

Minding the dynamic gap: measuring ultrafast processes in biomolecular systems

Neil T. Hunt

(University of York, UK)

Biological processes *in vivo* involve interactions between large, complex molecules under conditions that are best-approximated by a viscous, multicomponent solution at temperatures often somewhat above room temperature. This means that the molecules involved are dynamic — their structure changes — on many timescales in ways that range from very small, localized fluctuations to fundamental changes in their overall appearance. What roles do these dynamics play in the structure–function relationship? Do very fast, local motions have any impact on well-characterized, slower structure changes? How do we go about measuring the very fastest biomolecular fluctuations to find out? Here we discuss ultrafast multidimensional (2D-IR) spectroscopy and look at the complementary information that it provides as part of a raft of biophysical experiments.

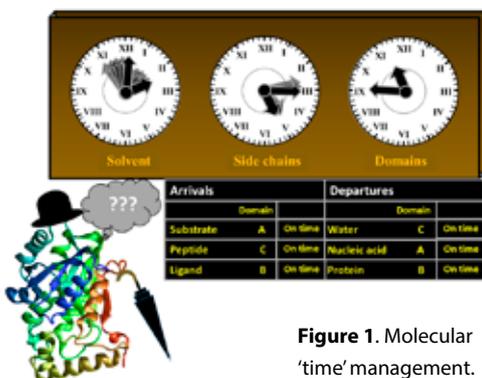


Figure 1. Molecular 'time' management.

Proteins are complex molecules. Beginning from the primary, covalently bonded, sequence of amino acids the functional protein molecular structure is constructed from secondary structural elements, such as α -helices or β -sheets that are held together by electrostatic interactions and hydrogen bonds. These secondary structures then form the basis of higher order, tertiary and quaternary structural arrangements. On top of this, even the simplest laboratory samples involve solvation of the protein in a bath of water molecules and ions. It is common to represent biological molecules visualized via crystallography experiments, as static three-dimensional structures. Indeed, the structure–function relationship is a central tenet of biological reasoning. However, at room temperature, these molecules are dynamic on timescales that span several orders of magnitude.

Dynamic processes begin with the biological solvent itself. At a chemical or physical level, water is more complex than other 'simpler' liquids. At the root of this complexity is the hydrogen bond, which accounts for water's unusually high boiling point, for such a small molecule, and its unique density-dependence upon temperature (the reason that ice floats). Indeed, the detailed molecular behaviour of water remains a topic of significant interest

to physical chemists and spectroscopists, but it is clear that, at room temperature, water molecules are constantly exchanging H-bond partners on the timescale of about a picosecond (10^{-12} s or a millionth of a millionth of a second). In addition, the amount of energy available at room temperature means that a large number of low frequency intermolecular vibrational modes of water are active, so water is a highly dynamic medium on ultrafast (picosecond and faster) timescales.

For a protein, interactions with water molecules are crucial, whether this is solvation of the protein's outer surface to prevent aggregation or specific H-bond interactions between water and amino acid residues and side chains within the structure. The fact that biological molecules require a certain level of hydration in order to become active demonstrates the importance of these solvent molecules. However, the same dynamic motions that are found in bulk water will influence protein structure. H-bond exchange and vibrational motion will lead to a range of seemingly-random ultrafast fluctuations of the protein on picosecond timescales. From the perspective of a potential energy surface, which defines the preferred equilibrium structure of the protein, this means that rather than a simple surface with a single minimum corresponding to one structure, the surface is rugged with many similar structures separated by energy barriers small enough to be surmounted at room temperature. This means that the protein structure fluctuates on small (Ångstrom: $1 \text{ \AA} = 0.1 \text{ nm}$) length scales on ultrafast timescales.

We know that proteins change structure on many different timescales, NMR spectroscopy has played an

important role in understanding conformational dynamics typically occurring in the millisecond (1000^{th} of a second) range. These can involve movement of disordered regions of protein structures or side chain motion. Experiments such as Förster resonance energy transfer (FRET) or dynamic light scattering can also report on variations in macromolecular structure, while crystallography highlights equilibrium structural distributions.

