Protein Vibrational Energy Transfer: New Sensors and Application in an Enzyme Active Site

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich Physik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
JAN GERRIT LÖFFLER
aus Hamburg

Frankfurt am Main 2023
(D 30)

vom Fachbereich Physik der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Roger Erb

Gutachter: Prof. Dr. Jens Bredenbeck
Prof. Dr. Josef Wachtveitl

Datum der Disputation:

„Welcome. And congratulations.
I am delighted that you could make it. Getting here wasn't easy, I know. In fact, I suspect it was a little tougher than you realize."

Bill Bryson, in
A Short History of *Nearly* Everything

Table of Contents

Motivation

After more than a century of protein research, our understanding of their functions has vastly improved, yet many questions remain to be answered. Already in the 19th century it was known that amino acids were the building blocks of proteins, but the first protein amino acid sequence and crystal structure were determined only in the mid of the 20th century [1–4]. Today these milestones have become much more accessible, and the protein structure remains the most important resource to prepare and interpret experiments. By now, it is increasingly recognized that the dynamics of protein molecules are crucial for understanding their function, and significant efforts are being made to elucidate these dynamics.

The relevant biomolecular dynamics span a broad range of timescales. Very fast events like bond breaking or reorganization, energy transfer and fast charge transfer processes happen on the timescale of femtoseconds to picoseconds. Side chain motions occur in picoseconds to nanoseconds, whereas large scale structure formation and changes can take between microseconds and seconds. The very fastest of these processes are typically investigated using multipulse techniques. Experimental time resolutions in the picosecond range can be achieved in relaxation nuclear magnetic resonance spectroscopy (NMR), ultrafast X-ray crystallography, or ultrafast laser spectroscopy, the latter reaching the femtosecond or even attosecond realm [5–7].

One of the very fast dynamic processes that is yet to be fully understood is vibrational energy transfer (VET). It is intrinsically anisotropic in proteins because of their folded structure with different types and strengths of couplings between the residues [8–11]. This allows a controlled flow of energy. Proteins evolved to create channels of energy flow that allow excess energy, e.g., after chemical or photo-induced reactions to dissipate in a controlled fashion [8, 12–15]. A connection between allostery, an important mechanism to control cellular biochemistry, has been proposed. There are some allosteric effects that do not involve large-scale structural rearrangements but are based on dynamic allostery. VET between such allosteric sites could be involved in information transfer and help to identify the coupled residues that facilitate the allosteric regulation [11, 16–19]. Furthermore, VET is facilitated by low-frequency modes (LFMs) that span the entire molecule. Some of these LFMs could be involved in rate-promoting vibrations (RPVs), a concept which is recently being used to finally explain the extraordinary catalytic power of enzymes [20–23].

Transient infrared (TRIR) and Raman spectroscopy have previously been used to study energy transfer phenomena in organic molecules and even proteins [24–28], but it was in the group of Prof. Bredenbeck that a pair of non-canonical amino acids (ncAAs) was developed with the specific goal to enable the characterization of VET through proteins [29]. These are β-(1-azulenyl)-L-alanine (AzAla) as VET donor and azido-L-homoalanine (Aha) as VET sensor.

The use of ncAAs enables the investigation of VET between any two chosen sites within the protein. In synthetic peptides this is straightforward, but in larger proteins, cotranslational approaches are required, and the simultaneous incorporation of more than one ncAA

is still challenging. It is more feasible to incorporate one functionality into the protein and add the second functionality via a ligand. Previously, AzAla could not be incorporated co-translationally, whereas selective pressure incorporation (SPI) could be used to incorporate Aha into proteins [30]. This however, limited the use of Aha to proteins that lack the amino acid Met.

The first protein system that was identified for the use of the AzAla/Aha VET pair was a PDZ domain of the post synaptic density protein 95 (PSD-95): it naturally does not contain Met and it binds peptide ligands. In first VET experiments on this system, the feasibility to track VET over longer distances has been shown [31, 32]. Furthermore, the data of a comprehensive study of VET in several PDZ mutants is the topic of a publication that is in preparation and is covered in another dissertation.

On this firm basis, the aim of the present dissertation was to expand the applicability of VET experiments. By identifying a wider range of VET-compatible ncAAs, new incorporation methods can be exploited to overcome the requirement of Met-free target proteins. Furthermore, with a more variable range of VET sensors, the best replacement for each amino acid could be used for experiments with mutants that are closer to the native state of the investigated protein than what is possible using only Aha. In Part II of this dissertation, the use of several alternative VET sensor ncAAs is tested. Their different spectral properties and VET responses are characterized and assessed with regard to VET experiments.

To get further insights into VET and its role in proteins, the VET investigations need to be extended to new protein systems. With a new system, several innovations could now be implemented that were not available for PDZ. AzAla can by now be incorporated into the protein, instead of Aha, and the new protein system for this project formate dehydrogenase (FDH) opened up the opportunity to introduce a new type of VET sensors, small molecule ligands. With AzAla mutants of FDH, the role of enzyme dynamics for catalysis can be investigated, adding a significant new aspect to the application portfolio of VET. The selection and development of a new target for VET and first experiments are described in Part III of this dissertation, which thereby contributes to the mutual effort to investigate and understand VET and protein dynamics on a fundamental level.

Part I

Introduction

1 Theory and Background

This chapter is intended to offer a short and very general introduction of the basic principles and techniques that the work presented in this thesis is based on. It provides by no means a comprehensive overview of the mentioned topics and the extensive literature, but may serve as a reminder or as incentive for further reading. Specific introductions to the projects in Parts II and III are given in chapters 3 and 6, respectively.

1.1 Vibrational Spectroscopy

Vibrational spectroscopy is the foundation for transient infrared spectrscopy, the technique that was used to gather the vibrational energy transfer (VET) data presented in this thesis. Most of the content of this section is based on P. W. Atkins *Atkins' Physical Chemistry* [33], M. Tasumi *Introduction to Experimental Infrared Spectroscopy* [34], and P. J. Larkin *Infrared and Raman Spectroscopy* [35].

Vibrational spectroscopy measures the interaction between IR light and a sample. The spectra can be interpreted in order to gather information about the molecules within that sample. The energy at which this interaction occurs can be approximated with the eigenvalues E_ν of a Hamiltonian describing a harmonic potential (Eq. 1.1) or more precisely with the eigenvalues for a Hamiltonian describing an anharmonic Morse potential (Eq. 1.2) between two point masses, which leads to an oscillation of these masses.

$$E_\nu = \hbar\omega \left(\nu + \frac{1}{2}\right) \tag{1.1}$$

$$E_\nu = \hbar\omega \left(\nu + \frac{1}{2}\right) - \hbar\omega\chi_e \left(\nu + \frac{1}{2}\right)^2 \text{ or in terms of wavenumbers } \widetilde{\nu}$$

$$E_\nu = hc\widetilde{\nu} \left(\nu + \frac{1}{2}\right) - hc\chi_e\widetilde{\nu} \left(\nu + \frac{1}{2}\right)^2 \tag{1.2}$$

$$\text{with } \nu = 0, 1, 2, ..., \nu_{max} \text{ and } \widetilde{\nu} = \frac{1}{2\pi c}\sqrt{\frac{k_F}{\mu}}$$

where ν is the quantum number of the vibrational state, ω is the angular frequency, and $\widetilde{\nu}$ the wavenumber of the system, k_F is the attractive force constant between two masses with the reduced mass μ, χ_e is the anharmonicity of the Morse potential. The remaining terms are physical constants, i.e., the reduced Planck constant $\hbar$, the Planck constant h, and the speed of light in vacuum c.

There are three important differences between harmonical and the anharmonical system: (i) The transition energies between adjacent states are equal for the harmonic oscillator, but decrease for excitation of higher states in the anharmonic case. (ii) There is a ν_{max} exclusively for the anharmonic potential. In the physicochemical context this reflects the dissociation of molecular bonds at energies exceeding the binding energy. (iii) Closer inspection

of the eigenstates of the systems reveals that the dipole operator – that defines the general interaction between a molecule and electromagnetic radiation – may facilitate the transition from one eigenstate to the other for cases of $\Delta\nu \neq \pm 1$ (overtones and combination bands) only in the anharmonic case.

A molecule must be at least diatomic in order to exhibit a vibrational mode and heteroatomic in order for this mode to be IR active.[1] But the outlined findings can be generalized for molecules with N atoms in which generally $3N - 6$ vibrational modes occur.[2] Excitation of the system from one eigenstate to a higher eigenstate can then be triggered by the interaction with a photon whose energy matches the energy difference between the involved states. With typical values for k_F, χ_e, and μ for small groups of atoms, these photon energies lie within the range of IR light. When a molecular vibration involves more or heavier atoms, the reduced mass increases, leading to transitions at lower wavenumbers.

The absorption can be quantified by measuring the relative amount of light transmitted by the sample. The Beer-Lambert law establishes a linear dependency between the exponential intensity-loss of light passing through a sample with a given thickness, concentration, and molar extinction coefficient ε. This extinction coefficient implicitly accounts for the probability of the absorption process described in the previous paragraphs.

From the equations above one would expect sharp absorption lines in an absorption spectrum, as only photons whose energy exactly matches the difference between two states (whose transition is allowed) could be absorbed. However, there are different effects that lead to a finite band width of the absorption instead, such as Doppler broadening (negligible in the condensed phase), lifetime broadening, and inhomogeneous broadening caused by the vibrational Stark effect or structural inhomogeneity. Especially the latter two can be used to interpret absorption features in a spectrum to infer molecular properties of the sample. With these observables, one can assign various subensembles of molecules, e.g., hydrogen-bonded (H-bonded) and non-H-bonded species [36–38] or molecules bound to a protein or not [39, 40] can be distinguished. A simple absorption measurement can therefore help to characterize specific properties and molecular interactions of a sample in equilibrium.

1.2 Transient Infrared Spectroscopy

Typically, linear IR spectroscopy is implemented in the time-domain, exploiting the ability to calculate a wavenumber spectrum from the interferogram of temporally shifted electromagnetic waves via the Fourier transform (FTIR) [41, 42]. In order to gather dynamic non-equilibrium information, rapid-mixing, or rapid-, and step-scan experiments can be realized [43–46]. With these techniques the temporal resolution is limited to the low ns regime with the step-scan approach [47]. However, there are many bio- and physicochemical processes that occur on picosecond or even femtosecond timescales [34, 48–50]. For the investigation of such processes other techniques are required.

1 In contrast to vibrational modes that can be excited by Raman scattering. This is founded on the different selection rules for these two processes. IR active modes require a change of the dipole moment during the vibration, whereas Raman active modes require a change in the polarizability.

2 In linear molecules $3N - 5$ vibrational modes exist but only $3N - 6$ in non-linear molecules.

With modern pulsed laser systems based on chirped pulse amplification [51] high-intensity laser pulses with short pulse duration can be generated routinely.[3] In combination with non-linear optical effects, these pulses can be tuned to cover virtually any frequency range [7]. With these short and tuneable laser pulses, TRIR spectroscopy can be realized.

In TRIR spectroscopy, a non-equilibrium event is triggered with a UV/Vis-pump pulse, and a second pulse, the probe pulse, in the IR is used to detect the induced changes. To that end, the pump pulse is tuned to be resonant with an absorption that initiates the process. Depending on the phenomenon of interest this could be a temperature-jump [57], a pH-jump [58], uncaging of a reactive group [59], isomerizations [60], local injection of energy [29] or other effects that can be triggered by the absorption of light.

The IR-probe pulse is tuned to cover a wavenumber region in which a characteristic vibrational absorption of the process lies. The absorption of the probe pulse with varying delay to the pump pulse is measured and compared to the absorption that occurs without the previous excitation. With this procedure the response of the probed vibration to the triggered event can be observed and interpreted in a time-resolved fashion, e.g., the irreversible reduction of absorption may originate from the cleavage of a bond, or a shift of the IR absorption may be due to a change of its environment [61].

1.3 Vibrational Energy Transfer in Proteins

The transfer of energy within proteins is highly important. Among the various functions of proteins are some that lead to excessive energy release especially upon the absorption of UV/Vis light or during chemical reactions. The energy from these processes is at first localized leading to high energy densities. In order for the proteins to remain functional, efficient redistribution of this energy is required. As discussed in this section, their spatial extent through the protein allows coupled low-frequency modes (LFMs) to serve as channels for VET.

One way to theoretically investigate VET in proteins are non-equilibrium molecular dynamics (MD) simulations [9, 22, 62, 63]. A residue or region of the minimized protein structure, equilibrated at a given temperature (often low to reduce the noise), is temporarily coupled to a thermal bath at a higher temperature. Subsequently, the mechanical couplings between atoms lead to energy dissipation through the protein, they are determined by the employed force-field (typically up to four-body interactions [64]). From these simulations one can directly calculate the kinetic energy of individual residues and analyze the dynamic redistribution of this energy.

In order to reduce the computational cost of energy transfer simulations a master equation ansatz has been employed [65]. This approach describes the energy transfer between residues in terms of transition rates in a kinetic model. These transition rates were initially calculated from the mode diffusivity and the results compared to all-atom simulations [65]. A direct fitting of the rate constants from the MD simulations would lead to overfitting due to the sheer amount of rate constants in a full protein system. To this end, scaling rules

3 With further compression techniques single-digit fs pulses in the near IR [52] and Vis [53, 54] down to as pulses in the extreme UV [55] and X-ray [56] region are attainable. The compressibility is ultimately limited by the band width and wavelength of the pulse.

(Eqs. 1.3 and 1.4) have been derived that reduce the dimensionality of the fit [66]:[4]

$$k_{ij}^{B} = \frac{D_B}{\langle x_{ij}^2 \rangle} \sqrt{\frac{f_j}{f_i}} \tag{1.3}$$

$$k_{ij}^{C} = \frac{D_C}{\langle q_{ij}^2 \rangle} \sqrt{\frac{f_j}{f_i}} \tag{1.4}$$

where the indices i and j denote two residues between which energy is transferred, k_{ij}^{B} and k_{ij}^{C} are the rate constants for backbone and contact energy transfer, respecively, D_B and D_C are the diffusion constants for backbone and contact energy transfer, respecively, f_x is the number of degrees of freedom of residue x, $\langle x_{ij}^2 \rangle$ is the mean square distance between all atoms of i and j along covalent bonds, and $\langle q_{ij}^2 \rangle$ is the mean square distance of the two atoms that mediate the non-covalent contact.

Instead of fitting each rate constant individually, the degrees of freedom could be reduced to two for the complete description of energy transfer within the protein, i.e., D_B and D_C. The former is only defined for neighbouring residues but the latter describes the various types of non-covalent contacts between all residues in the protein. Additional three rate constants are required to describe the heating process itself and solvent effects. The master equation analysis was fit to data of different systems resulting in very similar fit parameters, leading to a generalized model for arbitrary systems [66]. However, for a quantitative fit to experimental data a quantum correction factor is required [68].

Another way to describe VET is the use of protein normal modes. In an average protein with ~350 amino acids [69] and ~16 atoms per amino acid there are in principle 16800 vibrational modes. This number contains local vibrations, in which only a few atoms move considerably, as well as vibrations that involve numerous atoms of various residues distributed within the protein. According to Sec. 1.1, modes at low wavenumbers require participation of many residues. Indeed, it was found that up to $^2/_3$ of all residues in a protein contribute to the normal LFMs, vibrations below ~150 cm^{-1}, whereas fewer residues contribute to normal modes at higher wavenumbers and generally remain localized above 300 cm^{-1} independent of structural details [64, 70]. The normal modes allow the expression of the heat current operator and definition of the mode diffusivity that are used to identify modes which efficiently facilitate energy transfer [71]. These can then help to identify mechanically coupled ("communicating") regions of a protein via communication maps [14, 72, 73]. It should be noted, that not every vibration that occurs in a protein represents a single normal mode, but all vibrations can be described by a combination of the normal modes, i.e., they constitute an orthogonal basis. Since these modes can routinely be calculated, they generate an important framework for the characterization and interpretation of protein vibrations and concomitant effects like energy transfer [70, 71, 74, 75].

In the normal mode picture the intramolecular vibrational redistribution (IVR), the process that leads to VET through the protein, is thought of in a hierarchical model. For the IVR from a high-frequency mode towards LFMs the sum of the excitation energies of the LFMs must equal the energy of the relaxing mode. Instead of decaying into numerous LFMs at once, the high-frequency mode is believed to populate a few modes at medium energy (first tier), whose decays in turn populate a few lower-frequency modes (second tier) and

4 Shown are the revised scaling rules, slightly different equations have originally been introduced in [67].

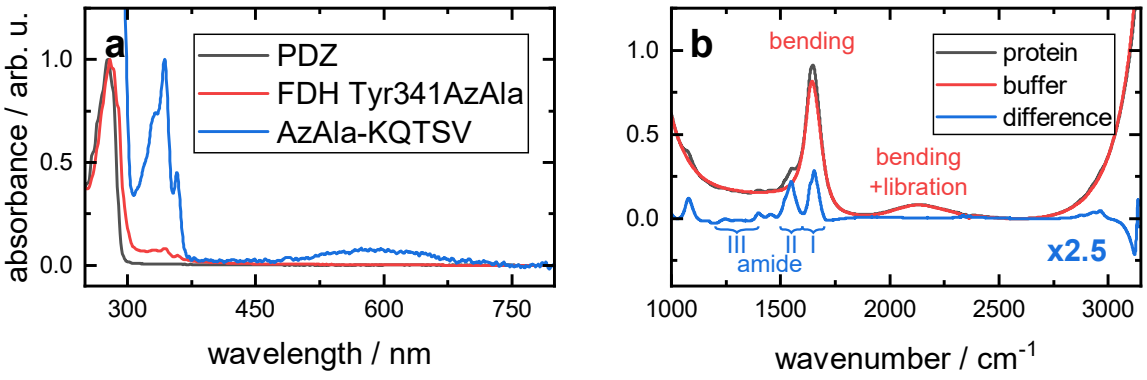

Figure 1.1: *UV/Vis spectra of different proteins and peptides (a). Native proteins without additional chromophores absorb around 280 nm. Specific UV/Vis labels should have their absorption above 300 nm. In the FTIR range (b), native proteins mainly absorb in the amide region whereas aqueous solutions show multiple vibrational bands. Specific IR labels should have their absorption in the transparent window between 1800 cm^{-1} and 2500 cm^{-1} where no native protein and only weak solvent absorptions occur.*

so on [10]. This way first the direct surrounding of the localized mode experiences a rise in the population density of lower-frequency modes and only the propagation to higher tiers leads to the VET over longer distances via increasingly delocalized LFMs.

Both models, the master equation ansatz as well as the descriptions of VET in the normal mode basis, can help to identify paths of energy transfer, though they are based on harmonic approximations and fall short of anharmonic contributions which significantly influence the VET in experiments. The IVR (and ultimately intermolecular vibrational redistribution to the solvent) that leads the system from a non-equilibrium situation back to a Boltzmann distribution of the populated vibrational states relies on this anharmonicity [24, 64, 76–79]. From the normal mode picture it is obvious that for energy flow from one mode to another a certain spatial overlap of the involved modes is required. For that reason, VET relies on local properties (the local density of states and local anharmonic coupling constants) and is in general anisotropic [10, 13, 64, 65, 80–82].

1.4 Protein Labeling for Spectroscopy

All proteins exhibit some general spectroscopic properties. In the UV region these are peptide bond absorptions around 190–230 nm and absorptions of the aromatic amino acids and disulfide bonds around 280 nm (Fig. 1.1a) [83–85]. In the infrared region it is mainly the various amide bands (amide A and B: 3300 and 3070 cm^{-1}; amide I: 1650 cm^{-1}; amide II: 1550 cm^{-1}; amide III: 1200–1400 cm^{-1}) that contribute to the protein spectrum and few side chains that absorb in the same regions or slightly redshifted (Fig. 1.1b) [86]. In order to gather structural information one can infer secondary structures from the circular dichroism (CD) spectroscopic response of the peptide bonds [87, 88] or from the band shapes of the amide bands [89–91]. Both methods rely on an observable that is caused from atom groups that are dispersed over the protein. The signals and information that can be obtained represent an average signal of the whole protein and allow to characterize the overall protein structure but contain no localized information.

If a protein contains a chromophore, prosthetic group or ligand with a distinct absorption, this may help to overcome this limitation. The absorption of the chromophore of green fluorescent protein can indicate local electronic fields [92]. The absorption of a heme group can indicate the oxygenation state and type of ligand bound to the protein [93, 94] and the IR absorption of the ligand itself can indicate different conformations or reveal local dynamics [95–97]. Furthermore, the chemical modification of a ligand can be used to artificially insert a distinct label absorption [98, 99]. But all these external labels restrict the investigations to one specific (binding-) site within the protein.

Vibrational labels are of special interest, as they are smaller and their absorption is more sensitive to the surrounding than that of electronic probes. Besides the amide bands, the absorption of a good vibrational label needs to be separated from intense IR bands of the solvent. In the case of proteins the typical solvent is an aqueous buffer which allows to control the pH of the solution and mimics the native environment of the protein. H_2O exhibits strong and broad absorption band around 3404 cm^{-1}, a weak and broad band around 2128 cm^{-1}, and two narrow bands at 1644 and 1550 cm^{-1} representing a stretching, a combination (bending and libration), and two bending modes, respectively [100]. For that reason most vibrational probes have their characteristic absorption in a transparent window between 1800 and 2500 cm^{-1}, only encountering the weak H_2O combination band [101–103].

To spectroscopically investigate proteins that do not bind exploitable ligands and for investigation of arbitrary protein sites a modification of the protein is inevitable. Either one can post-translationally modify the protein or insert non-canonical amino acids (ncAAs) co-translationally.[5] For flexible and site-specific labeling both methods require modification of the gene that translates to the target protein.

For the direct incorporation of ncAAs into short peptides chemical peptide synthesis (e.g., solid phase peptide synthesis [SPPS]) techniques can be used [106–108]. Despite the development of ligation techniques, the direct chemical synthesis of larger proteins is still challenging [109]. Cell-free expression systems are being developed but are not yet standard solutions for every protein expression [110]. To date the most popular incorporation strategy for ncAAs is the recombinant expression of proteins in a suitable host cell. Most often expression in *Escherichia coli* (*E. coli*) cells is possible, the system is simple to maintain and high yields are attainable at a low cost [111].

There are several reviews of post-translational labeling methods [112–114], especially for fluorescence labels. Various methods rely on Tyr [114], Met [115], Lys [116], and Cys [117, 118] modifications. The modification of cystein to form a thiocyanate is especially attractive for vibrational labeling due to its distinct IR absorption around 2160 cm^{-1} [119]. For all of these methods the target protein may only contain a single reactive amino acid at an accessible position. This restricts the proteins that can be investigated to those that completely lack a certain amino acid, either naturally or by targeted mutagenesis.

The selective pressure incorporation (SPI) method is based on the similarity of the ncAA and one of the proteinogenic amino acids and requires auxotrophic *E. coli* strains. These strains have been genetically modified to lack the ability to synthesize a certain essential component (for SPI an amino acid) and rely on its uptake from external sources [120]. If they are supplied with a sufficiently similar ncAA, the vacant aminoacyl-tRNA synthetase (aaRS) can load the corresponding tRNA with this surrogate and the ncAA is in-

5 For many post-translational labels previous incorporation of an ncAA is still required. This is especially important for specific labeling applications in complex biological environments like cell imaging [104, 105].

corporated during the ribosomal translation at the positions of the vacant codon. Because all occurring codons of the replaced amino acid will be translated to the ncAA, the target proteins are equally restricted as for chemical post-translational labeling.

Finally, one can resort to various types of genetic code expansion techniques [121–123], in this work only stop codon suppression (SCS) was used and is explained in detail.[6] The SCS approach exploits the usage of stop codons for the incorporation of ncAAs. While originally introduced *in vitro* [126] it is now routinely used *in vivo*.

At one of the three stop codons (amber, ochre, and opal) the ribosomal translation usually terminates. This process is initiated by binding of either release factor (RF) 1 in response to the amber or ochre codon or RF 2 in response to the ochre or opal codon [110]. To suppress the termination, a tRNA with the anticodon of one of the stop codons can be introduced to the cell, this tRNA will then compete with the RFs for the binding to the stop codon. The amber stop codon (TAG or UAG in terms of RNA bases) is the least abundant stop codon *E. coli* [123] and is therefore most frequently used for SCS.

Furthermore, an aaRS is required that loads the desired ncAA to the suppressor tRNA. The affinity of the aaRS to the ncAA and the tRNA needs to be high and at the same time highly specific. This prevents the loading of a natural tRNA with the ncAA and the loading of the suppressor tRNA with a natural amino acid. Only if this specificity is given the aaRS/tRNA pair is truly orthogonal to the translation of the expression host. There are different methods to design comprehensive libraries of aaRS/tRNA pairs which then get screened and further optimized for these properties [127]. With such an optimized orthogonal aaRS/tRNA pair the ncAA can then be incorporated co-translationally into a protein at any desired position by exchanging the native amino acid codon for the suppressed stop codon.

As the SCS approach leads to a competition between the RFs and the loaded suppressor tRNA, it always leads to the expression of truncated protein when the RF outcompetes the tRNA. Different aspects have been addressed to improve the SCS. Various strains have been developed that do not express RF 1. For this purpose the amber stop codon was eliminated from their genome [128] or partially removed to counteract reduced viability [129]. Various studies have presented proof for context effects, effects of adjacent or close-by bases and codons, but clear conclusions regarding the different mechanisms and a consensus on the optimal context have not yet emerged [130, 131].

1.5 Measuring Vibrational Energy Transfer with Transient Infrared Spectroscopy

Comprehensive investigation of VET requires the site-specific deposition of vibrational energy. *In silico* arbitrary groups of atoms can be coupled to an external thermostat, this can not be achieved experimentally. Similar to the detection of dynamics, as described above, a distinctive absorption of a localized atomic group is required to selectively initiate VET from a clearly defined position, this group is the VET donor. The mechanism by which vibrational energy is transferred from the VET donor is explained in this section and schematically represented in Fig. 1.2.

6 Other methods for genetic code expansion are the use of quadruplet/frameshift codons, or creation of artificial codons with unnatural base pairs. Furthermore, recently an *E. coli* strain devoid of certain codons in its genome has been synthetically developed, creating free codons not by expanding the genetic code but by reducing the codon usage of some natural codons [124, 125].

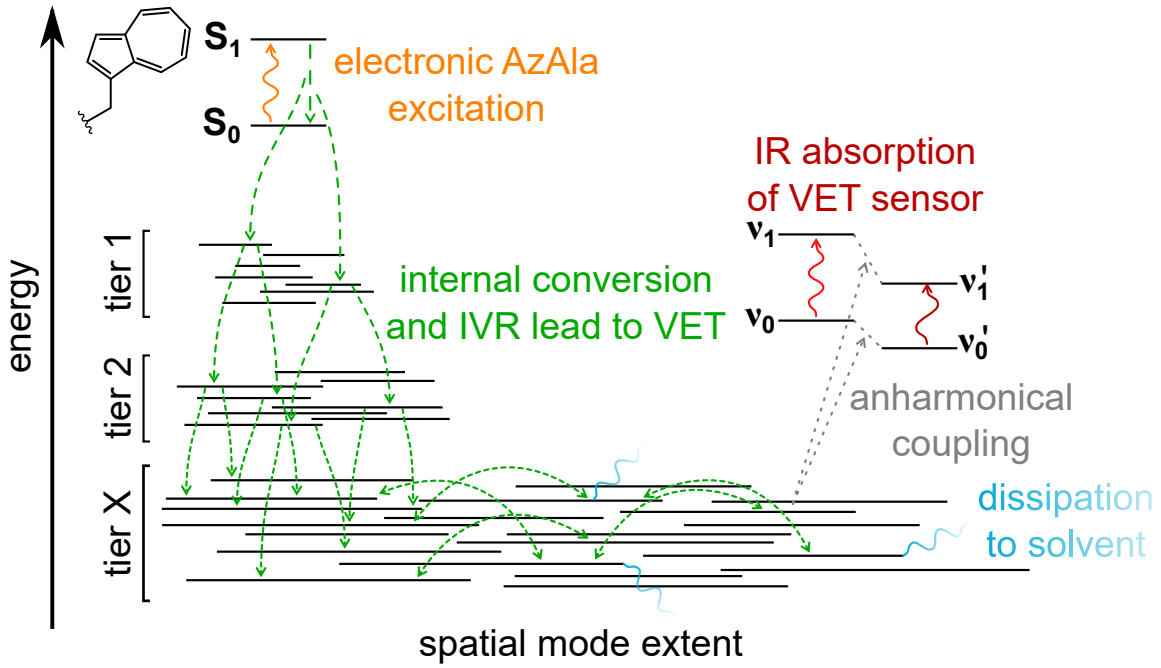

Figure 1.2: *Schematic of the VET mechanism as occurring in a TRIR experiment. Horizontal bars indicate molecular modes, only the electronic energies of AzAla and vibrational energies of the VET sensor mode are represented by two energy levels, the width of the bars indicates their localization. After electronic excitation (orange) AzAla undergoes ultrafast IC and IVR (green). The vibrational energy reaches LFMs after traversing multiple tiers of more localized higher-frequency modes. The less localized LFMs allow VET to distant sites of the molecule. Due to anharmonical couplings (gray) to the LFMs the energy levels of the VET sensor's IR mode shift and the transition frequency (red) gets redshifted. Due to collisions with the solvent and coupling to solvent modes energy dissipates from the molecule to the solvent (fading blue). Inspired by [27].*

For some VET experiments IR excitation of specific residues in small molecules has been used as VET donor and the energy transfer probed with the Raman or IR absorption of another atom group, the VET sensor [132]. The anti-Stokes Raman probe scheme can add quantitative information about the excited state populations, otherwise, the dynamic information accessible with these experiments is the same as with the TRIR method that was used in this work. The signal size and with it the maximal range between VET donor and VET sensor is limited by the photon energy that can be deposited by the excitation photon. For that reason the use of a UV/Vis-excitable VET donor is beneficial. In the Vis range heme groups in proteins have conveniently been used as VET donors but, as mentioned before, restrict the site of excitation [82]. A more flexible VET donor is the azobenzene moiety that can be incoporated into peptides and proteins, however, this donor additionally undergoes isomerization upon excitation which may be unwanted in a densely-packed protein environment [133].

For the experiments in this work an azulene moiety was used as VET donor. This small compound and its organic derivatives have a broad absorption feature around 600 nm off from the native protein absorption. Excitation in this band leads to an $S_1 \leftarrow S_0$ transition (orange in Fig. 1.2) after which ultrafast internal conversion (IC) occurs on a timescale of 1 ps [26]. The IC leads to a rapid conversion of the electronic energy into vibrational energy

that is clearly localized in the azulene moiety. With these properties the azulene provides a means to initialize VET with a sharp temporal onset from a defined spatial position.

For the application of this moiety in proteins, the Trp analog ncAA β-(1-azulenyl)-L-alanine (AzAla) has been developed [134]. Though originally intended as fluorescence label, this has paved the way for the use as flexible and small VET donor, as it shares the characteristic absorptions of azulene (Fig. A.3).[7] An enzymatic method for the synthesis of AzAla was recently developed that allows simple large-scale synthesis of AzAla [135]. By now AzAla has been incorporated as VET donor in various small synthetic peptides [29, 68, 136] and co-translationally in a protein domain after the development of an orthogonal tRNA/aaRS pair [137].[8]

After the IC in AzAla, IVR makes the vibrational energy spread through the molecule (green in Fig. 1.2). The energy levels of the anharmonic oscillator presented in Eq. 1.2 are only valid for a completely isolated anharmonic oscillator, but in a real molecule the vibrational modes are not independent. The frequency ω of a photon exciting the molecule to a higher level $\nu_k + 1$ of a mode k depends on the excitation level of the other modes ν_i via the off-diagonal anharmonic constants x_{ik} and is given as [24]:

$$\omega(\nu_k \rightarrow \nu_k + 1) = \omega_k + 2x_{kk} + \sum_{i \neq k} x_{ik}/2 + 2x_{kk}\nu_k + \sum_{i \neq k} x_{ik}\nu_i \tag{1.5}$$

where ω_k is the harmonic frequency of mode k.[9]

The explicit determination or calculation of the coupling constants is challenging or even unfeasible in large systems, such as proteins. Generally most of the coupling constants are negative, leading mostly to a redshift of a mode when the populations of coupled LFMs increase through heating of the molecule or through IVR [24]. Through this mechanism, the IR absorption (red in Fig. 1.2) of a VET sensor at a certain site shifts during a TRIR experiment (light gray in Fig. 1.2). It thereby reports the arrival of vibrational energy in LFMs that are spatially close and anharmonically coupled to its reporter mode.

The local population density of LFMs reaches a maximum after a certain time (peak time). But the IVR process continues and energy spreads further through the molecule. Leading to the decay of the local population density and concomitantly the IR absorption of the VET sensor shifts back to the original frequency. The vibrations of the molecule are not only coupled among each other, but also to vibrational modes of the solvent. On long timescales the energy that was originally deposited within the sample molecules reaches the solvent whose vibrational modes may also experience a shift and create a background signal (fading blue in Fig. 1.2). The back-transfer from the solvent to the solute is slow and negligible in the experimental time frame so that a background signal remains quasi-constant [66, 68, 137].

7 The fluorescence of AzAla occurs from the S_2 state which is reached by excitation at 342 nm, sufficiently separated from the excitation wavelength of native Trp residues.

8 Furthermore, a Trp auxotrophic strain of *Lactococcus lactis* was developed for SPI of AzAla [138].

9 The notation of the original source was adapted to match that of the formulas in Sec. 1.1.

2 Experimental Methods

2.1 TRIR Laser Setup

The experimental laser setup is based on a commercial titanium-sapphire (Ti:Sa) regenerative amplifier system (Spitfire Ace, Spectra Physics). The Ti:Sa laser medium is pumped by a 527 nm (doubled frequency) diode-pumped neodymium-doped yttrium lithium fluoride laser (Empower 45, Spectra Physics) and seeded with a 797 nm near-IR (NIR) pulse from a Ti:Sa oscillator (Mai Tai SP, Spectra Physics). The seed pulses are chirped (thereby temporally stretched) before amplification in the pumped gain medium, the amplified pulses are compressed again resulting in an amplifier output of 1.7 mJ pulses with 100 fs pulse duration at a repetition rate of 3 kHz.

The pulses generated in the amplifier system are separated into two beams of 0.25 and 1.5 mJ to pump the IR-probe OPA and the Vis-pump OPA, respectively.[1] Both home-built OPAs are based on a two-stage optical parametric amplification in a β-barium borate (BBO) crystal [141, 142]. For a small amount of the initial NIR-pump beam the polarization is turned by 90° with a half-wave plate and a white-light continuum is generated by focusing the beam into an yttrium aluminum garnet (YAG) crystal. The IR component of the white-light is used to seed the first amplification step. Another fraction of the NIR-pump light overlaps spatially and temporally (controlled by a manual delay stage) with the white-light seed in the BBO. By controlling the angle of the BBO the type II phase-matching conditions can be tuned to match different wavelengths of the white-light and thereby select the output wavelengths of signal and idler. The signal (controlled by a manual delay stage) is redirected into the BBO as seed for the second amplification step with the remaining NIR. Signal and idler are then separated in a Michelson interferometer with a variable path length for the idler beam. In the IR-probe OPA both are focused into a silver thiogallate (AGS) crystal for difference frequency generation completing the conversion of the initial NIR light into mid-IR (MIR). A long-pass filter is used to isolate the MIR from residual NIR. The MIR is spatially separated into a probe and a reference beam after passing a polarizer that ensures a predefined polarization. Both beams are focused (~80 µm full width at half maximum [FWHM]) into the sample by a parabolic gold mirror and afterwards recollimated by a second. They are dispersed on a spectrometer grating (Triax 180, Jobin Yvon) with 150, or 300 $^{\text{lines}}/_{\text{mm}}$ and detected on two arrays of four × sixteen mercury cadmium telluride pixels (InfraRed Associates) that are readout with an FPAS-0144 gated integrator (Infrared Systems Development).[2]

The pump OPA was originally designed and operated in the same manner (just not focusing into the AGS due to the high energies). The OPA and the subsequent beam path were modified to generate Vis light in the pump OPA and facilitate the TRIR experiments presented in this work, these modifications were implemented in collaboration with E. Deniz.

1 K. B. Eberl has built the IR-probe optical parametric amplifier (OPA) [139] and T. Vogt has built an IR-pump OPA (that was modified for Vis generation during this thesis) and has designed and implemented the detection setup and hardware [140] that were used.

2 The python-based software that was used to acquire the TRIR data was developed by A. Kondratiev and A. R. Thun [143].

For the generation of the Vis-pump light the second-harmonic of the signal beam is generated. The idler beam can selectively be blocked within the interferometer to keep only the signal beam propagating. Within a BBO type I phase-matching leads to the second-harmonic generation of the signal that results in a tunable wavelength in a range of ~550–750 nm. A variable neutral density filter allows the attenuation of the pump intensity. Propagation through 10 cm calcium fluoride (CaF_2) slightly stretches the pulse to allow the usage of higher pulse energies without damaging the sample. Control of the delay between the probe and pump pulses is granted by a mechanical delay stage (M-406.6PD, Physik Instrumente) in the pump path. Behind the stage a lens is used to focus the pump light into the sample (~125 µm FWHM) and a polarizer can be used to adjust the angle between the pump and probe polarization. The pump light transmitted by the sample is dumped in a filter to prevent higher orders of the pump light of reaching the IR detector and interfering with the measurement [31]. A chopper in the white-light path of the pump OPA blocks every other shot, allowing to consecutively measure pump-on and pump-off IR spectra to calculate the TRIR difference spectra.

2.2 VET Data Analysis Software

The analysis of the VET data was performed using various self-written matlab scripts and functions. Some of the implemented functionalities are different types of weighted averaging to TRIR or broadband pump/probe data including selection of certain scans or frames (specific delays within a scan) to be discarded. The data within the scans and of the different scans can be averaged in many combinations. For all spectra shown in this dissertation, the pump/probe difference signals were calculated from pairs of consecutive spectra within the scans. The average of multiple scans were calculated by pixel-wise weighting the spectra with the inverse of the variance within the respective scan (this is referred to as method 1 in ref. [143]). There are functions to load and save data in the various formats that are being used in the Bredenbeck group. Functions to perform the different types of background corrections (on single data sets or for split sample cell data). A graphical user interface allows to conveniently inspect the data. These scripts and functions have been documented and a descriptive tutorial of the workflow has been prepared and are available in the Bredenbeck group.

2.3 General Microbiological Procedures

2.3.1 Plasmid Preparation and Mutagenesis

For plasmid replication *E. coli* DH5α cells were transformed with the required plasmid via electroporation. The cells were grown and the plasmids prepared with the AxyPrep Plasmid Miniprep Kit (Axygen scientific) according to the product's spin protocol. Plasmid concentrations were determined by UV/Vis spectroscopy and stored at -25°C until use.

For site-directed mutagenesis the Q5 site-directed mutagenesis kit (New England Biolabs) was used; primers were designed with the associated NEBaseChanger online tool. Volumes were reduced 3–4 fold and aliquots of cells divided accordingly all other procedures were executed according to the products's protocol. Plasmid sequences and correct mutagenesis were controlled via Sanger sequencing conducted by commercial providers (Eurofins Genomics or Microsynth AG). The sequencing results were manually assessed and evaluated to select a single correct plasmid for downstream applications.

2.3.2 Protein Expression and Purification

The following paragraphs provide a general overview on the steps involved in the protein expression. The exact culture volumes of different expression runs varied, e.g., between small-scale test expressions and full-scale expressions for the production of large protein amounts, but a given description is sufficiently detailed to reproduce the results also in other labs with different standard protocols. The details that vary between the expressions of different proteins (e.g., plasmids and added ncAAs) are explained in the sections treating those respective proteins and methods.

For protein expressions the plasmid harboring the gene for the protein of interest (Tab. A.8) was transformed into an appropriate *E. coli* strain (Tab. A.9) via electroporation. For SCS-expressions these strains had previously been transformed with the plasmid encoding the aaRS/tRNA pair for the desired ncAA incorporation. After shaking incubation for 45–60 min at 37°C the cells were plated on an agar plate and incubated overnight at 37°C. Single colonies were picked with a sterile tooth pick or pipette tip and used to inoculate ~5–10 ml of Lysogeny broth (LB) which were shaken and incubated at 37°C overnight. The cells were transferred into LB medium containing 10% glycerol and shock frozen in liquid nitrogen (LN_2) before storage at -70°C until further use.

For an expression a ~50 ml starter culture in LB broth were incoculated from the frozen bacterial stock and shaken and incubated overnight at 37°C. A larger volume of expression broth (LB, terrific broth [TB] or ZYP auto-induction medium [144]) was inoculated by adding 1% starter culture and was shaken and incubated at 37°C. The optical density of the culture at 600 nm (OD_{600}) was regularly quantified until it was in the range of 0.6–1 OD, then 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce the expression of the target protein which was regulated by the lac operator. In ZYP medium no manual induction was required, this medium automatically induces the expression once the provided stock of glucose is depleted. The expression continued overnight at 37°C in TB and ZYP medium or at room temperature in LB medium.

After expression the cells were harvested by centrifugation at 8000 rpm at 4°C (Sorvall RC-5B), resuspended in buffer and lysed via ultrasonication on ice. 0.5% $^{w}/_{v}$ streptomycin sulfate were added to precipitate the DNA while stirring on ice for 45 min. The DNA and cell debris were removed via ultracentrifugation at 35000 rpm for 45 min at 4°C (L7-65 ultracentrifuge with type 45 Ti rotor, beckman) and filtered, before the purification with immobilized metal affinity chromatography on nickel-nitrilotriacetic acid (Ni-NTA) resin. After purification the proteins were dialyzed against a final storage buffer overnight, concentrated in a centrifugal concentrator, flash frozen in LN_2, and stored at -70°C until further use.

All culture media (except in the first growth phase post transformation) were supplemented with selective antibiotics, ensuring the exclusive growth of cells containing the appropriate plasmids (Tab. A.8). Antibiotic concentrations were: ampicillin (amp) 100 µg/ml, kanamycin (kan) 50 µg/ml (100 µg/ml in ZYP medium), streptomycin (str) 50 µg/ml, spectinomycin (spec) 100 µg/ml.

2.4 Dipeptides

2.4.1 Synthesis of Azulenylalanine

The synthesis of AzAla was first established in the Bredenbeck lab in the course of the master thesis of V. G. Franz that was supervised by the author [145]. The synthesis is based on the procedure published in [135] where further details can be found.

In short: the plasmid pET22b(+) *Tm*Azul carrying the gene for an engineered enzyme (based on the tryptophan synthase β-subunit [TrpB] of *Termotoga maritima* [*Tm*]) was a present from F. H. Arnold (Caltech, Pasadena). This plasmid was transformed into electrocompetent *E. coli* BL21 Star (DE3) cells. The expression of *Tm*Azul was done in TB overnight at room temperature. The harvested cells were thermally lysed at 75°C for 1 h, cell debris were removed by ultracentrifugation with 11000 rpm at 4°C for 20 min. The supernatant was used for the biochemical synthesis of AzAla from azulene and L-serine in potassium phosphate buffer (KPi) at 37°C for 72–96 h. The product solution was washed with ethyl acetate, this way remaining azulene could be removed in the organic phase and AzAla was retained as precipitate and viscous aqueous phase. The remaining solvent was evaporated *in vacuo* and residual dimethyl sulfoxide (DMSO) was removed with excess ethyl acetate and the product dried. In the original publication this crude product was further purified by reversed-phase chromatography. This last step was omitted in the present work. The formation of the correct product was established from NMR spectra recorded by the NMR service of the Schwalbe group at Goethe University. The crude AzAla could be used for SCS in *E. coli* or was forwarded for peptide synthesis.

