Using Azobenzene Photocontrol to Set Proteins in Motion

Olga Bozovic, Brankica Jankovic, *Peter Hamm

Department of Chemistry, University of Zurich, Zurich, Switzerland

*peter.hamm@chem.uzh.ch

Controlling the activity of proteins with azobenzene photoswitches is a potent tool for manipulating their biological function. With the help of light, one can change e.g. binding affinities, control allostery or temper with complex biological processes. Additionally, due to their intrinsically fast photoisomerisation, azobenzene photoswitches can serve as triggers to initiate out-of-equilibrium processes. Such switching of the activity, therefore, initiates a cascade of conformational events, which can only be accessed with time-resolved methods. In this Review, we will show how combining the potency of azobenzene photoswitching with transient spectroscopic techniques helps to disclose the order of events and provide an experimental observation of biomolecular interactions in real-time. This will ultimately help us to understand how proteins accommodate, adapt and readjust their structure to answer an incoming signal and it will complete our knowledge of the dynamical character of proteins.

I. INTRODUCTION

Proteins have a dynamical nature that governs their biological function, and understanding the underlying processes ultimately allows for a fine-tuning of the affinity, controlling allostery, or regulating the activity of target proteins. When discussing the dynamical nature of proteins, it is essential to distinguish between equilibrium and non-equilibrium dynamics. Equilibrium dynamics of proteins lie in the fact that native proteins, even when fully folded, are not rigid entities. The equilibrium dynamics of proteins is reflected by an ensemble of conformational sub-states and the continuous fluctuation between them. The prototype example illustrating the importance of protein fluctuations is myoglobin. The heme is deeply buried inside the protein and oxygen would not be able to access it if the structure of myoglobin would be the static X-ray representation. Studying the equilibrium dynamics of proteins is mostly the realm of NMR spectroscopy (see Table I).

Non-equilibrium dynamics, on the other hand, describe how proteins adapt and respond to an incoming event. This event can be the interaction with a signalling molecule or another protein, or the response to an allosteric signal. In order to experimentally follow the non-equilibrium adaptation of proteins to a signal, one needs an instantaneous perturbation that initiates a response of the system while following a time-dependent observable. The non-equilibrium dynamics portray how a system behaves in real-time, provides us with a "molecular movie", and broadens our in-equilibrium picture of a molecular mechanism. The static representation of biomolecular recognition, protein allostery or enzymatic activity, as for instance the "lock and key" mechanism, have long been known to be inadequate, given that they do not represent the natural dynamical fluidity of proteins and their interacting partners. Even for an event as "simple" as small ligand binding to a protein, many complex, transient and sometimes subtle redistributions of interactions and conformational changes do occur.

Complex biological processes, such as enzymatic and metabolic cascades within diverse signaling pathways, involve the communication between many players and as such require many underlying layers of control. Consequently, there are several mechanisms for biological regulation of protein’s activity, where the targets for regulation can be an orthosteric or an allosteric site. Orthosteric regulation represents a direct regulation of biomolecular recognition, i.e., of protein-ligand interactions at the active/binding site. It is manifested by the competitive binding of different ligands or by e.g., a reversible phosphorylation that alters the binding affinity. Allosteric regulation, on the other hand, implies binding of a signaling molecule on a site different from an active site. The communication between two non-overlapping sites regulates the activity of the protein by an "allosteric signal", the nature of which is hotly debated. Implementation of time-resolved spectroscopic techniques is of paramount importance for a full dynamical description of such biological processes.

Light triggering is an attractive way to initiate non-equilibrium dynamics. Such experiments require two ingredients: a structure-sensitive spectroscopic method with appropriate time resolution as well a photoswitchable component in the molecular system under study. As for the first, time-resolved IR spectroscopy (which is what we will focus on here, see Box 1) and time-resolved X-ray diffraction (XRD) are the methods of choice. Time-resolved XRD is based on synchrotrons, whose X-ray light is pulsed with a typical pulse duration of 100 ps and more recently on hard X-ray free electron lasers. Table I compares the different methods that are used to study the structural dynamics of proteins. It might seem that time-resolved XRD is the one perfect method on this list, however, it requires the protein system to be crystalized, which very often is a bottleneck (X-ray free-electron lasers have relaxed that situation a bit since "only" nanocrystals are needed). Furthermore, the availability of such instru-
TABLE I. Comparison of various structure sensitive spectroscopic methods to study the dynamics of proteins.

| Sample | Solution (≈ 100 µM) | Crystal | Solution (≈ 1 mM) | Time-resolved IR | Time-resolved XRD |
|--------|---------------------|---------|------------------|------------------|-------------------|
| Overall structure resolution power | high | low | high | high |
| Local structure resolution power | high | very high | high | high |
| Time-range for equilibrium dynamics | ≳100 fs–1 s | ≳100 fs–10 ps | ≳1 ps | ≳100 fs or ≳100 ps |
| Time-range for non-equilibrium dynamics | ≳100 ms | ≳1 ps | ≳100 fs or ≳100 ps |
| Availability | high | high | low |

a in connection with an assigned local mode and/or a specific IR label.
b free electron laser vs synchrotron, respectively.
c requires a large scale facility with limited access, such as a hard-X ray free electron laser or a synchrotron.

The vast majority of biological processes, however, work without light. Indirect photocontrol can be achieved with caged compounds, or pH jumps. On the other hand, Nature’s strategy of using light for controlling biochemical processes may serve as an inspiration for the more direct design of artificially photocontrollable proteins by involving photoswitchable molecules (or photoswitches for short). Photoswitches are small molecules that undergo a fast light-induced isomerization by properly incorporating them into a protein system, one may modulate its function or activity directly, or one may manipulate the binding affinity of a photoswitchable ligand that binds to a protein.

Unlike naturally abundant photo-responsive proteins, artificial photocontrol allows for a rational design, depending on the effect one wants to achieve, which can be applied to virtually any protein system. Photoswitches are versatile and fully customizable, and the induced effects are often larger than those achieved with e.g. temperature or pH jumps.

