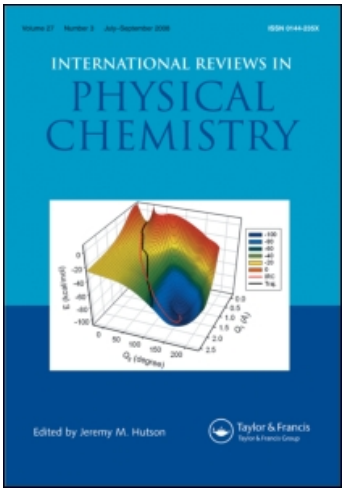

International Reviews in Physical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713724383

Biomolecule large-amplitude motion and solvation dynamics: modelling and probes from THz to X-rays

David M. Leitner[a]; Martina Havenith[b]; Martin Gruebele[c]
[a] Department of Chemistry, University of Nevada, Reno, NV 89557, USA [b] Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany [c] Departments of Chemistry and Physics, Center for Biophysics and Computational Biology, University of Illinois, Urbana, IL 61801, USA

Online publication date: 11 October 2010

To cite this Article Leitner, David M. , Havenith, Martina and Gruebele, Martin(2006) 'Biomolecule large-amplitude motion and solvation dynamics: modelling and probes from THz to X-rays', International Reviews in Physical Chemistry, 25: 4, 553 — 582
To link to this Article: DOI: 10.1080/01442350600862117
URL: http://dx.doi.org/10.1080/01442350600862117

International Reviews in Physical Chemistry,
Vol. 25, No. 4, October–December 2006, 553–582

Taylor & Francis
Taylor & Francis Group

Biomolecule large-amplitude motion and solvation dynamics: modelling and probes from THz to X-rays

DAVID M. LEITNER†, MARTINA HAVENITH‡ and MARTIN GRUEBELE*§

†Department of Chemistry, University of Nevada, Reno, NV 89557, USA
‡Lehrstuhl für Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany
§Departments of Chemistry and Physics, Center for Biophysics and Computational Biology, University of Illinois, Urbana, IL 61801, USA

Biomolecules are innately flexible, and undergo large-amplitude motions that affect the surrounding solvation shell. Dynamical X-ray scattering provides direct insight into global shape changes that the biomolecule undergoes during folding (1 nm and up length scale). THz spectroscopy directly probes solvation and collective motions on a somewhat smaller length scale (0.3–1 nm). Infrared spectroscopy looks at the influence of such motions on localized vibrational modes (up to 0.5 nm length scale). Molecular dynamics simulations and models of vibrational energy flow within biomolecules complement such experimental studies by providing a molecular-level explanation for the experimental observations. In this review, we consider the interplay between simulation and experiment across length scales for biomolecules such as carbohydrates and globular proteins.

Contents

*Corresponding author. Email: gruebele@scs.uiuc.edu

International Reviews in Physical Chemistry
ISSN 0144–235X print/ISSN 1366–591X online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01442350600862117

1. Introduction

The structure of biomolecules tells only half the story. From small carbohydrate molecules to large protein complexes, biomolecular function requires motion on many length scales [1, 2]. These motions influence, and in turn are influenced by, the solvation environment around the molecule [3, 4]. For the purpose of discussion, we broadly divide relevant structural fluctuations into three length scales (table 1). The smallest, below 0.5 nm, is described by localized vibrational motion, the making and breaking of covalent or hydrogen bonds, and the spatial extent of high-frequency normal modes of biomolecules. A quantum mechanical molecular description is most appropriate. The intermediate length scale, from about 0.3 to 1 nm, encompasses functional large-amplitude motions of protein loops, solvation layers around biomolecules, and delocalized low-frequency modes of biomolecules. On this length scale, collective interactions such as the hydrophobic effect make their first appearance. The largest length scale, about 1 nm and up, includes complete unfolding of proteins, large-scale diffusional migration in lipid membranes and large-scale motions of molecular motors. Here, the thermodynamic view becomes useful.

One of the largest questions in biophysical chemistry is how to move seamlessly from the molecular to the thermodynamic level [5–8]. When effective thermodynamic models are applied on too small of a length scale, empirical parameters end up having little meaning beyond fitting the data. When molecular level simulations are applied to very large systems, a thorough analysis of the dynamics is required to extract a mechanism.

The problem can also be viewed as one of time scales. Discounting activation barriers, molecular phenomena on the vibrational length scale occur on the femtosecond to picosecond time scale, those on the large-amplitude motion length scale occur in picoseconds to nanoseconds, and those on the length scale of biological macromolecules themselves extend from nanoseconds to microseconds. These time scales are considerably extended when activation barriers slow down the dynamics below the rate of 'elementary' vibrational or diffusional motions. Ironically, the largest objects can become the fastest: an unfolded protein that reacts to fold in an hour or longer would be considered a very slow folder by a biophysicist, yet an organic molecule that breaks apart and cannot be stored on a shelf for more than an hour would be considered an unstable reactive intermediate by an organic chemist.

The two linear spectroscopic experiments, scattering and absorption/emission, can provide much information to investigate these length and time scales, and to tie

Table 1. Measurement techniques and associated length and time scales.

Technique	Length scale	Dynamical time scale[d]	Current kinetic time scale limit
Infrared	0.001–0.5 nm[a]	0.02–0.5 ps	0.1 ps[e]
Terahertz	0.3–1 nm[b]	0.3–2 ps	0.5 ps[e]
SAXS	1–30 nm[c]	0.0001 ps	$0.1–10^6$ ps[f]

[a]Range of atomic displacement to size of a localized normal mode.
[b]Range of solvent modes to global biomolecule modes.
[c]Range from biomolecule subdomains to biomolecules.
[d]Vibrations for IR, collective modes for THz, and electronic motions for SAXS.
[e]Bandwidth-limited pump-probe experiments.
[f]Single pulse to full bunch duration on table-top or synchrotron experiments.

them together. Here we do not intend to provide a comprehensive overview of all biospectroscopy techniques. Rather, we pick representative techniques from our work associated with each length/time scale (table 1), and discuss how information about biomolecular large-amplitude motions is gathered, how information connecting different scales is obtained, and how theory and modelling on multiple length and time scales can inform such experiments. Infrared spectroscopy probes local motions or changes in bond strength during chemical reactions, which can report on larger amplitude motions indirectly [9]. Terahertz spectroscopy directly probes large-amplitude motions of biomolecules, as well as collective solvent dynamics surrounding biomolecules [10]. Small-angle X-ray scattering provides a direct look at the overall size of biomolecules, which can undergo dramatic changes during their folding or function [11–13].

For such experiments, modelling that combines a linear response framework with molecular dynamics simulations provides much useful interpretation. The time correlation function, or in the Fourier domain the spectral density function of a biomolecule and its solvent, can be directly related to linear spectroscopy and scattering data [14]. They are also amenable to both molecular and thermodynamic interpretations. This allows one to move from empirical formulas, such as the Clausius–Mossotti equation, to molecular-level unified models for the THz and IR spectra of solvated macromolecules [10]. Appropriate correlation functions and spectral densities can be computed, each associated with a particular observable such as the electric dipole moment, the total density of intermolecular and intramolecular vibrational states [15], or solvent shell order parameter. Thus, properly interpreted, a simulation can examine energy flow in biomolecules [16–18], optical properties [19–21], or the aqueous hydrogen bonding network around biomolecules [10, 22, 23], to determine how far out solvation water extends, and what its spectral density is compared to bulk water. The power of the molecular dynamics–linear response combination is its unified treatment of experimental observables and its applicability on many length and time scales, from local quantum mechanical observables to collective coordinate fluctuations. Even in the zero-frequency (long time) limit, properties continue to emerge that reveal a great deal about interactions of biomolecules with solvent molecules and energy flow within biomolecules. For example, analysis of protein structures in databases reveals that the

mass in a protein core scales as $L^{2.5}$, not as L^3, with core diameter L [24]. Protein structures are not quite three-dimensional objects because they contain empty pockets of varying size, some large enough for solvent molecules. The low dimensionality manifests itself in heat capacity and heat transport properties of proteins (see section 3).

After reviewing the spectroscopic and theoretical tools separately, we present case studies of how these methods provide fresh information about the dynamics of carbohydrates and proteins on different length and time scales. Our model systems will be simple carbohydrates such as lactose and galactose, and the small helix-bundle protein $\lambda_{6\text{-}85}$. We will see how the terahertz and infrared spectroscopies and X-ray scattering provide clues about solvation, large-amplitude motions, shapes, and large scale transitions in the shape of biomolecules. Much work remains to be done, as the current generation of modelling and experiments has just begun to scratch the surface of these problems.

2. Experimental tools

Our focus in this review is on linear spectroscopy techniques as applied to large-amplitude motions in biomolecules: one photon absorption/emission and elastic scattering experiments. Such experiments can be correlated directly with bilinear correlation functions emerging from linear response theory [25]. Except for the optical regime, nonlinear techniques and corresponding higher order correlations are still in their infancy: very few nonlinear THz experiments on biomolecules have been reported to date; two-dimensional IR experiments and their interpretation are just now coming on line [26, 27]; and practical nonlinear X-ray experiments are still years away from realization.

We discuss three techniques, each associated with a different biological length scale. Infrared spectroscopy (IR) has become a very useful tool for structural characterization of biomolecules because it directly monitors localized (usually <3 Å) vibrational dynamics on the sub-picosecond time scale, yet it can also be used indirectly to study biomolecule kinetics on the nanosecond to millisecond time scale, where many relevant biological processes occur [9]. For example, transient infrared absorption spectroscopy can be used to study vibrational relaxation and energy flow in proteins [28, 29]. At the other extreme, repeated collection of infrared spectra of a protein at microsecond to millisecond intervals maps out the changes in hydrogen bonding and secondary structure content that occur during the folding process [30]. We distinguish two time scales accessible to the experiment: the motions probed directly (here vibrations) occur on the 'dynamic' time scale; subsequently, the macromolecule reorganizes more slowly on the 'kinetic' time scale, but the dynamics of the motion probed change adiabatically during reorganization, allowing the reorganization to be monitored indirectly.

Terahertz (THz) spectroscopy is a very recent development that can also be split into 'dynamic' and 'kinetic' time scales. Here, the motions monitored directly lie in the picosecond range [31]. As a result, slower motions such as reorientation and translational modes in the solvent (as opposed to hydrogen bond vibration), or collective motion of secondary structure elements (e.g. vibrations of protein helices relative to one another, or twisting of an RNA loop) can be probed directly on a 3–10 Å

length scale. Like infrared light, terahertz light can also be used to monitor kinetics on a 0.1 nanosecond and slower time scale. It then reports more directly on changes of the large-amplitude motions the biomolecule undergoes during folding or function.

X-ray scattering operates on a very fast dynamical time scale, as electrons adjust rapidly to nuclear motion. For now, it is thus used exclusively in the 'kinetic' regime of biomolecular motions. In crystals, Debye–Waller factors report on the spatially resolved flexibility of macromolecular backbones and side chains [32]. Very recent work has also made possible the observation of large-scale motion of small molecules within crystallized proteins [33]. However, small-angle X-ray scattering (SAXS) is generally required to monitor large-scale motions of the biomolecules themselves because crystal packing prevents large-amplitude motions without fracture. SAXS can be used to measure biomolecular radii of gyration and higher structural moments [11], and with modern finite element fitting techniques, can determine rough (>10 Å resolution) structures of small biomolecules. In this regard it is complementary to cryoelectron microscopy, used to study rough structures of individual large complexes immobilized on surfaces [34].

2.1. *Infrared spectroscopy*

A number of excellent reviews on biological applications of infrared spectroscopy are available [9, 35]. The major classes of biomacromolecules have fingerprint signatures accessible in the IR, although in some cases deuterated water must be used as a solvent to avoid strong interference from bands in the 1500 and 3500 cm^{-1} regions. Proteins are generally probed in the amide I′ (the prime refers to D_2O solvent), II′ or III′ bands, centred near 1650, 1500 and 1300 cm^{-1}, respectively. The amide I′ band in particular has become an important fingerprint for protein secondary structure, with bands for helices, sheets, and coil/loop structures sufficiently separated so linear decomposition techniques can be applied to spectra [36, 37]. Nucleic acids are probed in the 1150 cm^{-1} region (phosphate stretching motions), allowing the distinction of A- and B-forms of nucleotide sequences in loops [35], or near 1660 cm^{-1}, where different base stackings contribute differently [38]. Lipids have characteristic methylene group absorptions in the 2900 cm^{-1} region [35]. However, interference from water bending overtones often makes Raman spectroscopy more attractive than infrared spectroscopy. Simple carbohydrates have characteristic peak patterns in the 1000–3500 cm^{-1} spectral region, where water absorption is relatively low [35].

