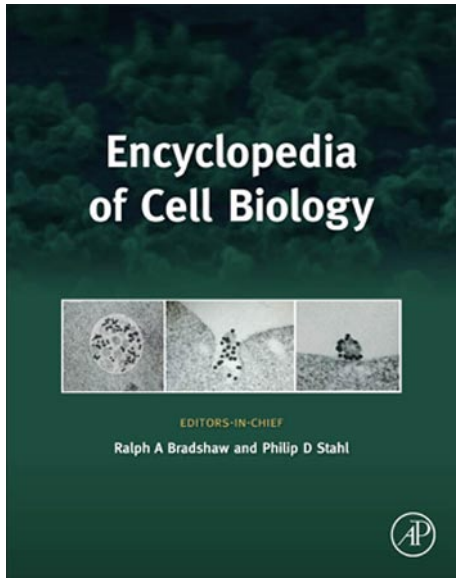

Marius Clore G., NMR in Structural and Cell Biology. In: Ralph A Bradshaw and Philip D Stahl (Editors-in-Chief), Encyclopedia of Cell Biology, Vol 1, Waltham, MA: Academic Press, 2016, pp. 98-107.

NMR in Structural and Cell Biology

G Marius Clore, National Institutes of Health, Bethesda, MD, USA

Published by Elsevier Inc.

Introduction

Nuclear magnetic resonance (NMR) is a powerful spectroscopic technique that permits the detailed study at atomic resolution of the three-dimensional structures and dynamics of macromolecules and their complexes in solution (Wüthrich, 1986; Clore and Gronenborn, 1989; Clore and Gronenborn, 1998a; Cavanagh *et al.*, 2007). The major source of structural information resides principally in a large number of short (<6 Å), approximate interproton distance restraints derived from nuclear Overhauser enhancement (NOE) measurements. These can be supplemented by torsion angle restraints derived from three-bond scalar couplings and backbone chemical shifts, orientational restraints in the form of residual dipolar couplings (RDCs) measured in weak alignment media, and long-range distance restraints (up to 35 Å) from paramagnetic relaxation enhancement (PRE) or paramagnetic pseudocontact shift (PCS) measurements. Dynamical information can be derived over a wide range of time scales ranging from picoseconds to seconds (Cavanagh *et al.*, 2007). Techniques include relaxation spectroscopy to measure dynamics of bond vectors in the picosecond to low nanoscond regime (faster than the rotational correlation time), relaxation dispersion spectroscopy in the microsecond to millisecond regime, RDCs potentially in the picosecond to millisecond regime, *z*-exchange spectroscopy for the millisecond to sub-second regime, and real time spectroscopy for the second regime upward. Recent interest in dynamics has largely focused on the application of relaxation dispersion spectroscopy to detect and characterize the kinetic properties of sparsely populated species (Korzhnev and Kay, 2008), PRE to detect and visualize such species (Clore *et al.*, 2007), and lifetime line broadening and dark-state exchange saturation transfer (DEST) spectroscopy to probe kinetics and dynamics of interactions of NMR visible molecules with large megadalton assemblies that are invisible to NMR (Fawzi *et al.*, 2011).

Brief Historical Perspective of NMR and Solution Structure Determination of Macromolecules

The development of two-dimensional ^{1}H-NMR spectroscopy (Ernst *et al.*, 1987) lead to the first three-dimensional solution structure determinations of small proteins in the mid-1980s (Williamson *et al.*, 1985; Clore *et al.*, 1986). Subsequent work incorporating larger numbers of interproton distance and torsion angle restraints together with stereospecific assignments lead to significant increase in both precision and accuracy (Clore and Gronenborn, 1998a,b; Clore *et al.*, 1986). Chemical shift overlap, however, limited the successful application of 2D ^{1}H-NMR techniques to proteins less than about 100 residues ($\sim$10 kDa). The late 1980s and early 1990s witnessed the development of 3D and 4D heteronuclear NMR spectroscopy which extended the range of applicability of the NMR method to significantly larger systems (Clore and Gronenborn, 1991). The first structure determination of a protein larger than 150 residues using 3D and 4D NMR was interleukin-1β ($\sim$18 kDa), which at the time was $\sim$50% larger than any previous NMR protein structure determination (Clore *et al.*, 1991). Over the next few years, these methods were extended to a variety of protein–peptide and protein–DNA complexes (Clore and Gronenborn, 1998a). Hybrid approaches combining existing high-resolution structures determined either by crystallography or NMR with sparse experimental NMR data, solution X-ray scattering, and computational techniques involving the application of conjoined rigid body/torsion angle-simulated annealing saw the structure determination of a number of significantly larger protein–protein complexes and proteins (up to $\sim$150 kDa) (Clore and Gronenborn, 1998a; Schwieters *et al.*, 2010). At the same time, deuteration combined with transverse relaxation optimized spectroscopy further increased the molecular weight range by significantly reducing linewidths (Peruvshin *et al.*, 1997; Kay, 2005). More recently, novel techniques, based on the application of PRE, have been developed to visualize sparsely populated, highly transient species that are undetectable by conventional biophysical and structural techniques, including crystallography, conventional NMR, cryoelectron microscopy, and single molecule spectroscopies (Clore *et al.*, 2007; Iwahara and Clore, 2006; Tang *et al.*, 2006, 2007; Clore and Iwahara, 2009).

Fundamentals of NMR Structure Determination

Macromolecular structure determination by NMR is intrinsically a highly specialized, labor-intensive and time-consuming technique. In addition, for a system of any reasonable size (say greater than about 70 residues) isotopic labeling with ^{15}N and ^{13}C is required. For even larger systems, additional labeling schemes are also necessary, including site specific isotope-labeling, deuteration and methyl protonation on a ^{13}C and ^{2}H background. Numerous reviews have been written on the subject detailing the experimental and computational methodologies involved (Wüthrich, 1986; Clore and Gronenborn, 1989, 1991, 1998a,b; Cavanagh *et al.*, 2007).

Determining the structure of a single protein by NMR can be broken down into essentially four steps: (1) sequential resonance assignment making use of a number of experiments to identify through-bond connectivities along the backbone and side chains (usually 3D triple resonance experiments); (2) assignment of cross-peaks in nuclear Overhauser enhancement spectra (usually 3D and 4D) to obtain short ($\leq$6 Å) interproton distance restraints which provide the main source of geometrical information; (3) measurement of additional NMR observables that provide useful conformational information (these may include three-bond scalar couplings that are related to torsion angles by simple empirical equations; backbone

 doi:10.1016/B978-0-12-394447-4.10019-7

chemical shifts which are related empirically to backbone ϕ/ψ torsion angles; long-range orientational restraints, such as RDCs measured in dilute liquid crystalline media; and (4) calculation of the three-dimensional structure from the experimental NMR restraints using simulated annealing. Generally an iterative refinement strategy is employed (Clore and Gronenborn, 1989, 1991, 1998b): calculations are initially carried out with a limited set of interproton distance restraints corresponding to NOE cross-peaks with unambiguous assignments; further interproton distance restraints from the remaining NOE cross-peaks are subsequently added in an iterative manner on the basis of a successively calculated series of structures. While improvements in spectrometer technology (e.g., the advent of cryoprobe technology that increases the signal-to-noise ratio three- to fourfold; higher field magnets that increase spectral resolution, thereby reducing spectral overlap) has reduced the measurement time to some extent, collecting all the data necessary to solve an NMR structure at high accuracy may still requires several months. Similarly, improvements in both spectral analysis software (Herrmann *et al.*, 2002a; Gerstein *et al.*, 2003; Yee *et al.*, 2003) and structure calculation algorithms (Schwieters and Clore, 2001; Linge *et al.*, 2003; Herrmann *et al.*, 2002b; Kuszewski *et al.*, 2004, 2008) has permitted the introduction of some degree of automation, but extensive human intervention is still necessary to fully and reliably interpret the data in all but the simplest of cases.

