Quarterly Reviews of Biophysics

www.cambridge.org/qrb

Dynamics and kinetics in structural biology: the example of DNA photolyase

Keith Moffat

Department of Biochemistry and Molecular Biology, Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL, USA

Review

Cite this article: Moffat K (2025). Dynamics and kinetics in structural biology: the example of DNA photolyase. *Quarterly Reviews of Biophysics*, **58**, e8, 1–15
https://doi.org/10.1017/S0033583524000222

Received: 30 September 2024
Revised: 23 December 2024
Accepted: 23 December 2024

Keywords:
chemical kinetic mechanisms; DNA photolyase; dynamics; energy landscapes

Corresponding author:
Keith Moffat;
Email: jkmoffat@uchicago.edu

Abstract

All biochemical reactions directly involve structural changes that may occur over a very wide range of timescales from femtoseconds to seconds. Understanding the mechanism of action thus requires determination of both the static structures of the macromolecule involved and short-lived intermediates between reactant and product. This requires either freeze-trapping of intermediates, for example by cryo-electron microscopy, or direct determination of structures in active systems at near-physiological temperature by time-resolved X-ray crystallography. Storage ring X-ray sources effectively cover the time range down to around 100 ps that reveal tertiary and quaternary structural changes in proteins. The briefer pulses emitted by hard X-ray free electron laser sources extend that range to femtoseconds, which covers critical chemical reactions such as electron transfer, isomerization, breaking of covalent bonds, and ultrafast structural changes in light-sensitive protein chromophores and their protein environment. These reactions are exemplified by the time-resolved X-ray studies by two groups of the FAD-based DNA repair enzyme, DNA photolyase, over the time range from 1 ps to 100 µs.

Table of contents

Introduction

The study of atomic level, static structures by X-ray crystallography and increasingly, cryo-electron microscopy (cryo-EM) forms the backbone today of the discipline of structural biology. The key word here is 'static'. The static nature of experimental crystallographic and cryo-EM structures has often been imposed by chemical constraints to trap the system in a stable chemical and structural state (Moffat and Henderson, 1995; Bock and Grubmueller, 2022). To study an enzyme crystal, for example, a key chemical reactant such as a substrate or an essential cofactor may be omitted, an inhibitor of enzymatic activity may be added, or an unfavorable pH used. Alternatively, the constraint may be physical, imposed by freezing a crystal or EM sample to cryogenic temperatures. In single-particle cryo-EM studies, quasi-static states related to authentic reaction intermediates can also be cryotrapped, imaged, and analyzed. However, the detailed relation between the structures of these trapped states and authentic (but unknown) intermediate states may be affected by specific details of the subsequent freezing process (Halle, 2004; Bock and Grubmueller, 2022). Cryo temperatures minimize radiation damage to the sample by intense electron or X-ray beams and dehydration damage by the high vacuum in which the beams propagate. Both cryocrystallography and cryo-EM deliver more precise structures at much higher spatial resolution, more speedily, using less sample. These compelling experimental advantages have led to the widespread adoption of cryo-based techniques.

The application of chemical or physical constraints to the sample imposes a fundamental limitation. All biochemical, chemical, and physical reactions require not merely static structure but structural changes. In physics and chemistry, these may involve motion at the electronic and atomic levels, and in biophysics at the atomic, molecular, and chemical levels. Time-dependent structural changes are fundamental to all reactions and activity. Chemically trapped or frozen samples are completely inactive. The slogan, common in the early days of protein crystallography when I was a research student, that '*structure* determines function' is too limiting. '*Changes in structure* determine function' is more accurate. Precise structures obtained in the presence of

chemical or physical constraints can only indirectly address the structural bases of activity - the mechanisms by which all biological or chemical functions such as transcription, catalysis, or ion translocation are achieved. Mechanism is inherently dynamic.

An artistic presentation of a ballet dancer in the transition between two quasi-static positions is shown in Figure 1, created decades before the scientific concept of 'transition state' was developed. The sculptor effectively captured the dancer in her dynamic transition between two static states.

Cryo structures are certainly more precise, but are they more accurate? Consider structure-based drug design. Could the usefulness of the experimental structures that inform the designs be enhanced if the samples from which the structures were obtained were demonstrably active rather than completely inactive? Can the structures and properties of authentic, biochemically active samples during their reactions be directly determined? Might short-lived intermediates be useful or even superior targets for drug design?

Questions like these date from the earliest days of protein crystallography: are crystal structures physiologically relevant (Chance et al., 1966)? Addressing them has returned to the fore in recent years. The number of structures obtained by cryo-based techniques has greatly expanded. The advent of highly accurate prediction of static structures, based on artificial intelligence by techniques such as AlphaFold (Jumper et al., 2021) and RosettaFold (Baek et al., 2021), suggests that the experimental determination of static structures by X-ray crystallography will no longer be at the forefront (Khusainov et al., 2024). Novel computational and experimental approaches are essential to explore time-resolved structural changes and critically, to explore mechanisms.

Figure 1. The dynamic transition of a ballet dancer between two positions. Sculpture c. 1913 by Henri Gaudier-Brzeska (1891–1915). Photo reproduced with permission from Kettle's Yard, Cambridge University.

Processes that involve structural changes and directly address mechanisms span a very wide range of time scales from fs to s or even minutes. In X-ray crystallography and solution scattering, earlier time-resolved experiments were based on X-ray pulses emitted by hard X-ray storage rings such as the European Synchrotron Radiation Facility (ESRF) and the Advanced Photon Source (APS). Their time resolution was limited by the X-ray pulse length to roughly 100 ps (Szebenyi et al., 1988; Srajer et al., 1996). The advent of exceptionally intense, fs to ps X-ray pulses at the first hard X-ray free electron laser source (XFEL), the Linac Coherent Light Source (LCLS) at the Stanford Linear Accelerator Center, greatly expanded the reach of time-resolved crystallography (Chapman et al., 2011). The ultrashort pulses at the LCLS extended the time resolution below 100 ps by two to three orders of magnitude, into the fs to few ps time range. Extension had critical scientific consequences. For biochemistry, chemistry, and materials science more generally, fs to ps spans the time range in which elementary chemical reactions such as isomerization, bond breaking, and ultrafast electron transfer occur. New areas opened up as the structural bases of such reactions became accessible to experiment. Even the attosecond transitions characteristic of atomic physics that underlie the reaction chemistry and primary radiation damage may become accessible (Graves et al., 2020).

The properties of the storage ring and XFEL X-ray sources, their beamlines, and the overall experimental approaches to time-resolved experiments in biophysics and structural biology were recently presented at book length (Moffat and Lattman, 2024). Successful time-resolved experiments at XFELs have been comprehensively tabulated and reviewed (Orville, 2020; Pearson and Mehrabi, 2020; Branden and Neutze, 2021; Malla and Schmidt, 2022; Khusainov et al., 2024) and time-resolved techniques emphasized (Hekstra, 2023). The book and these reviews cover a wide range of systems differing in complexity, method of reaction initiation, time resolution, and experimental approaches. Together, they form a useful guide for scientists planning their own time-resolved studies.

This article is personal and selective. A description of aspects of time-resolved crystallography defines key terminology, emphasizes the differences between dynamic and kinetic experimental approaches and their presentation on energy landscapes, and summarizes some key experimental aspects such as radiation damage arising from very intense pulses from a visible pump laser or probe X-ray laser. Two excellent recent articles (Maestre-Reyna et al., 2023; Christou et al., 2023) explore the mechanism of the light-dependent photoenzyme, DNA photolyase. They illustrate well the current state of time-resolved crystallography, its strengths and limitations.

Nevertheless, time-resolved structural techniques are by no means mature. New ideas and approaches are essential. This text is peppered with unanswered (or under-answered) questions intended to be provoking. Some will turn out to be misdirected or wrong, but others may be useful. A more general set of important questions in biophysics was offered by Norden (2021) as he became Editor in Chief of this journal. Several of his questions foreshadowed developments that have come to pass.

Dynamic and kinetic experiments

Dynamics and kinetics: terminology

Informally, the terms 'dynamics' and 'kinetics' indicate motion and time dependence of a process, to be contrasted with static and time-independent. Here, I prefer more formal uses of 'dynamics' and

https://doi.org/10.1017/S0033583524000222

'kinetics' in structural biophysics which draw attention to a scientifically significant difference between the terms. The difference lies both in the style of experiments and the representation of structural results in an energy landscape. The formal use restricts 'dynamics' to describe the time dependence of structural changes in a *statistically small number of molecules*. In practice, the statistically small number is usually one. Examples are an experiment on a single molecule (or a single cell), or a molecular dynamics calculation. Observations or calculations based on single molecules may of course be averaged later over many molecules. In contrast, the term 'chemical kinetics' or more simply 'kinetics' may be restricted to the time dependence of the properties of a *statistically large number of molecules*, in which the experimental observations directly average over all the many molecules in the sample. Examples are experimental X-ray scattering from a crystal or a solution in which all molecules are simultaneously illuminated, or thermodynamic measurements of equilibria and rate coefficients in bulk solution. Cryo-EM and single molecule spectroscopy experiments thus observe dynamics; crystallography, conventional optical and X-ray spectroscopies, and thermodynamics observe kinetics; and molecular dynamics simulations predict dynamics.

In time-resolved experiments in biochemistry and biophysics, the experimental goal is to identify and characterize all structures on the pathway or pathways between reactant and product. How do these structures interconvert? With what rate coefficients? Attainment of this goal determines the mechanism of the reaction at the structural and chemical levels. The experimental goal may be complemented by theoretical and computational goals, to understand *what* form the mechanism takes, *why* it occurs in this way, and *why* its steps have their specific rate coefficients. Dynamic and kinetic experiments differ greatly in the information they reveal and how it is presented and analyzed in terms of mechanism.