The amide I′ band has been probed extensively by transient IR spectroscopy to study energy flow [29, 39–41]. Lifetimes of the amide I band of order 1 ps have been observed, though the lifetime varies significantly within the band [41]. Most of the energy transfer is intramolecular, with relatively little energy transfer from the amide I′ band directly to the solvent, as concluded from experiments [29, 40] and later by calculations [15–17]. Temperature-dependent measurements carried out from 6 K to 310 K reveal decay rates ranging from $\approx$0.5 to 1 ps^{-1} over the amide I′ band of myoglobin [39]. This weak temperature dependence reveals that little energy flows directly to the low-frequency modes of the protein.

Thermodynamic perturbations are frequently used to extract information about biomolecule stability from infrared spectra [42]. When perturbing a biomolecule

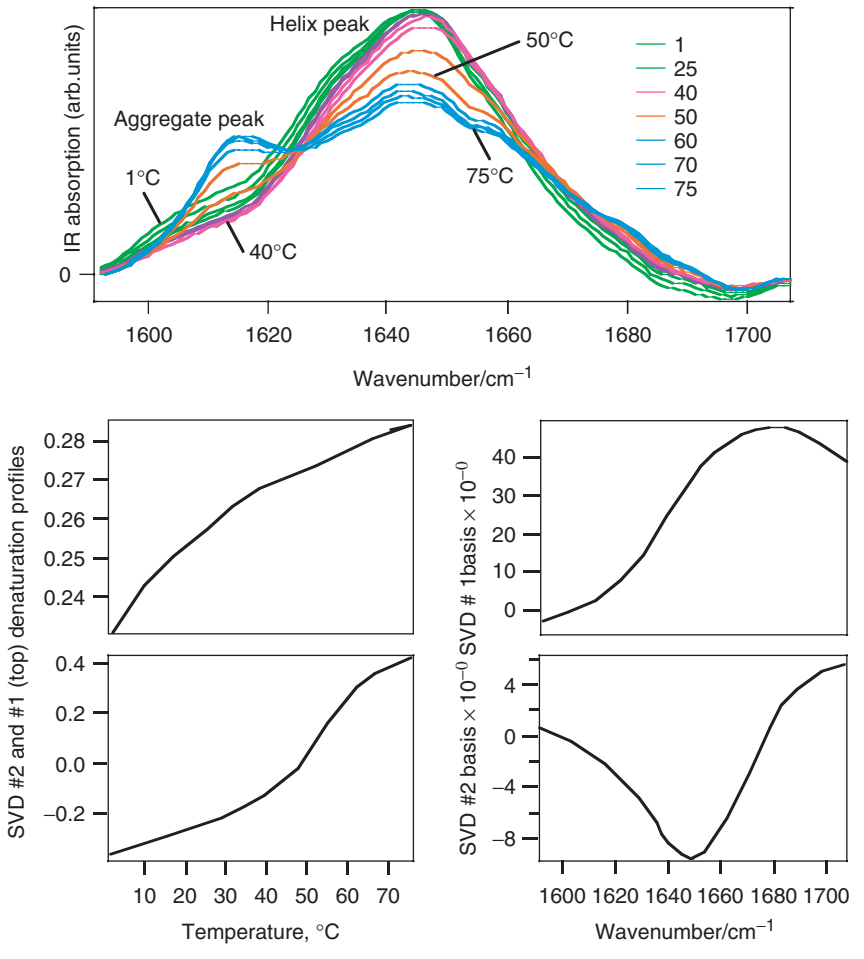

Figure 1. Top: Amide I′ IR spectrum of $\lambda_{6\text{-}85}$ mutant Y22W/Q33Y/A37,49G as a function of temperature, showing heat denaturation and formation of aggregates at high temperature. Bottom panels: Singular value basis functions and denaturation traces obtained form the top plot. The second SVD component shows two-state like (sigmoidal) denaturation in the bottom left panel.

(e.g. temperature, denaturant, or pressure scan), the resulting spectral shifts can be decomposed into separate species. Singular value decomposition [43], spectral cross-correlation [44], and other algebraic techniques are routinely used to obtain such spectral data. Figure 1 shows amide I′ IR spectra of the lambda repressor fragment $\lambda_{6\text{-}85}$ mutant Y22W/Q33Y/A37G/A49G, as well as the first and second singular value components and their temperature trend [45]. Singular value decomposition decomposes a series of spectra (here as a function of temperature) into ranked basis functions. $\lambda_{6\text{-}85}$ is a five-helix bundle; the declining peak at 1650 cm^{-1} indicates loss of helix upon thermal unfolding; the shoulder at 1670 cm^{-1} indicates increased coil structure. There is also a sharp peak below 1620 cm^{-1}, caused by heat-induced aggregation of this particular mutant above 50°C. The spectral density shifts arising from interchange of helices, sheets and coils can be simulated by molecular dynamics simulations comparing

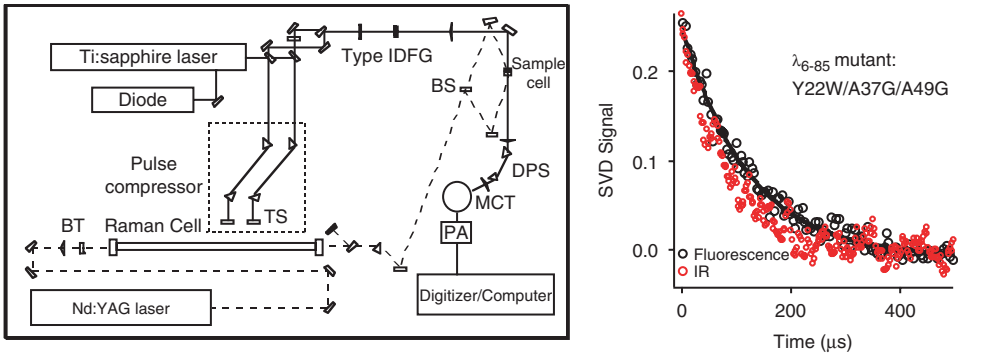

Figure 2. Left: Experimental set-up for single shot kinetics detected by nanosecond-resolved transient IR spectroscopy: a near-IR fs oscillator produces short pulses at two wavelengths, which are chirped in prism pairs and then mixed down to $1650 \pm 100\,\mathrm{cm}^{-1}$. The IR light passes through the sample and a prism monochromator before being detected by a 32 channel array detector. Right: Comparison of the IR (light circles) and fluorescence (black circles) detected kinetics of the $\lambda_{6\text{-}85}$ mutant Y22W/A37,49G.

the folded and unfolded states [46]. Formation of large aggregates is still currently beyond the computational power of molecular dynamics simulations.

Our research has focused on using infrared spectroscopy as a kinetic probe for large-amplitude motions of biomolecules: infrared spectra are collected at time slices as the molecule evolves, thus monitoring conformational changes as in the thermodynamic case, but with time resolution [47]. The large scale folding and functional motions of biomolecules, in the ns to μs range, are much slower than the vibrational motions themselves, which just adiabatically follow the macromolecular motion. Several experimental techniques have been developed to monitor kinetics via infrared spectra. They use either Fourier-transform spectroscopy [48–51], wavelength scanning of high-resolution lasers (the kinetics being induced at each wavelength to produce a time-evolving spectrum) [30], or intense broadband IR sources to obtain full spectra in a single shot while kinetics are acquired [47].

We discuss in detail the latter, developed in one of our labs [47]. Figure 2 shows the set-up used to obtain transient IR spectra of folding proteins or nucleic acids: a broadband near-IR femtosecond laser is operated simultaneously at two wavelengths, 750 and 860 nm. These essentially jitter-free pulses are then mixed with one another to generate a $200\,\mathrm{cm}^{-1}$ wide mid-IR pulse centred at $1650\,\mathrm{cm}^{-1}$. After the pulse is absorbed by the sample, a multichannel infrared array detector detects the full dispersed IR spectrum with up to 150 ns time resolution. Thus the change of various bands characteristic of secondary structure can be monitored in the amide I′ fingerprint region of a protein, for example. Figure 2 shows a comparison between kinetics detected by infrared spectroscopy and by fluorescence spectroscopy for the Y22W/A37G/A39G mutant of lambda repressor fragment $\lambda_{6\text{-}85}$. The IR transient monitors in real time the decline of the α-helical band at $1650\,\mathrm{cm}^{-1}$ discussed above. This particular molecule is nearly a two-state folder, as the IR and fluorescence kinetics are almost identical. However, faster-folding mutants of this protein do not obey two-state kinetics, instead folding downhill on a rough free energy surface. The spectral shift from 1650 to

$1670\,cm^{-1}$ indicates that the hydrogen bonding is switching from helix–helix bonds to helix–solvent bonds.

2.2. *THz absorption spectroscopy*

THz spectroscopy inherently monitors larger amplitude and more delocalized motions than IR spectroscopy, on a picosecond rather than a 50–100 femtosecond time scale. It also can be used in a 'dynamical' mode to study such motions directly, and in a 'kinetic' mode where these motions report on even larger scale complete folding and unfolding of biomolecules occurring on a 10 ps or longer time scale. Steady-state THz spectra in aqueous solvent are sensitive to the dynamics of the solvent shell around biomolecules, allowing solvation water and bulk water to be distinguished.

The THz region is currently a technologically difficult region between microwave and infrared spectroscopy. By developing new laser radiation sources and sensitive detectors, a new experimental window between photonics and microwave technology has been opened. The novel THz sources include: time domain THz systems, quantum cascade lasers, the p-doped Germanium laser (p-Ge laser), THz generation by photo mixing, and the free electron laser [31, 52–59]. The basic features and first applications of THz time domain spectrometers (THz-TDS) have been discussed in detail in a review by [60].

Future applications of THz spectroscopy for the study of protein dynamics aim in two directions:

THz spectroscopy is of special interest for the spectroscopy of proteins since it covers the region of the collective motions in proteins, including large-amplitude motions of the entire protein or of large parts of it. This could include breathing modes of helices, local displacements of tertiary packing, or movements of groups of amino acid residues. Terahertz (THz) absorption spectroscopy should be able to probe theses modes directly in unfolded, partially folded, and folded proteins. Coupled with site-directed mutagenesis and fast relaxation techniques, it is expected to provide site-specific information on the time evolution of collective modes during folding or function.

In addition to directly probing the motions on a ps-to sub-ps time scale, THz spectroscopy accesses global solvent network changes by probing the re-orientational and translational dynamics of the solvent [61]. As has been shown in simulations, and discussed below, the relaxation of the protein–water hydrogen network, in particular due to solvent translational displacement, is largely responsible for protein structural relaxation [52]. We discuss below how relevant motions of the solvent can be directly probed by THz spectroscopy, which turns out to be very sensitive to changes in the water network dynamics.

Due to the high absorption coefficient of water (increasing from ca. $\alpha = 150\,cm^{-1}$ at 1 THz to ca. $\alpha = 450\,cm^{-1}$ at 3 THz) the transmitted power decreases rapidly with increasing frequency which makes it an experimentally challenging frequency range. Using low-power radiation sources it is simply impossible to penetrate thick layers of water. As a result, many previous studies have been restricted to biomolecules embedded in polyethylene, which is transparent in the THz region [62], hydrated films of biomolecules [63], and only rarely hydrogen-bonding solvents [61, 64]. It is well known that solvation water near the surface of a biomolecule exhibits dynamics distinct

from that of bulk water [65]. For the study of biomolecules in water powerful THz radiation sources, such as the p-doped Germanium (p-Ge) laser [66–68], will be essential.

Using a p-Ge laser spectrometer [67], which combines the important elements of high power, tunability, and broadband emission in a table top experiment, we demonstrated the possibility of accurately measuring absorption of liquid samples, and showed that it is possible to penetrate thick layers up to 150 μm of liquid water or buffer [66, 68, 69]. It is thus possible to study proteins in their natural surrounding.

Figure 3 shows the experimental set-up of the p-Ge spectrometer: The p-Ge laser serves as radiation source and is hosted in a closed cycle 3 K cryostat. It is tuneable between 40 and 140 cm^{-1} by varying the crossed magnetic and electric fields that are applied to the crystal. Even for fixed fields, the emitting frequency will change during the laser pulse because the temperature of the crystal increases during the pulse. Such transient temperature tuning results in broadband emission in an interval of about 10 cm^{-1} around the centre frequency. The radiation is passed through the sample when varying the sample length between 10 and 250 μm. The frequency dependence of the transmitted radiation is measured using a grating monochromator. Photoconductors can be used for the THz detector, such as a photoconducting aluminium-doped germanium crystal placed on the cold plate of a liquid helium cryostat. Bolometers are also suitable as detectors.