Owing to many attractive features, the most widely used class of photoswitches are derivatives of azobenzene (see Box 2). They are relatively small molecules, chemically very stable, and isomerize around the central N=N bond with high quantum yield. The optical spectra of the two isomers, cis and trans, differ sufficiently to be able to steer the molecule to both states with high purity. The light-induced isomerisation proceeds through...
Box 2 Azobenzene-based photoswitches

Azobenzene-based molecules are the most widely used molecular photoswitches for controlling biological function. Unlike many other approaches, azobenzene photoswitches undergo a reversible photoisomerisation upon illumination, and many cycles of photo-switching can be achieved. The UV/Vis absorption spectra of the trans and cis isomers are different from each other (panel a), most prominently at \( \approx 360 \) nm with the strong \( \pi-\pi^* \) transition for the trans-state (the exact position of the absorption maximum varies a bit depending on azobenzene derivative), and essentially no absorption in the cis-state. In contrast, the \( \pi-\pi^* \) transition around 450 nm exists for both isomers.

The trans-isomer of azobenzene is by \( \approx 10 \) kJ/mol more stable than the cis-isomer and therefore dominates if the molecule is kept in the dark. Upon illumination of the \( \pi-\pi^* \) transition, the photoswitch undergoes a trans-to-cis isomerisation, eventually resulting in a cis-sample with purities of \( \geq 85\% \). The trans-sample can then be re-established by keeping it in the dark and allowing for the equilibrium to be restored by thermal back-isomerisation, which happens on a minute to hour timescale. The sequence of spectra shown in the figure represents this process, each taken at a different time after preparing a cis-sample.

Time-resolved experiments with azobenzene controlled systems can be performed in both directions. For trans-to-cis switching, one starts from a dark-adapted sample and excites either the \( \pi-\pi^* \) or the \( n-\pi^* \) transition of trans-azobenzene with a short laser pulse. For cis-to-trans switching, one first accumulates the cis-state by pre-illuminating the \( \pi-\pi^* \) transition of trans-azobenzene, and then excites the \( n-\pi^* \) transition of cis-azobenzene with a short laser pulse. For time-resolved experiments, it is important to have only one photoswitch per protein. If more than one are used, one could not guarantee that all switch with the one pump laser pulse, since the excitation probability and the isomerisation quantum yield are smaller than 1.

Woolley and coworkers have pioneered a chemical concept, with which azobenzene photoswitches can selectively be linked to two cysteines of a peptide or a protein in a simple, one-step reaction (panel b). Derivatives of the azobenzene photoswitch have been designed with additional -SO\(_3^-\) groups that render it water-soluble.

A major motivation to incorporate azobenzene photoswitches into proteins is to photocontrol biological activity. Examples reported in literature range from the regulation of enzymatic activity, ion-channels and receptors all the way to cytoskeleton regulation, cell-cell adhesion, as well as in vivo studies of various physiological processes. But azobenzene photocontrol can also serve as a trigger to set proteins in motion, and this is where the full potential of this type of photocontrol comes in place. Exploiting the practically instantaneous isomerization of azobenzene photoswitches incorporated into proteins and peptides, and combining it with time-resolved spectroscopic techniques, can reveal what happens to proteins once they are pushed out of the equilibrium. This combination will be the main topic of the present review article. In addition, we will discuss relevant applications of photocontrollable systems and how rational design can be used to modulate biological structure and function. Ultimately, we will reconcile the two aspects of protein dynamics and show how fundamental questions can be addressed, which require a time dimen-

![Absorbance, OD](image-url)
The first realizations of photoswitchable proteins linked the azobenzene moiety to a single site or incorporated an azobenzene-moiety bearing amino acid phenylazophenylalanine. This is called “monofunctional control” and illustrated in Fig. 1a. The azobenzene-moiety interferes sterically with the natural agonist and thereby modulates the activity of the protein system. Furthermore, azobenzene derivatives mimicking an amino acid, with a carboxyl group on one ring and an amino group on the other, have been incorporated directly into a peptide backbone by modified peptide synthesis, in essence introducing a kink in the backbone, (c) ligand switching, as well as bifunctional (d) α-helix switching and (e) cross-linking different secondary structure motifs of a protein.

The concept of azobenzene photoswitching is in fact much more versatile, as one can control any secondary structure element, such as β-sheets or loops. Furthermore, two secondary structure motifs within a protein have been linked by a photoswitch in order to modulate their relative distance and/or orientation (Fig. 1d). Control of biological activity has also been achieved via photocontrollable small molecule ligands or inhibitors that interact with a target protein (Fig. 1).

III. TIME-RESOLVED STUDIES ON SMALL PEPTIDES: SIMPLE MODEL SYSTEMS

Time-resolved studies of azobenzene controlled biomolecules are quite scarce, and began with investigating the folding dynamics of small peptides. In a time-resolved experiment, one switches the azobenzene photoswitch either from the cis- to the trans-state, or vice versa, and observes the response of the photoswitch and the peptide by transient Vis or transient IR spectroscopy, see Box 1. In the first attempt to understand the conformational dynamics of azobenzene-based peptides, Ref. 102 employed a small cyclic peptide with the photoswitch embedded in the peptide backbone. The ultrafast time-resolved response was followed in the UV/Vis spectral region, using the azobenzene molecule as a spectroscopic observable. The study showed that the majority of conformational rearrangement of the photoswitch happens within the first 50 ps after excitation. In follow-up work on the the same cyclic peptide, the peptide backbone response has been monitored employing transient IR spectroscopy of the amide I band. This enabled the observation of the subsequent peptide relaxation, which extends up to 16 ns.

With the help of time-resolved optical rotatory dispersion, Ref. 110 studied the folding dynamics of a 16 amino acid long α-helical peptide. By cross-linking the photoswitch between two cysteines separated by 11 amino acids, the system had a higher helical propensity in the trans-state of the photoswitch, while isomerisation to cis-state lead to disruption (unfolding) of the helix (see Fig. 1). It was seen that the forced unfolding of the helix occurs within 55 ns. The same α-helical construct was also investigated by transient IR spectroscopy. Taking advantage of the photoswitching reversibility, both the
folding and unfolding direction has been investigated and that in a site-selective manner with the help of C$^{13}$=O$^{18}$ isotope labelling. The typical timescale of helix folding is significantly slower, 300 ns to 3 $\mu$s depending on temperature.

The folding and unfolding of a tryptophan zipper analogue containing a $\beta$-hairpin has been addressed as well with the construct shown in Fig. 1. Photoswitching initiated a conformational transformation of the system between a $\beta$-hairpin and an unfolded hydrophobic cluster. Unfolding takes a few nanoseconds only, while the reverse folding of this system requires more sampling of the conformational space and occurs on a 30 $\mu$s timescale.