2.4.2 Synthesis of 4-Cyano-Tryptophan

The synthesis of the ncAA 4-cyano-tryptophan (CNTrp) was first established in the Bredenbeck lab in the course of the master thesis of C. Feid that was supervised by the author [146]. The synthesis is based on the procedure published in [147] where further details can be found.

The plasmid pET22b(+) *Tm*9D8* carrying the gene for another engineered variant of TrpB was a present from F. H. Arnold (Caltech, Pasadena). The preparation of *Tm*9D8* for the synthesis of CNTrp was done analogous to that of *Tm*Azul. However, the synthesis of CNTrp requires 4-cyanoindole instead of azulene. The synthesis mixture was incubated at a higher temperature of 55°C. The CNTrp was precipitated by cooling the mixture on ice. The CNTrp was filtered and washed twice with ethyl acetate before drying. The formation of the correct product was established from NMR spectra recorded by the NMR service of the Schwalbe group at Goethe University. The final product was forwarded for peptide synthesis.

2.4.3 Peptide Synthesis

All dipeptides used in this work were synthesized via standard SPPS techniques by commercial sources (sb-PEPTIDE [formerly Smart-Bioscience], Saint Egrève) and used without further purification. The ncAAs AzAla and CNTrp were synthesized in-house and provided to the commercial suppliers for the use in the custom peptide synthesis.

2.4.4 Sample Preparation and Experimental Procedures

Sample Preparation. The peptide samples were prepared by dissolving 1 mg aliquots (provided by the supplier) of the dipeptides in 60 µl of DMSO, or tetrahydrofuran (THF). Aqueous samples were prepared by dissolving 1 mg in 20 µl 250 mM NaOH, and dilution with 40 µl H_2O.

For VET experiments the samples were filled into a recess cut into a 100 µm polytetrafluoroethylene (PTFE) spacer between two CaF_2 windows in a brass cell.

UV/Vis Spectroscopy. UV/Vis spectra were recorded on a Hitachi U-2000 spectrometer against the respective solvent background. The sample concentration was determined from the absorption at 612 nm with extinction coefficients ε_{612} of 328, 360, and 317 $M^{-1}cm^{-1}$ in DMSO, THF, and H_2O, respectively).

The extinction coefficients of AzAla in the different solvents were determined from the AzAla-Aha dipeptide. The dipeptide was dissolved in H_2O and the UV/Vis spectrum was recorded in triplicate for three dilutions of this stock solution with DMSO, THF, and H_2O ($^3/_{700}$, $^1/_{140}$, $^1/_{100}$); the triplicates were averaged. The concentration in the dilutions was calculated from the aqueous spectra with the literature value of the extinction coefficient of ~4200 $M^{-1}cm^{-1}$ at 342 nm [134]. A linear fit to the absorption at the different concentrations yielded the extinction coefficient at each wavelength (Fig. A.3).Note that the UV/Vis spectrum of AzAla experiences significant changes after prolonged time in solution in H_2O or 0.5 M HCl, but not in DMSO or 0.5 M KOH (Fig. A.4). Similar changes have been observed for azulene in solution and were attributed to changes in the protonation state, stacking and other effects [148].

FTIR Spectroscopy. The FTIR spectra of the dipeptides in THF were recorded from the 100 µm laser samples, those of samples in H_2O or DMSO were recorded with a 60 µm groove cuvette. The FTIR spectra were recorded on a Bruker Tensor 27 FTIR spectrometer with 1 cm^{-1} resolution. The spectra were corrected for the background by subtracting a spectrum of the respective solvent. Subsequently, a low-order polynomial function was fit to the spectral region around the label absorption (which was omitted from the fit) to finalize a flat baseline. Second derivatives were calculated from the spectra and the solvent spectra separately using the OriginPro 2021b software. The derivatives were smoothed with a Savitzky-Golay filter with a 20 datapoint window and a quadratic polynomial before the subtraction of both derivatives.

VET Experiments. Ultrafast TRIR experiments were carried out on the laser setup described in Sec. 2.1. The output of the pump OPA was regulated with a variable neutral density filter wheel to 6 µJ pulse energy at the sample. The sample cells were mounted in a two-axis scanning device. During the experiment the sample was moved within the focal plane of the laser beams in a Lissajous curve, for regular exchange of the sample.

For all TRIR experiments the background spectrum (pump-on minus pump-off) at a pump/probe delay of -40 ps was subtracted from all other delays. The data in THF and DMSO were analyzed without further corrections. The TRIR data from samples in H_2O show a strong water signal at long delays. These signals arise from energy dissipation from the protein into the solvent leading to difference signals of the combination band in the probe region. This background was subtracted by averaging the spectra at delays after the decay of the probe signal, this spectrum was then fit to each other spectrum with a scaling

and an offset parameter, the pixels carrying the VET signal were excluded for the fitting of the parameters. At each delay the background spectrum was scaled and shifted according to the fit and subtracted from the original spectrum.

The dynamics of the TRIR signals were analyzed by integrating the absolute signal amplitudes along the probe axis in the range of the probe signal at each delay. To pinpoint the time of the maximum signal, the resulting trace was normalized to the maximum and a biexponential function was fit to the region around that peak with a relative intensity >0.65 allowing to interpolate the maximum to determine the peak time.

The shift analysis for the dipeptides was based on the equations described in reference [149], the given numbers refer to the numbering of equations in that article. The following procedure was implemented in a matlab script. For given beam parameters, the pump photon flux $\bar{J}$ was calculated according to:

$$\bar{J}(D) = \frac{4\ln(2)P_0}{E_0\pi\Delta^2}\exp\left(-\frac{D_x^2 + D_y^2}{2s^2}\right) \tag{15}$$

with the pulse power P_0, photon energy E_0, and the FWHM Δ of the beam.[3]

The beam radius that was considered was 3 times the FWHM of the pump beam with a step size of 10 nm to ensure high numerical accuracy. At every point along the beam radius the fraction of excited molecules n_1 was calculated (assuming a quantum yield φ of 1) and integrated over the polarization orientations x via the `trapz` command with a step size of 0.001 to determine the ensemble-averaged photolyzed fraction $\langle n_1 \rangle$:

$$\langle n_1 \rangle = \frac{1}{2}\int_{-1}^{1}\left(1 - \exp\left(\frac{-3\bar{J}\sigma\varphi\left(1-x^2\right)}{2}\right)\right)dx \tag{13}$$

with the absorption cross section σ which can be computed from the molar extinction coefficient.

The distribution of the probe beam was determined along the radius and the area normalized, this way the increased influence of the inner part of the beam is correctly considered. Weighting of the pump beam excited fraction with the probe beam distribution yields the percentage of molecules that are both pumped and probed. Integration along the beam radius yields the size of the difference signal relative to the FTIR signal.

Along the sample, the pump photon flux is reduced due to the sample absorption. The pump beam attenuation was determined via microelement analysis of 100 steps taken along the sample depth [150]. At each step the calculations described above were performed anew and the photon flux adjusted for the next step according to the relative excitation and sample concentration.

3 Note that the factor ε_{2D} in the publication was omitted. It is already accounted for by the substitution $s = \Delta/(2\sqrt{2ln(2)})$ and the modified version of the formula perfectly reproduces the data shown in the publication.

2.5 PDZ Domain

2.5.1 Protein Expression and Labeling

PDZ Wild Type (WT). The PDZ domains were expressed in *E. coli* BL21 (DE3) according to the general protocol in Sec. 2.3.2. The protein was expressed from a pGDR11 plasmid (from this plasmid all PDZ-mutant plasmids were derived via site-directed mutagenesis; Tab. A.8) together with an N-terminal histidine tag (HisTag) linked to the protein via a recognition site of tobacco etch virus protease (TEV) of the sequence ENLYFQ|G. After harvest, the cells were resuspended in 50 mM 2-amino-2-(hydroxymethyl)propane-1,3-diol (Tris)/HCl pH 8 buffer with 100 mM NaCl and 20 mM imidazole for ultrasonication and DNA precipitation. The filtered protein was applied to a Ni-NTA column. Washed with with 50 mM imidazol and eluted with 300 mM imidazole in the otherwise same buffer. The fractions of the elution were pooled, 5 mM β-mercaptoethanol and a HisTagged TEV was added, and the mixture dialyzed overnight against the Tris/HCl buffer without imidazole. The solution was centrifuged and filtered and again applied to the Ni-NTA, removing the tagged TEV and the cleaved HisTag. The flowthrough was pooled, concentrated, and further purified and the buffer exchanged via size exclusion chromatography (SuperDex PG 75) with 50 mM NaPi pH 6.8. The protein was concentrated, flash frozen in LN_2, and stored at -70°C until further use.

PDZ Aha Variants. For the incorporation of Aha the SPI-approach was utilized [30]. To this end, the Met auxotrophic strain *E. coli* B834 (DE3) was used with a plasmid carrying a Met codon at the incorporation site. After growth in LB medium at 37°C overnight, the cells were washed with and transferred into M9 medium supplemented with riboflavin, nicotinamide, pyridoxine monohydrate, thiamine each at a concentration of 1 $^{mg}/_{l}$. The culture was incubated at room temperature for ~3 h before adding 0.5 mM Aha. At an OD_{600} of 1.5, Lys, Thr, Phe, Leu, Ile, and Val were added as in [151], and the overnight expression was induced with 1 mM IPTG. Cell harvest and purification was conducted as described for PDZ WT.

PDZ CNPhe and N_3Phe Variants. The incorporation of CNPhe and N_3Phe into PDZ was achieved via SCS. The plasmid with an amber stop codon at the target sites of the ncAA was transformed into *E. coli* BL21 Star (DE3) cells containing the plasmid pULTRA-CNF [152, 153] with an additional D286R mutation for increased suppression efficiency [154]. The expression was completed in TB medium, 1 mM N_3Phe or CNPhe were added at an OD_{600} of 0.15. After induction the culture was further incubated overnight at 37°C. Cell harvest and purification was conducted equally to PDZ WT.

PDZ Thiocyanate (SCN) Variant. The expression and purification of PDZ Ile327SCN[4] was done as described for PDZ WT with a plasmid carrying an Ile327Cys mutation. After the purification, the sole Cys residue was cyanylated using 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) and isotopically labeled $K^{13}C^{15}N$ using centrifugal concentrators as described in [155], this approach is based on [119].

4 Quantitative labeling of Phe325Cys did not succeed [145].

2.5.2 Sample Preparation and Experimental Procedures

Crystallography. Protein crystals of PDZ Phe325N_3Phe and Phe325CNPhe were grown as part of a bachelor thesis by J. Lee [156]. The crystals grew with the hanging drop method within 30 days at room temperature. The hanging drops were prepared in a ratio of 1:1 from 15 $^{mg}/_{ml}$ and 5 $^{mg}/_{ml}$ protein stock solutions, respectively, in 10 mM NaCl, 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.2 and the reservoir solution. The reservoir solution was 1.28 M sodium citrate, 0.1 M HEPES pH 7.5.

X-ray diffraction experiments on these crystals have been conducted by G. Wille and the author of this dissertation at beamline MX 14.2 at BESSY II [157]. For the experiments, the protein crystals were fished from the solution with a nylon loop, dipped into the reservoir solution supplemented with 20% glycerol and flash frozen and stored in LN_2 until they were mounted for measurement. The data were analyzed using the software XDS [158] at the site, and Phaser-MR and the automatic refinement tool integrated in the Phenix software suite(Ver. 1.19.2-4158) [159, 160].

FTIR Spectroscopy. FTIR spectra were recorded from the concentrated protein before and after adding the ligand to the solution. The spectra were recorded as described for the dipeptides in H_2O. The second derivatives were corrected for the derivative of the background using the spectrum of PDZ WT after addition of the AzAla-KQTSV ligand.

VET Experiments. For laser experiments a stock solution of the peptide ligand AzAla-KQTSV (synthesized by Peptides & Elephants, Berlin; or sb-PEPTIDE [formerly SmartBioscience], Saint Egrève) was prepared in 50 mM NaPi pH 6.8. The concentration of the stock solution was determined via the absorption at 341 nm. The protein sample was concentrated with a centrifugal concentrator, the concentration was determined by the 280 nm absorption of the protein, and the ligand added in ~20% excess, ensuring >98% saturation of the protein.[5] A PDZ WT sample was simultaneously prepared with concentrations closely matching those of the VET sample (containing the IR probe). Final protein concentrations were ~8–11 mM , only for PDZ Ile327SCN the concentration was 18 mM.

Both samples were filled into the two compartments of a 100 µm split sample cell as described in more detail in [161]. The sample cell was moved in a Lissajous curve during the measurement as described for the dipeptides. After completing one scan of the sample compartment, a linear actuator moved the sample cell to measure the signal in the WT compartment for one scan. The compartment was switched for every other scan to acquire a highly correlated background signal for subsequent correction of the TRIR data.

For both sample compartments, a background correction was performed as described for the aqueous samples of the dipeptides. To better isolate the significantly smaller signals in the proteins, the data from the sample compartment were additionally corrected by fitting the data from the WT compartment with an offset and a scaling parameter to each delay (omitting the region of the VET signal). From the resulting data, the trace was calculated, and interpolated as described above.

5 Based on K_D values determined by isothermal calorimetry (ITC) experiments in the course of two master theses supervised by the author and reported therein [145, 146].

2.6 Cyclophilin A

2.6.1 Protein Expression

The gene for CypA was originally provided by C. Schiene-Fischer (Martin-LutherUniversität, Halle) in a pQE-70 plasmid. The gene was cloned into pET28a by T. Baumann (Technische Universität, Berlin), adding a C-terminal HisTag. The author used this plasmid to express CypA in *E. coli* BL21 Star (DE3) in ZYP medium. The Ni-NTA purification was done in 50 mM NaPi pH 8, 300 mM NaCl, and 20 mM imidazole and eluted with a gradient of imidazole. The product was dialyzed overnight against 100 mM HEPES pH 8, 150 mM NaCl. The protein was concentrated, flash frozen in LN_2 and stored at -70°C until further use.

For the expression of the Gly14AzAla mutant T. Baumann replaced the Gly14 codon of CypA with the amber stop codon via a megaprimer mutagenesis approach [162], and guided the protein expression by the author. The plasmid bearing the mutation was transformed into BL21 Star (DE3) which already contained the pBU16 *Mj*AzuRS plasmid [137]. The starter culture contained 1% glucose, a measure for increased suppression of background expression before induction. 2.5 h after inoculation of the ZYP expression culture 1 mM AzAla was added to the culture.[6] After harvest, the resuspended cells were homogenized (IKA-T10 basic) and disrupted (Microfluidics M-110L Microfluidizer). 1 $^{mg}/_{ml}$ lysozyme and 0.01 $^{mg}/_{ml}$ and 2 mM iron(II) sulfate were added and the solution incubated on ice for 15 min before centrifuging at 17500 rpm for 60 min. The purification was performed in the same way and with the same buffers as described above.

2.6.2 Sample Preparation and Experimental Procedures

Cyclophilin Activity Assay. A chymotrypsin-coupled activity assay was used to observe the activity of CypA [163]. The protein was diluted with 50 mM HEPES pH 8, 100 mM NaCl. The concentration of the protein stock solution was determined from the absorption at 280 nm, assuming ε_{280} of 40 $mM^{-1}cm^{-1}$ for AzAla [164]. A chymotrypsin stock solution was prepared with 100 $^{mg}/_{ml}$ in 1 mM HCl. A 1.9 $^{mg}/_{ml}$ stock solution of the substrate peptide succinyl-AAPF-p-nitroanilide(pNA) was prepared in trifluoroethanol (TFE).

Buffer, chymotrypsin (10 $^{mg}/_{ml}$), and CypA (3.5 nM) were mixed in a 1 cm cuvette and equilibrated to 6°C. The assay was initiated by adding 32 µl substrate solution, resulting in a final reaction volume of 1.3 ml. Under continuous stirring, the absorption of pNA at 390 nm was monitored using a USB4000 fiber spectrometer (Ocean Optics) and a Xe light source. During the experiment, the complete setup was flushed with dry air that was blown onto the cuvette to prevent fogging. The reported reaction speeds were determined from an exponential fit to the data.

ITC Experiments. The dissociation constants of the peptide ligands with the amino acid sequences HAGPIA, Ac-FGPDLP-NH_2, and Ac-PVHAGPIAP-NH_2 (synthesized by GenicBio, Shanghai; used without further purification) were determined via ITC.[7] For the experiments, the protein was diluted in 100 mM HEPES pH 8, 150 mM KCl, 1.5 mM MgCl, 2 mM tris-(2-carboxyethyl)phosphine (TCEP), and 10% glycerol. The lyophilized ligand was dissolved in the same buffer and the pH (influenced by residual acidic salts from the SPPS) adjusted to match that of the protein solution. The concentration of the protein was deter-

6 Note for clarity: 1 mM L-AzAla was added from a 200 mM stock solution of a racemic mixture of AzAla HCl salt from chemical synthesis, for all other purposes biochemically synthesized AzAla (see above) was used.

7 Ac indicates N-terminal acetylation, NH_2 indicates C-terminal amidation.

mined by the absorbance at 280 nm and that of the ligand by its absorbance at 205 nm [85]. The final concentrations were 170–300 µM for the protein and 3.6–5 mM for the ligand. Both solutions were degassed for ~10 min before loading the sample cell and titration syringe of a Malvern Panalytical VP-ITC MicroCalorimeter for the experiment. Injections of 8–10 µl were run with an injection speed of 0.5 $\mu l/s$, the injections were spaced by 300 s to ensure reequilibration between the injections. The temperature was set to 28°C (HAGPIA) or 25°C (other peptides). The data were anlyzed using the AFFINImeter ITC software.

2.7 Formate Dehydrogenase

2.7.1 Protein Expression

Formate Dehydrogenase (FDH) WT and Val123Gly variant. The gene for *Candida boidinii* (*Cb*) FDH (European Nucleotide Archive accession number: KM454879) was synthesized by the Invitrogen GeneArt services (ThermoFisher Scientific) flanked by a BsaI and an XhoI restriction site in a pMK cloning vector. It was subcloned into the multiple cloning site of an empty pET28b plasmid between the XhoI and NcoI site using ElectroLigase (all cloning enzymes were purchased from New England Biolabs). Afterwards, the C-terminal HisTag encoded on the plasmid was joined with the open reading frame by mutating the ochre stop to a Ser (TCA) codon.[8] From this plasmid all mutant plasmids were derived.

The proteins were expressed in *E. coli* BL21 Star (DE3) in TB medium at 37°C overnight. The purification on the Ni-NTA was conducted with 50 mM KPi pH 8, 100 mM KCl, and 20 mM imidazole and eluted with 150 mM imidazole. The pooled fractions of the elution were dialyzed against 100 mM KPi pH 7.5 before flash freezing in LN_2 and storage at -70°C.

FDH AzAla Variants. For the expression of FDH proteins that contain AzAla several strains of *E. coli*, optimized for amber suppression by deletion of RF1 and (selective) deletion of amber codons in the *E. coli* genome, have been used. These strains were *E. coli* B-95.ΔA (RIKEN), *E. coli* B-95.ΔAΔ*fabR* (RIKEN), and *E. coli* C321.ΔA.exp (DE3) (provided by N. Budisa [University of Manitoba, Winnipeg]). For amber SCS these strains were transformed with pBU16 *Mj*AzuRS, providing the optimized aaRS and tRNA for AzAla [137]. The expression of the mutants were tested in all of these strains at a small scale to chose the one strain that was then used for all full-size expressions of each mutant, this was done in the course of the master thesis of Y. Shen that was supervised by the author [168]. The expression medium was supplemented with 1 mM AzAla from a 100 mM stock solution, the stock solution was prepared by dissolving crude AzAla (see Sec. 2.4.1) in 1 M HCl and subsequently diluting the solution 10-fold by addition of H_2O.

Prior to the purification on the Ni-NTA, the proteins were denatured with 7 M urea. The protein was refolded on the column by gradually removing the urea. This additional step was required to remove the remaining truncated FDH, which on its one could not bind to the resin but formed heterodimers with the complete expression product, this behavior was first observed in the course of the bachelor thesis of N. Joachim which was supervised by the author [169]. All subsequent steps were performed as described for FDH WT.

8 This was done, after purification on a blue sepharose 6 fast flow column with previously published protocols did not yield good purification results [165–167].

2.7.2 Sample Preparation and Experimental Procedures

FTIR Spectroscopy. FTIR spectra of the laser samples were recorded in a 60 µm groove cuvette on a Tensor 27 FTIR spectrometer with 1 cm^{-1} resolution. The background of the H_2O combination band was removed by fitting the FTIR spectrum of the reference sample (without N_3^-), a linear baseline and the Lorentzians representing the N_3^- absorption to the spectra. The Lorentzian representing unbound N_3^- had fixed central wavenumber (2048.1 cm^{-1}) and FWHM (27 cm^{-1}) which were determined from the FTIR spectrum of N_3^- in buffer. Protein concentrations were determined from UV/Vis spectra at 280 nm using an extinction coefficient of 40 $mM^{-1}cm^{-1}$ for AzAla [164]. Final N_3^- concentrations were determined from the total area of the fitted peaks, assuming an identical integrated extinction coefficient of 65.5 $mM^{-1}cm^{-2}$ for all bands, which was determined by a concentration series.

VET Experiments. For the VET experiments the FDH proteins were concentrated to up to ~2 mM. Nicotinamide adenine dinucleotide (NAD^+) was added in excess from a 100 mM stock solution. The solution was divided in two fractions. N_3^- (from a 100 mM stock solution) was added to one fraction, the other was topped up with the same volume of buffer to maintain equal protein concentrations in both. A 100 µm or 150 µm thick split sample cell was filled with the two solution (see PDZ), the N_3^--free acting as reference for the background correction. The measurements were done using the 300 $^{lines}/_{mm}$ grating.

In contrast to the PDZ data, the sample compartment was directly corrected with the data from the reference compartment, without previous compartment-wise correction of the solvent signal. For each delay, the signal of the reference compartment was fit to the signal-free pixels (1–10 and 42–64) of the sample compartment with a scaling and an offset paramter. The additional step of individually correcting for the solvent signal was not required, presumably because of the higher intensity of the VET signal relative to the solvent signal. The trace and its dynamics have been analyzed as described for the dipeptides.

FDH Activity Assay. The Michaelis-Menten kinetics with FDH were performed using the same setup as described for CypA above. The assay was performed at room temperature in 100 mM KPi pH 7.5 with varying concentrations of NAD^+ and a saturating concentration of 200 mM sodium formate. In an Eppendorf-tube 989 µl of buffer and substrates were prepared and 11 µl of FDH solution (200 nM final concentration) added. It was thoroughly mixed and quickly transferred into a spectroscopic cell to record the increase of product absorption at 340 nm. The initial reaction rate was determined by a linear fit of the first 10 s of the data using an extinction coefficient of 6220 $M^{-1}cm^{-1}$ for NADH. Each concentration was reproduced in triplicate and parameters of Michaelis-Menten equation were fit to match the data.

For the experiments comparing the activity with and without illumination, the substrate concentrations were 1.5 mM and 200 mM of NAD^+ and sodium formate, respectively. The solution was equilibrated to $3\pm0.4°C$ under constant stirring with an external thermostat and the actual temperature of the sample was measured using a K201 digital thermometer (voltcraft) with its probe submerged in the solution. The reaction was initiated by adding the protein to a final concentration of 150 nM (WT, Tyr196AzAla, Tyr341AzAla), 2.5 µM (Val121AzAla), or 4.5 µM (Tyr341AzAla+VG). For the latter mutants the concentration was increased to improve the detectibility of the product formation, which was slow because of the low activity. The data of 9 min of reaction, after a short mixing time, were analyzed. For experiments under illumination, the sample was pumped with pulses of a Coherent Evolu-

tion laser (526 nm, ~250 ns pulses at 1 kHz) with 3.2 W, the scattered light was blocked with a shortpass filter with 500 nm cut-off (Edmund Optics, #84-693). Though this excitation is in the lower flank of the broad AzAla absorption band, it was the best light source available to reach high excitation powers.

For the analysis, the absorbance readings between 338 nm and 342 nm were averaged and the average from 786 nm to 790 nm (above 770 nm the filter is transmissive) subtracted to reduce noise caused by variations in the probe intensity over the complete spectrum (Fig. 2.1a). Subsequently the data were corrected for upward outliers caused by the occasional fogging of the cuvette. Data points off from a linear fit by more than three standard deviations were cropped iteratively until the standard deviation of the residuals was below 1 mOD, on average 2% of the data were removed by this procedure (Fig. 2.1b). The activity is then determined by the slope of the linear fit.

The absolute activities measured in these experiments on different days varied by about 10% even for nominally identical samples. Presumably because of unidentified differences in the sample conditions (e.g., additional freeze-thaw cycles). Because unilluminated and illuminated activity of the same mutants are compared, the activity was normalized to the unilluminated experiments of each day and experiments from multiple days combined. The results presented are the average of 5 repetitions for FDH WT, 6 repetitions for Val121AzAla and Tyr341AzAla+VG, and 3 repetitions for Tyr196AzAla and Tyr341AzAl, for each the unilluminated and illuminated case.

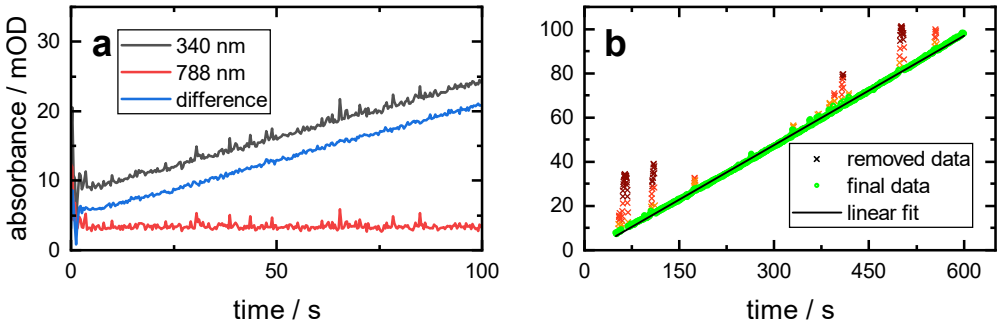

Figure 2.1: *Noise suppression by subtracting the 788 nm data from the 340 nm data, note the fluctuations at early times due to addition of FDH WT and the mixing (a). Example of removing spikes caused by occasional fogging of the cuvette (b). After three iterations of outlier removal (red and orange crosses) no major spikes remain and the final data, in this case 92% of the original data, can be fit to extract the slope which represents the steady-state reaction rate.*

Part II

Labels for Vibrational Energy Transfer

3 Introduction

In this part of the dissertation, the experimental characterization and assessment of six potential VET sensors is presented. Most of the results presented here have been published in the author's publication *Versatile Vibrational Energy Sensors for Proteins* [136].

To investigate the VET dynamics of a protein and the connectivity between two sites, one has to equip the system with a VET donor and a VET sensor. By measuring the peak times and amplitudes of VET signals from VET sensors at various positions, one could ultimately experimentally reproduce the communication maps mentioned in Sec. 1.3. The pair of ncAAs that was originally published to enable such experiments in proteins was the combination of AzAla as VET donor and Aha as VET sensor [29]. The successful site-directed incorporation of the amino acids has two requirements: (i) a method for the incorporation of the amino acid at arbitrary positions, and (ii) the incorporation may not harm the protein fold and function. To map the connectivity in a whole protein, successful incorporation is required at multiple sites.

Since the introduction of the VET pair, the incorporation of AzAla has been improved by the development of an aaRS/tRNA pair that allows the convenient incorporation of the ncAA using SCS [137], thus creating a way to position the VET donor at virtually any site without conceptual restrictions. To fulfill the second criterion, a conservative incorporation should be chosen. The Trp analog constitutes an aromatic residue, the lowest impact of the mutation is therefore expected for the exchange of Trp, Tyr, or Phe residues.

Aha, on the other hand, can only be introduced into proteins by means of SPI [30]. With this approach, all Met in a protein would be exchanged for Aha. In a VET experiment one can not necessarily distinguish between VET signals of multiple Ahas within one protein, a site-resolved observation of VET would not be feasible. This restricts the usage of the Aha VET sensor to proteins that do not contain any Met. The similarity to Met makes Aha substitutions less disturbing. However, a larger set of VET sensors that can be incorporated more flexibly would allow to pick the one sensor that best fits each site in a protein. Only with such a complete set of VET sensors can this technique become applicable to all proteins and the specific requirements of the unique environment of each residue.

The most common IR labels are based on the asymmetric stretching vibration of an azide (N_3) moiety, or on the stretching mode of a nitrile (CN) or thiocyanate (SCN) moiety [101, 102]. The absorption of these modes occurs within the transparent window (see Sec. 1.4) allowing convenient IR spectroscopic investigations. They can be incorporated into peptides and proteins because multiple ncAAs exist that carry, or can chemically be equipped with, one of those moieties. Azido compounds have a favorably high extinction coefficient in comparison to nitriles, which makes them attractive IR labels [103, 170]. In 2D-IR spectroscopy, this extinction coefficient is especially important, because the signal amplitude depends on its *square*. At the same time its short vibrational lifetime of 1 ps limits the timescales that can be covered with this label [39]. Additionally, the band shape of the azido vibration is often distorted from Fermi resonances [171–173], leading to ambigu-

ous interpretation of experimental results. The well-defined nitrile bands have a generally longer lifetime (especially SCN) but a significantly lower extinction coefficient.

Different criteria need to be considered to find a good VET sensor. Fermi resonances or otherwise contorted band shapes are of lesser concern for the analysis of VET dynamics because of the integration of the signal. The amplitude of the VET signal is, however, influenced by the band shape. The presence of multiple species can lead to asymmetric signals, with one contribution (reduced absorption or induced absorption) being more intense than the other. The extinction coefficient influences the amplitude linearly. The coupling strength between the probed mode and the LFMs determines the strength of the shift and thereby the amplitude of the VET signal [24]. Furthermore, the band width of the probed mode determines the signal amplitude. If the VET-induced shift is equally strong in two probes, the amplitude of the VET signal will be larger for that probe with a more narrow band. The vibrational lifetime of the probe does not directly influence the VET signal, as its dynamic is governed by the lifetime of the coupling LFMs, but a longer lifetime may be indicative of weaker coupling constants which also govern the vibrational decay of an excited mode [174]. Finally the sheer size of the probe may influence the dynamics as well as the amplitude of the VET signal. The local LFMs and the density of states will be influenced by the label, an effect that is thought to be less important in proteins but may play a role in smaller peptide systems.

The PDZ protein domain is a recurring protein motif that is often featured in signal transduction and binds C-termini of target proteins [175]. Because it participates in many cellular processes, a fundamental understanding of its functionality is of high interest. How exactly it propagates signals is still a matter of discussion. Interestingly, the structure of this domain is almost unaltered by ligand binding [176], ruling out allostery in the classical sense, i.e., binding of an effector leads to a structural change at a distant protein site and thereby alters chemical properties and ligand interactions at that site. Many recently reviewed studies highlight the role of protein dynamics for allosteric behavior [177, 178]. These include the PDZ domain and the influence of its distal α3 helix on ligand binding [179], employing a broader understanding of allostery as any change (as opposed to only effector binding) that leads to a distal change of protein properties. Furthermore, evolutionary analysis of PDZ [180] indicate a conserved pathway of energetic connectivity within PDZ, that was later similarly identified by thermal diffusion simulations [9]. While the investigation of allostery and energy flow in PDZ in itself is therefore interesting, it is in this thesis employed as a stable protein scaffold to characterize the VET sensors in a full-scale protein system.

Because of the complex interplay of different effects on the VET signal and the involvement of the difficult-to-calculate coupling constants, an experimental characterization of vibrational probes is inevitable for a profound decision for one VET sensor. In this part of the dissertation, and the corresponding publication [136], this characterization is realized for the following IR labels and potential VET sensors:

Azidoalanine (N_3Ala) is the smallest available azido ncAA, the azide moiety is linked to the backbone by a single CH_2 group. So far N_3Ala can not be incorporated into expressed proteins, but only into peptides with SPPS. For that reason it was rarely used in IR spectroscopy [181]. Furthermore, a recent study confirmed that the asymmetric FTIR band shape of this ncAA is caused by a Fermi resonance that can be reduced by isotopic labeling of only the central N atom [172].

Aha has one more CH_2 group in the aliphatic linker than N_3Ala. This amino acid can be incorporated into proteins using SPI [30]. While still restricting the flexible use to Met-free proteins this allowed the study of, e.g., protein folding and ligand binding [39, 182]. This ncAA was the VET sensor originally proposed for VET in proteins [29] and has since been used for this purpose in different peptides and proteins [68, 137, 161, 183].

N_3Phe presents a chance to incorporate an azide moiety with aromatic character because of the phenyl ring within the linker. This ring gives rise to pronounced Fermi resonances [171]. Isotopic labeling of only the central N atom can not sufficiently uncouple the azide mode from the phenyl modes to generate a single band [173], which was the case for phenylazide and 3-azidopyridine [184, 185]. Anharmonic frequency calculations, however, suggest that moving the azide moiety from the para to the meta position may eliminate the Fermi resonance [171]. This ncAA can freely be incorporated into proteins using SCS with an amber codon and an aaRS/tRNA pair that was originally designed for CNPhe [153].

CNPhe is the nitrile analog to N_3Phe. It can equally be incorporated into proteins by amber SCS [152]. It is commonly used as 2D-IR label because of the 4 ps lifetime, which can be prolonged by isotopic labeling [186, 187]. The longer lifetime and the simple band shape comes at the cost of the lower extinction coefficient inherent to nitriles.

CNTrp is a Trp analog with a nitrile moiety. In comparison to CNPhe it features a shorter vibrational lifetime of 1.2–1.7 ps [188] and is somewhat larger because of the indole ring. CNTrp and its isomers have been proposed as fluorescence labels for biological samples, because the additional nitrile group leads to a redshift of its excitation wavelength from that of native Trp [189–191]. Together with their IR properties they may adopt a double role for complementary spectroscopic investigations [188, 192] and we want to investigate if it may also be a valid VET sensor. Until now no SPI or SCS incorporation into recombinant proteins was published, but requested [193]. However, based on preliminary findings, SCS of 5-cyano-tryptophan is possible and first attempts of CNTrp SCS showed a low incorporation yield which may further be increased [194]. Furthermore, CNTrp has recently become broadly available due to the development of a bioenzymatic synthesis technique [147].

SCN is another common IR label for proteins. This label is incorporated by post-translational modification of cystein residues [119]. The incorporation of this label is similarly restricted as that of Aha, due to the requirement of only a single exposed Cys residue. The extraordinarily long lifetime of this label of tens of picoseconds in a protein allows the extended observation of protein dynamics [195]. The extinction coefficient of the SCN in a protein is typically in the range of 30–60 $M^{-1}cm^{-1}$ [195, 196].

For the characterization of the VET sensors dipeptides of the VET donor AzAla and either N_3Ala, Aha, N_3Phe, CNPhe or CNTrp were used (Fig. 3.1a).[1] With these minimal systems the general suitability of the different sensors was investigated in three solvents (H_2O, DMSO, THF) to account for variable protein environments. In such a system, the individual traits of each sensor will have a strong influence on the resulting measurements. For the prospect of using these sensors in a protein environment, however, these fundamental effects may be less influential. In order to assess the sensors under circumstances equal to the future application a PDZ domain was equipped with either Aha, N_3Phe, or CNPhe at

1 Cyanylation of an AzAla-Cys dipeptide was not attempted, because previously a very small compound, acetyl-cysteine, could not be cyanylated, supposedly due to the formation of disulfide bridges between the acetyl-cysteines [197].

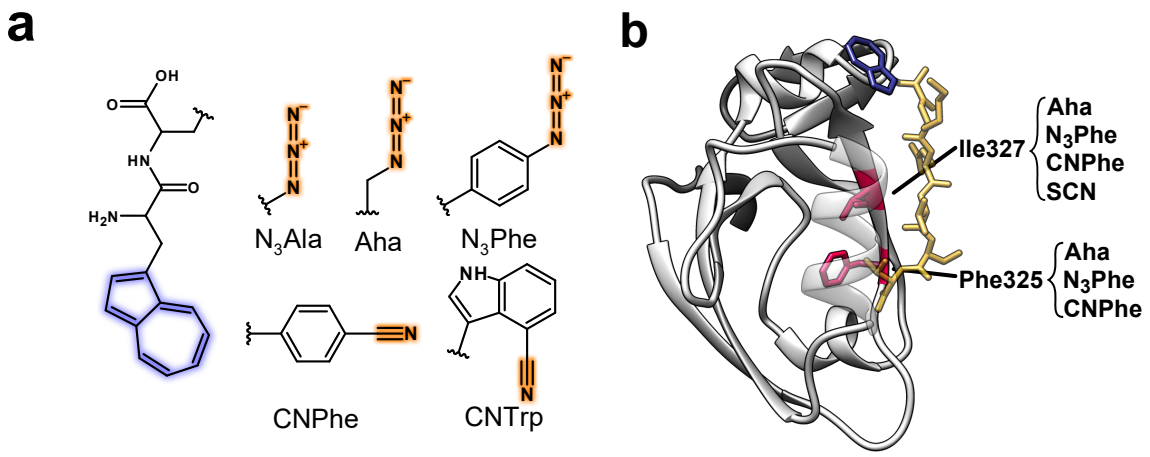

Figure 3.1: *a) Chemical structures of the dipeptide scaffold and the different vibrational probes. Each dipeptide consists of the AzAla moiety (highlighted in blue) for VIS excitation and one of the ncAAs for IR probing (probed atom groups are highlighted in orange). b) Schematic of the PDZ structure (gray) with the ligand (yellow) containing AzAla (blue). The indicated residues (purple) have each individually been replaced by Aha, N_3Phe and CNPhe, Ile327 was additionally replaced with SCN labeled cysteine (structure based on RCSB protein data bank (pdb): 1BE9[176]). Adapted from [136], licensed under CC BY.*

either Ile327 or Phe325, or SCN at the site of Ile327. In these more complex systems, the AzAla VET donor was introduced as the C-terminus of a peptide ligand (Fig. 3.1b).

4 Results and Discussion

4.1 FTIR

4.1.1 Dipeptides

The IR extinction coefficients of the IR labels in the dipeptides are shown in Fig. 4.1, see Tab. A.2 for a complete list of FTIR parameters. The maxima of the azide moieties are situated near 2113 cm^{-1} and those of the nitrile moieties around 2224 cm^{-1}. The determined extinction coefficients of the azide sensors is lower by about ~40% in THF and up to 30% lower for the nitrile sensors in THF. A hypsochromic shift of up to 13 cm^{-1} is observed in H_2O as compared to the aprotic solvents DMSO and THF, in which the central frequency is very similar (within 2 cm^{-1}). This shift is accompanied by an increase of the band width (except for N_3Ala) by up to 80% for CNPhe.

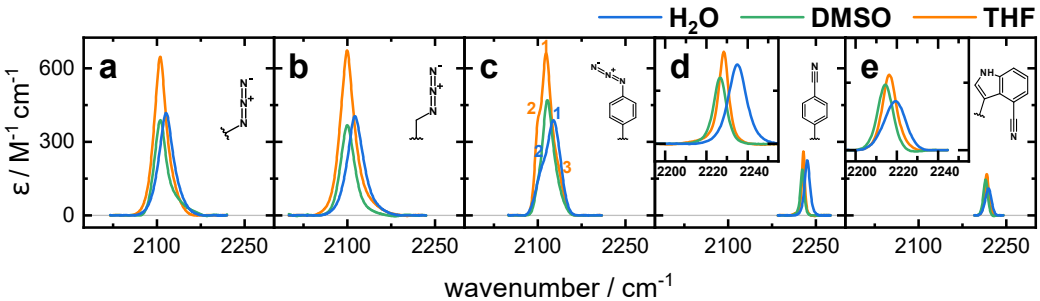

Figure 4.1: *Extinction coefficients of the IR band of the dipeptides containing N_3Ala, Aha, N_3Phe, CNPhe, or CNTrp (indicated by the structures) in different solvents. The spectral range covered in all panels is identical, to better visualize the difference in band width of the different IR probes, the insets show the narrow nitrile band shape in detail. Small numbers in the FTIR spectra of AzAla-N_3Phe indicate the presence of multiple subbands. Adapted from [136], licensed under CC BY.*

For organic azides, it was shown that the wavenumber shift is especially sensitive to the H-bonding and insensitive to the electric fields in the surrounding. Even compared to other protic solvents the shift is most pronounced in H_2O due to the high H-bond density in this solvent [198]. Thus explaining the blueshift of the labels in H_2O as compared to the aprotic solvents, whose different polarities and permittivities only weakly influence the central wavenumber. Both properties, H-bonding and the electric field, have a strong influence on the the nitrile stretching frequency. Therefore, knowledge about the H-bonding status of a nitrile group is required to quantify the electric field it experiences [199]. For CNTrp and CNPhe we observe an incresing redshift from H_2O to THF to DMSO (CNPhe: 2235 cm^{-1},

2228.4 cm^{-1}, 2226.7 cm^{-1}; CNTrp: 2219.3 cm^{-1}, 2216.4 cm^{-1}, 2214.4 cm^{-1}). This behavior has been reported in several previous studies [188, 199, 200].

The band shape of the nitrile absorption are single, symmetric peaks. The bands of the azide labels consist of up to three subbands. The subbands are best visible for N_3Phe in THF and H_2O, for which they are labeled in Fig. 4.1c). These additional bands occur because of Fermi resonances between combination bands (involving also the symmetric N_3 stretch vibration) and the phenyl ring modes [171]. In the N_3Ala and Aha dipeptides no clear subbands occur (not even in the second derivative, Fig.A.5a and b), only slight asymmetries of the absorption bands are visible. Using 2D-IR spectroscopy it was shown that this asymmetry in the band of N_3Ala is also caused by a Fermi resonance [172].

From these spectra alone, it appears that the azide sensors are better suited as VET sensors. They generally feature peak extinction coefficients larger by a factor of ~4 and integrated extinction coefficients larger by a factor of 5–16 than those of the nitriles. Furthermore an incorporation in a hydrophobic core of a protein appears to be advantageous in order to maximize the extinction coefficient. The Fermi resonances do not generally pose a problem for the analysis of the VET dynamics, because the analysis is based on integrated signals. The prediction and comparison of signal amplitudes, however, the Fermi resonances will complicate the prediction and direct comparison of signal amplitudes.

4.1.2 Proteins

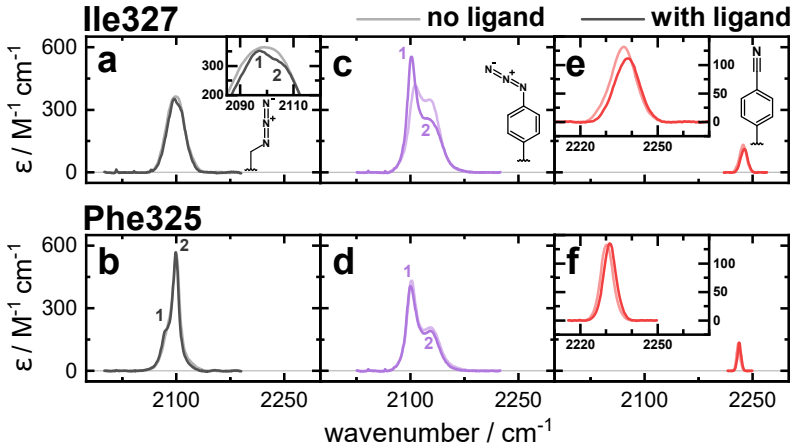

Figure 4.2: *Extinction coefficients of PDZ with Aha (gray), N_3Phe (purple), and CNPhe (red), as indicated by the structures in the top panels, inserted at positions Ile327 (top) and Phe325 (bottom). Each panel shows the extinction coefficient without the ligand (dim colors) and with the ligand (intense colors). Small numbers in the spectra of Aha and N_3Phe containing mutants indicate the presence of multiple subbands. The structure in each panel indicates the VET sensor of the data in the respective panel. Adapted from [136], licensed under CC BY.*

The insertion of Aha into the PDZ scaffold leads to a two-peak structure of the absorption band, which was not observed in the dipeptides. In the spectrum of PDZ Ile327Aha without ligand the two-peak structure is not clearly visible. Ligand binding leads to a relative reduction of the high-wavenumber component making the two peaks more recognizable (inset in Fig. 4.2a). However, the second derivative of the spectrum reveals the presence of both components already before the addition of the ligand (Fig. A.6a). The two-peak structure in the Phe325Aha mutant without ligand is also visible mainly in the second derivative

of the spectrum (Fig. A.6b), whereas the addition of the ligand leads to the formation of a clear shoulder resulting of a -2 cm^{-1} shift of the less intense low-frequency component. This strong component may report on the apolar, hydrophobic core in which the side chain is embedded. In comparison to the similarly intense spectrum of the AzAla-Aha dipeptide the main feature is much more narrow (27.8 cm^{-1} and 9.9 cm^{-1}, respectively; see Tab A.3 for a complete list of FTIR parameters). This extreme band shape exemplifies how the presence of Fermi resonances can distort band shapes and lead to signals that can not be predicted from the spectrum of the IR label alone but only occur in the defined environment of a certain protein site.

