Frequency Modes in the Raman Spectra of Proteins*, Biopolymers **21**, 1469 (1982)

[Pal97]: K. Palmo, S. Krimm, *Electrostatic Model for Infrared Intensities in a Spectroscopically Determined Molecular Mechanics Force Field*, Journal of Computational Chemistry **19** (7), 754 (1997)

[Pal98]: E. Palik, *Handbook of optical constants of solids*, Academic Press, San Diego, (1998)

[Pal02]: S. K. Pal, J. Peon, B. Bagchi, A. H. Zewail, *Biological Water: Femtosecond Dynamics of Macromolecular Hydration*, J. Phys. Chem. B **106**, 12376 (2002)

[Pal04]: S. K. Pal, A. H. Zewail, *Dynamics of Water in Biological Recognition*, Chem. Rev.**104**, 2099 (2004)

[Pal07]: K. Palmo, S. Krimm, *Group Polarizability Model for Molecular Mechanics Energy Functions*, J. Chem. Theory Comput. **3**, 2120 (2007)

[Pao07]: M. Paolantoni, P. Sassi, A. Morresi, S. Santini, *Hydrogen bond dynamics and water structure in glucose-water solutions by depolarized Rayleigh scattering and low-frequency Raman spectroscopy*, J. Chem. Phys. **127**, 024504 (2007)

[Pap96]: R. V. Pappu, W. J. Schneller, D. L. Weaver, *Electrostatic Multipole Representation of a Polypeptide Chain: An Algorithm for Simulation of Polypeptide Properties*, Journal of Computational Chemistry **17**, 1033 (1996)

[Par81]: F. Parak, E. N. Frolov, R. L. Mössbauer, V. I. Goldanskii, *Dynamics of Metmyoglobin Crystals Investigated by Nuclear Gamma Resonance Absorption*, J. Mol. Biol. **145**, 825 (1981)

[Par83]: F. S. Parker, *Applications of infrared, raman, and resonance raman spectroscopy in biochemistry*, Plenum Pr., New York, (1983)

[Par03]: F. G. Parak, *Physical aspects of protein dynamics*, Rep. Prog. Phys. **66**, 103 (2003)

[Pau67]: L. Pauling, *The Nature of the Chemical Bond*, Cornell University Press, Ithaca, (1967)

[PDB08]: Protein Data Bank, www.pdb.org, (2008)

[Ped92]: J. E. Pedersen, S. R. Keiding, *THz Time-Domain Spectroscopy of Nonpolar Liquids*, IEEE Journal of Quantum Electronics **28** (10), 2518 (1992)

[Per60]: M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, A. C. T. North, *Structure of Haemoglobin*, Nature **185**, 416 (1960)

[Per66]: M. F. Perutz, F. S. Mathews, *An X-Ray Study of Azide Methaemoglobin*, J. Mol. Biol. **21**; 199 (1966)

[Per82]: W. B. Person (Ed.), *Vibrational intensities in infrared and Raman spectroscopy*, Elsevier, Amsterdam, (1982)

[Per86]: S. J. Perkins, *Protein volumes and hydration effects*, Eur. J. Biochem. **157**, 169 (1986)

[Pet79]: Peticolas WL. *Low frequency vibrations and the dynamics of proteins and polypeptides*. Methods Enzymol. **61**, 425 (1979)

[Pet84]: G. A. Petsko, D. Ringe, *Fluctuations in Protein Structure from X-Ray Diffraction*, Ann. Rev. Biophys. Bioeng.**13**, 331 (1984)

[Pet87]: R. Pethig, D. B. Kell, *The passive electrical properties of biological systems: their significance in physiology, biophysics and biotechnology*, Phys. Med. Biol. **32** (8), 933 (1987)

[Pet92]: R. Pethig, *Protein-Water Interactions Determined by Dielectric Methods*, Annu. Rev. Phys. Chem. **43**, 177 (1992)

[Plu03]: D. F. Plusquellic, T. M. Korter, G. T. Fraser, R. J. Lavrich, E. C. Benck, C. R. Bucher, A. R. H. Walker, J. L. Domenech, *Continuous-Wave Terahertz Spectroscopy of Plasmas and* Biomolecules, Intern. J. High Speed Electronics and Systems **13** (4), 1287 (2003)

[Pon06]: V. V. Ponkratov, J. Friedrich, J. M. Vanderkooi, *Hole burning experiments with proteins: Relaxations, fluctuations and glass-like features*, Journal of Non-Crystalline Solids **352**, 4379 (2006)

[Poo83a]: P. L. Poole, J. L. Finney, *Sequential Hydration of a Dry Globular Protein*, Biopolymers **22**, 255 (1983)

[Poo83b]: P. L. Poole, J. L. Finney, *Hydration-induced conformational and flexibility changes in lysozyme at low water content*, Int. J. Biol. Macromol. **5**, 308 (1983)