The works on small peptides not only contributed to a better understanding of the protein folding problem, but also paved the way for subsequent studies on more complex artificially photocontrollable protein systems, which will be discussed in the next Chapter.

IV. TIME-RESOLVED STUDIES OF PROTEIN SYSTEMS

Larger proteins typically do not completely fold/unfold upon photoswitching, since the configurational change of the azobenzene moiety is too small to initiate such a large transition (Ref. 116 is an exception in this regard). Rather, one changes a local structure element and thereby the function of the protein system. There are two approaches to artificially photocontrol a protein and subsequently investigate it with time-resolved methods. First, one can decide to photocontrol a binding partner for the system in question (Fig. 2b,d). These partners usually are relatively short peptide ligands, which can be synthesized using standard solid state peptide synthesis and subsequently cross-linked with an azobenzene molecule. This approach is significantly less demanding than the second one, where a photoswitch molecule is inserted postranslationally to a full length protein (Fig. 2a,c). Deciding on the method depends on the general properties of the system as well as the scientific questions posed. Furthermore, one needs to keep in mind that proteins need water as solvent, in contrast to the peptide systems discussed in the previous Chapter. This typically requires a photoswitch that is made water-soluble by adding polar groups to the azobenzene rings.

Time-resolved studies of artificially photocontrollable proteins focused on two classes of proteins, PDZ domains (Fig. 2a-c) and the RNase S complex (Fig. 2d). PDZ domains are ubiquitous protein interaction domains found in a wide range of organisms from bacteria to mammals (Postsynaptic density protein/Drosophila disc large tumour suppressor/Zonula occludens-1 protein). PDZ domains are relatively small protein domains that consist of approximately 100 amino acids. They share a common fold and are usually integral parts of larger, multidomain proteins. Given that PDZ domains are an important class of protein interaction domains, they have been extensively studied in the context of protein allostery and protein-ligand binding by time-resolved spectroscopy. (a) PDZ2 domain with the binding-groove cross-linked by a photoswitch (b) PDZ2 domain with a photoswitchable peptide ligand (c) PDZ3 domain with a photoswitchable $\alpha$-helix and (d) the RNase S complex with a photoswitchable S-peptide.

FIG. 2. Various artificially photoswitchable proteins that have been studied in the context of protein allostery and protein-ligand binding by time-resolved spectroscopy. (a) PDZ2 domain with the binding-groove cross-linked by a photoswitch (b) PDZ2 domain with a photoswitchable peptide ligand (c) PDZ3 domain with a photoswitchable $\alpha$-helix and (d) the RNase S complex with a photoswitchable S-peptide.
complex consists of the S-protein and the shorter S-peptide fragment. The isolated S-protein does not possess any catalytic activity, given that the second histidine residue necessary for hydrolysis of RNA is located on the S-peptide. Nevertheless, when co-dissolved, S-protein and S-peptide re-associate with nanomolar affinity and restore the full enzymatic activity. Another peculiarity of this system lies in the fact that the isolated S-protein preserves its folded structure to a certain extent, as evidenced from circular dichroism spectra (no NMR spectra or crystal structures are available), while the isolated S-peptide is an unstructured random coil. Upon re-association of the S-peptide with the S-protein, the RNase S complex restores a structure that is practically the same as that of RNase A with the S-peptide being α-helical. The RNase S system is a widely used model system to study protein-peptide binding mechanisms, in particular in the context of “induced fit” vs “conformational selection”.

Utilizing these two protein systems as examples, we will discuss in the following how proteins respond to a perturbation induced by a photoswitch, determine the speed of an allosteric signal, and study the full sequence of events during ligand unbinding.

A. Protein Response

As a first step towards understanding the non-equilibrium allosteric nature of proteins, Ref. presented a PDZ2 domain, in which the photoswitch was linked across its binding groove (Figs. [1] and [2]). Isomerization of the photoswitch forces the opening or closing of the binding groove. Albeit an artificial construct by default, it was shown by solving NMR structures for the photoswitchable protein in the cis- and the trans-state that the perturbation closely mimics the structural change of a native PDZ2 domain upon ligand binding/unbinding. Transient infrared spectroscopy revealed three major phases of the overall process. Initial photo-
excitation is followed by fast photoisomerization and heat dissipation on a 10 ps timescale, an effect widely seen in these type of experiments.\textsuperscript{140,149} Subsequently, the photoswitch relaxes structurally on a 10 ns timescale, as judged from a vibrational mode that is localized on the linker between the azobenzene moiety of the photoswitch and the protein backbone. The structural relaxation slows down when attaching the photoswitch to the protein, owing to the strain imposed by the protein, which cannot accommodate the structural change of the photoswitch immediately. This phase was thus assigned to the perturbation of the binding groove of the PDZ2 domain. Finally, the perturbation of the binding groove propagates through the whole protein within 10 µs, as deduced from the amide I band, which is an extraordinarily sensitive reporter of protein structure.\textsuperscript{32}

Exactly this propagation of a perturbation through the protein was investigated more closely in a follow-up study\textsuperscript{52} utilizing a less artificial PDZ2 construct (Fig. 2b). That is, the PDZ2 domain was kept intact in this case, while a newly introduced peptide ligand was made photoswitchable. \textit{Cis-to-trans} isomerisation of the peptide ligand stabilizes its \textit{β}-strand structure in the binding groove of the PDZ2 domain and thus increases its binding affinity by 5 fold. It also injects an allosteric signal into the PDZ2 domain, that has been investigated by transient infrared spectroscopy in connection with isotope labelling of the whole protein. Fig. 3 shows a typical example of such an experiment, plotting the absorption change of the protein at various probe frequencies within the amide I band as a function of time; the latter on a logarithmic scale to be able to cover a wide range of timescales. Fig. 4 shows the corresponding timescale analysis (see Box 3 for an explanation), emphasizing the very rich information content of such an experiment.

Molecular dynamics (MD) simulations were performed to understand this complex response.\textsuperscript{53} They revealed that PDZ2 exists in a couple of conformational sub-states, in essence reflecting what is known as the “rugged energy landscape” of proteins.\textsuperscript{141–143} Photoswitching of the ligand redistributes the population between those sub-states to a relatively small extent, in accordance with an emerging new view of allostery.\textsuperscript{111} The overall change of protein structure is only 0.3 Å, illustrating the extraordinary sensitivity of transient IR spectroscopy, which is able to detect such small changes.