The N_3Phe label has a band shape different to those observed in the AzAla-N_3Phe dipeptide. In literature, the central absorption feature was assigned to the asymmetric N_3 stretching vibration [171], matching the spectra of the dipeptide, in which the central feature is most intense in H_2O and THF. In DMSO only two features can be discerned in the second derivative, from which the high-wavenumber feature (2116.5 cm^{-1}) is dominant (Fig. A.5c). Interestingly, the spectrum of N_3Phe at both labeling sites in PDZ shows a two-peak structure in which the low-wavenumber contribution is dominant. In Ile327N_3Phe the binding of the ligand leads to an increase of the low-wavenumber and a broadening of the high-wavenumber component. The increased size of N_3Phe compared to the native Ile residue may have lead to the azide moiety pointing into the binding pocket instead of the protein core. Ligand binding could then reduce the access to the solvent leading to a less polar and more aprotic environment.

In Phe325N_3Phe the ligand binding does not induce a change in the band shape. Protein crystals with the mutations Phe325N_3Phe and Phe325CNPhe in the absence of the ligand have been experimentally resolved (see Chap. i). From these data it is clear, that both ncAAs occupy a single conformation and are buried within the protein. We therefore conclude that the band shape of Phe325N_3Phe is indeed a result of Fermi resonance and not conformational inhomogeneity. Accordingly, it is likely that the very similar band shape of Ile327N_3Phe is also caused by the Fermi resonance. An additional structure of this mutant is required to certainly explain the increased sensitivity to ligand binding.

The single-band CNPhe band at both protein sites is only slightly affected by the ligand binding, only a redshift of 1 cm^{-1} occurs. The band shape of Ile327CNPhe mostly resembles that of the AzAla-CNPhe dipeptide in H_2O. Contradictory to Ile327N_3Phe, the ligand binding leads to a slight blueshift of the band, in principle indicating a more polar or protic environment. However, the shift is not expressive enough for certain interpretation, furthermore, the same shift occurs for Phe325CNPhe whose position was less sensitive to ligand binding with the other ncAAs and which is clearly buried within the protein judged by the crystal structure.

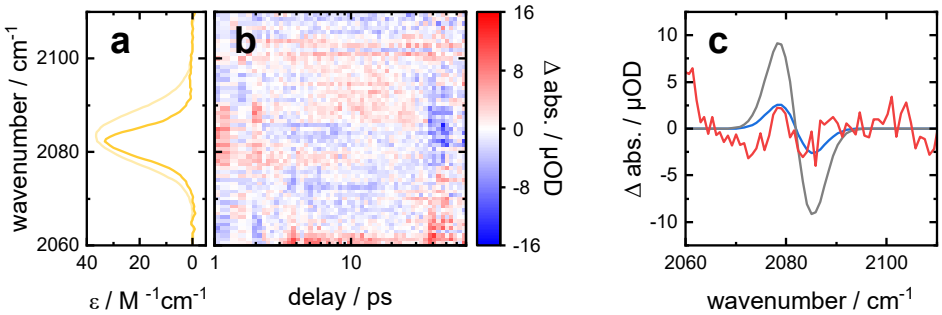

Figure 4.3: *Extinction coefficients of PDZ Ile327SCN before (light yellow) and after adding the AzAla-KQTSV ligand (yellow; a). VET signal of PDZ Ile327SCN (b). Expected VET signal assuming the same shift (gray) or a 3.5 fold lower shift (blue) than observed in Ile327CNPhe and average of the VET signal at 4–8 ps delays (red; c). Adapted from [136], licensed under CC BY.*

Ile327 was also substituted with SCN. The extinction coefficient of this mutant is 33–36 $M^{-1}cm^{-1}$. The FWHM of the absorption is reduced from 12.6 cm^{-1} to 7.9 cm^{-1} upon ligand binding (Fig. 4.3a). These parameters are similar to the ranges observed in PYP mutants [196].

4.2 VET Signals

4.2.1 Dipeptides

The dipeptides show intense, time-dependent VET signals (Fig. A.7–A.8). The spectral position of the signals and also their band shapes reflect those featured by the FTIR spectra. In Fig. 4.4 the signal amplitudes at the maximum signal have been scaled to the peptide concentration. For each dipeptide, the amplitude is lowest in H_2O and the signal in DMSO and THF is larger by a factor of 2–4.

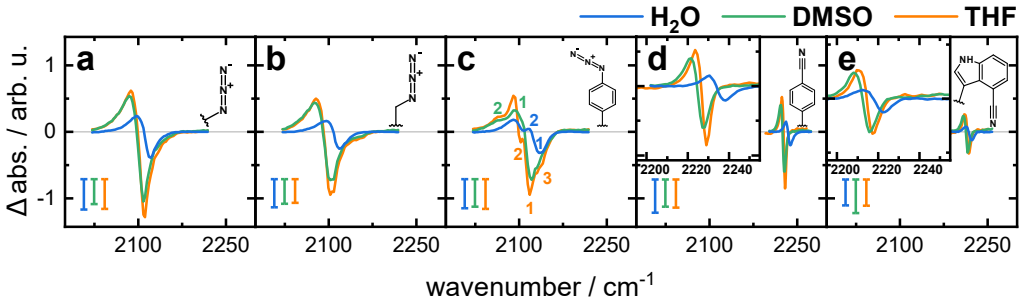

Figure 4.4: *Difference spectra of the VET experiments at the time of maximum signal intensity of the dipeptides containing N_3Ala, Aha, N_3Phe, CNPhe, and CNTrp (same order as in Fig. 4.1) in different solvents. The signal intensity is scaled to the peptide concentration of each experiment, so that signal amplitudes can be compared. Colored bars indicate 1 mOD difference signal in the VET experiments. The spectral range covered in all panels is identical, to better visualize the difference in band width of the different VET sensors, the insets show the narrow nitrile band shape in detail. Small numbers in the VET spectra of AzAla-N_3Phe indicate the presence of multiple subbands. Adapted from [136], licensed under CC BY.*

Despite the lower extinction coefficient of the nitrile probes, their VET signal is only somewhat smaller. The peak-to-peak amplitude of CNPhe is on average over the three solvents 36%, 18%, and 13% lower than that of N_3Ala, Aha, and N_3Phe, respectively. The signal of CNTrp is 50% lower than that of CNPhe.

In N_3Phe the multi-band structure of the absorption band influences the signal amplitude. The shift of the band leads to the overlap of induced absorption from one subband with the reduced absorption of the other. Besides broadening of the induced absorption, the local cancellation effects can lead to highly asymmetric reduced and induced absorption.

From the traces of the dipeptides (Fig. A.9a–e) the peak times can be determined (Fig. 4.5 top). In the different solvents, the peak times shift comparable to the amplitudes of the signals. The signals climax earliest in H_2O, between 2.4 ps and 3.6 ps. Relative to H_2O, the peak time shifts by 59–95% (76% on average) and 74–98% (85% on average) in DMSO and THF, respectively. The influence of the solvents on the peak time and the maximum signal amplitude, likely arises through the different couplings between the solvents and the peptides.

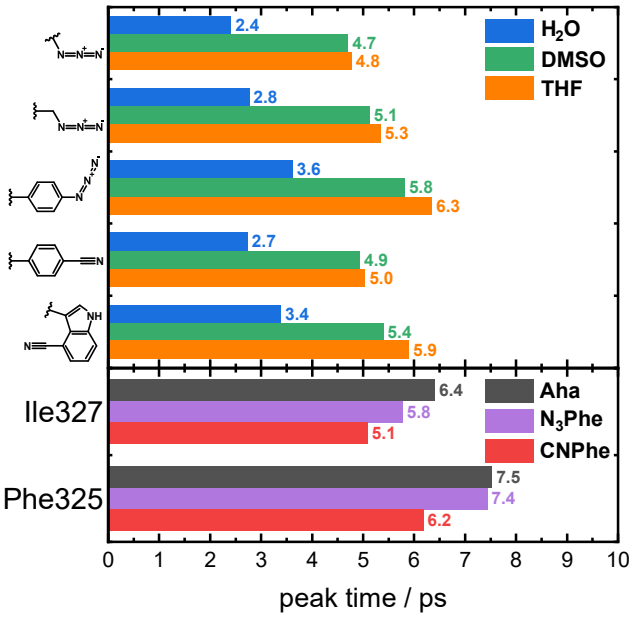

Figure 4.5: *Bar plot of the VET peak times of the dipeptides in H_2O, DMSO, and THF (top). The VET sensor of the respective group of three bars is indicated at the left. Bar plot of the VET peak times of the PDZ variants with Aha, N_3Phe, and CNPhe at the sites indicated at the left (bottom). Taken from [136], licensed under CC BY.*

The direct influence of the side chain on the peak time is best judged by comparing the peak times of different dipeptides in the same solvents. By increasing the size of the linker between the peptide backbone and the reporter group, we can observe an increase of the peak time. The additional methyl group in Aha increases the peak time by 12% compared to N_3Ala, the addition of the phenyl ring in N_3Phe increases the peak time by 35%. The exchange of the phenyl ring against the larger indole ring in CNTrp also increases the peak time by 16%. Generally one could expect the increased density of states to increase the amount of pathways along the LFMs that lead to a VET signal, but apparently this effect is out competed by the increase in distance that needs to be covered.

The experiments on N_3Phe and CNPhe allow for a direct comparison of the reporter groups and the coupling to the LFMs within the phenyl ring. The peak time decreases by 20% when the azide is exchanged for a nitrile. This contradicts the expectation that one could have from the vibrational lifetimes of these IR labels. The lifetime of 0.8 ps for

N_3Phe [201] is significantly lower than that for CNPhe of 4 ps [186]. From the long lifetime one could have expected a higher peak time, because the coupling between the reporter mode and the LFMs appears to be weak. This however, assumes that the same LFMs are responsible for the vibrational relaxation and the occurrence of the VET signal, which is not necessarily the case. Though the system will on the long run equilibrate to a population of LFMs determined by the Boltzmann distribution, the intermediate steps lead to non-Boltzmann populations. Therefore a very different population distribution of LFMs gives rise to the VET dynamics of a mode than to the characteristic decay dynamics of that mode.

Though the extinction coefficients of the nitrile sensors is considerably lower than of the azido sensors, the signal amplitudes are only slightly lower. The amplitude of the difference signal depends linearly on the band width of the sensor band and its shift. In order to compare the different probes, the relative shift $\Delta\tilde{\nu}_{rel}$, the quotient of the shift and FWHM, is introduced. The signal amplitude is proportional to $\Delta\tilde{\nu}_{rel}$ for small $\Delta\tilde{\nu}_{rel}$, 10% deviation from linearity is reached at $\Delta\tilde{\nu}_{rel} \approx 0.5$ for Gaussians and at $\Delta\tilde{\nu}_{rel} \approx 0.4$ for Lorentzians. For larger shifts it approaches the total amplitude of the peak, the total separation of both contributions (Fig. 4.6).

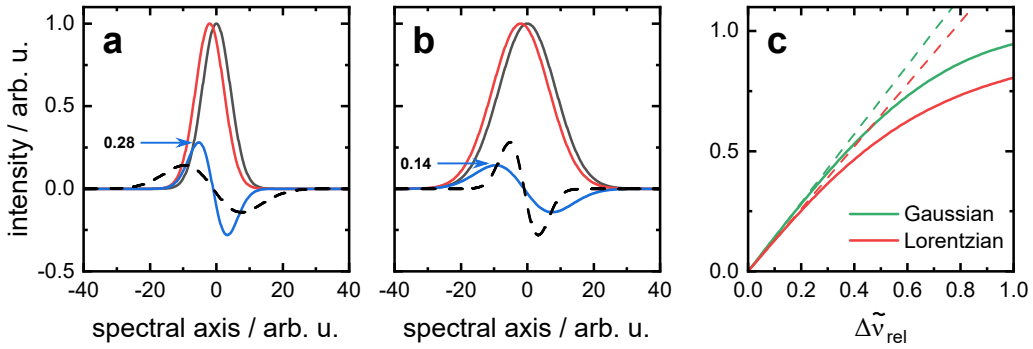

Figure 4.6: *The difference signal (blue) of two slightly shifted Gaussian functions (gray and red) in inversely proportional to their FWHM. Examplified for a shift of two units for Gaussians with FWHM of 10 and 20 units in panels a and b, respectively. The signal of the other panel is shown for comparability (black dashed line). (c) The signal amplitude of Gaussians and Lorentzians (solid lines) deviates from proportionality (dashed lines) at large $\Delta\tilde{\nu}_{rel}$. Adapted from [136], licensed under CC BY.*

To determine the shifts that occur during the VET experiments, the amplitude of the shifted peak must be fixed because of the dependency between the signal amplitude and the shift. The same signal amplitude can be generated either by increasing the shift or the amplitude of the base absorption. To this end the amplitude was determined relative to the FTIR signal of the samples using the formulas given in [149] (see Sec. 2.4.4 for further details).

Note that this analysis ignores stimulated emission, the probe beam polarization, and attenuation. Because of the lack of stimulated emission, the delta pump pulse (whose compression in time further increases the calculated excited fraction) could even lead to population inversion. With the conditions of the dipeptide experiments there is no polarization orientation that is excited over the theoretical limit of 50%. The relative polarization between pump and probe beam is irrelevant for the magic angle geometry in the experiments. Furthermore, the intensity dependent polarization dependence of the absorption can lead to a narrowing of the intensity profile of the probe beam. This effect, dominated by the solvent absorption, would change the probe beam weight distribution along the sample depth. Finally, the beams are assumed to be perfectly parallel and perfectly aligned to each

other, though the pump beam hits the sample at a low angle. Therefore the calculated signal size (relative to the FTIR amplitude) can only be an upper limit and the corresponding shifts a lower limit. Because the deviations of the idealized calculations are the same for all dipeptides the comparison of the resulting values is justified.

Based on these calculations the amplitude was fixed to 12% of the FTIR signal, the upper limit of the excitation. This signal with fixed amplitude was then shifted against itself to reproduce the VET signal at each delay. This method results in systematic deviations from the experimental data, too weak reduced absorption and too strong induced absorption. This is overcome by fitting a Voigt peak to the FTIR spectrum, keeping the parameters of the reduced absorption constant (in accordance to the FTIR spectrum), and allowing the Voigt function of the induced absorption to change its band shape but keeping its area equal to that of the reduced absorption. The additional parameters of the Voigt function can account for band shape changes that originate from the anharmonic couplings [24]. The fit from the Voigt functions does better recreate the VET signal shape, but both methods nicely reproduce the dynamics of the VET trace and result in nearly identical shifts A.10. The resulting shifts at the peak of the VET signal are indicated in Tab. 4.1.

$\Delta\widetilde{\nu}_{rel}$ is small for all labels in H_2O, 2.5% for Aha and ≈3.5% for the other sensors (Tab. 4.1). In DMSO $\Delta\widetilde{\nu}_{rel}$ of N_3Ala, Aha, and CNTrp is very similar (8–9%), but for CNPhe it is 13%. In THF the relative shift of CNPhe similarly exceeds that of the other labels. The relative shift therefore explains the slightly lower signal of CNTrp in comparison to CNPhe, it furthermore explains how the signal amplitude of CNPhe can become similar to the amplitudes of the azide sensors even with its lower extinction coefficient.

Table 4.1: *VET-induced blueshifts (in cm^{-1}) of the dipeptides, as determined by the shifted Voigt peak (first value) and the shifted FTIR (second value). The values in brackets are the corresponding values of $\Delta\widetilde{\nu}_{rel}$ (in %). The signal of N_3Phe was not fit, because of the more complex band shape.*

	H_2O	**DMSO**	**THF**
N_3Ala	1.03 / 0.99	2.34 / 2.23	1.76 / 1.68
	(3.6 / 3.5)	(9.8 / 9.3)	(7.2 / 6.8)
Aha	0.78 / 0.77	2.17 / 2.15	1.52 / 1.46
	(2.5 / 2.4)	(7.9 / 7.8)	(5.5 / 5.3)
CNPhe	0.36 / 0.35	0.94 / 0.93	0.70 / 0.70
	(3.5 / 3.5)	(12.9 / 12.8)	(12.5 / 12.5)
CNTrp	0.43 / 0.49	0.73 / 0.73	0.67 / 0.66
	(3.5 / 4.0)	(8.5 / 8.5)	(8.4 / 8.3)

4.2.2 Proteins

The VET signals measured in PDZ are considerably smaller because of the increased distance between VET donor and VET sensor A.11, their shapes (Fig. 4.7) match the observations from the FTIR spectra (Fig. A.12). F325Aha, I327N_3Phe, and F325N_3Phe clearly show the presence of two subbands and even the very weak double structure in I327Aha can be recognized.

The amplitude of the azide sensors varies strongly because of the variable band shape. The narrow and intense features in I327N_3Phe and F325Aha lead to a signal enhancement of these variants. The extraordinary band shapes of these variants lead to an enhancement

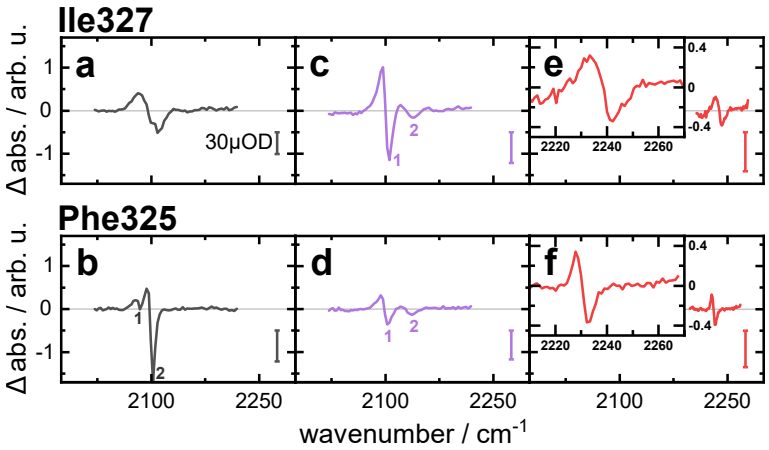

Figure 4.7: *Difference spectra of the VET experiments at the time of maximum signal intensity of PDZ with Aha (gray), N_3Phe (purple), and CNPhe (red), in the same order as in Fig. 4.2, inserted at positions Ile327 (top) and Phe325 (bottom). The VET signal intensity is scaled to the protein concentration and to the difference signal of the water at 2212 cm^{-1} for each experiment leading to comparable signal amplitudes. Colored bars indicate 30 μOD difference signal in the VET experiments. Small numbers in the spectra of Aha and N_3Phe containing mutants indicate the presence of multiple subbands. Adapted from [136], licensed under CC BY.*

of their VET signals, both with a relative peak-to-peak amplitude of 21.6. The signals of I327Aha and F325N_3Phe appear to be more representative with relative peak-to-peak amplitudes of 9.2 and 6.7, respectively. Similar to the dipeptides, the signal amplitudes of the CNPhe mutants are in a similar range to the azide signals, with relative peak-to-peak amplitudes of 6.6 and 7.1 at Ile327 and Phe325, respectively.

In I327N_3Phe the side feature is blueshifted, nevertheless, its induced absorption has only minor influence on the reduced absorption of the main feature, preserving its symmetry. In F325Aha the side feature is redshifted. In this mutant the two-peak structure leads to a highly asymmetric VET signal with two peaks visible in the induced absorption and only a single, very intense peak of the reduced absorption (3.6 times stronger than the induced absorption).

From the peak times of the dipeptides one would expect the peak of Aha and CNPhe to occur on the same time and that of N_3Phe 1 ps later. But instead we observe an increase of the peak time from CNPhe to N_3Phe and another increase to Aha (Fig. 4.5 bottom, Fig. A.9f and g) at both protein sites. This difference in the order, as compared to that of the dipeptides, emphasizes the importance of a thorough characterization of the potential sensors. In the dipeptides VET has to occur quasi-linearly, this is not the case within the complex structure of a protein. In the protein a combination of numerous VET pathways, that contain covalent bonds as well as non-covalent contacts, connecting the VET donor and VET sensor are conceivable. These contribute in parallel to a greater or lesser extent to the total observed VET [13, 63, 65]. The multitude of pathways reduces the impact of each single step, so that the experiments in the protein offer a comparison more specific to the VET sensors than the dipeptides.

The peak time between Ile327 and Phe325 mutants increases by 1.1 ps, 1.6 ps, and 1.1 ps for Aha, N_3Phe, and CNPhe, respectively. This systematic increase reflects the longer distance between the VET donor AzAla in the peptide ligand and the VET sensors within the protein, a correlation that has been observed in related experiments and simulations [63,

202, 203]. At both sites, the peak times of all sensors lie within a window of 1.3 ps. If one uses different sensors within one system, this uncertainty needs to be considered. The influence of structural differences (e.g., orientation of the azide group on N_3Phe) which may open up additional pathways for the VET can not be examined from this small set of mutants. Effects like these may be the reason for the 0.5 ps larger increase between the two N_3Phe mutants as compared to the other ncAAs. Knowledge of additional structures, i.e., structures of the different mutants ideally in complex with the ligand, further experiments with VET sensors at other positions within the protein, and the combination with simulations are required for a complete dissection of the different contributions.

An additional VET experiment was conducted on PDZ Ile327SCN. The experiment and the data processing was done identical to the other mutants, but a VET signature could not be reliably identified in the data (Fig. 4.3b). It is safe to assume the same amount of energy is transferred from AzAla to the sensors. If this energy induced the same shift in SCN as in Ile327CNPhe, the expected peak-to-peak amplitude in this experiment was 18 µOD. A signal of that size should have been visible under the given experimental conditions (the mean of the absolute values is 2 µOD). If the observed band shape between 4 ps and 8 ps includes a genuine VET signal, the observed shift of SCN is ~3.5 times lower than that of CNPhe at the same position (Fig. 4.3c). It was observed that the vibrational lifetime of various vibrational labels could be extended, by inserting a heavy atom (e.g., S, Se, Si, Sn) as linker [204–207]. This is attributed to an insulation effect of the heavy atom, that uncouples the probe mode from the vibrational modes of the scaffold. The sulfur that increases the lifetime of the SCN label insulates it from VET along the backbone and prevents a strong VET signal. The SCN label is therefore regarded as bad VET sensor. It can not generate a reliable VET signal under circumstances at which the other three sensors generated signals in the order of tens of µOD.

5 Conclusion

Experimental investigation of VET in proteins is possible using TRIR spectroscopy with ncAAs, that allow spatially resolved experiments. The originally proposed VET sensor, Aha, could only be incorporated into Met-free proteins and may not be a suitable sensor to replace some other amino acids (e.g., aromatic amino acids). To facilitate the experimental characterization of VET in every protein and at every site, it was therefore essential to extend the collection of well characterized and readily available VET sensors.

The ncAAs that have been presented now offer a wider variety of incorporation methods (SPPS: all; SPI: Aha; SCS: N_3Phe, CNPhe). This vastly increases the number of target proteins. Additionally they include larger ncAAs with aromatic side chains, which can be more appropriate exchange partners for certain residues, reducing the impact of the VET sensor incorporation into a protein scaffold.

The presented data demonstrate that not only azides can be used as VET sensors, but also nitriles with their lower extinction coefficient lead to reasonably large signals. This is because the lower extinction coefficient is accompanied by a narrow band whose relative shift is larger than that of the azide. Signal shapes and amplitudes are also more predictable for the nitrile probes, because of Fermi resonances in the organic azides. Under certain conditions these give rise to intense, narrow bands that can favor the VET signal amplitude. At the same time, the occurrence of multiple subbands results in asymmetric VET signals, because of cancellation effects between the different contributions.

These results indicate that a variety of VET sensors can be used to investigate VET in the diverse protein environments. When combining different labels in one investigation, it is important to consider the specific peak time differences that lie within 1.3 ps, approximately the time it takes the energy to travel from residue i to residue $i + 2$ in a protein, for the investigated sensors.

The comprehensive characterization of five labels in dipeptides in different solvents and three of these labels in a protein-ligand system allows to assess the influence of various effects on VET experiments. They can be the basis for the optimal design of future VET experiments in more diverse protein systems. With these more versatile VET sensors it will be feasible to follow VET through the complete structure of minimally perturbed proteins.

Part III

Vibrational Energy Transfer and Low-Frequency Modes in Enzymes

6 Introduction

In Sec. 1.3 it was outlined how LFMs facilitate VET through large protein structures. Though thermally accessible, these modes tend to be overdamped in solvated samples due to their frequencies [208], and only recent studies could identify certain underdamped LFMs in different proteins [209–211]. Nevertheless, their potential role for enzymatic reactions has been discussed for several years [212], but direct experimental evidence to prove such a hypothesis is still lacking.

The dual role of LFMs, as vibrational energy conductors and engine for enzymatic reactions [22], opens up a great chance to elucidate this proposal. The experiments described in this part of this thesis have been designed for this purpose. Two underlying models and the systems that were investigated are introduced within this chapter.

6.1 Interplay of LFMs and Enzymatic Reactions

The function of an enzyme is to catalyze a chemical reaction, by accelerating the crossing of an energy barrier that separates reactants from products of the reaction. But the exact mechanism by how this occurs remains a matter of active research [213]. It has been proposed that the enzyme does not only generate a setting in which the reactants can readily cross to the product side but they may actively be involved in the reaction by dynamical modulation of this reaction environment. Two models that describe this dynamic enzymatic catalysis are presented in the following sections, along the lines of a recent review by C. M. Cheatum [214].

6.1.1 Activated Tunneling Model

Enzymatic hydrogen transfer reactions strongly depend on the tunneling rate of the hydrogen nucleus to the H-acceptor. But even semiclassical models with tunneling corrections were not sufficient to explain some experimental phenomena, leading to the development of new tunneling models [21]. In the framework of the activated tunneling model (ATM; various names are used throughout the literature), the overall reaction rate can be written as [214, 215]

$$k(T) = C(T)\frac{|V|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}e^{-\frac{(\Delta G^\circ+\lambda)^2}{4\lambda k_B T}}\int_0^\infty F\left(m,\mathrm{DAD}\right)e^{-\frac{E(\mathrm{DAD})}{k_B T}}d\mathrm{DAD} \tag{6.1}$$

where $C(T)$ accounts for the fraction of reactive enzymatic complexes, V is the coupling between the reactant and product state, λ is the reorganization energy in the heavy atom coordinate, ΔG° is the driving force of the reaction, and $F\left(m,\mathrm{DAD}\right)$ is the tunneling probability for a particle of mass m at the donor-acceptor distance DAD. The first term (before the integral) describes the rate of reaching the tunneling ready state (TRS) within the heavy atom coordinate, including the motions of all atoms except the hydrogen. This first term dominates the temperature dependency of $k(T)$ but is identical for the transfer of different isotopes. At the TRS the zero-point energies of the hydrogen bound to the reactant or the

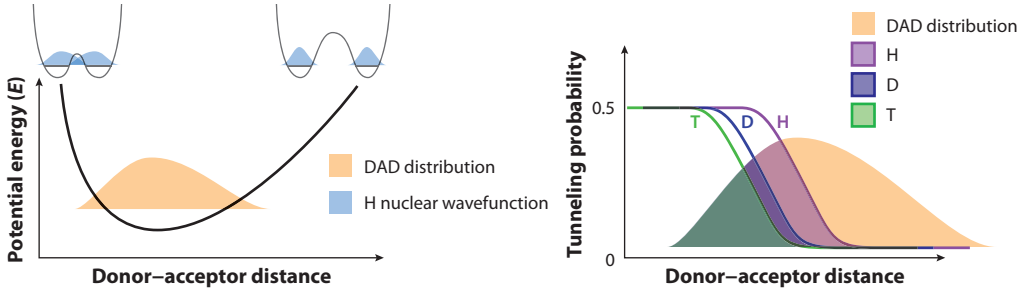

Figure 6.1: *Left: At the TRS the thermal distribution of the DADs (brown surface) is determined by the DAD potential. The tunneling probability for low DADs is higher because of a stronger overlap of the nuclear wavefunctions of the tunneling hydrogen. Right: At the TRS the tunneling probability (colored lines) depends on the DAD and on the mass of the tunneling isotope. The integral of the tunneling probability and the DAD distribution over all DADs (colored areas) gives the thermally averaged tunneling probability. This probability is different for each hydrogen isotope, which results in a measurable KIE. Adapted with permission of Annual Reviews, Inc., from [214]; permission conveyed through Copyright Clearance Center, Inc.*

product potential are degenerate (small double-well potentials in Fig. 6.1) and tunneling between both wells is possible. The kinetics at the hypersurface of the TRS are represented by the integral. At the TRS the DAD distribution is determined by the Boltzmann factor (final term) according to the potential energy along the DAD coordinate (Fig. 6.1 left), this sampling of the DAD distances is sometimes referred to as gating [216]. The integration yields the averaged tunneling probability (colored areas in Fig. 6.1 right) for a certain hydrogen isotope and temperature.

According to the ATM the kinetic isotope effect (KIE) can be used to probe DAD distributions. It monitors the change of a reaction rate for different isotopes. Because the first part of Eq. 6.1 is independent of the tunneling isotope, the KIE purely represents the integral over the DAD distribution. A narrow DAD distribution centered at low DADs will have no or a very low KIE (identical integrals for the different isotopes, the deep tunneling regime) that is essentially temperature-dependent. If the DAD distribution is broader and spans the DAD range in which the tunneling probabilities of the isotopes vary (this situation is shown in Fig. 6.1), the KIE will be larger and temperature dependent because of the temperature sensitivity of the DAD distribution. A large and temperature-dependent KIE can therefore be indicative of an enzymatic process, that is highly sensitive to the current DAD distribution at the TRS. The thermal sampling of this DAD distribution is facilitated by the thermally accessible motions which link it to the enzyme LFMs.

Reasonably it is assumed that enzymes have evolved to optimize the catalyzed reaction. This includes the development of presumably narrow DAD distributions centered at small DADs. For that reason, the KIE of many WT proteins is temperature independent under native conditions and changes to temperature-dependent behavior when exposed to different types of perturbations [21]. Often mutations have been introduced into proteins to elucidate the role of dynamics in the enzymatic process. These mutations give (non additive) rise to temperature-dependent KIEs [217–219], presumably shifting the DAD distribution to longer DADs. Though these studies cannot directly show which LFMs are involved in the catalysis in the native protein, they may help assessing whether LFMs can principally be important for catalysis. If so, this could help to understand the development from their ancestors to the highly efficient enzymes that we encounter today and shape the future of how we design and affect enzymes [220–222].

6.1.2 Rate-Promoting Vibrations

Another approach to include protein dynamics in enzymatic processes are rate-promoting vibrations (RPV), a concept developed in the lab of S. D. Schwartz [20, 223].

This model implies a direct coupling between a certain, well defined protein mode and the reaction coordinate. These modes could be identified by *in silico* transition-path sampling methods [224]. Such methods allow the identification of rarely-occurring events (such as the barrier crossing in an enzymatic reaction) and focus the computational effort on trajectories including these events. In this way they can be used to thoroughly characterize the trajectories that lead to the transition and from them the dynamic nature of the transition itself, a futile endeavor if totally randomized trajectories were sampled. An important difference to the ATM is that the high-dimensional reaction coordinate determined in this fashion cannot be reflected by the Marcus-like ATM [225].

These trajectories can be used to define a hypersurface, from which the probabilities of reaching the reactant state or the product state are equal (50%), the separatrix, that represents all possible transition states [226]. From these trajectories all motions that are required for the reaction to occur can be identified and by definition the reaction coordinate is orthogonal to this surface. In lactate dehydrogenase it was additionally shown, that the disruption of a certain protein motion, which was present in all relevant trajectories, would stop the transition towards the product state even in the previously successful trajectories [227]. This result emphasizes that these protein motions not only accompany the reaction but are an essential part of the enzymatic catalysis [213]. In contrast to the classical picture of enzyme catalysis, in which the reaction barrier is reduced (the transition state stabilized or the ground state destabilized), the RPVs lower and simultaneously narrow the reaction barrier which exponentially accelerates tunneling and classical proton transfer [223, 228].

6.2 Experimental Concept

Using VET, we can investigate the energetic connectivity between distant protein sites. The idea of identifying allosteric sites that are connected to the active part of the protein by preferred VET channels can be applied to all proteins. Though the connections may be identified with VET, it is not necessarily the mechanism by which information is transferred during protein operation. In enzymes on the other hand, the excitation of certain modes may actively affect the catalysis.

For human lactate dehydrogenase (LDH) the connection between RPVs and the catalytic reaction has been shown exceptionally nicely *in silico*. This enzyme features a subpicosecond compressive RPM that is essential for the reaction, i.e., part of the reaction coordinate [229]. Then it was shown that the residues that participate in the RPV constitute a preferred channel for vibrational energy [22], which should be verifiable in VET experiments. Based on the RPV, a non-competitive inhibitor that uncouples the reaction from the protein motion was identified [230]. For homologous porcine LDH the actual inhibition was subsequently proved for two inhibitors very similar to that one originally proposed *in vitro* [231].

Previous experiments have shown that AzAla can be used to inject vibrational energy into the protein site-specifically. It should be possible to use this property of AzAla to specifically excite an RPV and thereby increase the catalytical function of an enzyme. By varying the position of AzAla, the amount of energy reaching the RPV is expected to change, the positioning could then be used to tune the strength of the reaction enhancement. The en-

ergy reaching a VET sensor in the acitve site is expected to vary accordingly. VET experiments could then be used to identify RPVs and simultaneously allow for a new method to remotely, specifically, and reversibly steer the activity of enzymes by illumination.

6.3 Enzyme Dynamics in Cyclophilin A

Cyclophilin A (CypA) has been discovered in 1984 as a cyclosporin A-binding protein [232] and simultaneously a peptidyl-prolyl cis-trans isomerase (PPIase) was found in porcine kidney tissue [163]. These two were identified as the same protein five years later, simultaneously by two independent groups [233, 234]. Since this discovery many studies have aimed at uncovering the details of the function of cyclophilins and PPIases, many of which have repeatedly been collected in numerous review papers focusing on various aspects [235–238].

The isomerization of proline peptide bonds is an important step in protein folding. In equilibrium, the trans-isomer of proline is favored, though the uncatalyzed conversion from the cis-isomer (7.1 a) is very slow [239]. CypA and other PPIases support the folding and refolding of proteins [240], most notably, human cyclophilins are crucial for the maturation and infectivity of human immunodeficiency virus (HIV) and are linked to numerous other human diseases [237]. Additional roles of PPIases have been assigned to gene transcription and cell signaling [238, 241].

The enzymatic catalysis increases the rate constant of the otherwise slow cis-trans isomerization by six orders of magnitude [242]. Most recent studies assume that conversion of the substrate occurs via a rotation of the N-terminal peptide chain while the C-terminus remains stationary, though the direction of the rotation is a matter of debate [243–245].

Various studies investigated the involvement of protein dynamics and the prospect of dynamic allostery in the CypA catalysis, characterizing important protein dynamics on timescales ranging from nanoseconds to milliseconds [246–251]. These theoretical, evolutionary, crystallographic, and NMR relaxation studies commonly identify sets of conserved, strongly-coupled amino acids that connect distal parts of the protein to the active site. These networks of coupled amino acids confer equilibrium structural transitions and changes in dynamics, induced by distant mutations.

More specifically, it was shown that two catalytically important active-site residues (Arg55 and Phe113) slightly rotate between a major active and a minor inactive conformation. The mutation Ser99Thr slows down this interconversion, a possible explanation for the reduced activity of this mutant [247, 250]. More distant mutations, that were part of motional clusters but not interacting with the substrate, could partially increase the active-site dynamics and recover the activity of the CypA mutant without altering the WT dynamics [252]. NMR relaxation studies identified still more distant mutations that influence the active-site dynamics [249]. By correlating the effects on dynamics of numerous mutations two types of response were identified which are clustered within two distinct networks of dynamic allostery. Mutations of residues within these networks led to different binding, release, and isomerization rates which could be tuned by combining mutations. Additionally, networks of conserved H-bonds have been identified by comparing the amino acid sequences of cyclophilins from different organisms, these H-bond networks also contribute to vibrational modes that are associated with the substrate conversion [246].

All of the identified networks of coupled amino acids are potential preferred channels for vibrational energy. By inserting the VET donor AzAla at different sites in CypA, VET through different regions can be investigated. For these experiments a VET sensor is required. The peptide binding of CypA can be exploited to insert a VET sensor without the need to incorporate a second ncAA into the protein, but instead in the substrate ligand. In principle the tight-binding inhibitor cyclosporin A ($K_d < 100$ nM [253]) would be a good choice as ligand substrate, but because of its untypical structure[1] a customized version of this peptide including Aha is not readily available. Instead linear peptides derived from the CypA binding site of the HIV capsid protein (HAGPIA) and identified from a randomized peptide library (FGPDLPAGD) were intended to be used as VET sensor carrier ($K_D \leq 50$ µM) [255–257].

The activity of the same mutant proteins can be monitored to investigate the role of protein dynamics in catalysis. By exciting the AzAla, dynamics can locally be enhanced. The networks of coupled motions can propagate the local effects over larger distances and into certain protein modes. If dynamic effects or certain modes are coupled to the catalysis, it is expected that different enzyme activity can be observed with or without excitation of AzAla.

6.4 Low-Frequency Oscillations in Formate Dehydrogenase

The enzyme FDH catalyzes the reduction of NAD^+ to NADH by direct hydride tunneling from formate [258]. The regeneration of NAD(P)H cofactors in biotechnological applications is one reason for the vast efforts undertaken to understand and optimize this enzyme and its catalytic function [259].

Candida boidinii (*Cb*) FDH was the first enzyme in which 2D-IR spectroscopy was used to observe underdamped LFMs in the active site, a highly challenging endeavor [210, 219, 260]. In these experiments, the enzyme formed a ternary complex with NAD^+ and the azide anion (N_3^-), a tight-binding inhibitor (K_D=43 nM) and transition state analog [167]. The N_3^- has a strong transition dipole that makes it a convenient 2D-IR label, though it has a vibrational lifetime of only 2 ps [261]. In these experiments, the center line slope (CLS) of the time-dependent signals was analyzed and revealed additional exponentially damped oscillatory features with frequencies corresponding to 10 cm^{-1} and 24 cm^{-1} [210]. The CLS analysis can identify the time scales at which the frequency-frequency correlation of an oscillator is lost, i.e., how fast the molecules sample the environmental substates. The oscillatory dynamics indicate that the protein active site samples the same configuration repeatedly with distinct frequencies (as opposed to random motions through configuration space). These oscillatory dynamics cease in the ternary complex with NADH instead of NAD^+. The occurrence and especially the modulation of these ultrafast motions in the active site raises the question whether these dynamics have catalytical relevance.

To tackle this question the KIE in FDH was thoroughly investigated [167, 219, 262]. In WT FDH the KIE is temperature independent [167], as expected for a well-evolved protein. If however, one introduces slight perturbations, the KIE becomes temperature dependent. This could be achieved either by introducing heavy isotopes into the protein [262] or by mutating an active-site residue (Val123Ala or Val123Gly) [219]. The introduction of ^{15}N and ^{13}C reduces the amplitude of the 10 cm^{-1} oscillatory component modulating the CLS decay, ad-

1 Cyclosporin A is a cyclic undecapeptide that includes D-Ala, non-proteinogenic amino acids, and methylated amino acids [254].

ditional incorporation of ^{2}H increases the amplitude of both oscillatory components (even above the WT) [260]. The temperature dependence of the KIE increases with an increase of the protein mass, but its amplitude and the activation energy differences $\Delta E_{a(T-H)}$ (between tunneling ^{1}H and ^{3}H) do not. These non-linear effects (in the CLS dynamics and the KIE) indicate that the isotopic labeling does not purely increase the mass of the protein, but also influences its electrostatics. The implications of heavy-enzyme KIEs in the context of enzyme dynamics therefore need to be reconsidered. For the Val123, mutants clear correlations between disturbance (size of the active-site residue: Val123>Val123Ala>Val123Gly), the KIE, and the enzyme dynamics are observed [219]. From Val123 to Val123G $\Delta E_{a(T-H)}$ increases linearly, the same is true for the amplitude of the 10 cm^{-1} oscillatory component on the CLS dynamics.[2]

Two groups performed theoretical studies in the context of the experiments, in which they were not involved. One study focused on the KIE of FDH WT (based on the structure of *Pseudomonas* sp. FDH) and on the interpretation of the KIE in the context of enzymatic reactions [263]. The other study was building on the effects observed from Val123 mutation [225]. Both studies identified protein motions that occur at ultrashort times (<1 ps) before the reaction event. These motions that are part of the reaction coordinate suggest that the protein takes an indispensable, active role in the reaction.

After exploring the suitability of different ncAAs as VET sensors in Part II, FDH allows one to test if the non-covalent, small molecule ligand N_3^- can also act as VET sensor. In previous studies, VET along non-covalent bonds between different amino acids was established. Compared to amino acids (and peptides or proteins) N_3^- is considerably smaller, only a minuscule number of modes may couple to the reporter mode. This may negate the positive effect of a large extinction coefficient. If N_3^- can detect VET, FDH mutants with differently positioned AzAla can be used to investigate the energetic connectivity of the active site and distal protein sites.

The same mutants can be used to investigate the proposed role of enzyme dynamics in catalysis. As described for CypA, the enzymatic activity is expected to be modulated by AzAla excitation and the subsequent processes. The effect should vary depending on the position of AzAla because of the distance to the reactive center and the importance of motions in certain regions for the reaction coordinate.

2 This correlation gets lost for wavenumbers >20 cm^{-1}, the authors attribute this to the faster dephasing of the smaller amino acids, which is also evidenced by band broadening.