Figure 3 shows the absorption spectrum of liquid water in the THz region. We observe a monotonic increase in absorption with increasing frequency. This increase is part of a broader absorption feature of water with a peak centred at 190 cm^{-1}, which has been attributed to either the stretches of the hydrogen bonds alone or from translational–orientational motions [70]. When we study the concentration dependence of the absorption spectra of solute–solvent systems such as lactose in water or proteins in buffer we find a more complex behaviour due to the direct response of the solute on the solvent dynamics. THz spectra of such mixtures cannot be analysed any longer in terms of a two-component model. From a thermodynamic point of view, these systems are really effective media, and an effective medium approach, such as the Clausius–Mossotti approximation and its extensions to composite media and solutions [71, 72], could potentially be applied here. However, the solvation water around a biological molecule such as lactose is known to exhibit distinct dynamical properties compared to bulk water. Molecular dynamics simulations can directly probe this microscopic picture, as discussed below, permitting analysis at the molecular level with separate contributions of the solute, the bulk solvent and the solvation shell. It was found that it is possible to decompose the distinct contributions of biomolecule, solvation water, and bulk water, and directly determine the size of the solvation shell and its specific THz absorption. THz spectroscopy is turning out to be a sensitive experimental tool for detecting solute induced changes in the water network.

2.3. *SAXS*

Of the three techniques discussed here, SAXS studies the largest scale rearrangements of biomolecules, typically on a nanometre and upward length scale. It can also be used

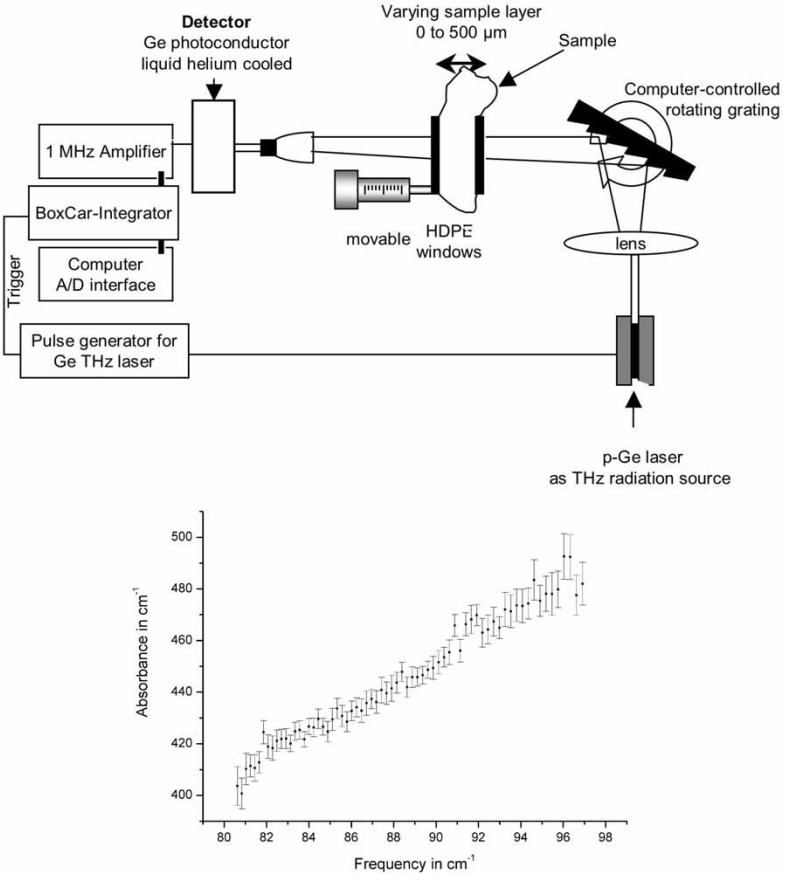

Figure 3. Top panel: Terahertz absorption spectrometer. The radiation emitted by the p-Ge laser is sent through an adjustable path-length sample and dispersed by a rotatable grating onto the p-Ge detector. A Beer's law scan of the sample path-length is used to extract the absorption coefficient. Bottom panel: Terahertz absorption coefficient of bulk water obtained with the apparatus shown above.

to study unfolded proteins and the relationship of their sequence length to radius of gyration [73]. Although the dynamical time scale of SAXS is in principle sub-femtosecond, the large amplitude motions it indirectly studies often occur on a microsecond to millisecond time scale because they correspond to the longest length scales discussed here [74]. In practice, SAXS beamlines are often configured to produce nanosecond duration bunches, averaging over a set of instantaneous configurations of the sample on the nanosecond time scale.

In a SAXS experiment, the solvated biomolecule is subjected to an X-ray beam, typically $\lambda \approx 1\,\text{Å}$ in wavelength, and the scattering signal is recorded. Unlike crystallography experiments, molecules are randomly oriented and produce a cylindrically symmetric diffraction pattern. Since distance and scattering angle are

Fourier-conjugate variables, small scattering parameters

$$q = \frac{4\pi \sin\theta}{\lambda} \tag{2.1}$$

correspond to large molecular length scales. Folding/unfolding signals for protein or RNAs typically contribute in the region $q \approx 0.04\,\text{Å}^{-1}$. As shown by Guinier [11], the scattering intensity for an object with radius of gyration R_g scales as

$$\ln\frac{I}{I_0} = -\frac{(qR_g)^2}{3}. \tag{2.2}$$

At small values of q, deviations may occur because of inter-protein interactions (aggregation); at large values of q, deviations occur because higher moments of the shape contribute. A modern complement to the simple Guinier analysis is to model the shape of a biomolecule via finite elements, then fit the finite-element shape to SAXS data to obtain shape information beyond the average radius [75]. Much progress has been made in steady-state applications, but kinetics measurements do not yet have sufficient signal-to-noise ratio for such analysis. However, the advent of a new generation of bright sources, such as the Advanced Photon Source at Argonne National Laboratory, will make kinetic resolution below the mean radius possible in the near future.

Figure 4 illustrates the acquisition of data showing large-amplitude shape changes of the Y22W mutant of lambda repressor fragment $\lambda_{6\text{-}85}$, our example protein molecule [76]. A stopped flow apparatus is placed in the vacuum lines guarding the X-ray beam, collimated to about $100 \times 100\,\mu$m at the sample. Water strongly absorbs X-rays, so cells with <1 mm path length and $\leq 50\,\mu$m sapphire windows are required. The data show that the mutant compactifies from an extended state (24 Å, the random coil would be at 26 Å) to a native radius of gyration (the native value from the X-ray crystal structure is 14 Å, slightly smaller) with a half-life of about 300 ms. Considering the experiments were carried out in a cryosolvent of much higher viscosity than water, this rate is much faster than that extrapolated from room temperature data of the same protein, showing that low temperature speeds up folding when viscosity is factored out.

A slight discrepancy between dynamically obtained R_g and 'dry' R_g from crystal structures is common in SAXS measurements [77]. It reflects scattering of solvation shell water molecules that differ from bulk water. The thickness of the water layer thus obtained around a protein's surface residue typically is about 3 Å. The upper end of that range is compatible with solvation layer thicknesses around carbohydrates discussed in section 4. Programmes for the estimation of solvation water layers from SAXS data have been developed, and generally yield good agreement between experiment and 'dry' structures [78]. It will be very interesting to compare careful SAXS, and THz measurements, as the dynamical time scales of the two experiments (and others, such as NMR) cover many orders of magnitude: nanoseconds for SAXS

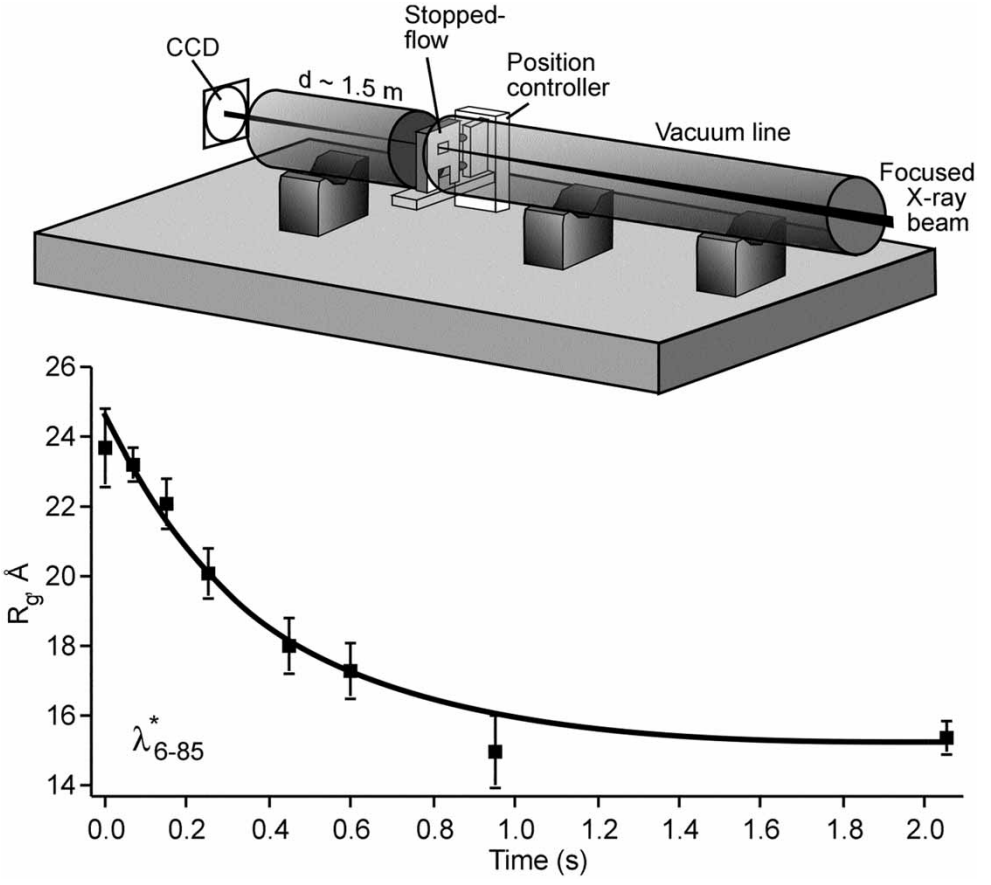

Figure 4. Top: Small-angle X-ray scattering set-up for kinetics detection: a stopped flow apparatus with 5 ms time resolution is placed in the collimated X-ray beam that propagates through vacuum pipes. A CCD array detector ca. 1.5 m from the sample detects the scattered X-rays. Bottom: Kinetic trace of the $\lambda_{6\text{-}85}$ mutant Y22W obtained by SAXS. The radius of gyration of this protein smoothly decreases from the near random coil value (25 Å) to the native value including solvation layer (14 Å; bare protein: 11 Å).

averaged over a pulse train, and ps for direct THz electric field detection described in section 2.2. This will provide important clues for the length scale over which solvation layers remain static, vs. the length scale over which water molecule dynamically interchange.

3. Theoretical tools

A number of theoretical tools aid in describing biological molecules over a variety of length and time scales. These are often based on and at least begin with molecular dynamics (MD) simulations of the biomolecule. Over short length scales, quantum

mechanical methods may be needed to describe the transfer of energy from localized vibrational modes to other parts of the biomolecule. We discuss computation of the vibrational density of states and the absorption spectra of biomolecules, important for interpreting IR and THz spectra. We then discuss methods for computing rates of vibrational energy transfer from the vibrational states quantum mechanically, and compare with results of pump-probe IR spectroscopy. Over intermediate length scales, it is important to also address properties of the solvent molecules, how they interact with the biomolecule and with each other in the vicinity of the biomolecule. We shall summarize methods we have used to track the hydrogen bond reorganization between water molecules around a carbohydrate or protein, which is crucial for interpreting the THz spectra of solvated biomolecules. Finally, we discuss the scaling properties of biomolecular geometry, which are intimately connected with protein dynamics and energy flow over both short and long length and time scales. The fractal nature of protein molecules governs the subdiffusive and anisotropic flow of energy throughout a biomolecule on time scales of tenths to tens of picoseconds, as well as the fluctuations in the distances between protein residues that can occur over μs time scales or longer.