The question is how do these large-scale dynamic processes, taking milliseconds and longer, relate to the ultrafast, Ångstrom-level fluctuations of H-bonds? Alternatively, does the dynamic motion of proteins control ligand or substrate binding, or do ligands and substrates modulate protein dynamics in a meaningful manner? As yet, we don't have a unified picture, though rather than these processes on disparate time and length-scales being completely de-coupled, evidence is growing that the fast fluctuations on the rugged ultrafast dynamic landscape represent a conformational searching process that ultimately acts as a trigger for larger-scale functional changes. Similarly, evidence is beginning to emerge that ligand or drug binding alters the fast dynamics of the protein molecule, which may have a downstream impact on their conformational searching or flexibility, potentially acting to modulate function remotely as in allosteric processes.

Currently, computational molecular dynamics (MD) simulations dominate our knowledge of the behaviour of proteins on ultrafast timescales. These methods simulate molecular motion using femtosecond (1000^{th} of a picosecond, 10^{-15} s) steps, but while powerful and widely applied, experimental validation is difficult to come by and so the predictive potential of such tools is not being fully realized. Furthermore, these methods place high demands on computing resources, meaning that high accuracy simulations are often limited to nanosecond (10^{-9} s) durations. The result is a dynamic gap between the end of computational simulations and the start of our experimental methods (Figure 2). Here we describe an experimental approach, two-dimensional infrared spectroscopy, which may assist with probing this dynamic gap and reveal how the dynamic motions of proteins link together over time and length scales to give a new picture of biological function.

Infrared (IR) spectroscopy is often applied to study proteins because the C=O bonds that occur in every peptide linkage between amino acid residues exhibit stretching vibrational modes (so-called 'amide-I modes') that are intense and sensitive to the structural and solvation environment. By this it is meant that the precise frequency of the absorption varies diagnostically upon whether the C=O bond is involved in an α -helix, β -sheet or random coil, while its linewidth reports on the range of H-bonding environments to which the bond is subjected. As described above, the local water molecules and the rest

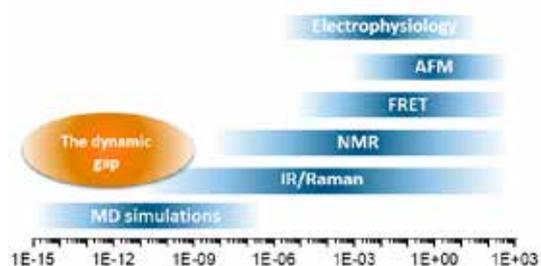


Figure 2. Schematic diagram of the dynamic gap. The timescale ranges of biophysical experiments are shown in blue (in seconds) while the orange region highlights the shortfall between experimental methods and MD simulations.

of the protein architecture near a given peptide residue will be moving constantly, generating a range of different protein structures. If we were to take a single snapshot of the protein sample, we would find molecules in each of these structures, but if the snapshots were joined together to form a 'molecular movie' an individual protein molecule would explore them all within a few picoseconds.

IR absorption spectroscopy thus contains a lot of information, but it cannot deliver time-resolved insight, recording as it does a time-averaged picture of the sample. By contrast, 2D-IR methods enable us to probe beneath the broad, sometimes featureless, IR absorption profile to gain more insight. 2D-IR is the infrared equivalent of the 2D-NMR measurements that have proved immensely powerful in structural biology. Similarly to 2D-NMR, a series of two, or more, IR laser pulses are directed into the sample with precisely defined time delays between them. The result is to spread the 1D-IR absorption spectrum out over a 2nd frequency axis (Fig. 3). Like 2D-NMR, the IR absorption spectrum is found along the diagonal of the 2D-IR plot while new information appears in the off-diagonal region. In the case of 2D-IR, these off-diagonal peaks tell us about vibrational coupling, molecular structure and energy transfer, all of which relate to either protein conformation or chemical connectivity.