7 Cyclophilin A

7.1 AzAla Incorporation Sites

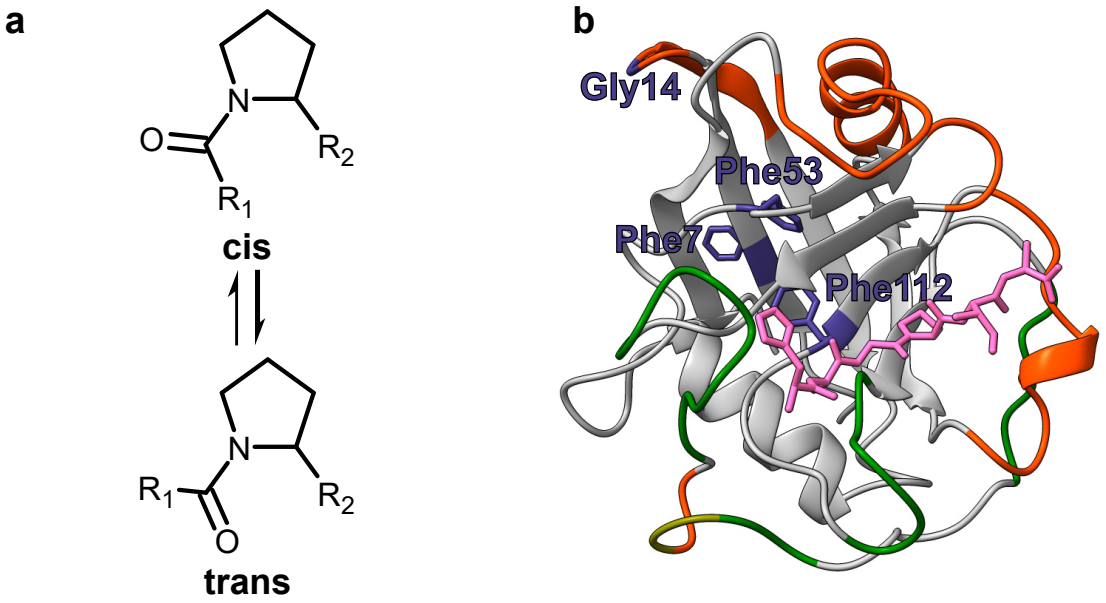

Figure 7.1: *Cis- and trans-isomer of the peptide bond preceding a Pro resiude (a). CypA catalyzes the isomerization this bond in peptides. It accelerates the speed at which the equilibrium of ~5–10% cis-isomer is reached. Crystal structure of CypA with a peptide ligand (HAGPIA) in pink (b). The positions at which AzAla was meant to be incorporated are labeled and highlighted in dark blue. Colored residues take part in the reaction-coupled modes A (orange) or B (green), residues 80 and 81 (olive) participate in both modes. Structure based on RCSB protein data bank (pdb): 1AWQ [255].*

The only Trp residue in CypA (Trp121) is situated within the active site and is involved in ligand binding [264]. The indole-N forms an H-bond with the carbonyl-O of the amino acid following the substrate Pro and the ring system is involved in hydrophobic interaction with the substrate. The software tool YASARA together with the FoldX plugin [265, 266] was used to asses the destabilizing effect of the AzAla incorporation by *in silico* mutation of various residues to Trp. The mutation sites Phe7, Gly14, Phe53, and Phe112 were selected. Two of these residues (Phe50 and Phe112) are close to the active site and situated between the residues that are involved in the reaction-coupled modes A and B (7.1 b) [246]. Gly14 participates in mode A and Phe7 was intended as control site, as it is far from the active site and no part of modes A–C.

7.2 Results and Discussion

CypA WT was successfully expressed and purified. CypA Gly14AzAla was expressed and purified in the laboratories of N. Budisa (at the Technische Universität Berlin) by the author guided by T. Baumann. Even after gathering experience with the expression and incorporation protocols for SCS with AzAla, the expression of CypA AzAla variants could not be established in Frankfurt. Expression of CypA Gly14AzAla with the same protocols, plasmids and *E. coli* strains that were utilized in Berlin did not succeed. It is assumed that the reason for the failure was the AzAla that was used. In both locations racemic AzAla-HCl salt was used, but the ncAAs were of different batches prepared by different synthesists. While the AzAla-HCl that was used in Berlin did readily dissolve to a high concentration (>200 mM) in H_2O, this was not the case for the ncAA that was present in Frankfurt at that time. Either in this batch of AzAla-HCl the true content of L-AzAla was lower than expected, or the reduced solubility led to lower final concentrations within the expression cultures, hampering the incorporation of the ncAA.

To observe the activity of the successfully prepared enzymes a chymotrypsin-coupled assay was performed. This assay is based on the substrate specificity of chymotrypsin towards succinyl-AAPF-pNA, only in the trans-configuration of the Pro residue does chymotrypsin cleave off pNA at the C-terminus, shifting the pNA absorption from ~310 nm to ~390 nm [267]. Because this is the preferred form, a rapid increase of the product absorption is observed upon addition of the substrate to a pre-equilibrated reaction mixture of chymotrypsin or CypA with chymotrypsin. The high concentration of chymotrypsin in the reaction mixture allows quasi-instantaneous cleavage of the trans-isomer. Because the digestion constantly depletes the concentration of the trans-isomer, the equilibrium is disturbed and the cis-to-trans isomerization continues. Therefore, only the kinetics after the initial burst phase reflect the time constant of the thermal or catalyzed isomerization, without or with CypA, respectively.

To reduce the background of thermal isomerization, the assay was conducted at 6°C. At this temperature, the thermal isomerization rate k_{therm} is $5x10^{-3}$ s^{-1}. A clear acceleration of this rate can be seen when 4 nM CypA WT are present in the reaction mixture (Fig. 7.2 a). In this case, the observed isomerization rate is k_{WT}=$17x10^{-3}$ s^{-1}, which is the sum of k_{therm} and the reaction constant for the catalyzed isomerization pathway. This result is a clear proof for the successful recombinant expression and purification of active CypA. However, no catalytic activity was observed for CypA G14AzAla (Fig. 7.2 b).

To screen for suitable ligands, that could be used to introduce Aha into the protein system, isothermal titration calorimetry (ITC) experiments were conducted on CypA WT. From these experiments, dissociation constants (K_D) of 80 µM, 350 µM, and 8.3 mM were determined for the peptides HAGPIA, Ac-FGPDLP-NH_2, and Ac-PVHAGPIAP-NH_2, respectively (A.13).

In a previous study, K_D of the complete HIV capsid protein (with the main recognition sequence HAGPIA) and CypA was determined as 16 µM [256]. In comparison to this value, the K_D of the HAGPIA peptide determined herein is 5 times higher. For the peptide Ac-PVHAGPIAP-NH_2, which contains additional amino acids of the capsid sequence and terminal modifications,[1] K_D is even larger by two orders of magnitude. The K_D of the

1 The N-terminal acetylation and C-terminal amidation remove charges from the termini, increasing the similarity of the peptide to the full-length protein, in which the amino acid sequence continues on both sides.

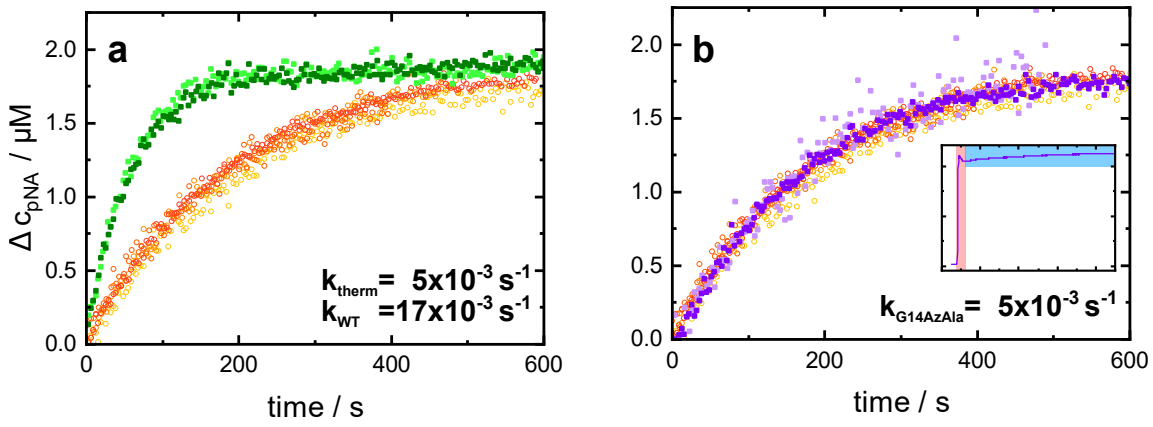

Figure 7.2: *Chymotrypsin-coupled activity assay for CypA at 6°C. The accumulation of free p-NA after the cleavage from trans succinyl-AAPF-pNA as determined from the rise of absorbance at 390 nm is shown. The thermal isomerization of the peptide is shown for four independent measurements (open yellow/red circles). The isomerization is accelerated in the presence of 4 nM CypA WT (duplicate, green squares in a). The isomerization rate in the presence of 6 nM CypA G14AzAla was unchanged from the pure thermal isomerization (duplicate, purple squares in b). The inset (b) exemplifies how most of the original substrate is instantly cleaved in the burst phase upon addition (red highlight), only ~10% remain for the observation of the isomerization kinetics (blue highlight).*

Ac-FGPDLP-NH_2 peptide is 7-fold increased compared to that of a similar peptide (Ac-FGPDLPAGD-NH_2) [257]. The K_D of Ac-VHAGPIAP-NH_2 appears implausible in comparison to that of HAGPIA. For HAGPIA and Ac-FGPDLP-NH_2 the differences to the literature are reasonably low when considering the experimental conditions.

The low protein concentration compared to the K_D results in a low value $c = c_{CypA}/K_D$ of 0.03-3.5, which is ideally between 10 and 100 for reliable results. In this range of c, the binding curve probed by the ITC experiment can optimally be fit to determine the thermodynamic parameters. In this range the complete sigmoidal shape can be resolved and the increase is not too steep, but spreads over multiple injections. The low concentrations were chosen to be sufficient for the expected K_D-values, whereas appropriate concentrations for the determined values would require very large amounts of protein. The determination of K_D values from ITC is highly sensitive to the concentrations of both components, as well as the exact composition of the solutions being titrated. The concentration of the protein solution can easily be determined spectroscopically, while the concentration of the peptide is purely based on the weighed portion, because they do not contain amino acids with a distinct absorption. For this value errors can occur from the scale itself, as the amount of peptide used per experiment is $\leq$2 mg. Moreover, the SPPS peptides contain trifluoroacetic acid in an undefined amount. This can lead to deviations between the real and the calculated peptide concentration and in the composition of the peptide solution.

After these preliminary experiments, this project was discontinued for three reasons: (i) The difficulties of incorporating AzAla into the protein sequence. (ii) The activity assay that can only cover low substrate concentrations, in which observing small illumination-induced changes in the activity becomes exponentially more challenging. (iii) The low affinity of the protein to the peptide ligands, even before the introduction of the ncAAs that are required for the intended VET experiments. The subsequent search for an alterna-

tive enzymatic protein, in which VET could be utilized to uncover energetic connectivity pathways and clarify the role of LFMs in the turnover step resulted in the FDH project, whose results are summarized in the following chapter.

8 Formate Dehydrogenase

8.1 Mutation Sites

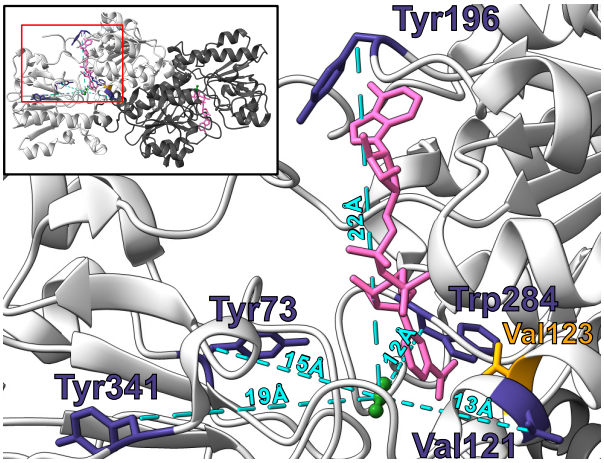

Figure 8.1: *Location of the AzAla (dark blue) and Gly (orange) incorporation sites relative to the active site with NAD$^+$ (pink) and N_3^- (green). The distances between C_β of the amino acids and the central N_3^- atom are labeled (cyan). The inset indicates the location of the magnified region (red rectangle) in the context of the dimer. Based on RCSB pdb: 5DN9 [167].*

Five mutation sites in different directions from the active site were chosen for the incorporation of AzAla into FDH. Four conservative mutation sites and a single non-conservative mutation close to the important Val123 residue were selected based on *in silico* mutations of all Tyr and Phe residues and residues near Val123 to Trp in YASARA using the FoldX plugin and their distance towards the bound azide [265, 266].

The mutation sites Tyr73, Val121, Trp284, Tyr341, and N_3^- are within one plane. The amino acids are distributed in four different directions with distances of 12–19 Å to N_3^- in the active site (Fig. 8.1).[1] Another mutation site is Tyr196, at a distance of 22 Å to N_3^- perpendicular to the aforementioned plane. For all mutants the distance to the azide bound in the second subunit is at least twice as large as the distance to that in the same subunit and at least 32 Å. if a VET signal of the second subunit can be observed it should be very weak and temporally seperated from that of the primary binding site. These five AzAla mutants allow for the investigation of VET towards N_3^- from five different directions over different distances. In Tyr196AzAla a VET pathway along NAD$^+$ is highly probable, which requires efficient VET between three molecules, from FDH to NAD$^+$ and from NAD$^+$ to N_3^-.

In the previous study on *Pseudomonas* FDH, motions of various active-site residues were identified as integral part of the reaction coordinate [263]. Because of the high sequential and structural similarity to *Cb* FDH, these residues can clearly be identified as Val93,

1 Because of the lack of mutant structures, distances are measured from the C_β of the substituted amino acid. The position of this atom is expected to be more conserved upon substitution with AzAla than the side chain position, which can easily change by a mere rotation along the C_α-C_β bond.

Asn119, Thr256, Arg258, Asp282, Gln287, His311, and Ser313 in the present FDH sequence. These residues are either in close proximity to or lie in one line between the former four AzAla incorporation sites and N_3^-. This situation improves the chance for the selective VET-driven excitation of an RPV involving one or multiple of these residues. Tyr196 is positioned at one end of the NAD^+ hydride acceptor, remote from the redox-active nicotinamide. Vibrational excitation in this region of the enzyme may lead to a motion that reduces the distance and/or the angle between the hydride donor and acceptor, two roles that have also been attributed to the motions of residue Val123 [219, 225]. To specifically excite dynamics in this residue, without directly perturbing its influence on the active site, the only non-conservative mutation Val121AzAla was proposed. It is in close proximity to and on the same helix as Val123. Because this helix is part of the dimer interface and contains no aromatic amino acids, this site, that purely interacts hydrophobically with the second subunit, was selected for AzAla incorporation.

Because both models, ATM and RPVs, have been consulted to explain the experimental results in FDH, both possible explanations should be explored. If the ATM is more appropriate to describe the processes in FDH, the Val123Ala/Val123Gly mutation, which leads to a temperature-dependent KIE, may be essential to observe an effect of the AzAla excitation on the protein kinetics. Therefore, each AzAla mutant should be combined with a Val123Gly (VG) mutation, as this is the mutation that leads to the stronger temperature dependence.

8.2 Results and Discussion

8.2.1 Protein Characterization

Of the ten FDH AzAla mutants (including the VG double mutants), eight were successfully expressed and purified, but the amount of Tyr341AzAla+VG was not sufficient for a VET experiment. The yield of full-length FDH mutants varied strongly between different expressions, despite efforts to optimize and maintain the expression conditions. It is assumed that this variability arises partially from context effects during expression. The efficiency of the SCS does not only depend on the tRNA/aaRS pair and the supply of loaded suppressor tRNA. The codon context, i.e., the bases in the vicinity of the stop codon on the mRNA has widely been recognized as important factor [130, 131, 268]. The variable amount of full-length FDH affects the relative purity that can be attained through the Ni-NTA chromatography. The most abundant contamination that remains after purification is a protein that runs at approximately 27 kDa in a sodium dodecyl sulfate (SDS) acrylamide gel (Fig. A.15). It is assumed that this protein is the 21 kDa SlyD, a common contamination in recombinantly expressed His-tagged proteins [269, 270].[2] This *E. coli* PPIase naturally binds to Ni-NTA columns because of an unordered C-terminus in which a stretch of 31 amino acids contains 14 His residues and could not be eliminated during this work.

In order to improve the purity of the FDH protein product in the continuation of this project, the following approaches can be pursued: (i) Most desirably, the amount of expressed full-length protein should be increased. Since various conditions of the expression have already been optimized and the expression results still vary considerably, this may be hard to attain. Additionally, a first test to improve expression through optimized codon

2 Because of the mismatch between the apparent and the actual mass, this protein was also called wondrous histidine-rich protein (WHP) [269].

context (changing the codon following TAG to ACT [Thr] or TCT [Ser] [130]) did not lead to a clear enhancement. (ii) A change of the purification methods may increase the protein purity. Insertion of a protease site (e.g., a TEV site) before the HisTag would allow for a second, orthogonal purification step, that is expected to remove SlyD. After a first Ni-NTA purification, FDH (with SlyD) could be treated with a protease to cleave-off the HisTag. In a subsequent Ni-NTA step, the HisTag, His-tagged TEV, and SlyD could be removed as successfully cleaved FDH does no longer bind to the nickel column. To keep eliminating truncated FDH (if not using a strain devoid of RF1), the HisTag should remain C-terminal. This will result in a remainder of the protease recognition sequence (e.g., ENLYFQ for TEV) instead of the HisTag. Alternatively further optimization of the current purification protocol may open up a new way to remove more SlyD (e.g., a salt gradient was reported to effectively elute SlyD from a Ni-NTA column [271]). (iii) The use of a SlyD-optimized *E. coli* strain could reduce the amount of SlyD in the final product. For this purpose the LOBSTR strain was already acquired as a gift from A. Möller (Osnabrück University) [270]. Before use, the RIL plasmid (used to increase the number of rare tRNAs in *E. coli*) needs to be purged from these cells. It shares the p15A origin of replication with pBU16 *Mj*AzuRS. However, the presence of RF1 in LOBSTR may increase the amount of truncation product, negating the gain from reduced SlyD contamination.

For a general assessment of the FDH construct on which the mutants are based, the WT protein was expressed and purified without a HisTag, for a proper comparison with the existing literature. The CD spectrum (Fig. A.14a) clearly indicates a well-structured protein containing α-helices and β-sheets and matches the published spectrum of this FDH isoform [260]. A double minimum between 205 nm and 225 nm is a feature of α-helical structures. In the CD spectrum of FDH this is combined with the negative feature of the β-sheet, which reduces the feature to one broad minimum and dominates the maximum at 195 nm, which would be at a lower wavelength for a purely helical protein [87, 272]. A fit of the CD spectrum for the secondary structure determination is in good agreement with that predicted from the crystal structure (Tab. A.4), ensuring the correct fold of FDH.

The steady-state kinetic parameters of FDH for the NAD^+ substrate have been determined via Michaelis-Menten analysis (Fig. A.14b). The Michaelis constant K_M and the catalytic rate constant k_{cat} are in good agreement with values in the literature [167, 273]. The correct fold of the enzyme and the observation of catalytic activity verify that the chosen approach of recombinant expression of FDH in *E. coli* is feasible. The present construct therefore meets the prerequisites for VET experiments (a correct fold) and for the modulation of the enzyme activity by excitation of LFMs and/or RPVs (general activity).

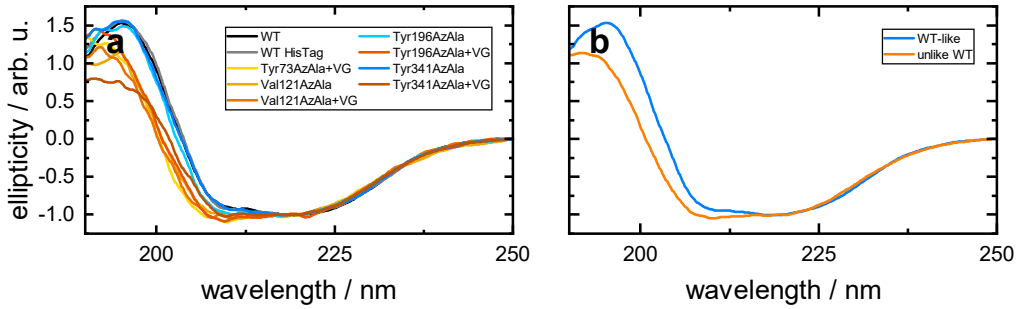

Figure 8.2: *Normalized CD spectra of different FDH variants (a). Only here the HisTag is explicitly mentioned for the WT protein. The spectra cluster into two groups, WT-like (hues of blue and gray) and unlike the WT (hues of orange). Mean spectra of the two groups (b).*

The CD spectrum of the His-tagged FDH WT closely matches that of the WT without tag (Fig. 8.2a). This confirms, that the tag does not disturb the fold of the protein. The CD spectra of the Tyr196AzAla and Tyr341AzAla mutants also nicely match the spectrum of FDH WT. But the spectra of all other mutants form a distinct group. For them the slope and the maximum is blueshifted to 192 nm and its amplitude reduced. This may be indicative of an increased ratio of partially unfolded protein. However, from the CD spectra a confident interpretation which part of the structure was influenced in which way is not possible. It is striking that all VG double mutants are in the latter group of spectra. Tyr196AzAla and Tyr341AzAla have WT-like CD spectra, but their double mutants do not. Though, it is known from X-ray crystal structures, that the Val123Ala (pdb: 6D4B) and Val123Gly (pdb: 6D4C) have no influence on the protein fold [219]. The disturbed CD spectrum is therefore not caused by one of the mutations itself. Large relative amounts of differently-folded protein impurities could dominate the CD spectra, but judged from representative SDS gels this is not the case (Fig. A.15). This leaves only sample specific causes (e.g., during the refolding step in the purification, which in principle could be abolished in RF1 devoid strains) for the disturbed CD spectra, which, from the available data, can not further be specified.

For the VET experiments, NAD^+ and N_3^- were added to the protein. These binding partners have dissociation constants of 92 µM and 43 nM, respectively, they establish a highly stable ternary complex with the WT FDH [167]. The FTIR spectra of the complex can be used to ensure that the proteins binds N_3^-. In other protein/N_3^- complexes, the binding may be indicated by a clear shift of the absorption band of N_3^- [211], however, this is not the case in FDH.

The asymmetric stretching vibration (which should be used as VET reporter mode) of free N_3^- has a Lorentzian band shape centered at 2048 cm^{-1} and a FWHM of 27 cm^{-1} (Fig. 8.3a). In the ternary complex with FDH, the main absorption feature blueshifts by 1.5 cm^{-1}, but the FWHM reduces significantly to 10 cm^{-1}. A fit with two Lorentzians, one of which represents unbound N_3^- with a fixed position and FWHM, systematically underfits the data. It can not sufficiently describe the asymmetry of the band (which can be seen, but is not discussed in reference [97]).[3] An additional feature in the range of around 2038 cm^{-1} (FWHM: ~16 cm^{-1}; Tab. A.6) is required. It is required when spectrally no free N_3^- can be identified (Fig. 8.3b–c), but also when N_3^- is in excess (leading to significant amounts of free N_3^-;

3 Such a fit underestimates the intensity at the red flank and overestimates the intensity at the blue flank (Fig. A.16).

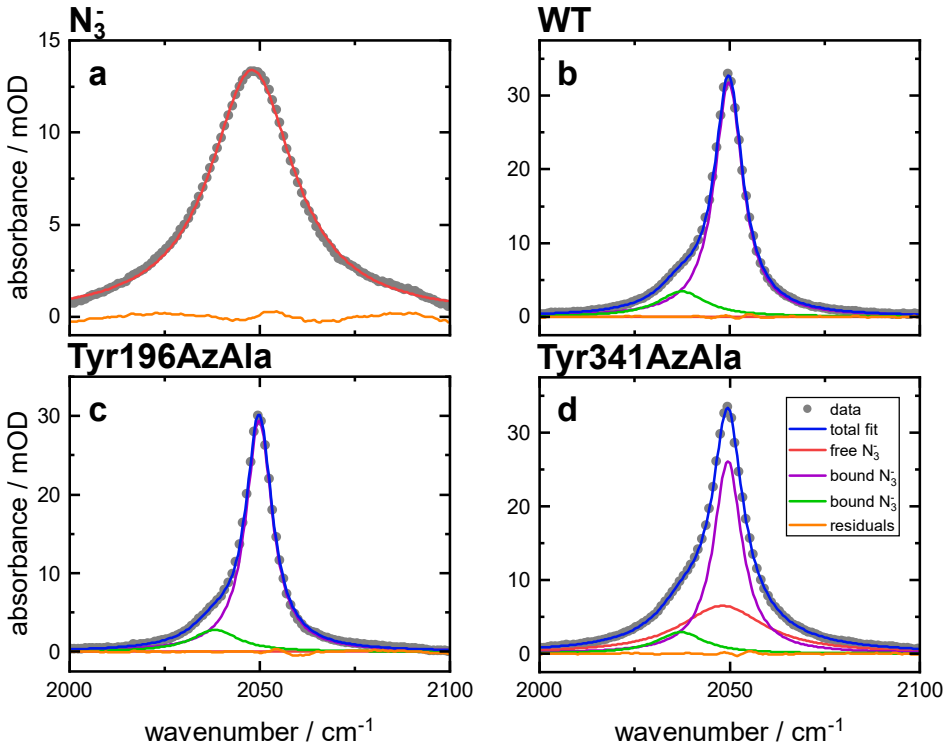

Figure 8.3: *FTIR spectra of free N_3^- (a) and of FDH WT (b), FDH Tyr196AzAla (c), and FDH Tyr341AzAla (d) in complex with NAD^+ and N_3^-. Sums of up to three Lorentzians are required to explain the complete absorption band.*

Fig. 8.3d). It is also evident in the second derivative of the raw absorption spectrum. For FDH Tyr73+VG, Val121AzAla+VG, Tyr196AzAla+VG, and Trp284AzAla no N_3^- binding could be observed, and the side feature was not visible. Importantly, these were all mutants whose CD spectra differed from that of FDH WT, confirming that the CD spectrum is a great indicator for the suitability of an FDH sample for VET experiments (Fig. A.17).

From the crystal structure (and its electron density) of the FDH complex, only a single conformation of N_3^- is expected. From the spectroscopic data, the origin of the side feature can not clearly be identified. The binding of N_3^- to an occasional protein contamination is excluded from the highly pure WT sample. Furthermore, it would be expected to occur in some of the samples, in which the N_3^- did not bind to FDH. As this is not the case, also nonspecific interactions at the protein surface can be ruled out. Supposedly the peak therefore originates from a fraction of N_3^-, that binds to FDH albeit in an altered geometry.

For the VET experiments, it is generally beneficial to have high sample concentrations. The low extinction coefficient of AzAla can partially be compensated with high pump pulse energies (~20 µJ at 612 nm), which are routinely attainable with the setup. The use of the N_3^- anion as VET sensor is a further improvement compared to the organic ncAAs that have been presented in Part II, as its extinction coefficient is 3–10 times higher. However, the long distances between VET donor at one site and VET sensor at a distinct site in a protein lead to significant losses of energy, that reduce the signal size that can be reached. But for FDH, and even more so the VG mutant, it was reported that it is prone to aggregation at concentrations exceeding 1.5 mM [219]. In this work, it was even possible to reach concentrations

of 3 mM with the WT for short periods of time. The AzAla mutants were less stable. Only Tyr196AzAla and Tyr341AzAla could be concentrated above 1 mM (Tab. A.7).

Many stabilizing agents have been used to increase protein solubility [274]. Whether this approach can be used to reach higher concentrations of FDH in the future, was tested by concentrating Val121AzAla in a buffer containing 1 M trehalose. Indeed the concentration of Val121AzAla+VG could be increased by from 0.7 mM to 1.1 mM. As Val121AzAla did not bind N_3^- in the first place, Tyr196AzAla was used to test if the high concentrations of the sugar interfere with N_3^- binding, which is not the case (Fig. A.17f). But the presence of high concentrations of trehalose alter the shape of the H_2O combination band, that gets narrower and develops an association band through interactions with the sugar [275] (Fig. A.18), which can be corrected for using the appropriate background. Therefore, 1 M trehalose may be used in future experiments to reach higher concentrations of sensitive protein mutants. How the trehalose influences the VET experiments themselves is addressed in the following section.

8.2.2 VET Experiments in FDH

It was ensured that the detected signals originate from the bound N_3^- and are invoked by the VET from AzAla. While the negative control for N_3^- is recorded simultaneously with each experiment in the split sample cell, as the reference compartment does not include the IR label. The negative control for AzAla is represented by the experiment with FDH WT (Fig. A.19 top). No structured signals can be observed in this experiment, only the spectrally flat decay of the coherent artifact, as it is expected in the absence of a specific absorber of the pump light. Furthermore, experiments with the mutants that did not indicate N_3^- binding do not show a N_3^- signal, as exemplified on Val121AzAla+VG (Fig. A.19 bottom). At early delays, the coherent artifact decays quickly, before the energy of the AzAla excitation reaches the solvent, which gives rise to a broad negative solvent feature at wavenumbers too high for N_3^-. This representative experiment illustrates how the solvent signal can reliably indicate a generally operative VET experiment with AzAla present, AzAla successfully excited (overlap between pump and probe pulses), and transfer of energy that finally reaches the solvent (redshift of the H_2O combination band upon heat uptake [276]). The split sample cell approach allows for the complete correction of these background signals. The absence of the N_3^- signal also in the difference between both compartments prove that the free N_3^- in the solvent does not generate a detectable signal.

VET experiments with Tyr196AzAla, Tyr341AzAla, and Tyr196AzAla with 1 M trehalose did result in clear VET signals only of the main feature of bound N_3^- (Fig. 8.4). The peak times of the maxima are 9.8 ps, 8.5 ps, and 9 ps, respectively. The signals have amplitudes of 20–30 µOD, highly similar when considering the sample thickness of 150 µm for Tyr196AzAla in standard buffer against 100 µm of the other samples (Tab. A.5).

As expected (and observed in PDZ) the peak times of Tyr196AzAla and Tyr341AzAla correlate with the distance between VET donor and VET sensor. The distance increased by 3 Å in Tyr196AzAla leads to a delay of the peak time by 0.8 ps. For Tyr196AzAla in 1 M trehalose, the peak time shifts significantly by -1.3 ps. Though the mechanism of protein stabilization by trehalose is not yet fully understood, its presence clearly alters the interaction between the solvent and protein [277]. Like H_2O, trehalose can act as H-bond donor and acceptor, furthermore, its more complex structure increases the number of LFMs. Therefore, it is assumed that trehalose at the protein surface accelerates the rate of energy transfer out

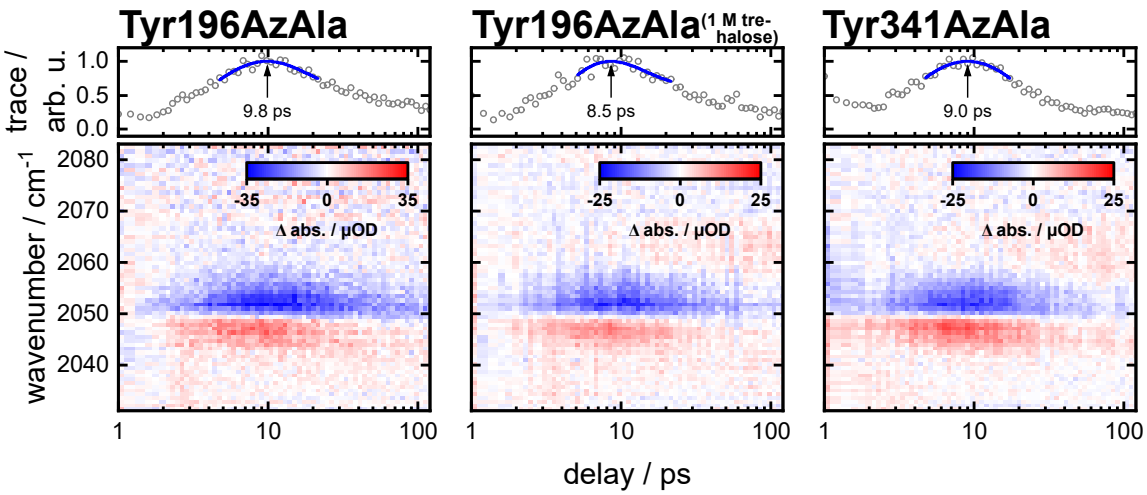

Figure 8.4: *VET data and traces of FDH Tyr196AzAla (a; and in buffer with 1 M trehalose b), and Tyr341AzAla (c), traces of the signals and fits for the determination of the indicated peak times (top).*

of the protein, which shifts the peak time to shorter delays. Because of the dynamics of the VET process, this shift will not be constant but vary for each position of the VET donor (in absolute and relative value). For the prospect of using trehalose for protein stabilization in future experimental series, this needs to be considered. Furthermore, the association band of the solvent and trehalose can be observed in the solvent heating signal. In contrast to the overall redshift of the solvent absorption, this feature experiences a blueshift upon heating, as is the case for most vibrational bands upon weakening of hydrogen bonding (Fig. A.18b). As this signal is present in the sample and reference compartment in a split sample cell and shares the same dynamics as the solvent heating, it can equally be corrected for. With these results, it can be concluded that VET experiments with trehalose are generally feasible. However, all experiments should be conducted under the same conditions, or further investigation should take place to satisfy a compromise between the predictability and error of the peak times when comparing experiments with or without trehalose.

In these VET experiments the signals are clearly visible in the sample compartment data even before applying any corrections (not shown). This is in stark contrast to the experiments on the different PDZ mutants that were discussed in Part II. PDZ could be concentrated to much higher concentrations than FDH, this largely compensated the lower extinction coefficient of Aha as compared to N_3^-. But in the VET experiments, the higher concentration of PDZ leads to an increased concentration of AzAla and a higher amount of total energy that is deposited in the sample.[4] This leads to a lower contrast between the solvent background signal and the signal of the VET sensor. In the PDZ Aha mutants the solvent signal is 10–20 times more intense than the VET signal, in FDH only 2–3 times. Furthermore, the signal of N_3^- is redshifted compared to the azido ncAAs. This shifts the signal to the nodal line of the solvent signal, reducing the spectral overlap of both contributions.[5] Additionally, it is expected that a larger part of the photon energy is distributed within the protein before reaching the solvent, due to the buried AzAla. The more exposed the AzAla, the more energy is lost to the solvent, before contributing to the VET signal of

4 The energy within each protein remains the same, the pump photon energy.
5 The blueshifted nitrile ncAAs are completely within the broad, negative feature of the solvent signal.

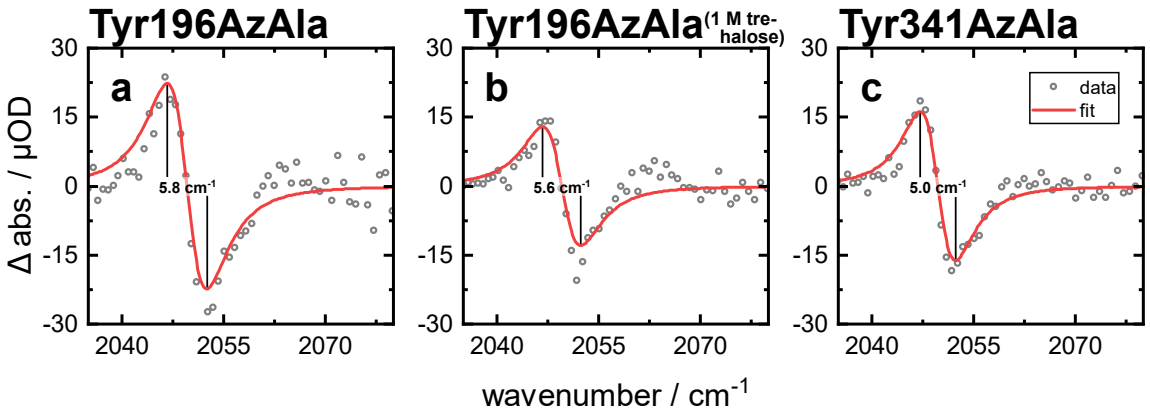

Figure 8.5: *Peak separation in the VET signals of FDH Tyr196AzAla (a; and in buffer with 1 M trehalose b), and Tyr341AzAla (c). The fit, for the determination of the separation, is the sum of two shifted Lorentzians with amplitude ±1 OD and FWHM matching of the respective FTIR band.*

the VET sensor. This is supported by the scaling factors that result from applying the solvent correction with the long-term spectrum, instead of the split sample cell. One can see how the saturation of the solvent signal systematically slows down with reducing solvent exposure of the AzAla between the dipeptides, PDZ (with AzAla in the ligand), and AzAla within FDH (Fig. A.20).

From the control experiments with FDH mutants that did not bind N_3^- (Fig. A.19 bottom), it is probable that the signal observed with Tyr196AzAla and Tyr341AzAla indeed belongs to the bound N_3^-. The confidence can further be boosted by the dynamics of the signal. A signal of unbound N_3^- is expected to have the same dynamics as the background signal. Instead, we observe a rise-and-decay behavior of the transient signal. For an assignment based on the spectral position of the signal, the two components are close, however, the FWHM provides further evidence. Besides the narrow appearance, the separation $\Delta\tilde{\nu}$ of the positive and negative extrema of the signal can be analyzed. This separation has a minimum value of $\Delta\tilde{\nu}_{min} = \mathrm{FWHM}/\sqrt{3}$ for Lorentzian band shapes (Eq. 10.1) and $\Delta\tilde{\nu}_{min} = \mathrm{FWHM}/\sqrt{\ln 4}$ for Gaussian band shapes, which correspond to the separation of the inflection points of the original band. For the free N_3^- $\Delta\tilde{\nu}$ must be larger than 15.6 cm^{-1}, because of the FWHM of 27 cm^{-1}. Instead, the observed $\Delta\tilde{\nu}$ is 5–6 cm^{-1}, matching the minimum separation of the narrower bound N_3^- (Fig. 8.5).

Note, that in principle the peak separation could be used as an indicator of the actually occurring shift, especially since it is straightforward to determine. However, in the previous part, the expected amplitude (which needs to be calculated from the experimental conditions) of the shifted signal was used to determine this shift. This is because the VET generally causes a weak shift of only ≤13% in the dipeptides (Tab. 4.1), and less with increasing distance. In the small-shift regime, the signal amplitude is more sensitive to the shift than the peak separation. Therefore, a separation-based analysis of the VET shift cannot resolve the subtle shifts that can be observed from the amplitude-based approach.

The presented experiments impressively demonstrate that VET experiments in AzAla-carrying FDH mutants with N_3^- as VET sensor are feasible, despite the restriction to low concentrations. From the perspective of the VET donor, this approach can potentially increase the energy that contributes to the signal by reducing the direct loss to the solvent,

compared to the highly solvent exposed AzAla in a peptide ligand. The only drawback of the VET sensor is its restricted applicability to proteins that naturally bind N_3^-. But bound to FDH it combines the advantages of the ncAAs Aha and CNPhe. It has an extinction coefficient that even surpasses that of Aha and at the same time its typical FWHM is comparable to that of CNPhe. The spectral overlap of free and bound N_3^- does not impede its use as VET sensor in FDH. Furthermore, by varying the site of the VET donor instead of the sensor, it can be ensured that the band shape and amplitude remains identical.

8.2.3 Steering FDH Activity by Illumination

To investigate the role of LFMs in FDH catalysis, the steady-state protein activity was assayed unilluminated and under illumination with a 526 nm laser. The AzAla $S_1 \leftarrow S_0$ transition was excited at 612 nm in the VET experiments, but the broad absorption band does still allow for excitation at 526 nm, a wavelength that is available with high powers from the pump lasers of the ultrafast laser systems. Just like in the VET experiments, the internal conversion is expected to excite LFMs in the vicinity of the AzAla label, that, if coupled to the reaction coordinate, can then alter the reaction kinetics.

These experiments were conducted with FDH WT, Val121AzAla, Tyr196AzAla, Tyr196AzAla+VG, Tyr341AzAla, and Tyr341AzAla+VG, other mutants could not be used because of low sample amounts that were depleted by the VET experiments. The sample temperature was lowered to 3°C to minimize the reaction speed in the absence of illumination, in order to improve the contrast to possible light-induced effects. Despite flushing the setup with dry air, an issue that could not be resolved was the occasional fogging of the cuvette, which was visually confirmed. This induced light scattering and thereby spikes in the absorption experiments, which needed to be corrected in the post processing (Sec. 2.7.2). Otherwise, the custom implementation allowed for reproducible experiments.

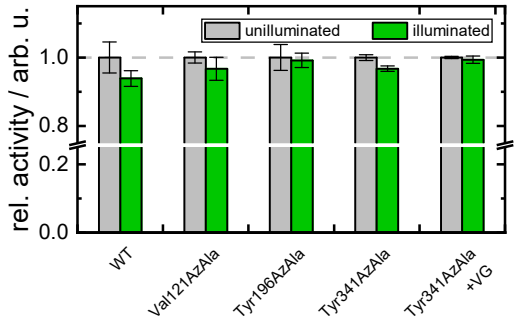

Figure 8.6: *FDH mutant activity unilluminated and under illumination. The activity was normalized to the unilluminated activity for each variant. The error bars indicate ±two standard errors of the mean (SEM) of 3–6 experiments per condition (SEM<2.5% for each condition).*

For all investigated FDH variants, the enzyme activity (in terms of product formed per second) is reduced by 6% for FDH WT and by 3% or 1% for the mutants under illumination (Fig. 8.6). However, this reduction is within two SEM for all variants except Tyr341AzAla. The apparent k_{cat} at 3°C without illumination was between 0.1 s^{-1} and 0.2 s^{-1} but much lower for the active site mutants Val121AzAla, Tyr196AzAla+VG (no activity detected), and Tyr341AzAla+VG. The CD spectra of these mutants did not match that of the WT and for them no or little azide binding was detected (Tab. A.7).

Interestingly, even the activity of FDH WT is reduced under illumination. Because of the lack of AzAla, no excitation-induced effect is expected. In these experiments, the sam-

ple temperature was determined using a digital thermometer immersed in the reaction solution. An external thermostat was set to meet the target sample temperature. Though the temperature probe was placed directly beneath the solution surface, ~5 mm above the pump beam, a direct influence of the light on the read temperature cannot completely be ruled out. If this was the case, the real temperature in the illuminated experiments may have been lower than intended, which would reduce the observed activity.

If one assumes that the temperature offset was the same in all illuminated experiments, and the effect of the temperature is the same for all FDH variants, one could normalize the illuminated experiments to FDH WT. Under these circumstances, one indeed observes an increase 3–6% of the mutant activities upon illumination (not shown). But only the activity of Tyr341AzAla and its double mutant are significantly different from the unilluminated case. This is because their SEMs are especially low.

A reliable assessment of the light-induced effect is only possible after establishing a method that assures matching sample temperature without and with illumination. A simple solution may be to further shield the probe from the incident light, e.g., with a small tube that is not in direct contact to the probe. It should be ensured that the shield is not in direct contact with the probe, and the frequent exchange of the solution surrounding it needs to be maintained (achieved by constant stirring). While there is room for improvement, this preliminary result may hint towards a direct observation of the rate-enhancing role of LFMs in enzymatic catalysis.

To put these experiments into perspective, they are compared to recently published results of accelerated enzyme catalysis in dihydrofolate reductase (DHFR) [278]. Like FDH, DHFR catalyzes a hydride transfer, and various studies suggested networks of coupled motions and allosteric sites that may be involved in the catalytic step [217, 279–281]. In the aforementioned experiment, DHFR was covalently bound at four different binding sites to gold nanoparticles (AuNPs) which were then excited at 527 nm with pulsed or continuous wave (cw) lasers. Steady-state rate enhancement of 50–75% was observed for two of the binding sites upon excitation with a pulsed lasers. Instead of a gradual increase of the activity with increasing laser power, threshold behavior was observed. The authors interpret these findings in terms of an accelerated sampling of the protein energy landscape together with the non-Boltzmann excitation of collective breathing modes of the protein that modulate the active site structure [278].

The VET experiments have shown that the excitation energy of AzAla completely dissipates into the solvent within tens of ps, whereas the excessive heat released by the AuNPs needs up to 2 ns to fully dissipate [278]. Despite using higher laser powers to excite the AzAla, the large extinction coefficient of the AuNPs of 367×10^6 $M^{-1}cm^{-1}$, six orders of magnitude higher than that of AzAla at the excitation wavelengths (~200 $M^{-1}cm^{-1}$), leads to much higher excitation densities. Taking into account sample concentrations and illumination conditions, only assuming that the total volume of sample was the same in the published experiments and those presented here, we can estimate that for each enzyme in the solution the energy of 12000 photons is absorbed per second by the AuNPs. Practically every laser pulse (1 kHz and 5 kHz in the published experiments) leads to an excitation. Because of the lower extinction coefficient of AzAla, each FDH molecule only receives approximately 7 photons per second (at 1 kHz repetition rate).