3.1. *Molecular dynamics simulations*

Molecular dynamics (MD) simulations, coupled with other computational techniques including analysis of vibrational modes (normal mode analysis, NMA) (see, for example, [79]), principal components (see, for example, [80–84]), or local flexibility within a biomolecule [46, 85], have for some time provided a wealth of information about biomolecular dynamics [86, 87]. For the purposes of this review, we consider the application of MD simulations and other computational techniques for analysing dynamics of biomolecules with focus on areas most closely connected to THz and IR spectroscopy, including calculation of the spectral density, energy flow in biomolecules, analysis of biomolecule–water interactions, and water–water interactions near a biomolecule. We also discuss connections to biomolecular geometry, i.e., the fractal nature of biomolecules, relevant to measurements over short and long length and time scales.

3.2. *Spectral density*

The spectral density is connected to the vibrational density of states of the biomolecule, which is often convenient to estimate by a normal mode analysis [18–21, 88]. Corrections due to anharmonicity can also be included [19, 20]. As an example, consider the vibrational density of the small helix-bundle protein $\lambda_{6\text{-}85}$, shown in figure 5. A number of peaks in the density are seen at higher frequency, above $3000\,\mathrm{cm}^{-1}$, corresponding to CH, OH and NH stretches, and in the IR, for instance the amide I vibrations, from about $1600\,\mathrm{cm}^{-1}$ to $1700\,\mathrm{cm}^{-1}$. Pump-probe measurements in recent years have provided much information about energy flow from the amide I′ band of peptides and proteins [29, 39–41], via analysis of line broadening and shifts. In the future, 2-D IR experiments will explore homogeneous and inhomogeneous broadening of IR lineshapes further. Moving down to THz frequencies, below about $100\,\mathrm{cm}^{-1}$,

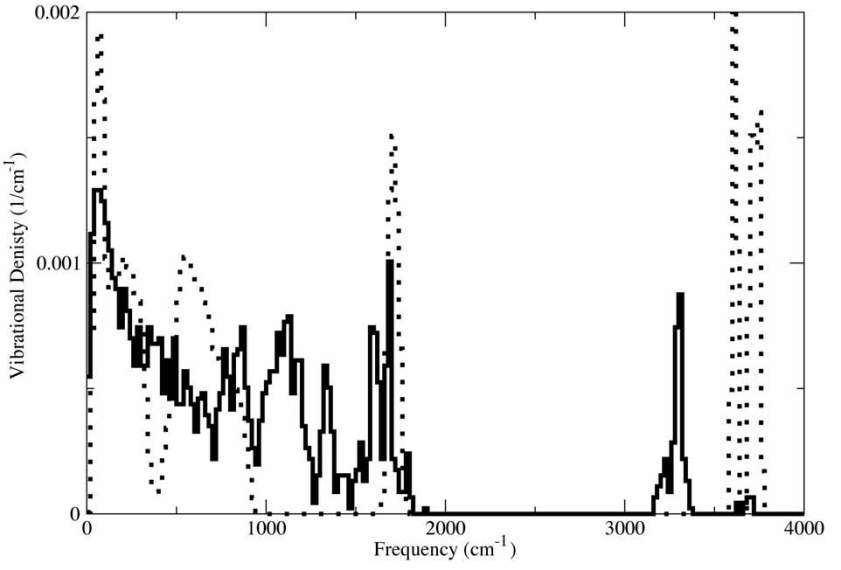

Figure 5. Histogram of the vibrational density of states of $\lambda_{6\text{-}85}$ (solid) and TIP3P water (dotted) for comparison.

we see a relatively high density of protein vibrations that lie in this 'fracton' region (see subsection 3.5).

Absorption spectra for biomolecules have been computed, oftentimes using the vibrational modes of the molecule [10, 19–21]. In terms of the vibrational modes, the IR absorbance for a transition in mode α, with n_α vibrational quanta to $n_\alpha + 1$ vibrational quanta can be estimated in terms of the mode specific dipole derivative [20, 21]. We take the charges on the atoms to be fixed. The absorbance, A_α, from the ground state for a mode of frequency ν_α is then proportional to $\nu_\alpha(d_x^2 + d_y^2 + d_z^2)$, where the transition dipole moments d_x, d_y, and d_z are calculated by expansion in normal mode transition moments, $\langle \Psi_\alpha^{(1)} | Q_\alpha | \Psi_\alpha^{(0)} \rangle$. Due to the extensive inhomogeneous broadening of protein vibrational spectra, the results are broadened with Gaussians that have a full-width at half-maximum of a few wave numbers.

In figure 6 we plot the absorption spectrum so calculated for solvated $\lambda_{6\text{-}85}$ together with the absorbance of a comparable volume of bulk water to 90 cm^{-1}. In the harmonic approximation the solvated protein is seen to absorb more strongly than the water over this frequency range. This trend can be reversed when an harmonicity is included. We also plot the spectrum of the dehydrated protein and a similar volume of bulk water, which we observe to be similar up to 90 cm^{-1}. The absorbance of solvation water is thus different from that of bulk water. We have found this to be the case for the solvation water around lactose [10], too, which is also shown in figure 6. We find that the absorbance of the solvation water around lactose is greater than that of the same number of water molecules in the bulk due to collective, coherent oscillations of the solvation water and biomolecule in the THz region that are absent in bulk water. This is in good agreement with the experimentally obtained THz spectra, as can be seen in the bottom of figure 6 and will be discussed in more detail in section 4.

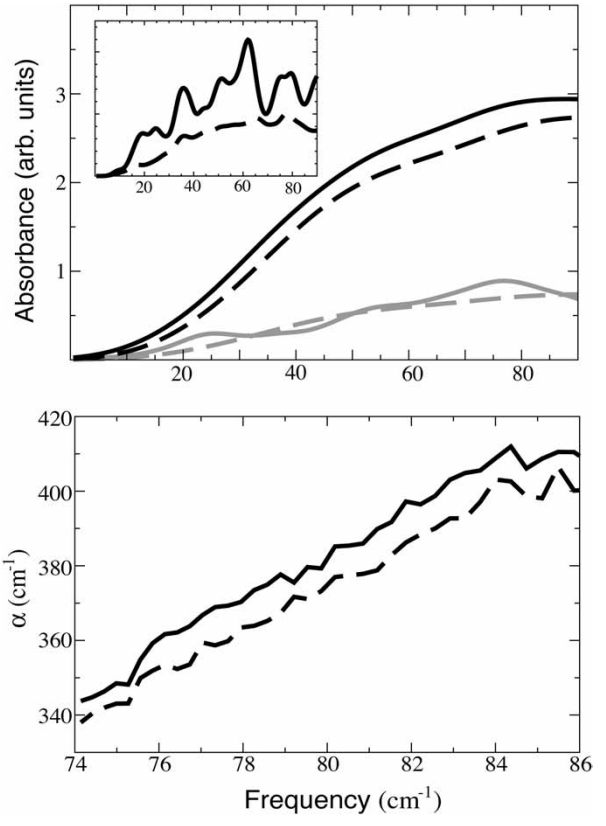

Figure 6. Top panel: Absorbance calculated in the harmonic approximation for the solvated five-helix bundle $\lambda_{6\text{-}85}$ (black solid curve) and for a comparable volume of bulk water (black dashed) are plotted together with the absorbance calculated for dehydrated $\lambda_{6\text{-}85}$ (grey solid) and a comparable volume of bulk water (grey dashed). The inset shows the absorbance of solvation water within 7 Å of a lactose molecule (solid) compared to the absorbance for the same number of bulk water molecules (dashed). Bottom panel: Measured THz absorption of aqueous protein as a function of frequency for buffer solution (magnesium acetate, pH = 7.32) at 3.41 mg/ml concentration of pseudo-wild-type λ-repressor (M.W. ≈ 9.9 kDa). Dashed curve is the measured absorption of the buffer solution without protein.

3.3. *Computational studies of energy flow*

Energy flow in proteins has been studied using a variety of approaches, including MD simulations with the aim of describing specific pathways for energy to flow from a cofactor or binding site through the protein [89–93], or describing generic properties of energy flow in biomolecules [16, 90, 94, 95]. Quantum mechanical approaches have also been used to determine how fast and by what pathway energy flows among vibrational states [16–18, 96].

Quantum methods are generally needed to compute vibrational lifetimes of relatively high-frequency modes of biomolecules. Following a golden rule approach that has been used for solid state systems [97], we can address the influence of low-order anharmonic interactions on the lifetime of the vibrational modes, starting with the cubic terms,

an assumption that is valid at sufficiently low temperatures and often for energy transfer from the relatively high-frequency modes. The energy transfer rate from mode α, W_α, can be calculated as the sum of terms that can be described as decay and collision [97], the former typically much larger except at low frequency where both terms are comparable. The anharmonic decay rate of mode α is $W_\alpha = W_\alpha^{decay} + W_\alpha^{coll}$, where

$$W_\alpha^{decay} = \frac{h\pi}{8\omega_\alpha} \sum_{\beta,\gamma} \frac{|\Phi_{\alpha\beta\gamma}|^2}{\omega_\beta \omega_\gamma} (1 + n_\beta + n_\gamma)\delta(\omega_\alpha - \omega_\beta - \omega_\gamma), \tag{3.1}$$

$$W_\alpha^{coll} = \frac{h\pi}{4\omega_\alpha} \sum_{\beta,\gamma} \frac{|\Phi_{\alpha\beta\gamma}|^2}{\omega_\beta \omega_\gamma} (n_\beta - n_\gamma)\delta(\omega_\alpha + \omega_\beta - \omega_\gamma), \tag{3.2}$$

where n_i is the occupation number of mode $i = \alpha, \beta, \gamma$, which at temperature T we take to be $n_i = (\exp(h\omega_i/k_B T) - 1)^{-1}$. The matrix elements $\Phi_{\alpha\beta\gamma}$ appear as the coefficients of the cubic terms in the expansion of the interatomic potential in normal coordinates, computed numerically as

$$\Phi_{\alpha\beta\gamma} = \frac{1}{2\delta Q_\gamma} \left(\partial^2 V / \partial Q_\alpha \partial Q_\beta |_{Q_0 + \delta Q_\gamma} - \partial^2 V / \partial Q_\alpha \partial Q_\beta |_{Q_0 - \delta Q_\gamma} \right) \tag{3.3}$$

where $Q_{\alpha,\beta,\gamma}$ are mass-weighted normal coordinate, and Q_0 is the equilibrium position of the protein in normal coordinates. As mentioned above, low order anharmonic terms have also been computed for small proteins to evaluate the contribution of anharmonicity to vibrational spectra [19, 20]. Other, non-Markovian perturbative techniques have been applied recently to the study of vibrational energy transfer in proteins and peptides [17].

Energy transfer rates computed with such quantum perturbative approaches have in general agreed well with the fast times observed by pump-probe IR measurements [16–18, 96]. For example, the vibrational energy transfer rates of roughly 1 ps measured for the amide I′ band, discussed above, have been reproduced by these calculations, as has the small temperature dependence of the rate. The calculations and interpretation of them reveal a fairly weak temperature dependence for rates of vibrational energy transfer from most relatively high-frequency modes of proteins [16, 18, 96], not just the amide I band, indicating that little energy flows directly from these modes into the low-frequency modes of the protein, but instead first to vibrational modes of several hundred wave numbers that correspond to other vibrations of the peptide bond.

3.4. *Hydrogen bond dynamics in solvation shells*

A number of computational studies have addressed how the structure and dynamics of water in the bulk differs from that of water near a biomolecule and how water activates dynamics of a protein molecule [1, 22, 23, 81, 96, 98–104]. Recent THz measurements on biomolecules [10, 21, 105, 106] indicate that solvation water contributes to the

observed spectra. This work has led to the analysis of water dynamics in the neighborhood of the solute, specifically a hydrogen bond correlation function, $c(t) = \langle h(0)h(t)\rangle/\langle h\rangle$, similar to others proposed earlier [22, 107, 108], which is a measure of lifetimes of hydrogen bonds due to their reorganization. Here $h(t)$ is a hydrogen bond population operator, which is 1 for a given hydrogen bond between a donor (D)–acceptor (A) pair at time t, and 0 otherwise; the angular brackets denote an average over all D–A pairs. The function $c(t)$ is the probability that a random D–A pair that is hydrogen bonded at time 0 still has the same hydrogen bond at time t, regardless of whether or not the bond was broken at intermediate times. The lifetime of a tagged hydrogen bond, τ_{HB}, is defined by $c(\tau_{HB}) = e^{-1}$ (although the decay is not usually single-exponential). We have adopted commonly used geometric criteria for hydrogen bonds, where bonds are formed when the distance between D and A is less than 3.5 Å and the angle between donor, hydrogen, and acceptor (D–H–A) is greater than 150°.