Practically, the 2D-IR experiment can be thought of as the first laser pulse exciting the sample at a particular frequency while the second pulse measures the effect of this on the molecule. By scanning the frequency of the first pulse, a 2D plot of excitation versus detection frequency is built up. Most importantly for studies of dynamics, the time delay between the laser pulses allows us to examine how the protein's structure changes with time. The time resolution (the fastest resolvable event) of a 2D-IR experiment is defined mainly by the duration of the laser pulses, which is typically 50 fs. Therefore, 2D-IR provides structural information in a way that mimics a high-speed camera, capable of observing dynamics as fast as H-bond exchange.

There are two main ways in which 2D-IR spectroscopy has been used to study protein dynamics, using a specialized probe group to measure dynamics in a specific place of interest within a protein structure or in an unlabelled way, studying the protein as a whole. We describe examples of each on the following pages.

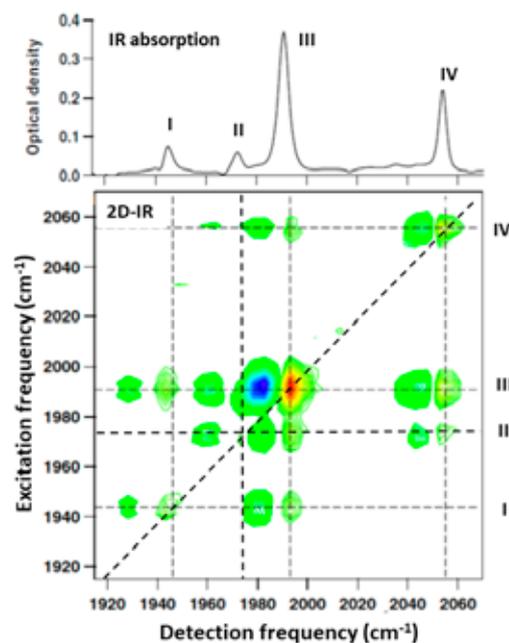


Figure 3. Upper plot shows an IR absorption spectrum of $(\mu\text{-C}(\text{CH}_3)(\text{CH}_2\text{S})_2(\text{CH}_2\text{S}(\text{CH}_2)_2\text{Ph})\text{Fe}_2(\text{CO})_5$, an organometallic model of the active site of the [FeFe]-hydrogenase enzyme, in the region of the IR spectrum associated with CO-stretching vibrational modes. Lower plot shows a 2D-IR spectrum of the same molecule. The four IR active modes in the absorption spectrum (I-IV) all give rise to a peak on the diagonal of the 2D-IR plot. In addition, off-diagonal peaks can be seen linking the diagonal peaks (see intersections of dashed guide lines). These peaks indicate the presence of vibrational coupling of the modes and energy transfer between them. Polarization dependence of the peaks obtained by changing the polarization of the input laser pulses can be used to obtain structural information, while the 2D shape reports on molecular dynamics.

Probing local dynamics:

This method involves implanting a molecule, or functional group within a protein structure at a specific location to act as a ‘vibrational label’. This ‘label’ will display an IR absorption that is well-separated from the absorptions of the rest of the protein so that its response can be isolated. Most often, this is achieved using a small molecule that acts as a natural ligand to the protein (e.g. CO or NO for a haem protein) or the use of isotopic labelling to change a specific $^{12}\text{C}=^{16}\text{O}$ peptide link to a $^{13}\text{C}=^{18}\text{O}$ unit. In the latter case, the heavier mass means that the labelled peptide link will exhibit a lower vibrational frequency than the rest of the peptide groups in the protein backbone. Non-natural amino acid replacement, using e.g. azido homoalanine or para cyanophenylalanine, has also been used to implant labels (N_3 or CN units). In all cases, the vibrational label groups provide dynamic information via their unique IR absorption bands at specific locations within the protein structure, while their small size generally reduces problems arising from fundamental changes in protein structure.