Place of NMR Spectroscopy in Structural and Cell Biology

In this light, what contribution can NMR make to structural and cell biology? There are two major methods for deriving high-resolution structural information at atomic resolution: NMR spectroscopy in solution and single crystal X-ray diffraction. In rare instances, cryoelectron microscopy is also capable of providing high-resolution information in the solid state. In addition, mass spectrometry in combination with cross-linking data is potentially capable of providing low-resolution structural information when combined with the computational techniques conventionally employed to derive structures from NMR data. If crystals can be rapidly obtained, there is little doubt that crystallography, particularly with the advent of synchrotron X-ray sources, offers the fastest route to high-resolution structure determination. However, complexes are generally more difficult to crystallize than isolated proteins, and it is usually the case that weak complexes (with K_D's in the 1–100 μM range) are extremely difficult to co-crystallize, while very weak complexes (K_D's > 1 mM) are virtually impossible to crystallize. In the case of NMR, complexes are amenable to structural investigation providing exchange is either fast (weak binding) or slow (tight binding) on the chemical shift time scale. If exchange, however, is intermediate on the chemical shift time scale, the signals are broadened out precluding any detailed structural work.

A full structure determination of a protein–protein complex by NMR is extremely time consuming. For example, in the case of the 40 kDa EIN · HPr complex from the bacterial phosphotransferase system, the total NMR measurement time alone was ~3500 h (or 4.8 months) (Garrett *et al.*, 1999). Clearly, therefore, the conventional approach is not suitable for high throughput. Fortunately, new developments have significantly shortened the amount of time required by making full use of prior knowledge in the form of existing high-resolution crystal or NMR structures of the free proteins (Schwieters and Clore, 2001; Clore, 2000; Clore and Bewley, 2002). Measurement of RDCs can quantitatively confirm that the structures of the components within the complex are either unchanged from that in the free-state or exhibit specific regions with identifiable structural changes. Similarly, chemical shift perturbations can also be used qualitatively in this regard since small chemical shift perturbations do not entail any significant structural changes. With this information in hand it is then possible to derive high-resolution structures of complexes using limited intermolecular NOE data to provide translation (as well as orientational) information and, if measurable, RDC data (Bax *et al.*, 2001; Prestegard *et al.*, 2000) to generate very accurate orientational information. In addition, strategies based on orientational information from RDCs have been developed whereby translational information from NOE data can be entirely replaced in suitable cases by highly ambiguous intermolecular distance restraints derived from $^{15}N/^{1}H_N$ chemical shift perturbation mapping (Clore and Schwieters, 2003). Long-range distance restraints from a paramagnetic label attached to an engineered surface cysteine on one partner to nuclei of the other partner can also be very helpful and can replace the measurement of intermolecular NOE data completely if multiple paramagnetic labels are used (Pintacuda *et al.*, 2007).

Experimental and Computational Considerations

Given that structure determination of proteins has been extensively reviewed, we will briefly review the main experimental restraints and computational techniques used in the structure determination of complexes.

Intermolecular Distance Restraints

As noted above, the NOE is the primary source of geometric information for NMR-based structure determination (Wüthrich, 1986; Clore and Gronenborn, 1989, 1998a,b). The NOE (in the initial rate approximation) is proportional to the sixth root of the distance between two protons. The upper limit for interproton distances that can be detected using the NOE is 5–6 Å. The key to deriving intermolecular NOE-derived interproton distance restraints lies in combining various isotope (^{15}N and ^{13}C)-labeling strategies with isotope-filtering experiments that permit one to detect NOEs on protons attached to specific isotopes of nitrogen and carbon (i.e., NMR active such as ^{15}N or ^{13}C, or NMR inactive such as ^{14}N and ^{12}C) (Clore and Gronenborn, 1998a). For example, in a complex comprising one protein labeled uniformly with ^{13}C and the other at natural isotopic abundance (i.e., ^{12}C), one can selectively detect NOEs from protons attached to ^{13}C to protons attached to ^{12}C.

Paramagnetic-Based Distance Restraints

It is possible to derive intermolecular distance restraints using another NMR-based approach which involves derivatizing (one at a time) suitable surface accessible cysteines (which may have to be introduced by site-directed mutagenesis) on one protein with either a nitroxide spin label or a metal-binding site (such as EDTA) and measuring the resulting PRE or PCS on the other protein to yield long-range (15–35 Å) distance restraints (Clore *et al.*, 2007; Pintacuda *et al.*, 2007). Because in most cases the paramagnetic label is attached to the protein by a linker involving several rotatable bonds, it is essential to consider the conformational space sampled by the paramagnetic label in order to obtain accurate results (Clore and Iwahara, 2009). Two types of measurement can be made with paramagnetic labels, depending on the nature of the label. For paramagnetic labels with an isotropic *g*-tensor (e.g. Mn^{2+}, Gd^{3+}, nitroxide radical), pseudocontact shifts (PCSs) are not observed and PRE measurements can be carried out to determine the PRE rates for each paramagnetic-^{1}H interaction from the difference in relaxation rates (usually transverse) between the paramagnetic and diamagnetic samples. When the *g*-tensor is anisotropic (e.g., many lanthanide ions), PCSs are observed.

Both PRE and PCS effects arise from the dipole-dipole interaction between the unpaired electron of the paramagnetic spin and a nuclear spin. Because of the large magnetic moment of the electron these effects are very large and measurable over much larger distances (up to 35 Å in suitable cases) than the dipolar interaction between two protons that gives rise to the NOE. The magnitude of the PRE is related to the sixth root of the distance between the paramagnetic center and a proton and can be directly used in refinement. The PCS has the same functional form as RDCs and is dependent on both the cube root of the distance between the paramagnetic center and the nuclei of interest and the orientation that the paramagnetic center-nucleus vector makes relative to the χ paramagnetic tensor.

In general, paramagnetic effects arising from an extrinsic paramagnetic center (as opposed to one that is intrinsic as is the case for metalloproteins) can only be applied in a rational manner if one already has a good idea of the interaction surfaces involved in complex formation. Such information can be derived rather easily by either $^{15}N/^{1}H_N$ chemical shift perturbation mapping (Walters *et al.*, 2001) or cross-saturation experiments (Takahashi *et al.*, 2000). The latter experiment is far more challenging experimentally since it necessitates that one of the protein is not only ^{15}N-labeled but fully deuterated as well.