[Poo84]: P. L. Poole, J. L. Finney, *Sequential Hydration of Dry Proteins: A Direct Difference IR Investigation of Sequence Homologs Lysozyme and alpha-Lactalbumin*, Biopolymers **23**, 1647 (1984)

[Por96]: D. Porschke, *Electrostatics and Electrodynamics of Bacteriorhodopsin*, Biophysical Journal **71**, 3381 (1996)

[Por97]: D. Porschke, *Macrodipoles Unusual electric properties of biological macromolecules*, Biophysical Chemistry **66**, 241 (1997)

[Pow87]: J. W. Powell, G. S. Edwards, L. Genzel, F. Kremer, A. Wittlin, *Investigation of far-infrared vibrational modes in polynucleotides*, Physical Review A **35** (9), (1987)

[Pow91]: J. W. Powell, W.L. Peticolas, L. Genzel, *Observation of the far-infrared spectrum of five oligonucleotides*, J. of Molecular structure **247**, 107 (1991)

[Pre93]: S. J. Prestrelski, N. Tedeschi, T. Arakawa, J. F. Carpenter, *Dehydration-induced Conformational Transitions in Proteins and Their Inhibition by Stabilizers*, Biophysical Journal **65**, 661 (1993)

[Ric77]: F. M. Richards, *Areas, Volumes, Packing, and Protein Structure*, Ann. Rev. Biophys. Bioeng. **6**, 151 (1977)

[Roc98]: C. Rocchi, A. R. Bizzarri, S. Cannistraro, *Water dynamical anomalies evidenced by molecular-dynamics simulations at the solvent-protein interface*, Physical Review E **57** (3), 3315 (1998)

[Rog03]: T. Rog, K. Murzyn, K. Hinsen, G. R. Kneller, *nMoldyn: A Program Package for a Neutron Scattering Oriented Analysis of Molecular Dynamics Simulations*, Journal of Computational Chemistry **24** (5), 657 (2003)

[Roh05]: J. H. Roh, V. N. Novikov, R. B. Gregory, J. E. Curtis, Z. Chowdhuri, A. P. Sokolov, *Onsets of Anharmonicity in Protein Dynamics*, PRL **95**, 038101 (2005)

[Roi95]: A. Roitberg, R. B. Gerber, R. Elber, M. A. Ratner, *Anharmonic Wave Functions of Proteins: Quantum Self-Consistent Field Calculations of BPTI*, Science **268** (5215), 1319 (1995)

[Roi97]: A. Roitberg, R. B. Gerber, M. A. Ratner, *A Vibrational Eigenfunction of a Protein: Anharmonic Coupled-Mode Ground and Fundamental Excited States of BPTI*, J. Phys. Chem. B **101** (9), 1700 (1997)

[Ron97]: C. Ronne, L. Thrane, P. O. Astrand, A. Wallqvist, K. V. Mikkelsen, S. R. Keiding, *Investigation of the temperature dependence of dielectric relaxation in liquid water by THz reflection spectroscopy and molecular dynamics simulation*, J. Chem. Phys. **104** (14), 5319 (1997)

[Ros02]: F. Rosca, A. T. N. Kumar, D. Ionascu, X. Ye, A. A. Demidov, T. Sjodin, D. Wharton, D. Barrick, S. G. Sligar, T. Yonetani, P. M. Champion, *Investigations of Anharmonic Low-Frequency Oscillations in Heme Proteins*, J. Phys. Chem. A **106** (14), 3540 (2002)

[Ros05]: J. Rosgen, B. M. Pettitt, D. W. Bolen, *Protein Folding, Stability, and Solvation Structure in Osmolyte Solutions*, Biophysical Journal **89**, 2988 (2005)

[Rot93]: C. M. Roth, A. M. Lenhoff, *Electrostatic and van der Waals Contributions to Protein Adsorption: Computation of Equilibrium Constants, Langmuir **9**, 962 (1993)*

[Rot95]: C. M. Roth, A. M. Lenhoff, *Electrostatic and van der Waals Contributions to Protein Adsorption: Comparison of Theory and Experiment, Langmuir **11**, 3500 (1993)*

[Rot96]: C. M. Roth, B. L. Neal, A. M. Lenhoff, *Van der Waals Interactions Involving Proteins*, Biophysical Journal **70**, 977 (1996)

[Row80]: S. P. Rowland (Ed.), *Water in Polymers*, ACS, Washington D. C., (1980)

[Rup80]: J. A. Rupley, S. P. Rowland (Ed.), *Water in Polymers*, ACS, Washington D. C., (1980)

[Rup83]: J. A. Rupley, E. Gratton, G. Careri, *Water and globular proteins*, Trends in Biochemical Sciences **8** (1), 18 (1983)

[Rus04]: D. Russo, G. Hura, T. Head-Gordon, *Hydration Dynamics Near a Model Protein Surface*, Biophysical Journal **86**, 1852 (2004)

