1

Introduction

The collective vibrations in proteins and excitons in molecular aggregates lead to structural sensitivity of optical spectra and nontrivial energy transport mechanisms. Understanding the nature of the collective modes requires detailed modeling of the couplings between individual oscillators and the interaction with the environment. Direct verification, as well as parametrization, of these models is provided by a comparison to experimental optical spectra. Especially nonlinear optical experiments are rich in detailed information about the nature and dynamics of the excitations.

1.1 Proteins

Proteins are large molecules which enable biological systems to function, to the extent that they are essential in almost any biochemical process. Their structure is a paradox: complex yet simple. The simple part is the notion that all natural proteins are constructed from only 20 basis building blocks: the amino acids. Each protein is a chain of a few to many thousands of amino acids, which are linked together by covalent bonds. The sequence of amino acids uniquely characterizes the protein. In principle, properties of the protein, for example its structure, should then be predictable from knowledge of the sequence. Here, however, the simple part ends. The amino acid chains fold into complex three-dimensional conformations, and although some progress has been made [Zhang 2008], it remains a formidable, if not impossible, task to predict the protein structure from knowledge of the sequence. A description of the structure is complicated further when the protein conformation is not

---

1 Many authors explain the structure and function of proteins, see for example Berg et al. [2002].

2 There are, actually, a few extra building blocks, which cannot be formed from the genetic code.
Figure 1.1: Illustration of the protein structure as a chain of amino acids. The box encloses a single amino acid. The difference between the amino acids, which have the same basic structure, lies in the various side groups, denoted “R”. There are 20 different choices for R. In proteins, the linear chain shown here folds into a complex three-dimensional conformation.

constant. In a solvent, which is normally water, parts of especially smaller proteins can fluctuate considerably, and the average structure is only a part of the complete picture. Even more interesting are non-equilibrium fluctuations in the structure. The best-known case is protein folding and unfolding, where the structure changes drastically in response to a change in the temperature or the concentration of certain chemical substances. Another obvious example are proteins whose function is a direct consequence of structural changes, such as membrane proteins which can open and close to permit the passage of for example ions through a cell membrane [Manor et al. 2006]. There is great progress in techniques which can be used to measure the three-dimensional structure, as well as conformational fluctuations and dynamics.

Two experimental methods capable of structure determination are X-ray diffraction and nuclear magnetic resonance (NMR). If the protein can be crystallized, the diffraction pattern of X-rays which have passed through the crystal can often be inverted to obtain the protein structure [Drenth 1994]. This technique is powerful and allows for the reconstruction of atom positions with a high resolution. However, crystallization, if at all possible, “freezes” the protein, and conformational fluctuations, let alone non-equilibrium dynamics in the natural environment, cannot be seen. NMR experiments, and, in particular, its multidimensional versions [Ernst et al. 1987], are capable of finding the protein structure under biologically relevant conditions - in a liquid. The flexibility around the average structure can be determined as well with NMR experiments. There is also a large interest in obtaining dynamical information on the picosecond to second time scale from time-resolved NMR [Mittermaier and Kay 2006] and X-ray [Schotte et al. 2003] experiments.

In this thesis, we will describe models for different, optical, probes of biomolecules. To be more precise, we are interested in experiments performed with infrared laser pulses. The need for these methods is not immediately clear. If X-ray diffraction can be used to find detailed structures, and NMR to find the dynamical structure in a solvent, why do we need the optical techniques? Three main reasons lead to interest in optical spectroscopy: the ultrafast time resolution, the possibility to study a protein in its natural solvated environment combined with a sensitivity to the local solvent, and the possibility to probe the redistribution of vibrational energy.
1.1: Proteins

Figure 1.2: X-ray structures of the proteins myoglobin [Scouloudi and Baker 1978, PDB id: 1MBS] and concanavalin A [Ahmed et al. 2007, PDB id: 2UU8]. Myoglobin transports oxygen in the muscles of humans and animals (in this particular case, seals). Concanavalin A is a sugar binding protein. The drawn line follows the sequence of amino acids. Ribbons highlight the α-helical (in myoglobin) and β-sheet (in concanavalin A) structural elements. In the myoglobin molecule, the two ends of the chain can be identified in the upper left and right corners. The pictures were generated with VMD [Humphrey et al. 1996].

Optical experiments with pulsed lasers have an intrinsic time scale which can be as short as a few femtoseconds, and is ultimately limited only by the length of the optical cycle. Many of the ideas behind sophisticated multidimensional NMR schemes can, in principle, also be used to design experiments with infrared laser pulses. Over the past years, technological advances have made this possible in practice, and multiple-pulse experiments in which the phase as well as the amplitude of the signal is measured have been developed [de Boeij et al. 1995]. The response of the sample to multiple pulses is nonlinear in the external field, and is used to produce nonlinear optical spectra. The basic idea behind multiple-pulse spectroscopy is “more is better”, and, indeed, nonlinear optical spectra contain fundamentally more information than their linear counterparts [Hochstrasser 2007b], which can be measured with a single excitation pulse. One of the most general ways to display the nonlinear response as it is measured in a three-pulse infrared experiment is two-dimensional infrared spectroscopy [Hamm et al. 1998].

