Protein-protein interactions

Di Savino, A.; Ubbink, M.; Luchinat, C.; Parigi, G.; Ravera, E.

Paramagnetism in Experimental Biomolecular NMR

CHAPTER 5

Protein-Protein Interactions

M. Ubbink[a*] and A. Di Savino[a]

[a] Leiden University, Institute of Chemistry, Einsteinweg 55, 2333 CC Leiden, The Netherlands

*Corresponding contributor. E-mail: m.ubbink@chem.leidenuniv.nl

Abstract

Paramagnetic NMR methods are excellently suited for the study of protein-protein complexes in solution. Intermolecular pseudocontact shifts (PCSs), residual dipolar couplings (RDCs) and paramagnetic relaxations enhancements (PRE) can be used, ideally in combination, for docking proteins and determining their orientation in the complex. PCSs can be used for breaking the structure symmetry in dimer complexes. PCSs also can be applied to detect structural differences in proteins and protein complexes in solution in comparison to crystal structures. RDCs are sensitive to the degree of alignment of both partners in a protein complex and are thus very useful to detect dynamics within complexes. PREs can detect states in which nuclei approach a paramagnetic centre closely, even if it exists only for a small fraction of the time. Thus, PREs are used to detect minor states and characterize ensembles. PRE studies have been the foundation for characterizing encounter states and the process of protein complex formation. In weak complexes, such as found in electron transfer chains, proteins can be in an encounter state for a large fraction of the complex lifetime. Paramagnetic NMR tools thus have found many applications for studying protein complexes, and more may be on the horizon.

5.1 Introduction

Proteins function by interactions with other molecules. These can be small, such as enzyme substrates and allosteric modulators or large, such as lipids, polysaccharides, nucleic acids and other proteins. Protein-protein interactions (PPI) are ubiquitous in living organisms and vital to life. The nature of the interaction is tuned to biological function, from very strong and semi-permanent to weak and highly-transient. Structural characterization of protein-protein complexes can be done with X-ray diffraction of crystals, electron microscopy (EM), mass spectrometry or NMR. For stable, long-lived complexes the approach is largely the same as for single proteins and in those cases paramagnetic NMR is complementary (see Chapter 3). It can provide additional distance and angle restraints for structure calculations, solve symmetry problems in dimers and assist in mapping of binding sites. For weak complexes, which are difficult to crystallize and too unstable for EM, NMR is the only method that can provide atom-level insight. Not only can the stereospecific, ground state structure of the complex be characterized, also the dynamic encounter state can be studied. For PPI, the relevant paramagnetic effects are PCSs, PREs[1-4] and pRDCs.

5.2 Protein-protein interactions

The affinities of PPI range from picomolar to millimolar. For a simple complex of two proteins, the dissociation constant, K_D (in M), is the ratio of the dissociation rate constant (k_{off}, in s^{-1}) and the association rate constant (k_{on}, in $M^{-1}s^{-1}$):

$$K_D = \frac{k_{off}}{k_{on}} \qquad\qquad (5.1)$$

The association of two proteins into a well-defined, stereospecific complex requires that the partners find each other via diffusion. The chance that they meet in the correct orientation to form the complex is very small, which implies that many collisions are non-productive. The number of collisions is limited by diffusion and is of the order of 10^9 $M^{-1}s^{-1}$. However, for non-optimized PPI, the k_{on} –the rate constant for *productive* complex formation– will only be around $10^4 – 10^5$ $M^{-1}s^{-1}$, assuming that no major conformational changes are required to form the complex. If rapid formation of the complex is not required for its function, such association rates will be found. Association can be accelerated by electrostatic

complementarity to enhance the chance of collision and the lifetime of the encounter. Also, if the charges are distributed non-symmetrically on the protein, creating a dipole, electrostatic pre-orientation during the encounter can favour certain collisional orientations over others. Thus, with complementary charged patches near the stereospecific binding sites on both proteins, the chance of colliding in approximately the correct orientation can be enhanced. In this way, the fraction of productive encounters, which lead to the active complex, can be dramatically enhanced, to k_{on} values near to the diffusion limit in favourable cases.[5, 6]

The biological function of the complex often determines the k_{off}. The lifetime of the complex is defined as $1/k_{off}$, so complexes that require a long life-time must have a low k_{off}, for example of the order of 10^{-3} s^{-1}. With $k_{on} = 10^5$ M^{-1}s^{-1}, the affinity is then already in the low nanomolar range ($K_D = 10^{-8}$ M, see Equation 5.1). For complexes that require even higher affinity, for example toxin-inhibitor complexes, the association rate constant is raised by electrostatic interactions, bringing the K_D in the picomolar range.[7] On the other end of the spectrum are the weak interactions, which are meant to exist fleetingly, such as electron transfer and cell signalling complexes. Their biological function is to act quickly, so k_{off} values are of the order of 10^3 s^{-1}. To still get some affinity –and thus specificity– electrostatic interactions are often invoked, resulting in K_D values in the micromolar range. Even weaker interactions, with $K_D > 1$ mM, are observed in the case of self-association, i.e. dimerisation or polymerisation of proteins at high concentrations. Such interactions can be relevant in fibril formation.[6]

The association and dissociation rates have a direct bearing on the NMR spectra, because they determine the exchange regime. In very tight complexes, the amount of free protein can be neglected and exchange effects between free and bound forms will be minimal. However, in weaker complexes, the fraction of free protein will be significant. In 'classic' NMR the effects of exchange on the observed chemical shift and linewidth are well established and described by the McConnell equations.[8] In general, any NMR observable will exhibit averaging effects when a nucleus exchanges between states in which that observable has different values. As is explained below, in the case of paramagnetic NMR on protein-protein complexes, the paramagnetic centre is usually on one protein and the effects are observed on the other. Thus, for the complex, the paramagnetic effect is observed, whereas the free protein does not experience such an effect. In the case that the exchange rate (k_{ex}) is very fast, of the order of 10^3 times faster than the difference between the frequencies of the relevant NMR observable(s) in the different states (expressed in rad s^{-1} for shifts and s^{-1} for relaxation), the measured value of the observable will be the average of the values of the individual states,

weighted with their populations. This assumption is often made, but it should be realized that, in particular for PREs, it may not be valid, because PREs can be very large indeed for nuclei that are close to the paramagnetic centre, resulting in underestimation of the PREs in that state.[9, 10] Therefore, it is common to use lower bounds only for restraints derived from large PREs.
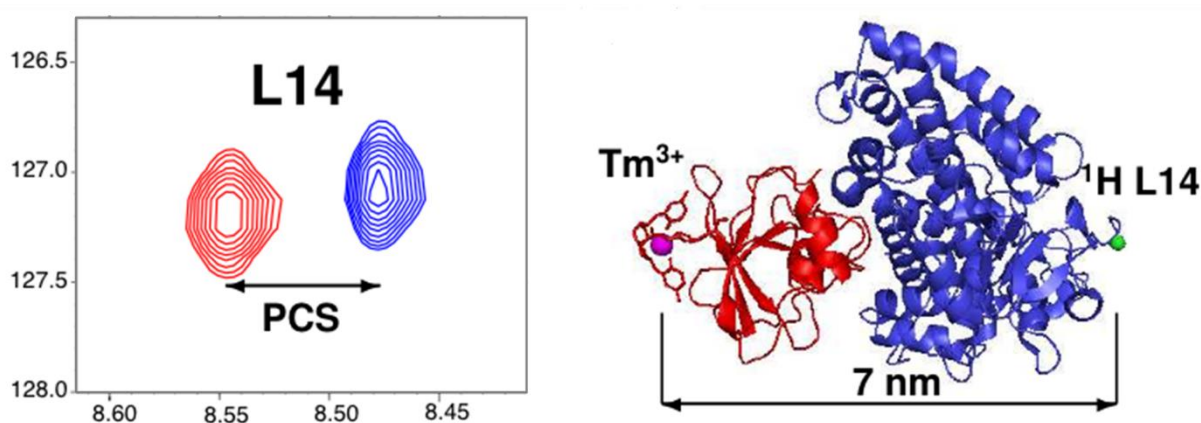

Figure 5.1 Illustration of the long-range effect of PCS. A clearly significant intermolecular PCS was observed for the amide proton of Leu14 of cytochrome P450cam caused by a Tm^{3+} ion in the paramagnetic tag CLaNP-7 [64] attached to putidaredoxin over a distance of 70 Å (7 nm) [63]. Reprinted from [25], Copyright (2014), with permission from Elsevier.

5.3 The new toolbox

For weak complexes, conformational changes often do not occur upon complex formation and only the binding site and orientation of the proteins need to be determined to establish the overall structure of the complex, provided the structures of the free proteins are available. In such cases, backbone labelling of the smaller protein and paramagnetic tagging of the larger one can provide sufficient information to obtain a model of the complex in solution. Paramagnetic NMR is particularly suited for the study of such protein complexes. The paramagnetic effects can usually be measured in simple 2D HSQC-like spectra, requiring only low sample concentrations. Due to the strong spin of an unpaired electron, long-range effects can be observed when measuring paramagnetic nucleus-electron interactions.