[San89]: W. S. Sandberg, T. C. Terwilliger, *Influence of Interior Packing and Hydrophobicity on the Stability of a Protein*, Science **245**, 54 (1989)

[Sar94]: G. Sartor, E. Mayer, G. P. Johari, *Calorimetric Studies of the Kinetic Unfreezing of Molecular Motions in Hydrated Lysozyme, Hemoglobin, and Myoglobin*, Biophysical Journal **66**, 249 (1994)

[Sca63]: B. K. P. Scaife, *A New Method of Analysing Dielectric Measurements*, Proc. Phys. Soc. **81**, 124 (1963)

[Sca89]: B. K. P. Scaife, *Principles of dielectrics*, Clarendon Press, Oxford, (1998)

[Sch85]: J. E. Schinkel, N. W. Downer, J. A. Rupley, *Hydrogen exchange of lysozyme powders. Hydration dependence of internal motions*, Biochemistry **24** (2), 352 (1985)

[Sch94]: S. Scheiner, *Ab Initio Studies of Hydrogen bonds: The Water Dimer Paradigm*, Annu. Rev. Phys. Chem. **45**, 23 (1994)

[Sch01]: J. Schlichter, J. Friedrich, *Glasses and proteins: Similarities and differences in their spectral diffusion dynamics*, J. Chem. Phys. **114** (19), 8718 (2001)

[Sch04]: P. F. Schofield, K. S. Knight, J. A. M. van der Houwen, E. Valsami-Jones, *The role of hydrogen bonding in the thermal expansion and dehydration of brushite, di-calcium phosphate dihydrate*, Phys Chem Minerals **31**, 606 (2004)

[Sci92]: F. Sciortino, A. Geiger, H. E. Stanley, *Network defects and molecular mobility in liquid water*, J. Chem. Phys. **96** (5), 3857 (1992)

[Ser06]: I. N. Serdyuk, N. R. Zaccai, J. Zaccai, *Methods in Molecular Biophysics*, Cambridge University Press, Cambridge, (2006)

[Set92]: M Settles, F. Post, D. Müller, A. Schulte, W. Doster, *Solvent damping of internal processes in myoglobin studied by specific heat spectroscopy and flash photolysis*, Biophysical Chemistry **43**, 107 (1992)

[Set96]: M. Settles, W. Doster, *Anomalous diffusion of adsorbed water: A neutron scattering study of hydrated myoglobin*, Faraday Discuss. **103**, 269 (1996)

[Sha98]: S. L. Shamblin, B. C. Hancock, G. Zografi, *Water vapor sorption by peptides, proteins and their formulations*, European Journal of Pharmaceutics and Biopharmaceutics **45**, 239 (1998)

[Sha05]: M. Sharma, R. Resta, R. Car, *Intermolecular Dynamical Charge Fluctuations inWater: A Signature of the H-Bond Network*, PRL **95**, 187401 (2005)

[Sha07a]: M. Sharma, R. Resta, R. Car, *Dipolar Correlations and the Dielectric Permittivity of Water*, PRL **98**, 247401 (2007)

[Sha07b]: M. Sharma, R. Resta, R. Car, *Erratum: Intermolecular Dynamical Charge Fluctuations inWater: A Signature of the H-Bond Network*, PRL **100**, 019901 (2007)

[She83]: G. F. Sheats, L. S. Forster, *Fluorescence Lifetimes in hydrated Bovine Derum Albumin Powders*, Biochemical and Biophysical Research Communications **114** (3), 901 (1983)

[She00]: D. Sheehan, Physical Biochemistry, Wiley, Weinheim (2000)

[She03]: Y. C. Shen, P. C. Upadhya, E. H. Linfeld, A. G. Davies, *Temperature-dependent low-frequency vibrational spectra of purine and adenine*, Applied Physics Letters **82** (14), 2350 (2003)

[She07]: S. C. Shen, L. Santo, L. Genzel, *THz Spectra for Some Bio-molecules*, International Journal of Infrared and Millimeter Waves **28** (8), 595 (2007)

[Sho74]: W. J. Shotts, A. J. Sievers, *The Far-Infrared Properties of Polyamino Acids*, Biopolymers **13**, 2593 (1974)

[Sim96]: T. Simonson, D. Perahia, *Polar fluctuations in proteins: Moleculardynamic studies of cytochrome c in aqueous solution*, Faraday Discuss. **103**, 71 (1996)

[Sim03]: T. Simonson, *Electrostatics and dynamics of proteins*, Rep. Prog. Phys. **66**, 737 (2003)

[Smi86]: J. Smith, S. Cusack, U. Pezzeca, B. Brooks, M. Karplus, *Inelastic neutron scattering analysis of low frequency motion in proteins: A normal mode study of the bovine pancreatic trypsin inhibitor*, J. Chem Phys. **85** (6), 3636 (1986)