When an infrared laser pulse hits the sample, the light interacts with vibrational modes in the molecules. In a molecule with $N$ atoms, there are $3N - 6$ vibrational modes, and many of them can be observed using infrared light. To reduce the amount of information, experiments usually focus on one or a few of these modes. In nonlinear infrared spectroscopy of peptides and proteins, the amide I mode is popular, because it is seen as an intense and isolated peak in the absorption spectrum. The infrared spectra of amide I modes are known to be sensitive to the protein

---

3Small proteins and protein fragments are called peptides.
structure. This is a consequence of the detailed dependence of the properties of the excited vibrations, such as their energies and damping constants, on the structure of the molecule. An example of such a correspondence in the context of proteins, which has been known for a long time, is the sensitivity of the amide I mode to ubiquitous structural elements: $\alpha$-helices show one peak at an energy of about $1650 \text{ cm}^{-1}$ in the absorption spectrum, while $\beta$-sheets absorb at $1630 \text{ cm}^{-1}$ and $1690 \text{ cm}^{-1}$ [Miyazawa 1960]. Although the solution of a complete protein structure from optical experiments is normally impossible, some structural information can, thus, be extracted from the spectra. Because of the ultrafast time resolution that comes with optical experiments, structural dynamics can be followed in real time [Hamm et al. 2008]. A nice example of the realization of this idea is the unfolding of $\beta$-sheets in proteins [Ganim et al. 2008]. The specific spectral signature of folded $\beta$-sheets disappears after rapid heating of the sample, revealing the unfolding process.

The nature of the vibrational modes in a protein does not only depend on the protein structure, but also on the dynamical structure of the local environment. The dynamical behavior of the environment, which is usually water at room temperature, leads to fluctuations in the parameters that describe the vibration. This is observed in the infrared spectrum. The nonlinear optical response is much more sensitive to the nature of the fluctuations than linear infrared spectroscopy. One of its strong points is, therefore, the ability to study the solvent close to the protein [DeFlores and Tokmakoff 2006; Mukherjee et al. 2006b].

Finally, optical spectroscopy gives direct information about the vibrational excitations themselves. The energy transport that is a consequence of vibrational dynamics plays an important role in protein function [Leitner 2008]. In a nonlinear experiment, one can prepare an excitation in a well-defined vibrational mode and look at its state after a certain waiting time has passed. The initially prepared excitation will not be stable for a long time, but will be transported through the protein, and decay into other vibrational modes in the protein and the solvent. By correlating the initial with the final state as a function of the waiting time, the vibrational transport and relaxation can be followed [Backus et al. 2008].

For the detailed interpretation of nonlinear optical spectra, comparison with simulations is vital. Models aimed at the description of optical spectra of proteins in solution must include three ingredients. First of all, understanding the structural sensitivity of the optical response requires a model that relates the vibrational states in the protein to the structure. Secondly, the effect of the environment on the vibrations must be taken into account. Not only is this essential for understanding the role of the environment, it is also turns out to be necessary for a correct description of the structural sensitivity of vibrations. Finally, if we are interested in vibrational relaxation, the vibrational dynamics must be included in the model.

By virtue of the weak interaction with other modes, a model of the amide I vibrations can be of a relatively simple building block type. Each amino acid contains one amide I vibration, which forms a building block of the model [Cho 2008]. The vibrations in different building blocks interact, and the nature of the collective vibrations seen in the spectrum depends on the strengths of the interactions. These strengths depend sensitively on the structure. The attractive point of the building block model
is the ability to predict spectral features of arbitrary structures once the interactions are known as a function of the relative orientations of blocks. This approach turns out to work, provided we include the competition of the interactions with the effect of the environment in the model.

1.2 Molecular aggregates

Molecular aggregates are supramolecular systems of aggregated molecules, which are often formed by self-assembly. They are observed in various shapes: there are linear dye molecules, and rings and cylinders are found in dyes and photosynthetic complexes. The initial interest in aggregates was triggered by a special optical property: the existence of a strong and narrow absorption line [Jelley 1936, 1937; Scheibe 1936, 1937], which is a useful characteristic in applications. Current research on molecular aggregates focuses on the nature of the excited states in various geometries, as well as on the dynamics of the excitations. The Frenkel exciton model commonly used in the description of optical excitations in molecular aggregates is of the building block type. Electrodynamic interactions couple the excitations in different blocks, while coupling to the environment destroys the coherence on larger length scales [Fidder et al. 1991a]. The excitations can transport energy through the aggregate. The theme of energy transfer as a result of long-range electromagnetic interactions can be applied in many physical systems [Agranovich and Galanin 1982]. In aggregates, the mechanism of transport, which is a key to the function, is strongly influenced by the interaction with the environment [Malyshev et al. 2007], and can be studied in detail using nonlinear optical spectroscopy. Two-dimensional optical spectroscopy with visible laser pulses, which was demonstrated around the same time as two-dimensional infrared experiments [Hybl et al. 1998], is a natural tool for such studies. In the past years, it has been applied to study natural light-harvesting complexes [Engel et al. 2007], as well as artificial cylindrical systems [Nemeth et al. 2009] and excitons in semiconductors [Zhang et al. 2007].