The kinetic consequence of the rugged energy picture can be understood in terms of a Markov State Model (MSM), which is explained in more detail in Box 4. Most importantly, one expects a relatively small number of discrete timescales, the “implied” timescales of the MSM, as indeed observed in the averaged dynamical content of the transient IR data shown in Fig. 3c. MD simulations of the first PDZ2 construct discussed above (Fig. 2b) revealed qualitatively the same behaviour.\textsuperscript{112,141} and we have also observed it in a different PDZ domain\textsuperscript{53} as well as in a naturally photoswitchable protein.\textsuperscript{146} It thus appears to be an universal property of photoswitchable proteins.

### B. Allosteric Signal

The PDZ3 domain differs from the rest of the PDZ family by an additional \textalpha-helix at its C-terminus.\textsuperscript{111,112} This helical extension is distant from the binding groove and was shown to be an allosteric element of the protein.\textsuperscript{122,123,124} In Ref. 68, a photoccontrollable PDZ3 protein was designed by cross-linking the photoswitch to that \textalpha3-helix (Fig. 2c). Isomerization of the photoswitch leads to the perturbation of the helical structure and changes the binding affinity for a peptide ligand up to 120-fold, depending on temperature. Moreover, the perturbation of the helix not only changes the affinity in the binding groove, but the binding of a ligand also speeds up the thermal \textit{cis-to-trans} isomerization rate of the photoswitch, introducing the concept of an “allosteric force”.

### Box 3 Timescale Analysis and Dynamical Content

The timescales contained in a transient data set such as that of Fig. 3a are determined by fitting each kinetic trace at probe-frequency \( \omega_i \) to a multienzymeal function:

\[
S(t, \omega_i) = a_0 - \sum_k a(\omega_i, \tau_k) e^{-t/\tau_k},
\]

where a maximum entropy method has been applied for a regularisation of the otherwise ill-posed inverse Laplace transformation.\textsuperscript{137–139} In this fit, the timescales \( \tau_k \) were fixed and equally distributed on a logarithmic scale with 10 terms per decade, while the amplitudes \( a(\omega_i, \tau_k) \) are the free fitting parameters. In Ref. 140 we have compared this approach to more conventional global fitting. The amplitudes \( a(\omega_i, \tau_k) \) are called “timescale spectra” or sometimes “dynamical content”\textsuperscript{53}.

Fig. 3b exemplifies the concept for one particular probe frequency \( \omega_i \). On a logarithmic time axis, an exponential process exhibits a sigmoidal shape. Whenever such a sigmoidal step occurs in the data, which can be up or down, the corresponding lifetime spectrum shown in blue reveals a peak. Fig. 3c exemplifies the concept for one particular probe-frequency \( \omega_i \). The timescales contained in a transient data set such as that of Fig. 3d. MD simulations as indeed observed in the averaged dynamical content of the transient IR data shown in Fig. 3b. MD simulations of the first PDZ2 construct discussed above (Fig. 2b) revealed qualitatively the same behaviour.\textsuperscript{112,141} and we have also observed it in a different PDZ domain\textsuperscript{53} as well as in a naturally photoswitchable protein.\textsuperscript{146} It thus appears to be an universal property of photoswitchable proteins.

The PDZ3 domain differs from the rest of the PDZ family by an additional \textalpha-helix at its C-terminus.\textsuperscript{111,112} This helical extension is distant from the binding groove and was shown to be an allosteric element of the protein.\textsuperscript{122,123,124} In Ref. 68, a photoccontrollable PDZ3 protein was designed by cross-linking the photoswitch to that \textalpha3-helix (Fig. 2c). Isomerization of the photoswitch leads to the perturbation of the helical structure and changes the binding affinity for a peptide ligand up to 120-fold, depending on temperature. Moreover, the perturbation of the helix not only changes the affinity in the binding groove, but the binding of a ligand also speeds up the thermal \textit{cis-to-trans} isomerization rate of the photoswitch, introducing the concept of an “allosteric force”.

The PDZ3 domain differs from the rest of the PDZ family by an additional \textalpha-helix at its C-terminus.\textsuperscript{111,112} This helical extension is distant from the binding groove and was shown to be an allosteric element of the protein.\textsuperscript{122,123,124} In Ref. 68, a photoccontrollable PDZ3 protein was designed by cross-linking the photoswitch to that \textalpha3-helix (Fig. 2c). Isomerization of the photoswitch leads to the perturbation of the helical structure and changes the binding affinity for a peptide ligand up to 120-fold, depending on temperature. Moreover, the perturbation of the helix not only changes the affinity in the binding groove, but the binding of a ligand also speeds up the thermal \textit{cis-to-trans} isomerization rate of the photoswitch, introducing the concept of an “allosteric force”.

\[
D(\tau_k) = \sqrt{\sum_i a(i, \tau_k)^2}.
\]
Panels (a) and (b) sketch the rugged energy landscape of a protein. That is, proteins do not adopt a single rigid structure, rather an ensemble of closely related, but not identical structures, each of which characterized by a local minimum of the free energy surface of the protein. According to an emerging new view, the allosteric signal is related to a remodelling of the free energy landscape, i.e., the depths of the various free energy minima change upon an allosteric signal (compare panels a and b), and thereby shift the populations between substates.