PCSs have been reported up to 70 Å (Figure 5.1) using a rigid tag and Tm^{3+} as the paramagnetic ion. Protein complexes are usually rather large and tags should not be close to

the interface, so long-range distance restraints are necessary for structural characterisation. PCSs are obtained by determining resonance positions in the spectra of the paramagnetic sample and a diamagnetic control that should be as similar as possible. Usually, this is achieved by using a sample with the same components except that the paramagnetic metal is replaced by a diamagnetic one with similar properties (e.g. Lu^{3+} for lanthanoids). PCSs, like chemical shifts, can be measured with high precision and are described very well by the standard point-dipole approximation (Section 1.2.2.1, Equation 1.50) for nuclei further than 5 Å from the paramagnetic centre. Contact contributions (Section 1.2.1) can generally be neglected for nuclei more than four bonds away and certainly in complexes in which one protein is tagged and the other observed. Thus, with the proper diamagnetic control PCSs can be measured with high accuracy, and for larger PCSs, also high precision. PCSs of ^1H nuclei are preferred over those of heavy atoms because the latter are more affected by residual chemical shift anisotropy and pRDCs. It is non-trivial to determine the orientation and size of the $\Delta\chi$ tensor and the location of the metal in the case of protein complexes, because the nuclei are in one protein and the centre is on the other. Ideally, these parameters are determined independently using a sample in which the protein with the centre is isotope-labelled. Then, the tensor is determined in the regular way described for free proteins (Section 3.3). If it is not possible, the tensor parameters must be optimized iteratively during the structure calculations.[11]

Paramagnetic centres with an anisotropic magnetic susceptibility (described by the $\Delta\chi$ tensor) not only cause PCSs but also induce partial alignment of the protein or protein complex in a strong magnetic field (Section 1.2.4). Similar to alignment caused by external media (bicelles, gels, bacteriophages), RDCs are observed as a consequence (Section 1.3). In rigid complexes, the degree of alignment in the partner with the paramagnetic centre and the one without is the same. Intermolecular RDCs can be used along with the PCSs and PREs for the determination of the structure of the complex. Perhaps even more interesting is the use of such pRDCs to demonstrate the presence of dynamics in a protein complex, in analogy to multidomain dynamics (Chapter 4). RDCs are not dependent on the distance between paramagnetic centre and the nuclear vector, only on the degree of alignment. If the partners in a complex show dynamics, the alignment of the protein without the paramagnetic centre will be reduced relative to that of the protein that contains the centre (see Section 4.2.2 and below).

PRE effects at 35-45 Å can be measured in cases where the complex is big, resulting in a long rotational correlation time (τ_r), and Gd^{3+} is used as the paramagnetic centre (Section 3.2).[1, 6]

PREs are normally generated by paramagnetic centres with long electronic relaxation times (> 100 ns), such as spin labels and Gd^{3+} ions. Such systems are isotropic, implying the absence of PCSs and pRDCs. Relaxation is then dominated by Solomon dipole interactions (Section 1.4.1.1, Equations 1.84 and 1.85). An advantage is that only peak intensity reductions occur due to PREs but no resonance shifts, so reassignment of the spectrum is not required. As already discussed in Sections 4.2.6 and 4.3.4, PREs are particularly sensitive to minor states, such as present in ensembles of the encounter state. Apart from the stereospecific complex, other orientations of the proteins in the complex occur, collectively forming the encounter state. In particular in a weak complex with strong electrostatic interactions, the fraction of the encounter state can be prominent, as will be discussed in more detail later. The encounter state is best described as an ensemble of orientations that rapidly interconvert. The orientations in which an observed nucleus approaches a paramagnetic centre closely will result in (very) high PRE for that nucleus. Consequently, even orientations within the complex with a low population can be detected due to the PRE. Thus, this paramagnetic effect has been instrumental in the description of encounter complexes.

Similarly, paramagnetic centres that are not attached to a protein but are used as co-solutes cause PREs on the protein nuclei near to the solvent (see Chapter 10). By comparing such 'solvent PRE' (sPRE) for free protein and protein in a complex, binding sites can be identified due to a shielding effect in the complex from the sPRE effect.[12] The sPRE data are dependent on the kind of probe, its concentration and solubility limit. Usually sPRE are expressed as the increase in relaxation rates or decrease of signal intensities per mM of paramagnetic probe. Metal ions, organic radicals or metal chelates have been used as probes.[13] For example, [Gd-DOTAM]$^{3+}$ and [Gd-DOTP]$^{5-}$ are two charged paramagnetic probes tested in the study of the protein-protein complex formed by rubredoxin and cytochrome c_3, two proteins with opposite charges at their binding sites. sPRE experiments based on the interactions of the two probes with the surface of the proteins allowed to gain information about charged patches on the protein surface and to characterize the binding site in the protein complex.[14]

As each of the paramagnetic effects has its peculiarities, it is emphasized that combining restraints based on PCSs, RDCs and PREs is a safeguard against biased results. For example, structure calculations based only on PREs may be prone to bias toward minor states, because the PREs are so sensitive to those. Such minor states can also represent weak non-specific interactions in concentrated protein solutions. PCSs and RDCs are not sensitive to such interactions because the effects will be weaker and, anyway, average out for dynamic systems.[15] PCSs are dominated by their distance dependence, whereas RDCs are distance

independent and sensitive to rotation but not translation. RDCs are also intolerant toward variations in the structures of the free proteins.[16] All three effects are degenerate because none determine a unique position of a nucleus relative to the paramagnetic centre. For example, the PRE depends on the distance r in an isotropic way, thus a PRE only defines a sphere with radius r around the centre on which the nucleus will lie. Similarly, multiple combinations of two angles and a distance of the position of a nucleus in the frame of the $\Delta\chi$ tensor yield the same PCS (Figure 1.2). It is advisable to use data from more than a single paramagnetic centre for the purpose of triangulation. Also the use of data from techniques other than paramagnetic NMR (NOE, SAXS, EPR, FRET) can help to reduce bias.[17, 18] It is noted, however, that paramagnetic NMR is capable of delivering a wealth of information because paramagnetic spectra yield hundreds of data points in one experiment, one for each nucleus, whereas FRET and EPR usually produce a single distance for one sample.[19] Figure 5.2 shows schematically which set of experiments is required to obtain a complete set of PCSs, RDCs and PREs for structure determination of a protein complex.

As discussed in Chapter 2, various tags can be used to introduce paramagnetic centres. Clearly, for long-range effects, the stronger paramagnetic metals (Tm^{3+}, Tb^{3+}, Dy^{3+}, Er^{3+}) should be used. Also, the more rigid the centre is relative to the protein, the less averaging effects will occur, resulting in larger and more accurate paramagnetic effects.
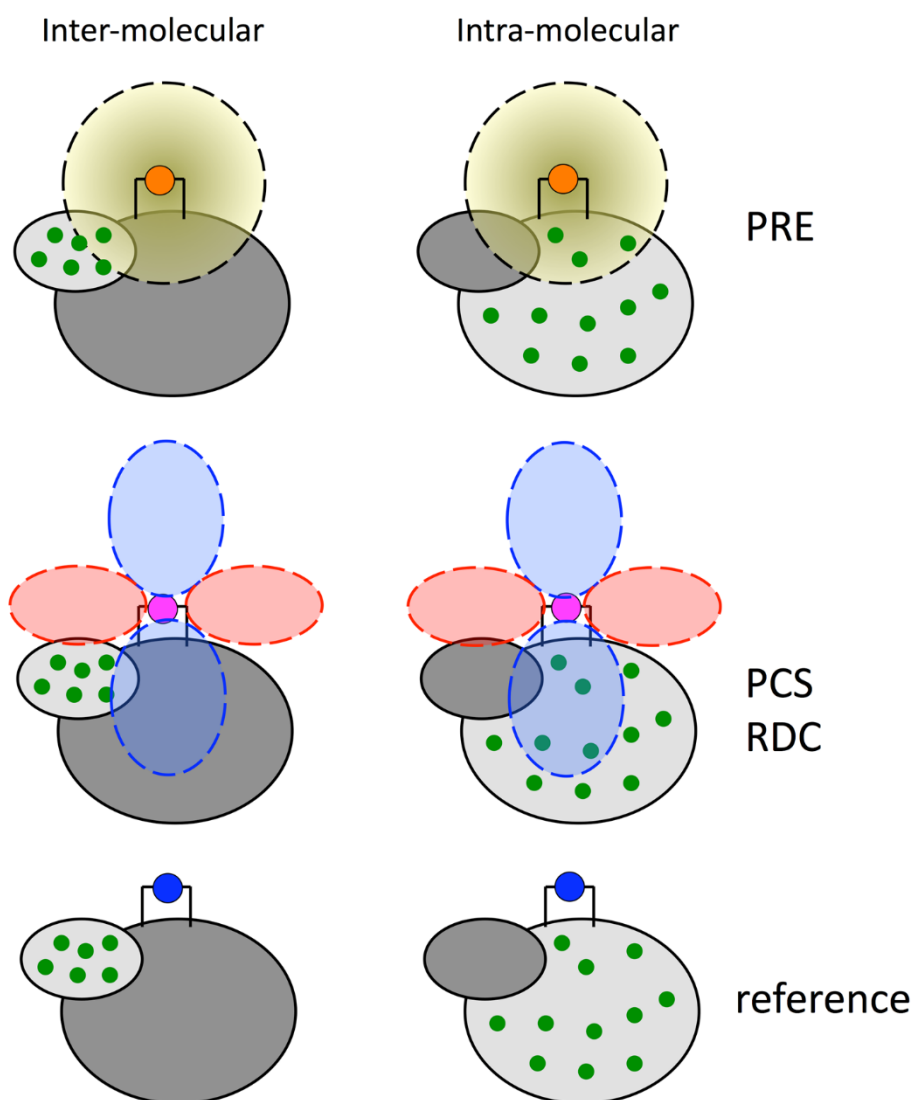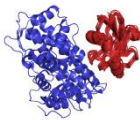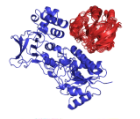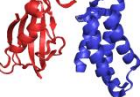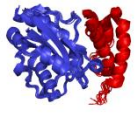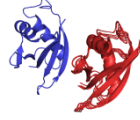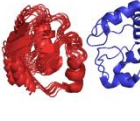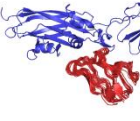
Figure 5.2 Strategy for measuring PCSs, PREs and pRDCs for structure determination of protein–protein complexes. Six samples are produced. PREs are extracted from a complex tagged with an isotropic paramagnetic centre, for example, Gd^{3+}. PCSs and RDCs are obtained from complex tagged with an anisotropic paramagnetic centre, for example, Tm^{3+}. A diamagnetic control complex, for example tagged with Lu^{3+}, serves a reference. To decrease signal overlap the inter-molecular and intra-molecular are determined independently with different samples in which either one or the other proteins is isotope labelled (green dots). The intra-molecular restraints are used to position the paramagnetic centre and orient the magnetic susceptibility tensor. A combination of ten 2D spectra (such as TROSY and HSQC) is sufficient to obtain a complete set of intermolecular and intra-molecular PREs, RDCs, and PCSs. Reprinted from [25] Copyright (2014), with permission from Elsevier.