1.3 Models for coupled oscillators

The central theme in this thesis is the description of multiple coupled oscillators, which are the building blocks of the models, in contact with an environment. The coupling leads to collective behavior: the movement of the oscillators is choreographed through well-defined phase relations. Many of the interesting properties of large systems are a direct consequence of this collective behavior. Without coupling, there would be no structural sensitivity of the spectrum, and no transport of excitations. On the other hand, there is interaction with the environment, for example the solvent. This interaction gives rise to disorder: each oscillator is slightly different because of its local environment. Collective behavior works best when everything is the same and, consequently, disorder tends to destroy it. Only some of the oscillators “dance together”, in the sense that their phase relations are sharply defined.
It is the competition between coupling and environment-induced disorder that leads to nontrivial dynamics and makes the description of coupled oscillators so complex - and interesting. Main questions are: what is the nature of the collective excitations in the presence of disorder? How do the states depend on the structure of the system, and what is their size, defined as the number of oscillators that participate in a collective mode? Does their collective nature influence the transport of excitations through the system and their relaxation? Ultimately, answering these questions in real systems requires a comparison with experiments. Two-dimensional infrared and visible spectra are ideal to make the comparison. They contain a wealth of information about the collective excitations and their evolution in time.

In this thesis, we study theoretical models to understand the nature of collective excitations in proteins and molecular aggregates, their dynamics, and their signatures in the two-dimensional spectrum.

1.4 Outline

To answer the questions posed in the previous section, we will need to model the vibrations in proteins and the excitons in molecular aggregates, the couplings between individual oscillators, and the interaction with the environment. Such models are introduced in Chapter 2, starting with the amide I and II vibrations in proteins. The interactions between these oscillators in different building blocks are of electrodynamic or mechanical origin. The effect of the environment on the amide vibrations is considered in two models. First, we will treat the sum over states limit, in which all effects of the environment are either much faster or much slower than the time scale of the experiment. We then turn to a high-temperature model in which the environment leads to time-dependent fluctuations of the oscillator energies. At the end of the chapter, we discuss how a sum over states model can be used to describe excitons in molecular aggregates. Actual observation of the oscillators in our models is possible because they can be excited by light. The nonlinear optical response, which is particularly sensitive to their properties, is discussed in Chapter 3. The treatment will focus on the calculation of two-dimensional infrared and visible spectra of the models introduced in Chapter 2.

In Chapter 4 we study the effect of an inhomogeneous environment on the two-dimensional infrared spectra of protein $\beta$-sheets in the amide I region. We introduce a method to visualize the nature of the collective states and study them in the presence of disorder. Using the sum over states approach, we explain how the interplay between couplings and disorder leads to the characteristic Z-shape observed in experimental spectra. In the calculated two-dimensional spectra, we find a measure for the localization size of the collective vibrations.

Chapters 5 and 6 are focused on the dynamics - transport and relaxation - of excited amide vibrations. In contrast to some earlier studies, we find that the amide II mode plays a major role in the relaxation of the amide I vibration. We study the environment-assisted relaxation between these two modes in NMA-d, a small molecule that is a model for a single peptide unit, in Chapter 5. We find that this mechanism accounts for a large part of the lifetime of the amide I vibration. The pop-
ulation transfer can be followed in the two-dimensional infrared spectra. The same amide modes are studied in a peptide with multiple amide groups in Chapter 6, where we consider vibrational dynamics in a model α-helix. In this system, the relaxation from amide I to amide II modes occurs together with vibrational energy transport through the helix. This chapter also contains the discussion of a map for the mechanical coupling between amide I and II units in neighboring amide groups.

Two-dimensional optical spectra of linear molecular aggregates at low temperature are the topic of Chapter 7. We are interested in the collective excitations which arise from the competition between coupling and disorder, and find a measure for the energy-dependent exciton localization size in the two-dimensional spectrum. Furthermore, we study the spectral signature of exciton dynamics.

Finally, in Chapter 8, we perform a theoretical study of stochastic Schrödinger equations, which can be used to model the exact quantum dynamics of a system coupled linearly to a harmonic bath. This treatment aims at a more detailed description of the interaction of system oscillators with the environment.