From these numbers, the absence of a clear activating effect is less surprising. The AuNP-bound enzymes are more often excited for longer periods of time. Furthermore, within a single pulse the AuNPs accumulate and release the energy of multiple photons per bound enzyme. AzAla can not provide these amounts of energy. However, the non-equilibrium

state that is created through the IVR is just required to specifically populate those LFMs that are involved in the reaction coordinate, whose energies are well below the photon energy. With the highly site-specific excitation of AzAla (which can not only be placed at the protein surface), a more selective injection of vibrational energy is expected that, if appropriately located, could still induce acceleration of the catalysis. It is instead required to increase the chance of coincident excitation and chemical reaction, therefore it would be beneficial to induce higher excitation densities in the sample. In the present configuration, the pump power was limited by a rapid increase of the experimental noise. The effective extinction coefficient could be increased by ~65% by exciting AzAla closer at its maximum near 600 nm, but a light source with a high power at that wavelength would be required. The optimal solution for this type of experiments could be a VET donor ncAA with a significantly higher extinction coefficient. This option is currently being investigated in cooperation with other groups from Goethe university. A 50-fold increase of the extinction coefficient could be reached, but a method for the incorporation into a protein remains to be implemented [282].

9 Conclusion

The aim of this part of the thesis was to establish a new protein system to investigate VET and the role of protein dynamics in the context of enzymatic catalysis. Because the VET donor AzAla can inject the energy of a photon as vibrational energy into a protein at the site of incorporation, it can also be used to specifically excite LFMs. If the motion of these modes is involved in the enzymatic reaction, a direct influence on the activity is expected.

The first target protein that was used for this purpose was CypA. The WT of the protein was successfully expressed and its enzymatic activity was proven with after establishing a colorimetric assay. The local expression of AzAla mutants could not be established, and the one mutant that was externally expressed by the author did not exhibit enzymatic activity.

FDH was chosen to replace CypA. FDH WT and FDH AzAla mutants were locally expressed. In VET experiments with this protein, the exceptionally small VET sensor N_3^-, a natural inhibitor to FDH, was tested. In these experiments, it was shown that the non-covalent ligand is a well suited VET sensor. In contrast to the ncAAs, its application is restricted to proteins that bind N_3^- and its bandwidth and position may vary from protein to protein. On the other hand, the unchanged position of the VET sensor has the advantage of an unchanged band shape in different mutants. This increases the comparability of the VET signals, which is not granted for Aha and N_3Phe at different sites of a protein. Though the absorption of free and FDH-bound N_3^- are not spectrally separated, the signal can clearly be identified as originating from the bound species because of its bandwidth. In the protein the FWHM is reduced from 27 cm^{-1} to 9 cm^{-1}. Therefore the N_3^- combines an even higher extinction coefficient than the organic azides with the narrow bandwidth of nitriles, making it the perfect VET sensor for proteins with low solubility.

The enzyme activity of the AzAla mutants was measured and compared between the unilluminated case and under illumination with a 526 nm laser. From these results no clear effect of the illumination was visible. In the current setup, the activity of all FDH variants decreased upon illumination. Even the negative control (WT) was influenced by the illumination. When this effect, that may be caused by a too low temperature, is corrected for, an enhancement of 3–6% is observed in the mutants. However, the reliability of this effect remains elusive until the uncertainty about the temperature is fully eliminated.

10 Summary and Outlook

VET experiments in proteins contribute to our understanding of ubiquitous biological processes such as allostery, dissipation of excess energy, and possibly enzymatic catalysis. All of these have been under investigation for extended periods of time, but still many questions remain unanswered. The aim of this thesis was to extend the applicability of the existing techniques, to find tailor-made solutions for the great variety of proteins and protein sites. Furthermore, new target proteins were meant to be identified and established, that could expand the repertoire of VET targets investigated in the Bredenbeck group towards the role of VET and LFMs in enzymes.

Labels for Vibrational Energy Transfer (Part II)

For the spectroscopic investigation of VET two functionalities are needed: a VET donor, that converts the energy of an absorbed photon into vibrational energy and injects it site-specifically into the protein, and a VET sensor, that can report on the arrival of vibrational energy through a change in the IR absorption. The originally published pair of ncAAs that assume these functions are AzAla and Aha, respectively [29]. This VET sensor could be incorporated into proteins by means of SPI, but this method requires a protein sequence that lacks Met residues [30], a serious constraint. Furthermore, the ncAA Aha is a great surrogate for Met in terms of chemical and steric properties, but not for most other amino acids which it could replace in proteins.

In this part of the thesis, the results of which were peer-reviewed and published [136], the applicability of several new ncAAs as VET sensors was assessed. The ncAAs contained azide or nitrile groups, whose stretching vibrations absorb IR light near 2120 cm^{-1} and 2220 cm^{-1}, respectively. The VET signal amplitude is linearly dependent of the extinction coefficient, which is typically 2–4 times lower for nitriles than for azides [103, 170]. But the VET signal originates from the shifting of the absorption band due to coupling to populated LFMs [24]. The VET sensors can therefore not solely be assessed by the extinction coefficient. If two bands of the same amplitude shift equally, the difference signal is larger for the narrower band. For small shifts, as is the case in VET experiments, the signal amplitude rises linearly with the relative shift, the ratio between the shift and the FWHM of the band. Because of the non-Boltzmann population distributions that are caused by VET, the VET shift is not necessarily the same as the thermal shift that can be measured in temperature dependent FTIR. But the individual populations and coupling constants are not readily obtainable, therefore the experimental assessment of the VET sensors is imperative.

Dipeptides. The ncAAs N_3Ala, Aha, N_3Phe, CNPhe, and CNTrp were coupled to AzAla in dipeptides. Their spectral properties have been compared from the FTIR spectra in H_2O, DMSO, and THF. Besides the lower extinction coefficient, the absorption bands of the nitriles (FWHM 6–12 cm^{-1}) are significantly narrower than those of the azide compounds (FWHM 19–34 cm^{-1}). Furthermore, the organic azides also exhibited the Fermi resonances already described in the literature [171–173].

The solvent has a strong influence on the VET signals that are measured. For all dipeptides, the signal amplitude in H_2O was significantly lower than in the other solvents. This

was accompanied by a strongly delayed peak time in THF and a less delayed peak time in DMSO compared to H_2O. Both effects can be explained by the interaction of the solvent with the dipeptides. This leads to the fast dissipation of excess energy into the solvent, reducing the maximum shift that can be induced and shifting the peak time to shorter delays. This effect is strongest in H_2O, which is protic and polar.

In the subgroup of azide and nitrile sensors, the peak time increased with the size of the side chain. This demonstrates the dependence between peak time and the distance between VET donor and sensor. The peak time of N_3Phe is 0.9 ps larger than that of CNPhe. These two ncAAs share the same side chain structure, so that this difference is caused solely by the different reporter groups and their coupling to the LFMs. This result is exactly the opposite of what could be expected on the basis of the vibrational lifetimes, which is shorter for N_3Phe than for CNPhe.

In the VET experiments of the dipeptides, it was observed that the signal amplitudes of the nitriles and of the azides are more similar than the lower extinction coefficient of the nitriles would suggest. The effect of the extinction coefficient is indeed compensated by the fact that the shifted band is narrower. To further solidify this findings, it was required to determine the spatial distribution of excited molecules and their contribution to the IR difference signal. Using calculations that consider these influences, it was determined that 12% of the FTIR signal contribute to the experimental difference signal. A fit to the VET signal could then be performed to yield *absolute* values for the shift at each delay. Two approaches for the fit were compared: (i) directly shifting the scaled FTIR signals and (ii) fitting the FTIR signal with a Voigt line shape, whose area remains constant during the shift, but allowing its shape to change. A better agreement between data and fit is obtained from method (ii). The additional parameters can account for changes in the band width, that come along with the shift through the coupling of the LFMs. However, both methods result in similar shifts and nicely reproduce the dynamics of the VET signal as seen from the traces. From this analysis, it was confirmed, that the relative shift of CNPhe exceeded that of the other ncAAs.

Proteins. From the dipeptides alone, the general suitability of the ncAAs as VET sensors could be judged. But for the future prospect of investigating VET in proteins Aha, N_3Phe, and CNPhe were incorporated into the PDZ protein domain at two distinct sites.

In these more complex systems, the importance of the azide Fermi resonances was underlined. The line shapes of these sensors in the proteins were different from those of the dipeptides. Phe325Aha had a very narrow and intense IR absorption band, whereas a pronounced two-band structure could be seen for N_3Phe at both protein sites. This makes the band shape and intensity of azide sensors in a protein unpredictable and complicates the comparison of signals. For CNPhe the band width and position changed in the protein, but no additional features emerged.

The complex band shapes of the FTIR signals did fully transfer into the VET signals. All features that were seen in the VET signals could be recreated from a shifted difference signal of the FTIR bands. The signal amplitudes were dominated by the extraordinary band shapes of the azide sensors. But the comparison of the well-behaved Ile327AzAla and the CNPhe variants confirmed that the lower extinction coefficient of CNPhe does not necessarily lead to a lower VET signal, also in the protein.

The peak times are still larger for increasing donor-sensor distances, but the order of the peak times is different for the VET sensors in the protein than in the dipeptides. In the proteins, CNPhe peaks first, N_3Phe second, and Aha last (with a difference of 1.3 ps between CNPhe and Aha). In the small dipeptides, the dynamics are largely influenced

by the interaction with the solvent, which is less dominant in the protein. On the other hand, in the protein a large side chain can also shorten the effective distance and there are countless paths between the VET donor and VET sensor in the protein. The influence of all these pathways makes the VET experiments in the protein less susceptible to each single step in the VET process and extends the characterization of the VET sensors to more general experimental conditions.

In conclusion, these experiments indicate that nitrile groups can serve as VET sensors without strongly impairing the signal amplitude, because the lower extinction coefficient is largely compensated by the narrower band. The new VET sensors combine the use of more diverse incorporation techniques with different chemical and sterical properties. With these findings, the influence of different side chains on the peak time can be estimated, so that in a single protein different sensors can be used at different sites. This increases the number of potential target proteins and allows the sensor incorporation to be less disturbing to different protein sites, paving the way to investigate VET in native-like proteins at every single site.

Vibrational Energy Transfer and Low-Frequency Modes in Enzymes (Part III)

Aim of this part of the thesis was to establish a new protein system to investigate VET and the role of protein dynamics in the context of enzymatic catalysis. Because the VET donor AzAla can inject the energy of a photon as vibrational energy into a protein at the site of incorporation, it can also be used to specifically excite LFMs. If the motion of these modes is involved in the enzymatic reaction, a direct influence on the activity is expected. This hypothesis has existed for many years, but could not yet be verified [21, 283, 284].

Cyclophilin A. The first enzyme that was chosen for this task was CypA. For this a well studied PPIase different studies proposed that the protein dynamics are catalytically relevant [246, 250, 285]. The WT protein was successfully expressed and purified. A chymotrypsin-coupled enzyme protein assay was implemented with which the chemical activity of CypA was proven. Based on previous publications, several peptide ligands were designed, that should act as scaffold to carry an ncAA VET sensor into the active site of the protein, but the previously reported high affinities between CypA and the peptides could not be verified.

The first local attempts of cotranslational AzAla incorporation, using a newly developed aaRS/tRNA pair [137], were carried out to express CypA mutants containing AzAla at various sites. With the methods and materials present at that time, this goal was not achieved. A single CypA mutant (Gly14AzAla) was successfully expressed by the author in the labs of N. Budisa at Technische Universität Berlin. However, this mutant did not exhibit enzymatic activity. Because of the various difficulties encountered, the project was discontinued in favor of another enzyme system.

Formate dehydrogenase. The next enzyme that was chosen was FDH. It was the first enzyme in which active site LFMs were detected using 2D-IR spectroscopy [210]. Several studies have found evidence for the involvement of LFMs in FDH katalysis [167, 219, 225, 262, 263]. This specific system allows the use of a natural VET sensor. It forms a stable complex with NAD^+ and N_3^-, a superb IR label. It offers the chance to test small non-covalent ligands as VET sensors.

An important improvement that enabled this project, was the supply with AzAla. The locally used AzAla was originally laboriously produced via chemical synthesis in the group

of N. Budisa in Berlin. But in 2020 a simple enzymatic synthesis method was published [135]. As part of this thesis, this method was locally established and by the time the FDH project was launched, crude AzAla could be produced in large quantities and with consistent quality. This allowed for a constant supply with AzAla during expression testing and for large scale expressions.

AzAla was successfully incorporated into FDH at several sites, under these circumstances. CD spectra of these mutants were compared to the spectrum of FDH WT. Only the spectra of two mutants (Tyr196AzAla and Tyr341AzAla) matched the spectrum of FDH WT. The CD spectra are not expressive enough to pinpoint the exact structural reasons for the deviations of the other mutants. The most probable cause is an untraceable effect during protein purification. In includes unfolding/refolding step. This was required to remove truncated protein from the sample (which otherwise remained because of the FDH dimerization). This step may be redundant while using expression strains without RF1, making a higher rate of functional AzAla mutants possible.

FTIR spectra revealed that only those variants whose CD spectra matched that of the WT could bind N_3^-. Despite spectral overlap of free (2048 cm^{-1}) and bound N_3^- (2050 cm^{-1}), the latter could be distinguished by its 3 fold narrower FWHM of 9 cm^{-1}.

For most non-binding variants, the maximum concentration was well below that of FDH WT, Tyr196AzAla, or Tyr341AzAla, which could be concentrated up to 2 mM. 1 M trehalose was added to the buffer to test whether this could increase the solubility without impeding the experiments [274]. Because this was the case, this measure should be considered in the future of this project.

Clear VET signals were observed for Tyr196AzAla and Tyr341AzAla, with peak times of 9.8 ps and 9 ps, respectively. In agreement with the 3 Å increased distance between AzAla and bound N_3^- in Tyr196AzAla as compared to Tyr341AzAla. It was proven that the observed VET signals were only caused by the bound N_3^-. In buffer containing 1 M trehalose, the Tyr196AzAla peak time was reduced to 8.5 ps. The altered solvent interactions need to be considered when the trehalose buffer is adopted. For optimal comparability all experiments should be performed with the same buffer composition, as its influence on the VET dynamics is non-negligible.

The activity of FDH WT and four AzAla mutants was assayed under saturating substrate conditions without and with excitation of AzAla. In these experiments, a slight reduction of the activity was observed under illumination, even for the WT. So far, a mismatch in the sample temperature can not be excluded as the reason for this decrease. These results were put into perspective by comparing them to a similar experiment with DHFR and AuNPs, in which significantly stronger excitations were possible [278]. However, it may be possible to slightly (but significantly) increase the enzyme activity even with the current experimental design, but the final verification requires improved control over the sample temperature.

The experiments with FDH highlight the potential of small molecule ligands as VET sensors. N_3^- is especially attractive because of its simplicity, which prevents Fermi resonances, and its high extinction coefficient. Its use adds many metalloproteins to the potential targets of VET experiments [286–289]. These findings could also bring attention to other alternative VET sensor ligands, such as IR labeled flavins or NAD^+ [98, 99].

Indications for future improvements in the expression and purification strategies of FDH were given, which can potentially lead to a more reliable sample preparation. Newly designed FDH AzAla mutants will allow to analyze VET pathways and groups of dynamically coupled residues by characterizing VET from different directions towards the active site step by step. Furthermore, the optimized FDH activity assay could show a direct acceleration of an enzymatic reaction, caused by the specific activation of protein LFMs. In this way it will help finding an answer to the long-standing questions about the outstanding catalytic power of enzymes.

Zusammenfassung

Schwingungsenergietransfer (engl. vibrational energy transfer, VET) Experimente in Proteinen tragen zu unserem Verständnis grundlegender biologischer Prozesse wie Allosterie, Dissipation von Überschussenergie und möglicherweise enzymatischer Katalyse bei. All diese Prozesse werden schon seit langem untersucht, doch viele Fragen bleiben nach wie vor unbeantwortet. Das Ziel dieser Arbeit war es, die Anwendungsbereiche der bestehenden spektroskopischen Technik zur Untersuchung von VET zu erweitern, um maßgeschneiderte Lösungen für die Vielfalt von Proteinen und Aminosäureumgebungen zu finden. Außerdem sollten neue Zielproteine identifiziert und etabliert werden, die das Repertoire der in der Bredenbeck-Gruppe untersuchten Proteine in Richtung der Rolle von VET und niederfrequenter Moden (engl. low-frequency modes, LFM) in Enzymen erweitern.

Sensoren für Schwingungsenergietransfer (Teil II)

Für die spektroskopische Untersuchung von VET werden zwei Funktionalitäten benötigt: ein VET-Donor, der die Energie eines absorbierten Photons in Schwingungsenergie umwandelt und diese ortsspezifisch in das Protein leitet, und ein VET-Sensor, der das Eintreffen von Schwingungsenergie durch eine Änderung seiner IR-Absorption melden kann. Das ursprünglich veröffentlichte Paar von nichtkanonischen Aminosäuren (engl. non-canonical amino acids, ncAAs), die diese Funktionen übernehmen, sind β-(1-Azulenyl)-L-alanin (AzAla) bzw. Azido-L-Homoalanin (Aha) [29]. Dieser VET-Sensor kann als Methioninanalogon in Abwesenheit von Methionin unter Selektionsdruck in Proteine eingebaut werden (engl. selective pressure incorporation, SPI), jedoch erfordert diese Methode eine Proteinsequenz ohne Methionin [30], was eine große Einschränkung darstellt. Aha ist zwar hinsichtlich der chemischen und sterischen Eigenschaften ein guter Ersatz speziell für Methionin, aber nicht für die meisten anderen Aminosäuren, an deren Stelle es in Proteine eingebaut werden könnte.

In Teil II dieser Arbeit, deren Ergebnisse fachlich begutachtet und veröffentlicht wurden [136], wurde die Verwendbarkeit mehrerer neuer ncAAs als VET-Sensoren untersucht. Diese ncAAs enthielten Azid- oder Nitrilgruppen, deren Streckschwingungen infrarotes Licht bei 2120 cm^{-1} bzw. 2220 cm^{-1} absorbieren. Die Amplitude des VET-Signals hängt linear vom Extinktionskoeffizienten ab, der für Nitrile typischerweise 2–4-fach niedriger ist als für Azide [103, 170]. Das VET-Signal entsteht jedoch durch die Verschiebung einer Absorptionsbande aufgrund der Kopplung an besetzte LFMs [24]. Die VET-Sensoren können daher nicht ausschließlich anhand des Extinktionskoeffizienten beurteilt werden. Wenn sich zwei Banden gleicher Amplitude gleich stark verschieben, ist das Differenzsignal für die schmalere Bande größer. Bei kleinen Verschiebungen, wie es bei VET-Experimenten der Fall ist, steigt die Signalamplitude linear mit der relativen Verschiebung, dem Verhältnis zwischen der Verschiebung und der Halbwertsbreite (engl. full width at half maximum, FWHM) der Bande, an. Aufgrund der nicht-Boltzmann'schen Besetzungsverteilungen, die durch den VET verursacht werden, ist die VET-Verschiebung nicht zwangsläufig mit der thermischen

Verschiebung identisch, die in der temperaturabhängigen Infrarotspektroskopie gemessen werden kann. Die genauen Besetzungen und Kopplungskonstanten sind nicht ohne weiteres bestimmbar, weshalb ein experimenteller Vergleich der VET-Sensoren unerlässlich ist.

Dipeptide. Die ncAAs Azidoalanin (N_3Ala), Aha, p-azido-Phenylalanin (N_3Phe), p-cyano-Phenylalanin (CNPhe) und 4-cyano-Tryptophan (CNTrp) wurden in Dipeptiden an AzAla gekoppelt. Ihre spektralen Eigenschaften wurden anhand der Fourier-Transform-Infrarot-Spektren (FTIR) in H_2O, Dimethylsulfoxid (DMSO) und Tetrahydrofuran (THF) verglichen. Neben den niedrigeren Extinktionskoeffizienten sind die Absorptionsbanden der Nitrile deutlich schmaler (FWHM 6–12 cm^{-1}) als die der Azidverbindungen (FWHM 19–34 cm^{-1}). Darüber hinaus wiesen die organischen Azide die schon in der Literatur beschriebenen Fermi-Resonanzen auf [171–173].

Das Lösungsmittel hat einen starken Einfluss auf die gemessenen VET-Signale. Für alle Dipeptide war die Signalamplitude in H_2O deutlich geringer als in den anderen Lösungsmitteln. Dies ging einher mit einer stark verspäteten Peakzeit (Zeitpunkt der maximalen Amplitude) in THF und einer weniger verspäteten Peakzeit in DMSO im Vergleich zu H_2O. Beide Effekte können durch die Wechselwirkung des Lösungsmittels mit den Dipeptiden erklärt werden. Diese führt zu einer schnellen Dissipation überschüssiger Energie in das Lösungsmittel, wodurch die maximale Verschiebung, die induziert werden kann, verringert und die Peakzeit zu kürzeren Verzögerungen verschoben wird. Dieser Effekt ist am stärksten in H_2O, das protisch und polar ist.

In den Untergruppen der Azid- bzw. Nitrilsensoren nahm die Peakzeit mit der Größe der Seitenkette zu. Dies zeigt die Abhängigkeit zwischen der Peakzeit und dem Abstand zwischen VET-Donor und Sensor. Die Peakzeit von N_3Phe ist um 0,9 ps später als die von CNPhe. Diese beiden ncAAs haben die gleiche Seitenkettenstruktur, sodass dieser Unterschied ausschließlich durch die unterschiedlichen Reportergruppen und deren Kopplung an die LFMs verursacht wird. Dieses Ergebnis ist genau das Gegenteil von dem, was man auf Grundlage der Schwingungslebensdauer erwarten könnte, welche für N_3Phe kürzer ist als für CNPhe.

Bei den VET-Experimenten der Dipeptide wurde beobachtet, dass die Signalamplituden der Nitrile und der Azide ähnlicher sind, als der deutlich kleinere Extinktionskoeffizient des Nitrils erwarten ließe. Der Effekt des Extinktionskoeffizienten wird dadurch kompensiert, dass die verschobene Bande schmaler ist. Um diesen Befund weiter zu untermauern, war es erforderlich, die räumliche Verteilung der angeregten Moleküle und ihren Beitrag zum IR-Differenzsignal zu bestimmen. Durch Berechnungen, die diese Einflüsse berücksichtigen, konnte bestimmt werden, dass 12% des FTIR-Signals zum gemessenen Differenzsignal beitragen. Hierdurch konnte eine Anpassung an das VET-Signal durchgeführt werden, um *absolute* Werte für die Verschiebung bei jeder Verzögerung zu erhalten. Zwei Ansätze für die Anpassung wurden verglichen: (i) die direkte Verschiebung der skalierten FTIR-Signale und (ii) die Anpassung des FTIR-Signals mit einem Voigt-Profil, dessen Fläche während der Verschiebung konstant bleibt, deren Form sich aber ändern kann. Eine bessere Übereinstimmung zwischen Daten und Anpassung ergibt sich bei Methode (ii). Deren zusätzliche Parameter können Änderungen der Bandbreite berücksichtigen, die mit der Verschiebung durch die Kopplung der LFMs einhergehen. Beide Methoden führen jedoch zu ähnlichen Verschiebungen und geben die Dynamik des VET-Signals, wie sie im Verlauf der integrierten Signale zu sehen ist, gut wieder. Mit dieser Analyse wurde bestätigt, dass die relative Verschiebung von CNPhe die der anderen ncAAs übersteigt.

Proteine. Anhand der Dipeptide konnte die allgemeine Eignung der ncAAs als VET-Sensoren nachgewiesen werden. Für den direkten Vergleich unter den Bedingungen zu-

künftiger VET-Experimente in Proteinen wurden Aha, N_3Phe und CNPhe zusätzlich an zwei verschiedenen Stellen in die PDZ-Proteindomäne eingebaut.

In diesem komplexeren System fiel die besondere Bedeutung der Azid-Fermi-Resonanzen auf. In der Proteinumgebung unterschieden sich die Bandenformen der Sensoren von denen in den Dipeptiden. Phe325Aha hatte eine sehr schmale und intensive IR-Absorptionsbande, während für N_3Phe an beiden Proteinstellen eine ausgeprägte Zweibanden-Struktur zu erkennen war. Dies macht die Bandenform und -intensität von Azidsensoren in einem Protein unvorhersehbar und erschwert den Vergleich der Signalamplituden. Bei CNPhe änderten sich die Bandenbreite und die Position im Protein leicht, aber es traten keine zusätzlichen Banden auf.

Die komplexen Bandenformen der FTIR-Signale übertrugen sich vollständig auf die VET-Signale. Alle Merkmale, die in den VET-Signalen zu sehen waren, konnten aus einem verschobenen Differenzsignal der FTIR-Signale rekonstruiert werden. Die Signalamplituden wurden jedoch von den außergewöhnlichen Bandenformen der Azidsensoren dominiert. Der Vergleich von Ile327Aha und den CNPhe-Varianten bestätigte jedoch auch im Protein, dass der niedrigere Extinktionskoeffizient von CNPhe nicht zu einem kleineren VET-Signal führt.

Weiterhin sind die Peakzeiten für größere Abstände zwischen VET-Donor und VET-Sensor größer, jedoch ist die Reihenfolge für die verschiedenen VET-Sensoren im Protein anders als in den Dipeptiden. Im Protein erreicht CNPhe als erstes, N_3Phe als zweites und Aha als letztes sein Maximum (mit einem Unterschied von 1,3 ps zwischen CNPhe und Aha). Bei den kleinen Dipeptiden wird die Dynamik weitgehend von der Wechselwirkung mit dem Lösungsmittel bestimmt, die im Protein weniger dominant ist. Andererseits kann eine große Seitenkette im Protein auch den effektiven VET-Abstand verkürzen und es gibt unzählige Wege zwischen dem VET-Donor und dem VET-Sensor. Der Einfluss all dieser Wege macht die VET-Experimente im Protein weniger anfällig für die einzelnen Schritte im VET-Prozess und erweitert die Charakterisierung der VET-Sensoren auf allgemeine experimentelle Bedingungen.

Insgesamt zeigen diese Experimente, dass Nitrilgruppen als VET-Sensoren dienen können, da der niedrigere Extinktionskoeffizient durch die schmalere Bandenbreite weitgehend kompensiert wird. Die neuen VET-Sensoren kombinieren den Einsatz vielfältigerer Inkorporationstechniken mit unterschiedlichen chemischen und sterischen Eigenschaften. Mit diesen Erkenntnissen kann der Einfluss verschiedener Seitenketten auf die Peakzeit abgeschätzt werden, sodass in einem einzigen Protein verschiedene Sensoren an verschiedenen Stellen eingesetzt werden können. Dadurch erhöht sich die Zahl der potenziellen Zielproteine, und der Sensoreinbau kann an den unterschiedlichen Stellen des Proteins mit weniger Störung der Struktur umgesetzt werden, was den Weg zur Untersuchung von VET in nativen Proteinen an jeder einzelnen Stelle ebnet..

Schwingungsenergietransfer und niederfrequente Moden in Enzymen (Teil III)

Ziel von Teil III dieser Arbeit war es, ein neues Proteinsystem zu etablieren, um VET und die Rolle der Proteindynamik im Kontext der enzymatischen Katalyse zu untersuchen. Da der VET-Donor AzAla die Energie eines Photons als Schwingungsenergie ortsspezifisch in ein Protein einspeisen kann, kann er auch zur gezielten Anregung von LFMs verwendet werden. Wenn die Bewegung dieser Moden an der enzymatischen Reaktion beteiligt ist,

wird ein direkter Einfluss auf die Aktivität erwartet. Diese Hypothese besteht bereits seit langem, konnte jedoch noch nicht endgültig verifiziert werden [21, 283, 284].

Cyclophilin A. Das erste Enzym, das hierfür ausgewählt wurde, war Cyclophilin A (CypA). Für diese gut untersuchte Peptidyl-Prolyl-cis-trans-Isomerase (PPIase) geben verschiedene Studien Hinweise darauf, dass die Proteindynamik von Relevanz für die Katalyse ist [246, 250, 285]. Der CypA Wildtyp (WT) wurde erfolgreich exprimiert und gereinigt. Es wurde ein Chymotrypsin-gekoppeltes Enzym-Assay durchgeführt, mit dem die chemische Aktivität von CypA nachgewiesen werden konnte. Basierend auf früheren Veröffentlichungen wurden mehrere Peptidliganden entworfen, die als Gerüst dienen sollten, um einen ncAA-VET-Sensor in das aktive Zentrum des Proteins zu bringen. Jedoch konnten die zuvor berichteten hohen Affinitäten zwischen CypA und den Peptiden nicht bestätigt werden.

Die ersten lokalen Versuche des kotranslationalen AzAla-Einbaus wurden unter Verwendung eines neu entwickelten aminoacyl-tRNA-Synthetase/tRNA-Paares durchgeführt, um CypA-Mutanten zu exprimieren, die AzAla an verschiedenen Stellen enthalten. Mit den damals zur Verfügung stehenden Methoden und Materialien konnte dieses Ziel nicht erreicht werden. Lediglich eine einzelne CypA-Mutante (Gly14AzAla) wurde vom Autor in den Laboren von N. Budisa an der Technischen Universität Berlin erfolgreich exprimiert. Diese Mutante wies jedoch keine enzymatische Aktivität auf. Aufgrund verschiedener Schwierigkeiten wurde dieses Projekt letztlich zugunsten eines anderen Enzymsystems eingestellt.

Formiatdehydrogenase. Das Enzym das anschließend ausgewählt wurde, war Formiatdehydrogenase (FDH). Es war das erste Enzym, bei dem LFMs im aktiven Zentrum mithilfe der zweidimensionalen IR-Spektroskopie nachgewiesen werden konnten [210]. Verschiedene Studien haben außerdem Beweise für die Beteiligung von LFMs an der Katalyse in FDH gefunden [167, 219, 225, 262, 263]. Dieses spezifische System ermöglicht die Verwendung eines natürlichen VET-Sensors. FDH bildet einen stabilen Komplex mit NAD^+ und N_3^-, einem hervorragende IR-Marker. Es bot damit die Möglichkeit, niedermolekulare nicht-kovalente Liganden als VET-Sensoren zu testen.

Eine wichtige Verbesserung, die dieses Projekt ermöglichte, war die Versorgung mit AzAla. Das örtlich verwendete AzAla wurde ursprünglich in der Gruppe von N. Budisa in Berlin synthetisiert. Im Jahr 2020 wurde jedoch eine einfache enzymatische Synthesemethode veröffentlicht [135]. Im Rahmen dieser Arbeit wurde diese Methode vor Ort etabliert sodass bis zum Start des FDH-Projekts AzAla in großen Mengen und mit gleichbleibender Qualität hergestellt werden konnte. Dies ermöglichte eine konstante Versorgung mit AzAla während der Expressionstests und für groß angelegte Expressionen.

Unter diesen Umständen wurde AzAla erfolgreich an mehreren Stellen in FDH eingebaut. Die Zirkulardichroismus-Spektren (CD) dieser Mutanten wurden mit dem Spektrum von FDH WT verglichen. Nur die Spektren von zwei Mutanten (Tyr196AzAla und Tyr341AzAla) stimmten mit dem Spektrum von FDH WT überein. Um die genauen strukturellen Gründe dafür zu ermitteln, sind die CD-Spektren nicht aussagekräftig genug. Die wahrscheinlichste Ursache ist ein nicht reproduzierbarer Effekt während der Proteinreinigung. Bei der Reinigung ist ein Entfaltungsschritt vorgesehen, der erforderlich ist, um das verkürzte Protein aus der Probe zu entfernen (das ansonsten aufgrund der FDH-Dimerisierung in der Probe verbleibt). Bei der Verwendung von Expressionsstämmen ohne Release-Faktor 1 (RF1) könnte dieser Schritt jedoch überflüssig sein, sodass eine bessere Ausbeute funktionaler Mutanten möglich wäre.

Die FTIR-Spektren zeigten, dass nur die Varianten deren CD-Spektren mit dem des WT übereinstimmten N_3^- binden konnten. Trotz des spektralen Überlapps von freiem und gebundenem N_3^- (2048 cm^{-1} bzw. 2050 cm^{-1}) konnte letzteres durch seine dreifach schmalere FWHM von 9 cm^{-1} nachgewiesen werden.

Für die meisten nicht bindenden Varianten lag die maximal erreichbare Proteinkonzentration deutlich unter der von FDH WT, Tyr196AzAla und Tyr341AzAla, welche bis zu 2 mM aufkonzentriert werden konnten. 1 M Trehalose wurde dem Puffer hinzugefügt, um zu testen, ob dies die Löslichkeit erhöhen könnte [274], ohne die Experimente zu stören. Da dies der Fall war, sollte diese Maßnahme in Zukunft bei diesem Projekt in Betracht gezogen werden.

Deutliche VET-Signale wurden für Tyr196AzAla und Tyr341AzAla beobachtet. In Übereinstimmung mit dem um 3 Å verlängerten Abstand zwischen AzAla und gebundenem N_3^- betragen die Peakzeiten 9,8 ps und 9 ps für Tyr196AzAla bzw. Tyr341AzAla. Es konnte nachgewiesen werden, dass die beobachteten VET-Signale ausschließlich durch das gebundene N_3^- verursacht wurden. Im Puffer mit 1 M Trehalose war die Peakzeit von Tyr196AzAla auf 8,5 ps verringert. Wenn der Trehalosepuffer verwendet wird, müssen daher die veränderten Wechselwirkungen mit dem Lösungsmittel berücksichtigt werden. Für eine optimale Vergleichbarkeit sollten alle Experimente mit der gleichen Pufferzusammensetzung durchgeführt werden, da ihr Einfluss auf die VET-Dynamik nicht vernachlässigbar ist.

Die Aktivität von FDH WT und vier AzAla-Mutanten wurde unter sättigenden Substratbedingungen ohne und mit Anregung von AzAla untersucht. In diesen Experimenten wurde eine leichte Verringerung der Aktivität unter Beleuchtung sogar für den WT beobachtet. Bisher kann ein Unterschied der Probentemperatur als Grund für diesen Rückgang nicht ausgeschlossen werden. Die Ergebnisse wurden durch den Vergleich mit einem ähnlichen Experiment mit Dihydrofolat Reduktase und Gold Nanopartikeln relativiert, bei dem deutlich stärkere Anregungen möglich waren [278]. Möglicherweise kann man die Enzymaktivität auch mit dem derzeitigen Versuchsaufbau geringfügig (aber signifikant) erhöhen. Das erfordert aber wegen der erwarteten kleinen Effekte absolute Gewissheit über die Probentemperatur.

Die Experimente mit FDH verdeutlichen das Potenzial von niedermolekularen Liganden als VET-Sensoren. Die einfache Struktur, die Fermi-Resonanzen verhindert, und der hohe Extinktionskoeffizient machen N_3^- besonders attraktiv. Seine Nutzung kann viele Metalloproteine zu den potenziellen Zielen für VET-Experimente hinzufügen [286–289]. Mit diesen Erkenntnissen kommen auch andere VET-Sensor-Liganden für die Nutzung infrage, z.B. IR-markierte Flavine oder NAD^+ [98, 99].

Es wurden Hinweise für zukünftige Verbesserungen der Expressions- und Reinigungsstrategien von FDH gegeben, die zu einer zuverlässigeren Probenvorbereitung führen können. Neu entwickelte FDH AzAla-Mutanten werden es außerdem ermöglichen, VET-Pfade und Gruppen von dynamisch gekoppelten Aminosäuren zu identifizieren, indem VET aus verschiedenen Richtungen zum aktiven Zentrum Schritt für Schritt charakterisiert wird. Darüber hinaus könnte der optimierte FDH-Aktivitätsassay eine direkte Beschleunigung einer enzymatischen Reaktion zeigen, die durch die spezifische Anregung von Protein-

LFMs verursacht wird. Auf diese Weise kann dieses Projekt dazu beitragen, eine Antwort auf die seit langem bestehenden Fragen nach der außerordentlichen katalytischen Effizienz von Enzymen zu finden.

Part IV

Appendix

i Crystal Structures of PDZ Phe325N_3Phe and Phe325CNPhe

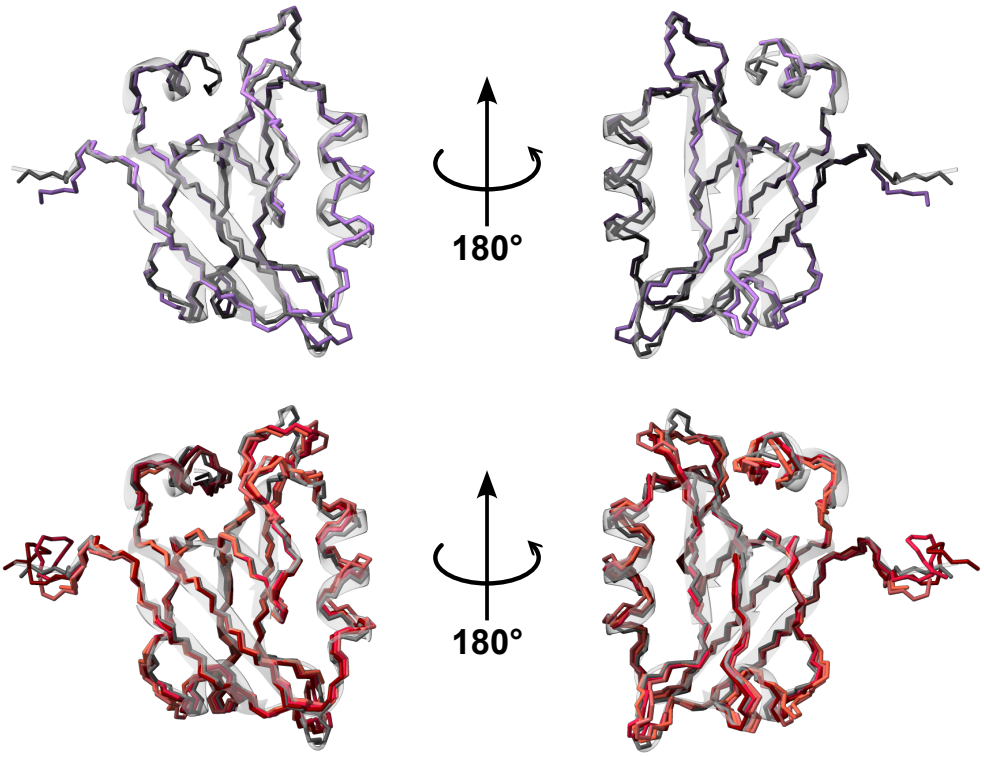

Figure A.1: *Top: Overlay of the backbone atoms of PDZ WT (gray) and PDZ Phe325N_3Phe (purple). Bottom: Overlay of the backbone atoms of PDZ WT (gray) and the four chains of PDZ Phe325CNPhe (red hues). An additional, semitransparent ribbon representation for the PDZ WT is shown for orientation. PDZ WT data are pdb entry 1BFE [176].*

The asymmetric unit of the Phe325N_3Phe crystal contains one PDZ molecule, the asymmetric unit of Phe325CNPhe contains four. The four chains of Phe325CNPhe are very similar to each other (RMSD $\leq$0.71 Å for each pair). The structures of both mutants reveal that the mutation does not alter the overall fold of the protein. Comparison with the WT structure (Fig. A.1) yields RMSD of $\leq$0.84 Å and 0.68 Å for CNPhe and N_3Phe, respectively. The largest differences between the WT and the mutants as well as the different chains of Phe325CNPhe occur at the N-terminus, which is too flexible to model the first few residues, and in the loops around residue Ser320 and Asp332. These loop regions are less rigid (reflected also in the associated B-factors) and are partially involved in crystal contacts.

The electron density of the ncAAs is well-defined. Interestingly, a model of Phe325N_3Phe with a C-N bond length of 1.4 Å [290] results in a higher electron density close to the phenyl ring than observed (Fig. A.2 a). The observed electron density is better explained with an extraordinary long bond of 1.7 Å (Fig. A.2 b), a length that is also present in pdb entry

6HCS [291], or even longer bonds.[1] Within this work the reasonably good structure with a bond length of 1.4 is used for referencing distances.

The only clearly mutation-induced change occurs in the structure of Phe325N_3Phe. In this structure one can observe, that the αB helix shifts to accommodate the longer side chain of N_3Phe, whose space would otherwise be restricted by Leu379 (Fig. A.2 d). Judged by the crystallographic electron density both ncAAs occupy a single conformation.