5.4 Tight complexes: breaking symmetry

Structure determination of tight obligate or non-obligate complexes[20] can be performed via the classic approaches (X-ray crystallography, EM, NMR). Paramagnetic NMR can provide distance and angle restraints additional to NOE and chemical shift information. A particular application of paramagnetic NMR is in breaking the symmetry in dimers. Solving the structure of a dimer composed of symmetric units can be very difficult using conventional NMR because of the necessity to distinguish between NOEs generated by nuclei within one subunit (intra-subunit NOEs) and those generated by nuclei in different subunits (inter-subunit NOEs). Several solutions have been proposed, such as asymmetric isotope labelling in combination with filtered NOESY experiments,[21] or by computationally extracting the correct orientation of one of the two units in the complex using a function describing the contributions generated by intra- and intermolecular NOEs.[22] Another solution is offered by paramagnetic NMR, which can give an unambiguous answer to the symmetry problem and at the same time provide long-range intermolecular restraints for the determination of the structure of the dimer. The solution consists of tagging of just one of the two subunits with a paramagnetic centre. By tagging only one of the subunits, asymmetry is introduced that is reflected in the generation of non-equivalent paramagnetic effects, e.g. PCSs for nuclei in the tagged subunit (intra-subunit PCSs) and in the untagged subunit (inter-subunit PCSs).[23] CylR2, a protein involved in the cytolysin production in *Enterococcus faecalis*, forms a symmetric homodimer of 15.4 kDa. The structure in solution was determined using different sources for experimental intersubunit information, followed by rigid-body docking. The following sets of restraints were compared: PREs, PREs and RDCs, NOEs and RDCs, and NOEs, PREs and RDCs. PREs were collected using two samples of single mutants of CylR2 tagged with MTSL at different locations. The structure obtained with only PREs deviated from the X-ray structure used as reference with a backbone root means square deviation (RMSD) of 3.0 Å, and the addition of RDCs reduced the deviation to 1.5 Å. The combination of PREs, RDCs and NOEs increased the accuracy of the model slightly, whereas restraints from only NOEs and RDCs showed a deviation of 2 Å from the crystal structure. Docking without experimental restraints yielded an accuracy comparable to the one obtained with only the set of experimental PREs. The inclusion of RDC restraints gave a high-resolution structure.[24]

Table 5.1 Structure of protein complexes solved using restraints from anisotropic paramagnetic centres. Reprinted from [25], Copyright (2014), with permission from Elsevier.

Molecule 1 (mass)	Molecule 2 (mass)	Paramagnet	Tag/Me-ligand	Structure	Reference (PDB ID)	Comment
Cyt P450cam (45 kDa)	putidaredoxin (11 kDa)	Tm^{3+}, Gd^{3+}	ClaNP-7		[24] (2M56)	
ADR (51 kDa)	Adrenodoxin (14 kDa)	Tm^{3+}, Eu^{3+}, Gd^{3+}	ClaNP-5		[34] (not in PDB)	
FKBP12 (12 kDa)	mTOR FRB (11 kDa)	Dy^{3+}, Tb^{3+}	LBT-peptide		[28] (2RSE)	
ε186 (21 kDa)	Θ (9 kDa)	Dy^{3+}, Er^{3+}, Ho^{3+}	Native Mg^{2+}/Mn^{2+} site		[29,39] (2XY8)	
PB1 (10 kDa)	PB1 (10 kDA)	Tb^{3+}, Dy^{3+}, Er^{3+}, Tm^{3+}	LBT-peptide		[35] (2KTR)	Asymmetric homodimer
Cyt c (12 kDa)	Adrenodoxin (12 kDa)	Fe^{3+}	Native heme		[43](2JQR)	Covalently cross-linked
STAT-4-NT (14 kDa)	STAT-4-NT (14 kDa)	Co^{2+}	EDTA tag		[38] [58][a] (1BGF)[a]	Symmetric dimer
Cyt f (27 kDa)	plastocyanin (11 kDa)	Fe^{3+}	Native heme		[36] (2PCF, 1TU2,1TKW,2JXM)	Solved for different species

[a] Crystal structure, no solution structure deposited

5.5 Ground states structures of protein complexes

In cases where determination of the structure of a protein complex is difficult using X-ray crystallography and classic NMR approaches, paramagnetic NMR can offer a valid alternative, at least for cases in which the individual structures are available and no backbone conformational changes occur upon complex formation. Paramagnetism can provide long-

range restraints for rigid-body docking to determine the location and orientation of the partners in the complex. In principle, also conformational changes upon complex formation can be modelled, provided sufficient restraints are obtained. However, inherent in measuring long-range distance restraints is that they are less sensitive for small structural variations, so contrary to NOEs, paramagnetic restraints obtained with strong paramagnetic centres are not suitable for modelling details of structures. Weak paramagnetic centres provide information over shorter distances (as contact contributions do), but they require the centre to be close to the nuclei of interest. Thus, the structures of protein complexes reported so far have usually been obtained via rigid body docking on the basis of the paramagnetic restraints. Table 5.1 gives examples of such structures based on anisotropic paramagnetic centres (PCS/RDC data).[25] The idea to determine a protein-protein complex structure through PCSs was applied for the first time to the transient complex of two photosynthetic electron transfer proteins from plants, plastocyanin and cytochrome *f*.[26] The low-spin paramagnetic heme Fe^{3+} ion in cytochrome *f* was used to generate intermolecular PCSs in ^{15}N labelled plastocyanin. The structure showed how a short electron transfer pathway can be formed between the heme in the cytochrome and the copper in plastocyanin, in line with the rapid electron transfer that occurs between these proteins, solving a long-standing debate about the main site of interaction of plastocyanin. The use of naturally occurring paramagnetic ions is, unfortunately, only possible in few cases. In the structure determination of the 30 kDa complex of the Θ subunit and the N-terminal domain of the ε proof-reading exonuclease subunit, ε186, from *Escherichia coli* polymerase III, tagging of the proteins was also not necessary.[16] In this case, metal-substitution was used to create a paramagnetic centre within the ε subunit. The pair of divalent cations (Mn^{2+}/Mg^{2+}) in its active site was substituted by lanthanoid ions. Dy^{3+} or Er^{3+} were used to generate inter-subunit PCSs, whereas the diamagnetic control was the apo-complex. The rigid-body structure of the complex obtained using PCSs based restraints was verified with NOEs measured in complexes composed of $^{15}N/^{2}H$ labelled ε186 and unlabelled Θ as well as with ^{15}N labelled Θ and unlabelled ε186.