Paramagnetic effects, however, do have to be used with some caution. In the case of very tightly binding complexes where exchange is slow on the paramagnetic relaxation time scale, the paramagnetic effects will arise solely from the specific complex (Clore *et al.*, 2007; Clore and Iwahara, 2009). However, when exchange is fast, the footprints from paramagnetic effects arising from minor species and configurations are apparent. This is more marked for PRE measurements than PCS ones owing to the respective sixth versus cube root distance dependencies. It is precisely these effects that have permitted states that are undetectable by conventional structural and biophysical techniques to be studied.

Other Sources of Distance Information

NMR is not the only method that can be used to derive intermolecular distance restraints. It is also possible to derive distance restraints using a combination of cross-linking, proteolytic digestion, and mass spectrometry (Bennett *et al.*, 2000; Sinz and Wang, 2001; Schulz *et al.*, 2004). In many cases, however, the data will not yield unique cross-linking partners but multiple possibilities.

Fluorescence energy transfer (FRET) through non-radiative dipolar–dipolar coupling from the fluorophore, the energy donor, to a second chromophore, the energy acceptor, scales as the sixth root of the distance between the two chromophores and can probe separations ranging from 10 to 100 Å (Hillisch *et al.*, 2001). Likewise double nitroxide spin-labeling coupled with pulsed electron paramagnetic resonance (EPR) methods such as double electron–electron resonance (DEER) based on the magnitude of the magnetic dipolar coupling of the unpaired nitroxide electrons which scales as the cube root of the separation between the two nitroxide labels, can yield remarkably accurate distances in the 20–60 Å range (Altenbach *et al.*, 2008). FRET and EPR are not limited by the molecular weight of the system being studied, but suffer from a major drawback in so far that only a single pairwise distance can be measured per sample (i.e., each distance requires a new double spin-labeled or double chromophore labeled sample, with the labels in different positions). Thus, although FRET and EPR can yield very specific information they do not afford a practical approach for solving three-dimensional structures of proteins or their complexes.

Orientational Restraints

Long-range orientational restraints can be derived from the measurement of RDCs (Bax *et al.*, 2001; Prestegard *et al.*, 2000) and chemical shift anisotropy (Wu *et al.*, 2001; Tjandra *et al.*, 1997) in liquid crystalline media, and in suitable cases from heteronuclear T_1/T_2 data (Tjandra *et al.*, 1997). The characteristic feature of these various parameters is that they yield direct geometric information on the orientation of an interatomic vector(s) with respect to an external axis system (e.g., the alignment tensor in liquid crystalline media, the diffusion tensor for relaxation measurements) expressed in terms of two angles: θ, the angle between the interatomic vector and the *z* axis of the tensor, and ϕ, the angle which describes the position of the projection of the interatomic vector on the *x*-*y* plane of the tensor.

For most practical purposes, RDCs provide the easiest method for deriving orientational information. In an isotropic medium, the dipolar couplings average to zero. In the solid state, the maximum value of the ^{15}N-^{1}H dipolar coupling is 20.7 kHz. To effectively measure dipolar couplings in solution, therefore, it is necessary to devise means of inducing only a small ($\sim 10^{-3}$) degree of order such that the ^{15}N-^{1}H dipolar couplings lie in the ± 20 Hz range. Experimentally, this is achieved by dissolving the protein or protein complex of interest in a dilute, water soluble, liquid crystalline medium. Examples of such media include lipid bicelles, filamentous phages such as fd or pf1, rod-shaped viruses such as tobacco mosaic virus, polyethylene glycol/hexanol, and stretched polyacrylamide gels.

Some Computational Methods

In many instances, protein complex formation involves no significant changes in backbone conformation. Thus, if the structures of the individual proteins are already known at high resolution and it can be shown that the backbone conformation remains essentially unchanged upon complex formation (e.g., by comparison of dipolar coupling data measured on the complex with the X-ray structures of the free proteins), one can then make use of conjoined rigid body/torsion angle dynamics to rapidly solve the structure of the complex on the basis of intermolecular NOE data and backbone NH dipolar couplings (Schwieters and Clore, 2001; Clore, 2000). In this procedure, only the interfacial side chains are allowed to alter their conformation. The backbone and non-interfacial side chains of one protein are held fixed, while those of the second protein are only allowed to rotate and translate as a rigid body. This has been applied with considerable success in the case of the 30–70 kDa protein–protein complexes of the bacterial phosphotransferase system (Clore and Venditti, 2013), as well as to the 42 kDa ternary Oct1 · Sox2 · *Hoxb1*-DNA ternary transcription factor complex (Williams *et al.*, 2004).

It should be emphasized that conjoined rigid body/torsion angle dynamics can readily be extended to cases where significant changes in backbone conformation are localized to specific regions of the protein, such as the binding interface. In such cases, both the interfacial side chains and the relevant portions of the protein backbone would be given torsional degrees of freedom, and the experimental data would also have to include intramolecular NMR restraints (NOE, dipolar coupling, etc.) relating to that portion of the backbone. This, for example, is the strategy that was employed to solve the structure of the IIA^{Mtl} · HPr complex (Cornilescu *et al.*, 2002). This was necessitated because the crystal structure of IIA^{Mtl} (Van Montfort *et al.*, 1998) which contains multiple copies of IIA^{Mtl} in the unit cell, revealed alternate conformations for four loops in relatively close proximity to the putative interaction surface with HPr.

Providing the complex under study can be aligned in a suitable liquid crystalline medium, the measurement of dipolar couplings is straightforward and permits one to determine the relative orientation of two proteins in a complex. Dipolar couplings, however, do not yield any translational information which is essential for docking. Clearly, NOE-derived intermolecular interproton distance restraints provide the most useful and reliable source of translational information. However, intermolecular NOEs are not always easy to observe and their unambiguous assignment is still difficult and time consuming, particularly for larger complexes. Backbone $^{1}H_{N}$ and ^{15}N chemical shifts, on the other hand, are highly sensitive to environment and have been extensively used to rapidly map interaction surfaces on proteins (Walters *et al.*, 2001). Not surprisingly, examination of the NMR literature reveals hundreds of examples of chemical shift mapping studies; to date, however, only a handful of structures of macromolecular complexes have been determined by NMR. Recently, it has been shown that it is possible to convert chemical shift perturbation maps into highly ambiguous intermolecular distance restraints which, in combination with orientational restraints from dipolar couplings, can reliably and accurately dock the partner proteins in a complex by means of rigid body/torsion angle dynamics calculations (Clore and Schwieters, 2003). Clearly, this methodology provides a powerful tool for high throughput structural proteomics and, moreover, can greatly accelerate the determination of higher accuracy NMR structures of complexes (including the detailed placement of interfacial side chains) by providing a good starting point for the assignment of intermolecular NOE data.