[Smi90a]: J. Smith, K. Kuczera, M. Karplus, *Dynamics of myoglobin: Comparison of simulation results with neutron scattering spectra*, Proc. Nati. Acad. Sci. USA **87**, 1601 (1990)

[Smi02]: A. L. Smith, H. M. Shirazi, S. R. Mulligan, *Water sorption isotherms and enthalpies of water sorption by lysozyme using the quartz crystal microbalance/heat conduction calorimeter*, Biochimica et Biophysica Acta **1594**, 150 (2002)

[Smi04]: J. C. Smith, F. Merzel, A. N. Bondar, A. Tournier, S. Fischer, *Structure, dynamics and reactions of protein hydration water*, Phil. Trans. R. Soc. Lond. B **359**, 1181 (2004)

[Sne05]: K. Sneppen, G. Zocchi, *Physics in molecular biology*, Cambridge Univ. Press, Cambridge, (2005)

[Son04]: X. Song, X. Zhao, *The van der Waals interaction between protein molecules in an electrolyte solution*, Journal of Chemical Physics **120** (4), 2005 (2004)

[Sop00]: A. K. Soper, *The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa*, Chemical Physics **258**, 121 (2000)

[Sou96]: M. Souaille, J. C. Smith, *Molecular dynamics analysis of charge fluctuations associated with far-infrared absorption in water*, Molecular Physics **87** (6), 1333 (1996)

[Ste91]: P. J. Steinbach, R. J. Loncharich, B. R. Brooks, *The effects of environment and hydration on protein dynamics: a simulation study of myoglobin*, Chemical Physics **158**, 383 (1991)

[Ste93a]: H. J. Steinhoff, B. Kramm, G. Hess, C. Owerdieck, A. Redhardt, *Rotational and Translational Water Diffusion in the Hemoglobin Hydration Shell: Dielectric and Proton Nuclear Relaxation Measurements*, Biophysical Journal **65**, 1486 (1993)

[Ste93b]: P. J. Steinbach, B. R. Brooks, *Protein Hydration Elucidated by Molecular Dynamics Simulation*, Proc. Nat. Acad. Sci. USA **90** (19), 9135 (1993)

[Ste96]: P. J. Steinbach, B. R. Brooks, *Hydrated myoglobin's anharmonic fluctuations are not primarily due to dihedral transitions*, Proc. Natl. Acad. Sci. USA **93**, 55 (1996)

[Ste01]: F. Sterpone, M. Ceccarelli, M. Marchi, *Dynamics of Hydration in Hen Egg White Lysozyme*, J. Mol. Biol. **311**, 409 (2001)

[Sti92]: JD. F. Stickle, L. G. Presta, K. A. Dill, G. D. Rose, *Hydrogen Bonding in Globular Proteins*, J. Mol. Biol. **226**, 1143 (1992)

[Sti01]: F. H. Stillinger, P. G. Debenedetti, T. M. Truskett, *The Kauzmann Paradox Revisited*, J. Phys. Chem. B **105**, 11809 (2001)

[Suk05]: S. Sukumaran, K. Hauser, A. Rauscher, W. Mäntele, *Thermal stability of outer membrane protein porin from Paracoccus denitrificans: FT-IR as a spectroscopic tool to study lipid–protein interaction*, FEBS Letters **579**, 2546 (2005)

[Sun89]: M. Sundaralingam, Y. C. Sekharudu, *Water-Inserted α-Helical Segments Implicate Reverse Turns as Folding Intermediates*, Science **244** (4910), 1333 (1989)

[Sur88]: W. K. Surewicz, H. H. Mantsch, *New insight into protein secondary structure from resolution-enhanced infrared spectra*, Biochimica et Biophysica Acta **952**, 115 (1988)

[Sur93]: W. K. Surewicz, H. H. Mantsch, D. Chapman, *Determination of protein secondary structure by Fourier transform infrared spectroscopy: A critical assessment*, Biochemistry **32** (2), 389 (1993)

[Sut98]: G. Sutmann, *Structure formation and dynamics of water in strong external electric fields*, Journal of Electroanalytical Chemistry **450**, 289 (1998)

[Sve98]: D. I. Svergun, S. Richards, M. H. J. Koch, Z. Sayers, S. Kuprin, G. Zaccai, *Protein hydration in solution: Experimental observation by x-ray and neutron scattering*, Proc. Natl. Acad. Sci. USA **95**, 2267 (1998)

[Swa82]: S. Swaminathan, T. Ichiye, W. Van Gunsteren, M. Karplus; *Time dependence of atomic fluctuations in proteins: analysis of local and collective motions in bovine pancreatic trypsin inhibitor*, Biochemistry **21**, 5230 (1982)

[Tai97]: S. Tai, K. H. Illinger, S. Papasavva, Infrared Intensities, *Atomic Charges, and Dipole Moments in the Fluoroethane Series Using Atomic Polar Tensor Analysis*, J. Phys. Chem. A **101** (50), 9749 (1997)