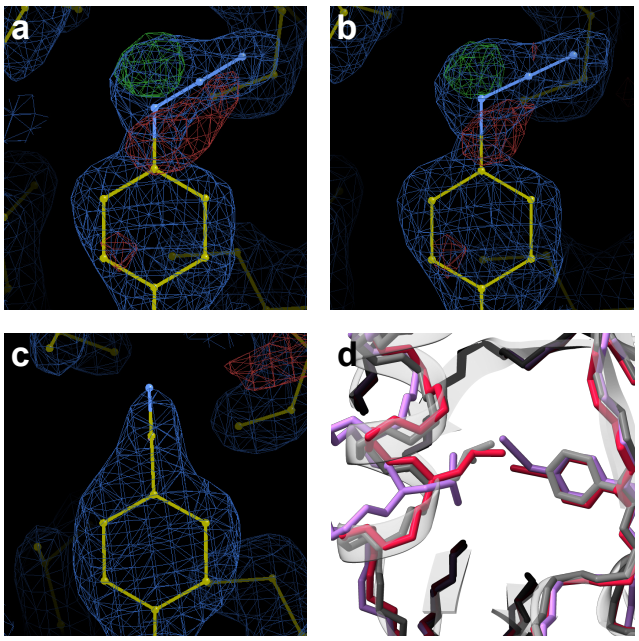

Figure A.2: *Electron density around Phe325N_3Phe with a typical C-N bond length (a). Electron density around Phe325N_3Phe with an increased C-N bond length of 1.7 Å(b). Electron density around Phe325CNPhe of chain A. The correctly modeled electron densities (blue mesh) are mapped at 1 rmsd, excessive electron density (red mesh) and electron density missing in the model (green mesh) are mapped at 3 rmsd (c). Comparison of the structure around the mutation site in an overlay as in Fig. A.1 (d). Side chain atoms are only shown for residue 325 and Leu379, which is clearly displaced in PDZ Phe325N_3Phe. PDZ Phe325CNPhe is represented only by chain A. For clarity objects in the background fade out.*

Table A.1: *Parameters of the crystal structures of PDZ Phe325N_3Phe and Phe325CNPhe.*

	Phe325N_3Phe	**Phe325CNPhe**
Space group	C 2 2 2_1	C 1 2 1
Resolution / Å	1.86	1.85 Å
Outer shell completeness / %	99.66	98.76
R-value free	0.2855	0.2775
R-value work	0.2474	0.2283

1 To the best of my knowledge, the ncAA N_3Phe is currently only rarely represented in the structures in the pdb, either with chemical ID 4II (one entry) or 4OO (two entries).

ii Derivation of $\Delta\tilde{\nu}_{min}$

Given a Lorentzian function $f(\tilde{\nu})$ centered around 0 with area a and FWHM w:

$$f(\tilde{\nu}) = \frac{2a}{\pi}\frac{w}{4\tilde{\nu}^2 + w^2} = c\frac{w}{4\tilde{\nu}^2 + w^2}$$

The difference signal $d(\tilde{\nu})$ generated by subtracting this Lorentzian from one shifted by $-\tilde{\nu}_s$ is:

$$d(\tilde{\nu}) = c\left(\frac{w}{4(\tilde{\nu}+\tilde{\nu}_s)^2 + w^2} - \frac{w}{4\tilde{\nu}_s^2 + w^2}\right)$$

The roots of the first derivative $^{\delta}/_{\delta\tilde{\nu}}\, d(\tilde{\nu})$ are at the positions $\tilde{\nu}_i$ of the extrema.

$$\frac{\delta}{\delta\tilde{\nu}}d(\tilde{\nu}) = c\left(\frac{8w\tilde{\nu}}{(4\tilde{\nu}^2 + w^2)^2} - \frac{8w(\tilde{\nu}+\tilde{\nu}_s)}{(4(\tilde{\nu}+\tilde{\nu}_s)^2 + w^2)^2}\right) \stackrel{!}{=} 0$$

$$\tilde{\nu}_{1/2} = \pm\sqrt{\frac{\sqrt{\tilde{\nu}_s^4 + \tilde{\nu}_s^2 w^2 + w^4}}{6} + \frac{\tilde{\nu}_s^2 - w^2}{12}} - \frac{\tilde{\nu}_s}{2}$$

$$\tilde{\nu}_{3/4} = \pm\sqrt{-\frac{\sqrt{\tilde{\nu}_s^4 + \tilde{\nu}_s^2 w^2 + w^4}}{6} + \frac{\tilde{\nu}_s^2 - w^2}{12}} - \frac{\tilde{\nu}_s}{2}$$

The separation of the real-valued solutions $\tilde{\nu}_{1/2}$ in the limit of small shifts $\Delta\tilde{\nu}_{min}$ is:

$$\begin{aligned}\Delta\tilde{\nu}_{min} &= \lim_{\tilde{\nu}_s \to 0}\sqrt{\frac{2\sqrt{\tilde{\nu}_s^4 + \tilde{\nu}_s^2 w^2 + w^4}}{3} + \frac{\tilde{\nu}_s^2 - w^2}{3}} \\ &= \frac{w}{\sqrt{3}}\end{aligned} \tag{10.1}$$

iii Supplementary Figures for Part II

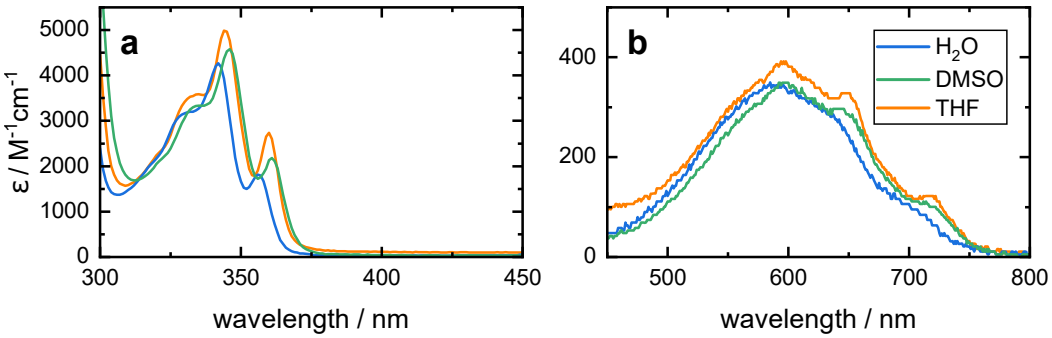

Figure A.3: *Extinction coefficient of AzAla in the wavelength range 300–450 nm (a) and 450–800 nm (b) in different solvents. The steps in b) originate from the mOD resolution of the data transferred via the serial port of the spectrometer.*

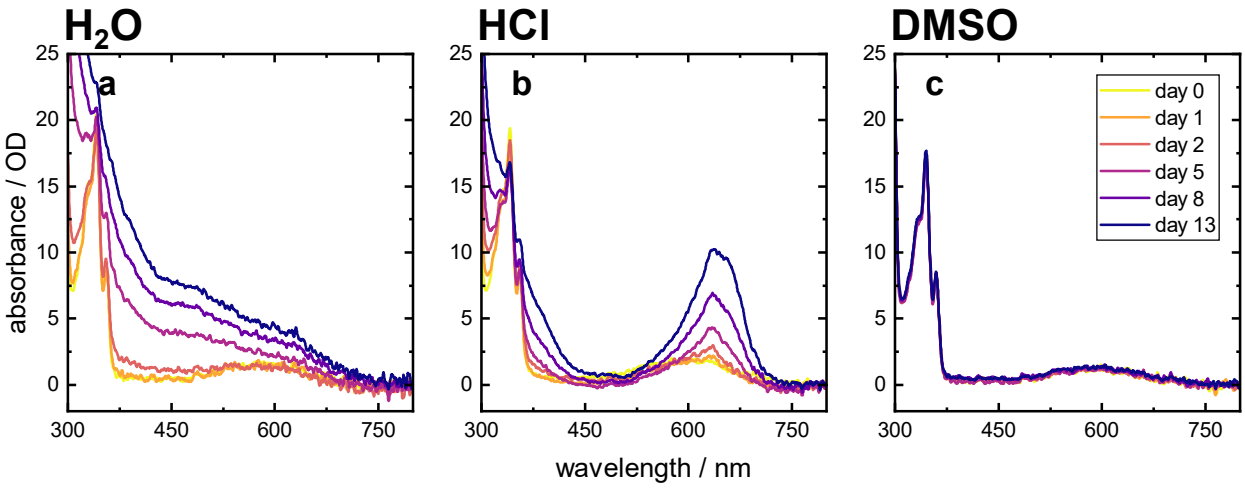

Figure A.4: *After prolonged periods in aqueous (a), or acidic (b) solutions, the UV/Vis spectrum of AzAla undergoes significant changes. This is not the case in DMSO (c) or basic solutions (not shown). These data have been recorded on a NanoDrop One spectrometer, the absorbance values were scaled to 1 cm path length using the internal auto-ranging function.*

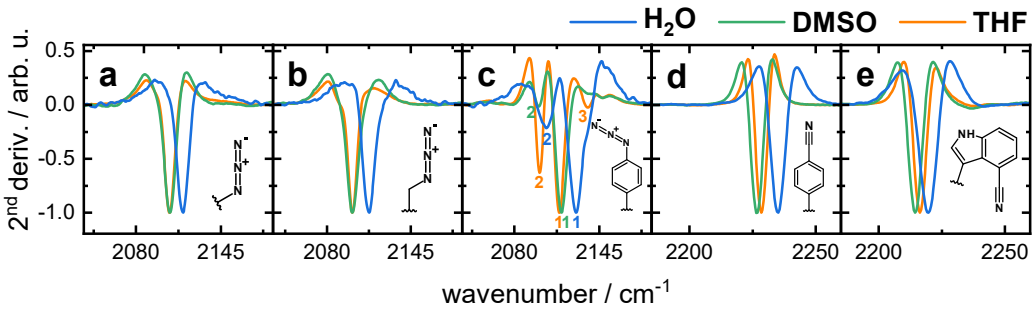

Figure A.5: *Normalized second derivative of the AzAla-VET sensor dipeptide FTIR spectra. Small numbers in the graph of AzAla-N_3Phe (c) indicate the presence of multiple subbands. The structure in each panel indicates the VET sensor of the data in the respective panel.*

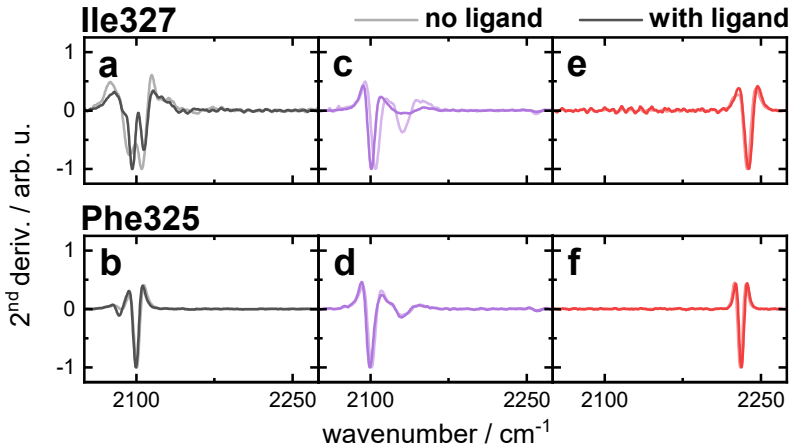

Figure A.6: *Normalized second derivative of the FTIR spectra of PDZ with Aha (gray), N_3Phe (purple), and CNPhe (red) inserted at positions Ile327 (top) and Phe325 (bottom). Each panel shows the second derivative of the spectrum without the ligand (dim colors) and with the ligand (intense colors).*

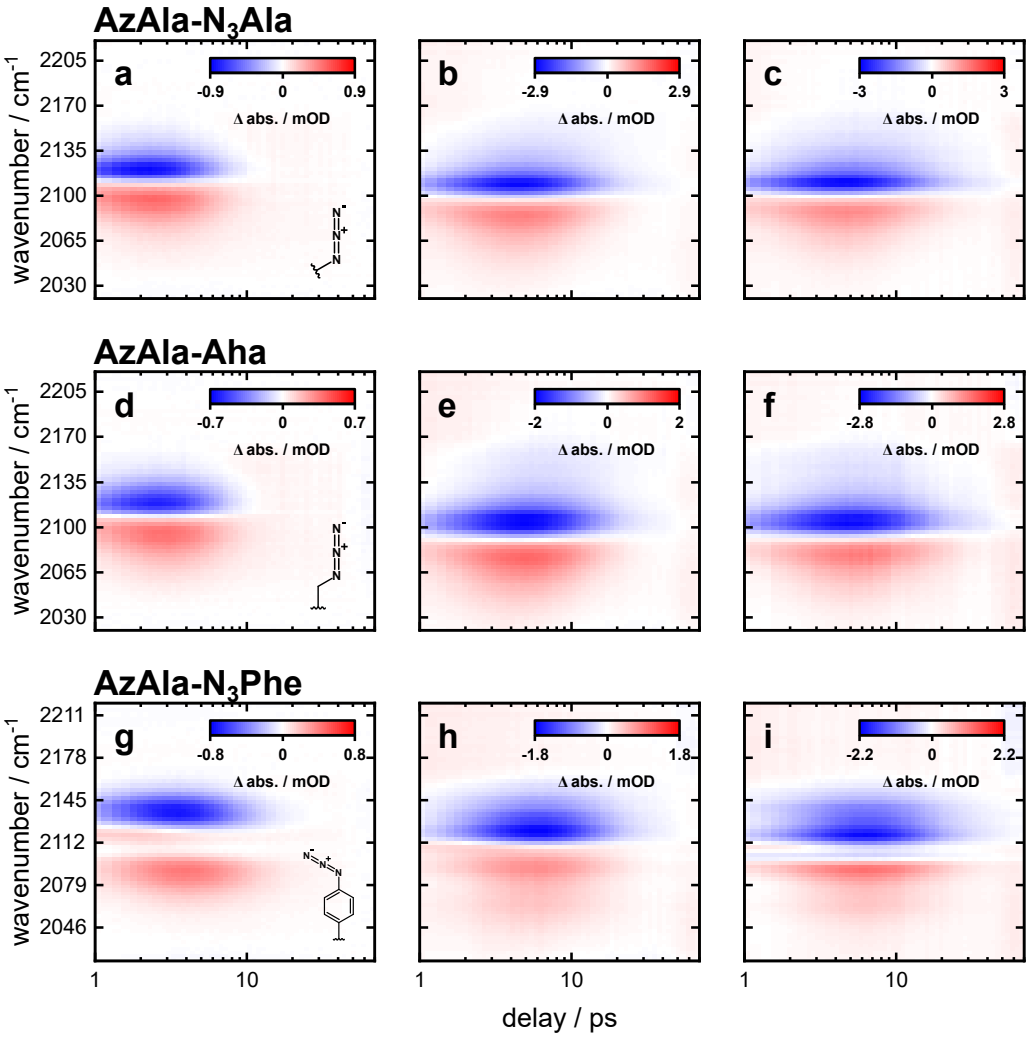

Figure A.7: *VET data of the AzAla-AzAla-N_3Ala (a–c), AzAla-Aha (d–f), and AzAla-N_3Phe (g–i) dipeptides in H_2O (left), DMSO (middle column), and THF (right).*

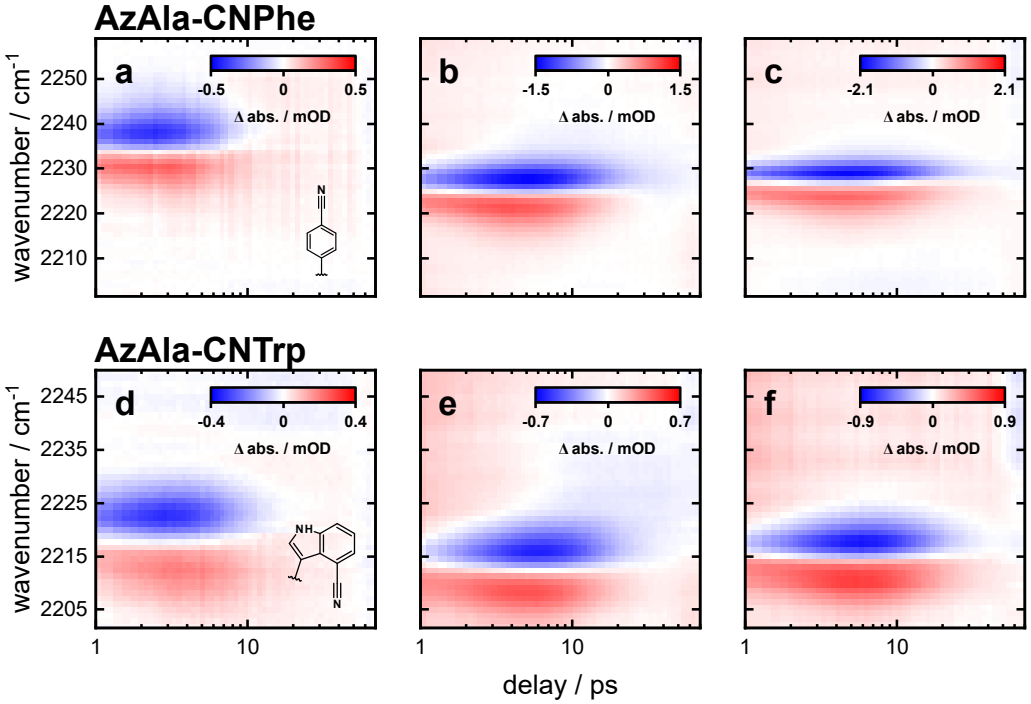

Figure A.8: *VET data of the AzAla-CNPhe (a–c), and AzAla-CNTrp (d–f) dipeptides in H_2O (left), DMSO (middle column), and THF (right).*

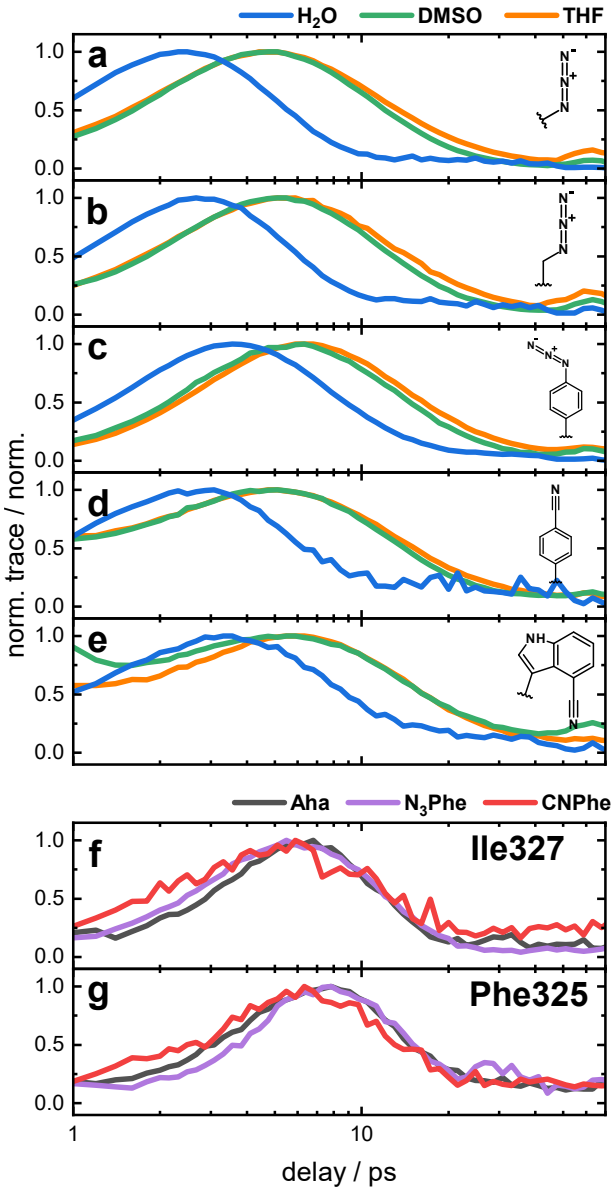

Figure A.9: *Normalized VET traces of the AzAla-VET sensor dipeptides in H_2O, DMSO, and THF (a–e). The side chain structure of the respective VET sensor is shown within the panels. Normalized VET traces of the PDZ variants with Aha, N_3Phe, and CNPhe at the positions indicated within the panels (f,g). Taken from [136], licensed under CC BY.*

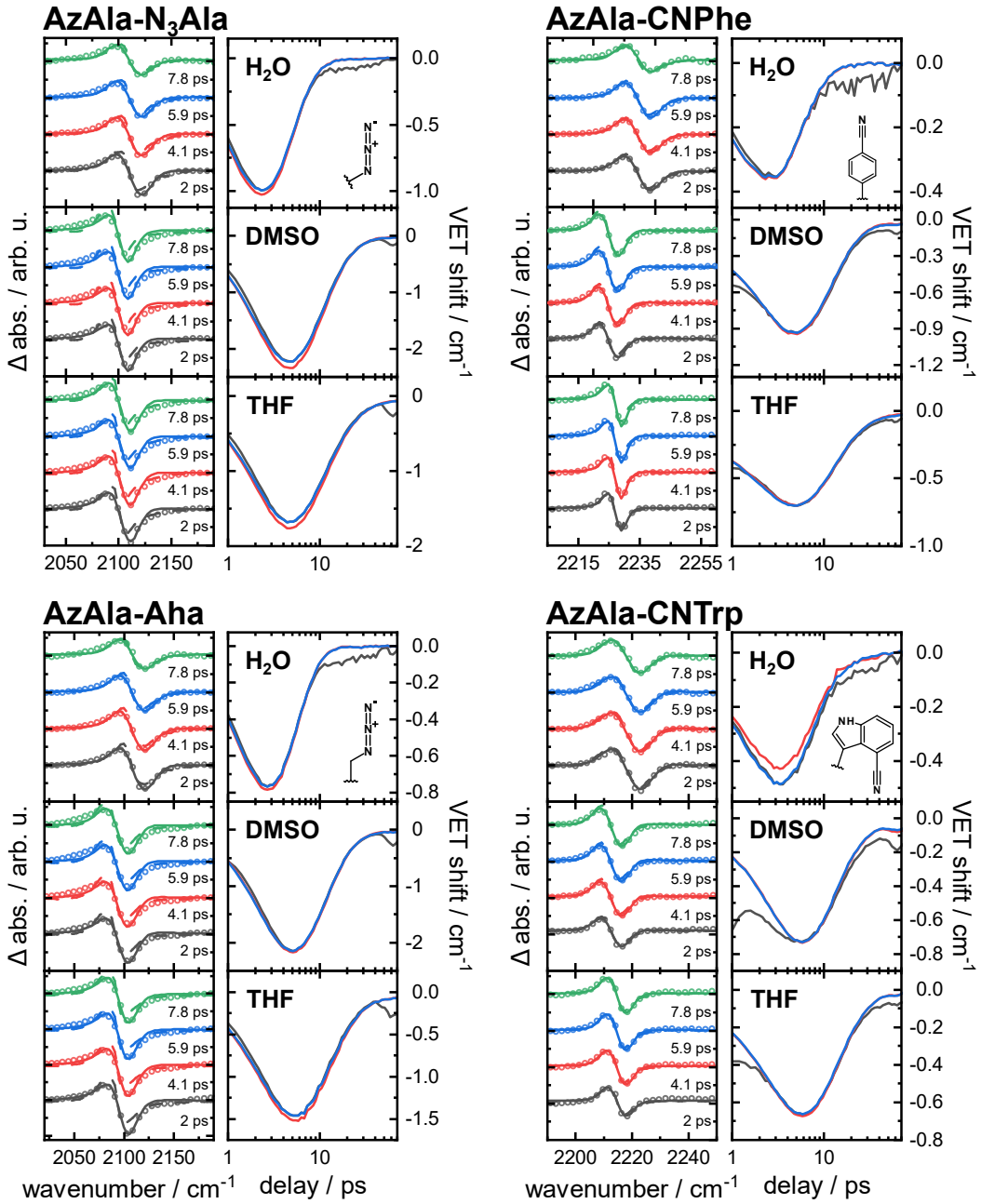

Figure A.10: *Quantitative analysis of the VET induced shifts of AzAla-N_3Ala (top left), AzAla-Aha (bottom left), AzAla-CNPhe (top right), and AzAla-CNTrp (bottom right). The left column in each group shows the normalized and offset VET data at the indicated delays (circles), the fit to the data using the FTIR spectrum (dashed lines), and the fit to the data using a shifting Voigt profile (solid lines). The right column of each group shows the fitted absolute shifts using the FTIR spectrum (blue), the fitted absolute shifts using a Voigt profile (red; see text for further information), and the trace calculated from the VET data scaled to match the scale (grey).*

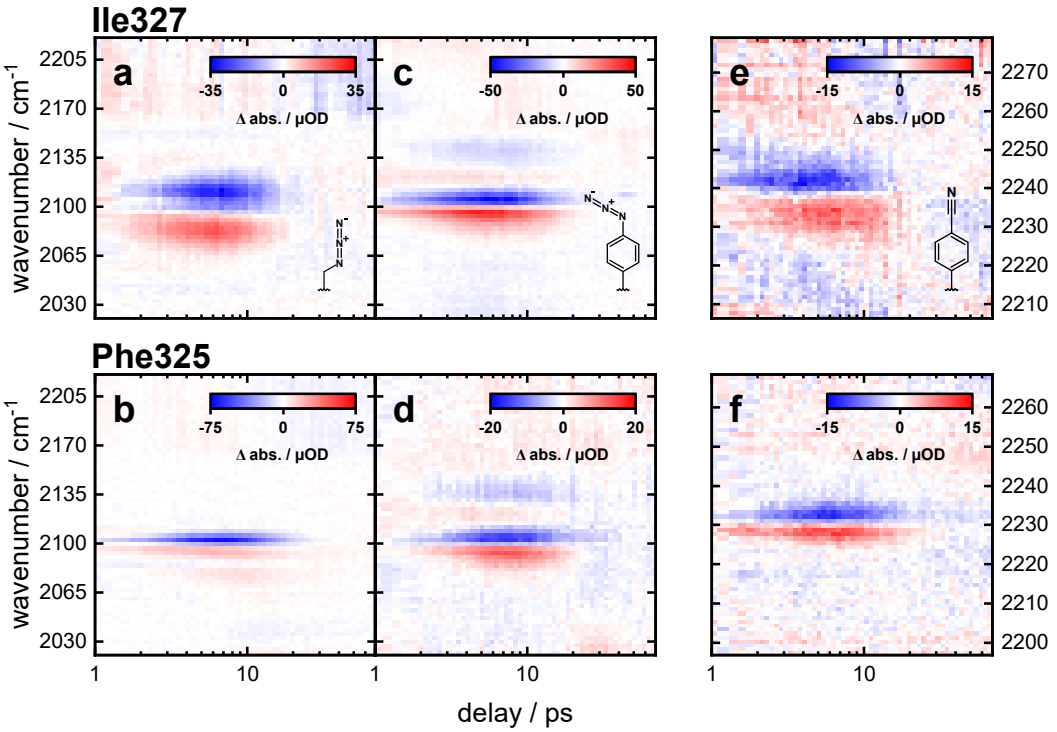

Figure A.11: *VET data measured in PDZ with VET sensors inserted at position Ile327 (top) or at position Phe325 (bottom). The structures (top) indicate the VET sensor within each column.*

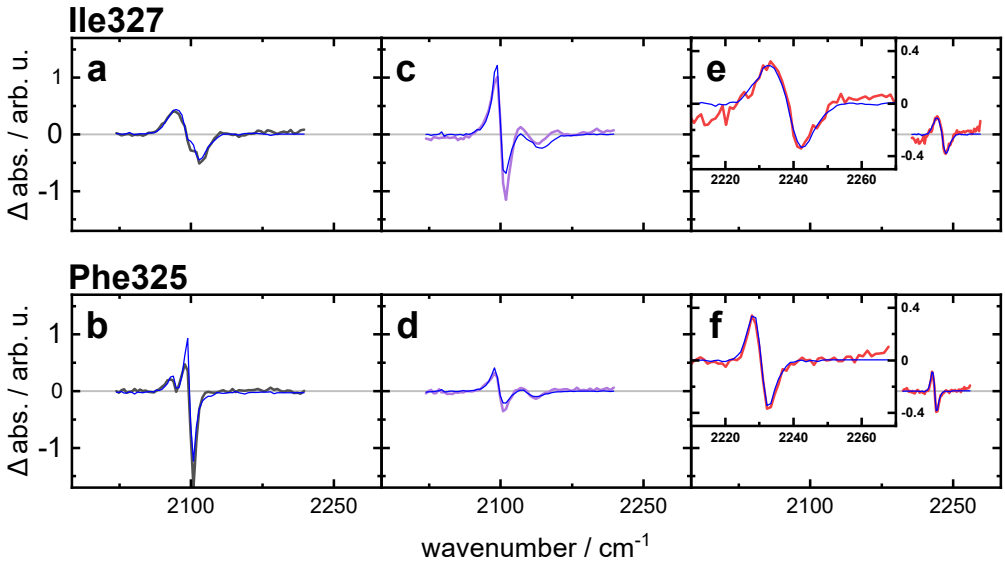

Figure A.12: *The shape and features of the VET signals (as shown in Fig. 4.7) are well reproduced by the difference signal of the FTIR spectrum and a shifted FTIR spectrum (blue lines). Adapted from [136], licensed under CC BY.*

iv Supplementary Figures for Part III

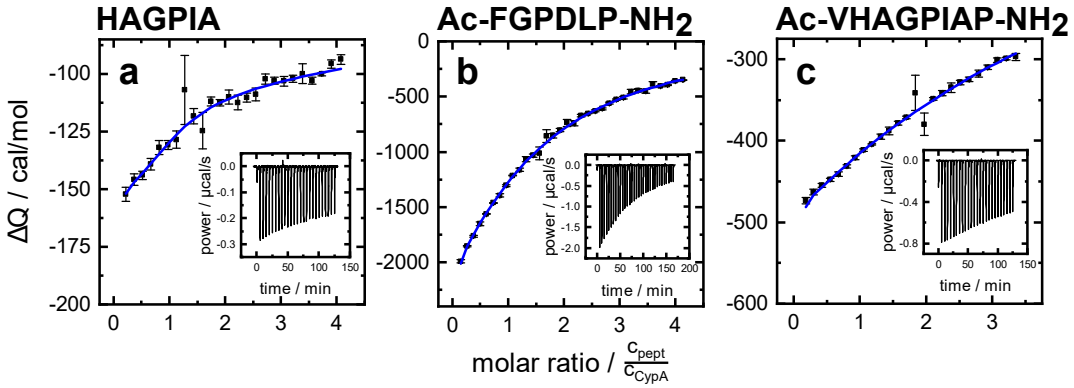

Figure A.13: *ITC data of CypA with the ligand peptides HAGPIA (a), Ac-VGPDLP-NH_2 (b), and Ac-VHAGPIAP-NH_2 (c). Each panel shows the integrated heat after each injection (black squares), and a fit to the data (blue line) as a function of the ratio between the concentrations of CypA and the peptide. The insets show the power during each measurement for reference.*

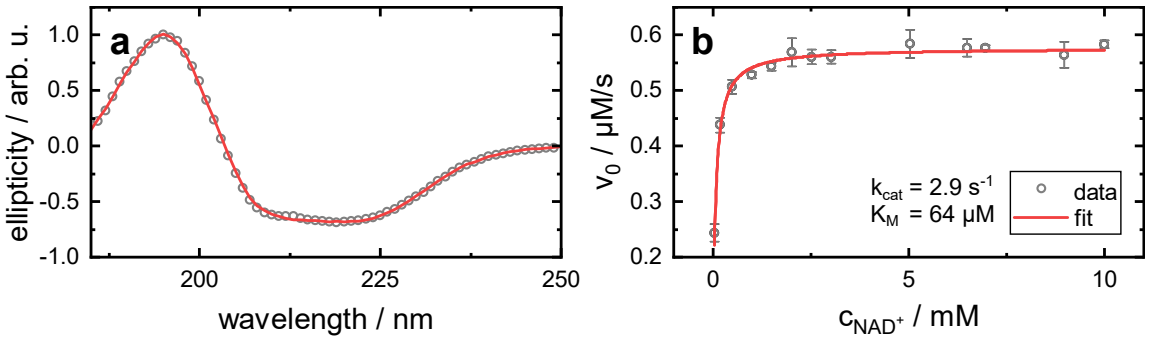

Figure A.14: *Normalized CD spectrum of FDH WT and fit of this spectrum for secondary structure determination (a). The fit was conducted using the BeStSel web server [88, 272]. Michaelis-Menten plot and fit of the initial reaction velocity v_0 at varying NAD^+ concentrations with 200 nM FDH present in the reaction mixture (b). The data points are the mean and the error bars the standard deviation of three independent measurements.*

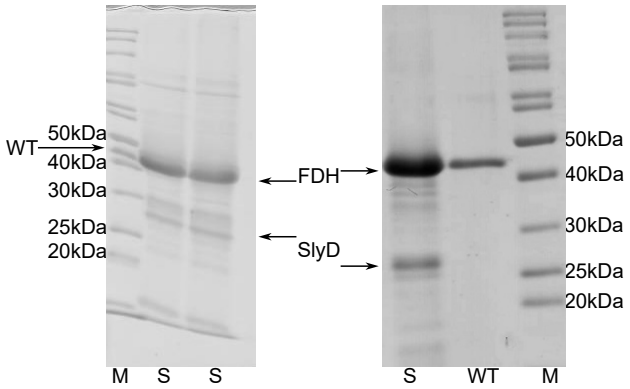

Figure A.15: *SDS gels of FDH Tyr196AzAla (left) and Tyr196AzAla+VG (right). Both samples show comparable degrees of impurities, when excessive amounts of sample are applied to the gel. The lanes contain marker (M), FDH WT for reference (WT), and mutant sample (S). Marker bands are labeled with the corresponding protein mass. The WT reference in the Gel of Tyr196AzAla was added to the marker lane.*

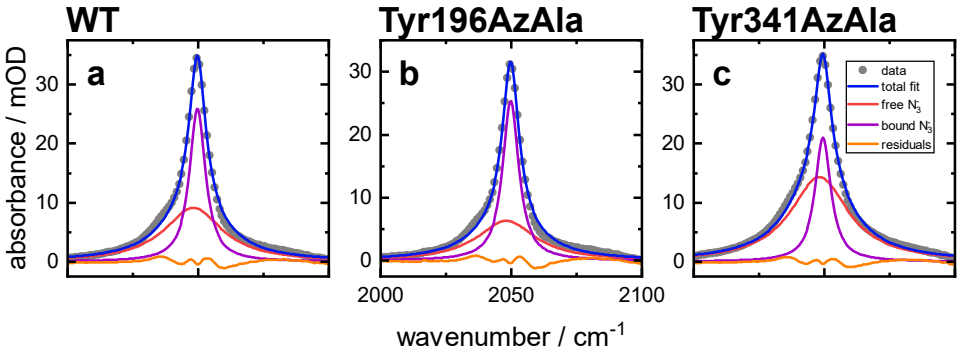

Figure A.16: *FTIR spectra of FDH WT (a), FDH Tyr196AzAla (b), and FDH Tyr341AzAla (c) in complex with NAD^+ and N_3^-. Using one Lorentzian with the characteristics of free N_3^- and one unconstrained results in underfitting of the flanks. Using two independent Lorentzians result in a shift of -10cm^{-1} of the free species to account for the shoulder in the red flank (not shown).*

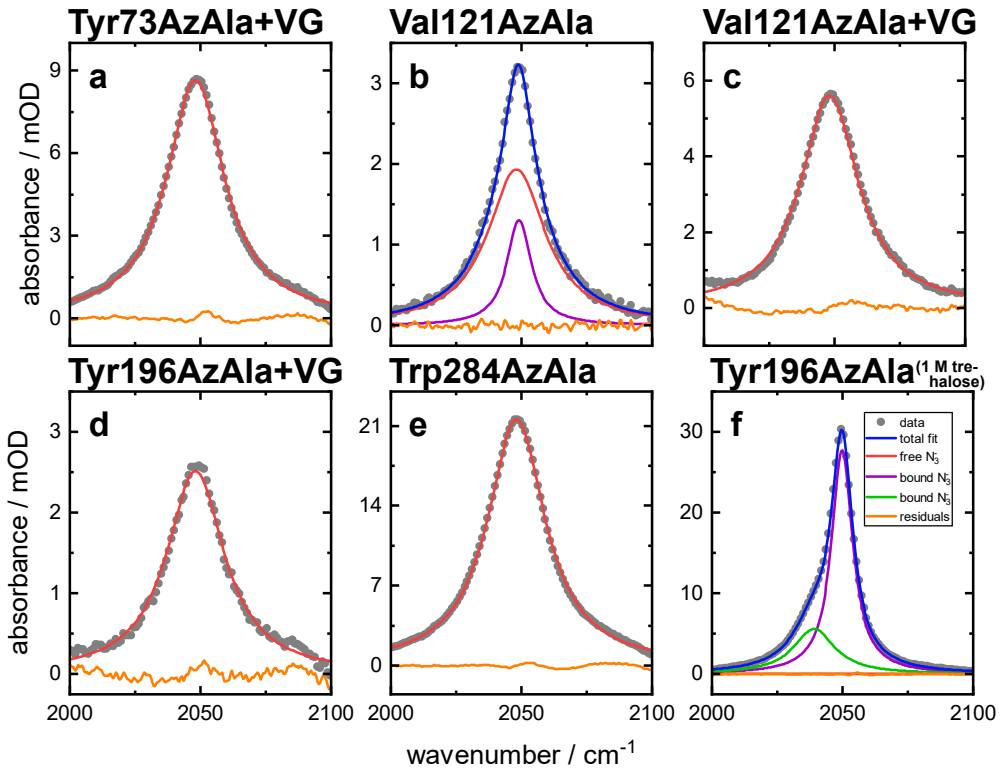

Figure A.17: *FTIR spectra of N_3^- with FDH Tyr73AzAla+VG (a), Val121AzAla (b), Val121AzAla+VG (c), Tyr196AzAla+VG (d), Trp284AzAla (e). Only Val121AzAla weakly binds a fraction of the N_3^- (adding a third small Lorentzian only minimally improves the fit). In none of these spectra the side feature is visible. The addition of 1 M trehalose to the sample buffer does not interfere with N_3^- binding in FDH Tyr196AzAla (f).*

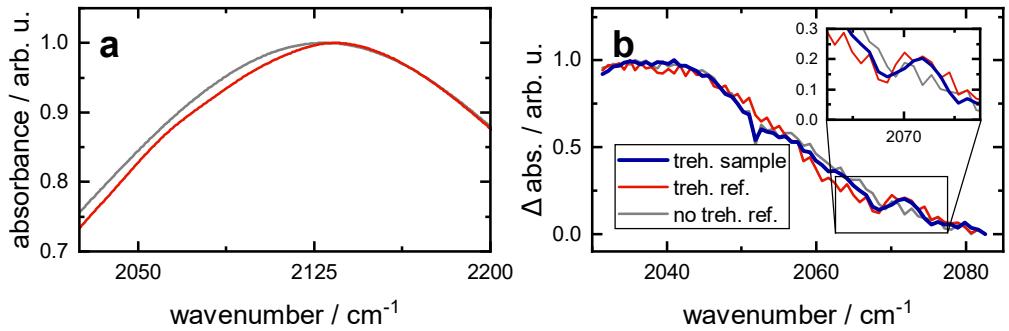

Figure A.18: *Normalized FTIR (a) and VET (b) data of samples without and with 1 M trehalose. All data are of the reference compartment or the sample compartment of FDH Tyr196AzAla samples. The VET data at 120 ps delay are shown in b, and the inset is focused on the additional feature in the solvent heating signal due to the addition of 1 M trehalose.*

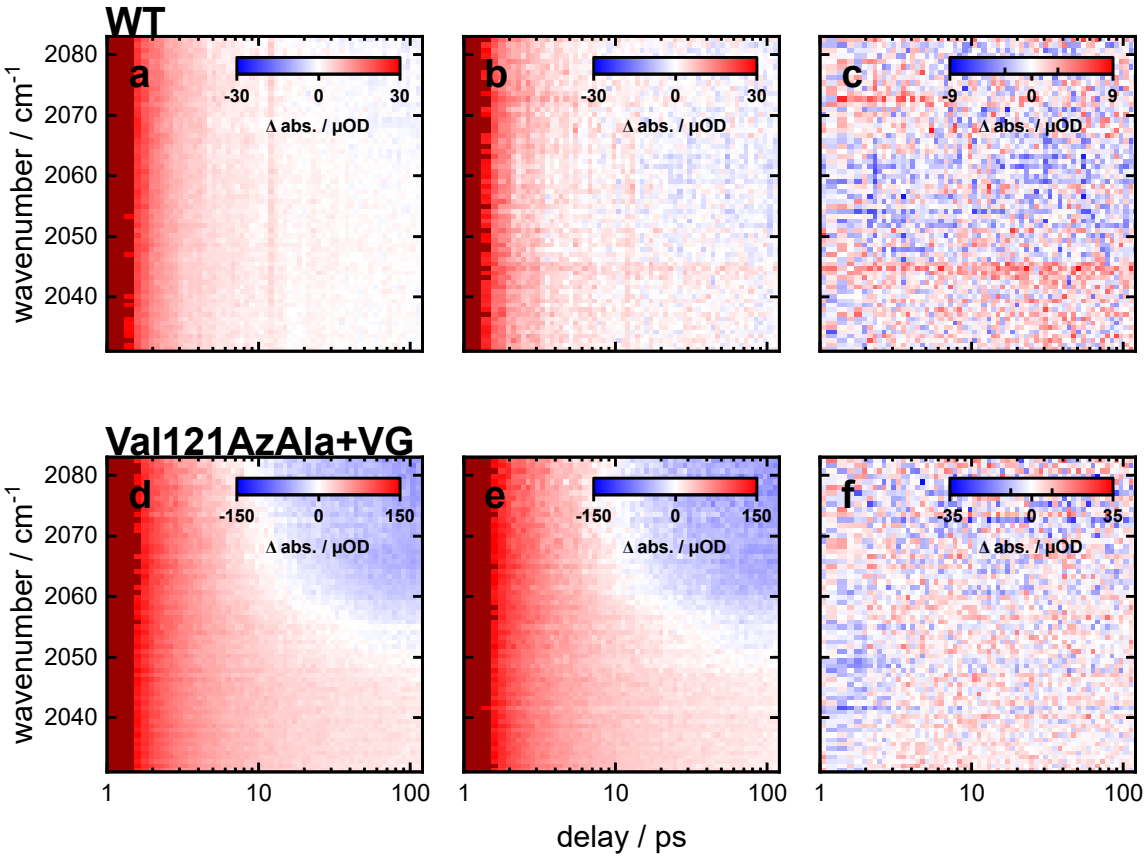

Figure A.19: *VET data of FDH WT (a–c) and FDH Val121AzAla+VG (d–f). The VET data of the reference compartment without N_3^- (left), the sample compartment with N_3^- (center), and the results of correcting the sample compartment with the data of the reference compartment (right). The dynamic range of the color bars are adapted to show the water signal, higher intensities (at early delays) are uniformly dark red. The inside ticks in the color bars of panels c and f contain 80% of the data.*

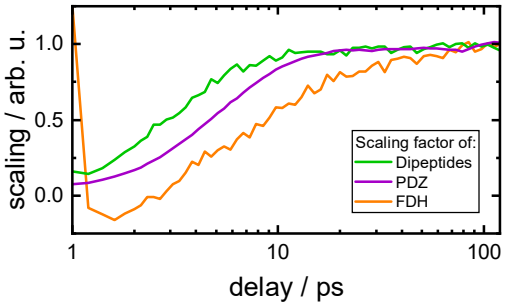

Figure A.20: *Solvent scaling factor for dipeptides, PDZ, and FDH. The scaling parameters were averaged over all aqueous dipeptide experiments, the WT compartments of PDZ mutants, and the reference compartments of all FDH experiments without trehalose.*

v Supplementary Tables

FTIR Band Parameters for Part II

Table A.2: *Parameters determined from the FTIR spectra of the AzAla-ncAA dipeptides in different solvents ($\tilde{\nu}_{max}$: wavenumber of the maximum; ε_{max}: extinction coefficient at the maximum; B: integrated extinction coefficient).*

		N_3Ala	Aha	N_3Phe	CNPhe	CNTrp
H_2O	$\tilde{\nu}_{max}$ / cm^{-1}	2115.7	2113.1	2026.9	2235	2219.3
	FWHM / cm^{-1}	28.6	31.4	34.5	10.2	12.2
	ε_{max} / $M^{-1}cm^{-1}$	417	405	388	224	109
	B / $mM^{-1}cm^{-2}$	14.0	15.0	13.7	2.6	1.4
DMSO	$\tilde{\nu}_{max}$ / cm^{-1}	2106	2100	2116.5	2226.7	2214.4
	FWHM / cm^{-1}	23.9	27.4	23.4	7.3	8.6
	ε_{max} / $M^{-1}cm^{-1}$	388	367	470	187	146
	B / $mM^{-1}cm^{-2}$	11.7	11.6	12.8	1.6	1.4
THF	$\tilde{\nu}_{max}$ / cm^{-1}	2105.7	2099.7	2114.4	2228.4	2216.4
	FWHM / cm^{-1}	24.6	27.8	28.7	5.6	8.0
	ε_{max} / $M^{-1}cm^{-1}$	647	673	663	261	168
	B / $mM^{-1}cm^{-2}$	19.1	23.1	20.0	1.9	2.1

Table A.3: *Parameters determined from the FTIR spectra of the PDZ variants without (-) and with (+) AzAla-KQTSV ligand. Only the spectrum of Phe325N_3Phe was fit with a Voigt profile, the other spectra were fit with Gaussians (FWHM: full width at half maximum; B: integrated extinction coefficient; $\tilde{\nu}_{max}$: wavenumber of the maximum).*

			PDZ Ile327			PDZ Phe325		
			$\tilde{\nu}_{max}$ / cm^{-1}	FWHM / cm^{-1}	B / $mM^{-1}cm^{-2}$	$\tilde{\nu}_{max}$ / cm^{-1}	FWHM / cm^{-1}	B / $mM^{-1}cm^{-2}$
Aha	-	Peak 1	2099.0	29.6	11.7	2088.1	16.3	2.2
		Peak 2				2100.0	11.5	9.7
	+	Peak 1	2097.2	25.6	8.6	2085.9	16.1	2.7
		Peak 2	2110.5	20.8	1.8	2099.4	9.9	8.3
N_3Phe	-	Peak 1	2105.2	20.5	7.2	2101.0	12.9	4.1
		Peak 2	2128.3	29.4	10.2	2121.0	45.0	10.1
	+	Peak 1	2101.2	12.7	5.3	2100.4	15.9	5.9
		Peak 2	2121.3	44.8	12.9	2125.1	33.3	6.8
CNPhe	-		2236.7	11.6	1.6	2230.3	5.8	0.8
	+		2238.1	11.4	1.4	2231.5	5.8	0.8
SCN	-		2083.4	12.6	0.5			
	+		2082.0	7.9	0.3			

CD Secondary Structure of FDH WT

Table A.4: *Secondary structure analysis of FDH based on the CD spectrum and on the crystal structure (pdb: 5DN9 [167]). The analysis was conducted using the BeStSel web server [88, 272].*

	helix / %	antiparallel β-sheet / %	parallel β-sheet / %	turn / %	other / %
CD spectrum	26.7	3.6	15.4	11.1	43.2
structure	30.4	1.5	14.7	13.8	39.6

Experimental Conditions of FDH

Table A.5: *Experimental conditions of the FTIR and VET experiments with FDH.*

variant	FDH concentration / mM	N_3^- concentration / mM	thickness / µm	pulse energy / µJ
Tyr196AzAla	1.6	1.2	150	19
Tyr196AzAla (1 M trehalose)	1.6	1.5	100	16
Tyr341AzAla	1.7	1.8	100	20
WT	1.5	1.4	100	18
Y73AzAla+VG	1.1	1.1	150	22
Val121AzAla	0.3	0.3	150	22
Val121AzAla+VG	0.7	0.6	100	20
Tyr196AzAla+VG	0.3	0.3	150	19
Trp284AzAla	0.6	2.3	150	17

FTIR Band Parameters of N_3^- in FDH

Table A.6: *Band shape parameters, rounded to the first decimal place, of the Lorentzians used to fit the FTIR absorption of N_3^- bound to the FDH mutants, and the average thereof. The indices 1 and 2 indicate the main feature and the side feature (not detected for Val121AzAla), respectively.*

variant	$\widetilde{\nu}_1$ / cm^{-1}	$FWHM_1$ / cm^{-1}	$\widetilde{\nu}_2$ / cm^{-1}	$FWHM_2$ / cm^{-1}
WT	2049.7	9.0	2037.3	15.4
Val121AzAla	2049.1	13.1		
Tyr196AzAla	2049.8	8.5	2038.2	14.9
Tyr196AzAla (1 M trehalose)	2049.8	9.8	2039.1	19.3
Tyr341AzAla	2049.5	9.0	2037.1	15.3
average	2049.6	9.9	2037.9	16.2

Overview of FDH Results

Table A.7: *Overview of the experimental results of the different FDH variants.*

variant	max. concentration / mM	CD spectrum	binds N_3^-	VET peak time / ps	k_{cat} (3°C) / s^{-1}
WT	3	like WT	yes	–	0.19
Y73AzAla+VG	1.1	unlike WT	no	–	N/A
Val121AzAla	0.3	unlike WT	weakly	–	0.0163
Val121AzAla+VG	0.7	unlike WT	no	–	N/A
Val121AzAla+VG (1 M trehalose)	1.1	N/A	no	–	N/A
Tyr196AzAla	1.6	like WT	yes	9.8	0.17
Tyr196AzAla (1 M trehalose)	1.6	N/A	yes	8.5	N/A
Tyr196AzAla+VG	0.3	unlike WT	no	–	–
Trp284AzAla	0.6	N/A	no	–	N/A
Tyr341AzAla	1.8	like WT	yes	9	0.11
Tyr341AzAla+VG	N/A	unlike WT	no	N/A	0.0037

N/A: no measurement was conducted. –: no signal could be observed.