An early example of structure determination of a non-metalloprotein complex using paramagnetic tagging with a double-arm LBP (see Chapter 2) is provided by the p62 PB1 domain.[27] p62 is a multidomain protein involved in autophagy.[28] It forms multimers through PB1 domain interactions in a front-to-back fashion but it was possible to obtain a 20 kDa dimer of the PB1 domain through mutagenesis. PCSs generated by four different lanthanoid ions were used to solve the structure of the dimer.[29]

The power of PCSs in the determination of protein-protein complexes was also illustrated by the large complex (65 kDa) of adrenoxin reductase (AdR) and adrenodoxin (Adx) using the CLaNP-5 tag.[30, 31] This complex is part of the electron transfer pathway from NADH to mitochondrial cytochromes P450 in the adrenals. In this study, AdR (51 kDa) was tagged at two positions and Adx (14 kDa) was isotope-labelled with ^{15}N and ^{2}H for observation. Inter-subunit PCSs, generated by Tm^{3+}-CLaNP-5, and PREs (Gd^{3+}), were used to dock Adx on AdR on the basis of the structure of the free proteins.[11] The $\Delta\chi$ tensor parameters for Tm^{3+}-CLaNP-5 were optimized during the docking in an iterative approach. No NMR information on AdR was used in this study. A complicating factor was the naturally occurring paramagnetic FeS cluster in Adx. In this case, this cluster could not be used to generate restraints, as no assignments were available for AdR nuclei. Rather, it complicated the study because its relaxation properties broaden out the signals from many hydrogens in its environment. The cluster is close to the interface in the complex, so most of the resonances of Adx that could be expected to exhibit chemical shift perturbations upon complex formation were not observed. However, the inter-subunit PCSs are long-range (up to 56 Å in this case) and could be observed for Adx nuclei far from the cluster and the binding site and the derived restraints could be used for rigid-body docking. The calculated location and orientation of Adx on AdR were in good agreement with a crystal structure of the complex obtained via cross-linking.[32]
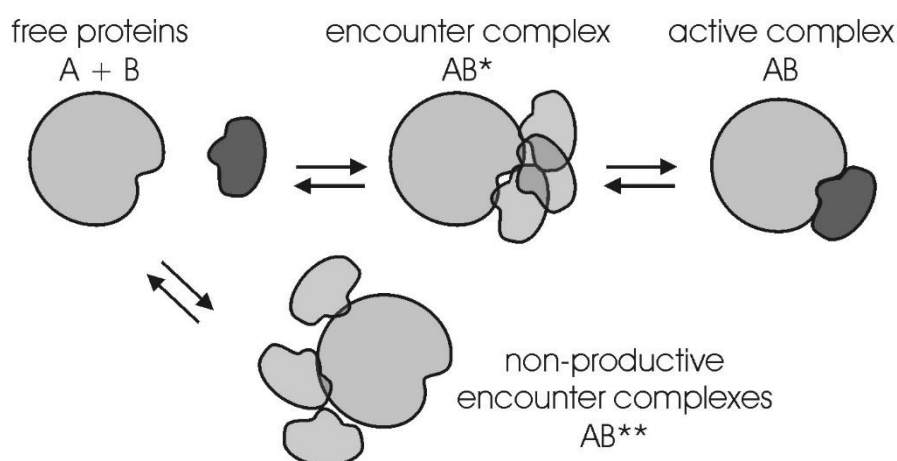

Figure 5.3 Model for protein complex formation. The free proteins A and B can associate in productive or non-productive encounter complexes. The former can evolve into the active, stereospecific complex.

5.6 Dynamics and encounter states

Figure 5.3 gives a schematic model of the process of protein complex formation. When two freely diffusing protein molecules approach each other, they will, generally, not collide in the correct orientation to form the stereospecific complex. First, an encounter complex is formed in which the partners sample multiple orientations by rotational diffusion and weak interactions. Encounters can be futile, i.e. not lead to the stereospecific complex. The proteins part again without having formed a productive complex.[33] In other cases, the encounter complex will proceed to the stereospecific complex, a productive encounter. In that case the association is successful. The association rate constant k_{on} reports successful encounters. As explained above, the fraction of productive encounters can be increased by prolonging the encounter state and directing the partner toward the stereospecific binding site. Prolongation of the encounter state enables the proteins to sample a larger area of the partner, in what has been considered as a reduced dimensionality search for the binding site.[34, 35] Directing the partner to the binding site reduces the sampling of irrelevant areas. Both effects can be achieved by complementary charge interactions and also hydrophobic contacts in the encounter state can lead to enhanced association rates.

Paramagnetic NMR techniques have played an essential role in providing evidence for the model of protein complex formation and the characterisation of the encounter states. These states usually (though not always, see below) represent a small fraction of the complex, with the proteins being the stereospecific complex the majority of the time. Furthermore, it appears from many studies that chemical shift perturbations in the encounter state are minimal. They are mostly caused by the interactions in the stereospecific state. The chemical shift perturbations observed in ^{15}N-^{1}H HSQC-like spectra are due to changes in the polarization and hydrogen bond formation propensity of the amide groups, requiring changes in solvation. It is likely that in the encounter state the proteins are still fully solvated, so no large chemical shift changes are expected. The encounter state consists of an ensemble of many orientations in fast exchange, making it hard to detect NOEs or indeed PCSs and impossible to crystallize. PREs, with their sensitivity for minor states, are excellently suited for the characterization of encounter states. Orientations in which nuclei are brought close to a spin label will result in a PRE, even if the population of that orientation is very low (< 1%). However, the observed PRE is an average of all orientations, provided they are in fast exchange relative to the largest PRE (which can several thousands s^{-1}). As an infinite number of combinations of orientations can yield the same average PRE, it is not possible to determine the orientations present in the

encounter ensemble directly from the PRE. Additional assumptions and models must be invoked to visualize the encounter state (see also Chapter 4 for the discussion of ensemble descriptions). By using PRE tags on multiple sites on the protein surface and observing the effects on the partner protein it is possible to map the areas that are *not* visited. In this way, an exclusion map can be made, limiting the extent of the encounter state.[36-40] The absence of inter-molecular PREs is the most solid evidence obtained in such experiments. Observing PRE indicates that the region where the tag is localized is being visited but not to what extent. It is possible that nuclei that experience the PRE get very close to the tag very briefly or spend more time at a somewhat larger distance. The effect on the average PRE will be the same. It should also be realized that tags can influence the complex. For structure determination of ground states, its effects are easily spotted, because chemical shift perturbations or the K_D of the complex with the tagged protein may differ from the wild type complex. Such tag positions should then be discarded. However, tags may also affect the encounter state and such effects may go unnoticed. It is thus advisable to use small tags, such as MTSL, and use data from many locations on the protein surface. Several examples of encounter state complex studies will be discussed later.

Some processes, like the formation of crystals, viral capsids, amyloid fibrils or amorphous aggregates, are driven by the spontaneous association of macromolecules, identical subunits that form a complex of a high level of order. During the formation of these structures, the molecules shift from the free state in solution to a macromolecular cluster by employing ultra-weak intermolecular interactions ($K_D > 1$ mM). Visualization of the complexes formed by ultra-weak interactions is particularly difficult with most common biophysical techniques because their population is extremely low. In principle PREs can be used to explore such complexes and elucidate the nature of the ultra-weak association because of its high sensitivity to lowly populated states. This approach was employed to investigate the properties of the formation of non-fibrillar oligomers formed by the amyloid β (Aβ) peptide [41], which could be the primarily responsible agent for neurotoxicity in the Alzheimer's disease. While Aβ fibrils are extremely stable, it was found that the Aβ(1-40) peptide forms large, dynamic oligomers, in which the peptide state changes continuously from the free to the bound state. No transverse PREs were measured above the reliable detection limit of about 5 s^{-1}. Since the PRE experiment is able to visualize low populated states in fast exchange, it was concluded that the population of transient oligomers amounts to less than 1-2%, or that the exchange between bound and free state is slow on the PRE time scale (<ca. 1 ms).

5.7 Examples of applications

5.7.1 Breaking the symmetry of the STAT4 complex

Paramagnetic NMR was used to break the symmetry and determine the structure of the 28 kDa symmetric complex formed by two copies of the N-terminal domain of STAT4 (STAT4$_{NT}$).[23] STAT4 is involved in signal transduction and activation of transcription in hematopoietic cells, required for the differentiation of T helper 1 cells and optimal IFN-γ production.[42] The complex was tagged sub-stoichiometrically with a thiol-reactive EDTA bound to Co^{2+} as paramagnetic probe, labelling just one subunit. It was then possible to obtain PCSs and RDCs, providing long-range restraints, both for distance and orientation, to understand how the two subunits bind. The PCSs observed for the nuclei in one subunit were mostly of the opposite sign compared to those for nuclei in the second subunit, which facilitated the analysis. Moreover, the comparison between the observed PCSs and those calculated on basis of the crystal structure and the NOE-derived structure of the subunit showed a good correlation. The PCSs were then used as unambiguous intermolecular orientational and distance restraints to study the molecular dynamics and establish the orientation of the two subunits in the dimer in solution, obtaining a structure that agreed well with the PCS restraints. The RDCs were used to determine the relative orientation of the two subunits independently from the PCSs, obtaining the same result. This approach is a good solution for symmetry problems in dimers, in particular for large complexes for which perdeuteration is necessary.