Distinction between dynamics and kinetics: the energy landscape

A useful and widely used way to represent the thermodynamics of a reaction is to plot the free energy of the system versus the reaction coordinate. The last is a pseudo-coordinate which for purposes of illustration, compresses the very large number of authentic chemical/structural coordinates of the system into a one- or two-dimensional representation that can be readily visualized: an *energy landscape* (Schon, 2024). In a landscape, each valley bottom of low free energy represents a reactant, intermediates, or product state. Each saddle point represents a *transition state*, the highest point of free energy between two adjacent states. The pathway of a reaction is illustrated by a *reaction trajectory* in this landscape. The trajectory begins at the valley bottom associated with the reactant, and then may continue on the reaction trajectory via several intermediates. The progress between two adjacent intermediates/states constitutes an *elementary step*. The reaction as a whole typically contains several elementary steps. The reaction trajectory ends at the valley bottom associated with the product. In principle, the structures in the energy landscape populate a Boltzmann distribution. The least populated, rarest structures are those of the highest free energy, which are likely to include the transition states of each elementary step in the overall reaction. The most populated structures in the energy landscape are those of the lowest free energy, the local valley bottoms corresponding to the reactant, and all reaction intermediates and products. Figure 2 illustrates a relatively simple energy

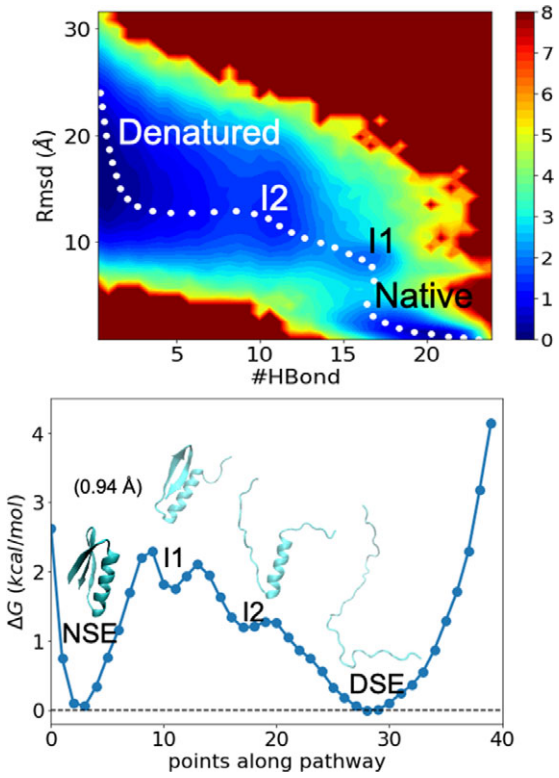

Figure 2. An energy landscape representation of the protein unfolding reaction of a designed mini-protein at its melting temperature, generated using the Upside coarse-grain molecular dynamics simulation. The native state (NSE; lower panel) proceeds via two distinct intermediates I1 and I2 to the denatured state (DSE; lower panel). The scale bar in the heat map (upper panel) is ln (population). The white dots (upper panel) trace the minimum energy trajectory as the reaction proceeds from the native state ensemble (NSE) to the denatured state ensemble (DSE). Details of similar representations are given by Peng et al. (2022). Images by kind permission of Tobin Sosnick.

landscape based on the simulation of a protein folding reaction (details in Peng et al., 2022), where the dotted line represents a highly likely reaction trajectory.

The ability to address the experimental goal is limited in very different ways in kinetic and dynamic experiments. In kinetic experiments, the overall time dependence of, for example, X-ray structure amplitudes originates in the rise and fall of the populations of the intermediate states that constitute the mechanism: the reactant state, the array of intermediate states, and the product state. Each of these structures is time-independent. Those of the reactant (prior to reaction initiation) and product are at equilibrium and indefinitely stable, but those of all intermediates are metastable. They form, persist, and decay with defined rates and at defined times after reaction initiation. In the absence of spatial synchronization (Khusainov et al., 2024), no kinetic experiment can yield a movie in which structures are continuously varying. Despite this limitation, the use of 'movie' in manuscripts and grant applications is surprisingly frequent. In a kinetic experiment, the average electron density of atoms in the sample declines in one state and correspondingly increases in the next state. No motion of atoms is directly visualized; no atomic positions along the reaction trajectory between adjacent states are revealed. Although a movie-like presentation of the data can morph computationally from one time-independent structure to the next in the reaction, the presentation contains only the information present in the two time-independent structures.

https://doi.org/10.1017/S0033583524000222

As a practical matter, the mechanism identified by kinetic experiments determines the number of intermediates present, the structure and properties of each, and how they are interconnected (Schmidt, 2023). This number may be estimated from time-dependent data by analytical approaches such as singular value decomposition (SVD; Schmidt et al., 2003; Ren and Yang, 2024) and non-negative matrix factorization (Lee and Seung, 1999). The Supplementary Text of Christou et al. (2023) details a recent application of SVD.

A limitation in the interpretation of data from kinetic experiments is that several mechanisms are compatible with an experimentally determined number of intermediates. Failure to recognize that identical kinetic data may arise from several distinct mechanisms is again surprisingly frequent. The preferred mechanism must be distinguished by other, external data such as the structural homogeneity of each intermediate in a candidate mechanism, chemical plausibility, theoretical computation – or plain simplicity, as in a linear reaction mechanism in which all steps are irreversible.

A second limitation is that short-lived states that follow long-lived states are very hard to detect. If their peak population does not rise significantly above the error level, the short-lived state will not be identifiable in the kinetic data nor incorporated in mechanisms derived from that data. However, independent chemical data may suggest that this short-lived state is plausible as a product of the decay of the long-lived state and a precursor of the next state. In that case, the presence of the mechanism of the short-lived state is neither supported nor disproved by the kinetic experiments.

This limitation has practical consequences. Chemical reactions such as isomerization or bond breaking are typically ultrafast and occur in fs or a few ps; but many tertiary structural changes in proteins are much slower, requiring ns or even μs. Chemical reactions that *precede* tertiary changes are readily detectable if the time resolution is sufficiently high. Fast chemical reactions that *follow* a slow tertiary structural change are not separately detectable; the chemical reaction and the tertiary changes are coupled and proceed with (very nearly) identical time courses. For example, a common observation is that kinetic data on the photocycles of light-sensitive systems are interpreted with first-order rate coefficients that progressively decrease in magnitude. That is, the lifetimes of intermediates progressively increase as the photocycle proceeds. This does not mean that no fast chemical steps occur late in the photocycle. If they do occur, they cannot be separately distinguishable from the preceding, much slower tertiary or quaternary structural changes in the protein.

Dynamic experiments confront two limitations which make them very challenging to conduct. First, a single molecule scatters electrons or X-rays much more weakly than a large population of molecules in a kinetic experiment. The signal in each measurement in a dynamic experiment is inherently very weak, the signal-to-noise is low and the signal is liable to be masked by noise or corrupted by systematic error. Second, the amount of experimental data necessary to survey the entire energy landscape, or even the region containing the trajectory of a single elementary step, is extremely large. A lower bound on the size of the data set depends on the number of sample points necessary to define a trajectory in each elementary step in the mechanism, and the number of such steps in all candidate mechanisms. A further extension may be necessary to account for any time dependence on the energy landscape itself. The landscape may alter as the overall reaction proceeds. For example, when charges move in a protein, their electrostatic interactions with all neighboring charges are affected, the positions of those charged atoms shift and the energy landscape alters. As a rough estimate, the very large data required in a comprehensive dynamic experiment exceeds the limited data in a kinetic experiment by three to four orders of magnitude.

At present, the most powerful experimental approach to dynamic studies is offered by cryo-EM of protein complexes or increasingly, of individual large proteins. A set of cryo-EM images of a reaction in progress records separately the structures of a very large number of individual complexes, each freeze-trapped and located in a different orientation and initially unknown position on the reaction coordinate. Individual structures can be successfully determined to high spatial resolution and classified into separate structural bins (Scheres, 2012; Frank and Ourmazd, 2016). If a reaction is initiated in a sample immediately prior to freezing, the structures in each bin are assumed to arise from a structural intermediate or from a point on the functional trajectory of an elementary step. In accord with the Boltzmann distribution, the least represented structures in the entire population should arise from the structures of the highest free energy. Perhaps the most interesting of these are at or near the transition states for each elementary step.

Spatially similar structures are likely to be close in the reaction coordinate, but explicit timing information is absent from the measurement. If structures α and β are close, does structure α precede structure β in the reaction coordinate, or does structure β precede structure α? More generally, can each structural bin be confidently associated with timing information? In a major analytical feat, the structures from all bins are assembled into a time series in the widely-used program RELION (Scheres, 2012). The overall structures are *time-dependent*. The time dependence constitutes a *dynamic reaction trajectory:* an authentic movie of continuous structural change where, for example, the dotted line in Figure 2 represents the most likely trajectory. An aspect of dynamic experiments that distinguishes them from kinetic experiments is that provided their time resolution is sufficient, they may be able to visualize short-lived intermediates that follow long-lived intermediates.

Experiments and analysis are not always straightforward. An experimental assumption is that the mounting of the protein complex on the cryo-EM grid, its history immediately prior to freezing, and the dehydration of molecules near the air-water interface of the droplet on the grid have not significantly affected the individual structures that populate the mechanism. The free energy landscape of molecules on the frozen grid is assumed to retain the same form of valleys and peaks as the landscape at physiological temperature, although shifted by the cryo temperature to much lower free energy. The overall freezing process requires ms to reach the temperature at which motion is frozen. However, extensive tertiary and even quaternary structural changes can and do occur in proteins on much shorter time scales, μs and below.

Lorenz (2024) is developing a novel approach that transiently 'melts' individual freeze-trapped structures on a cryoEM grid. Melting allows the initiation and progress of dynamics in the melted structure to be followed by electron microscopy in the time window between melting and refreezing, which terminates the reaction.

A complication is that dynamics and kinetics are in the end governed by free energy differences, which often are coupled only loosely to structural differences. Two tertiary structures can differ substantially in a local region, but the structures may have closely similar free energy. Conversely, two closely similar tertiary structures may lie on a large energy gradient and differ substantially in free energy. This suggests that the key dimension of structural

https://doi.org/10.1017/S0033583524000222

change is not length (root mean square displacement measured in nm) but rather, free energy (kJ/mol).

The energy landscape formalism dramatically illustrates the difference between kinetic and dynamic experiments. Kinetic experiments are relatively straightforward to conduct, but yield data limited to an important, but very sparse, array of points in the landscape. As noted above, these points correspond either to a thermodynamic state (reactant, intermediates, and product) at a valley bottom, or to a transition state at a saddle point. No data are obtainable at other points. In contrast, a complete dynamic experiment would yield data at an extremely large number of points in the chemical/structural space, at the few discrete points of the kinetic experiment, and importantly, at all points in between. That is, it reveals the energetic, reaction trajectory via a transition state for all elementary steps in the mechanism. It also yields all more distant points, not near or visited by any reaction trajectory – the entire energy landscape.

A complete dynamic experiment would reveal much more information than a kinetic experiment. Is the characterization of functional trajectories by dynamic approaches essential to a full understanding of mechanism? If the answer to this key question is 'yes', this justifies the value of the much larger experimental effort required for dynamic experiments. An evolutionary perspective on this question may be helpful. Evolution has selected the states, structures, and rate coefficients that populate a biochemical mechanism. These directly influence all cellular metabolism, the interactions between cells, and responses to their environment. The nature and quantitative properties of the states associated with these processes clearly can be, and are, judged fit or unfit in evolutionary terms. A fundamental question arises: has evolution *also* selected for the properties of the functional trajectories that connect each adjacent pair of states in a mechanism? As noted above, the states, structures, and rate coefficients are accessible to both kinetic and dynamic studies. However, functional trajectories are accessible only to dynamic studies. They are invisible to kinetic studies. To frame the question more casually, do the detailed pathways between states contribute to fitness in evolutionary terms? Or, is it sufficient for evolution that at least one energetically accessible pathway between states exists? From the energy landscape formalism, the existence of many accessible pathways between two states is certain and at the very least, confers redundancy. Continuing the landscape analogy, there are many pathways by which a hiker may cross a mountain pass. Some are arduous, some are easier, one is easiest, some may be hindered by a landslide, and some detour from the pass to explore a nearby peak. The time taken and energy required vary from pathway to pathway, but all offer success and allow the hiker to proceed from one valley over the pass to the next valley.