Plasmids

Table A.8: *Plasmids that have been used in this work. For the protein variants only the parent plasmid is listed.*

Plasmid	Resistance	Gene Products (Control)	Source
pET22b(+) *Tm*Azul	amp	*Tm*Azul-HisTag (T7+Lac)	F. H. Arnold (Caltech, Pasadena) [135]
pET22b(+) *Tm*9D8*	amp	*Tm*9D8*-HisTag (T7+Lac)	F. H. Arnold (Caltech, Pasadena) [147]
pGDR11-KOD PDZ	amp	HisTag-TEV-PDZ (T5+Lac)	Bredenbeck Group
pULTRA-CNF	str/spec	polyspecific *Mj* tyrosyl aaRS (tacI) & $tRNA_{CUA}^{Tyr}$ (proK) optimized for CNPhe	Addgene (#48215; [152, 153])
pULTRA-CNF D286R	str/spec	polyspecific *Mj* tyrosyl aaRS with D286R (tacI) & $tRNA_{CUA}^{Tyr}$ (proK) optimized for CNPhe	this work
pQE-70 CypA WT	amp	CypA (T5+Lac)	C. Schiene-Fischer (Martin-Luther-Universität, Halle)
pET28a CypA WT	kan	CypA-HisTag (T7+Lac)	Subcloned from pQE-70 CypA WT by T. Baumann (Technische Universität, Berlin)
pMK-FDH	kan	not applicable for cloning vectors	ThermoFisher Scientific
pET28b FDH	kan	*Cb*FDH (T7+Lac)	this work
pET28b FDH-HisTag	kan	*Cb*FDH-HisTag (T7+Lac)	this work
pBU16 *Mj*AzuRS	amp	*Mj* tyrosyl aaRS (glnS') & $tRNA_{CUA}^{Tyr}$ (proK) optimized for AzAla	N. Budisa (University of Manitoba, Winnipeg) [137]

E. coli Strains

Table A.9: *E. coli strains that have been used in this work.*

Strain Name	Genotype	Source
E. coli BL21 (DE3)	*E. coli* B F$^-$ *ompT gal dcm lon hsdS$_B$* (r_B^-, m_B^-) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)	Bredenbeck Group
E. coli BL21 Star (DE3)	*E. coli* BL21 (DE3) *rne*131	Thermo Fisher Scientific
E. coli DH5α	*E. coli* K-12 F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *hsd*R17(r_K^-, m_K^+), λ^-	Bredenbeck Group
E. coli NEB5α	*E. coli* K-12 F$^-$ *fhuA2*Δ(*argF-lacZ*)U169 *phoA glnV44* φ80*lacZ*M15 *recA1 relA1 gyrA96 endA1 thi-1 hsd*R17	New England Biolabs
E. coli B834 (DE3)	*E. coli* BL21 (DE3) *metE*	Bredenbeck Group
E. coli B-95.ΔA	*E. coli* BL21 (DE3) Δ*prfA*; 95 chromosomal amber stop codons replaced	RIKEN (#RDB13711); K. Sakamoto (RIKEN, Yokohama) [129]
E. coli B-95.ΔAΔ*fabR*	*E. coli* B-95.ΔA Δ*fabR*	RIKEN (#RDB13712); K. Sakamoto (RIKEN, Yokohama) [129]
E. coli C321.ΔA.exp (DE3)	*E. coli* K-12 F$^-$ λ$^-$ *ilvG*$^-$ *rfb*-50 *rph*-1 Δ(*ybhB-bioAB*)::*zeoR* Δ*prfA* λ(DE3); no chromosomal amber stop codons	N. Budisa (University of Manitoba, Winnipeg); based on [128]
E. coli LOBSTR	*E. coli* BL21(DE3) *arnA*(H359S H361S H592S H593S) *SlyD*Δ(His151_His196)	A. Möller (Osnabrück University); based on [270]

References

[1] F. Sanger and H. Tuppy. The amino-acid sequence in the phenylalanyl chain of insulin. 1. The identification of lower peptides from partial hydrolysates. *The Biochemical journal*, 49(4):463–481, 1951. doi: 10.1042/bj0490463.

[2] J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*, 181(4610):662–666, 1958. doi: 10.1038/181662a0.

[3] A. O. W. Stretton. The first sequence: Fred Sanger and insulin. *Genetics*, 162(2):527–532, 2002. doi: 10.1093/genetics/162.2.527.

[4] S. de Chadarevian. John Kendrew and myoglobin: Protein structure determination in the 1950s. *Protein science*, 27(6):1136–1143, 2018. doi: 10.1002/pro.3417.

[5] V. Šrajer and M. Schmidt. Watching proteins function with time-resolved X-ray crystallography. *Journal of physics D: Applied physics*, 50(37), 2017. doi: 10.1088/1361-6463/aa7d32.

[6] F.-A. Chao and R. A. Byrd. Protein dynamics revealed by NMR relaxation methods. *Emerging topics in life sciences*, 2(1):93–105, 2020. doi: 10.1042/etls20170139.

[7] M. Maiuri, M. Garavelli, and G. Cerullo. Ultrafast spectroscopy: State of the art and open challenges. *Journal of the American Chemical Society*, 142(1):3–15, 2020. doi: 10.1021/jacs.9b10533.

[8] D. E. Sagnella and J. E. Straub. Directed energy “funneling" mechanism for heme cooling following ligand photolysis or direct excitation in solvated carbonmonoxy myoglobin. *The journal of physical chemistry. B*, 105(29):7057–7063, 2001. doi: 10.1021/jp0107917.

[9] N. Ota and D. A. Agard. Intramolecular signaling pathways revealed by modeling anisotropic thermal diffusion. *Journal of molecular biology*, 351(2):345–354, 2005. doi: 10.1016/j.jmb.2005.05.043.

[10] D. M. Leitner. Energy flow in proteins. *Annual review of physical chemistry*, 59:233–259, 2008. doi: 10.1146/annurev.physchem.59.032607.093606.

[11] G. Li, D. Magana, and R. B. Dyer. Anisotropic energy flow and allosteric ligand binding in albumin. *Nature communications*, 5:3100, 2014. doi: 10.1038/ncomms4100.

[12] P. M. Champion. Following the flow of energy in biomolecules. *Science*, 310(5750):980–982, 2005. doi: 10.1126/science.1120280.

[13] T. Ishikura and T. Yamato. Energy transfer pathways relevant for long-range intramolecular signaling of photosensory protein revealed by microscopic energy conductivity analysis. *Chemical physics letters*, 432 (4-6):533–537, 2006. doi: 10.1016/j.cplett.2006.10.092.

[14] Y. Xu and D. M. Leitner. Communication maps of vibrational energy transport through photoactive yellow protein. *The journal of physical chemistry. A*, 118(35):7280–7287, 2014. doi: 10.1021/jp411281y.

[15] J. P. T. Zaragoza, A. R. Offenbacher, S. Hu, C. L. Gee, Z. M. Firestein, N. Minnetian, Z. Deng, F. Fan, A. T. Iavarone, and J. P. Klinman. Temporal and spatial resolution of distal protein motions that activate hydrogen tunneling in soybean lipoxygenase. *Proceedings of the National Academy of Sciences of the United States of America*, 120(10):e2211630120, 2023. doi: 10.1073/pnas.2211630120.

[16] Y. Kong and M. Karplus. Signaling pathways of PDZ2 domain: A molecular dynamics interaction correlation analysis. *Proteins*, 74(1):145–154, 2009. doi: 10.1002/prot.22139.

[17] L. Martínez, A. C. M. Figueira, P. Webb, I. Polikarpov, and M. S. Skaf. Mapping the intramolecular vibrational energy flow in proteins reveals functionally important residues. *The journal of physical chemistry letters*, 2(16):2073–2078, 2011. doi: 10.1021/jz200830g.

[18] W. B. Wang, Y. Liang, J. Zhang, Y. D. Wu, J. J. Du, Q. M. Li, J. Z. Zhu, and J. G. Su. Energy transport pathway in proteins: Insights from non-equilibrium molecular dynamics with elastic network model. *Scientific reports*, 8(1):9487, 2018. doi: 10.1038/s41598-018-27745-y.

[19] H. Poudel and D. M. Leitner. Locating dynamic contributions to allostery via determining rates of vibrational energy transfer. *The Journal of chemical physics*, 158(1):015101, 2023. doi: 10.1063/5.0132089.

[20] D. Antoniou and S. D. Schwartz. Internal enzyme motions as a source of catalytic activity: Rate-promoting vibrations and hydrogen tunneling. *Journal of Physical Chemistry B*, 105(23):5553–5558, 2001. doi: 10.1021/jp004547b.

[21] Z. D. Nagel and J. P. Klinman. A 21st century revisionist's view at a turning point in enzymology. *Nature chemical biology*, 5(8):543–550, 2009. doi: 10.1038/nchembio.204.

[22] A. Davarifar, D. Antoniou, and S. D. Schwartz. The promoting vibration in human heart lactate dehydrogenase is a preferred vibrational channel. *The journal of physical chemistry. B*, 115(51):15439–15444, 2011. doi: 10.1021/jp210347h.

[23] Y. Chalopin. The physical origin of rate promoting vibrations in enzymes revealed by structural rigidity. *Scientific reports*, 10(1):17465, 2020. doi: 10.1038/s41598-020-74439-5.

[24] P. Hamm, S. M. Ohline, and W. Zinth. Vibrational cooling after ultrafast photoisomerization of azobenzene measured by femtosecond infrared spectroscopy. *The Journal of chemical physics*, 106(2):519–529, 1997. doi: 10.1063/1.473392.

[25] Y. Mizutani and T. Kitagawa. Direct observation of cooling of heme upon photodissociation of carbonmonoxy myoglobin. *Science*, 278(5337):443–446, 1997. doi: 10.1126/science.278.5337.443.

[26] D. Schwarzer, C. Hanisch, P. Kutne, and J. Troe. Vibrational energy transfer in highly excited bridged azulene-aryl compounds: Direct observation of energy flow through aliphatic chains and into the solvent. *The journal of physical chemistry. A*, 106(35):8019–8028, 2002. doi: 10.1021/jp0210576.

[27] D. V. Kurochkin, S. R. G. Naraharisetty, and I. V. Rubtsov. A relaxation-assisted 2D IR spectroscopy method. *Proceedings of the National Academy of Sciences of the United States of America*, 104(36):14209–14214, 2007. doi: 10.1073/pnas.0700560104.

[28] N.-H. Seong, Y. Fang, and D. D. Dlott. Vibrational energy dynamics of normal and deuterated liquid benzene. *The journal of physical chemistry. A*, 113(8):1445–1452, 2009. doi: 10.1021/jp809679y.

[29] H. M. Müller-Werkmeister and J. Bredenbeck. A donor-acceptor pair for the real time study of vibrational energy transfer in proteins. *Physical chemistry chemical physics*, 16(7):3261–3266, 2014. doi: 10.1039/c3cp54760d.

[30] K. L. Kiick, E. Saxon, D. A. Tirrell, and C. R. Bertozzi. Incorporation of azides into recombinant proteins for chemoselective modification by the staudinger ligation. *Proceedings of the National Academy of Sciences of the United States of America*, 99(1):19–24, 2002. doi: 10.1073/pnas.012583299.

[31] H. M. Müller-Werkmeister. *Unnatural Amino Acids as Novel Probes for Ultrafast 2D-IR Spectroscopy of Proteins: Towards Real-Time Investigation of Biomolecular Dynamics and Vibrational Energy Flow*. Dissertation, Goethe-University, Frankfurt (Germany), 2014.

[32] K. B. Eberl. *Applying Non-Canonical Amino Acids for Investigation of Vibrational Energy Transfer and Dynamic Allostery in a Synaptic Protein Domain*. Dissertation, Goethe-University, Frankfurt (Germany), 2020.

[33] P. W. Atkins, J. de Paula, and J. J. Keeler. *Atkins' physical chemistry*. Oxford University Press, Oxford and New York, 11. edition, 2018. ISBN: 9780198769866.

[34] M. Tasumi. *Introduction to Experimental Infrared Spectroscopy: Fundamentals and Practical Methods*. Wiley, s.l., 1. edition, 2014. ISBN: 978-0-470-66567-1.

[35] P. J. Larkin. *Infrared and raman spectroscopy: Principles and spectral interpretation*. Elsevier, Amsterdam and Boston, 2011. ISBN: 9781493301270.

[36] J. Zheng and M. D. Fayer. Hydrogen bond lifetimes and energetics for solute/solvent complexes studied with 2D-IR vibrational echo spectroscopy. *Journal of the American Chemical Society*, 129(14):4328–4335, 2007. doi: 10.1021/ja067760f.

[37] K.-I. Oh, K. Rajesh, J. F. Stanton, and C. R. Baiz. Quantifying hydrogen-bond populations in dimethyl sulfoxide/water mixtures. *Angewandte Chemie (International ed. in English)*, 56(38):11375–11379, 2017. doi: 10.1002/anie.201704162.

[38] R. J. Fick, A. Y. Liu, F. Nussbaumer, C. Kreutz, A. Rangadurai, Y. Xu, R. D. Sommer, H. Shi, S. Scheiner, and A. L. Stelling. Probing the hydrogen-bonding environment of individual bases in DNA duplexes with isotope-edited infrared spectroscopy. *The journal of physical chemistry. B*, 125(28):7613–7627, 2021. doi: 10.1021/acs.jpcb.1c01351.

[39] R. Bloem, K. Koziol, S. A. Waldauer, B. Buchli, R. Walser, B. Samatanga, I. Jelesarov, and P. Hamm. Ligand binding studied by 2D IR spectroscopy using the azidohomoalanine label. *The journal of physical chemistry. B*, 116(46):13705–13712, 2012. doi: 10.1021/jp3095209.

[40] S. Ramos, R. E. Horness, J. A. Collins, D. Haak, and M. C. Thielges. Site-specific 2D IR spectroscopy: a general approach for the characterization of protein dynamics with high spatial and temporal resolution. *Physical chemistry chemical physics*, 21(2):780–788, 2019. doi: 10.1039/c8cp06146g.

[41] C. Berthomieu and R. Hienerwadel. Fourier transform infrared (FTIR) spectroscopy. *Photosynthesis research*, 101(2-3):157–170, 2009. doi: 10.1007/s11120-009-9439-x.

[42] K. M. Fahelelbom, A. Saleh, M. M. A. Al-Tabakha, and A. A. Ashames. Recent applications of quantitative analytical FTIR spectroscopy in pharmaceutical, biomedical, and clinical fields: A brief review. *Reviews in analytical chemistry*, 41(1):21–33, 2022. doi: 10.1515/revac-2022-0030.

[43] T. J. Johnson, A. Simon, J. M. Weil, and G. W. Harris. Applications of time-resolved step-scan and rapid-scan FT-IR spectroscopy: Dynamics from ten seconds to ten nanoseconds. *Applied Spectroscopy*, 47(9): 1376–1381, 1993. doi: 10.1366/0003702934067397.

[44] M. Muthusamy, E. A. Ambundo, S. J. George, S. J. Lippard, and R. N. F. Thorneley. Stopped-flow fourier transform infrared spectroscopy of nitromethane oxidation by the diiron(iv) intermediate of methane monooxygenase. *Journal of the American Chemical Society*, 125(37):11150–11151, 2003. doi: 10.1021/ja036081r.

[45] H. Fabian and D. Naumann. Methods to study protein folding by stopped-flow FT-IR. *Methods*, 34(1): 28–40, 2004. doi: 10.1016/j.ymeth.2004.03.004.

[46] J. D. Tolland and R. N. F. Thorneley. Stopped-flow fourier transform infrared spectroscopy allows continuous monitoring of azide reduction, carbon monoxide inhibition, and ATP hydrolysis by nitrogenase. *Biochemistry*, 44(27):9520–9527, 2005. doi: 10.1021/bi050453m.

[47] R. A. Palmer, J. L. Chao, R. M. Dittmar, V. G. Gregoriou, and S. E. Plunkett. Investigation of time-dependent phenomena by use of step-scan FT-IR. *Applied Spectroscopy*, 47(9):1297–1310, 1993. doi: 10.1366/0003702934067568.

[48] J. A. MacCammon and S. C. Harvey. *Dynamics of proteins and nucleic acids*. Univ. Pr, Cambridge, repr. edition, 1989. ISBN: 9780521307505.

[49] M. C. Thielges and M. D. Fayer. Protein dynamics studied with ultrafast two-dimensional infrared vibrational echo spectroscopy. *Accounts of chemical research*, 45(11):1866–1874, 2012. doi: 10.1021/ar200275k.

[50] T. Elsaesser. Introduction: Ultrafast processes in chemistry. *Chemical reviews*, 117(16):10621–10622, 2017. doi: 10.1021/acs.chemrev.7b00226.

[51] D. Strickland and G. Mourou. Compression of amplified chirped optical pulses. *Optics communications*, 55(6):447–449, 1985. doi: 10.1016/0030-4018(85)90151-8.

[52] S. in Hwang, S. B. Park, J. Mun, W. Cho, C. H. Nam, and K. T. Kim. Generation of a single-cycle pulse using a two-stage compressor and its temporal characterization using a tunnelling ionization method. *Scientific reports*, 9(1):1613, 2019. doi: 10.1038/s41598-018-38220-z.

[53] G. Cerullo, M. Nisoli, S. Stagira, and S. de Silvestri. Sub-8-fs pulses from an ultrabroadband optical parametric amplifier in the visible. *Optics letters*, 23(16):1283–1285, 1998. doi: 10.1364/ol.23.001283.

[54] M. Seo, K. Tsendsuren, S. Mitra, M. Kling, and D. Kim. High-contrast, intense single-cycle pulses from an all thin-solid-plate setup. *Optics letters*, 45(2):367, 2020. doi: 10.1364/OL.382592.

[55] K. Zhao, Q. Zhang, M. Chini, Y. Wu, X. Wang, and Z. Chang. Tailoring a 67 attosecond pulse through advantageous phase-mismatch. *Optics letters*, 37(18):3891–3893, 2012. doi: 10.1364/ol.37.003891.

[56] M. Hentschel, R. Kienberger, C. Spielmann, G. A. Reider, N. Milosevic, T. Brabec, P. Corkum, U. Heinzmann, M. Drescher, and F. Krausz. Attosecond metrology. *Nature*, 414(6863):509–513, 2001. doi: 10.1038/35107000.

[57] G. M. Greetham, I. P. Clark, B. Young, R. Fritsch, L. Minnes, N. T. Hunt, and M. Towrie. Time-resolved temperature-jump infrared spectroscopy at a high repetition rate. *Applied Spectroscopy*, 74(6):720–727, 2020. doi: 10.1177/0003702820913636.

[58] M. L. Donten, S. Hassan, A. Popp, J. Halter, K. Hauser, and P. Hamm. pH-jump induced leucine zipper folding beyond the diffusion limit. *The journal of physical chemistry. B*, 119(4):1425–1432, 2015. doi: 10.1021/jp511539c.

[59] L. J. G. W. van Wilderen, C. Neumann, A. Rodrigues-Correia, D. Kern-Michler, N. Mielke, M. Reinfelds, A. Heckel, and J. Bredenbeck. Picosecond activation of the DEACM photocage unravelled by VIS-pump-IR-probe spectroscopy. *Physical chemistry chemical physics*, 19(9):6487–6496, 2017. doi: 10.1039/c6cp07022a.

[60] J. Bredenbeck, J. Helbing, J. R. Kumita, G. A. Woolley, and P. Hamm. Alpha-helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7):2379–2384, 2005. doi: 10.1073/pnas.0406948102.

[61] E. T. J. Nibbering, H. Fidder, and E. Pines. Ultrafast chemistry: Using time-resolved vibrational spectroscopy for interrogation of structural dynamics. *Annual review of physical chemistry*, 56:337–367, 2005. doi: 10.1146/annurev.physchem.56.092503.141314.

[62] P. H. Nguyen, S.-M. Park, and G. Stock. Nonequilibrium molecular dynamics simulation of the energy transport through a peptide helix. *The Journal of chemical physics*, 132(2):025102, 2010. doi: 10.1063/1.3284742.

[63] A. Gulzar, L. Valiño Borau, S. Buchenberg, S. Wolf, and G. Stock. Energy transport pathways in proteins: A non-equilibrium molecular dynamics simulation study. *Journal of chemical theory and computation*, 15 (10):5750–5757, 2019. doi: 10.1021/acs.jctc.9b00598.

[64] D. M. Leitner and T. Yamato. Mapping energy transport networks in proteins. In A. L. Parrill and K. B. Lipkowitz, editors, *Reviews in Computational Chemistry, Volume 31*, Reviews in Computational Chemistry, pages 63–113. Wiley, 2018. ISBN: 9781119518020. doi: 10.1002/9781119518068.ch2.

[65] D. M. Leitner, S. Buchenberg, P. Brettel, and G. Stock. Vibrational energy flow in the villin headpiece subdomain: master equation simulations. *The Journal of chemical physics*, 142(7):075101, 2015. doi: 10.1063/1.4907881.

[66] L. Valiño Borau, A. Gulzar, and G. Stock. Master equation model to predict energy transport pathways in proteins. *The Journal of chemical physics*, 152(4):045103, 2020. doi: 10.1063/1.5140070.

[67] S. Buchenberg, D. M. Leitner, and G. Stock. Scaling rules for vibrational energy transport in globular proteins. *The journal of physical chemistry letters*, 7(1):25–30, 2016. doi: 10.1021/acs.jpclett.5b02514.

[68] E. Deniz, L. Valiño-Borau, J. G. Löffler, K. B. Eberl, A. Gulzar, S. Wolf, P. M. Durkin, R. Kaml, N. Budisa, G. Stock, and J. Bredenbeck. Through bonds or contacts? Mapping protein vibrational energy transfer using non-canonical amino acids. *Nature communications*, 12(1):3284, 2021. doi: 10.1038/s41467-021-23591-1.

[69] R. Milo. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays*, 35(12):1050–1055, 2013. doi: 10.1002/bies.201300066.

[70] X. Yu and D. M. Leitner. Vibrational energy transfer and heat conduction in a protein. *The journal of physical chemistry. B*, 107(7):1698–1707, 2003. doi: 10.1021/jp026462b.

[71] P. B. Allen and J. L. Feldman. Thermal conductivity of disordered harmonic solids. *Physical review. B, Condensed matter*, 48(17):12581–12588, 1993. doi: 10.1103/physrevb.48.12581.

[72] D. M. Leitner. Frequency-resolved communication maps for proteins and other nanoscale materials. *The Journal of chemical physics*, 130(19):195101, 2009. doi: 10.1063/1.3130149.

[73] R. Gnanasekaran, J. K. Agbo, and D. M. Leitner. Communication maps computed for homodimeric hemoglobin: Computational study of water-mediated energy transport in proteins. *The Journal of chemical physics*, 135(6):065103, 2011. doi: 10.1063/1.3623423.

[74] K. Moritsugu, O. Miyashita, and A. Kidera. Vibrational energy transfer in a protein molecule. *Physical review letters*, 85(18):3970–3973, 2000. doi: 10.1103/PhysRevLett.85.3970.

[75] T. Yamato and O. Laprévote. Normal mode analysis and beyond. *Biophysics and physicobiology*, 16:322–327, 2019. doi: 10.2142/biophysico.16.0{_}322.

[76] D. J. Nesbitt and R. W. Field. Vibrational energy flow in highly excited molecules: Role of intramolecular vibrational redistribution. *The Journal of physical chemistry*, 100(31):12735–12756, 1996. doi: 10.1021/jp960698w.

[77] M. Gruebele and R. Bigwood. Molecular vibrational energy flow: Beyond the golden rule. *International reviews in physical chemistry*, 17(2):91–145, 1998. doi: 10.1080/014423598230117.

[78] H. Fujisaki and J. E. Straub. Vibrational energy relaxation in proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19):6726–6731, 2005. doi: 10.1073/pnas.0409083102.

[79] B. M. Foley, C. S. Gorham, J. C. Duda, R. Cheaito, C. J. Szwejkowski, C. Constantin, B. Kaehr, and P. E. Hopkins. Protein thermal conductivity measured in the solid state reveals anharmonic interactions of vibrations in a fractal structure. *The journal of physical chemistry letters*, 5(7):1077–1082, 2014. doi: 10.1021/jz500174x.

[80] R. Bigwood, M. Gruebele, D. M. Leitner, and P. G. Wolynes. The vibrational energy flow transition in organic molecules: Theory meets experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 95(11):5960–5964, 1998. doi: 10.1073/pnas.95.11.5960.

[81] L. Bu and J. E. Straub. Simulating vibrational energy flow in proteins: Relaxation rate and mechanism for heme cooling in cytochrome c. *The journal of physical chemistry. B*, 107(44):12339–12345, 2003. doi: 10.1021/jp0351728.

[82] N. Fujii, M. Mizuno, H. Ishikawa, and Y. Mizutani. Observing vibrational energy flow in a protein with the spatial resolution of a single amino acid residue. *The journal of physical chemistry letters*, 5(18):3269–3273, 2014. doi: 10.1021/jz501882h.

[83] R. K. Scopes. Measurement of protein by spectrophotometry at 205 nm. *Analytical biochemistry*, 59(1):277–282, 1974. doi: 10.1016/0003-2697(74)90034-7.

[84] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, and T. Gray. How to measure and predict the molar absorption coefficient of a protein. *Protein science*, 4(11):2411–2423, 1995. doi: 10.1002/pro.5560041120.

[85] N. J. Anthis and G. M. Clore. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein science*, 22(6):851–858, 2013. doi: 10.1002/pro.2253.

[86] A. Barth. Infrared spectroscopy of proteins. *Biochimica et biophysica acta*, 1767(9):1073–1101, 2007. doi: 10.1016/j.bbabio.2007.06.004.

[87] N. J. Greenfield. Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols*, 1(6):2876–2890, 2006. doi: 10.1038/nprot.2006.202.

[88] A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 112(24):E3095–103, 2015. doi: 10.1073/pnas.1500851112.

[89] D. M. Byler and H. Susi. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers*, 25(3):469–487, 1986. doi: 10.1002/bip.360250307.

[90] A. Dong, P. Huang, and W. S. Caughey. Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry*, 29(13):3303–3308, 1990. doi: 10.1021/bi00465a022.

[91] K. E. Wilcox, E. W. Blanch, and A. J. Doig. Determination of protein secondary structure from infrared spectra using partial least-squares regression. *Biochemistry*, 55(27):3794–3802, 2016. doi: 10.1021/acs.biochem.6b00403.

[92] M. Drobizhev, P. R. Callis, R. Nifosì, G. Wicks, C. Stoltzfus, L. Barnett, T. E. Hughes, P. Sullivan, and A. Rebane. Long- and short-range electrostatic fields in GFP mutants: Implications for spectral tuning. *Scientific reports*, 5:13223, 2015. doi: 10.1038/srep13223.

[93] W. J. Bowen. The absorption spectra and extinction coefficients of myoglobin. *The Journal of biological chemistry*, 179(1):235–245, 1949. doi: 10.1016/S0021-9258(18)56832-0.

[94] S. J. Millar, B. W. Moss, and M. H. Stevenson. Some observations on the absorption spectra of various myoglobin derivatives found in meat. *Meat science*, 42(3):277–288, 1996. doi: 10.1016/0309-1740(94)00045-X.

[95] J. Vojtěchovský, K. Chu, J. Berendzen, R. M. Sweet, and I. Schlichting. Crystal structures of myoglobin-ligand complexes at near-atomic resolution. *Biophysical journal*, 77(4):2153–2174, 1999. doi: 10.1016/S0006-3495(99)77056-6.

[96] J. Bredenbeck, J. Helbing, K. Nienhaus, G. U. Nienhaus, and P. Hamm. Protein ligand migration mapped by nonequilibrium 2D-IR exchange spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 104(36):14243–14248, 2007. doi: 10.1073/pnas.0607758104.

[97] J. N. Bandaria, S. Dutta, M. W. Nydegger, W. Rock, A. Kohen, and C. M. Cheatum. Characterizing the dynamics of functionally relevant complexes of formate dehydrogenase. *Proceedings of the National Academy of Sciences of the United States of America*, 107(42):17974–17979, 2010. doi: 10.1073/pnas.0912190107.

[98] V. Massey, S. Ghisla, and K. Yagi. 6-azido- and 6-aminoflavins as active-site probes of flavin enzymes. *Biochemistry*, 25(24):8095–8102, 1986. doi: 10.1021/bi00372a045.

[99] S. Dutta, R. J. Cook, J. C. D. Houtman, A. Kohen, and C. M. Cheatum. Characterization of azido-NAD^+ to assess its potential as a 2D IR probe of enzyme dynamics. *Analytical biochemistry*, 407(2):241–246, 2010. doi: 10.1016/j.ab.2010.08.008.

[100] S. Venyaminov and F. G. Prendergast. Water (H_2O and D_2O) molar absorptivity in the 1000-4000 cm^{-1} range and quantitative infrared spectroscopy of aqueous solutions. *Analytical biochemistry*, 248(2):234–245, 1997. doi: 10.1006/abio.1997.2136.

[101] H. Kim and M. Cho. Infrared probes for studying the structure and dynamics of biomolecules. *Chemical reviews*, 113(8):5817–5847, 2013. doi: 10.1021/cr3005185.

[102] R. Adhikary, J. Zimmermann, and F. E. Romesberg. Transparent window vibrational probes for the characterization of proteins with high structural and temporal resolution. *Chemical reviews*, 117(3):1927–1969, 2017. doi: 10.1021/acs.chemrev.6b00625.

[103] M. C. Thielges. Transparent window 2D IR spectroscopy of proteins. *The Journal of chemical physics*, 155(4):040903, 2021. doi: 10.1063/5.0052628.

[104] K. J. Lee, D. Kang, and H.-S. Park. Site-specific labeling of proteins using unnatural amino acids. *Molecules and cells*, 42(5):386–396, 2019. doi: 10.14348/molcells.2019.0078.

[105] N. Elia. Using unnatural amino acids to selectively label proteins for cellular imaging: a cell biologist viewpoint. *The FEBS journal*, 288(4):1107–1117, 2021. doi: 10.1111/febs.15477.

[106] T. Kimmerlin and D. Seebach. '100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies. *The journal of peptide research : official journal of the American Peptide Society*, 65(2):229–260, 2005. doi: 10.1111/j.1399-3011.2005.00214.x.

[107] J. M. Palomo. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. *RSC advances*, 4(62):32658–32672, 2014. doi: 10.1039/C4RA02458C.

[108] A. Sharma, A. Kumar, B. G. de La Torre, and F. Albericio. Liquid-phase peptide synthesis (LPPS): A third wave for the preparation of peptides. *Chemical reviews*, 2022. doi: 10.1021/acs.chemrev.2c00132.

[109] I. Coin, M. Beyermann, and M. Bienert. Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nature protocols*, 2(12):3247–3256, 2007. doi: 10.1038/nprot.2007.454.

[110] Z. Cui, W. A. Johnston, and K. Alexandrov. Cell-free approach for non-canonical amino acids incorporation into polypeptides. *Frontiers in bioengineering and biotechnology*, 8:1031, 2020. doi: 10.3389/fbioe.2020.01031.

[111] G. L. Rosano and E. A. Ceccarelli. Recombinant protein expression in *Escharichia coli*: advances and challenges. *Frontiers in microbiology*, 5:172, 2014. doi: 10.3389/fmicb.2014.00172.

[112] O. Boutureira and G. J. L. Bernardes. Advances in chemical protein modification. *Chemical reviews*, 115 (5):2174–2195, 2015. doi: 10.1021/cr500399p.

[113] J. A. Shadish and C. A. DeForest. Site-selective protein modification: From functionalized proteins to functional biomaterials. *Matter*, 2(1):50–77, 2020. doi: 10.1016/j.matt.2019.11.011.

[114] D. Alvarez Dorta, D. Deniaud, M. Mével, and S. G. Gouin. Tyrosine conjugation methods for protein labelling. *Chemistry (Weinheim an der Bergstrasse, Germany)*, 26(63):14257–14269, 2020. doi: 10.1002/chem.202001992.

[115] J. Kim, B. X. Li, R. Y.-C. Huang, J. X. Qiao, W. R. Ewing, and D. W. C. MacMillan. Site-selective functionalization of methionine residues via photoredox catalysis. *Journal of the American Chemical Society*, 142 (51):21260–21266, 2020. doi: 10.1021/jacs.0c09926.

[116] S. T. Larda, D. Pichugin, and R. S. Prosser. Site-specific labeling of protein lysine residues and N-terminal amino groups with indoles and indole-derivatives. *Bioconjugate chemistry*, 26(12):2376–2383, 2015. doi: 10.1021/acs.bioconjchem.5b00457.

[117] C. Canovas, M. Moreau, C. Bernhard, A. Oudot, M. Guillemin, F. Denat, and V. Goncalves. Site-specific dual labeling of proteins on cysteine residues with chlorotetrazines. *Angewandte Chemie (International ed. in English)*, 57(33):10646–10650, 2018. doi: 10.1002/anie.201806053.

[118] T. Schlatzer, J. Kriegesmann, H. Schröder, M. Trobe, C. Lembacher-Fadum, S. Santner, A. V. Kravchuk, C. F. W. Becker, and R. Breinbauer. Labeling and natural post-translational modification of peptides and proteins via chemoselective Pd-catalyzed prenylation of cysteine. *Journal of the American Chemical Society*, 141(37):14931–14937, 2019. doi: 10.1021/jacs.9b08279.

[119] A. T. Fafarman, L. J. Webb, J. I. Chuang, and S. G. Boxer. Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins. *Journal of the American Chemical Society*, 128 (41):13356–13357, 2006. doi: 10.1021/ja0650403.

[120] T. Iwasaki, Y. Miyajima-Nakano, R. Fukazawa, M. T. Lin, S.-I. Matsushita, E. Hagiuda, A. T. Taguchi, S. A. Dikanov, Y. Oishi, and R. B. Gennis. *Escharichia coli* amino acid auxotrophic expression host strains for investigating protein structure-function relationships. *Journal of biochemistry*, 169(4):387–394, 2021. doi: 10.1093/jb/mvaa140.

[121] M. A. Shandell, Z. Tan, and V. W. Cornish. Genetic code expansion: A brief history and perspective. *Biochemistry*, 60(46):3455–3469, 2021. doi: 10.1021/acs.biochem.1c00286.

[122] H. Tang, P. Zhang, and X. Luo. Recent technologies for genetic code expansion and their implications on synthetic biology applications. *Journal of molecular biology*, 434(8):167382, 2022. doi: 10.1016/j.jmb.2021.167382.

[123] M. Kimoto and I. Hirao. Genetic code engineering by natural and unnatural base pair systems for the site-specific incorporation of non-standard amino acids into proteins. *Frontiers in molecular biosciences*, 9: 851646, 2022. doi: 10.3389/fmolb.2022.851646.

[124] J. Fredens, K. Wang, D. de La Torre, L. F. H. Funke, W. E. Robertson, Y. Christova, T. Chia, W. H. Schmied, D. L. Dunkelmann, V. Beránek, C. Uttamapinant, A. G. Llamazares, T. S. Elliott, and J. W. Chin. Total synthesis of *Escharichia coli* with a recoded genome. *Nature*, 569(7757):514–518, 2019. doi: 10.1038/s41586-019-1192-5.

[125] W. E. Robertson, L. F. H. Funke, D. de La Torre, J. Fredens, T. S. Elliott, M. Spinck, Y. Christova, D. Cervettini, F. L. Böge, K. C. Liu, S. Buse, S. Maslen, G. P. C. Salmond, and J. W. Chin. Sense codon reassignment enables viral resistance and encoded polymer synthesis. *Science*, 372(6546):1057–1062, 2021. doi: 10.1126/science.abg3029.

[126] C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith, and P. G. Schultz. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science*, 244(4901):182–188, 1989. doi: 10.1126/science.2649980.

[127] X. Fu, Y. Huang, and Y. Shen. Improving the efficiency and orthogonality of genetic code expansion. *BioDesign Research*, 2022:1–13, 2022. doi: 10.34133/2022/9896125.

[128] M. J. Lajoie, A. J. Rovner, D. B. Goodman, H.-R. Aerni, A. D. Haimovich, G. Kuznetsov, J. A. Mercer, H. H. Wang, P. A. Carr, J. A. Mosberg, N. Rohland, P. G. Schultz, J. M. Jacobson, J. Rinehart, G. M. Church, and F. J. Isaacs. Genomically recoded organisms expand biological functions. *Science*, 342(6156):357–360, 2013. doi: 10.1126/science.1241459.

[129] T. Mukai, H. Hoshi, K. Ohtake, M. Takahashi, A. Yamaguchi, A. Hayashi, S. Yokoyama, and K. Sakamoto. Highly reproductive *Escherichia coli* cells with no specific assignment to the UAG codon. *Scientific reports*, 5:9699, 2015. doi: 10.1038/srep09699.

[130] H. Xu, Y. Wang, J. Lu, B. Zhang, Z. Zhang, L. Si, L. Wu, T. Yao, C. Zhang, S. Xiao, L. Zhang, Q. Xia, and D. Zhou. Re-exploration of the codon context effect on amber codon-guided incorporation of noncanonical amino acids in *Escharichia coli* by the blue-white screening assay. *Chembiochem : a European journal of chemical biology*, 17(13):1250–1256, 2016. doi: 10.1002/cbic.201600117.

[131] Y. Chemla, E. Ozer, I. Algov, and L. Alfonta. Context effects of genetic code expansion by stop codon suppression. *Current opinion in chemical biology*, 46:146–155, 2018. doi: 10.1016/j.cbpa.2018.07.012.

[132] I. V. Rubtsov and A. L. Burin. Ballistic and diffusive vibrational energy transport in molecules. *The Journal of chemical physics*, 150(2):020901, 2019. doi: 10.1063/1.5055670.

[133] V. Botan, E. H. G. Backus, R. Pfister, A. Moretto, M. Crisma, C. Toniolo, P. H. Nguyen, G. Stock, and P. Hamm. Energy transport in peptide helices. *Proceedings of the National Academy of Sciences of the United States of America*, 104(31):12749–12754, 2007. doi: 10.1073/pnas.0701762104.

[134] Y. S. Moroz, W. Binder, P. Nygren, G. A. Caputo, and I. V. Korendovych. Painting proteins blue: β-(1-azulenyl)-l-alanine as a probe for studying protein-protein interactions. *Chemical communications (Cambridge, England)*, 49(5):490–492, 2013. doi: 10.1039/c2cc37550h.

[135] E. J. Watkins, P. J. Almhjell, and F. H. Arnold. Direct enzymatic synthesis of a deep-blue fluorescent noncanonical amino acid from azulene and serine. *Chembiochem : a European journal of chemical biology*, 21 (1-2):80–83, 2020. doi: 10.1002/cbic.201900497.

[136] J. G. Löffler, E. Deniz, C. Feid, V. G. Franz, and J. Bredenbeck. Versatile vibrational energy sensors for proteins. *Angewandte Chemie (International ed. in English)*, 2022. doi: 10.1002/anie.202200648.

[137] T. Baumann, M. Hauf, F. Schildhauer, K. B. Eberl, P. M. Durkin, E. Deniz, J. G. Löffler, C. G. Acevedo-Rocha, J. Jaric, B. M. Martins, H. Dobbek, J. Bredenbeck, and N. Budisa. Site-resolved observation of vibrational energy transfer using a genetically encoded ultrafast heater. *Angewandte Chemie (International ed. in English)*, 58(9):2899–2903, 2019. doi: 10.1002/anie.201812995.

[138] J. Shao, I. V. Korendovych, and J. Broos. Biosynthetic incorporation of the azulene moiety in proteins with high efficiency. *Amino acids*, 47(1):213–216, 2015. doi: 10.1007/s00726-014-1870-4.

[139] K. B. Eberl. *Construction of a low-energy OPA for ultrafast infrared spectroscopy, redox-IR spectroscopic characterization of an artificial amino acid and transient IR measurements of GFP*. Master thesis, Goethe-University, Frankfurt (Germany), 2014.

[140] T. Vogt. *Construction of a Femtosecond Infrared Light Source and a Detection System for its Spectral and Temporal Characterization by Frequency-Resolved Optical Gating*. Master thesis, Goethe-University, Frankfurt (Germany), 2016.

[141] P. Hamm, M. Lim, and R. M. Hochstrasser. Vibrational energy relaxation of the cyanide ion in water. *The Journal of chemical physics*, 107(24):10523–10531, 1997. doi: 10.1063/1.474216.

[142] P. Hamm, R. A. Kaindl, and J. Stenger. Noise suppression in femtosecond mid-infrared light sources. *Optics letters*, 25(24):1798–1800, 2000. doi: 10.1364/ol.25.001798.

[143] A. Kondratiev and A. R. Thun. *Development of Instrumentation and Software for 2D-IR Spectroscopy and Implementation of FT-VIPER Spectroscopy*. Master thesis, Goethe-University, Frankfurt (Germany), 2020.