Structural Proteomics of the Bacterial Phosphotransferase System

In bacteria, carbohydrate transport across the membrane is mediated by the phosphoenolpyruvate:sugar phosphotransferase system (PTS) which provides tight coupling of translocation and phosphorylation (Deutscher *et al.*, 2006). The PTS is a classical example of a signal transduction pathway involving phosphoryl transfer whereby a phosphoryl group originating on phosphoenolpyruvate is transferred to the translocated carbohydrate via a series of bimolecular protein–protein complexes. The first two steps of the PTS are common to all sugars: enzyme I (EI) is autophosphorylated by phosphoenolpyruvate and subsequently donates the phosphoryl group to the histidine phosphocarrier protein HPr. The proteins downstream from HPr comprise the sugar-specific enzymes II which fall into four distinct families: glucose (Glc), mannitol (Mtl), mannose (Man) and lactose/chitobiose (Chb). Although the four families bear no sequence or structural similarity, they do possess similar organizations consisting of two cytoplasmic domains A and B, and one or two membrane-bound domains, C and D, which may or may not be covalently linked to one another. The active site residue of the A domains is always a histidine which accepts the phosphoryl group from HPr on its Nε2 atom and donates a phosphoryl group to either a cysteine residue (in the case of IIB^{Glc}, IIB^{Mtl} and IIB^{Chb}) or to the Nδ1 atom of a histidine residue (in the case of IIB^{Man}). Subsequently the phosphoryl group is transferred onto the incoming sugar on the cytoplasmic side of the membrane-bound C domain (also known as the sugar permease).

The complexes in this pathway are rather weak with K_D's ranging from 1 μM to 3–6 mM. The K_D's in the millimolar range relate to complexes involving isolated domains that are connected by 20–30 residue long flexible linkers in the intact protein (Clore and Venditti, 2013). Although binding in such instances is very weak, it is in fact perfectly tuned to the system. In particularly, it can readily be calculated, based on the expected average end-to-end distance for the linkers, that these millimolar equilibrium dissociation constants correspond to 50–85% probabilities of the two linked domains interacting with one another at any given time. Although high-resolution crystal structures and NMR structures have been determined for many of the individual proteins of the PTS, crystallization of these protein–protein complexes has proven to be refractory, despite many years of trying. Thus, this system provides a showcase for the impact of NMR in structural proteomics.

Figure 1 shows ribbon diagrams of structures of all nine cytoplasmic PTS complexes solved in our laboratory

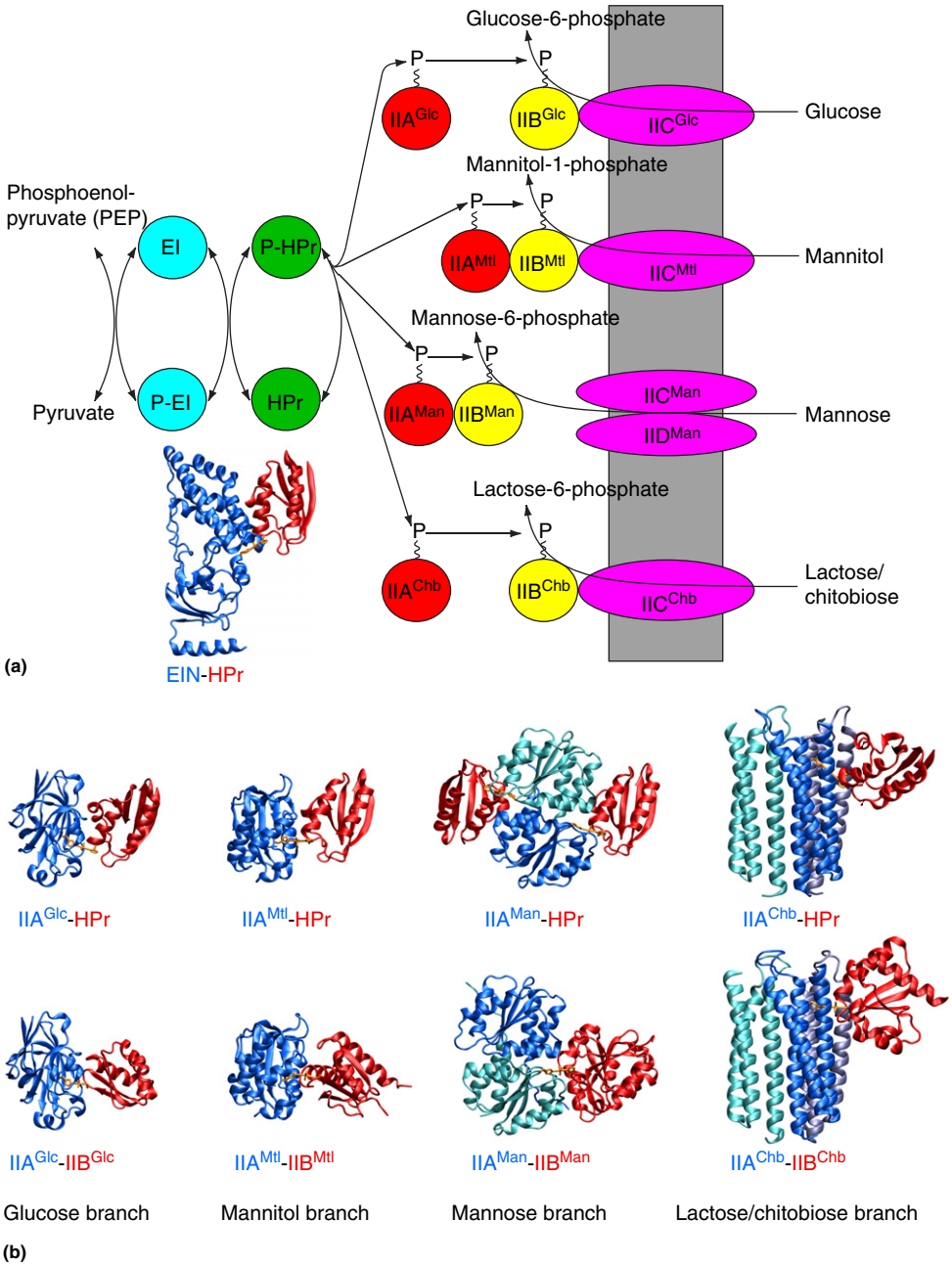

Figure 1 Structural biology of the bacterial phosphotransferase signal transduction pathway. (a) Diagrammatic representation of the pathway. The first two steps are common to all branches, and thereafter the pathway splits into four sugar-specific branches (glucose, mannitol, mannose, and lactose/chitobiose). (b) Ribbon diagrams of the nine protein–protein complexes of the *Escherichia coli* phosphotransferase system.

(Clore and Venditti, 2013). These complexes shed light on understanding fundamental aspects of protein–protein recognition, mechanisms for phosphoryl transfer between proteins, and the diversity of structural elements recognized by a single protein. Specificity of the protein–protein interaction surfaces is characterized by geometric and chemical complementarity, coupled with extensive redundancy to permit the effective recognition of multiple partners. There is little or no conformational change in the protein backbone before and after association. Some interfacial side chains, however, adopt different conformations (side chain conformational plasticity) depending on the interacting partner so as to achieve optimal intermolecular interactions. A consequence of these properties is increased velocity in signal transduction by eliminating any unnecessary time delay required for significant conformational change.