[Tak89]: K. Takeda, A. Wada, K. Yamamoto, Y. Moriyama, K. Aoki, *Conformational Change of Bovine Serum Albumin by Heat Treatment*, Journal of Protein Chemistry **8** (5), 653 (1989)

[Tak93]: K. Takeda, S. Hamada, A. Wada, *Secondary Structural Changes of Large and Small Fragments of Bovine Serum Albumin in Thermal Denaturation and in Sodium Dodecyl Sulfate Denaturation*, Journal of Protein Chemistry **12** (2), 223 (1993)

[Tak96]: S. Takashima, *Measurement and computation of the dipole moment of globular proteins III: Chymotrypsin*, Biophysical Chemistry **58**, 13 (1996)

[Tak99]: S. Takashima, *Computation of the dipole moment of protein molecules using protein databases. Bacteriophage T4 lysozyme and its mutants*, Colloids and Surfaces A: Physicochemical and Engineering Aspects **148**, 95 (1999)

[Tak00]: S. Takashima, *The Structure and Dipole Moment of Globular Proteins in Solution and Crystalline States: Use of NMR and X-Ray Databases for the Numerical Calculation of Dipole Moment*, Biopolymers **58**, 398 (2000)

[Tam00]: F. Tama, F. X. Gadea, O. Marques, Y. H. Sanejouand, *Building-Block Approach for Determining Low-Frequency Normal Modes of Macromolecules*, Proteins: Structure, Function, and Genetics **41**, 1 (2000)

[Tam01]: F. Tama, Y. H. Sanejouand, *Conformational change of proteins arising from normal mode calculations*, Protein Engineering **14** (1), 1 (2001)

[Tan87]: H. Tanaka, I. Ohmine, *Large Local Energy Fluctuations in Water*, J. Chem. Phys. **87** (10), 6128 (1987)

[Tar00]: M. Tarek, G. J. Martyna, D. J. Tobias, *Amplitudes and Frequencies of Protein Dynamics: Analysis of Discrepancies between Neutron Scattering and Molecular Dynamics Simulations*, J. Am. Chem. Soc. **122**, 10450 (2000)

[Tar06]: S. N. Taraskin, S. I. Simdyankin, S. R. Elliott, J. R. Neilson, T. Lo, *Universal Features of Terahertz Absorption in Disordered Materials*, PRL **97**, 055504 (2006)

[Tas82]: M. Tasumi, H. Takeuchi, S. Ataka, A. M. Dwivedi, S. Krimm, *Normal vibrations of proteins: Glucagon*, Biopolymers **21** (3), 711 (1982)

[Tee90]: M. M. Teeter, D. A. Case, *Harmonic and quasiharmonic descriptions of crambin*, J. Phys. Chem. **94** (21), 8091 (1990)

[Tee91]: M. M. Teeter, *Water-Protein Interactions: Theory and Experiment*, Annu. Rev. Biophys. Biophys. Chem. **20**, 577 (1991)

[The79]: T. M. Theophanides (Ed.), *Infrared and Raman spectroscopy of biological molecules*, Reidel, Dordrecht, (1979)

[Tho47]: A. M. Thorndike, A. J. Wells, E. B. Wilson, *The Experimental Determination of the Intensities of Infra-Red Absorption Bands*, Journal of Chemical Physics **15** (4), 157 (1947)

[Til92]: R. F. Tilton, J. C. Dewan, G. A. Petsko, *Effects of temperature on protein structure and dynamics: x-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320K*, Biochemistry **31** (9), 2469 (1992)

[Tip98]: P. A. Tipler, *Physik*, Spektrum, Heidelberg, (1998)

[Tow95]: J. K. Towns, *Moisture content in proteins: its effects and measurement*, Journal of Chromatography A **705**, 115 (19959

[Tse87]: J. S. Tse, M. L. Klein, *Molecular Dynamics Calculations of the Infrared and Raman spectra of Ice IX in the translational mode region*, Chem. Phys. Lett. **142** (3), 175 (1987)

[Tse07]: K. Tsemekhman, L. Goldschmidt, D. Eisenberg, D. Baker, *Cooperative hydrogen bonding in amyloid formation*, Protein Science 16, 761 (2007)

[Tu99]: Y. Tu, A. Laaksonen, *On the effect of Lennard-Jones parameters on the quantum mechanical and molecular mechanical coupling in a hybrid molecular dynamics simulation of liquid water*, J. Chem Phys. 111 (16), 7519 (1999)

[Tus07]: J. A. Tuszynski, *Molecular and Cellular Biophysics*, Chapman & Hall/CRC, Boca Raton, (2007)

[Twa94]: J. Twardowski, P. Anzenbacher, *Raman and IR spectroscopy in biology and biochemistry*, Ellis Horwood, New York, (1994)