5.7.2 Synaptotagmin-1-SNARE complex

A recent example of application of PCSs to study dynamics in a protein-protein complex is the work on the synaptotagmin-1 C$_2$B domain bound to the SNARE complex, with a total size of the complex of 51 kDa.[43] One of the processes which allow a neurotransmitter to be released is Ca^{2+} dependent and the mechanism is controlled by proteins like synaptotagmin-1 (Syt1) and the SNARE complex, comprising sinaptobrevin, syntaxin-1 and SNAP-25. The interactions between Syt1 bound to Ca^{2+}, the SNARE complex and complexins allow Syt1 to rapidly release the neurotransmitter, but a molecular model of the mechanism was still lacking, because no high-resolution structure of Syt1-SNARE complex was available. PCSs induced by tagging the SNARE complex at two sites with the Dy^{3+}-C2 probe[44] (Section 2.3.2.1.5.2) were used to investigate the nature of the Syt1:SNARE complex. The paramagnetic effects on Syt1 show that the basic region on the C$_2$B domain of Syt1 binds to

an acidic region of the SNARE complex, constituted by residues of syntaxin-1 and SNAP-25. On the basis of these and other results a model for the role of Syt1 in membrane fusion was proposed. Moreover, the complex was found to be dynamic, as it was not possible to explain the PCS data by a single location and orientation of Syt1 with respect to the SNARE complex. It is possible that the dynamic nature of the Syt1-SNARE complex is necessary for the complex function, to allow a fast release of the neurotransmitter.

5.7.3 Cytochrome f and plastocyanin

In photosynthesis, inter-protein electron transfer needs to be fast to avoid limiting the overall rate of the light reactions that are responsible for the production of NADPH and ATP, required for carbon fixation. A key electron transfer step is the one from cytochrome f (Cytf) in the cytochrome b_6f complex to the copper protein plastocyanin (Pc). Pc shuttles the electrons, one at a time, to photosystem I. In different branches of the tree of life, it appears that different solutions have been found to ensure rapid formation and dissociation of the Cytf-Pc complex. To obtain fast electron transfer from one redox centre (the heme in Cytf) to the next (the copper ion in Pc), Marcus theory[45] states that the distance between the centres should be short (< 16 Å)[46]. As long as the distance is short and the proteins make contact to reduce the penalty of through-space exchange coupling, the exact orientation of the proteins does not matter for electron transfer, contrary to reactions in which atoms are transferred. Thus, the aim in electron transfer complexes is not necessarily to reach a single stereospecific complex. Rapid formation and dissociation are more important. Consequently, it is observed that in such complexes often the encounter state represents a large fraction of the complex.

The first Cytf-Pc complex studied was that of plants. The plant Pc has a hydrophobic patch close to one of the ligands of the copper ion and a negative patch on the side of the protein (e.g. PDB ID 4PCY[47]), illustrated in Figure 5.4. Plant Cytf is mostly negative but has positive residues in a patch close to the heme group.[48] Thus, it is no surprize that complex formation depends much on the ionic strength, indicating a favourable electrostatic interaction.[49] Intermolecular PCSs from the Cytf heme iron on isotope labelled Pc could be observed and used to dock Pc in a well-defined orientation on Cytf.[26] The results suggested the presence of an orientation that was populated for a considerable fraction of the time, but the data also provided evidence that the complex visited an encounter state. The well-defined state (PDB ID 2PCF) clearly showed that fast electron transfer was possible from the iron to the copper via the His ligand located in the hydrophobic patch of Pc. See Figure 5.4, in which the cyan surface represents His and the hydrophobic areas are coloured green.

17

Subsequent work on the same complex from the cyanobacterium *Phormidium laminosum* showed a different picture.[50] This complex was much less sensitive to the presence of salt, suggesting a reduced role of electrostatic interactions. This finding was in line with the surface of *P. laminosum* Pc, which lacks the negative patch (Figure 5.4). Cytf from *P. laminosum*, on the contrary, is overall very negatively charged, indicating that the lack of electrostatic interactions is due to the properties of Pc. Again, intermolecular PCSs could be detected, however, a single structure could not describe the data. An ensemble of orientations was necessary to obtain a reasonable fit and in this ensemble Pc only binds via its hydrophobic patch, contrary to plant Pc.
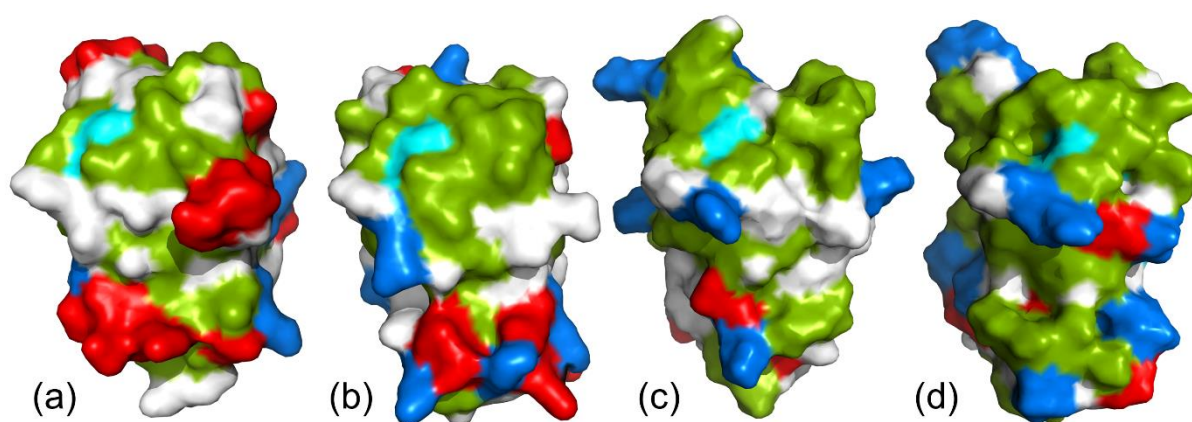

Figure 5.4 Surfaces of Pc for which the complex with Cytf was determined. (a) Spinach [91] (PDB ID 1AG6); (b) *Phormidium laminosum* (PDB ID 2Q5B); (c) *Prochlorothrix hollandica* [92] (PDB ID 1B3I); (d) *Nostoc* sp. (*Anabaena variabilis*) [93] (PDB ID 1NIN). Pc is shown in surface representation coloured according to surface charge, with red for Asp and Glu and blue for Lys and Arg residue. His residues are in cyan. Aliphatic and aromatic groups are in green. The figure was produced using PyMol (The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC).

Results intermediary between plant and *Phormidium* were obtained for the complexes from *Nostoc*[51] and *Prochlorothrix hollandica*.[52] The Pc surfaces are not highly charged, similar to Phormidium Pc (Figure 5.4). As the studies indicated the presence of a substantial fraction of encounter complex, *Nostoc* Cytf was tagged with MTSL radicals to generate intermolecular PREs on Pc and sample the Cytf surface area that is visited.[53-55] It was observed that Pc samples a large area of Cytf in approximately the same orientation, involving a prominent positive residue (Arg) and the Pc hydrophobic patch. Electrostatic calculations clearly demonstrate that charge interactions alone cannot explain the encounter orientations observed.