The interaction surfaces for HPr on enzyme I and the four sugar-specific enzymes IIA are very similar despite the fact that their underlying structures are completely different in terms of linear sequence, secondary structure and topological arrangement of structural elements. HPr makes use of essentially the same surface to interact with both its upstream and

downstream partners (cf. Figure 1). Concomitantly, the binding sites for the sugar-specific enzymes IIB and HPr on the corresponding sugar-specific enzymes IIA overlap extensively. One might therefore anticipate that the enzymes IIB could also interact with EIN. However, NMR data indicate that there is absolutely no interaction between EIN and any of the sugar-specific enzymes IIB at millimolar concentrations. From a functional perspective this is important since it ensures that the PTS cascade is not bypassed. In addition, prevention of any potential shortcuts circumventing HPr and the sugar-specific enzymes IIA is also necessary since these proteins in different states of phosphorylation also regulate the functions of proteins in other pathways (Deutscher *et al.*, 2006). The structural basis for specificity and discrimination lies in the different charge distributions on the interaction surfaces of HPr and the sugar-specific enzymes IIB such that binding of a sugar-specific enzyme IIB to EIN is precluded by electrostatic repulsion (Clore and Venditti, 2013).

Exploring Sparsely Populated States of Proteins and Their Complexes

Many biological processes proceed via sparsely populated species. Examples include the initial formation of encounter complexes in macromolecular association, target searching in specific protein–DNA recognition, conformational selection in ligand binding, conformational transitions associated with allostery, intermediates along the protein folding pathway or in the course of enzyme catalysis, and early events in self-assembly processes. In general the populations of these states at equilibrium are low and their lifetimes are short. Consequently, transient states arising from rare but rapid excursions between the global free energy minimum and higher free energy local minima are extremely challenging to study at atomic resolution under equilibrium conditions since they are effectively invisible to most structural and biophysical techniques including crystallography, conventional NMR spectroscopy, electron microscopy and single molecule spectroscopy. Recent developments in NMR have rendered short-lived, sparsely populated states accessible to spectroscopic analysis, yielding considerable insights into their kinetics, thermodynamics, and structures.

Principal NMR Methods to Probe Transient Sparsely Populated States

Three main NMR methods have been developed over the last few years to probe rare invisible states of macromolecules and their complexes at equilibrium: namely PRE, relaxation dispersion spectroscopy and lifetime line broadening coupled with DEST spectroscopy (see Anthis and Clore (2015) for a recent in-depth review).

The PRE requires that the distances between a paramagnetic label and the monitored spins (usually protons) are significantly shorter in the sparsely populated state than in the major species, and that the lifetime of the minor species is less than ~250–500 μs (Iwahara and Clore, 2006; Tang *et al.*, 2006, 2007; Clore and Iwahara, 2009). In this exchange regime, the footprint of sparsely populated states can be observed on PRE profiles measured on the resonances of the major species, thereby yielding structural information that is directly related to paramagnetic center-nuclei distances, from which it is possible, under suitable circumstances, to compute a structure or ensemble of structures for the minor species (Tang *et al.*, 2006, 2007). Relaxation dispersion spectroscopy is dependent on the existence of significant chemical shift differences between the NMR active nuclei (^{1}H, ^{15}N or ^{13}C) in the various states, and in general can be used to probe events occurring on time scales ranging from about 50 μs to 10 ms (Korzhnev and Kay, 2008; Palmer *et al.*, 2001; Loria *et al.*, 2008; Baldwin and Kay, 2009). Lastly, lifetime line broadening and DEST probe exchange dynamics at atomic resolution between NMR visible molecules and large (in excess of 1 MDa) NMR invisible 'dark' states on time scales ranging from 0.5 ms to 1 s (Fawzi *et al.*, 2011, 2014; Libich *et al.*, 2013). The DEST experiment relies entirely on large differences in transverse relaxation rates between the NMR visible and invisible species. These three complementary techniques are capable of detecting states with populations as low as 0.5%.

Basis of PRE for the Study of Sparsely Populated States

The PRE yields structural information directly but cannot be used to obtain kinetic information (i.e., rate constants). The underlying theory of the PRE for static systems dates back to the late 1950s (Solomon, 1955; Bloembergen and Morgan, 1961) and the PRE has long been used in the study of paramagnetic metalloproteins. The potential of the PRE for structure determination of single proteins, however, was first demonstrated in the mid-1980s (Kosen, 1989) but then largely neglected until about 10 years ago with the advent of straightforward biochemical methods for introducing paramagnetic labels at specific sites in proteins (Battiste and Wagner, 2000). Moreover, the quantitative use of the PRE for structure determination was thwarted until the introduction of the appropriate theoretical framework and computational methods to take into account the large conformational space sampled by a paramagnetic label attached to the protein via a linker with multiple rotatable bonds (Iwahara *et al.*, 2004). By representing the paramagnetic label by an ensemble of states and taking care to calculate PRE order parameters from the coordinates during the course of structure refinement, it is possible to directly refine against the PRE relaxation rates and obtain accurate structures where agreement between the model and the experimental data is quantitatively assessed by a *Q*-factor analogous to a crystallographic R-factor (Iwahara *et al.*, 2004).

The key insight into using the PRE to detect transient low-population species lies in rapid exchange phenomena whereby the transverse PRE observed on a major species is modulated by the presence of the minor species (Iwahara and Clore, 2006). In a two-site exchange system comprising two species A and B that interconvert on a time scale that is fast on the PRE time scale, the observed PRE measured on either resonance will be the population weighted average of the PRE rates for the two species. Thus, if one has, for example, a system where a particular paramagnetic center-proton distance is 30 Å for the major species and 8 Å for the minor species, the corresponding PRE rates (for a system ~30 kDa in size with Mn^{2+} as the

paramagnetic center) will have values of $\sim 2\,s^{-1}$ and $\sim 5600\,s^{-1}$, respectively. If the major and minor species are populated at 99% and 1%, respectively, the minor species will be invisible in the NMR spectrum. But in the fast exchange limit the observed PRE measured on the resonance of the major, NMR visible, species will be the population weighted average of the PREs for the major and minor species, in this instance $\sim 50\,s^{-1}$, much larger than that expected for the major species alone. Therefore, providing distances between the paramagnetic center and the protons of interest are significantly shorter in the minor species than the major one, and the interconversion rate between the two species is fast, the PRE profiles observed on the major species will reveal the footprint of the minor species. The PRE profiles can be analyzed quantitatively to derive structural information if the PRE profile for the major species is either known or can be calculated from a known structure (Tang *et al.*, 2006, 2007; Clore and Iwahara, 2009). As the exchange rate decreases, the influence of the minor species on the observed PRE profile for the major species will be reduced until in the slow exchange limit the PRE profile for the major species will be unaffected by the presence of the minor species. Thus, the use of the PRE to detect and characterize sparsely populated states is limited to rapidly exchanging systems, typically with lifetimes less than about 250–500 μs (Clore and Iwahara, 2009).