[Upa04]: P. C. Upadhya, Y. C. Shen, A. G. Davies, E. H. Linfeld, *Far-infrared vibrational modes of polycrystalline saccharides*, Vibrational Spectroscopy **35**, 139 (2004)

[Ura98]: H. Urabe, Y. Sugawara, M. Ataka, A. Rupprecht, *Low-Frequency Raman Spectra of Lysozyme Crystals and Oriented DNA Films: Dynamics of Crystal Water*, Biophysical Journal 74, 1533 (1998)

[Ven01]: D. S. Venables, K. Huang, C. A. Schmuttenmaer, *Effect of Reverse Micelle Size on the Librational Band of Confined Water and Methanol*, J. Phys. Chem. B. **105** (38), 9132 (2001)

[Vij04]: J. K. Vij, D. R. J. Simpson, O. E. Panarina, *Far infrared spectroscopy of water at different temperatures: GHz to THz dielectric spectroscopy of water*, Journal of Molecular Liquids **112**, 125 (2004)

[Vil00]: J. A. Vila, D. R. Ripoll, H. A. Scheraga, *Physical reasons for the unusual a-helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides*, PNAS **97** (24), 13075 (2000)

[Vin71]: S. N. Vinogradov, R. H. Linnell, *Hydrogen Bonding*, Van Nostrand Reinhold, New York, (1971)

[Vit00]: D. Vitkup, D. Ringe, G. A. Petsko, M. Karplus, *Solvent mobility and the protein 'glass' transition*, Nature Structural Biology **7** (1), 34 (2000)

[Vli99]: H. W. T. van Vlijmen, M. Karplus, *Analysis of Calculated Normal Modes of a Set of Native and Partially Unfolded Proteins*, J. Phys. Chem. B **103** (15), 3009 (1999)

[Vos06]: N.R. Voss, M. Gerstein, T.A. Steitz, P.B. Moore, *The Geometry of the Ribosomal Polypeptide Exit Tunnel*, Journal of Molecular Biology **360** (4), 893 (2006)

[Waa73]: J. D. van der Waals, *Die Kontinuität des gasförmigen und flüssigen Zustands*, Diss., Leiden, 1873

[Wal86]: G. E. Walrafen, M. R. Fischer, M. S. Hokmabadi, W. H. Yang, *Temperature dependence oft he low- and high-frequency Raman scattering from liquid water*, J. Chem. Phys. **85** (12), 6970 (1986)

[Wal89]: G. E. Walrafen, M. S. Hokmabadi, W. H. Yang, Y. C. Chu, B. Monosmith, *Collision-induced Raman scattering from water and aqueous solutions*, J. Phys. Chem. **93** (8), 2909 (1989)

[Wal96]: G. E. Walrafen, Y. C. Chu, G. J. Piermarini, *Low-Frequency Raman Scattering from Water at High Pressures and High Temperatures*, J. Phys. Chem.**100** (24), 10363 (1996)

[Wal00]: M. Walther, B. Fischer, M. Schall, H. Helm, P. U. Jepsen, *Far-infrared vibrational spectra of all-trans, 9-cis and 13-cis retinal measured by THz time-domain spectroscopy*, Chem. Phys. Letters **332**, 389 (2000)

[Wal02]: M. Walther, P. Plochocka, B. Fischer, H. Helm, P. U. Jepsen, *Collective Vibrational Modes in Biological Molecules Investigated by Terahertz Time-Domain Spectroscopy*, Biopolymers **67**, 310 (2002)

[Wan00]: J. Wang, P. Cieplak, P. A. Kollman, *How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules?*, Journal of Computational Chemistry **21** (12), 1049 (2000)

[Wan02]: X. Wang, G. Minasov, B. K. Shoichet, *Noncovalent Interaction Energies in Covalent Complexes: TEM-1 b-Lactamase and b-Lactams*, Proteins: Structure, Function, and Genetics **47**, 86 (2002

[War74]: A. Warshel, M. Karplus, *Calculation of .pi..pi.* excited state conformations and vibronic structure of retinal and related molecules*, J. Am. Chem. Soc. **96** (18), 5677 (1974)

[War84]: S. G. Warren, *Optical constants of ice from the ultraviolet to the microwave*, Appl. Optics **23** (8), 1206 (1984)

[Wat61]: F. Watari, S. Kinumaki, *Enhanced Infrared Intensity of Benzene-Iodine Complex, Science reports of the Research Institutes*, Tohoku University. Ser. A, Physics, chemistry and metallurgy **13**, 257 (1961)

[Wax54]: N. Wax (ed.), *Selected papers on noise and stochastic processes*, Dover Publ., New York, (1954)

[Web95]: D. Weber, *Technische Feuchtemessung in Gasen und Festkörpern*, Vulkan-Verlag, Essen, (1995)