MD simulations carried out by us on the solvation dynamics near saccharides were performed with the MD simulation programme MOIL [109], which uses a combination of AMBER, OPLS, and CHARMM force fields. For our study of the lactose molecule [10], we obtained force field parameters from the CHARMM version 22 parameter set [110]. Details of the MD simulations have been provided elsewhere [10]. Those simulations found τ_{HB} for bonds between water molecules within 6 Å of lactose to be distinctly larger than for bonds between water molecules in the bulk. We have similarly computed the hydrogen bond correlation time for bonds between water molecules near $\lambda_{6\text{-}85}$. These MD simulations have been carried out using the GROMACS 3.2 package [111, 112], with the GROMOS96 force field [113], where the protein is embedded in a box containing simple point charge (SPC) model water [114]. For $\lambda_{6\text{-}85}$ we find, as seen for lactose, that the reorientation dynamics of hydrogen bonds between water molecules is significantly slower than for bulk water out to about 6 Å from $\lambda_{6\text{-}85}$, as we show in figure 7. Specifically, we find a reorientation time of 1.5 ps for bulk water, compared to 1.9 ps for hydrogen bonds between water molecules within 6 Å of $\lambda_{6\text{-}85}$.

3.5. *The fractal nature of biomolecules*

The fractal geometry of large biomolecules provides a unified framework for describing protein dynamics over many length and time scales. Studies of energy flow in proteins have for some time revealed anisotropic flow [89, 91–93], which is a general feature of energy flow on a fractal object. Indeed, evidence that protein molecules may be characterized by fractal geometry has appeared in a variety of measurements, such as the anomalous temperature dependence seen in spin echo experiments [115, 116]. Recently Granek and Klafter have shown that structural fluctuations between different parts of a protein [2], seen in single molecule experiments [117–119], can be described by exploiting connections between the fractal geometry of a protein molecule and dynamics, properties that are related by the "fracton" models of Alexander and Orbach [120, 121].

To address the connection between vibrational energy flow and the fractal geometry of a biomolecule, we consider how the mean square displacement of vibrational energy, $\langle R^2\rangle$, varies with time, as, for example, a vibrational wave packet

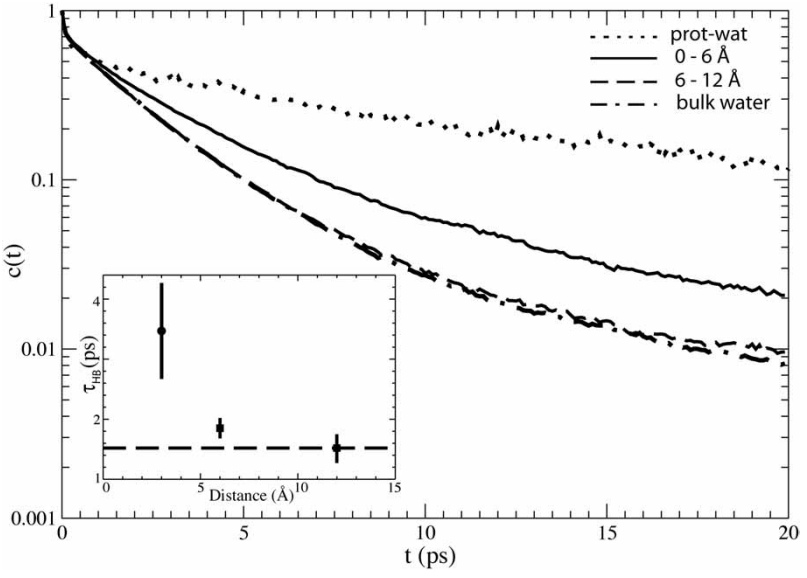

Figure 7. Hydrogen bond correlation function, $c(t)$, for hydrogen bonds between water molecules within 6 Å of $\lambda_{6\text{-}85}$ (solid curve) and between 6 and 12 Å from $\lambda_{6\text{-}85}$ (dashed). The result for bulk water (dot-dashed) lies just below the dashed curve. Also shown is $c(t)$ for hydrogen bonds between the protein and water (dots). The inset shows the correlation times, τ_{HB}, for bonds between protein and water (circle, shown at 3 Å) and for bonds between water molecules to 6 Å and between 6 and 12 Å from $\lambda_{6\text{-}85}$ (squares). Error bars corresponding to two standard deviations are indicated. The dashed line is the correlation time for hydrogen bonds between water molecules in the bulk.

spreads through the molecule following an impulsive excitation. For a diffusive process,

$$\langle R^2 \rangle \sim t^{2\nu}, \tag{3.4}$$

where $\nu = 1/2$ corresponds to normal diffusion. For proteins one finds $\nu < 1/2$ so that the process is subdiffusive, i.e., the spread of energy is less rapid than a Brownian process. Alexander and Orbach [120] and Rammal and Toulouse [122] showed that

$$\nu = \frac{\bar{d}}{2D}, \tag{3.5}$$

where D is the fractal dimension of the object, which describes how the mass, M, scales with length, L, $M \sim L^D$; and $\bar{d}$ is the spectral dimension, which characterizes how the vibrational mode density varies with mode frequency, $\rho_L(\omega) \propto \omega^{\bar{d}-1}$. The spectral dimension, which reflects the connectivity or bonding between atoms of the object [123, 124] and is generally smaller than the fractal dimension of the object, can be found by computing the vibrational density of states. As an example, we show in figure 8 the

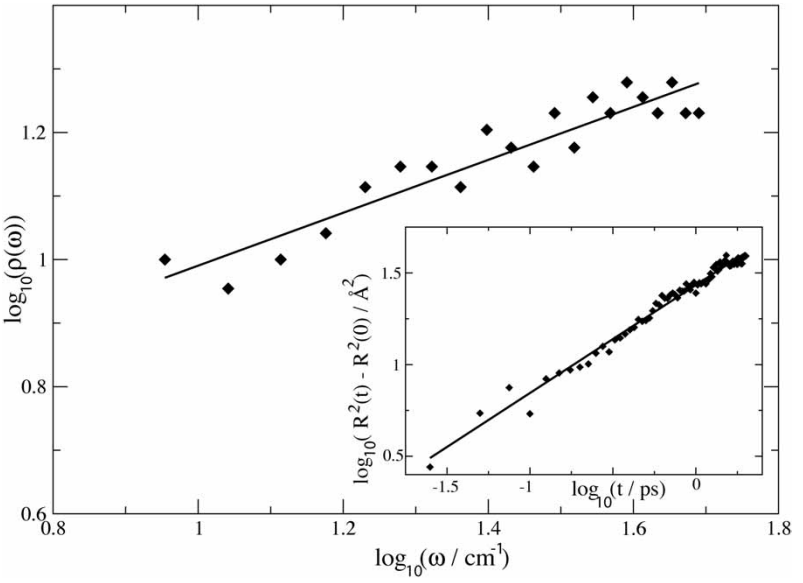

Figure 8. Log of the vibrational density vs. log of the vibrational frequency for green fluorescent protein (GFP) to about 2 THz. A power-law variation is seen, where the slope of the line fit to the data gives $\bar{d}-1$, where $\bar{d}$ is the spectral dimension. For GFP we find $\bar{d}=1.45$. The fractal dimension is $D=2.43$ [59]. The variance of a wave packet spreads with time as $\langle R^2\rangle \sim t^{2\nu}$, where $\nu=\bar{d}/2D=0.30$. Since ν is smaller than 0.5, energy flows subdiffusively. The inset shows results of energy flow simulations, where subdiffusive energy flow is observed with $\nu=0.30$. The spectral dimension and fractal dimension also determine the power-law decay of the autocorrelation function for the distance between two residues of a protein over much longer time scales [2].

dependence of the vibrational density of states on mode frequency in the THz region for green fluorescent protein (GFP), which has fractal dimension $D=2.43$ [94]. The result was obtained by normal mode analysis. The scaling appears as expected for a fractal object, yielding a spectral dimension for GFP, $\bar{d}=1.45$. We also show in figure 8 how $\langle R^2\rangle$ obtained from simulations varies with t, where we observe in the figure subdiffusive flow of energy, with $\nu=0.30$, which indeed corresponds to $\bar{d}/2D$. Values of $\bar{d}$ and D have been computed for many proteins. In a recent study of 58 proteins ranging from 100 to 3600 residues, Burioni *et al.*, computed $\bar{d}$ and found it to range from about 1.3 to 2.0, appearing to increase logarithmically with protein size [125]. The fractal dimension, D, can be found by examining how the mass contained inside a sphere varies with radius, L. D has been computed for 200 proteins from the protein data bank with 100 to 10 000 residues and found to be on average $\approx$2.5 [39]. Since the spectral dimension was found to be $\approx$1.3 [125] and $D\approx 2.3$ [94] for proteins of $\approx$100 amino acids, then $\nu\approx 0.28$ in this case. For proteins with at least 1000 amino acids, $\bar{d}\approx 2$ [125], and $D\approx 2.6$ [94], so that $\nu\approx 0.38$.

The fractal nature of large biomolecules and the corresponding subdiffusive and anisotropic flow of energy have been observed in molecular dynamics simulations [91, 93, 94], such as that shown in figure 8. There are also consequences over much longer times, not directly accessible to simulation. Granek and Klafter [2] showed recently that the autocorrelation function for the distance between two residues of a

protein, which can be measured by single molecule experiments [117–119], exhibits a power-law decay, where the power again depends on D, $\bar{d}$ and their ratio.

4. Two cases: carbohydrate solutions and protein folding/solvation dynamics

4.1. *Carbohydrates*

We now consider two cases where experiments and computational modelling have mapped out dynamics on multiple length and time scales. In the first case, a small biomolecule, a simple sugar, is examined in aqueous solution. Molecular dynamics simulations allow a microscopic interpretation of terahertz absorption experiments in terms of a solvation shell that extends well beyond the carbohydrate. In the second example, a small protein, $\lambda_{6\text{-}85}$, is examined by IR, THz, SAXS and MD techniques to reveal directly solvation dynamics, as well as kinetics of large-amplitude motion leading to folding and unfolding. Analysis of molecular dynamics simulations sheds light on the observed dynamics.

The special importance of the solvent on protein dynamics has been discussed in context with sugars. Solvated mono- and disaccharides are found to have a significant stabilization effect on proteins and membranes in a dehydrated or frozen state [126, 127]. Stability in extreme conditions has been found to increase with increasing trehalose concentrations. It was speculated that slowing of the water dynamics induced by the carbohydrates might lead to a reduced mobility of proteins, though the exact underlying molecular mechanism is not yet understood. One possible mechanism is that the sugar substitutes for water in the hydrogen bond network near the protein and thereby helps preserve the secondary and tertiary structure [128]. A second explanation could be that sugars influence the dynamics of the surrounding water, which in turn interacts with the protein. In a more glassy state, the molecular motions slow down, affecting the protein dynamics inasmuch as they are coupled to the solvent dynamics. This idea is supported by a recent combined experimental and modelling study which presented indirect evidence that certain trehalose motions are coupled to those of the solvation shell [129, 130]. When the time-dependent fluctuations of protein structure were probed by the dephasing of the heme-bound CO, the dynamics on time scales of tens of picoseconds were found to be significantly changed in dry trehalose glasses compared to an aqueous surrounding.

In order to further investigate the question as to whether water dynamics is affected by the presence of a carbohydrate molecule we have recently carried out measurements in the THz frequency range, measuring the absolute absorption coefficient of a lactose solution as a function of the lactose concentration. The individually determined absorption coefficients could be fitted to a three-component model, which assumes a distinct absorption coefficient of the water around the sugar due to the distinct properties of solvation water. For details see [10].

This analysis, combined with molecular dynamics modelling, seeks a direct molecular-level explanation for carbohydrate solvation, in contrast to dielectric effective medium approaches previously used to interpret low-frequency measurements. We model the total absorption as a volume-weighted average of the absorption of the

solute, the solvation water, and the bulk water:

$$\alpha_{total}(\omega) = \frac{V_{solute}}{V}\alpha(\omega)_{solute} + \frac{V_{sw}}{V}\alpha(\omega)_{sw} + \left(1 - \frac{V_{sw}}{V} - \frac{V_{solute}}{V}\right)\alpha(\omega)_{bulk\ water} \qquad (4.1)$$

with $\alpha_{total}(\omega)$, $\alpha(\omega)_{solute}$, $\alpha(\omega)_{sw}$, $\alpha(\omega)_{bulk\ water}$, being the absorption coefficients of the solution, the absorption coefficient of the solute, the absorption coefficient of the water in the solvation layer, and the absorption coefficient of the bulk water, respectively. V, V_{solute}, V_{sw} are the total volume probed by the THz beam, the volume of solute molecules, and the volume occupied by the solvation shells.