Basis of Relaxation Dispersion Spectroscopy

In an exchanging system between multiple states, the transverse relaxation rate (R_2) is given by the sum of the intrinsic transverse relaxation rate R_2^0 and an exchange contribution R_{ex}. The R_{ex} term is a function of the exchange rate k_{ex} (which in the case of a two-site exchange system is simply the sum of the forward and backward rate constants) and the chemical shift difference for the nucleus in the two distinct chemical environments. In the slow exchange regime, when k_{ex} is much (20-fold or more) smaller than the chemical shift difference measured in radians s^{-1}, two distinct resonances will be observed. In the fast exchange regime when k_{ex} is much (≥ 20) larger than the chemical shift difference, a single resonance will be observed at a position corresponding to the population weighted mean of the chemical shifts in the two states. In the extreme fast and slow exchange limits, the exchange contribution to the linewidth is negligible. In the intermediate regime, however, the R_{ex} term results in line broadening which is most marked when k_{ex} is equal to the chemical shift difference. Under these conditions, even the presence of a state populated at the 0.5% level can cause significant line broadening of the resonances of the major species. The key to relaxation dispersion spectroscopy lies in the use of special pulse sequences to progressively attenuate the R_{ex} contribution to the measured R_2 rate which can be achieved by applying a train of refocusing pulses while magnetization evolves under the influence of a chemical shift that varies stochastically as a result of the exchange process (Palmer *et al.*, 2001; Loria *et al.*, 2008; Baldwin and Kay, 2009). Because each nucleus follows a slightly different trajectory, dephasing of magnetization occurs resulting in larger R_{ex} rates and hence broader linewidths. By reducing the interval between the refocusing pulses (i.e. increasing the number of pulses during a fixed period T), dephasing is decreased, the R_{ex} term is reduced and the linewidths become narrower. Plots of the observed R_2 rate as a function of the interval between the refocusing pulses yields what is known as a relaxation dispersion curve with large observed R_2 rates at low repetition rates and small R_2 rates at high repetition rates. The detailed shape of the relaxation dispersion curve is a complex function of the exchange rate k_{ex}, the populations of the states, the chemical shift difference between the states and the repetition rate. For data recorded at a single magnetic field, the contribution of population and chemical shift difference to R_{ex} cannot be separated *a priori*. Thus, unless the population of the species or the chemical shift difference between these species is already known, deconvolution of these two terms necessitates recording relaxation dispersion data at different magnetic field strengths since the species populations are independent of magnetic field while the chemical shift difference in frequency units is linearly proportional to the magnetic field. Typically relaxation dispersion experiments using refocusing pulses can probe exchange processes with lifetimes ranging from ~ 50 μs to ~ 10 ms and occupancies for the minor species as low as 0.5–1%.

The key feature of NMR that distinguishes it from all other forms of spectroscopy is that interactions involving many sites of known identity can be probed simultaneously. The sites comprise NMR observable nuclei (^{1}H, ^{15}N, ^{13}C) whose resonance assignments in the major species are readily obtained using modern triple resonance NMR spectroscopy. Using global fitting procedures in which all the relaxation dispersion data at the observed sites are fitted simultaneously, it is possible to dissect kinetic pathways. The most commonly employed nucleus for relaxation dispersion experiments is the ^{15}N of the backbone amide groups. Yet, the structural information provided by ^{15}N chemical shifts alone is generally quite limited unless reference ^{15}N chemical shifts are already available for the various states being studied. More recent developments have extended relaxation dispersion measurements to all ^{1}H, ^{13}C and ^{15}N backbone atoms, as well as to side chain methyl groups (Sekhar and Kay, 2013). This opens the way to obtain highly reliable backbone ϕ/ψ torsion angle restraints for minor states derived from complete backbone chemical shifts, and even to compute 3D structures under suitable circumstances, thereby providing a potential avenue for obtaining full 3D structural information on sparsely populated states. More recent developments have shown that one can use relaxation dispersion experiments to obtain bond vector orientation information on minor states (Vallurupalli *et al.*, 2007). Two NMR observables are available, RDCs and residual chemical shift anisotropy (RCSA). Both involve the use of weakly aligned media (such as dilute solutions of bicelles and filamentous phages) to reintroduce anisotropic magnetic interactions that are otherwise averaged to zero in isotropic solution. Because these effects are small, highly accurate relaxation dispersion measurements are required and the experiments are extremely demanding.

Basis of Lifetime Line Broadening and DEST

Exchange dynamics between molecules free in solution and bound to the surface of a large supramolecular structure, a polymer, a membrane or solid support are important in many

phenomena in biology and material science. These interactions can be probed by lifetime line broadening and DEST spectroscopy to probe exchange dynamics at atomic resolution between NMR visible molecules and large (in excess of 1 MDa) NMR invisible 'dark' states on time scales ranging from 0.5 ms to 1 s (Fawzi *et al.*, 2010, 2011, 2014; Libich *et al.*, 2013).

Exchange line broadening can arise from either differences in chemical shifts (chemical exchange line broadening) or transverse relaxation rates (lifetime line broadening) between the free and bound states. When a molecule binds to a high (>700 kDa) molecular weight entity, the reduced rate of molecular tumbling leads to a marked increase in transverse (R_2) relaxation rates (i.e., severe line broadening) which precludes direct observation of the bound state by standard NMR techniques. If the dissociation rate constant k_{off} is considerably smaller (by two orders of magnitude or more) than the R_2 in the bound state, the difference in R_2 values (ΔR_2) for the NMR visible species in the presence and absence of the large molecular weight species will be equal to the pseudo-first order association rate constant (Fawzi *et al.*, 2010, 2011). If k_{off} is comparable to or larger than R_2 in the bound state, ΔR_2 will be dependent upon both the association and dissociation rate constants as well as the R_2 in the bound state (Libich *et al.*, 2013).

The essence of the DEST experiment is that the large R_2 values in the bound state that preclude direct observation by NMR allow for efficient partial saturation of longitudinal magnetization of bound state resonances by a weak radio-frequency field, even at offsets where the magnetization of the free species is completely unaffected (Fawzi *et al.*, 2011). In other words, even though the bound resonances are completely broadened out beyond detection (i.e. they effectively lie in the baseline) they can be perturbed and partially saturated. Saturation of the bound resonances is then transferred back to the corresponding resonances of the free species by chemical exchange and subsequently measured as attenuation of the easily observed resonances of the NMR visible species. Operationally, the DEST experiment involves the creation of an action profile by applying weak saturation at set intervals from say +35 kHz to −35 kHz, and measuring the cross-peak intensities in a 2D-correlation spectrum as a function of the frequency offset of the saturation pulse. The resulting profiles are dependent upon the association and dissociation rate constants, and the transverse relaxation rates in the bound state.

Relevance to Cell Biology

The PRE has been used to study how the opposing constraints of speed and specificity are optimized in biological interactions. Examples include the first direct demonstration of intra- and intermolecular translocation of transcription factors to enhance specific site searching (Iwahara and Clore, 2006); the first experimental visualization of encounter complexes in protein–protein association (Tang *et al.*, 2006); the elucidation of conformational selection of a very low-population of correctly configured dimer in auto-processing of the HIV-1 protease precursor monomer (Tang *et al.*, 2008), a phenomenon of fundamental practical importance in the design of novel HIV protease inhibitors; and the dissection of the complementary interplay between conformational selection and induced fit, exemplified by the characterization of a transient open-to-closed transitions in apo maltose binding protein (Tang *et al.*, 2007) and calmodulin that facilitate ligand-induced formation of the holo state (Anthis *et al.*, 2011).

Relaxation dispersion has shed fundamental insights into a range of biological problems of considerable significance (Korzhnev and Kay, 2008; Palmer *et al.*, 2001; Loria *et al.*, 2008; Baldwin and Kay, 2009; Sekhar and Kay, 2013). In the case of protein folding, this includes the structure determination of folding intermediates, the elucidation of the kinetics of their interconversion and the delineation of on- and off-pathway events. Relaxation dispersion has also been used to probe allosteric mechanisms associated with ligand binding in very large assemblies including chaperones, aspartate transcarbamoylase and the proteasome.