Posing these questions to evolutionary biologists and biophysicists has not yielded an answer - at least, thus far. There's clear agreement that states, structures, and rate coefficients are critically important and selected for by evolution, but there's no agreement on whether selection extends to functional trajectories and pathways.

Experimental aspects

Reaction initiation and monitoring: pump-probe

Pump – probe approaches to reaction initiation and monitoring originated in spectroscopy and have been widely adopted in biophysics (Moffat and Lattman, 2024). The reaction and its structural changes are initiated uniformly in the sample by a rapid chemical or physical process known as the pump, thus defining time zero. The pump is followed after a time delay by a structural probe, to monitor for example the X-ray structure amplitudes or solution scattering as a function of scattering angle. The course of the reaction is followed by varying the time delay between the pump and probe. The overall reaction is assembled from the data in a set of delays, time point by time point: a *time series.* This is tedious and also error-prone since data at different time points is acquired on different samples. However, it is necessitated by two factors. As discussed below, radiation damage severely limits the amount of good data that can be acquired on one sample, often to a single diffraction pattern at a single time point. And, the rate at which X-ray detectors can acquire, initially analyze, and store data may be much lower than the rate at which probe pulses can be delivered. With the wide availability of short, intense X-ray probe pulses and well-controlled timing jitter between the pump and probe pulses, the time resolution is largely set by the duration of the pumping process.

The mode of reaction initiation depends on the questions to be addressed, the properties of the sample, the time resolution required, the time range to be explored, and how the sample is introduced to the probe. The simplest form initiates the reaction by mixing reactants in solution, or for example by adding a substrate solution to a tiny liquid droplet containing a crystal. The time resolution is determined by the hydrodynamics of mixing of liquids, or by the rates of diffusion of a reactant into a crystal. Millisecond time resolution is routine and fortunately, well matched to the relatively low turnover number of enzymes. However, the hydrodynamics of even tiny samples is relatively slow, which means that μs or better time resolution in the sub-ps chemical timescale cannot be achieved. Relatively slow mixing processes are associated with small free energy changes, insufficient to drive structural changes that require large amounts of free energy. Thus, the energy landscape for reactions initiated by mixing is not rugged. It consists of valley bottoms separated by transition states that do not differ greatly in free energy from that of the valley states: valleys separated by low passes and hills. Since the reaction proceeds near equilibrium throughout, it can be characterized by equilibrium thermodynamics.

Figure 3 illustrates an energy landscape reaction based on mixing as the means of reaction initiation. The mixing reaction requires a transition between two energy landscapes, calculated without and with a set of ligands (calcium, ATP, and caffeine) present that bind to the ryanodine receptor (Dashti et al., 2020). The reaction trajectory between reactant (START, upper landscape with no ligands, before mixing) passes through a set of intermediates S1 through S6 and the point of highest free energy (HOT) corresponding to the transition state to reach the product with bound ligands (FINISH, lower landscape with ligands, after mixing). The curved path shows a pathway of high probability between reactant and product: a candidate reaction trajectory.

Time resolution in the ultrafast ps to fs region is required to access the time scale of individual chemical reactions. These are the critical chemical steps in a mechanism, directly coupled to changes in protein tertiary structure. Reaction initiation with time resolution in the ps or sub-ps range that directly probes chemical processes is readily achievable by a laser pulse in the visible region of the spectrum. Indeed, ultrafast time-resolved crystallography was developed and initially applied to naturally light-sensitive proteins such as CO-myoglobin (Srajer et al., 1996; Barends et al., 2015, 2024), signaling photoreceptors such as photoactive yellow protein (Perman et al., 1998; Schotte et al., 2012;

https://doi.org/10.1017/S0033583524000222

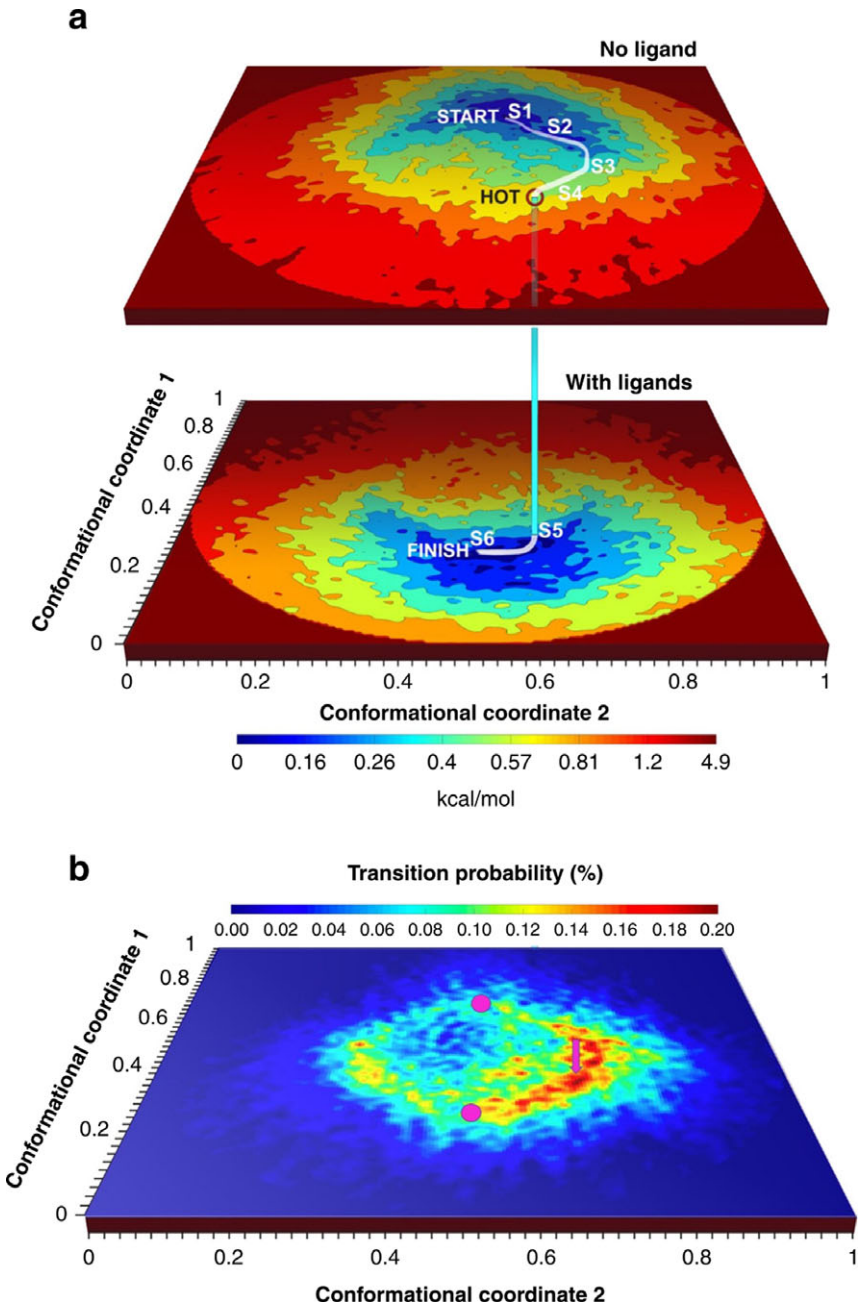

Figure 3. Two energy landscapes represent the binding of ligands to the ryanodine receptor type 1. In panel a, the upper landscape is with no ligands, and the lower landscape is with ligands after mixing. Panel b shows the probability of transition between the two landscapes. Details are given by Dashti et al. (2020). Reproduced with permission of Springer Nature/Licensed under CC BY 4.0.

Tenboer et al. (2014) and bacteriophytochrome (Edlund et al., 2016); integral membrane proteins such as bacteriorhodopsin (Nogly et al., 2018), rhodopsin (Gruhl et al., 2023), a sodium pump (Skopintsev et al., 2020) and the photosynthetic reaction center (Dods et al., 2021); and flavoenzymes such as DNA photolyase (Maestre-Reyna et al., 2023; Christou et al., 2023). In these systems, light is absorbed by a non-protein, small organic component known as the chromophore such as heme, biliverdin, retinal or flavin mononucleotide. The chromophore is embedded in a protein pocket and coupled covalently or non-covalently to the protein. Absorption of a highly energetic photon in the visible region of the spectrum excites the chromophore to a higher electronic and vibrational state. Absorption drives the system far from equilibrium and changes its electronic and three-dimensional structure. The chromophore de-excites as the system cools, and associated structural changes of the chromophore pocket propagate throughout the tertiary structure of the protein and change its interfaces with other proteins. The early stages of de-excitation are strongly downhill in free energy and characterized by non-equilibrium thermodynamics. These stages have a free energy change sufficient to rapidly drive substantial structural changes such as isomerization or long-range electron transfer, processes that are generally inaccessible to the much smaller free energy available through reaction initiation by mixing. The energy landscape for reactions initiated by light is rugged. Descent from the high mountaintop achieved by photon absorption

https://doi.org/10.1017/S0033583524000222

occurs via steep slopes across mountain passes between peaks, towards a final valley. This valley is identical to the initial valley if the photocycle is fully reversible.

The two processes of reaction initiation, by mixing or by photon absorption, therefore fundamentally differ in time resolution, in their ability to drive energetic structural changes, and in whether equilibrium or non-equilibrium thermodynamics describes the subsequent reactions.

Radiation damage by X-ray probe and visible laser pump pulses

The quasi-monochromatic, extremely intense X-ray pulses emitted by XFEL sources are well suited to probing samples that scatter X-rays very weakly, such as tiny crystals, dilute solutions, and single particles. A potentially serious limitation is that radiation damage arising from the absorption of a single pulse of X-rays completely destroys the sample. Fortunately, a brief time window exists prior to the destruction of the sample. In this window, one X-ray diffraction pattern (from a single crystal) or one continuous scattering pattern (from a solution sample) of high quality can be obtained. The existence of the window was predicted (Neutze et al., 2000), experimentally confirmed (Chapman et al., 2006), and made possible 'diffraction before destruction '(Chapman et al., 2011). Although it is often stated that XFEL scattering data in this window is damage-free, this is not correct. Primary radiation damage is immediate and results from the direct absorption of an X-ray photon by an atom and the prompt ejection of a photoelectron. As a fundamental consequence of atomic physics, primary damage cannot be evaded and at the least, reduces the scattering power of the atom, alters its charge, and affects its chemical bonding. The photoelectron ejected by the absorbing atom causes substantial, secondary radiation damage along its path, which occurs over fs to a few ps. The fs to few ps time window is long enough to allow secondary damage to be greatly minimized by using very brief X-ray pulses. Longer and more intense pulses cause significant primary and secondary damage; shorter pulses reduce secondary damage (Lomb et al., 2011). More recent studies therefore use shorter, somewhat less intense X-ray probe pulses (Khusainov et al., 2024). XFEL scattering data obtained during the window is substantially reduced in radiation damage but is not damage-free.