The contribution of the solute to the total absorption coefficient, as well as the decrease in total absorption due to the water volume displaced by lactose, scale linearly with increasing lactose concentration. The contribution of the solvation water does not. For very dilute solutions an increase in the concentration of the solute will result in a linear increase of solvation water. Yet as soon as solvation shells start to overlap, one expects a deviation from this linear behaviour. The onset is especially sensitive to the size of the layer and allows one to determine the size of the solvation layer. If we increase the concentration even far beyond this point, the absorption of the solute will finally dominate the total absorption.

In figure 9 we display for each lactose concentration the mean value of all experimentally determined THz absorption coefficients at a fixed frequency ($80\,cm^{-1}$), together with their statistical error (the error of individual measurements is greater). The straight line corresponds to a linear change of absorption coefficient, as would be expected if we take the known absorption coefficient of bulk water ($\alpha = 385.3\,cm^{-1}$ at $\nu = 80\,cm^{-1}$) and lactose, and predict the resulting absorption coefficient by just adding these two components. The best-fit curve, which shows the results of the fit to a three-component model, including the solvation water resembles the measured data much better. These data allow us to deduce the size and the absorption coefficient of the solvation water.

The frequency dependent absorption coefficient for bulk water and the (fitted) solvation water are also displayed in figure 9. We see that the absorption of the solvation water is greater than that of bulk water. This finding is qualitatively consistent with the absorbance calculated in the THz region for water in the bulk and water in the solvation layer around lactose, as shown in figure 6 [10]. Examination of the dynamics of the solute–solvent interactions reveals that the enhancement of the absorbance of water near lactose is due to coherent oscillations of the solute and nearby water molecules in the THz region that are absent in bulk water. The volume of the solvation water, e.g., the water which has an increased absorption coefficient for motions on the ps and sub-ps time scale is determined to be $3660\,Å^3$. This corresponds to 123 water molecules per lactose molecule. The solvation shell thereby extends ($5.13 \pm 0.24\,Å$) from the surface of the lactose molecule, well beyond the first solvation shell [10]. The extent of the solvation layer deduced from the THz spectra is consistent with the influence of hydrogen bond reorientation dynamics deduced by molecular dynamics simulations. We find the hydrogen bond correlation time to be larger for water molecules within $6\,Å$ of lactose than for water molecules in the bulk, similar to the case for the water around a protein shown in figure 7. This measure of the solvation layer is consistent with

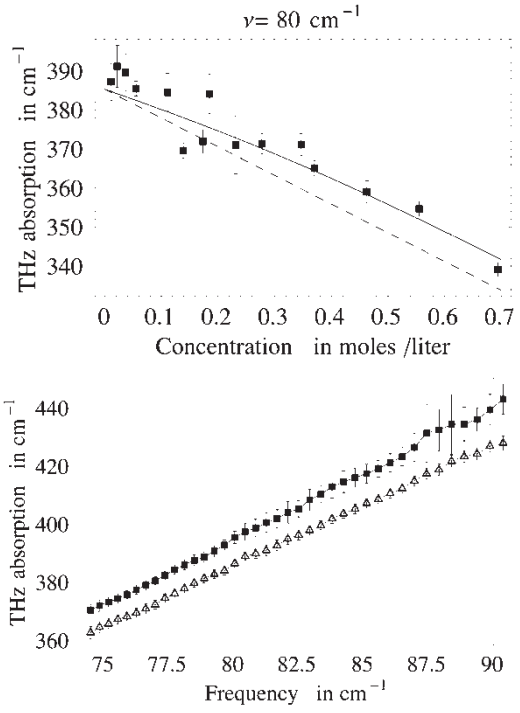

Figure 9. Top panel: Absorbance of lactose solution as a function of lactose concentration. Displayed is the mean value of all measured THz absorption coefficients for each lactose concentration together with the statistical error. Shown are a reference fit with bulk water lactose only (dashed) and the best fit accounting for solvation water. Bottom panel: Comparison of the pure water (lower) and solvation water (upper) terahertz absorption derived from the measurements in the upper panel.

other dynamical measures of the solvation layer around sugar molecules used in simulations, specifically the local diffusion coefficient for water molecules around the solute, which yielded a similar value for the solvation layer around other saccharides [23].

The experimentally determined size of the solvation layer exceeds considerably that reported in previous studies. For example, Magazù *et al.*, studied the influence of different carbohydrates on solvation dynamics by measuring the quasi-elastic peak width using the IRIS backscattering facility at the Rutherford Appleton Laboratory [130]. They indeed observed a retardation of water dynamics, but the solvation numbers they deduced ranged only from 7.5 to 9 for sucrose, maltose and trehalose, respectively. Moreover, a more recent neutron diffraction study stated that the long-range structure of the water is not significantly affected by the sugar solute [131]. These values are in reasonable agreement with results of ultrasonic measurements (yielding solvation numbers between 14 and 15) but all these values are much smaller than the solvation numbers obtained by THz spectroscopy.

We must bear in mind, however, that the results of these studies depend on the time scale of the observations, which begs the questions: which motions are actually probed by the experiment, and how far out does the carbohydrate perturb these motions?

It is known that water molecules move between the bulk and the solvation layer on a time scale of 10–100 ps, so any technique that provides data averaged over more than the ps time scale will smear out the solvation layer, rendering it largely "invisible" to the observer. This would explain the smaller measured layer size of the ultrasonic and the neutron diffraction studies, which have longer averaging times. On the other hand, the size of the solvation layer will also depend on the frequency range that is probed since the experiment can either probe mostly the re-orientation dynamics (which is true for the dielectric measurements in the GHz range) or the translational dynamics (which is found to dominate the low end of the THz range). Different measurements probing distinct kinds of motion may thus lead to different conclusions concerning the size of the solvation layer.

THz spectroscopy allows a quantitative determination of the size of the solvation shell characterized by perturbed absorption of water probed on the ps and sub-ps time scale. Molecular modelling simulations reveal that the hydrogen bond reorientation dynamics of water is retarded around the lactose molecule on this time scale [10]. The observed changes in the THz absorption can thus be taken as an indication of the slowing of the hydrogen bond rearrangement dynamics for water molecules near lactose.

Although it is not yet clear which motions are most responsible for the observed bio-protection mechanism discussed at the outset, the translational motion of water near the solute has been found to be the primary mechanism of protein–water hydrogen bond network relaxation, which makes THz spectroscopy a promising tool for the investigation of the underlying molecular mechanism. Our measurements further support the idea that bioprotection by sugars is achieved via solute induced changes in the solvation shell affecting the dynamics of water, as we find the sugar molecule to significantly affect the dynamics of the surrounding water.

4.2. *Proteins*

Larger biomolecules, such as proteins, afford more length and time scales over which dynamics can be observed. This is true even for so-called "two-state" folders, proteins with a sufficiently large activation barrier so that the population can be approximately divided into a "reactant" (unfolded) and "product" (folded) population.

To put this in perspective, we can compare activation barriers along a collective reaction coordinate with activation barriers for a small molecule along a reaction coordinate [132]. Consider activated chemical reactions with an Arrhenius-type rate expression $k = \nu^{\dagger} \exp[-\Delta G^{\dagger}/kT]$. For a typical S_N2 reaction, the characteristic vibrational frequency of the reaction coordinate at the transition state is $\nu^{\dagger} \approx (100\,\text{fs})^{-1}$, and the barrier is $\Delta G^{\dagger} \approx 100\,\text{kJ/mole}$, about 25% of a covalent bond. For a protein folding in 1 second behaving like a Zimm polymer, the prefactor is approximately†

$$\nu^{\dagger}_{\text{Zimm}} = \frac{4\omega_{well}\omega_{TS}m}{3\pi^2\eta b\sqrt{6\pi N}}, \tag{4.2}$$

†In the Zimm model, the polymer backbone is a freely jointed chain of beads of mass m. ω_{well} is the frequency of collective motions in the unfolded well, ω_{TS} the same at the barrier, η is the solvent vicosity, b the separation between monomers (ca. 3.8 Å in proteins), and N the length of the polypeptide chain.

or $(100\,\text{ns})^{-1}$ in $\eta = 1\,\text{mPa s}$ viscosity solution (aqueous room temperature), and the barrier along the collective reaction coordinate is 40 kJ/mole. However, for a protein folding at record speeds, such as the three-helix bundle α-3D (ca. 0.5 μs) or $\lambda_{6\text{-}85}$ (ca. 10 μs), with diffusive prefactors estimated in the $(0.1\text{–}2\,\mu\text{s})^{-1}$ range, the collective free energy barrier may lie below 5 kJ/mole [133, 134]. The absence of a large barrier removes the coarse-graining to which experiments are usually subject, and populations can no longer be divided into "reactant" and "product", nor can the reaction coordinate be well represented by a single principal component [135].

Molecular dynamics simulations have already revealed the underlying complexity of protein free energy surfaces [136, 137], which extends from the fractal nature of the native state to large numbers of saddle points connecting local minima. Sampling strategies have been put forth to characterize the free energy surfaces of proteins and clusters [138, 139], providing information about minima and the transition states that separate them on a complex energy landscape. The landscape can be conveniently represented by disconnectivity graphs [140, 141], which graphically reveal the permutationally distinct minima and the barriers that separate them, thereby displaying the complexity of the energy landscape [138]. One important question that can be answered by such simulations regards the nature of motions required for folding: are the important motions composed of very large numbers of high-frequency modes, whose collective motion can be approximated by a diffusive process [98], or are coherent large-amplitude (low-frequency) modes responsible, whose motion becomes diffusive only by solvent damping? In the latter case, solvent friction should dominate over internal friction of the biomolecule.

Consider for example mutants of the lambda repressor fragment 6-85, originally studied by Oas [142]. Figure 10 shows a comparison of refolding rates derived from infrared data such as from figure 2, and SAXS data such as in figure 4 [76, 143, 144]. Rates are shown for several mutants of the five-helix bundle $\lambda_{6\text{-}85}$, making the two-state assumption that

$$k_{obs} = k_{fold} + k_{unfold}, \quad k_{eq} = k_{fold}/k_{unfold} \tag{4.3}$$

The infrared data measuring secondary structure were obtained in aqueous buffer near room temperature (30–60°C), while the SAXS data were obtained at −28°C. The Y22W mutant behaves like a two-state folder at room temperature. As shown in figure 10, the collapse of the Y22W and Y22W/A37,49G mutants is faster than expected from room temperature measurements of mutants of similar stability. Only the downhill-folding mutant Y22W/Q33Y/G46,48A has an IR-detected rate at room temperature that is also comparably faster. Secondary structure formation measured by circular dichroism at −28°C is even faster than collapse measured by SAXS, and yields a third set of rates in addition to the IR and SAXS rates [76]. Thus the cryosolvent conditions uncouple the collapse of the protein from the formation of secondary structure.

The fact that three or more rates measured by different probes do not match, shows that ultimately, the two-state assumption is not robust for this protein. Krantz *et al.* have discussed how nonspecific contraction can add a second rate coefficient while the two state scenario is preserved [146]. However, if the additional phases form large amounts of structure (as in Y22W under cryogenic conditions), or if there are many

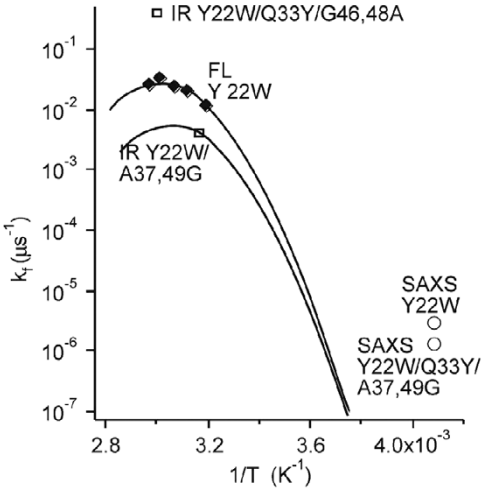

Figure 10. Folding rates of λ-repressor fragment mutants, obtained from relaxation data using a two-state assumption: the collapse of the protein measured by SAXS is much faster in ethylene glycol cryosolvent than extrapolated from room temperature. Only downhill folders like Y22W/Q33Y/G46,48A show a comparably faster rate relative to the room temperature wild type.

additional phases, then their nonspecific contraction + barrier scenario becomes either a downhill, downhill + barrier or an intermediates + barrier scenario. If the time scale of any phases becomes comparable to equation (4.2), then that stage of folding should be considered downhill. It has been proposed that downhill folding gives way to barrier limited folding mainly because of evolutionary optimization of proteins for function and against aggregation, so substantial ($\gg kT$) barriers are not obligatory for folding [132]. Recent molecular dynamics simulations of $\lambda_{6\text{-}85}$ are compatible with the very rapid formation of some secondary structure in the molecule. Replica exchange simulations show that in particular, helix 1 (residues 9–25) and helix 4 (residues 59–69) have a very high propensity for forming secondary structure and for orienting correctly in the unfolded state under native conditions [147].