Lastly, lifetime line broadening and DEST have been used to probe exchange processes between monomer and protofibril-bound states of amyloid β (Fawzi *et al.*, 2010, 2011, 2014) and between intrinsically disordered proteins and the chaperonin GroEL (Libich *et al.*, 2013) on time scales of 0.5–100 ms, imprinting the residue-by-residue footprint of the NMR invisible protofibril-bound and GroEl-bound states on the easily observed monomer. The experiments on amyloid β shed light on protofibril formation, the structure and dynamics of protofibrils and exchange processes occurring on the surface and ends of protofibrils, which are of interest since the accumulation of toxic, aggregated forms of amyloid β are implicated in the etiology of Alzheimer's disease. The experiments on the interaction of intrinsically disordered proteins with GroEL unveil the complex molecular recognition process whereby GroEL recognizes a large array of sequences and structures. The demonstrated ability of the DEST technique to examine dynamics at single residue resolution of otherwise NMR invisible 'dark' states has the potential to revolutionize many areas of current interest in both biology and materials science.

Acknowledgments

This work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. This review is adapted from Clore, G.M., 2014. NMR in structural biology. In: Howard, G.C., Brown, W.E., Auer, M. (Eds.), Imaging Life: Biological Systems from Atoms to Tissues. Oxford: Oxford University Press, pp. 51–73.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis; Protein Domains: Structure, Function, and Methods

References

Altenbach, C., Kusnetzow, A.K., Ernst, O.P., Hoffmann, K.P., Hubbell, W.L., 2008. High-resolution distance mapping in rhodopsin reveals the pattern of helix

movement due to activation. Proceedings of the National Academy of Sciences of the United States of America 105, 7439–7444.
Anthis, N.J., Clore, G.M., 2015. Visualizing transient dark states by NMR spectroscopy. Quarterly Reviews of Biophysics 48 (1), 35–116. doi:10.1017/S0033583514000122.
Anthis, N.J., Doucleff, M., Clore, G.M., 2011. Transient sparsely-populated compact states of apo and calcium-loaded calmodulin probed by paramagnetic relaxation enhancement: interplay of conformational selection and induced fit. Journal of the American Chemical Society 133, 18974–18976.
Baldwin, A.J., Kay, L.E., 2009. NMR spectroscopy brings invisible protein states into focus. Nature Chemical Biology 5, 808–814.
Battiste, L., Wagner, G., 2000. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. Biochemistry 39, 5355–5365.
Bax, A., Kontaxis, G., Tjandra, N., 2001. Dipolar couplings in macromolecular structure determination. Methods in Enzymology 339, 127–174.
Bennett, K.L., Kussmann, M., Bjork, P., *et al.*, 2000. Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping: a novel approach to assess intermolecular protein contacts. Protein Science 9, 1503–1518.
Bloembergen, N., Morgan, L.O., 1961. Proton relaxation times in paramagnetic solutions: effects of electron spin relaxation. Journal of Chemical Physics 34, 842–850.
Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton, N.J., 2007. Protein NMR Spectroscopy: Principles and Practice, second ed. Amsterdam: Elsevier Academic Press.
Clore, G.M., 2000. Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization. Proceedings of the National Academy of Sciences of the United States of America 97, 9021–9025.
Clore, G.M., Bewley, C.A., 2002. Using conjoined rigid body/torsion angle simulated annealing to determine the relative orientation of covalently linked protein domains from dipolar couplings. Journal of Magnetic Resonance 143, 329–335.
Clore, G.M., Gronenborn, A.M., 1989. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. CRC Critical Reviews in Biochemistry and Molecular Biology 24, 479–564.
Clore, G.M., Gronenborn, A.M., 1991. Structures of larger proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. Science 252, 1390–1399.
Clore, G.M., Gronenborn, A.M., 1998a. Determining structures of larger proteins and protein complexes by NMR. Trends in Biotechnology 16, 22–34.
Clore, G.M., Gronenborn, A.M., 1998b. New methods of structure refinement for macromolecular structure determination by NMR. Proceedings of the National Academy of Sciences of the United States of America 95, 5891–5898.
Clore, G.M., Iwahara, J., 2009. Theory, practice and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. Chemical Reviews 109, 4108–4139.
Clore, G.M., Nilges, M., Sukuraman, D.K., *et al.*, 1986. The three-dimensional structure of α1-purothionin in solution: Combined use of nuclear magnetic resonance, distance geometry and restrained molecular dynamics. EMBO Journal 5, 2729–2735.
Clore, G.M., Schwieters, C.D., 2003. Docking of protein-protein complexes on the basis of highly ambiguous distance restraints derived from $^{1}H_{N}/^{15}N$ chemical shift mapping and backbone ^{15}N-^{1}H residual dipolar couplings using conjoined rigid body/torsion angle dynamics. Journal of the American Chemical Society 125, 2902–2912.
Clore, G.M., Tang, C., Iwahara, J., 2007. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. Current Opinion in Structural Biology 17, 603–616.
Clore, G.M., Venditti, V., 2013. Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. Trends in Biochemical Sciences 38, 515–530.
Clore, G.M., Wingfield, P.T., Gronenborn, A.M., 1991. High resolution three-dimensional structure of interleukin-1β in solution by three- and four-dimensional nuclear magnetic resonance spectroscopy. Biochemistry 30, 2315–2323.
Cornilescu, G., Lee, B.R., Cornilescu, C.C., *et al.*, 2002. Solution structure of the phosphoryl transfer complex between the cytoplasmic A domain of the mannitol transporter $II^{Mannitol}$ and HPr of the *Escherichia coli* phosphotransferase system. Journal of Biological Chemistry 277, 42289–42298.
Deutscher, J., Francke, C., Postma, P.W., 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiology and Molecular Biology Reviews 70, 939–1031.
Ernst, R.R., Bodenhausen, G., Wokaun, A., 1987. Principles of Nuclear Magnetic Resonance in One and Two Dimensions. Oxford: Clarendon Press.
Fawzi, N., Ying, J., Torchia, D.A., Clore, G.M., 2010. Kinetics of amyloid β monomer to oligomer exchange by NMR relaxation. Journal of the American Chemical Society 132, 9948–9951.
Fawzi, N.L., Libich, D.S., Ying, J., Tugarinov, V., Clore, G.M., 2014. Characterizing methyl-bearing side chain contacts and dynamics mediating amyloid beta protofibril interactions using $^{13}C_{methyl}$-DEST and lifetime line broadening. Angewandte Chimie 53, 10345–10349.
Fawzi, N.L., Ying, J., Ghirlando, R., Torchia, D.A., Clore, G.M., 2011. Atomic resolution dynamics on the surface of amyloid Ab protofibrils probed by solution NMR. Nature 480, 268–272.
Garrett, D.S., Seok, Y.J., Peterkofsky, A., Gronenborn, A.M., Clore, G.M., 1999. Solution structure of the 40,000 Mr phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. Nature Stuctural Biology 6, 166–173.
Gerstein, M., Edwards, A., Arrowsmith, C.H., Montelione, G.T., 2003. Structural genomics: current progress. Science 299, 1663.
Herrmann, T., Güntert, P., Wüthrich, K., 2002a. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. Journal of Biomolecular NMR 24, 171–189.
Herrmann, T., Güntert, P., Wüthrich, K., 2002b. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics program DYANA. Journal of Biomolecular NMR 319, 209–227.
Hillisch, A., Lorenz, M., Diekmann, S., 2001. Recent advances in FRET: distance determination in protein-DNA complexes. Current Opinion in Structural Biology 11, 201–207.
Iwahara, J., Clore, G.M., 2006. Detecting transient intermediates in macromolecular bnding by paramagnetic NMR. Nature 440, 1227–1230.
Iwahara, J., Schwieters, C.D., Clore, G.M., 2004. Ensemble approach for NMR structure refinement against ^{1}H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. Journal of the American Chemical Society 126, 5879–5896.
Kay, L.E., 2005. NMR studies of protein structure and dynamics. Journal of Magnetic Resonance 173, 192–207.
Korzhnev, D.M., Kay, L.E., 2008. Probing invisible low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: Application to protein folding. Accounts of Chemical Research 41, 442–451.
Kosen, P.A., 1989. Spin labeling of proteins. Methods in Enzymology 177, 86–121.
Kuszewski, J., Schwieters, C.D., Garrett, D.S., *et al.*, 2004. Completely automated, highly error-tolerant macromolecular structure determination from multidimensional nuclear Overhauser enhancement spectra and chemical shift assignments. Journal of the American Chemical Society 126, 6258–6273.
Kuszewski, J., Thottungal, R.A., Clore, G.M., Schwieters, C.D., 2008. Automated error-tolerant macromlecular structure determination from multidimensional nuclear Overhauser enhancement spectra and chemical shift assignments: improved robustness and performance of the PASD algorithm. Journal of Biomolecular NMR 41, 221–239.
Libich, D.S., Fawzi, N.L., Ying, J., Clore, G.M., 2013. Probing the transient 'dark-state' of substrate binding to GroEL by relaxation-based solution NMR. Proceedings of the National Academy of Sciences of the United States of America 110, 11361–11366.
Linge, J.P., Habeck, M., Rieping, W., Nilges, M., 2003. ARIA: Automated NOE assignment and NMR structure calculation. Bioinformatics 19, 315–316.
Loria, J.P., Berlow, R.B., Watt, E.D., 2008. Characterization of enzyme motions by solution NMR relaxation dispersion. Accounts of Chemical Research 41, 214–221.
Palmer 3rd, A.G., Kroenke, C.D., Loria, J.P., 2001. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods in Enzymology 339, 204–238.
Peruvshin, K., Riek, R., Wider, G., Wüthrich, K., 1997. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proceedings of the National Academy of Sciences of the United States of America 94, 12366–12371.
Pintacuda, G., John, M., Su, X.C., Otting, G., 2007. NMR structure determination of protein-ligand complexes by lanthanide labeling. Accounts of Chemical Research 40, 206–212.
Prestegard, J.H., al-Hashimi, H.M., Tolman, J.R., 2000. NMR structures of biomolecules using field-oriented media and residual dipolar couplings. Quarterly Reviews of Biophysics 33, 371–424.
Schulz, D.M., Ihling, C., Clore, G.M., Sinz, A., 2004. Mapping the topology of the calmodulin · melittin complex by chemical cross-linking and high resolution FTICR mass spectrometry. Biochemistry 43, 4703–4715.