The authentic signal in a light-initiated pump-probe experiment (and in physiological systems) results from the absorption of only one visible photon from the pump pulse by each chromophore. This causes excitation to its first excited state and subsequent decay to the ground state, accompanied by immediate structural changes in the chromophore and protein and followed by an array of slower changes. A second source of potentially significant radiation damage originates in the pump pulse when a very high fluence (energy per unit area), ultrafast, visible laser pulse is used as the pump to initiate the reaction. If a chromophore absorbs more than one photon from a single pump pulse, it may be driven into more highly excited, non-physiological electronic states (Grunbein et al., 2020; Besaw and Miller, 2023).

In many XFEL experiments, the peak fluences (intensities) of pump pulses were sufficiently high to generate two-photon absorption (see Table 1 of Khusainov et al., 2024), However, experimental evidence for significant damage from this source, such as unexpected features in DED maps that lack a clear biological or chemical explanation, has been surprisingly limited (Branden and Neutze, 2021). A powerful way to explore possible damage is to conduct a power titration of the pump pulse. The lowest fluences can be chosen to maximize single-photon excitation, then progressively increased to higher fluences. The higher the fluence, the greater the probability that two photons will be absorbed per chromophore, and the extent, chemical, and structural nature of any damage will be increased.

To conduct such an experiment, the desired size and nature of crystals, pump wavelength, pulse duration, and delay time are established. Time-resolved crystallographic data are then collected under conditions in which only the fluence of the pulse is varied, and the scattering data are carefully analyzed by structural refinement and DED maps. This description is easy to present, but the experiments are hard to conduct and the data even harder to analyze. For example, at the lowest fluences, the overall photolysis yield is reduced, all structural changes are diminished in magnitude and DED maps have low signal-to-noise. At higher fluences, the yield is increased but accompanied by an even larger temperature jump, whose peak magnitude approaches many hundreds of degrees K – but cools rapidly after the pulse.

Barends et al. (2024) recently took this approach to revisit the most widely studied light-dependent reaction in structural biophysics, the photolysis of tiny single crystals of the CO complex of myoglobin. Their paper is accompanied by a commentary by Neutze and Miller (2024), whose two authors present rather different views on the conclusions of Barends et al. (2024). The rupture of the covalent Fe-CO bond liberates CO, which diffuses in the heme pocket and eventually dissociates completely from the globin. Rupture also initiates structural changes in the heme group and in the surrounding helices. Earlier results by the same group (Barends et al., 2015) used high fluences, likely to generate significant absorption of two photons by each heme. The new results (Barends et al., 2024) identified substantial structural changes at and near the heme in the fs – few ps time range, whose magnitude and nature were dependent on the fluence. Most prominently, the photodissociated CO diffuses within the distal pocket, and surrounding helices move to varying extents.

The first fundamental question is: are the details of the structural changes influenced by high fluence? The comprehensive and inevitably complicated data analysis of Barends et al. (2024) strongly suggests that the answer is 'yes'. The conclusion clearly applies to the chemistry and biophysics of the myoglobin reaction. The authors infer that it extends to the details of all XFEL experiments at high fluence on crystals of other proteins. Indeed, they conclude 'We consider this to be a starting point for reassessing both the design and the interpretation of ultrafast TR-SFX pump-probe experiments such that mechanistically relevant insight emerges'.

If fully reversible, artefactual structural changes resulting from high fluence will disappear as the changes decay. If irreversible, these changes will remain even in the longest time points when authentic changes have ended and are no longer apparent in DED maps. Careful evaluation of earlier XFEL data on widely diverse proteins has observed only changes that are compatible with apparently authentic biological processes (Branden and Neutze, 2021). This suggests that any high fluence, two-photon damage is reversible. The data of Barends et al. (2024) do not extend to longer time points that could directly address reversibility.

The second fundamental question is: are these fluence-dependent structural details of biological (not just chemical and physical) significance? In the evolution of light-dependent reactions, the systems are likely to have been exposed only to low light fluences. Selection has harnessed the high energy available from absorbing a single visible photon from a low fluence source and used it to drive biologically desirable reactions such as isomerization or electron transfer. Equally, selection pressure may have mitigated any damaging effects arising from the absorption of more than one photon

https://doi.org/10.1017/S0033583524000222

from any source of exceptionally high light fluence. Indeed, Miller et al. (2020) suggest that 'It seems that no matter the starting point, … [bacteriorhodopsin] … is poised to channel energy into forming the structural intermediates most intimately involved in the key proton transfer step'. Barends et al. (2024) express a similar idea rather differently, citing the old adage 'all roads lead to Rome'.

Two-photon absorption is very likely to have occurred in almost all light-driven time-resolved XFEL experiments to date, conducted at high fluence. However, the structural details of fluence-dependent alterations to the mechanisms may decay in less than a few ps. That is, fluence-dependent mechanisms converge rapidly. Experiments that directly explore the effects of higher fluence at longer time points are essential to establish this critical issue. Such experiments will support – or oppose – the positive view that mechanisms at longer time points are not significantly affected and that earlier biological conclusions remain sound.

Laue and monochromatic diffraction by single crystals

Two features distinguish Laue diffraction from the classical diffraction of monochromatic X-rays. In Laue diffraction, the beam is polychromatic with an X-ray bandwidth that can be narrow or broad, and the crystal is stationary throughout the exposure. In monochromatic diffraction, the bandwidth is narrow and the crystal rotates during the exposure. During a single X-ray pulse at a storage ring or XFEL source (or a short pulse train of a few pulses at a storage ring), no significant rotation is possible. Thus *all* fast time-resolved crystallographic studies that utilize a single pulse or short pulse train give rise to Laue diffraction.

A typical Laue experiment at a storage ring source has a significant X-ray bandwidth of 1–5% if the entire first harmonic from an undulator is used. Each Laue diffraction pattern records a substantial volume of reciprocal space (Moffat et al., 1984) and contains many Laue spots, whose intensities are integrated over X-ray energy. The bandwidth may be reduced by X-ray optical elements in the beam such as a wide bandwidth monochromator (that is, a polychromator). In contrast, the X-ray beam at XFEL sources has a very narrow bandwidth of around 0.1% arising from the self-amplified stimulated emission (SASE) process at the heart of lasing (Emma et al., 2017). This bandwidth may be too narrow to integrate the intensities over energy. However, integration is required if accurate structure amplitudes are to be extracted. In the absence of integration, the intensity of each Laue spot is partial, less than the integrated intensity. Experimentally, the X-ray beam is tightly focused by a transmission lens of short focal length to illuminate small samples such as a microcrystal, which also minimizes background scattering from the crystal mount or liquid. Focusing introduces angular crossfire into the X-ray beam at the sample, which reduces or may even eliminate the partiality of each Laue spot. Novel approaches to diffraction on the as-to-fs time scale by a highly convergent X-ray beam are presented by Chapman et al. (2024).

The geometry of Laue diffraction also differs from that of monochromatic diffraction. Each mono spot arises from a single point (hkl) in reciprocal space. In contrast, each Laue spot arises from a ray, a central line in reciprocal space that contains many points between the origin of reciprocal space and the diffraction limit of the crystal. Although most Laue spots contain only one point, a minority of spots contain several points (Cruickshank et al., 1987). However, the structure amplitudes associated with individual reciprocal lattice points along the ray can be accurately determined by several algorithms and software packages. The important consequence is that Laue diffraction from storage ring sources can yield X-ray structure amplitudes as accurate as those from monochromatic rotation patterns (Ren and Moffat, 1995). Fortunately, this is also true for XFEL sources (Kirian et al., 2011).

Systematic errors

Static X-ray scattering from a tiny crystal or a dilute protein solution is inherently weak. The structural changes during the initial, fast stages of a biochemical reaction generally involve only small motions of a few atoms. This further reduces the time-resolved signal and may mask it by systematic errors. The structural changes are ultimately represented by difference electron density (DED) maps, also known as difference Fourier maps, or by differences in solution scattering. However, any systematic errors affecting the signal at individual time points are often correlated between time points. The highly desirable consequence is that the systematic errors in DED maps are much reduced from those in static maps. For this and other reasons (Moffat and Lattman, 2024), the signal-to-noise in DED maps is far superior to that in conventional, static electron density maps (Figure 4). A further result is that the systematic errors in time-resolved XFEL data are often superior to those in storage ring data, probably due to the smaller crystals and a larger extent of reaction initiation in each crystal.

Since the desired signal is time-dependent, all components of the experimental approach should minimize any time dependence of systematic errors in the background, such as trace X-ray scattering from slits and other beamline components. This constitutes noise. Radiation damage arising from X-ray absorption is an unavoidable, dose- and time-dependent systematic error that is generally irreversible. It is therefore superimposed on the authentic time-dependent structural changes throughout the entire time course. However, the time dependence of the primary component

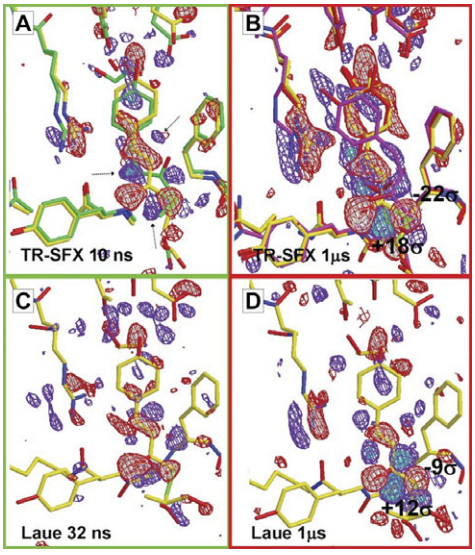

Figure 4. Time-resolved difference electron density (DED) maps of the chromophore region of photoactive yellow protein. DED maps at 1.6A resolution were obtained at the Linac Coherent Light Source XFEL (panel A at 10 ns after reaction initiation; panel B at 1 µs) and at BioCARS/Advanced Photon Source storage ring (panel C at 32 ns; panel D at 1 µs). Although positive (red) and negative (blue) features are in closely similar positions in the XFEL and BioCARS DED maps, those from XFEL data have superior signal-to-noise. Adapted from Figure 3 of Tenboer et al. (2014); reproduced with permission from the American Association for the Advancement of Science.

https://doi.org/10.1017/S0033583524000222

of radiation damage depends linearly on the absorbed dose. This differs from the expected exponential (or sum of exponentials) time dependence of the authentic signal and may enable radiation damage and signal to be distinguished.