The kinetic consequence of that picture can be understood in terms of a Markov State Model (MSM), which in essence describes a set of states connected by microscopic rates $k_{ij}$ (see panel c) that are determined by the barriers connecting the local minima $i$ and $j$. The corresponding kinetic equation is:

$$\frac{dp}{dt} = K \cdot p,$$

where $p$ is a vector whose components $p_i$ contain the populations of state $i$, and $K$ a kinetic matrix:

$$K = \begin{pmatrix}
-k_{12} & k_{13} & k_{31} & \cdots \\
 k_{12} & \ddots & \ddots & \ddots \\
 k_{13} & \ddots & \ddots & \ddots \\
 \vdots & \ddots & \ddots & \ddots \\
 \end{pmatrix},$$

exemplifying here only the microscopic rates shown in panel (c). The free-energy difference $\Delta F_{ij}$ between pairwise two states $i$ and $j$ is related to the microscopic rates by detailed balance:

$$k_{ij} = e^{-\Delta F_{ij} / k_B T}.$$

The kernel (nullspace) of the kinetic matrix $K$, setting the left-side of Eq. 3 to zero, determines the population $p_{eq}$ in equilibrium. Initially, the protein will be in equilibrium with respect to a kinetic matrix $K_b$ (where the subscript $b$ stands for “before”). Upon an allosteric signal (here upon light-triggering), the free energy surface changes and correspondingly also the kinetic matrix $K_a$ (for “after”), along the lines of Eq. 5. Immediately after the trigger, the populations will still be same as before, hence the system all of the sudden will be out of equilibrium. It will subsequently equilibrate according to Eq. 3 with the kinetic matrix $K_a$.

The dynamical content (see Box 3) of a transient experiment will not contain the microscopic rates $k_{ij}$, but rather the eigenvalues of the kinetic matrix $K_a$, which are sometimes called the “implied timescales” of the MSM. The number of implied timescales equals the number $n$ of states in the MSM minus 1, which is a significantly smaller number than that of microscopic rates (of the order $n^2$). The relatively small number of implied timescales facilitates the discreteness of the dynamical content shown in Fig. 3d. It is important to stress that the same set of implied timescales is revealed, regardless of the observable one is looking at, but different observables may weigh the various implied timescales differently. For example, each probe frequency in Fig. 3b reveals a different response, yet the averaged dynamical content in Fig. 3d emphasizes that there is a common set of timescales.

This is allostery in its literal sense, as one may consider the cis-to-trans isomerization of the azobenzene moiety a chemical reaction. The PDZ3 domain thus is arguably the smallest fully allosteric protein with a well-defined allosteric and effector site.

The unique properties of the photoswitchable PDZ3 domain make it an ideal candidate to shed light on the very nature of the allosteric signal. Employing a combination of transient UV/Vis and IR spectroscopies, the propagation of the allosteric signal could be reconstructed. Similar to the kinetics of the forced unfolding of isolated helices discussed above, the $\alpha$-helix unfolds (partially) on a 5 ns timescale after photoswitching. The protein responds to that perturbation on multiple time-scales up to 10 $\mu$s, again representing the ruggedness of its free energy landscape. However, by double-difference spectroscopy, the response of the peptide ligand could be singled out, revealing one dominating timescale of 200 ns, which has been attributed to the speed of the allosteric signal within the protein.
C. Ligand Unbinding

Biomolecular recognition is a major mechanism of protein regulation. Protein ligand interactions are fundamental in any biochemical process and have been studied in great depth for a plethora of systems. The RNase S complex (Fig. 2) represents one such system, which has been widely used as a model for studying coupled binding and folding of the S-peptide during the interaction with the S-protein. Various photoswitchable variants of a S-peptide have been designed with different spacings and positions of the anchoring points of the photoswitch, and their binding affinities to the S-protein have been measured with CD spectroscopy. If the spacing of the two anchoring points is 9 (S-pep(6,15)), photoswitching has virtually no effect on the binding affinity (Fig. 3), in accordance with the observation for isolated helices that this spacing is the dividing point between the two regimes, in which either the cis or the trans state stabilizes the helix. For shorter spacings, the helical structure of the S-peptide is stabilized in the cis-state of the photoswitch, facilitating binding to the S-protein. The binding affinity changes by more than 20 fold when the spacing is 4 (S-pep(6,10), see Fig. 4), in the same order as observed for other protein constructs.

S-pep(6,13) with spacing 7 is unique in this regard, as it binds in the cis-state with reasonable binding affinity, while no specific binding could be detected in the trans-state (Fig. 4). Molecular dynamics (MD) simulations confirmed that the underlying mechanism for the modulation of the binding affinity does lie in the disruption of the helical content of the S-peptide setting off the possibility of a rational design of this type of molecular systems.

The research of protein-ligand binding circles around the long-lasting debate of two limiting scenarios - induced fit vs conformational selection. The necessary prerequisite for such a classification rests on the temporal sequence of binding event vs conformational changes of the binding partners. The on/off behaviour of S-pep(6,13) made it an interesting model system to study the complete sequence of events during the unbinding of the ligand by transient fluorescence and IR spectroscopy. Upon cis-to-trans switching, unbinding of the ligand is essentially a barrier-less process and proceeds in the fastest possible manner, revealing the “speed limit” of ligand unbinding. As such, it is less of a rate-limiting step and enables the observation of slower relaxation processes of the protein.

Ref. uncovered the full sequence of events during unbinding. Transient IR spectroscopy revealed that the helical S-peptide, still inside the binding groove, unfolds within 20 ns; a similar timescale as that observed for isolated α-helices or for the α3-helix in the PDZ3 system. The peptide remains bound, either in the binding pocket or to other parts of the protein, for quite some time and leaves only after about 300 µs; four orders of magnitude later in time. The protein subsequently responds to that on a 3 ms timescale and an even slower process that could not be observed with the experimental setup of Ref. Encompassing all the underlying events before, during and after the peptide unbinding, we offered a comprehensive explanation of the RNase S binding mechanism. That is, from the perspective of the S-peptide, induced fit seems to be the predominant mechanism, since it can explore its conformational space on timescales much faster (20 ns) than it remains bound to the protein, even for a ligand that has been designed to unbind as quickly as possible. To understand this argument, one has to keep in mind that the experiment of Ref. investigated peptide unbinding, while the induced fit scenario argues from a binding. On the other hand, the behaviour of the S-protein is better explained as conformational selection, as its conformational dynamics is slower than unbinding. These studies showcased the importance of the non-equilibrium approach to unequivocally discover the sequence of events on all relevant timescales and ultimately clarify the binding mechanisms.
FIG. 5. Full sequence of events of allosteric propagation and ultimate ligand unbinding together with their typical timescales.

D. Typical Timescales

Compiling the results obtained from the experiments we just discussed, the picture shown in Fig. 5 emerges. The results have been obtained for different protein systems (Fig. 2), not all of them revealing all processes shown in Fig. 5. Nevertheless, the different protein systems share quite a few common features with respect to the type of response and their typical timescales, and we can put them into the unified representation of Fig. 5.