Schwieters, C.D., Clore, G.M., 2001. Internal coordinates for molecular dynamics and minimization in structure determination and refinement. Journal of Magnetic Resonance 152, 288–302.

Schwieters, C.D., Suh, J.-Y., Grishaev, A., *et al.*, 2010. Solution structure of the 128 kDa Enzyme I dimer from *Escherichia coli* and its 146 kDa complex with HPr using residual dipolar couplings and small and wide angle X-ray scattering. Journal of the American Chemical Society 132, 13026–13045.

Sekhar, A., Kay, L.E., 2013. NMR paves the way for atomic level descriptions of sparsely-populated, transiently formed biomolecular conformers. Proceedings of the National Academy of Sciences of the United States of America 110, 12867–12874.

Sinz, A., Wang, K., 2001. Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: Application to nebulin-calmodulin complexes. Biochemistry 40, 7913–8903.

Solomon, I., 1955. Relaxation processes in a system of two spins. Physical Review 99, 559–565.

Takahashi, H., Nakanishi, T., Kami, K., Arata, Y., Shimada, Y., 2000. A novel NMR method for determining the interfaces of large protein-protein complexes. Nature Structural Biology 7, 220–223.

Tang, C., Iwahara, J., Clore, G.M., 2006. Visualization of transient encounter complexes in protein-protein association. Nature 444, 383–386.

Tang, C., Louis, J.M., Aniana, A., Suh, J.-Y., Clore, G.M., 2008. Visualizing transient events in amino-terminal auto-processing of HIV-1 protease. Nature 455, 693–696.

Tang, C., Schwieters, C.D., Clore, G.M., 2007. Open-to-closed transition in apo maltose-bindng protein visualized by paramagnetic NMR. Nature 449, 1078–1082.

Tjandra, N., Garrett, D.S., Gronenborn, A.M., Bax, A., Clore, G.M., 1997. Defining long-range order in NMR structure determination from the dependence of heteronuclear relaxation times on rotational diffusion anisotropy. Nature Structural Biology 4, 443–449.

Vallurupalli, P., Hansen, D., Stollar, E., Meirovitch, E., Kay, L.E., 2007. Measurement of bond vector orientations in invisible excited states of proteins. Proceedings of the National Academy of Sciences of the United States of America 104, 18473–18477.

Van Montfort, R.L., Pijning, T., Kalk, K.H., *et al.*, 1998. The structure of the *Escherichia coli* phosphotransferase $IIA^{Mannitol}$ reveals two conformations of the active site. Structure 6, 377–388.

Walters, K.J., Ferentz, A.E., Hare, N.J., *et al.*, 2001. Characterizing protein–protein complexes and oligomers by nuclear magnetic resonance spectroscopy. Methods in Enzymology 339, 238–258.

Williams, D.C., Cai, M., Clore, G.M., 2004. Molecular basis for synergistic activation by Oct1 and Sox2 revealed from the solution structure of the 42 kDa Oct1 · Sox2 · Hoxb1-DNA ternary transcription factor complex. Journal of Biological Chemistry 279, 1449–1457.

Williamson, M.P., Havel, T.F., Wüthrich, K., 1985. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1 H nuclear magnetic resonance and distance geometry. Journal of Molecular Biology 182 (295), 315.

Wu, Z., Tjandra, N., Bax, A., 2001. ^{31}P chemical shift anisotropy as an aid in determining nucleic acid structure in liquid crystals. Journal of the American Chemical Society 123, 3617–3618.

Wüthrich, K., 1986. NMR of Proteins and Nucleic Acids. New York, NY: John Wiley.

Yee, A., Pardee, K., Christendat, D., *et al.*, 2003. Structural proteomics: Toward high-throughput structural biology as a tool in functional genomics. Accounts of Chemical Research 36, 183–189.