Planning

All time-resolved experiments are complicated and cannot be conducted remotely, unlike today's more routine static experiments at storage ring sources. Planning experiments by prospective users are highly desirable and sought after by the panels that review applications for storage ring and XFEL beamtime. Planning experiments should confirm the retention of activity in the sample, establish overall feasibility, select the best experimental approach, and identify time periods of highest priority that provide the greatest - and least - chemical and biochemical interest. Good planning experiments are firmly in the interests of both the facilities and the users, so smaller amounts of beamtime to conduct them are routinely available. Their informative results support the award of larger amounts of beamtime to conduct the main experiments.

The mechanism of the light-dependent DNA repair flavoenzyme, DNA photolyase

Two excellent recent papers (Maestre-Reyna et al., 2023; Christou et al., 2023) by independent groups on the structure and mechanism of the DNA repair flavor enzyme, DNA photolyase, illustrate the present state of the art in time-resolved crystallography. An associated editorial commentary (Vos, 2023) sets the scene.

Upon illumination of DNA by near-UV light, the large amount of energy associated with the absorption of a UV photon frequently drives the formation of a cyclobutane pyrimidine dimer (CPD), in which two covalent bonds are formed between adjacent pyrimidines in the same DNA chain. This damage disrupts the double helical structure of DNA. Damage is subsequently repaired by the light-dependent flavoenzyme, DNA photolyase (PL). The light thus plays two roles. Absorption of the first highly energetic UV photon causes the original DNA damage. Much later, absorption of a second photon by PL bound to damaged DNA activates repair by PL and restores the authentic structure of double-stranded DNA.

Both groups explore an important question: the chemical and structural mechanism of DNA repair by PL. Many prior spectroscopic, theoretical, and computational studies of the repair reaction (cited by Maestre-Reyna et al., 2023; Christou et al., 2023) have been conducted on PLs from several species. All PLs contain a flavin adenine dinucleotide (FAD) coenzyme which *in vivo* absorbs a single near-UV photon. These studies showed that repair subsequently involves several fast, chemically distinct, elementary steps over the ps to ns time range (Figure 5; Christou et al., 2023). An initial, ultrafast, light-driven electron transfer via two single electron transfer steps photoreduces oxidized FAD (FADox) to the active, reduced, anionic quinone state ($FADH^-$). Photoreduction is promptly followed by further electron transfer from this state to the CPD. Multistep splitting of the two covalent bonds that form the CPD dimer recovers the two pyrimidine monomers, and backward electron transfer regenerates the active coenzyme. Over the time range from 500 ns to hundreds of μs, the two thymine bases, the tertiary structure of the PL active site, and numerous hydrogen bonds in the active site involving both the protein and water molecules rearrange to form undamaged but distorted, double-stranded DNA. The final steps in the time range from hundreds of μs to a few seconds involve the recovery of undistorted DNA, its dissociation from the active site, and the return of PL to its unilluminated, dark structure.

Despite this wealth of spectroscopic and chemical information, information from any species of PL was limited on the structural time course of the overall reaction from initial photoactivation to release of the repaired DNA product. Earlier structural studies used 100 ps X-ray probe pulses from storage ring sources and lacked the ps time resolution necessary to characterize short-lived chemical intermediates. The course of the reaction was also complicated by X-ray-induced photoreduction of the FAD. To achieve the desired time resolution, both groups therefore used X-ray pulses from XFEL sources in the tens of fs range. Based on the well-established pump – probe experimental approach, they applied time-resolved, serial femtosecond crystallography over the time range from a few ps to hundreds of μs. This range was sufficient to span the ultrafast chemical steps and structural changes in the chromophore, the damaged DNA substrate as it was repaired, the tertiary structure of the PL enzyme itself, and water molecules in its active site. The final elementary steps involving dissociation of repaired, undistorted DNA from PL at times >200 μs had to be inferred. Probably due to limitations on XFEL beamtime, neither group directly observed dissociation. Follow-up experiments will surely probe it.

A laser pump pulse initiated the repair process, followed by XFEL probe pulses at defined time points after initiation. Both groups used the Alvra endstation of the Aramis hard X-ray beamline at the SwissFEL. This beamline delivers 65–70 fs X-ray pulses of 11.98 keV at 100 Hz. Christou et al. (2023) collected data at all ten time points from 3 ps to 100 μs there. Maestre-Reyna et al. (2023) collected data from 100 ps to 10 ns there, and from 10 ns to 200 μs at the Japanese XFEL, SACLA, for a total of 18 time points.

The experiments themselves and the data analysis are very demanding. As is so often the case, the experiments generate further challenges. The goal here is to identify and discuss challenges that remain, both for this specific system and the experimental and analytical approach as a whole.

Both groups purified PL from the archaea *Methanosarcina mazei* and studied single crystals of PL to which a short, double-stranded stretch of DNA containing a synthetic *cis-syn* CPD thymine dimer had been bound in the dark, with or without anaerobic incubation. Crystal growth conditions, cell dimensions in the orthorhombic space group, and intermolecular contacts between the two molecules in the asymmetric unit are closely similar between the groups, but not identical. The crystallization buffer was supplemented by different high-viscosity additives to decrease the velocity of the injection jet into the X-ray beam, increase the frequency of 'hits' of X-ray pulses on a crystal, and hence the efficiency of reaction initiation. The additives were cellulose (Christou et al., 2023) or hydrophobic grease matrix (Sugahara et al., 2015; Maestre-Reyna et al., 2023). The effects - if any - of these additives on the rates of the elementary steps in the repair process are unknown. The temperatures in which the crystals were held immediately prior to injection are not specified but are sufficiently high to allow essential structural changes to proceed. Taken together, experimental differences between the groups suggest that attempts at direct comparison or formation of a unified data set from publicly available data may be challenging, but testable.

Qualitative DNA repair activity in the crystals is unequivocally demonstrated by the extensive, chemically plausible structural changes in PL protein, FAD, and DNA after illumination by the

https://doi.org/10.1017/S0033583524000222

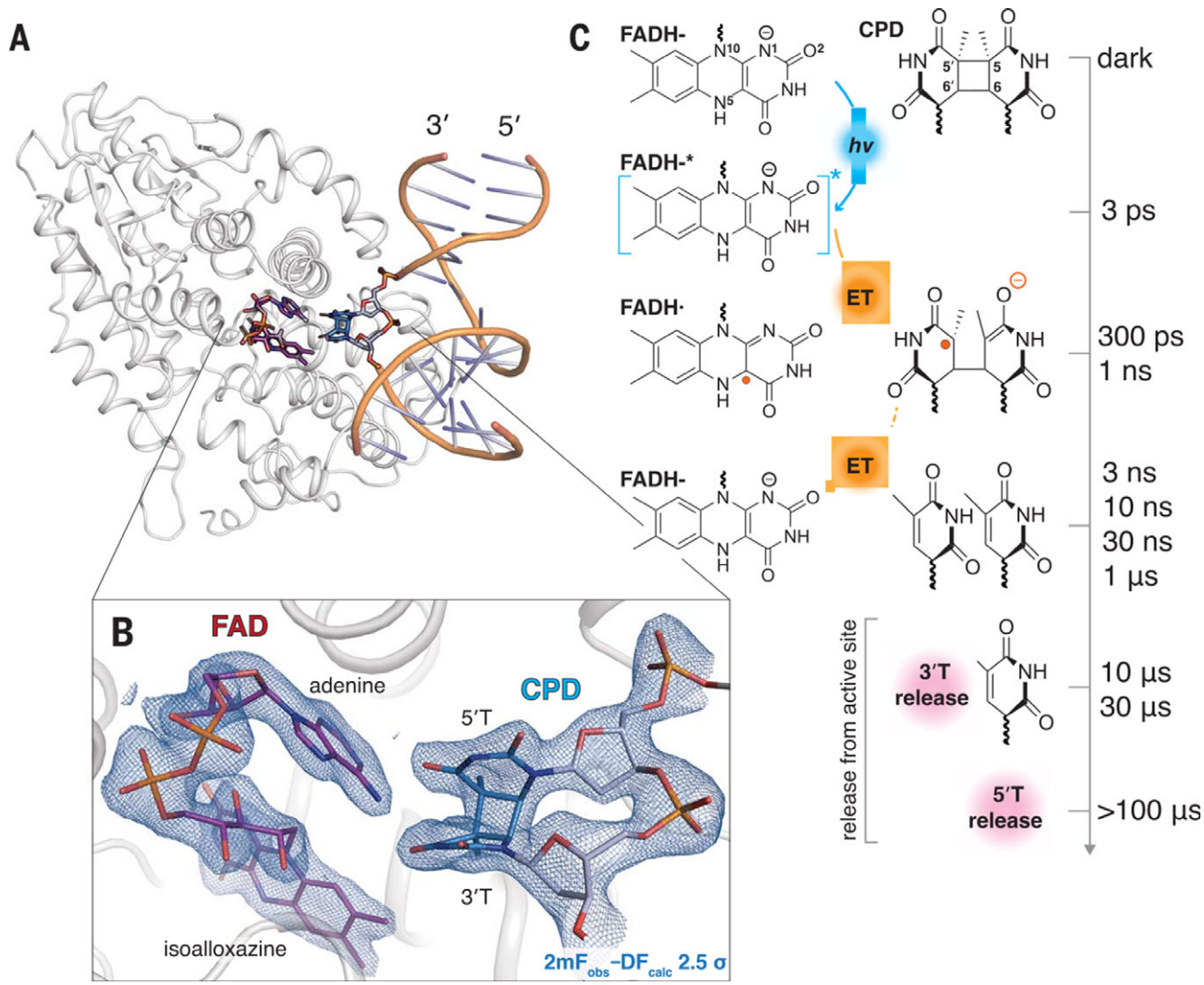

Figure 5. The DNA photolyase reaction. Panel A: Structure of photolyase PL cocrystallized with a dsDNA 14-mer containing a cyclobutane pyrimidine dimer (CPD). Panel B: electron density of the FAD cofactor (left) and the nearby CPD. Panel C: The FAD states (left column) and the CPD states (center column) as the repair reaction proceeds. Time points at which the reaction is probed are shown in the right column. Adapted from Fig. 1 of Christou et al. (2023); reproduced with permission from the American Association for the Advancement of Science.

pump light pulse. Quantitative aspects of the mechanism such as the identification of all elementary steps and the rate coefficients with which they interconvert are much more challenging to define. Refinement of candidate intermediate structures that comprise each elementary step is made very difficult by the coexistence at many time points of several structures whose populations vary with time. The structures themselves are time-independent and are assigned (the verb the authors use) to the chemical intermediates of the overall reaction proposed by prior spectroscopic, chemical, and computational studies in other species of PL (Figure 5). However, structural intermediates lack explicit labels. Although the overall mechanism is likely to be identical in all species, the rate coefficients with which intermediates form and decay, the time at which their peak concentrations are attained and their magnitude and detectability may differ significantly among the species, and be sensitive to apparently insignificant experimental details. Caution is necessary. The possibility that these prior studies on PL from different species mistook aspects of the mechanism is unlikely but cannot be discounted.