In figure 10, mutations switching glycines to alanines substantially speed up rates measured both by IR absorption spectroscopy and by SAXS. This is generally observed, and has two causes: alanines stabilize α-helices, and in addition the reduced flexibility of the backbone caused by the larger sidechain reduces the configurational entropy of the unfolded state. Site-directed mutagenesis of protein has proved very useful in folding studies [148, 149]. They can produce proteins with much lower activation barriers than the ''wild type'', as in the above example, or they can be used to probe transition state structure when there is a barrier. In the latter case, more conservative mutations than Gly → Ala are usually made, such as Ile → Val or Val → Ala, where only 1 or 2 methyl groups are truncated away. If such a mutation changes the folding rate coefficient such that $-RT\ln(k_{before}/k_{after}) \approx \Delta G^{\dagger}_{before} - \Delta G^{\dagger}_{after}$ is comparable to the change in native state free energy, then the amino acid is said to make a native-like contact in the transition state [150].

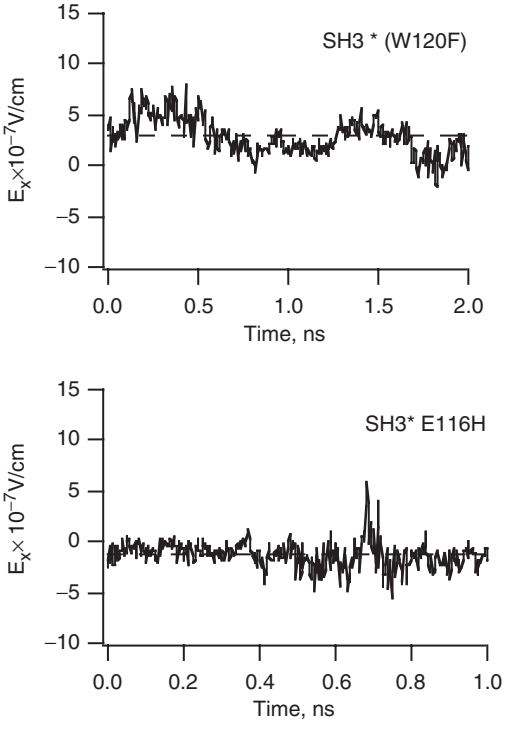

Figure 11. Molecular dynamics simulations (using the CHARMM22 force field and NAMD with full Tip3P solvation) of the electric field along the dipole axis of the tryptophan 119 indole ring in *fyn* SH3. The glutamate to histidine (charged) mutation results in a significant drop of the average electric field experienced at position 119.

Other variations are possible: replacement of NH groups by oxygen in amide linkages between amino acids deletes a hydrogen bond donor and reports on the relative importance of backbone–solvent in the unfolded state vs. backbone–backbone hydrogen bonds in the folded state [151]. Substitution of protonatable residues on the surface, such as histidines (whose $pK_a \approx 5$–6 leads to a switch of charge in that pH range) can affect the local solvation environment of the protein, which in turn can be studied by infrared or THz absorption spectroscopy [152]. Molecular dynamics can then be used to study variations in water structure caused by the mutation. An example is shown in figure 11, where two mutants of *fyn* SH3 domain, a mainly β-sheet protein, are simulated [153]. The mutants, differing by a single Glu $\rightarrow$ His mutation, are simulated for 2 ns using explicit TIP3P water and the CHARM22 force field in NAMD. The charge on the histidine and solvent restructuring lead to a reversal in the local electric field sensed at the position of residue 119.

THz spectroscopy is the ideal tool to probe local and global solvent shell dynamics of the type that can be differentiated by protein surface mutations. In addition, THz spectroscopy probes larger amplitude motions of the protein backbone and sidechains than is possible in the IR. We therefore have begun to measure the THz spectrum of the 5-helix bundle $\lambda_{6\text{-}85}$ [152].

Figure 6 displays the measurement of a solvated protein THz spectrum in the frequency range 74–86 cm^{-1}. Following the analysis presented for the solvated lactose molecule to obtain an average solvation layer absorbance and thickness, we first have to consider the decrease in absorption according to the replacement of water by the protein. At the protein concentration of 3.41 mg/ml in figure 6, the expected decrease due to exclusion of water is 0.002%, a trend that is the opposite of the increased absorption shown in figure 6. In fact, an increase in the absorption coefficient is consistent with the computed absorbance of solvated $\lambda_{6\text{-}85}$, also shown in figure 6, which we observe to be larger than the computed absorbance of a similar volume of bulk water. At higher concentrations, the protein solution absorbs even less than predicted by this model, so the explanation requires change in the water absorbance.

The simple exclusion picture may be a reasonable zeroth order approximation for a small carbohydrate, but clearly fails for proteins because of their larger THz spectral density. In fact, an increase in the absorption coefficient is consistent with the computed absorbance of solvated $\lambda_{6\text{-}85}$, also shown in figure 6, which we observe to be larger than the computed absorbance of a similar volume of bulk water.

The calculated absorbance of $\lambda_{6\text{-}85}$ shown in figure 6 (which also shows lactose solution for comparison) is similar to that of a comparable volume of bulk water. Thus, the increased absorption coefficient at low protein concentrations shown in figure 6 is consistent with a larger absorbance of the solvation layer compared to bulk water. This is in agreement with molecular-dynamics based spectral densities of individual solvated $\lambda_{6\text{-}85}$ molecules, where the water spectral density can be computed as a function of the distance from the protein (figure 6).

THz spectra are sensitive to changes in the water network and water dynamics. For example, a sizable and nonmonotonic variation in THz absorption coefficients with change in temperature from 4 to 50°C has been observed for bulk water [70]. We speculate that the observed THz spectrum of $\lambda_{6\text{-}85}$ is an indication of a water structuring effect of the protein, and that sizeable changes induced by proteins will be observed in the future as a function of pH, protein concentration, temperature, and side chain mutations at the surface of the protein.

5. Summary and outlook

Molecular dynamics simulations, analysed to yield spectral densities, protein structural fluctuations and vibrational motions, can provide a microscopic interpretation of biomolecule dynamics on multiple time and length scales. Such multiscale processes occur from the native state, where energy flow is governed by a fractal structure, to the unfolded ensemble, which consists of a complex network of free energy minima revealed when folding barriers are minimized by biomolecule engineering. Infrared spectroscopy, terahertz spectroscopy, and small-angle X-ray scattering illustrate the type of thermodynamic, dynamic and kinetic data that can be obtained and analysed. The key to understanding biomolecular dynamics will be to analyse multiple such experiments simultaneously within a single computational and theoretical framework.

Acknowledgments

This work was supported by a grant from the Human Frontiers Science programme to MH, DML and MG. MG was further supported by grant MCB 03-16925 from the National Science Foundation, MH by grant HA2394/8-3 from the Deutsche Forschungsgemeinschaft, and DML by grant CHE-0506020 from the National Science Foundation.

References

[1] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes, Science **254,** 1598 (1991).
[2] R. Granek and J. Klafter, Phys. Rev. Lett. **95,** 098106 (2005).
[3] G. N. Phillips Jr and B. M. Pettitt, Prot. Science **4,** 149 (1995).
[4] S. J. Gill, S. F. Dec, G. Olofsson, *et al*., J. Phys. Chem. **89,** 3758 (1985).
[5] P. Ferrara, J. Apostolakis, and A. Caflisch, J. Phys. Chem. B **104,** 5000 (2000).
[6] H. S. Chan, Proteins: Struct. Funct. Genet. **40,** 543 (2000).
[7] J. Gomez, V. J. Hilser, D. Xie, *et al*., Proteins **22,** 404 (1995).
[8] W. Y. Yang, J. Pitera, W. Swopes, *et al*., J. Mol. Biol. **336,** 241 (2004).
[9] H. Susi, Methods Enzym. **22,** 455 (1971).
[10] U. Heugen, G. Schwaab, E. Brundermann, *et al*., Proc. Natl. Acad. Sci. USA **103,** 12301 (2006).
[11] A. Guinier and G. Fournet, *Small-Angle Scattering of X-rays* (John, New York, 1955).
[12] E. Larios, J. S. Li, K. Schulten, *et al*., J. Mol. Biol. **340,** 115 (2004).
[13] G. Damaschun, H. Damaschun, K. Gast, *et al*., Biochemistry **32,** 7739 (1993).
[14] D. Chandler, *Modern Statistical Mechanics* (Oxford University Press, Oxford, 1989).
[15] X. Yu and D. M. Leitner, in *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*, edited by Q. Cui and I. Bahar (Chapman and Hall/CRC, Boca Raton, 2005).
[16] D. M. Leitner, Adv. Chem. Phys. **130B,** 205 (2005).
[17] H. Fujisaki, Y. Zhang, and J. E. Straub, J. Chem. Phys. **124,** 144910 (2006).
[18] X. Yu and D. M. Leitner, J. Chem. Phys. **122,** 054902 (2005).
[19] A. Roitberg, R. B. Gerber, and M. A. Ratner, J. Phys. Chem. **101,** 1700 (1997).
[20] A. Roitberg, R. B. Gerber, R. Elber, *et al*., Science **268,** 1319 (1995).
[21] S. E. Whitmire, D. Wolpert, A. G. Markelz, *et al*., Biophys. J. **85,** 1269 (2003).
[22] M. Tarek and D. J. Tobias, Phys. Rev. Lett. **88,** 138101 (2002).
[23] S. L. Lee, P. G. Debenedetti, and J. R. Errington, J. Chem. Phys. **122,** 204511 (2005).
[24] M. B. Enright and D. M. Leitner, Phys. Rev. E **71,** 011912 (2005).
[25] S. Mukamel, *Principles of Nonlinear Optical Spectroscopy* (Oxford University Press, New York, 1995).
[26] M. T. Zanni and R. M. Hochstrasser, Curr. Opinion. Struct. Bio. **11,** 516 (2001).
[27] H. S. Chung, M. Khalil, A. W. Smith, *et al*., Proc. Nat. Acad. Sci. USA **102,** 612 (2005).
[28] P. A. Anfinrud, C. Han, and R. M. Hochstrasser, Proc. Natl. Acad. Sci. USA **86,** 8387 (1989).
[29] P. Hamm, M. H. Lim, and R. M. Hochstrasser, J. Phys. Chem. B **102,** 6123 (1998).
[30] R. B. Dyer, F. Gai, W. H. Woodruff, *et al*., Acc. Chem. Res. **31,** 709 (1998).
[31] C. A. Schmuttenmaer, Chem. Rev. **104,** 1759 (2004).
[32] K. Tang and K. Dill, J. Biomolec. Struct. Dynam. **16,** 397 (1998).
[33] D. Bourgeois, B. Vallone, A. Arcovito, *et al*., Proc. Natl. Acad. Sci. USA **103,** 4924 (2006).
[34] L. G. Wang and F. J. Sigworth, Physiology **21,** 13 (2006).
[35] H. H. Mantsch and D. Chapman, *Infrared Spectroscopy of Biomolecules* (John Wiley, New York, 1996).
[36] P. Hamm, M. H. Lim, and R. M. Hochstrasser, J. Phys. Chem. B **102,** 6123 (1998).
[37] M. Goodman, Y. Masuda, and A. S. Verdini, Biopolymers **10,** 1031 (1971).
[38] E. B. Brauns and R. B. Dyer, Biophys. J. **89,** 3523 (2005).
[39] K. A. Peterson, C. W. Rella, J. R. Engholm, *et al*., J. Phys. Chem. B **103,** 557 (1999).
[40] A. H. Xie, A. F. G. van der Meer, and R. H. Austin, Phys. Rev. Lett. **88,** 018102 (2002).
[41] A. H. Xie, L. van der Meer, W. Hoff, *et al*., Phys. Rev. Lett. **84,** 5435 (2000).
[42] C. Tanford, Adv., Protein Chem. **24,** 1 (1970).
[43] E. R. Henry and J. Hofrichter, Methods Enzymol. **210,** 129 (1992).
[44] I. Noda and Y. Ozaki, *Two-Dimensional Correlation Spectroscopy* (John Wiley, New York, 2004).
[45] O. Chandrawinatha, *Chemistry*, B.S. Thesis (University of Illinois, Urbana, 2005).
[46] E. Larios, W. Y. Yang, K. Schulten, *et al*., Chem. Phys. **307,** 217 (2004).
[47] H. Ma, J. Ervin, and M. Gruebele, Rev. Sci. Instrum. **64,** 486 (2004).