[Wei07]: U. Weierstall, R. B. Doak, J. C. H. Spence, D. Starodub, D. Shapiro, P. Kennedy, J. Warner, G. G. Hembree, P. Fromme, H. N. Chapman, *Droplet streams for serial crystallography of proteins*, Exp Fluids **44**, 675 (2007)

[Whi03]: S. E. Whitmire, D. Wolpert, A. G. Markelz, J. R. Hillebrecht, J. Galan, R. R. Birge, *Protein Flexibility and Conformational State: A Comparison of Collective Vibrational Modes of Wild-Type and D96N Bacteriorhodopsin*, Biophysical Journal **85**, 1269 (2003)

[Wie30]: N. Wiener, *Generalized harmonic analysis*, Acta Mathematica **55**, (1930)

[Wik08]: Wikipedia, Online-Enzyklopädie, (2008)

[Wil46]: E. B. Wilson, A. J. Wells, *The Experimental Determination of the Intensities of Infra-Red Absorption Bands*, Journal of Chemical Physics **14** (10), 578 (1946)

[Wil55]: E. B. Wilson, *Molecular vibrations: the theory of infrared and Raman vibrational spectra*, McGraw-Hill, New York, (1955)

[Wil63]: G. R. Wilkinson, *Low frequency Infra-red Spectroscopy*, in *Infra-Red Spectroscopy and Molecular Structure*, M Davies (Ed.), Elsevier Publishing Company, Amsterdam, (1963)

[Wit86]: A. Wittlin, L. Genzel, F. Kremer, S. Häseler, A. Poglitsch, *Far-infrared spectroscopy on oriented films of dry and hydrated DNA*, Physical Review A **34** (1), (1986)

[Wol08]: P. G. Wolynes, V. Munoz (Ed.), *The Protein Folding Energy Landscape* in Protein Folding and Misfolding, Royal Society of Chemistry, Cambridge, (2008)

[Wor00]: F. J. Wortmann, P. Augustin, C. Popescu, *Temperature Dependence of the Water-Sorption Isotherms of Wool*, Journal of Applied Polymer Science **79**, 1054 (2000)

[Xie99]: A. Xie, Q. He, L. Miller, B. Sclavi, M. R. Chance, *Low Frequency Vibrations of Amino Acid Homopolymers Observed by Synchrotron Far-IR Absorption Spectroscopy: Excited State Effects Dominate the Temperature Dependence of the Spectra*, Biopolymers **49**, 591 (1999)

[Xie02]: A. Xie, A. F. G. van der Meer, R. H. Austin, *Excited-State Lifetimes of Far-Infrared Collective Modes in Proteins*, Phys. Rev. Letters **88** (1), (2002)

[Xu06a]: J. Xu, K. W. Plaxco, S. J. Allen, *Probing the collective vibrational dynamics of a protein in liquid water by terahertz absorption spectroscopy*, Protein Sci. **15**, 1175 (2006)

[Xu06b]: J. Xu, K. W. Plaxco, S. J. Allen, *Collective Dynamics of Lysozyme in Water: Terahertz Absorption Spectroscopy and Comparison with Theory*, J. Phys. Chem. B **110**, 24255 (2006)

[Xu06c]: J. Xu, K. W. Plaxco, S. J. Allen, *Absorption spectra of liquid water and aqueous buffers between 0.3 and 3.72 THz*, J. Chem. Phys. **124**, 036101 (2006)

[Xu07]: H. Xu, Y. Zhao, S. C. Vogel, L. L. Daemen, D. D. Hickmott, *Anisotropic thermal expansion and hydrogen bonding behavior of portlandite: A high-temperature neutron diffraction study*, Journal of Solid State Chemistry **180**, 1519 (2007)

[Yam02]: K. Yamamoto, K. Tominaga, H. Sasakawa, A. Tamura, H. Murakami, H. Ohtake, N. Sarukura, *Far-Infrared Absorption Measurements of Polypeptides and Cytochrome c by THz Radiation*, Bull. Chem. Soc. Jpn. **75**, 1083 (2002)

[Yam05]: K. Yamamoto, K. Tominaga, H. Sasakawa, A. Tamura, H. Murakami, H. Ohtake, N. Sarukura, *Terahertz Time-Domain Spectroscopy of Amino Acids and Polypeptides*, Biophysical Journal **89** (3), L22 (2005)

[Yan79]: P. H. Yang, J. A. Rupley, *Protein-Water Interactions. Heat Capacity of the Lysozyme-Water System*, Biochemistry **18** (12), 2654 (1979)

[You94]: A. C. M. Young, R. F. Tilton, J. C. Dewan, *Thermal Expansion of Hen Egg-white Lysozyme*, J. Mol. Biol. **235**, 302 (1994)

[You02]: A. N. Yousif, J. W. Kan, *Visualization of chicken ovomucoid in polyacrylamide gels*, Anal. Biochem. **311** (1), 93 (2002)

[Yu05]: H. Yu, W. F. van Gunsteren, *Accounting for polarization in molecular simulation*, Computer Physics Communications **172**, 69 (2005)