Prior spectroscopic information on the reaction of PL in solution and in a polycrystalline slurry under the exact sample conditions illuminated by the probe X-ray pulse would therefore have been extremely valuable. When available, this information enables informed planning of how to best distribute the limited XFEL beamtime among the time points. For example, the time course of substantial spectroscopic changes arising specifically from the FAD may be directly evident in the structure of the flavin and its immediate protein and water environment. However, the time courses of many spectroscopies and of X-ray scattering may differ significantly (Moffat and Lattman, 2024.) Many optical and X-ray spectroscopies (such as extended X-ray absorption fine structure, EXAFS) probe highly local aspects, sensitive only to the electronic structure of a cofactor such as FAD or the bonding of a metal. In contrast, X-ray diffraction and infrared spectroscopy are generally, sensitive to the electrons or vibrational transitions associated with all atoms in the asymmetric unit. The redox state of FAD may remain constant while its protein environment changes. If so, the time courses of a local spectroscopic probe and a general X-ray probe will differ.

https://doi.org/10.1017/S0033583524000222

Despite these limitations, prior spectroscopic studies identify time ranges when structures are likely to be easy to analyze. For example, an intermediate may be long-lived, structurally homogeneous over an extended time range, and thus easy to refine by conventional approaches from X-ray data over this range. Prior studies can also identify time ranges over which analysis is likely to be much more difficult. For example, when several structurally distinct intermediates co-exist in significant amounts, the data are heterogeneous and the component structures of each intermediate are particularly difficult to refine. Quantitative or semi-quantitative spectroscopic results thus aid in prioritizing the time points to be collected in the initial XFEL beamtime or postponed to later beamtime. Without these results, the experimenters are flying blind, as noted also by Vos (2023). Spectroscopic results provide an initial kinetic model to plan data collection, and analyze and refine the structural data. XFEL beamtime is hard to access and must be used to maximum effect. In the present case, it may be that the spirit of competition between the two groups and the need for prompt access to XFEL beamtime caused preliminary spectroscopy on this PL to be deferred.

Both groups assume without question that the overall mechanism is linear (state A to B to C to….). This is the simplest possible mechanism with an experimentally determined number of states and lacks complexities such as non-productive side or parallel reactions. The assumption cannot be established by the kinetic data alone. Both groups tacitly supplement their kinetic data with Occam's Razor: simplicity. In the absence of any data suggesting complexity, this assumption is very reasonable. A further assumption is that all elementary steps are irreversible. The back electron transfer step emphasized by Maestre-Reyna et al. (2023) is a separate elementary step at a later stage of the reaction, not a direct reversal of the forward electron transfer step. The assumption of irreversibility is very likely to be valid in the early stages of the overall reaction since absorption of a highly energetic pump photon drives the system far from equilibrium. The ensuing ultrafast reactions are strongly downhill energetically. Unless they arise from prompt, direct (and non-productive) decay of excited electronic states to the ground state, they are unlikely to be readily reversible. The two groups identify the same chemical intermediates and overall mechanism but not surprisingly, emphasize different features of the structural changes in each intermediate, and slightly different rate coefficients for their formation and decay. More experimental and theoretical research will be necessary to identify the key drivers in the mechanism.

The groups differ in their approaches to the initial structural analysis of the data, summarized in the comprehensive tables in their extensive Supplementary Materials. These Supplements are essential reading. Both use cautionary phrases, sometimes on topics where the main text is more confident. Both assemble the best structure amplitudes to a carefully determined limiting resolution, with extensive controls. The dark structures immediately prior to the pump pulse are refined and provide an initial set of phases. Raw DED maps identify the differences between each light structure at chosen times after the pump pulse and the relevant dark structure. Initially based on the simple DED approximation, the true phases of the difference structure factors are rapidly improved, ultimately obtaining phases for the light structure(s) at each time point, extrapolated to 100% occupancy from the experimental value of around 20% (Schmidt, 2019; Maestre-Reyna et al., 2022).

At this stage, the analysis strategies diverge. Christou et al. (2023) downplay the influence of structural heterogeneity at each time point and refine a single atomic structure into each extrapolated electron density map. Even if the structure at that time point is homogeneous, it's challenging to establish appropriate target geometries for the refinement. Structures may be highly strained in the FAD, the CPD, and intermediates in the sequential rupture of the two covalent bonds that constitute the dimer. This situation often occurs when refining very short-lived intermediates in which strain is to be expected. A further, potentially serious difficulty arises if, at that time point, the unknown intermediate is a heterogeneous mixture of two or more authentic, chemically plausible structures. A possible example here is the time-dependent bending and twisting of the isoalloxazine ring system of FAD, where a near-equimolar mixture of chemically plausible 'butterfly up' and 'butterfly down' structures could be closely approximated by a single, nearly planar structure. This situation is illustrated in Figure 2 of Christou et al. (2023). The approximation might reveal slightly larger B-factors for constituent atoms at the tips of the ring system.

In contrast, Maestre-Reyna et al. (2023) directly take structural heterogeneity into account, referring to candidate intermediate states with words like 'mixture' and phrases such as 'predominant species'. For example, they developed a kinetic model for CPD repair, presented in their Figs. S14B and S14C and reproduced here as Figure 6. The model is based on the integration of negative DED features in real space, specifically associated with the rupture of the two covalent bonds constituting the CPD and rearrangement of the thymine monomers during and after rupture. They label the resultant four intermediate structures as Int 2, Int 3, Int 4, and Int 5, which participate in extensive, time-dependent heterogeneity. Three species Int 2, Int 3, and Int 4 are simultaneously present in the time range around 450–650 ps, with Int 3 predominating near 650 ps; a different three species Int 3, Int 4, and Int 5 occur around 1.2 ns – 2.2 ns where rupture of the second bond occurs; and two species Int 4 and Int 5 occur at times greater than 2.5 ns as rearrangement of the thymine bases nears completion (Figures 5 and 6). Even though this chemically important region is spatially localized, confident refinement of its constituent atomic structures is not going to be straightforward.

The preliminary structural data, chemical plausibility, and computational studies suggest that most early time points are heterogeneous. The approach of Maestra-Reyna et al. (2023) is realistic and more challenging than that of Christou et al. (2023), but all is not lost. When more data at time points in the ps to tens of ns range is acquired, this will enable more confident refinement of reasonable structures that populate this range. The good news is that several adjacent time points contain excellent structural information on each of the homogeneous, intermediate structures present, whose populations vary smoothly over these time points. Authentic structural signals vary smoothly in time but noise flickers. In an ideal world and to accommodate uncertainties in initial kinetic models, time points should be collected uniformly in log t, with at least three points per decade across the range. Assuming initially a simple, linear, irreversible mechanism, steps in which rate coefficients are refined could be interspersed with structural refinement.

In the absence of this additional data, the present, refined sets of coordinates should be regarded as preliminary, and may not be satisfactory starting points for molecular dynamics calculations.

A remaining point of contention is the likelihood that the pump pulses used by both groups are sufficiently powerful that unphysiological two-photon excitation of the absorbing species occurs. As noted above, this could well generate artefacts. Both groups recognized the problem and conducted preliminary titrations in which the power/fluence of the pump pulse was varied. Unfortunately,