Hydrophobic interactions also contribute to the stability of the encounter complex in the *Nostoc* Cytf-Pc complex, in line with theoretical studies.[56, 57] This work led to a model in which the rate of the formation of an electron transfer active complex is enhanced by gradual increase of the hydrophobic overlap between the surface of Cytf and Pc to smooth out the transition barrier of desolvation (Figure 5.5).[54]
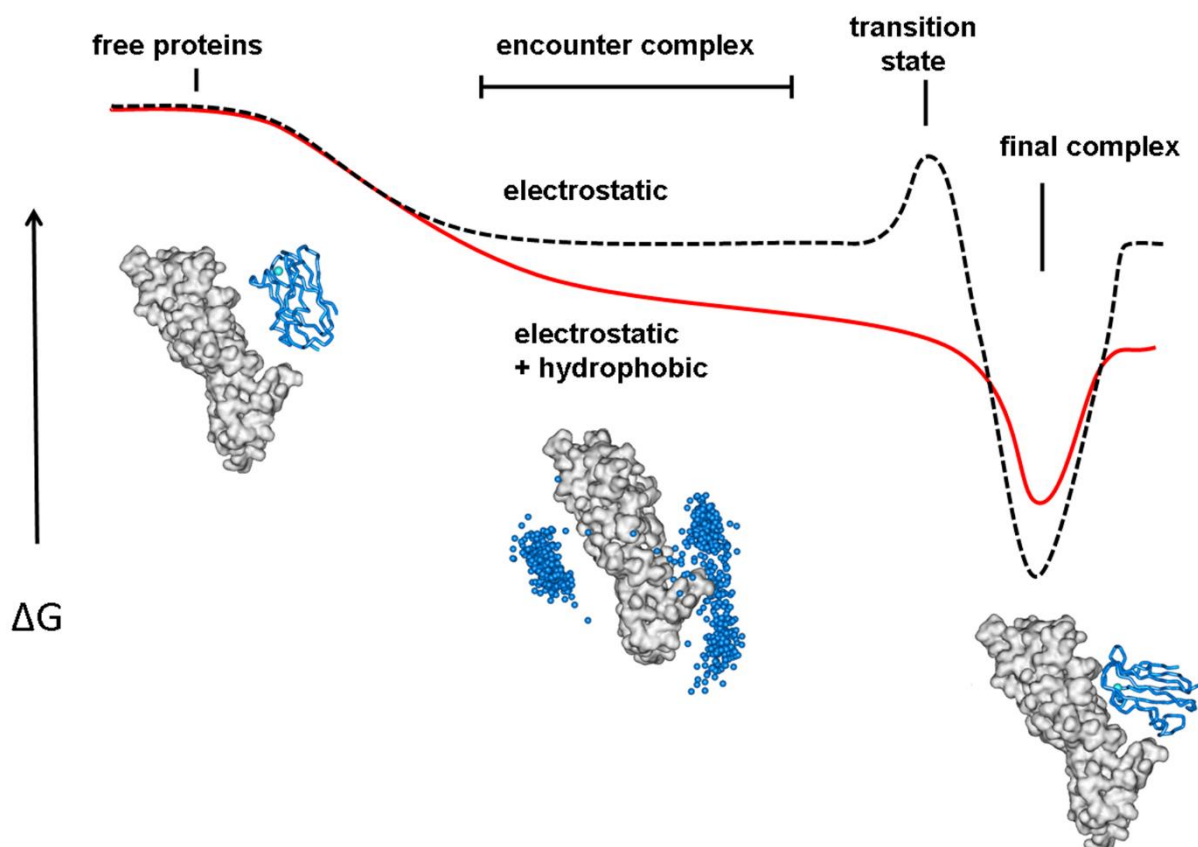
Figure 5.5 Free energy diagrams of association pathways with high (dashed black line) and no (solid red line) energy barrier for the transition state from the encounter to the final complex. Complexes based on electrostatic interactions are proposed to follow the dashed line because formation of the stereospecific complex requires the interaction interface to go from solvated to desolvated abruptly. If hydrophobic interactions gradually increase in the encounter state, the desolvation occurs more smoothly resulting in a smaller energetic barrier (red line), increasing the association rate. Reprinted with permission from [54]. Copyright 2010 American Chemical Society.

5.7.4 Cytochrome f and cytochrome c₆

Interestingly, in some organisms, in the photosynthetic redox chain Pc (see previous section) can be substituted by a cytochrome, cytochrome c_6 (Cytc6) to accept electrons from Cytf.

Cytc6 is a *c*-type cytochrome (with a covalent heme group) with a molecular mass of 10 kDa. The complex of Cytc6 and Cytf was reported to be dynamic and hydrophobic interactions are involved in the interaction.[58, 59] Intermolecular PRE were generated by tagging Cytf from *Nostoc* with MTSL on five positions surrounding the heme, and measured on isotope-labelled Cytc6. The resulting distance restraints were used for rigid-body docking calculations.[60] Large PREs were measured for each of the tagged Cytf molecules and the PRE effects were mostly found in the region on Cytc6 that also displays most of the chemical shift perturbations upon complex formation. This observation was explained by assuming that Cytc6 pre-orients upon its approach of the Cytf, due to long-range electrostatic interactions between the positive patch on Cytc6 and the wide-spread negative charges on Cytf. Moreover, it was not possible to explain the PRE data with a single orientation of Cytc6 with respect to Cytf. Instead, the PREs were well described by an ensemble of orientations in which Cytc6 samples mostly the hydrophobic patch close to the heme on Cytf, describing a form of "hydrophobic sliding" in which the desolvation of the hydrophobic patches increases gradually leading to the most stable orientations, very similar to what was described for Pc (Figure 5.5).[60]

5.7.5 Cytochrome c and adrenodoxin

Cytc is a small (12.5 kDa) heme protein found in mitochondria. Adx (11 kDa) was already mentioned in Section 5.5. Both proteins are electron transfer proteins, involved in shuttling electrons from one enzyme to the next. As they occur in different compartments of the cell, it is not expected that they form a physiologically relevant complex. Interestingly, however, electrons can be rapidly transferred from Adx to Cytc and this reaction has been used as a substitute for the natural reaction of Adx with cytochromes P450. In line with what was discussed on electron transfer complexes above (Section 5.7.3) it can be expected that between two small proteins, with the redox centres relatively close to the surface, electron transfer should occur in many orientations and a single stereospecific complex does not need to be formed for activity. This hypothesis was tested by establishing the degree of dynamics in the Cytc-Adx complex with the use of paramagnetic NMR. One study used the paramagnetic centres in Cytc (low-spin heme Fe^{3+}) and Adx (FeS cluster) to generate intermolecular PCSs and PREs, respectively on both the native complex and a cross-linked form that served as a model for a well-defined stereospecific complex.[61] In the cross-linked form, both intermolecular PCSs and PREs could be detected and also the chemical shift perturbations due to complex formation were significant. In the native complex, all these effects were eliminated due to averaging, indicative of high mobility. MTSL tagging of Cytc

caused PREs over a large area of the Adx surface, again indicating extensive mobility in the complex. Ultimate proof was given by attaching a CLaNP-5-Yb^{3+} tag on Cytc and measuring the RDCs on both Cytc and Adx in the native complex. The RDCs for Adx were insignificant, whereas for Cytc they were readily measured.[62] Thus, extensive mobility must be present in the complex, which was described as a pure encounter complex. This work demonstrates that indeed in the small redox proteins electron transfer can occur rapidly in the absence of a stereospecific complex, in line with Marcus theory.

5.7.6 Cytochrome P450cam and putidaredoxin

A combination of PCSs, RDCs and PREs was used in the recent determination of the solution structure of the complex of cytochrome P450cam (P450cam) and putidaredoxin (Pdx), which was subsequently confirmed by crystallography. It showed the power of paramagnetic NMR, as not only did it establish the orientations of the proteins in the complex in solution, the work also identified the presence of a lowly populated encounter complex.[63] P450cam is a soluble heme enzyme that catalyses the oxidation of camphor using molecular oxygen. It receives electrons from the small iron-sulphur protein Pdx. At the time, the structures of the free proteins were available but that of the complex was lacking. P450cam was tagged with CLaNP-7[64] at two sites and Pdx at one. Amide assignments for both proteins were available and extended.[65, 66] Intramolecular PCSs from CLaNP-7(Yb^{3+}) were used to establish the $\Delta\chi$ tensor orientations and intermolecular PCSs, RDCs (using Tm^{3+}) and PREs (Gd^{3+}) were obtained with CLaNP-7(Lu^{3+}) as the diamagnetic control. In this way, 446 restraints were obtained to dock Pdx as a rigid body on P450cam, resulting in an ensemble of structures with an average root mean square deviation of 1.3 Å for the heavy atoms from the mean. Independently, the crystal structure was determined by the same authors and by another group,[63, 67] showing Pdx in the same position as in the solution structure (Figure 5.6). All PCS and RDC restraints as well most of the PRE distances were satisfied very well by the calculated structure. However, the distances derived from the PRE of Pdx amide hydrogens from one of the tags on P450cam were much shorter that those calculated from the final structure (Figure 5.6). This observation suggested that Pdx assumes another state(s) in which it approaches that tag more closely than in the major, stereospecific complex. The PCS and RDC data were fitted well, suggesting that the additional state had a low population and was picked up by the PRE effects only because of their exquisite sensitivity for minor states. Further tagging with CLaNP-7 and MTSL on both Pdx and P450cam confirmed the presence of an encounter state with a population of only a few percent, in which Pdx samples the

surface of P450cam widely and in various orientations. Detailed calculations using the MaxOR approach (see Sections 4.3.2.2 and 4.3.4) were applied to an encounter complex for the first time, leading the identification of sites that were thought to represent productive encounters and others suggested to be futile encounters.[68]