The isomerisation process and the subsequent dissipation of the released heat universally occurs on a subnanosecond timescale. The photoswitch then perturbs the local structure element to which it is directly bound on a roughly 10 ns timescale, e.g. the opening of the binding groove of the PDZ2 construct shown in Fig. 2c, or the partial unfolding of an α-helix in the PDZ3 domain (Fig. 2d) or the RNase S complex (Fig. 2a). The perturbation subsequently propagates through the whole protein on multiple timescales, ranging from ≈10 ns to ≈10 µs, a characteristics we have observed for all protein systems we have studied so far. If it is an allosteric protein, that process includes the allosteric signal (found to reach the binding groove in the case of PDZ3 system within 200 ns). Ligand unbinding occurs only after a few 100 µs in the fastest possible cases, such as S-pep(6,13) bound to the S-protein, and the protein adapts to that change yet another time on a few millisecond timescale and beyond. To the best of our knowledge, this is the most comprehensive picture as of today of what is happening inside a protein during allosteric signalling. It explains how proteins adapt and respond to an allosteric signal, which ultimately leads to ligand unbinding.

V. OUTLOOK: AZOBENZENE PHOTOCONTROL OF BIOLOGICAL ACTIVITY

So far, we concentrated on structural aspects, but at the end of the day, the goal is to control biological activity. Many examples of such constructs have been described in literature, which all could be potential candidates for time-resolved studies, as they incorporate the necessary ultrafast switch - azobenzene. We will discuss a few of those examples more closely in the following.

Some of the earliest attempts to photocontrol enzyme activity was conducted in fact on the RNase S complex, albeit with a monofunctional construct as sketched in Fig. 1a. To that end, an unnatural azobenzene-bearing amino acid (phenylazophenylalanine) was incorporated into the S-peptide by means of a standard solid phase peptide synthesis at several positions, modulating the binding affinity between the S-peptide and S-protein and thereby the enzymatic activity of the RNase S complex. However, the effect was modest with less than a 5 fold difference in activity, supposedly due to the modest impact of a photoswitchable side chain that does not affect the backbone structure of the peptide. Bifunctional cross-linking as in Fig. 2d is expected to have a bigger effect on the enzymatic activity.

An illustrative example of such a bifunctional control has been demonstrated in the in vitro study of a photoswitchable calcium-binding protein, cadherin. In total 11 versions were synthesized, differing by the relative positions of the two cysteines to which the photoswitch was linked, and the most potent one was further characterized. A 18-fold difference in calcium binding affinity was achieved when affecting the structure of calcium binding loop. Trans-to-cis isomerization of the photoswitch translated directly into the dimerization propensity of cadherin. As cadherins play a direct role in cell-cell contacts, this study shows the potential for a control of cellular adhesion. Similar design strategies were applied...
FIG. 6. Photoswitchable ATPase, targeting the hinge motion of the α and β-subunits. The golden spheres represent the atoms of residues which are mutated to cysteines in the αA380C/βV409C variant, which revealed the biggest difference in activity. The picture has been produced from pdb entries 5dn6 and 3oaa.

in Ref. 62 with a motivation to control the enzymatic activity of PvuII restriction endonuclease.

In Ref. 66 an even larger protein construct was targeted. The ATPase molecular machinery was cross-linked with the azobenzene photoswitch between the α and β-subunits of the F_1 part of ATP synthase (see Fig. 6). The influence of light on the ATP-hydrolyzing activity was tested for 4 mutants with different positions of the cysteine pair. The highest effect was achieved when cis-to-trans photoswitching induces a large difference in the flexibility of the active site, enabling the opening and closing of the F_1 subunit.

In the in vivo study of Ref. 63, the biological activity of tubulin, the protein that constitutes the microtubule cytoskeleton, was controlled indirectly by modifying its polymerization inhibitor. To that end, an analogue of the small molecule inhibitor based on an azobenzene substituted with functional methoxy groups was used, important for inhibitory activity. Only in the cis-conformation of the azobenzene, methoxy groups were spatially arranged in an appropriate fashion to mimic the inhibitor. Upon switching to the trans-state, the inhibitory activity was abolished. This compound was applied to the cell culture and a complete “on/off” switching of the microtubule assembly was observed, allowing for a control of mitosis and cell death with single-cell precision.

The numerous efforts to control biological activity by azobenzene switches could be enriched by including a time-resolution, i.e., see how the time-dependent structural changes affect the activity of the proteins. All the processes illustrated above could be followed by time-resolved methods upon ultrafast photoswitching, addressing what happens with the protein in real-time. This would be a way to connect the hierarchy of timescales with the mechanisms of e.g. enzyme catalysis, microtubular assembly, endocytosis and many other biological processes. Presumably, as the size of a protein system becomes larger even slower timescales will become relevant, but we expect that fast timescales remain equally important. Connecting time-resolved studies with biological activities would go significantly beyond our current understanding of the structure/function paradigm.

VI. CONCLUSION

Understanding processes as complex as the dynamical nature of proteins requires a joint effort between several scientific fields. Introducing time-resolved techniques in various spectral regions is necessary to extend in-equilibrium studies of protein dynamics. Besides spectroscopic techniques, as discussed here, we have seen a tremendous advance in ultrafast time-resolved X-ray scattering in recent years, and various naturally photostable proteins have been studied thus far. The naturally photostable proteins are once again laying the foundation for studying artificially controllable systems, and we expect to see first results in the years to follow.

In the meantime, photoswitches suitable for in vivo applications, namely ones that isomerise on longer wavelengths (red to near-infrared) and are resistant to reductive environments and hydrolysis, are being developed. First attempts to apply azobenzene photocontrol in living systems have been made and its potential use in therapeutic targeting has been demonstrated.
As manipulating protein activity with azobenzene photoswitches is getting more adaptable and as its applications are broadening, optogenetics is emerging as a way of making proteins photosensitive by genetic manipulation of naturally occurring photosensitive proteins or domains. It usually designs novel photosensitive proteins based on, e.g., retinal (or other chroomophores) binding domains and their linking to other proteins on the genetic level. Neuronal activity can be regulated in this way, as well as many other aspects of cellular function. A strategy to involve azobenzene for optogenetic applications involves the method of codon expansion. It is based on an engineered pair of tRNA and tRNA synthetase, which incorporates an azobenzene-bearing unnatural amino acid during the protein translation directly in living cells.