In 2D-IR experiments using vibrational labels, the 2D shape of the signal from the label group reports on protein dynamics. The important factor is that the dynamic movement of the protein around the label will cause its vibrational frequency to change continuously at the same rate as the protein structure fluctuates. In a 2D-IR experiment, if the time between the excitation and detection pulses is short compared to the dynamic motion of the protein then the excitation and detection events both ‘see’ the same protein structure and the 2D-IR lineshape is diagonally-elongated 2D (Figure 4, left). If the time between the excitation and detection event is made progressively longer, it will eventually reach the point where it is longer than the rate of the molecular dynamics of the protein. In this case, by the time the detection pulse arrives, the protein structure has changed and the frequency of the probe group has changed with it.

This time, the label has lost the ‘memory’ of what the protein structure was when it was excited and, rather than a diagonal lineshape, this creates a round 2D-IR lineshape (Figure 4, right). Plotting the evolution of the 2D-IR lineshape from elongated to circular, produces an exponential function, the timescale of which reflects the timescales of the molecular motion of the protein surrounding the probe group.

In a collaborative study involving molecular and structural biologists, we have recently used this method to study the protein dynamics in the active site of the catalase enzyme from *Corynebacterium glutamicum*. Using a NO ligand attached to the haem in the active site, it was shown that the ultrafast dynamics were altered by changing the solvent from H_2O to D_2O . The heavier mass of the deuterated solvent slowing down the dynamics near the haem. This was a surprising result, because catalase was widely held to have a dry active site, with no solvent water present. However, X-ray crystallography showed that there is sufficient space for a small pool of dynamic water within the haem cavity, which can interact with the haem ligand and which may form part of the reaction mechanism of catalase. Intriguingly, the ultrafast IR signals were observed to oscillate at short delay times, suggesting that a low frequency vibrational mode of the local water molecules may influence the haem ligand dynamics, possibly indicating a ‘reactive vibration’ of the enzyme structure.

Whole molecule studies:

In another collaborative study involving molecular biologists and theoretical chemists, we used 2D-IR to study the binding of the anti-tuberculosis drug isoniazid to its target enzyme, InhA, an enoyl acyl carrier protein reductase enzyme involved in the cell-wall maintenance of *Mycobacterium tuberculosis*. Studying the InhA

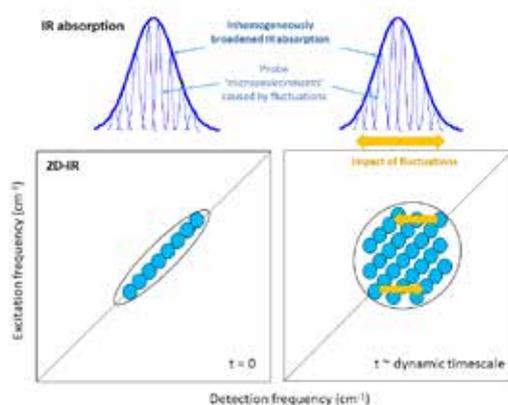


Figure 4. Extracting molecular dynamics from 2D-IR spectra. Upper row represents a schematic diagram of an IR absorption line shape of a vibrational label in a protein. A range of frequencies are caused by fluctuations of the protein structure. The schematic 2D-IR plots below show how the 2D-IR line shape changes from diagonally elongated to circular as the time delay between the laser pulses of the 2D-IR experiment progresses from shorter than the molecular fluctuations ($t=0$) to longer than the fluctuations ($t \sim$ dynamic timescale). Plotting the evolution of the peak shape with time quantifies the dynamic processes.

molecule as a whole, with no artificial labels implanted and combining our 2D-IR results with molecular modelling and spectral simulations, we showed that isoniazid binding to the wild-type enzyme induced changes in vibrational coupling between protein backbone residues in the binding site, which were visible in the off-diagonal part of the 2D-IR amide I lineshape. Importantly, the magnitude of the spectral changes correlated with the level of inhibition of InhA and the fast dynamics of the wild-type enzyme responded differently to those of the S94A mutation of InhA that is found in drug-resistant strains of *M. tuberculosis*. Of further interest, simulations predicted that the dynamic effects extend beyond the immediate drug binding site to the protein surface leading to the hypothesis that the changes in local dynamics caused by drug binding to wild-type InhA allosterically influence protein-protein binding in the InhA homotetramer.