[144] F. W. Studier. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification*, 41(1):207–234, 2005. doi: 10.1016/j.pep.2005.01.016.

[145] V. G. Franz. *VET Donor-Sensor Pairs in PDZ3: Enzymatic Synthesis, Spectroscopic Characterisation and Protein Labelling*. Master thesis, Goethe-University, Frankfurt (Germany), 2022.

[146] C. Feid. *Nitrile-Based Probes for Ultrafast IR Spectroscopy: Enzmatic Synthesis, Spectroscopic Characterization and Protein Labeling*. Master thesis, Goethe-University, Frankfurt (Germany), 2021.

[147] C. E. Boville, D. K. Romney, P. J. Almhjell, M. Sieben, and F. H. Arnold. Improved synthesis of 4-cyanotryptophan and other tryptophan analogues in aqueous solvent using variants of trpB from thermotoga maritima. *The Journal of organic chemistry*, 83(14):7447–7452, 2018. doi: 10.1021/acs.joc.8b00517.

[148] Y. A. Mikheev, L. N. Guseva, and Y. A. Ershov. Mechanism of the protonation of azulenes in aqueous solutions of acids. *Russian Journal of Physical Chemistry A*, 87(10):1645–1653, 2013. doi: 10.1134/S003602441309015X. PII: 760.

[149] L. J. G. W. van Wilderen, C. N. Lincoln, and J. J. van Thor. Modelling multi-pulse population dynamics from ultrafast spectroscopy. *PloS one*, 6(3):e17373, 2011. doi: 10.1371/journal.pone.0017373.

[150] Y. Fu, M. Huang, and W. Li. Spatial intensity distribution model of fluorescence emission considering the spatial attenuation effect of excitation light. *Optics express*, 29(5):6468–6480, 2021. doi: 10.1364/OE.416452.

[151] G. D. van Duyne, R. F. Standaert, P. A. Karplus, S. L. Schreiber, and J. Clardy. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *Journal of molecular biology*, 229 (1):105–124, 1993. doi: 10.1006/jmbi.1993.1012.

[152] K. C. Schultz, L. Supekova, Y. Ryu, J. Xie, R. Perera, and P. G. Schultz. A genetically encoded infrared probe. *Journal of the American Chemical Society*, 128(43):13984–13985, 2006. doi: 10.1021/ja0636690.

[153] D. D. Young, T. S. Young, M. Jahnz, I. Ahmad, G. Spraggon, and P. G. Schultz. An evolved aminoacyl-tRNA synthetase with atypical polysubstrate specificity. *Biochemistry*, 50(11):1894–1900, 2011. doi: 10.1021/bi101929e.

[154] T. Kobayashi, O. Nureki, R. Ishitani, A. Yaremchuk, M. Tukalo, S. Cusack, K. Sakamoto, and S. Yokoyama. Structural basis for orthogonal tRNA specificities of tyrosyl-trna synthetases for genetic code expansion. *Nature structural biology*, 10(6):425–432, 2003. doi: 10.1038/nsb934.

[155] J. M. Schmidt-Engler, L. Blankenburg, R. Zangl, J. Hoffmann, N. Morgner, and J. Bredenbeck. Local dynamics of the photo-switchable protein PYP in ground and signalling state probed by 2D-IR spectroscopy of -scn labels. *Physical chemistry chemical physics*, 22(40):22963–22972, 2020. doi: 10.1039/d0cp04307a.

[156] J. Lee. *Kristallisation und Kristallstrukturanalyse von Protein - Mutanten mit nicht kanonischen Aminosäuren*. Bachelor thesis, Goethe-University, Frankfurt (Germany), 2021.

[157] M. Gerlach, U. Mueller, and M. S. Weiss. The MX beamlines BL14.1-3 at BESSY II. *Journal of large-scale research facilities JLSRF*, 2, 2016. doi: 10.17815/jlsrf-2-64.

[158] W. Kabsch. Xds. *Acta crystallographica. Section D, Biological crystallography*, 66(Pt 2):125–132, 2010. doi: 10.1107/S0907444909047337.

[159] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. Phaser crystallographic software. *Journal of applied crystallography*, 40(Pt 4):658–674, 2007. doi: 10.1107/S0021889807021206.

[160] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, and P. D. Adams. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in phenix. *Acta crystallographica. Section D, Structural biology*, 75(Pt 10):861–877, 2019. doi: 10.1107/S2059798319011471.

[161] E. Deniz, J. G. Löffler, A. Kondratiev, A. R. Thun, Y. Shen, G. Wille, and J. Bredenbeck. High-precision background correction and artifact suppression for ultrafast spectroscopy by quasi-simultaneous measurements in a split-sample cell. *Review of Scientific Instruments*, 93(3):033001, 2022. doi: 10.1063/5.0079958.

[162] M. Forloni, A. Y. Liu, and N. Wajapeyee. Megaprimer polymerase chain reaction (PCR)-based mutagenesis. *Cold Spring Harbor protocols*, 2019(6), 2019. doi: 10.1101/pdb.prot097824.

[163] G. Fischer, H. Bang, and C. Mech. Nachweis einer enzymkatalyse für die cis-trans-isomerisierung der peptidbindung in prolinhaltigen peptiden. *Biomedica biochimica acta*, 43(10):1101–1111, 1984.

[164] M. Venanzi, A. Valeri, A. Palleschi, L. Stella, L. Moroder, F. Formaggio, C. Toniolo, and B. Pispisa. Structural properties and photophysical behavior of conformationally constrained hexapeptides functionalized with a new fluorescent analog of tryptophan and a nitroxide radical quencher. *Biopolymers*, 75(2):128–139, 2004. doi: 10.1002/bip.20110.

[165] N. E. Labrou, A. Karagouni, and Y. D. Clonis. Biomimetic-dye affinity adsorbents for enzyme purification: application to the one-step purification of candida boidinii formate dehydrogenase. *Biotechnology and bioengineering*, 48(3):278–288, 1995. doi: 10.1002/bit.260480314.

[166] N. E. Labrou. Improved purification of Candida boidinii formate dehydrogenase. *Bioseparation*, 9(2):99–104, 2000. doi: 10.1023/a:1008131320571.

[167] Q. Guo, L. Gakhar, K. Wickersham, K. Francis, A. Vardi-Kilshtain, D. T. Major, C. M. Cheatum, and A. Kohen. Structural and kinetic studies of formate dehydrogenase from candida boidinii. *Biochemistry*, 55(19):2760–2771, 2016. doi: 10.1021/acs.biochem.6b00181.

[168] Y. Shen. *Expression und Charakterisierung von Formiat-Dehydrogenase Mutanten für Energietransfermessungen sowie Verbesserung der Messtechnik*. Master thesis, Goethe-University, Frankfurt (Germany), 2022.

[169] N. Joachim. *Expression and Spectroscopic Characterisation of Formate Dehydrogenase Mutants Labelled with Para-Cyano-Phenylalanine*. Bachelor thesis, Goethe-University, Frankfurt (Germany), 2020.

[170] X. S. Gai, B. A. Coutifaris, S. H. Brewer, and E. E. Fenlon. A direct comparison of azide and nitrile vibrational probes. *Physical chemistry chemical physics*, 13(13):5926–5930, 2011. doi: 10.1039/c0cp02774j.

[171] J. Zhang, L. Wang, J. Zhang, J. Zhu, X. Pan, Z. Cui, J. Wang, W. Fang, and Y. Li. Identifying and modulating accidental fermi resonance: 2D IR and DFT study of 4-azido-l-phenylalanine. *The journal of physical chemistry. B*, 122(34):8122–8133, 2018. doi: 10.1021/acs.jpcb.8b03887.

[172] J. Y. Park, H.-J. Kwon, S. Mondal, H. Han, K. Kwak, and M. Cho. Two-dimensional IR spectroscopy reveals a hidden fermi resonance band in the azido stretch spectrum of β-azidoalanine. *Physical chemistry chemical physics*, 22(34):19223–19229, 2020. doi: 10.1039/D0CP02693J.

[173] J. Y. Park, S. Mondal, H.-J. Kwon, P. K. Sahu, H. Han, K. Kwak, and M. Cho. Effect of isotope substitution on the fermi resonance and vibrational lifetime of unnatural amino acids modified with ir probe: A 2D-IR and pump-probe study of 4-azido-L-phenyl alanine. *The Journal of chemical physics*, 153(16):164309, 2020. doi: 10.1063/5.0025289.

[174] H. D. Pandey and D. M. Leitner. Vibrational states and nitrile lifetimes of cyanophenylalanine isotopomers in solution. *The journal of physical chemistry. A*, 122(34):6856–6863, 2018. doi: 10.1021/acs.jpca.8b06300.

[175] H.-J. Lee and J. J. Zheng. Pdz domains and their binding partners: structure, specificity, and modification. *Cell communication and signaling : CCS*, 8:8, 2010. doi: 10.1186/1478-811X-8-8.

[176] D. A. Doyle, A. Lee, J. Lewis, E. Kim, M. Sheng, and R. MacKinnon. Crystal structures of a complexed and peptide-free membrane protein–binding domain: Molecular basis of peptide recognition by pdz. *Cell*, 85(7):1067–1076, 1996. doi: 10.1016/S0092-8674(00)81307-0.

[177] C.-J. Tsai, A. del Sol, and R. Nussinov. Allostery: absence of a change in shape does not imply that allostery is not at play. *Journal of molecular biology*, 378(1):1–11, 2008. doi: 10.1016/j.jmb.2008.02.034.

[178] H. N. Motlagh, J. O. Wrabl, J. Li, and V. J. Hilser. The ensemble nature of allostery. *Nature*, 508(7496):331–339, 2014. doi: 10.1038/nature13001.

[179] C. M. Petit, J. Zhang, P. J. Sapienza, E. J. Fuentes, and A. L. Lee. Hidden dynamic allostery in a PDZ domain. *Proceedings of the National Academy of Sciences of the United States of America*, 106(43):18249–18254, 2009. doi: 10.1073/pnas.0904492106.

[180] S. W. Lockless and R. Ranganathan. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science*, 286(5438):295–299, 1999. doi: 10.1126/science.286.5438.295.

[181] K.-I. Oh, J.-H. Lee, C. Joo, H. Han, and M. Cho. β-azidoalanine as an IR probe: application to amyloid aβ(16-22) aggregation. *The journal of physical chemistry. B*, 112(33):10352–10357, 2008. doi: 10.1021/jp801558k.

[182] S. Nagarajan, H. Taskent-Sezgin, D. Parul, I. Carrico, D. P. Raleigh, and R. B. Dyer. Differential ordering of the protein backbone and side chains during protein folding revealed by site-specific recombinant infrared probes. *Journal of the American Chemical Society*, 133(50):20335–20340, 2011. doi: 10.1021/ja2071362.

[183] M. Essig. *Using artificial amino acids to induce and measure energy transfer in a WW domain*. Master thesis, Goethe-University, Frankfurt (Germany), 2014.

[184] L. Lešetický, R. Barth, I. Nêmec, M. Štícha, and I. Tišlerová. Synthesis and spectra of n-15 labelled phenylazides. *Czechoslovak Journal of Physics*, 53(S1):A777–A782, 2003. doi: 10.1007/s10582-003-0101-0.

[185] J. S. Lipkin, R. Song, E. E. Fenlon, and S. H. Brewer. Modulating accidental fermi resonance: What a difference a neutron makes. *The journal of physical chemistry letters*, 2011(2):1672–1676, 2011. doi: 10.1021/jz2006447.

[186] J. M. Rodgers, W. Zhang, C. G. Bazewicz, J. Chen, S. H. Brewer, and F. Gai. Kinetic isotope effect provides insight into the vibrational relaxation mechanism of aromatic molecules: Application to cyanophenylalanine. *The journal of physical chemistry letters*, 7(7):1281–1287, 2016. doi: 10.1021/acs.jpclett.6b00325.

[187] A. L. Le Sueur, S. Ramos, J. D. Ellefsen, S. Cook, and M. C. Thielges. Evaluation of p-(^{13}c,^{15}n-cyano)phenylalanine as an extended time scale 2D IR probe of proteins. *Analytical chemistry*, 89(10):5254–5260, 2017. doi: 10.1021/acs.analchem.6b04650.

[188] L. J. G. W. van Wilderen, H. Brunst, H. Gustmann, J. Wachtveitl, J. Broos, and J. Bredenbeck. Cyanotryptophans as dual infrared and fluorescence spectroscopic labels to assess structural dynamics in proteins. *Physical chemistry chemical physics*, 20(30):19906–19915, 2018. doi: 10.1039/C8CP00846A.

[189] P. Talukder, S. Chen, B. Roy, P. Yakovchuk, M. M. Spiering, M. P. Alam, M. M. Madathil, C. Bhattacharya, S. J. Benkovic, and S. M. Hecht. Cyanotryptophans as novel fluorescent probes for studying protein conformational changes and DNA-protein interaction. *Biochemistry*, 54(51):7457–7469, 2015. doi: 10.1021/acs.biochem.5b01085.

[190] B. N. Markiewicz, D. Mukherjee, T. Troxler, and F. Gai. Utility of 5-cyanotryptophan fluorescence as a sensitive probe of protein hydration. *The journal of physical chemistry. B*, 120(5):936–944, 2016. doi: 10.1021/acs.jpcb.5b12233.

[191] M. R. Hilaire, I. A. Ahmed, C.-W. Lin, H. Jo, W. F. DeGrado, and F. Gai. Blue fluorescent amino acid for biological spectroscopy and microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 114(23):6005–6009, 2017. doi: 10.1073/pnas.1705586114.

[192] M. M. Waegele, M. J. Tucker, and F. Gai. 5-cyanotryptophan as an infrared probe of local hydration status of proteins. *Chemical physics letters*, 478(4):249–253, 2009. doi: 10.1016/j.cplett.2009.07.058.

[193] A. Acharyya, W. Zhang, and F. Gai. Tryptophan as a template for development of visible fluorescent amino acids. *The journal of physical chemistry. B*, 125(21):5458–5465, 2021. doi: 10.1021/acs.jpcb.1c02321.

[194] F. Schildhauer, T. Baumann, and N. Budisa. private communication, 2019.

[195] L. J. G. W. van Wilderen, D. Kern-Michler, H. M. Müller-Werkmeister, and J. Bredenbeck. Vibrational dynamics and solvatochromism of the label SCN in various solvents and hemoglobin by time dependent IR and 2D-IR spectroscopy. *Physical chemistry chemical physics*, 16(36):19643–19653, 2014. doi: 10.1039/C4CP01498G.

[196] J. M. Schmidt-Engler, R. Zangl, P. Guldan, N. Morgner, and J. Bredenbeck. Exploring the 2D-IR repertoire of the -SCN label to study site-resolved dynamics and solvation in the calcium sensor protein calmodulin. *Physical chemistry chemical physics*, 22(10):5463–5475, 2020. doi: 10.1039/c9cp06808b.

[197] D. Kern-Michler. *Introduction of a Thiocyanate Label into PDZ*. Bachelor thesis, Goethe-University, Frankfurt (Germany), 2012.

[198] M. P. Wolfshorndl, R. Baskin, I. Dhawan, and C. H. Londergan. Covalently bound azido groups are very specific water sensors, even in hydrogen-bonding environments. *The journal of physical chemistry. B*, 116 (3):1172–1179, 2012. doi: 10.1021/jp209899m.

[199] P. Deb, T. Haldar, S. M. Kashid, S. Banerjee, S. Chakrabarty, and S. Bagchi. Correlating nitrile IR frequencies to local electrostatics quantifies noncovalent interactions of peptides and proteins. *The journal of physical chemistry. B*, 120(17):4034–4046, 2016. doi: 10.1021/acs.jpcb.6b02732.

[200] M. Maj, C. Ahn, B. Błasiak, K. Kwak, H. Han, and M. Cho. Isonitrile as an ultrasensitive infrared reporter of hydrogen-bonding structure and dynamics. *The journal of physical chemistry. B*, 120(39):10167–10180, 2016. doi: 10.1021/acs.jpcb.6b04319.

[201] M. C. Thielges, J. Y. Axup, D. Wong, H. S. Lee, J. K. Chung, P. G. Schultz, and M. D. Fayer. Two-dimensional ir spectroscopy of protein dynamics using two vibrational labels: a site-specific genetically encoded unnatural amino acid and an active site ligand. *The journal of physical chemistry. B*, 115(38): 11294–11304, 2011. doi: 10.1021/jp206986v.

[202] N. I. Rubtsova, Z. Lin, R. T. Mackin, and I. V. Rubtsov. How intramolecular vibrational energy transport changes with rigidity and polarity of the environment? *High Energy Chemistry*, 54(6):427–435, 2020. doi: 10.1134/S0018143920060120.

[203] S. Yamashita, M. Mizuno, K. Takemura, A. Kitao, and Y. Mizutani. Dependence of vibrational energy transfer on distance in a four-helix bundle protein: Equidistant increments with the periodicity of α helices. *The journal of physical chemistry. B*, 126(17):3283–3290, 2022. doi: 10.1021/acs.jpcb.2c00956.

[204] K.-H. Park, J. Jeon, Y. Park, S. Lee, H.-J. Kwon, C. Joo, S. Park, H. Han, and M. Cho. Infrared probes based on nitrile-derivatized prolines: Thermal insulation effect and enhanced dynamic range. *The journal of physical chemistry letters*, 4(13):2105–2110, 2013. doi: 10.1021/jz400954r.

[205] D. E. Levin, A. J. Schmitz, S. M. Hines, K. J. Hines, M. J. Tucker, S. H. Brewer, and E. E. Fenlon. Synthesis and evaluation of the sensitivity and vibrational lifetimes of thiocyanate and selenocyanate infrared reporters. *RSC advances*, 43(6):36231–36237, 2016. doi: 10.1039/C5RA27363C.

[206] S. Ramos, K. J. Scott, R. E. Horness, A. L. Le Sueur, and M. C. Thielges. Extended timescale 2D IR probes of proteins: p-cyanoselenophenylalanine. *Physical chemistry chemical physics*, 19(15):10081–10086, 2017. doi: 10.1039/C7CP00403F.

[207] F. Chalyavi, A. J. Schmitz, N. R. Fetto, M. J. Tucker, S. H. Brewer, and E. E. Fenlon. Extending the vibrational lifetime of azides with heavy atoms. *Physical chemistry chemical physics*, 22(32):18007–18013, 2020. doi: 10.1039/D0CP02814B.

[208] R. E. Bruccoleri, M. Karplus, and J. A. McCammon. The hinge-bending mode of a lysozyme-inhibitor complex. *Biopolymers*, 25(9):1767–1802, 1986. doi: 10.1002/bip.360250916.

[209] D. A. Turton, H. M. Senn, T. Harwood, A. J. Lapthorn, E. M. Ellis, and K. Wynne. Terahertz underdamped vibrational motion governs protein-ligand binding in solution. *Nature communications*, 5:3999, 2014. doi: 10.1038/ncomms4999.

[210] P. Pagano, Q. Guo, A. Kohen, and C. M. Cheatum. Oscillatory enzyme dynamics revealed by two-dimensional infrared spectroscopy. *The journal of physical chemistry letters*, 7(13):2507–2511, 2016. doi: 10.1021/acs.jpclett.6b01154.

[211] J. M. Schmidt-Engler, S. von Berg, and J. Bredenbeck. Temperature-dependent low-frequency modes in the active site of bovine carbonic anhydrase ii probed by 2D-IR spectroscopy. *The journal of physical chemistry letters*, 12(32):7777–7782, 2021. doi: 10.1021/acs.jpclett.1c01453.

[212] D. Antoniou and S. D. Schwartz. Large kinetic isotope effects in enzymatic proton transfer and the role of substrate oscillations. *Proceedings of the National Academy of Sciences of the United States of America*, 94 (23):12360–12365, 1997. doi: 10.1073/pnas.94.23.12360.

[213] V. L. Schramm and S. D. Schwartz. Promoting vibrations and the function of enzymes. emerging theoretical and experimental convergence. *Biochemistry*, 57(24):3299–3308, 2018. doi: 10.1021/acs.biochem.8b00201.

[214] C. M. Cheatum. Low-frequency protein motions coupled to catalytic sites. *Annual review of physical chemistry*, 71:267–288, 2020. doi: 10.1146/annurev-physchem-050317-014308.

[215] C. M. Cheatum and A. Kohen. Relationship of femtosecond-picosecond dynamics to enzyme-catalyzed h-transfer. *Topics in current chemistry*, 337:1–39, 2013. doi: 10.1007/128_2012_407.

[216] J. P. Klinman. An integrated model for enzyme catalysis emerges from studies of hydrogen tunneling. *Chemical physics letters*, 471(4-6):179–193, 2009. doi: 10.1016/j.cplett.2009.01.038.

[217] L. Wang, N. M. Goodey, S. J. Benkovic, and A. Kohen. Coordinated effects of distal mutations on environmentally coupled tunneling in dihydrofolate reductase. *Proceedings of the National Academy of Sciences of the United States of America*, 103(43):15753–15758, 2006. doi: 10.1073/pnas.0606976103.

[218] Z. D. Nagel, C. W. Meadows, M. Dong, B. J. Bahnson, and J. P. Klinman. Active site hydrophobic residues impact hydrogen tunneling differently in a thermophilic alcohol dehydrogenase at optimal versus nonoptimal temperatures. *Biochemistry*, 51(20):4147–4156, 2012. doi: 10.1021/bi3001352.

[219] P. Pagano, Q. Guo, C. Ranasinghe, E. Schroeder, K. Robben, F. Häse, H. Ye, K. Wickersham, A. Aspuru-Guzik, D. T. Major, L. Gakhar, A. Kohen, and C. M. Cheatum. Oscillatory active-site motions correlate with kinetic isotope effects in formate dehydrogenase. *ACS catalysis*, 9(12):11199–11206, 2019. doi: 10.1021/acscatal.9b03345.

[220] J. P. Klinman and A. Kohen. Evolutionary aspects of enzyme dynamics. *The Journal of biological chemistry*, 289(44):30205–30212, 2014. doi: 10.1074/jbc.R114.565515.

[221] D. Petrović, V. A. Risso, S. C. L. Kamerlin, and J. M. Sanchez-Ruiz. Conformational dynamics and enzyme evolution. *Journal of the Royal Society, Interface*, 15(144), 2018. doi: 10.1098/rsif.2018.0330.

[222] J. W. Schafer and S. D. Schwartz. Directed evolution's influence on rapid density fluctuations illustrates how protein dynamics can become coupled to chemistry. *ACS catalysis*, 10(15):8476–8484, 2020. doi: 10.1021/acscatal.0c01618.

[223] D. Antoniou, S. Caratzoulas, C. Kalyanaraman, J. S. Mincer, and S. D. Schwartz. Barrier passage and protein dynamics in enzymatically catalyzed reactions. *European journal of biochemistry*, 269(13):3103–3112, 2002. doi: 10.1046/j.1432-1033.2002.03021.x.

[224] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annual review of physical chemistry*, 53(1):291–318, 2002. doi: 10.1146/annurev.physchem.53.082301.113146.

[225] D. Antoniou and S. D. Schwartz. Role of protein motions in catalysis by formate dehydrogenase. *The journal of physical chemistry. B*, 124(43):9483–9489, 2020. doi: 10.1021/acs.jpcb.0c05725.

[226] D. Antoniou and S. D. Schwartz. The stochastic separatrix and the reaction coordinate for complex systems. *The Journal of chemical physics*, 130(15):151103, 2009. doi: 10.1063/1.3123162.

[227] J. E. Basner and S. D. Schwartz. How enzyme dynamics helps catalyze a reaction in atomic detail: A transition path sampling study. *Journal of the American Chemical Society*, 127(40):13822–13831, 2005. doi: 10.1021/ja043320h.

[228] S. Hay, L. O. Johannissen, M. J. Sutcliffe, and N. S. Scrutton. Barrier compression and its contribution to both classical and quantum mechanical aspects of enzyme catalysis. *Biophysical journal*, 98(1):121–128, 2010. doi: 10.1016/j.bpj.2009.09.045.

[229] S. L. Quaytman and S. D. Schwartz. Reaction coordinate of an enzymatic reaction revealed by transition path sampling. *Proceedings of the National Academy of Sciences of the United States of America*, 104(30):12253–12258, 2007. doi: 10.1073/pnas.0704304104.

[230] M. W. Dzierlenga and S. D. Schwartz. Targeting a rate-promoting vibration with an allosteric mediator in lactate dehydrogenase. *The journal of physical chemistry letters*, 7(13):2591–2596, 2016. doi: 10.1021/acs.jpclett.6b01209.

[231] B. A. Andrews and R. B. Dyer. Small molecule cores demonstrate non-competitive inhibition of lactate dehydrogenase. *MedChemComm*, 9(8):1369–1376, 2018. doi: 10.1039/c8md00309b.

[232] R. E. Handschumacher, M. W. Harding, J. Rice, R. J. Drugge, and D. W. Speicher. Cyclophilin: A specific cytosolic binding protein for cyclosporin A. *Science*, 226(4674):544–547, 1984. doi: 10.1126/science.6238408.

[233] N. Takahashi, T. Hayano, and M. Suzuki. Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin. *Nature*, 337(6206):473–475, 1989. doi: 10.1038/337473a0.

[234] G. Fischer, B. Wittmann-Liebold, K. Lang, T. Kiefhaber, and F. X. Schmid. Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. *Nature*, 337(6206):476–478, 1989. doi: 10.1038/337476a0.

[235] J. Fanghänel and G. Fischer. Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. *Frontiers in bioscience : a journal and virtual library*, 9:3453–3478, 2004. doi: 10.2741/1494.

[236] P. Wang and J. Heitman. The cyclophilins. *Genome biology*, 6(7):226, 2005. doi: 10.1186/gb-2005-6-7-226.

[237] P. Nigro, G. Pompilio, and M. C. Capogrossi. Cyclophilin A: A key player for human disease. *Cell death & disease*, 4(10):e888, 2013. doi: 10.1038/cddis.2013.410.

[238] S. D. Hanes. Prolyl isomerases in gene transcription. *Biochimica et biophysica acta*, 1850(10):2017–2034, 2015. doi: 10.1016/j.bbagen.2014.10.028.

[239] J. F. Brandts, H. R. Halvorson, and M. Brennan. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry*, 14(22):4953–4963, 1975. doi: 10.1021/bi00693a026.

[240] E. R. Schönbrunner, S. Mayer, M. Tropschug, G. Fischer, N. Takahashi, and F. X. Schmid. Catalysis of protein folding by cyclophilins from different species. *The Journal of biological chemistry*, 266(6):3630–3635, 1991. doi: 10.1016/S0021-9258(19)67841-5.

[241] L. Min, D. B. Fulton, and A. H. Andreotti. A case study of proline isomerization in cell signaling. *Frontiers in bioscience : a journal and virtual library*, 10:385–397, 2005. doi: 10.2741/1536.

[242] J. L. Kofron, P. Kuzmic, V. Kishore, E. Colón-Bonilla, and D. H. Rich. Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. *Biochemistry*, 30(25): 6127–6134, 1991. doi: 10.1021/bi00239a007.

[243] P. Mark and L. Nilsson. A molecular dynamics study of cyclophilin A free and in complex with the Ala-Pro dipeptide. *European biophysics journal : EBJ*, 36(3):213–224, 2007. doi: 10.1007/s00249-006-0121-3.

[244] V. Leone, G. Lattanzi, C. Molteni, and P. Carloni. Mechanism of action of cyclophilin A explored by metadynamics simulations. *PLoS computational biology*, 5(3):e1000309, 2009. doi: 10.1371/journal.pcbi.1000309.

[245] C. Camilloni, A. B. Sahakyan, M. J. Holliday, N. G. Isern, F. Zhang, E. Z. Eisenmesser, and M. Vendruscolo. Cyclophilin A catalyzes proline isomerization by an electrostatic handle mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 111(28):10203–10208, 2014. doi: 10.1073/pnas.1404220111.

[246] P. K. Agarwal, A. Geist, and A. Gorin. Protein dynamics and enzymatic catalysis: investigating the peptidyl-prolyl cis-trans isomerization activity of cyclophilin A. *Biochemistry*, 43(33):10605–10618, 2004. doi: 10.1021/bi0495228.

[247] J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, and T. Alber. Hidden alternative structures of proline isomerase essential for catalysis. *Nature*, 462(7273):669–673, 2009. doi: 10.1038/nature08615.

[248] A. Ramanathan, A. J. Savol, C. J. Langmead, P. K. Agarwal, and C. S. Chennubhotla. Discovering conformational sub-states relevant to protein function. *PloS one*, 6(1):e15827, 2011. doi: 10.1371/journal.pone.0015827.

[249] M. J. Holliday, C. Camilloni, G. S. Armstrong, M. Vendruscolo, and E. Z. Eisenmesser. Networks of dynamic allostery regulate enzyme function. *Structure (London, England : 1993)*, 25(2):276–286, 2017. doi: 10.1016/j.str.2016.12.003.

[250] P. Wapeesittipan, A. S. J. S. Mey, M. D. Walkinshaw, and J. Michel. Allosteric effects in cyclophilin mutants may be explained by changes in nano-microsecond time scale motions. *Communications Chemistry*, 2(1), 2019. doi: 10.1038/s42004-019-0136-1.

[251] T. Haliloglu, A. Hacisuleyman, and B. Erman. Prediction of allosteric communication pathways in proteins. *Bioinformatics (Oxford, England)*, 38(14):3590–3599, 2022. doi: 10.1093/bioinformatics/btac380.

[252] R. Otten, L. Liu, L. R. Kenner, M. W. Clarkson, D. Mavor, D. S. Tawfik, D. Kern, and J. S. Fraser. Rescue of conformational dynamics in enzyme catalysis by directed evolution. *Nature communications*, 9(1):1314, 2018. doi: 10.1038/s41467-018-03562-9.

[253] M. Z. Wang, J. T. Shetty, B. A. Howard, M. J. Campa, E. F. Patz, and M. C. Fitzgerald. Thermodynamic analysis of cyclosporin A binding to cyclophilin a in a lung tumor tissue lysate. *Analytical chemistry*, 76 (15):4343–4348, 2004. doi: 10.1021/ac049536j.

[254] V. Mikol, J. Kallen, G. Pflügl, and M. D. Walkinshaw. X-ray structure of a monomeric cyclophilin A-cyclosporin A crystal complex at 2.1 Å resolution. *Journal of molecular biology*, 234(4):1119–1130, 1993. doi: 10.1006/jmbi.1993.1664.

[255] F. F. Vajdos, S. Yoo, M. Houseweart, W. I. Sundquist, and C. P. Hill. Crystal structure of cyclophilin A complexed with a binding site peptide from the HIV-1 capsid protein. *Protein science*, 6(11):2297–2307, 1997. doi: 10.1002/pro.5560061103.

[256] S. Yoo, D. G. Myszka, C. Yeh, M. McMurray, C. P. Hill, and W. I. Sundquist. Molecular recognition in the HIV-1 capsid/cyclophilin A complex. *Journal of molecular biology*, 269(5):780–795, 1997. doi: 10.1006/jmbi.1997.1051.

[257] K. Piotukh, W. Gu, M. Kofler, D. Labudde, V. Helms, and C. Freund. Cyclophilin A binds to linear peptide motifs containing a consensus that is present in many human proteins. *The Journal of biological chemistry*, 280(25):23668–23674, 2005. doi: 10.1074/jbc.M503405200.

[258] J. D. Hermes, S. W. Morrical, M. H. O'Leary, and W. W. Cleland. Variation of transition-state structure as a function of the nucleotide in reactions catalyzed by dehydrogenases. 2. Formate dehydrogenase. *Biochemistry*, 23(23):5479–5488, 1984. doi: 10.1021/bi00318a016.

[259] V. I. Tishkov and V. O. Popov. Protein engineering of formate dehydrogenase. *Biomolecular engineering*, 23(2-3):89–110, 2006. doi: 10.1016/j.bioeng.2006.02.003.

[260] C. Ranasinghe, P. Pagano, P. J. Sapienza, A. L. Lee, A. Kohen, and C. M. Cheatum. Isotopic labeling of formate dehydrogenase perturbs the protein dynamics. *The journal of physical chemistry. B*, 123(49):10403–10409, 2019. doi: 10.1021/acs.jpcb.9b08426.

[261] J. N. Bandaria, C. M. Cheatum, and A. Kohen. Examination of enzymatic H-tunneling through kinetics and dynamics. *Journal of the American Chemical Society*, 131(29):10151–10155, 2009. doi: 10.1021/ja902120t.

[262] C. Ranasinghe, Q. Guo, P. J. Sapienza, A. L. Lee, D. M. Quinn, C. M. Cheatum, and A. Kohen. Protein mass effects on formate dehydrogenase. *Journal of the American Chemical Society*, 139(48):17405–17413, 2017. doi: 10.1021/jacs.7b08359.

[263] M. Roca, J. J. Ruiz-Pernía, R. Castillo, M. Oliva, and V. Moliner. Temperature dependence of dynamic, tunnelling and kinetic isotope effects in formate dehydrogenase. *Physical chemistry chemical physics*, 20(40):25722–25737, 2018. doi: 10.1039/c8cp04244f.

[264] T. R. Gamble, F. F. Vajdos, S. Yoo, D. K. Worthylake, M. Houseweart, W. I. Sundquist, and C. P. Hill. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell*, 87(7):1285–1294, 1996. doi: 10.1016/s0092-8674(00)81823-1.

[265] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, and L. Serrano. The FoldX web server: An online force field. *Nucleic acids research*, 33(Web Server issue):W382–8, 2005. doi: 10.1093/nar/gki387.

[266] E. Krieger and G. Vriend. YASARA view - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics*, 30(20):2981–2982, 2014. doi: 10.1093/bioinformatics/btu426.

[267] G. Fischer, H. Bang, E. Berger, and A. Schellenberger. Conformational specificity of chymotrypsin toward proline-containing substrates. *Biochimica et biophysica acta*, 791(1):87–97, 1984. doi: 10.1016/0167-4838(84)90285-1.

[268] D. G. Schwark, M. A. Schmitt, and J. D. Fisk. Dissecting the contribution of release factor interactions to amber stop codon reassignment efficiencies of the *Methanocaldococcus jannaschii* orthogonal pair. *Genes*, 9(11), 2018. doi: 10.3390/genes9110546.

[269] C. Robichon, J. Luo, T. B. Causey, J. S. Benner, and J. C. Samuelson. Engineering *Escharichia coli* BL21(DE3) derivative strains to minimize e. coli protein contamination after purification by immobilized metal affinity chromatography. *Applied and environmental microbiology*, 77(13):4634–4646, 2011. doi: 10.1128/AEM.00119-11.

[270] K. R. Andersen, N. C. Leksa, and T. U. Schwartz. Optimized e. coli expression strain lobstr eliminates common contaminants from his-tag purification. *Proteins*, 81(11):1857–1861, 2013. doi: 10.1002/prot.24364.

[271] K. Sołtys and A. Ożyhar. Ordered structure-forming properties of the intrinsically disordered ab region of hrxrγ and its ability to promote liquid-liquid phase separation. *The Journal of steroid biochemistry and molecular biology*, 198:105571, 2020. doi: 10.1016/j.jsbmb.2019.105571.

[272] A. Micsonai, F. Wien, É. Bulyáki, J. Kun, É. Moussong, Y.-H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos. Bestsel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic acids research*, 46(W1):W315–W322, 2018. doi: 10.1093/nar/gky497.

[273] V. I. Tishkov and V. O. Popov. Catalytic mechanism and application of formate dehydrogenase. *Biochemistry (Moscow)*, 69(11):1252–1267, 2004. doi: 10.1007/PL00021765.

[274] D. J. Leibly, T. N. Nguyen, L. T. Kao, S. N. Hewitt, L. K. Barrett, and W. C. van Voorhis. Stabilizing additives added during cell lysis aid in the solubilization of recombinant proteins. *PloS one*, 7(12):e52482, 2012. doi: 10.1371/journal.pone.0052482.

[275] M. Malferrari, A. Nalepa, G. Venturoli, F. Francia, W. Lubitz, K. Möbius, and A. Savitsky. Structural and dynamical characteristics of trehalose and sucrose matrices at different hydration levels as probed by FTIR and high-field EPR. *Physical chemistry chemical physics*, 16(21):9831–9848, 2014. doi: 10.1039/c3cp54043j.

[276] P. K. Verma, A. Kundu, M. S. Puretz, C. Dhoonmoon, O. S. Chegwidden, C. H. Londergan, and M. Cho. The bend+libration combination band is an intrinsic, collective, and strongly solute-dependent reporter on the hydrogen bonding network of liquid water. *The journal of physical chemistry. B*, 122(9):2587–2599, 2018. doi: 10.1021/acs.jpcb.7b09641.

[277] N. K. Jain and I. Roy. Effect of trehalose on protein structure. *Protein science*, 18(1):24–36, 2009. doi: 10.1002/pro.3.

[278] R. Kozlowski, J. Zhao, and R. B. Dyer. Acceleration of catalysis in dihydrofolate reductase by transient, site-specific photothermal excitation. *Proceedings of the National Academy of Sciences of the United States of America*, 118(4):e2014592118, 2021. doi: 10.1073/pnas.2014592118.

[279] K. A. Reynolds, R. N. McLaughlin, and R. Ranganathan. Hot spots for allosteric regulation on protein surfaces. *Cell*, 147(7):1564–1575, 2011. doi: 10.1016/j.cell.2011.10.049.

[280] P. Singh, K. Francis, and A. Kohen. Network of remote and local protein dynamics in dihydrofolate reductase catalysis. *ACS catalysis*, 5(5):3067–3073, 2015. doi: 10.1021/acscatal.5b00331.

[281] J. W. McCormick, M. A. Russo, S. Thompson, A. Blevins, and K. A. Reynolds. Structurally distributed surface sites tune allosteric regulation. *eLife*, 10:e68346, 2021. doi: 10.7554/eLife.68346.

[282] C. Feid, L. Luma, T. Fischer, J. G. Löffler, J. Wachtveitl, A. Heckel, and J. Bredenbeck. *Iminothioindoxyl Donors with Exceptionally High Cross Section for Protein Vibrational Energy Transfer*. Unpublished manuscript, 2023.

[283] A. Kohen, R. Cannio, S. Bartolucci, and J. P. Klinman. Enzyme dynamics and hydrogen tunnelling in a thermophilic alcohol dehydrogenase. *Nature*, 399(6735):496–499, 1999. doi: 10.1038/20981.

[284] S. Caratzoulas and S. D. Schwartz. A computational method to discover the existence of promoting vibrations for chemical reactions in condensed phases. *The Journal of chemical physics*, 114(7):2910–2918, 2001. doi: 10.1063/1.1342817.

[285] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern. Intrinsic dynamics of an enzyme underlies catalysis. *Nature*, 438(7064):117–121, 2005. doi: 10.1038/nature04105.

[286] J. M. Young and J. H. Wang. The nature of binding of competitive inhibitors to alcohol dehydrogenases. *Journal of Biological Chemistry*, 246(9):2815–2821, 1971. doi: 10.1016/S0021-9258(18)62255-0.

[287] J. O. Alben and L. Y. Fager. Infrared studies of azide bound to myoglobin and hemoglobin temperature dependence of ionicity. *Biochemistry*, 11(5):842–847, 1972. doi: 10.1021/bi00755a025.

[288] S. Yoshikawa and W. S. Caughey. Infrared evidence of azide binding to iron, copper, and non-metal sites in heart cytochrome c oxidase. *Journal of Biological Chemistry*, 267(14):9757–9766, 1992. doi: 10.1016/S0021-9258(19)50157-0.

[289] R. Torres Martin de Rosales, M. Faiella, E. Farquhar, L. Que, C. Andreozzi, V. Pavone, O. Maglio, F. Nastri, and A. Lombardi. Spectroscopic and metal-binding properties of df3: an artificial protein able to accommodate different metal ions. *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry*, 15(5):717–728, 2010. doi: 10.1007/s00775-010-0639-9.

[290] A. G. Orpen, L. Brammer, F. H. Allen, O. Kennard, D. G. Watson, and R. Taylor. Appendix A: Typical interatomic distances in organic compounds and organometallic compounds and coordination complexes of the d– and f–block metals. In H.-B. Bürgi and J. D. Dunitz, editors, *Structure Correlation*, pages 752–858. Wiley-VCH, Hoboken, 2008. ISBN: 9783527290420. doi: 10.1002/9783527616091.app1.

[291] A. Creon, I. Josts, S. Niebling, N. Huse, and H. Tidow. Conformation-specific detection of calmodulin binding using the unnatural amino acid p-azido-phenylalanine (azf) as an ir-sensor. *Structural dynamics (Melville, N.Y.)*, 5(6):064701, 2018. doi: 10.1063/1.5053466.

List of Abbreviations

Aha — Azido-L-Homoalanine
amp — Ampicillin
ATM — Activated Tunneling Model
AuNP — Gold Nanoparticle
AzAla — β-(1-Azulenyl)-L-Alanine
BBO — β-Barium Borate
Cb — *Candida boidinii*
CD — Circular Dichroism
CLS — Center Line Slope
CNPhe — 4-Cyano-Phenylalanine
CNTrp — 4-Cyano-Tryptophan
cw — Continuous Wave
CypA — Cyclophilin A
DAD — Donor-Acceptor Distance
DHFR — Dihydrofolate Reductase
DMSO — Dimethyl Sulfoxide
DTNB — 5,5′-Dithiobis(2-Nitrobenzoic Acid)
E. coli — *Escherichia coli*
FDH — Formate Dehydrogenase
FTIR — Fourier Transform Infrared
FWHM — Full Width at Half Maximum
H-Bond — Hydrogen Bond
HEPES — 4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
HisTag — Histidine Tag
IC — Internal Conversion
IPTG — Isopropyl β-D-1-Thiogalactopyranoside
IR — Infrared
ITC — Isothermal Titration Calorimetry
IVR — Intramolecular Vibrational Redistribution
kan — Kanamycin
KIE — Kinetic Isotope Effect
LB — Lysogeny Broth
LDH — Lactate Dehydrogenase
LFM — Low-Frequency Mode
LN_2 — Liquid Nitrogen
MD — Molecular Dynamics
MIR — Mid-Infrared
Mj — *Methanocaldococcus jannaschii*
N_3^- — Azide Anion

NAD — Nicotinamide Adenine Dinucleotide
ncAA — Non-canonical Amino Acid
Ni-NTA — Nickel-Nitrilotriacetic Acid
NIR — Near-Infrared
NMR — Nuclear Magnetic Resonance
N_3Ala — Azidoalanine
N_3Phe — 4-Azido-Phenylalanine
OD_{600} — Optical Density at 600 nm
OPA — Optical Parametric Amplifier
pdb — RCSB Protein Data Bank
Pi — Phosphate Buffer
pNA — p-Nitroaniline
PPIase — Peptidyl-Prolyl Cis-Trans Isomerase
PTFE — Polytetrafluoroethylene
RF — Release Factor
RPV — Rate-Promoting Vibrations
SCN — Thiocyanate
SCS — Stop Codon Suppression
SDS — Sodium Dodecyl Sulfate
SEM — Standard Error of the Mean
spec — Spectinomycin
SPI — Selective Pressure Incorporation
SPPS — Solid Phase Peptide Synthesis
str — Streptomycin
Suc — Succinic Acid
TB — Terrific Broth
TCEP — Tris-(2-Carboxyethyl)phosphine
TEV — Tobacco Etch Virus Protease
TFE — Trifluoroethanol
THF — Tetrahydrofuran
Ti:Sa — Titanium-Sapphire
Tm — *Thermotoga maritima*
TRIR — Transient Infrared
Tris — 2-Amino-2-(Hydroxymethyl)Propane-1,3-diol
TrpB — Tryptophan Synthase β-Subunit
TRS — Tunneling Ready State
UV — Ultraviolet
VET — Vibrational Energy Transfer
VG — Val123Gly
Vis — Visible Light
WT — Wild Type
YAG — Yttrium Aluminum Garnet

List of Figures

List of Tables