In any case, azobenzene photocontrol currently offers the most versatile approach and offers answers to fundamental questions tackling the time propagation of protein-protein interaction mechanisms (conformational selection vs induced fit), and coupled binding/folding in intrinsically disordered proteins.

Acknowledgement: We thank all our coworkers in this project, in particular Brigitte Buchli, Steven Waldauer, Reto Walser, Oliver Zerbe, Rolf Pfister, Klemens Koziol, Philip J. M. Johnson, Claudio Zanobini, Jeannette Ruf, David Buhrke, and as well as the groups of Gerhard Stock, Amedeo Cafisch and Ben Schuler for their numerous contributions to this works. We also thank Andrew Woolley, who got us started in an initial phase of the project. The work has been supported by an European Research Council (ERC) Advanced Investigator Grant (DYNALLO) as well as the Swiss National Science Foundation (SNF) through the NCCR MUST and Grants 200021_165789/1 and 200020B_188694/1.

References:

1. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. Nature 450, 964–972 (2007).
2. Dyson, H. J. & Wright, P. E. Equilibrium NMR studies of unfolded and partially folded proteins. Nat. Struct. Biol. 5, 499–503 (1998).
3. Bernado, P. & Blackledge, M. Proteins in dynamic equilibrium. Nature 468, 1046–1048 (2010).
4. Elber, R. & Karplus, M. Enhanced sampling in molecular dynamics: use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. J. Am. Chem. Soc. 112, 9161–9175 (1990).
5. Palmer, A. G. NMR characterization of the dynamics of biomacromolecules. Chem. Rev. 104, 3623–3640 (2004).
6. Mittermaier, A. & Kay, L. E. New tools provide novel insights in NMR studies of protein dynamics. Science 312, 224–228 (2006).
7. Kay, L. E. NMR studies of protein structure and dynamics. J. Magn. Reson. 173, 193–207 (2005).
8. Olsson, M. H., Parson, W. W. & Warshel, A. Dynamical contributions to enzyme catalysis: Critical tests of a popular hypothesis. Chem. Rev. 106, 1737–1756 (2006).
9. Benkovic, S. J. & Hammes-Schiffer, S. A perspective on enzyme catalysis. Science 301, 1196–1202 (2003).
10. Csermely, P., Palotai, R. & Nussinov, R. Induced fit, conformational selection and independent dynamic segments: An extended view of binding events. Trends Biochem. Sci. 35, 539–546 (2010).
11. Swain, J. F. & Gierasch, L. M. The changing landscape of protein allostery. Curr. Opin. Struct. Biol. 16, 102–108 (2006).
12. Tsai, C. J., del Sol, A. & Nussinov, R. Allostery: Absence of a Change in Shape Does Not Imply that Allostery Is Not at Play. J. Mol. Biol. 378, 1–11 (2008).
13. Smock, R. G. & Gierasch, L. M. Sending signals dynamically. Science 324, 198–203 (2009).
14. Tsai, C. J. & Nussinov, R. A Unified View of "How Allostery Works". PLoS Comput. Biol. 10, e1003394 (2014).
15. Hilser, V. J., Wrabl, J. O. & Motlagh, H. N. Structural and Energetic Basis of Allostery. Annu. Rev. Biophys. 41, 585–609 (2012).
16. Schotte, F. et al. Watching a protein as it functions with 150-ps time-resolved X-ray crystallography. Science 300, 1944–1947 (2003).
17. Knapp, J. E., Pahl, R., Šrajer, V. & Royer, W. E. Allosteric action in real time: Time-resolved crystallographic studies of a cooperative dimeric hemoglobin. Proc. Natl. Acad. Sci. U. S. A. 103, 7649–7654 (2006).
18. Kern, J. et al. Simultaneous Femtosecond X-ray Spectroscopy and Diffraction of Photosystem II at Room Temperature. Science 340, 491–496 (2013).
19. Nogly, P. et al. Retinal isomerization in bacteriorhodopsin captured by a femtosecond x-ray laser. Science 361 (2018).
20. Standfuss, J. Membrane protein dynamics studied by X-ray lasers – or why only time will tell. Curr. Opin. Struct. Biol. 57, 63–71 (2019).
21. Skopintsev, P. et al. Femtosecond-to-nanosecond structural changes in a light-driven sodium pump. Nature 583, 314–318 (2020).
22. et al. Chapman, H. N. Femtosecond X-ray protein nanocrystallography. Nature 470, 73–77 (2011).
23. Wang, Q., Schoenlein, R. W., Peteanu, L. A., Mathies, R. A. & Shank, C. V. Vibrationally coherent photochemistry in the femtosecond primary event of vision. Science
Kneissl, S., Loveridge, E. J., Williams, C., Crump, M. P., Guerrero, L., Smart, O. S., Woolley, G. A. & Allemann, B., Bozovic, O., Jankovic, B. & Hamm, P. Sensing the Al-
nanosecond peptide conformational dynamics and validates molecular dynamics simulation. *Proc. Natl. Acad. Sci. USA* **99**, 7998–8002 (2002).

Bredenbeck, J. *et al.* Picosecond conformational transition and equilibration of a cyclic peptide. *Proc. Natl. Acad. Sci. USA* **100**, 6452–6457 (2003).

Aemisegger, A., Krautler, V., van Gunsteren, W. F. & Hilvert, D. A photoinducible beta-hairpin. *J. Am. Chem. Soc.* **127**, 2929–2936 (2005).

Rehm, S., Lenz, M. O., Mensch, S., Schwabbe, H. & Wachtveitl, J. Ultrafast spectroscopy of a photoswitchable 30-amino acid de novo synthesized peptide. *Chem. Phys.* **323**, 28–35 (2005).

Schrader, T. E. *et al.* Light-triggered beta-hairpin folding and unfolding. *Proc. Natl. Acad. Sci. USA* **104**, 15729–15734 (2007).

Rampp, M. S. *et al.* Time-resolved infrared studies of the unfolding of a light triggered beta-hairpin peptide. *Chem. Phys.* **512**, 116–121 (2018).

Flint, D. G., Kumita, J. R., Smart, O. S. & Woolley, G. A. Using an Azobenzene Cross-Linker to Either Increase or Decrease Peptide Helix Content upon Trans-to-Cis Photoisomerization. *Chem. Biol.* **9**, 391–397 (2002).