The crystal structures of the complex had suggested that P450cam opens its substrate entry channel upon binding of Pdx. EPR and modelling studies also provided evidence for this.[69-71] To test this model in solution at ambient temperature, a CLaNP-7 tag loaded with Yb^{3+} was attached to one of the α-helices that constitute the cover of the entry channel. PCS were obtained for amide nuclei in P450cam, selectively labelled with ^{15}N-Leu in the absence and presence of Pdx, clearly showing that no changes occurred upon Pdx binding. The observed PCSs fitted the closed structure much better than the open structure. These results suggested that in solution P450cam does not open its substrate entry channel, or at most very slightly.[72]
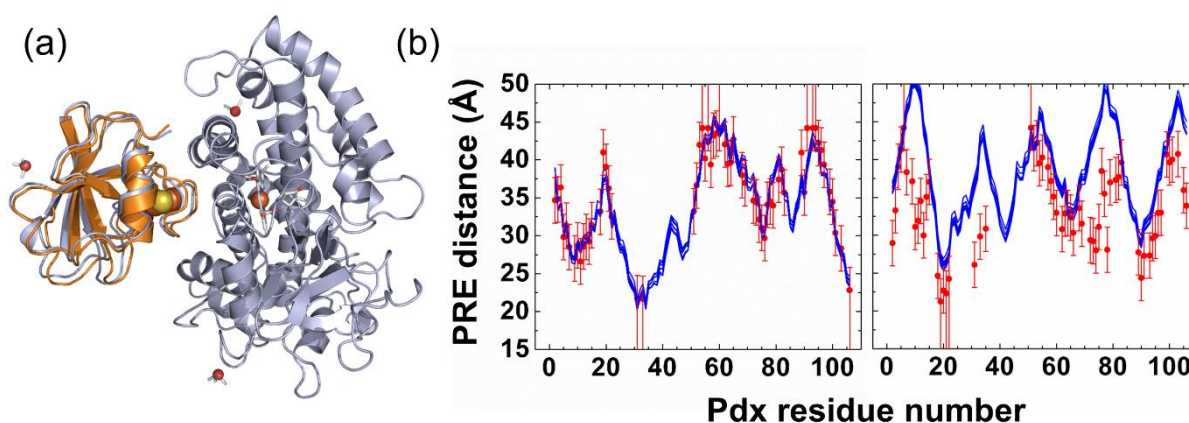

Figure 5.6 The complex of P450cam and Pdx. (a) The crystal structure (orange) and the solution structure closest to the mean (gray) of the oxidized Pdx–P450cam complex are shown in a ribbon representation with the P450cam structures aligned. The positions of Ln^{3+} ions, **Δχ** tensors and redox centers of Pdx and P450cam are depicted in sticks and spheres. (b) Violation analysis of PRE-derived distances plotted against Pdx residue numbers. Red circles and blue lines represent experimental and back-calculated effects for the 10 lowest-energy solutions, respectively. The left panel shows that the calculated structures fit the PREs obtained from probe A for Pdx nuclei. The right panel shows that Pdx nuclei experience larger PREs (shorted distances) from probe B than predicted by the 10 lowest-energy structures, presenting evidence for an encounter state. Reprinted from [63], Copyright (2013), with permission from Elsevier.

5.7.7 Ferredoxin, ferredoxin:thioredoxin reductase and thioredoxin

The study of the complex formed by ferredoxin (Fd), ferredoxin:thioredoxin reductase (FTR) and thioredoxin (Trx) is a good example of how X-ray crystallography and paramagnetic NMR can complement each other to obtain a more complete picture of how a protein complex is formed and functions. These three proteins form a ternary electron transfer complex, as part of the signalling pathway that links the light reactions of oxygenic photosynthesis with the carbon fixation reactions in the Calvin cycle. The structure of the ternary complex was visualized by X-ray crystallography showing that Fd and Trx bind on either side of the flat, elongated FTR surface, allowing them to bind simultaneously and form a redox chain (PDB ID 2PVO).[73] In the structure, FRT is cross-linked to a cysteine of Trx through a transient intermolecular disulphide bond. The formation of this intermediate complex was further investigated using paramagnetic NMR, exploiting the presence of a Fe_4S_4 cluster in FTR. Intermolecular PRE data for Trx were used as distance restraints for docking of Trx on FTR, obtaining a model of the intermediate that precedes the formation of the transient cross-link. The NMR-based model shows Trx in an orientation in the noncovalent complex that differs from that in the cross-linked crystal structure, suggesting it needs to rotate about 50° to proceed to the cross-linked complex.[74]

5.7.8 Cytochrome c peroxidase and cytochrome c

Cytochrome *c* peroxidase (CcP) is a mitochondrial heme enzyme from yeast that can reduce hydrogen peroxide. Cytochrome *c* (Cytc) acts as the electron donor. The CcP-Cytc complex is probably the best studied electron transfer complex. Paramagnetic NMR was important for our understanding of this complex, providing precious information about transient protein-protein complexes and encounter states. The crystal structure of the complex was already published in 1992.[75] The structure of the complex in solution was determined using solely PRE restraints, by tagging CcP at five different sites with MTSL and measuring the intermolecular relaxation changes of amide hydrogens in uniformly labelled Cytc.[37] The major form of the complex was similar to the one observed in the crystal structure. The work also demonstrated that a substantial fraction of the complex must be in an encounter state, because part of the PREs could not be explained by the major state of the complex. To visualize the encounter complex the number of tags was extended to 10 sites on CcP. In this way, it was shown that the encounter site is limited to the immediate environment of the stereospecific binding site. Structural ensembles from docking calculations based solely on

the electrostatic interactions between CcP and Cytc were in good agreement with the experimental PREs, assuming that the encounter state represented no less than 30% of the complex, with 70% being in the stereospecific complex observed in the crystal structure.[38] Using the ensemble obtained from these calculations, the encounter state could be visualized (Figure 5.7) and it was shown that Cytc samples just 15% of the CcP surface.[38] This can be explained by the charge distribution present on CcP and Cytc, which are highly dipolar, with strong negative and positive patches close to the respective binding sites. The fraction of encounter state could readily be changed by mutations in the binding site, in the range between 10% to 90% encounter state.[76] In the sections above, it was explained that for electron transfer complexes a stereospecific binding site is not a requirement, as long as the distance between the redox centres is short enough, e.g. in the complex of Cytc and Adx. In the CcP-Cytc complex, the stereospecific complex appears to be necessary to reduce the distance between the Cytc heme and the CcP Trp residue that acts as the first electron acceptor. The Trp and heme in CcP are deeply buried. The delicate balance between stereospecific complex and encounter state may be a consequence of competing interests. The encounter complex results from the strong electrostatic interactions that raise the association rate, the stereospecific complex is necessary for fast electron transfer. The complex cannot be too tight (K_D for Cytc is 5 μM[77]) because that would limit dissociation.

However, it is also possible that electron transfer occurs partly at another site that seems to bind Cytc, at least weakly under low ionic strength conditions. Recently, this binding site was localized using paramagnetic NMR methods. The high-affinity binding site on CcP first was blocked by cross-linking isotope-labelled CcP and unlabelled Cytc through an intermolecular disulphide bond.[78] In this way, it was possible to study the low-affinity interaction between the cross-linked complex and Cytc. PREs were obtained using Cytc tagged at three sites with an EDTA-based tag containing Mn^{2+}. The experiments revealed that the weak complex is composed of a dominant species, which is thought to be inactive in electron transfer since the hemes are too far apart (22 Å), and an ensemble of minor species, in which the hemes are close enough for electron transfer (< 16 Å). Therefore, it is possible that, under certain conditions, Cytc binding at the low affinity site contributes much to the electron transfer activity.[78]
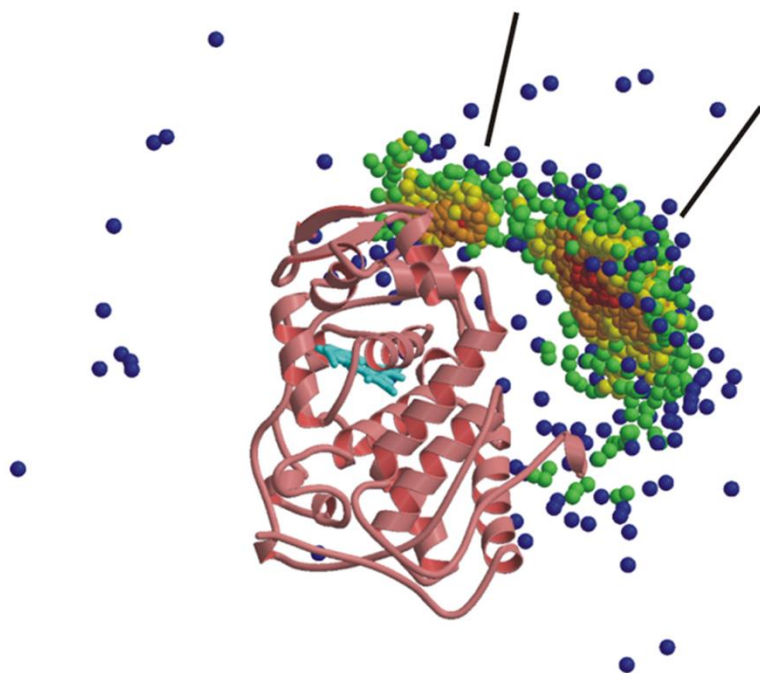
Figure 5.7 Simulated encounter complex of the CcP-Cytc complex based on MonteCarlo calculations of the electrostatic interactions. CcP is shown in pink ribbons and the heme in cyan sticks. The centres of mass of Cytc are shown as spheres, coloured to indicate the density of the distributions, decreasing from red to blue. The highest densities denote the most favorable electrostatic orientations. The lines indicate the two centers of high density. The model was shown to be in good agreement with extensive PRE data assuming that the encounter complex is populated for 30%. Reprinted with permission from [38]. Copyright 2010 American Chemical Society.