[Zak88a]: A. Zaks, A. M. Kilbanov, *Enzymatic Catalysis in Nonaqueous Solvents*, J. Biol. Chem. **263** (7), 3194 (1988)

[Zak88b]: A. Zaks, A. M. Kilbanov, *The effect of water on enzyme action in organic media*, J. Biol. Chem. **263** (17), 8017 (1988)

[Zan99]: J. M. Zanotti, M. C. Bellissent-Funel, J. Parello, *Hydration-Coupled Dynamics in Proteins Studied by Neutron Scattering and NMR: The Case of the Typical EF-Hand Calcium-Binding Parvalbumin*, Biophysical Journal **76**, 2390 (1999)

[Zas07]: A. Y. Zeasetsky, V. I. Gaiduk, *Study of Temperature Effect on Far-Infrared Spectra of Liquid H_2O and D_2O by Analytical Theory and Molecular Dynamic Simulations*, J. Phys. Chem. A **111**, 5599 (2007)

[Zel95]: H. R. Zelsmann, *Temperature dependence of the optical constants for liquid H_20 and D_20 in the far IR region*, Journal of Molecular Structure **350**, 95 (1995)

[Zha04]: C. Zhang, E. Tarhan, A. K. Ramdas, A. M. Weiner, S. M. Durbin, *Broadened Far-Infrared Absorption Spectra for Hydrated and Dehydrated Myoglobin*, J. Phys. Chem. B **108**, 10077 (2004)

[Zha06]: C. Zhang, S. M. Durbin, *Hydration-Induced Far-Infrared Absorption Increase in Myoglobin*, J. Phys. Chem. B **110**, 23607 (2006)

[Zho02]: D. Zhong, S. K. Pal, D. Zhang, S. I. Chan, A. H. Zewail, *Femtosecond dynamics of rubredoxin: Tryptophan solvation and resonance energy transfer in the protein*, PNAS **99** (1), 13 (2002)

[Zhu94]: L. Zhu, J. T. Sage, P. M. Champion, *Observation of Coherent Reaction Dynamics in Heme Proteins*, Science **266** (5185), 629 (1994)

[Zun96]: G. Zundel, *The far infrared vibration of hydrogen bonds with large proton polarizability*, Journal of Molecular Structure **381**, 23 (1996)

[Zwa65]: R. Zwanzig, *Time-Correlation Functions and Transport Coefficients in Statistical Mechanics*, Annu. Rev. Phys. Chem. **16**, 67 (1965)

7. Abbreviations

BPTI	Bovine Pancreatic Trypsin Inhibitor
BSA	Bovine Serum Albumin
DNA	Desoxyribonucleic Acid
FIR	Far-Infrared
GHz	Gigahertz
HEWL	Hen Egg White Lysozyme
IR	Infrared
MD	Molecular Dynamics
MERCEP	Model of Energy Related Correlation Effects in Proteins
MHz	Megahertz
MIR	Mid-Infrared
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
RDF	Radial Distribution Function
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid
THz	Terahertz
TI	Trypsion Inhibitor

8. Acknowledgement

First of all I want to thank Prof. Dr. Martin Dressel for his scientific advice throughout this PhD thesis and his organizational support for the trilateral project. This topic touched the field of physics, biology, and mathematics hence matching my personal study background so well and I'm very grateful to the scope of development which was given to me by him.
I'm very much obliged to Prof. Dr. Jörg Wrachtrup for taking over review and examination for this PhD thesis.
My sincere gratitude to Dr. Bruno Gompf for his support and advice during the past years in which these measurements have been performed this thesis has been written. He guided me through scientific backsets during the constructing phase of many different developed setups and we share the achievements that have been made in this emerging field of research.
Many thanks to Wasim Abuillan for friendship and bolstering each other during endless nights and weekends of measuring at the institute. We both took the chance of this trilateral project in getting to know each other and becoming familiar with our different countries, religions and cultures – a precious experience besides the scientific aspect of working together.
I want to thank Gabi Untereiner for her support in protein preparation, polishing a vast number of different window materials and so forth. Without her support and effort the large number of test runs and actual measurements woundn't have been reached and hence the scientific output would have been less.
Best thanks to Dr. Natalia Drichko for setting the basement of this thesis by introducing me to the spectrometer and the far-infrared techniques as well as showing me all these tricks which have been of so much value in the time afterwards when I was in charge for this equipment.
Many thanks to all collaborators and scientist I have been in contact to within the trilateral project and for the MD simulations especially Prof. Dr. Jürgen Pleiss, Dr. Ricardo Branco, Prof. Dr. Dan Davidov and Prof. Dr. Musa Abu-Teir. Also sincere gratitude to all members of the 1. Physikalisches Institut for their friendship during the time we worked together in the labs.
Last but not least a great big hug to my family which lacked most precious time together in completing this thesis.