[48] G. U. Nienhaus, J. R. Mourant, and H. Frauenfelder, Proc. Natl. Acad. Sci. USA **89,** 2902 (1992).
[49] J. A. Bailey, F. L. Tomson, S. L. Mecklenburg, *et al*., Biochemistry **41,** 2673 (2002).
[50] D. Reinstädler, H. fabian, J. Backmann, *et al*., Biochemistry **35,** 15822 (1996).
[51] H. Susi and D. M. Byler, Methods Enzym. **130,** 290 (1986).
[52] C. Fattinger and D. Grischkowsky, Appl. Phys. Lett. **54,** 490 (1989).
[53] R. Kohler, A. Tredicucci, C. Mauro, *et al*., Appl. Phys. Lett. **84,** 1266 (2004).
[54] E. Brundermann, in *Long-Wavelength Infrared Semiconductor Lasers*, edited by H. K. Choi (John Wiley, New York, 2004).
[55] T. Kleine-Ostmann, P. Knobloch, M. Koch, *et al*., Electron. Lett. **37,** 1461 (2001).
[56] S. Hoffmann, M. Hofmann, E. Brundermann, *et al*., Appl. Phys. Lett. **84,** 3585 (2004).
[57] A. Stohr and D. Jager, in *Microwave Photonics*, edited by S. Iezekiel (John Wiley, New York, 2006).
[58] W. B. Colson, E. D. Johnson, M. J. Kelley, *et al*., Phys. Today **55,** 35 (2002).
[59] G. L. Carr, M. C. Martin, W. R. McKinney, *et al*., Nature **420,** 153 (2002).
[60] M. C. Beard, G. M. Turner, and C. A. Schmuttenmaer, J. Phys. Chem. B **106,** 7146 (2002).
[61] D. S. Venables and C. A. Schmuttenmaer, J. Chem. Phys. **113,** 11222 (2000).
[62] M. Walther, B. Fischer, M. Schall, *et al*., Chem. Phys. Lett. **332,** 389 (2000).
[63] A. Wittlin, L. Genzel, F. Kremer, *et al*., Phys. Rev. A **34,** 493 (1986).
[64] J. T. Kindt and C. A. Schmuttenmaer, J. Phys. Chem. **100,** 10373 (1996).
[65] N. Nandi, K. Bhattacharyya, and B. Bagchi, Chem. Rev. **100,** 2013 (2000).
[66] F. Keilmann, V. N. Shastin, and R. Till, Appl. Phys. Lett. **58,** 2205 (1991).
[67] E. E. Haller and E. Brundermann, Patent No. 6, 011, 810, United States, 2000.
[68] E. Bründermann, D. R. Chamberlin, and E. E. Haller, Appl. Phys. Lett. **73,** 2757 (1998).
[69] A. Bergner, U. Heugen, E. Brundermann, *et al*., Rev. Sci. Instrum. **76,** 63110 (2005).
[70] J. K. Vij, D. R. J. Simpson, and O. E. Panarina, J. Molec. Liq. **112,** 125 (2004).
[71] O. Levy and D. J. Bergman, Phys. Rev. B Condens. Matt. **46,** 7189 (1992).
[72] T. P. Iglesias, J. M. Fornies-Marquina, and B. De Cominges, Molec. Phys. **103,** 2639 (2005).
[73] I. S. Millet, S. Doniach, and K. W. Plaxco, Adv. Protein Chem. **62,** 241 (2002).
[74] L. Pollack, M. W. Tate, N. C. Darnton, *et al*., Proc. Nat. Acad. Sci. USA **96,** 10115 (1999).
[75] D. I. Svergun, V. V. Volkov, M. V. Kozin, *et al*., J. Appl. Cryst. **30,** 798 (1997).
[76] C. Dumont, Y. Matsumura, S. J. Kim, *et al*., Protein Sci., in press (2006).
[77] Z. Qin, J. Ervin, E. Larios, *et al*., J. Phys. Chem. B **106,** 13040 (2002).
[78] D. Svergun, C. Barberato, and M. H. J. Koch, J. Appl. Cryst. **28,** 768 (1995).
[79] Q. Cui and I. Bahar, editors, *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems* (Chapman & Hall/CRC, Boca Raton, 2005).
[80] A. E. Garcia, R. Blumenfeld, G. Hummer, *et al*., Physica D **107,** 225 (1997).
[81] A. L. Tournier and J. C. Smith, Phys. Rev. Lett. **91,** 208106 (2003).
[82] B. Hess, Phys. Rev. E **65,** 031910 (2002).
[83] G. G. Maisuradze and D. M. Leitner, Chem. Phys. Lett. **421,** 5 (2006).
[84] Y. Mu, P. H. Nguyen, and G. Stock, Proteins: Struct. Funct. Bioinform. **58,** 45 (2005).
[85] J. Ervin, E. Larios, S. Osvath, *et al*., Biophys. J. **83,** 473 (2002).
[86] C. L. Brooks, M. Karplus, and B. M. Pettitt, Adv. Chem. Phys. **71,** 1 (1988).
[87] J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, New York, 1987).
[88] X. Yu and D. M. Leitner, J. Phys. Chem. B **107,** 1698 (2003).
[89] E. R. Henry, W. A. Eaton, and R. M. Hochstrasser, Proc. Natl. Acad. Sci. USA **83,** 8982 (1986).
[90] M. Tesch and K. Schulten, Chem. Phys. Lett. **169,** 97 (1990).
[91] N. Ota and D. A. Agard, J. Mol. Biol. **351,** 345 (2005).
[92] L. T. Bu and J. E. Straub, J. Phys. Chem. B **107,** 12339 (2003).
[93] D. E. Sagnella, J. E. Straub, and D. Thirumalai, J. Chem. Phys. **113,** 7702 (2000).
[94] M. B. Enright, X. Yu, and D. M. Leitner, Phys. Rev. E **73,** 051905 (2006).
[95] X. Yu and D. M. Leitner, J. Chem. Phys. **119,** 12673 (2003).
[96] X. Yu, J. Park, and D. M. Leitner, J. Phys. Chem. B **107,** 12820 (2003).
[97] A. A. Maradudin and A.A. Fein, Phys. Rev. Lett. **128,** 2589 (1962).
[98] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, *et al*., Proc. Natl. Acad. Sci. USA **99,** 16047 (2002).
[99] D. Vitkup, D. Ringe, G. A. Petsko, *et al*., Nat. Struct. Biol. **7,** 34 (2000).
[100] J. M. Gruschus and J. A. Ferretti, J. Biomolec. Nmr **20,** 111 (2001).
[101] D. Russo, G. Hura, and T. Head-Gordon, Biophys. J. **86,** 1852 (2004).
[102] A. R. Bizzarri and S. Cannistraro, J. Phys. Chem, B **106,** 6617 (2002).
[103] F. Piazza, P. De Los Rios, and Y.-H. Sanejouand, Phys. Rev. Lett. **94,** 145502 (2002).
[104] P. J. Steinbach and B. R. Brooks, Proc. Natl. Acad. Sci. USA **93,** 55 (1996).
[105] J. Knab, J. Y. Chen, and A. Markelz, Biophys. J. **90,** 2576 (2006).
[106] T. M. Korter and D. F. Plusquellic, Chem. Phys. Lett. **385,** 45 (2004).

[107] F. W. Starr, J. K. Nielsen, and H. E. Stanley, Phys. Rev. Lett. **82,** 2294 (1999).
[108] A. Luzar and D. Chandler, Phys. Rev. Lett. **76,** 928 (1996).
[109] R. Elber, A. Roitberg, C. Simmerling, *et al*., Comp. Phys. Comm. **91,** 159 (1995).
[110] O. Jaetaek, K. Yangmee, and W. Youngdo, Bull. Korean Chem. Soc. **16,** 1153 (1995).
[111] E. Lindahl, B. Hess, and D. van der Spoel, J. Molec. Model. **7,** 306 (2001).
[112] H. J. C. Berendsen, D. Van der Spoel, and R. Van Drunen, Comp. Phys. Comm. **91,** 43 (1995).
[113] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, *et al*., *Biomolecular Simulation: The GROMOS96 Manual and User Guide*, Zurich (1996).
[114] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, *et al*., in *Intermolecular Forces*, edited by B. Pullman (Reidel Publishing, Dordrecht, 1981).
[115] A. R. Drews, B. D. Thayer, H. J. Stapleton, *et al*., Biophys. J. **57,** 157 (1990).
[116] H. J. Stapleton, J. P. Allen, C. P. Flynn, *et al*., Phys. Rev. Lett. **45,** 1456 (1980).
[117] W. Min, G. B. Luo, B. J. Cherayil, *et al*., Phys. Rev. Lett. **94,** 198302 (2005).
[118] S. C. Kuo and X. S. Xie, Phys. Rev. Lett. **93,** 180603 (2004).
[119] H. Yang, G. B. Luo, P. Karnchanaphanurach, *et al*., Science **302,** 262 (2003).
[120] S. Alexander and R. Orbach, J. De Phy. Lett. **43,** L625 (1982).
[121] T. Nakayama, K. Yakubo, and R. L. Orbach, Rev. Mod. Phys. **66,** 381 (1994).
[122] R. Rammal and G. Toulouse, J. De Phy. Lett. **44,** L13 (1983).
[123] R. Elber and M. Karplus, Phys. Rev. Lett. **56,** 394 (1986).
[124] R. Elber, in *The Fractal Approach to Heterogeneous Chemistry*, edited by D. Avnir (John Wiley, New York, 1989).
[125] R. Burioni, D. Cassi, F. Cecconi, *et al*., Proteins: Struct. Funct. Bioinform. **55,** 529 (2004).
[126] J. H. Crowe, L. M. Crowe, and D. Chapman, Science **223,** 701 (1984).
[127] T. Y. Lin and S. N. Timasheff, Protein Sci. **5,** 372 (1996).
[128] J. H. Crowe, J. F. Carpenter, and L. M. Crowe, Ann. Rev. Physiol. **60,** 73 (1998).
[129] A. M. Massari, I. J. Finkelstein, B. L. McClain, *et al*., J. Am. Chem. Soc. **127,** 14279 (2005).
[130] S. Magazu, V. Villari, P. Migliardo, *et al*., J. Phys. Chem. B **105,** 1851 (2001).
[131] P. E. Mason, G. W. Neilson, J. E. Enderby, *et al*., J. Phys. Chem. B **109,** 13104 (2005).
[132] M. Gruebele, Comptes Rendue Biologies **328,** 701 (2005).
[133] W. Y. Yang and M. Gruebele, Nature **423,** 193 (2003).
[134] Y. Zhu, D. O. V. Alonso, K. Maki, *et al*., Proc. Natl. Acad. Sci. **100,** 15486 (2003).
[135] H. Ma and M. Gruebele, J. Comp. Chem. **27,** 125 (2005).
[136] J. Shea and C. L. Brooks, Ann. Rev. Phys. Chem. **52,** 499 (2001).
[137] V. Daggett, Curr. Opin. Struct. Biol. **10,** 160 (2000).
[138] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, 2003).
[139] K. D. Ball, R. S. Berry, R. E. Kunz, *et al*., Science **271,** 963 (1996).
[140] R. Czerminski and R. Elber, J. Chem. Phys. **92,** 5580 (1990).
[141] O. M. Becker and M. Karplus, J. Chem. Phys. **106,** 1495 (1997).
[142] G. S. Huang and T. G. Oas, Proc. Natl. Acad. Sci. USA **92,** 6878 (1995).
[143] W. Y. Yang and M. Gruebele, Biochemistry **43,** 13018 (2004).
[144] H. Ma and M. Gruebele, Proc. Nat. Acad. Sci. USA **102,** 2283 (2005).
[145] S. A. Hawley, Biochemistry **10,** 2436 (1971).
[146] B. A. Krantz and T. R. Sosnick, Biochemistry **39,** 11696 (2000).
[147] E. Larios, J. W. Pitera, W. C. Swope, *et al*., Chem. Phys. **323,** 45 (2006).
[148] M. Jäger, H. Nguyen, J. Crane, *et al*., J. Mol. Biol. **311,** 373 (2001).
[149] A. Matouschek, J. T. Kellis Jr, L. Serrano, *et al*., Nature **340,** 122 (1989).
[150] A. R. Fersht, *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (W.H. Freeman, New York, 1998).
[151] S. Deechongkit, H. Nguyen, E. T. Powers, *et al*., Nature **430,** 101 (2004).
[152] S. Ebbinghaus, S. J. Kim, X. Yu, *et al*., submitted (2006).
[153] S. Pappu, *Chemical Engineering* (University of Illinois, Urbana, 2005).