Bridging the dynamic gap?

We have discussed how 2D-IR can be used to obtain information relating to the picosecond vibrations (fluctuations) of the protein structure. How do we extend this to covering the gap between picoseconds, nanoseconds and the micro-millisecond domain of other dynamic measurements? One way is to incorporate an additional laser pulse into the 2D-IR experiment. Arriving before the 2D-IR pulse sequence, this new laser pulse excites vibrations of the water itself, rather than the protein, rapidly heating up the sample by up to 10–15°C. 2D-IR can then be used to observe the relaxation of the protein and any subsequent structural dynamics. This cooling process typically requires millisecond timescales meaning that using 2D-IR in this way allows fast dynamics to be linked to longer timescale processes. Such experiments have been demonstrated by the Tokmakoff group in the USA for studying DNA melting and protein folding dynamics and offer exciting new insights into the way in which biomolecule dynamics fit together into the bigger picture. ■

Further reading

- Hamm, P.; Zanni, M. T., *Concepts and Method of 2D Infrared Spectroscopy*. Cambridge University Press: Cambridge, 2011.
- Hunt, N. T., *Ultrafast 2D-IR Spectroscopy – Applications to Biomolecules*. *Chem. Soc. Rev.* 2009, **38**, 1837-1848.
- Adamczyk, K.; Candelaresi, M.; Robb, K.; Gumiero, A.; Walsh, M. A.; Parker, A. W.; Hoskisson, P. A.; Tucker, N. P.; Hunt, N. T., *Measuring Protein Dynamics with Ultrafast Two-Dimensional Infrared Spectroscopy*. *Meas. Sci. Tech.* 2012, **23**, 062001.
- Bonner, G. M.; Ridley, A. R.; Ibrahim, S. K.; Pickett, C. J.; Hunt, N. T., *Probing the effect of the solution environment on the vibrational dynamics of an enzyme model system with ultrafast 2D-IR spectroscopy*. *Faraday Discuss.* 2010, **145**, 429-442.
- Adamczyk, K.; Simpson, N.; Greetham, G. M.; Gumiero, A.; Walsh, M. A.; Towrie, M.; Parker, A. W.; Hunt, N. T., *Ultrafast Infrared Spectroscopy Reveals Water-mediated Coherent Dynamics in an Enzyme Active Site*. *Chem. Sci.* 2015, **6**, 505-516.
- Shaw, D. J.; Hill, R. E.; Simpson, N.; Hussein, F. S.; Robb, K.; Greetham, G. M.; Towrie, M.; Parker, A. W.; Robinson, D.; Hirst, J. D.; Hoskisson, P. A.; Hunt, N. T., *Examining the role of protein structural dynamics in drug resistance in Mycobacterium tuberculosis*. *Chem. Sci.* 2017, **8**, 8384-8399.
- Sanstead, P. J.; Stevenson, P.; Tokmakoff, A., *Sequence-Dependent Mechanism of DNA Oligonucleotide Dehybridization Resolved through Infrared Spectroscopy*. *J. Am. Chem. Soc.* 2016, **138** (36), 11792-11801.
- Chung, H. S.; Ganim, Z.; Jones, K. C.; Tokmakoff, A., *Transient 2D IR spectroscopy of ubiquitin unfolding dynamics*. *Proc. Nat. Acad. Sci.* 2007, **104** (36), 14237-14242.



Neil T. Hunt, Department of Chemistry and York Biomedical Research Institute, University of York. Neil obtained his PhD from the University of Cambridge (2000). Following Post-doctoral appointments at Rice University (USA) and the Universities of East Anglia and Strathclyde, he became an EPSRC Advanced Research Fellow (2006) and was appointed to a Professorship in Physics at Strathclyde in 2016. Neil moved to the University of York as Professor of Physical Chemistry in 2018. Email: neil.hunt@york.ac.uk