https://doi.org/10.1017/S0033583524000222

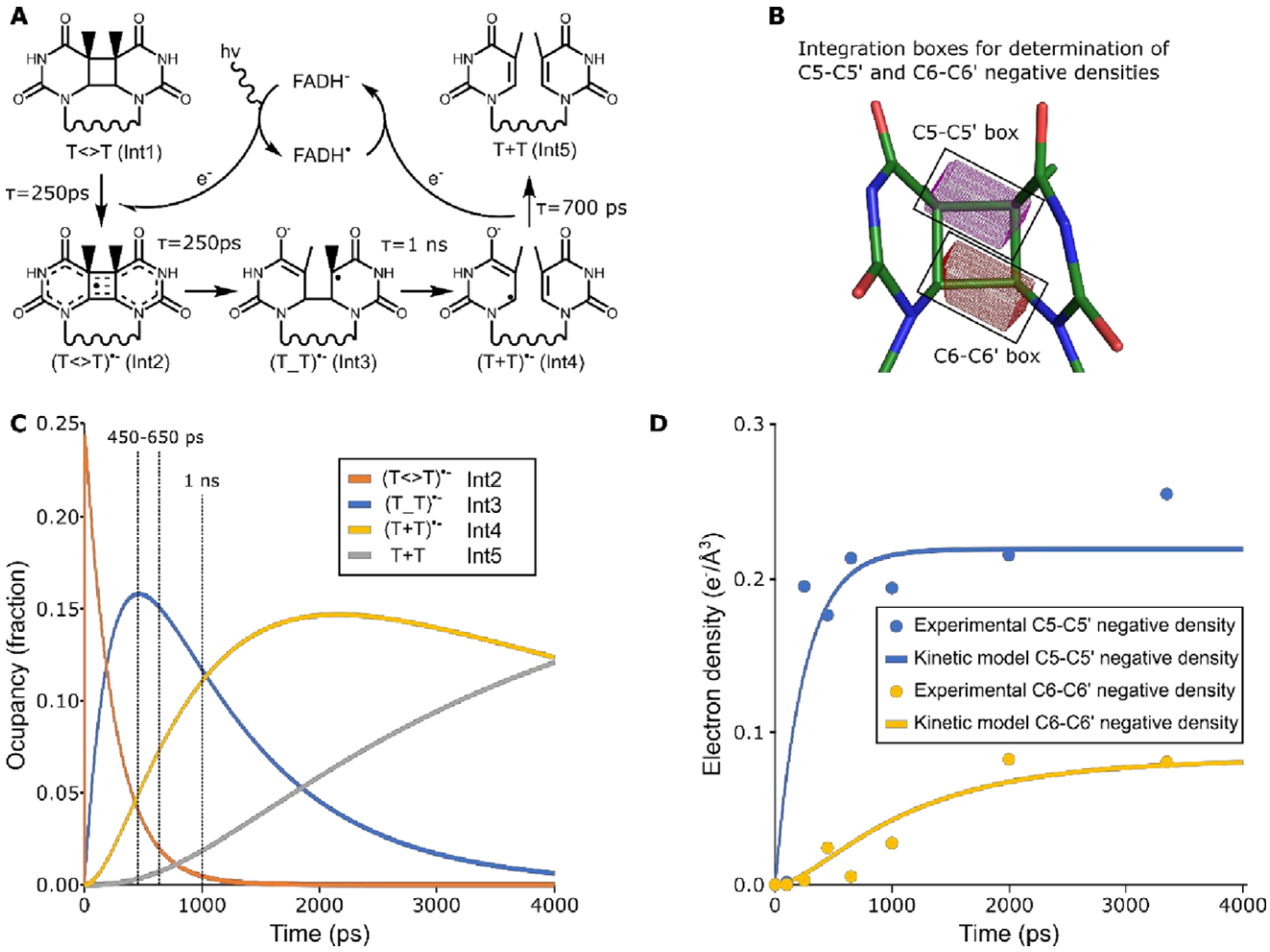

Figure 6. Modelled time course of the elementary steps involving rupture of the two covalent bonds forming the thymine dimer T<>T in DNA photolyase. Panel A: Kinetic model for the rupture, with time constants for each first-order reaction. Panel B: Location of integration boxes surrounding the DED features associated with rupture of the C5-C5' and C6-C6' bonds forming the thymine dimer. Panel C: Time-dependent concentration of intermediates I1 – I5 in panel A, based on the numerical solution of the system with the time constants in panel A. Panel D: Experimental (as in panel B) and modelled (as in panel C) DEDs associated with rupture of the covalent bonds. These data and modeling suggest that rupture of the C5-C5' bond occurs prior to that of the C6-C6' bond. Adapted from Fig. S13 of Maestre-Reyna et al. (2023); reproduced with permission from the American Association for the Advancement of Science.

constraints on X-ray beamtime limited the range that could be investigated. In common with studies on other proteins (reviewed by Branden and Neutze, 2021), no structural changes attributable to two-photon excitations were detected. As discussed above, any irreversible damage would be present at all time points; no such damage was seen. Fully reversible damage might have been limited to time points in the fs time range, which were not explored by either group.

The fundamental experimental dilemma is clear. At higher pump power/fluence, structural changes generate very significant features in DED maps (such as those in Figure 4 believed to be authentic), but two-photon excitation is much more probable. As the power is reduced, two-photon excitation becomes much less probable but the features diminish in magnitude and ultimately become insignificant. The ability to publish results, to be awarded more FEL beamtime, and to obtain continued grant support all depend on significant structural changes evident in the DED maps. The temptation is to increase the pump power until significant features appear in the DED maps, then back off slightly.

An unaddressed issue is that the probe pulse at the sample may retain partial coherence, despite the complexities introduced by X-ray optical elements such as tightly-focusing transmission mirrors. The source is, after all, a laser. The important question of how X-ray diffraction by a macroscopic crystal is affected by the coherence of the beam remains to be addressed.

Looking forward

The overall experimental design of the PL experiments by both groups is clearly sound. One obvious limitation can be addressed by conducting spectroscopic experiments of this specific PL on a dilute solution or even more informative, on a polycrystalline slurry. The slurry should match as closely as possible the experimental conditions to be used in subsequent XFEL experiments. Similar experiments could explore the effects of crystallization buffer, high-viscosity additives, and temperature. More substantial cooling is unlikely to substantially alter the mechanism but will change rate coefficients. This may reveal intermediates that were undetectable in the initial experiments.

The asymmetric unit in the PL crystals studied by both groups contains two molecules in the asymmetric unit that differ in their intermolecular contacts. One molecule is more ordered than the

https://doi.org/10.1017/S0033583524000222

other, and interpretation by both groups naturally focusses on the better-ordered. Are the time-dependent DED maps (roughly) identical for the two molecules? Or, are there differences in DED maps between the molecules, in particular time ranges (other than immediately prior to dissociation of the repaired DNA)? Significant differences in DED maps suggest differences in rate coefficients between the two molecules, which are likely to arise entirely from differences in their intermolecular contacts. That is, the two molecules in the asymmetric unit proceed at different rates down the reaction coordinate. A single physical time point contains molecules at two distinct, chemical points on the reaction coordinate.

The usual treatment of single crystal diffraction assumes an incoherent X-ray source that delivers an incoherent X-ray beam to the crystal. However, both the fs visible pump and the X-ray XFEL probe emit pulses that are at least partially coherent. The coherence and polarization properties of the pulses incident on the sample may be modified from those at the sources by X-ray or visible optical elements in the beamline, forexample by elements such as tightly-focusing reflection or transmission X-ray mirrors in the Aramis beamline at the Swiss XFEL and the SACLA XFEL used in the studies of DNA photolyase The coherence and polarization properties of the pump and probe pulses at the crystal and the extent of variation from pulse to pulse were apparently not measured experimentally. These properties and their variation could not be incorporated into the extraction of time-dependent structure amplitudes from the raw intensities of the Laue spots.

A further complication arises from the strongly anisotropic optical properties of many crystals that contain chromophores (see for example the effects of its p-hydroxycinnamic acid chromophore on the spectra of single crystals of photoactive yellow protein (Ng et al., 1995). The extent of reaction initiation of each crystal thus depends on the relationship between the polarization of the pump pulse and the orientation of the symmetry axes of that crystal. This factor could be extracted from each indexable diffraction pattern but at present is ignored.

The most serious limitation lies in the data analysis and thus is general to all time-resolved experiments. It must be actively explored. Interesting aspects of the mechanism are very likely to occur at time points where several structural species are simultaneously present, as in Figure 6. Although the sample at these time points is heterogeneous, the structures of its individual, homogeneous components are required (Ren and Yang, 2024). In conventional crystallography, heterogeneity is almost entirely absent and a single atomic structure can be satisfactorily refined against time-independent X-ray structure amplitudes. In time-resolved crystallography, heterogeneity in some important time ranges is very likely, even inevitable as with PL. This situation requires that several structures must be refined against structure amplitudes obtained at a series of adjacent time points, where the population of each time-independent structure differs from time point to time point. An advantage is that the populations vary smoothly with time and are correlated; the sum of the populations is constant. For example, if only two structures are present, a decrease in the first population is exactly matched by an increase in the second population.

Prior spectroscopic results may give sound estimates of the variation with time of a chromophore structure. A disadvantage is that the structures are closely related and the small structural differences between them will be difficult to identify. This suggests that refinement approaches that rely on differences between adjacent time points in DED maps may be more powerful than those that rely on extrapolated electron density maps at individual time points. A second disadvantage is that short-lived structures such as the excited flavin chromophore in PL may be highly strained in unusual ways which depend on details of the protein environment, and the nature of the strain varies from structure to structure. With chemical intuition supplemented by molecular dynamics simulation and calculation, structural constraints may be thoughtfully relaxed.

A limitation of ultrafast time-resolved crystallography is that the number of naturally light-sensitive target systems is limited. In all known examples, sensitivity to wavelengths found in sunlight is conferred by binding a non-protein chromophore such as retinal, flavin, chlorophyll, or tetrapyrrole. PL provides an excellent example of a flavin-binding enzyme that exploits its light-dependent properties to drive electron transfer and promote unusual chemistry. LOV and BLUF proteins are flavin-based signaling photosensors, in which absorption of light by the flavin generates a signal: a structural change that alters the affinity of the sensor domain for a downstream effector. In one example, replacing the oxygen-sensitive domain of a histidine kinase with a light-sensitive LOV domain generates a light-sensitive histidine kinase (Moeglich et al., 2009).

A relatively straightforward strategy to generate useful, chemically novel, light-sensitive targets is to modify the specificity of natural enzymes such as Old Yellow Enzyme, a flavin mononucleotide-dependent reductase. The initial stages of its mechanism closely resemble those of PL (Li et al., 2024). Photoinduced electron transfer from the flavin generates a substrate in high energy, unstable state, followed by the departure of a leaving group, radical formation, and capture. More generally, photoenzymes offer the promise of extending enzymatic catalysis into synthetically novel reactions (Harrison et al., 2022).

As more structural and detailed mechanistic information is obtained through studies of natural photosystems such as PL, it may become possible to design photoenzymes *de novo*. Flavins offer the most promising chromophore on which to base such molecules. They are readily bioavailable, naturally absorbed in the blue region of the spectrum, and undergo light-dependent redox chemistry that is beginning to be exploited.

All known signaling photosensors and photoenzymes are proteins. An admittedly speculative question (Moffat et al., 2013) is: do photosensors or photoribozymes based on RNA remain to be discovered? Flavin-based RNA riboswitches play a role in RNA metabolism (Serganov et al., 2009; Crielaard et al., 2022), but their sensitivity to light does not appear to have been looked for. There's no obvious chemical reason why RNA photosensors or photoribozymes do not exist. It would be fun to look.

A final, general issue is ethical rather than scientific. Access to XFEL beamtime worldwide is very competitive, the number of beamlines optimized for ultrafast crystallography is small, and the number of suitable light-sensitive systems is limited. Head-to-head scientific competition between user groups at the same XFEL beamline, studying the same system, is inevitable.

Beamline staff are extremely well informed about the technical aspects of users' experiments and often, are also highly knowledgeable about the scientific aspects. In their role as staff, they are expected to treat all users equally. If asked, they provide their best advice on details of how to conduct an experiment, or what details to avoid. The best advice to this week's user could be based on knowing exactly what last week's user discovered – but last week's user might be a competitor. Contributions by staff are frequently extensive enough to warrant inclusion as co-authors of the resultant

https://doi.org/10.1017/S0033583524000222

papers. Staff also play a role as essential collaborators, without whom the research could not be conducted.

A potential difficulty is that the dual roles of beamline staff as advisers to users and as a collaborator and prospective co-author may not align. The possibility of a conflict of interest between these roles arises. If so, how can the conflict be resolved? As a former beamline director and university administrator responsible for identifying and resolving conflicts of interest, similar issues occasionally arose. (There is NO evidence that any ethical issues arose between the two groups studying PL, or between the user groups and beamline staff. Indeed, both groups clearly received expert, well-informed, and essential advice from beamline staff at the Swiss FEL and SACLA, appropriately recognized by co-authorship.)

Acknowledgements. I thank two anonymous referees who provided thoughtful suggestions on how to address this fast-moving field. Andrew Nairne and Inga Frasier of Kettle's Yard (University of Cambridge) supplied Figure 1 and gave permission to use it; Tobin Sosnick (University of Chicago) supplied Figure 2 and gave permission to use it; Figure 3 is reproduced with permission of Springer Nature/Licensed under CC BY 4; permission to use Figure 5 was graciously given by TJ Lane (DESY); and permission to use Figure 6 was graciously given by Ming-Daw Tsai (National Taiwan University) and Lars-Oliver Essen (Marburg University). Figures 4, 5 and 6 are reproduced with permission from the American Association for the Advancement of Science. I thank all of them.