Woolley, G. A. Photocontrolling peptide alpha helices. *Acc. Chem. Res.* **38**, 486–493 (2005).

Morais Cabral, J. H. & Hamm, P. Alpha-Helix formation in a photoswitchable 30-amino acid de novo synthesized peptide. *Proc. Natl. Acad. Sci. USA* **106**, 18249–18254 (2009).

Ihalainen, J. A. *et al.* Folding and unfolding of a photoswitchable peptide from picoseconds to microseconds by time resolved IR spectroscopy. *Proc. Natl. Acad. Sci. USA* **102**, 2379–2384 (2005).

Ihalainen, J. A. *et al.* Alpha-Helix folding in the presence of structural constraints. *Proc. Natl. Acad. Sci. USA* **105**, 9588–9593 (2008).

Zhang, F. *et al.* Structure-based approach to the photo-control of protein folding. *J. Am. Chem. Soc.* **131**, 2283–2289 (2009).

Bredenbeck, J., Helbing, J., Kumita, J. R., Woolley, G. A. & Hamm, P. alpha-Helix formation in a photoswitchable 30-amino acid de novo synthesized peptide. *Proc. Natl. Acad. Sci. USA* **104**, 15729–15734 (2007).

Lee, H.-J. & Zheng, J. J. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal.* **8**, 8 (2010).

Ivarsson, Y. Plasticity of PDZ domains in ligand recognition and signaling. *FEBS Lett.* **586**, 2638–2647 (2012).

Kumar, A. T. N., Zhu, L., Christian, J. F., Demidov, A. A. & Champion, P. M. On the Rate Distribution Analysis of Kinetic Data Using the Maximum Entropy Method: Applications to Myoglobin Relaxation on the Nanosecond and Femtosecond Timescales. *J. Phys. Chem. B* **105**, 7847–7856 (2001).

Lorenz-Fonfria, V. A. & Kandori, H. Transformation of time-resolved spectra to lifetime-resolved spectra by maximum entropy inversion of the Laplace transform. *Appl.
Spectrosc. 60, 407–417 (2006).

Buhre, D., Oppelt, K. T., Heckmeier, P. J., Fernandez-Teran, R. & Hamm, P. Nanosecond protein dynamics in a red/green Cyanobacteriochrome revealed by transient IR spectroscopy. J. Chem. Phys. 153, 245101 (2020).

Shaw, D. E. et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. Science 330, 341–346 (2010).

Stock, G. & Hamm, P. A Nonequilibrium Approach to Allosteric Communication. Philos. Trans. R. Soc. B Biol. Sci. 373, 20170187 (2018).

Frauenfelder, H., Slijar, S. G. & Wolynes, P. G. The Energy Landscapes and Motions of Proteins. Science 254, 1598–1603 (1991).

Dill, K. A. & Chan, H. S. From Levinthal to Pathways to Funnels: The “New View” of Protein Folding Kinetics. Nat. Struct. Biol. 4, 10–19 (1997).

Bowman, G. R., Pande, V. S. & Noe, F. An Introduction to Markov State Models (Springer, Heidelberg, 2013).

Pande, V. S., Beauchamp, K. & Bowman, G. R. Everything you wanted to know about Markov state models but were afraid to ask. Methods 52, 99–105 (2010).

Prinz, J.-H. et al. Markov models of molecular kinetics: Generation and validation. J. Chem. Phys. 134, 174105 (2011).

Sengupta, U. & Strodel, B. Markov models for the elucidation of allosteric regulation. Philos. Trans. R. Soc. B Biol. Sci. 373, 20170178 (2018).

Ham, P., Ohline, S. M. & Zinth, W. Vibrational cooling after ultrafast photoisomerization of azobenzene measured by femtosecond infrared spectroscopy. J. Chem. Phys. 106, 519–529 (1997).

Baumann, T. et al. Site-Resolved Observation of Vibrational Energy Transfer Using a Genetically Encoded Ultrafast Heater. Angew. Chemie - Int. Ed. 58, 2899–2903 (2019).

Buchenberg, S., Sittel, F. & Stock, G. Time-Resolved Observation of Protein Allosteric Communication. Proc. Natl. Acad. Sci. USA 114, E6804–E6811 (2017).

Doyle, D. A. et al. Crystal structures of a complexed and peptide-free membrane protein- binding domain: Molecular basis of peptide recognition by PDZ. Cell 85, 1067–1076 (1996).

Ballif, B. A., Carey, G. R., Sunyaev, S. R. & Gygi, S. P. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. J. Proteome Res. 7, 311–318 (2008).

Zhang, J., Petit, C. M., King, D. S. & Lee, A. L. Phosphorylation of a PDZ domain extension modulates binding affinity and interdomain interactions in postsynaptic density-95 (PSD-95) protein, a membrane-associated guanylate kinase (MAGUK). J. Biol. Chem. 286, 41776–41785 (2011).

Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. Nat. Chem. Biol. 5, 789–796 (2009).

Morando, M. A. et al. Conformational Selection and Induced Fit Mechanisms in the Binding of an Anticancer Drug to the c-Src Kinase. Sci. Rep. 6, 24439 (2016).

Vogt, A. D. & DiCera, E. Conformational Selection or Induced Fit? A Critical Appraisal of the Kinetic Mechanism. Biochemistry 41, 5894–5902 (2012).

Gianni, S., Dogan, J. & Jemth, P. Distinguishing induced fit from conformational selection. Biophys. Chem. 189, 33–39 (2014).

Hammes, G. G., Chang, Y.-C. & Oas, T. G. Conformational selection or induced fit: A flux description of reaction mechanism. Proc. Natl. Acad. Sci. USA 106, 13737–13741 (2009).

Deisseroth, K. Optogenetics. Nat. Methods 8, 26–29 (2011).

Hausser, M. Optogenetics: The age of light. Nat. Methods 11, 1012–1014 (2014).

Hoppmann, C. et al. Genetically encoding photoswitchable click amino acids in Escherichia coli and mammalian cells. Angew. Chemie - Int. Ed. 53, 3932–3936 (2014).

Hoppmann, C., Maslennikov, I., Choe, S. & Wang, L. In Situ Formation of an Azo Bridge on Proteins Controllable by Visible Light. J. Am. Chem. Soc. 137, 11218–11221 (2015).