Financial support. Research by the author was funded over many decades by grants from the National Institute of General Medical Sciences, NIH. No grant from any funding agency, commercial or not-for-profit sectors funded preparation of this review.

Competing interest. The author declares no competing interests exist.

References

Baek MF, DiMaio F, Anishchenko I *et al.* (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**(6557), 871–876.

Barends TRM, Foucar L, Ardevol A *et al.* (2015) Direct observation of ultrafast collective motions in CO myoglobin upon ligand dissociation. *Science* **350**(6259), 445–450.

Barends TRM, Gorel A, Bhattacharya S *et al.* (2024) Influence of pump laser fluence on ultrafast myoglobin structural dynamics. *Nature* **626**(8000), 905–911.

Besaw JE and Miller RJD (2023) Addressing high excitation conditions in time-resolved X-ray diffraction experiments and issues of biological relevance. *Current Opinion in Structural Biology* **81**, 102624.

Bock LV and Grubmueller H (2022) Effects of cryo-EM cooling on structural ensembles. *Nature Communications* **13**, 1–13.

Branden G and Neutze R (2021) Advances and challenges in time-resolved macromolecular crystallography. *Science* **373**(6558), eaba0954.

Chance B, Ravilly A and Rumen N (1966) Reaction kinetics of a crystalline hemoprotein: An effect of crystal structure on reactivity of ferrimyoglobin. *Journal of Molecular Biology* **17**, 525–534.

Chapman HN, Barty A, Bogan MJ *et al.* (2006) Femtosecond diffractive imaging with a soft X-ray free electron laser. *Nature Physics* **2**(12), 839–843.

Chapman HN, Fromme P, Barty A *et al.* (2011) Femtosecond x-ray protein nanocrystallography. *Nature* **470**(7332), 73–77.

Chapman HN, Li C, Bajt S *et al.* (2024) Convergent-beam attosecond X-ray crystallography. *Structural Dynamics*, in press.

Christou NE, Apostolopoulou V, Meio DVM *et al.* (2023) Time-resolved crystallography captures light-driven DNA repair. *Science* **382**, 1015–1020.

Crielaard S, Maassen R, Vosman T *et al.* (2022) Affinity-based profiling of the flavin mononucleotide riboswitch. *Journal of the American Chemical Society* **144** (23), 10462–10470.

Cruickshank DWJ, Helliwell JR and Moffat K (1987) Multiplicity distribution of reflections in Laue diffraction. *Acta Crystallographica* **A43**, 656–674.

Dashti A, Mashayekhi G, Shekhar M *et al.* (2020) Retrieving functional pathways of biomolecules from single particle snapshots. *Nature Communications* **11** (1), 1–14.

Dods R, Bath P, Morozov D *et al.* (2021) Ultrafast structural changes within a photosynthetic reaction centre. *Nature* **589**, 310–314.

Edlund P, Takala H, Claesson E *et al.* (2016) The room temperature crystal structure of a bacterial phytochrome determined by serial femtosecond crystallography. *Scientific Reports (Nature)* **6**, 35279.

Emma C, Lutman A, Guetg M *et al.* (2017) Experimental demonstration of fresh bunch self-seeding in an X-ray free electron laser. *Applied Physics Letters* **110** (15), 154101.

Frank J and Ourmazd A (2016) Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. *Methods* **100**, 61–67.

Graves W, Chen J, Fromme P *et al.* (2020) ASU compact XFEL. In *Proceedings of the 38th International Free Electron Laser Conference*. Geneva, Switzerland.

Gruhl T, Weinert T, Rodrigues MJ *et al.*, (2023) Ultrafast structural changes direct the first molecular events of vision. *Nature* **615**, 939–944.

Grunbein ML, Stricker M, Nass Kovacs G *et al.* (2020) Illumination guidelines for ultrafast pump-probe experiments by serial femtosecond crystallography. *Nature Methods* **17**, 681–684.

Halle B (2004) Biomolecular crystallography: Structural changes during flash cooling. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4793–4798.

Harrison W, Huang X and Zhao H (2022) Photobiocatalysis for abiological transformations. *Accounts of Chemical Research* **55**, 1087–1096.

Hekstra DR (2023) Emerging time-resolved X-ray diffraction approaches for protein dynamics. *Annual Review of Biophysics* **52**, 255–274.

Jumper J, Evans R, Pritzel A *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596** (7873), 583–589.

Khusainov G, Standfuss J and Weinert T (2024) The time revolution in macromolecular crystallography. *Structural Dynamics* **11**, 020901.

Kirian RA, White TA, Holton JM *et al.* (2011) Structure-factor analysis of femtosecond micro-diffraction patterns from protein microcrystals. *Acta Crystallographica* **A67**, 131–140

Lee DD and Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791.

Li M, Yuan Y, Harrison W *et al.* (2024) Asymmetric photoenzymatic incorporation of fluorinated motifs into olefins. *Science* **385**, 416–421.

Lomb L, Barends TRM, Kassemeyer S *et al.* (2011) Radiation damage in protein serial femtosecond crystallography using an X-ray free-electron laser. *Physical Review B: Condensed Matter and Materials Physics* **84**, 214111.

Lorenz, UJ (2024) Microsecond time-resolved cryo-electron microscopy. *Current Opinion in Structural Biology* **98**, 102840.

Maestre-Reyna M, Yang CH, Nango E *et al.* (2022) Serial crystallography captures dynamic control of sequential electron and proton transfer events in a flavoenzyme. *Nature Chemistry* **14**, 677–685.

Maestre-Reyna M, Wang P-H, Nango E *et al.* (2023) Visualizing the DNA repair process by a photolyase at atomic resolution. *Science* **382**, eadd7795.

Malla TN and Schmidt M (2022) Transient state measurements on proteins by time-resolved crystallography. *Current Opinion in Structural Biology* **74**, 102376.

Miller RJD, Pare-Labrosse O, Sarracini A *et al.* (2020) Three-dimensional view of ultrafast dynamics in photoexcited bacteriorhodopsin in the multi-photon regime and biological relevance. *Nature Communications* **11**, 1240

Moeglich A, Ayers RA and Moffat K (2009) Design and signaling mechanism of light-regulated histidine kinases. *Journal of Molecular Biology* **385**, 1433–1444.

Moffat K and Henderson R (1995) Freeze trapping of reaction intermediates. *Current Opinion in Structural Biology* **5**, 656–663.

Moffat K and Lattman EE (2024) *Dynamics and Kinetics in Structural Biology: Unravelling Function through Time-resolved Structural Analysis*. Chichester: John Wiley and Sons Ltd.

Moffat K, Szebenyi D and Bilderback D (1984) X-ray Laue diffraction from protein crystals. *Science* **223** (4643), 1423–1425.

Moffat K, Zhang F, Hahn K *et al.* (2013) The biophysics and engineering of signaling photoreceptors. In Hegemann P and Segrist S (eds.), *Optogenetics*. Berlin: De Gruyter, pp. 7–22.

Neutze R and Miller RJD (2024) Energetic laser pulses alter outcomes of X-ray structures of proteins. *Nature* **626**, 720–722.

Neutze R, **Wouts R**, **van der Spoel D** ***et al.*** (2000) Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406** (6797), 752–757.

Ng K, **Getzoff ED and Moffat K** (1995) Optical studies of a bacterial photoreceptor protein, photoactive yellow protein, in single crystals. *Biochemistry* **34** (3), 879890.

Nogly P, **Weinert T**, **James D** ***et al.*** (2018) Retinal isomerization in bacteriorhodopsin captured by a femtosecond X-ray laser. *Science* **361** (6398), eaat0094.

Norden B (2021) Which are the 'Hilbert Problems' of biophysics? *Quarterly Reviews Biophysics Discovery* **2**, e1.

Orville AM (2020) Recent results in time resolved serial femtosecond crystallography at XFELs. *Current Opinion in Structural Biology* **65**, 193–208.

Pearson A and Mehrabi P (2020) Serial synchrotron crystallography for time-resolved structural biology. *Current Opinion in Structural Biology* **65**, 168–174.

Peng X, **Baxa M**, **Faruk N** ***et al.*** (2022) Prediction and validation of a protein's free energy surface using hydrogen exchange and (importantly) its denaturant dependence. *Journal of Chemical Theory and Computation* **18**, 550–561.

Perman B, **Srajer V**, **Ren Z** ***et al.*** (1998) Energy transduction on the nanosecond time scale: Early structural events in a xanthopsin photocycle. *Science* **279**, 1946–1950.

Ren Z and Moffat K (1995) Deconvolution of energy overlaps in Laue diffraction. *Journal of Applied Crystallography* **28** (5), 482–494.

Ren Z and Yang X-J (2024) Deconvolution of dynamic heterogeneity in protein structure. *Structural Dynamics* **11**, 041302.

Scheres, SHW (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* **180**(3), 519–530.

Schmidt M (2019) Time-resolved crystallography at pulsed x-ray sources. *International Journal of Molecular Sciences* **20** (6), 1401.

Schmidt M (2023) Practical considerations for the analysis of time-resolved x-ray data. *Structural Dynamics* **10**, 044303.

Schmidt M, **Rajagopal S**, **Ren Z** ***et al.*** (2003) Application of singular value decomposition to the analysis of time-resolved macromolecular x-ray data. *Biophysical Journal* **84**, 2112–2129.

Schon JC (2024) Energy landscapes past, present and future: A perspective. *The Journal of Chemical Physics* **161**, 050901.

Schotte F, **Cho HS**, **Kaila VRI** ***et al.*** (2012) Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19256–19261.

Serganov A, **Huang L and Patel DJ** (2009) Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* **458** (7235), 233–237 (2009).

Skopintsev P, **Ehrenberg D**, **Weinert T** ***et al.*** (2020) Femtosecond-to-millisecond structural changes in a light-driven sodium pump. *Nature* **583** (7815), 314–318.

Srajer V, **Teng T-Y**, **Ursby T** ***et al.*** (1996) Photolysis of the carbon monoxide complex of myoglobin: Nanosecond time-resolved crystallography. *Science* **274**, 1726–1729.

Sugahara M, **Mizohata E**, **Nango E** ***et al.*** (2015) Grease matrix as a versatile carrier of proteins for serial crystallography. *Nature Methods* **12**, 61–63.

Szebenyi D, **Bilderback D**, **LeGrand A** ***et al.*** (1988) 120 picosecond Laue diffraction using an undulator X-ray source. *Transactions of the American Crystallographic Association* **24**, 167–172.

Tenboer J, **Basu S**, **Zatsepin N** ***et al.*** (2014) Time-resolved serial crystallography captures high resolution intermediates of photoactive yellow protein. *Science* **346**, 1242–1246.

Vos MH (2023) Filming DNA repair at the atomic level. *Science* **382**, 996–997.

https://doi.org/10.1017/S0033583524000222