5.7.9 Enzyme I and the histidine-containing phosphocarrier protein

The histidine-containing phosphocarrier protein (HPr) and enzyme I are the two general components of the phosphotransferase system involved in transport of a variety of carbohydrates into the prokaryotic cell. PRE restraints were applied to gain insight into the mechanistic details of the interactions present in the transient encounter complex formed by the N-terminal domain of enzyme I (EIN) and HPr.[40, 79] The measurement of the sedimentation velocity of HPr, which is monomeric in solution, revealed that any higher-order species is populated less than 1-2% of the total soluble HPr.[79] Intermolecular PREs were collected on 1:1 solutions of ^{15}N-labelled HPr and HPr tagged with an EDTA-Mn^{2+} tag (or Ca^{2+} as diamagnetic ion) at three sites. PREs were observed for two of the labelled HPr, revealing ultra-weak self-association of HPr molecules in solution. The intermolecular PREs were abolished by adding EIN to form the EIN:HPr complex, which suggests that the EIN-

25

HPr interface is the same as the one used for HPr self-association. A similar effect was seen by increasing the salt concentration, evidence for an electrostatic component in the interaction, even though hydrophobic interactions are also present, as shown by the differences in the degree of reduction of the intermolecular PREs by the addition of salt for different regions. The quantitative analysis of the PRE data through the rigid body/torsion angle simulated annealing calculations suggested that an ensemble of self-association interactions can explain the observed PREs and that the K_D for the HPr-HPr self-association is $\geqslant$ 15 mM.[79]

HPr was investigated also for its interaction with EIN. PRE experiments were applied to study the encounter complexes formed by the two proteins in relation to the ionic strength.[80] EIN was ^{2}H and ^{15}N-labelled, while HPr was tagged with EDTA-Mn^{2+} in two positions. Observed PREs were divided in PREs caused by the stereospecific complex, which showed just a weak salt dependency, and PREs generated by the ensemble of nonspecific encounter complexes, which was much more affected by the salt concentration. In particular, the latter PREs decreased with an increasing salt concentration, indicating a shift in the equilibrium from encounter state to stereospecific complex. Other PRE experiments allowed further characterization of the encounter complex formed by HPr and EIN. Again, HPr was tagged with EDTA-Mn^{2+} at three different sites and PREs were measured on ^{15}N-labeled EIN.[40] The analysis of the energy landscape of the protein-protein interactions leading to the formation of the specific complex showed the presence of two classes of encounter complexes. It was suggested that the main function of one class is to guide HPr to the stereospecific site, in the first part of the productive complex formation, when the binding site is empty. The second class, instead, can form an ensemble of ternary complexes with the stereospecific complex and it is more populated when the binding site is occupied. For this reason, it was proposed that the function of this class is to "reload" the EIN with HPr to have a more efficient sugar transport.[40] It is estimated that ternary complex ensemble engages about the 1% of the enzyme I molecules *in vivo*, a percentage that could be higher due to intracellular crowding and compartmentalization.

5.7.10 NS2B-NS3 protease

Protein structure determination is often done for drug discovery (Chapter 9). This is the case for the complex formed by NS2B and NS3, two non-structural proteins of the dengue virus serotype 2. Many studies were conducted on segments of these two proteins linked together forming the NS2B-NS3 protease (NS2B-NS3pro). The development of an inhibitor for this

protein was problematic, and even though it is an established drug target, its exact structure was difficult to solve. It was not possible to solve the structure in solution with conventional NMR techniques because of low spectral dispersion, poor protein stability and line broadening caused by protein dynamics. The first crystal structure showed that the C-terminal segment of NS2B (NS2Bc) was far from the active site, in an open and inactive conformation,[81] but other structures obtained with NMR in solution showed that the complex is in a closed conformation, with NS2Bc near the substrate binding site both in presence and absence of inhibitors.[15, 82, 83] For example, one of these studies analysed the NS2B-NS3pro complex with bovine pancreatic trypsin inhibitor (BPTI), a complex of 35 kDa. In this study, NS3pro was tagged with the lanthanoid tag[44] $C2$-Tb^{3+} ($C2$-Y^{3+} as diamagnetic control) and PCS were measured of backbone amides in the complex with either partner being ^{15}N-labelled, demonstrating that the complex in solution and in presence of BPTI adopts a closed conformation. The quality of the model was confirmed by the good correlation between the experimental and back-calculated PCSs.[84] However, differences in peak intensities in the spectra indicated that NS2B-NS3pro experiences conformational exchange, which made it possible that an open conformation of NS2Bc of up to 10% would go undetected. On the other hand, it is also possible that the open conformation of the protease-BPTI complex could be a crystallization artefact.[84] Some studies have also been done on the unlinked NS2B-NS3pro construct. For example, the complex was tagged on three different positions with paramagnetic C1-and C2-lanthanoid tags[44,85] loaded with Tm^{3+} and Tb^{3+} and PCSs were measured.[15] The obtained model was consistent with the structure in the closed state detected in presence of inhibitors[44], even NS2Bc showed an increasing conformational exchange at the increasing of the pH and ionic strength. This did not result in the dissociation from NS3pro even at high concentration of salt.

5.8 Conclusions and perspective

Paramagnetic NMR offers a new toolbox for the study of protein-protein complexes in solution. Due to the strong spins of unpaired electrons, long-range interactions with nuclei are observed and the theory of PCS, RDC and PRE is well described, making it possible to extract accurate structural restraints from these observables. The advantage is that the effects are 'clean'. Usually, only a single paramagnetic centre is present in the sample, so, provided a suitable diamagnetic control sample is present, the paramagnetic effect can be isolated and readily attributed to the observed nucleus – electron(s) interaction. As compared to NOE, no

ambiguities of assignments and spin diffusion need to be considered and the absence of a paramagnetic effect is also very informative. On the other hand, a dense network as observed for NOEs will not be obtained for paramagnetic restraints, because only a single paramagnetic centre is used at a time. The longer distance range also precludes very detailed structural analysis. An exception is the immediate environment of the paramagnetic centre, however, in the case of protein-protein complex studies, the paramagnetic centre itself is not of interest.

The long-range nature makes it possible to study large systems, for example the PRE-based study of the complex of nitrite reductase and pseudoazurin of 152 kDa,[86] and the PCSs observed for the 90 kDa Lac repressor complex[87] and the 300 kDa transcarbomylase.[88] Provided that the structures of the free proteins are available and for at least one partner nuclei can be observed and assigned, paramagnetic restraints can be obtained relatively easily, using simple experiments and low protein concentrations. With ranges up to 70 Å, or perhaps even more, measurable with PCSs, very large systems can be investigated in principle.

So far, paramagnetic NMR has been used to determine protein-protein complexes of ground states based on rigid body docking. This is a relatively simple six-dimensional problem that can be solved with a limited number of restraints. Such an approach does not take into account conformational changes in the proteins upon complex formation and generally also does not consider side-chain orientations. Thus, such structures are no match yet for crystallography. To obtain high resolution structures, a very large set of restraints, for side-chain as well as backbone nuclei, is required. It can be questioned whether this is a goal worth striving for. Paramagnetic NMR seems to play its most important role in the study of ensembles and dynamics (see Chapter 4). PRE analysis has opened the avenue to experimental studies of the encounter complexes and with that to the process of protein complex formation. It has made us realize that the stereospecific complex is not the only representation of the complex. In many protein complexes, especially the weaker ones that are so important in processes such as signalling and electron transfer, the encounter state is an essential part of the complex. Paramagnetic NMR, in particular PCSs, can also be very useful to detect structural changes, i.e. to validate crystal structure models in solution. PCSs can readily distinguish between open and closed forms of proteins and protein complexes that may appear in crystals, but for which it is unknown whether they exist in solution. Finally, there may be future for paramagnetic relaxation dispersion studies. Relaxation dispersion NMR is well-established for measurement of μs – ms dynamics in proteins, based on exchange effects that cause line-broadening. The underlying chemical shift differences between the exchanging states are, however, difficult to interpret in structural terms. It has

been shown that combining relaxation dispersion with PCS data can potentially be very powerful but paramagnetic tag mobility is still a large problem, because it causes additional, undesired relaxation dispersion effects.[89, 90]

In conclusion, paramagnetic NMR has shown to have many applications for protein NMR and it is expected that its star will brighten in the future. PRE has been used widely, but the power of PCS and pRDC is still to be appreciated more by many in the NMR field.

References

1. G. Otting, in *Annu. Rev. Biophys.*, ed. D. C. Rees, K. A. Dill and J. R. Williamson, Annu. Rev. Biophys., Palo Alto, 2010, vol. 39, p. 387.

2. P. H. J. Keizers and M. Ubbink, in *Protein NMR Spectroscopy: Practical Techniques and Applications*, ed. L.-Y. L. a. G. Roberts, John Wiley & Sons, Ltd, Chichester,UK, 2011, p. 193.

3. P. M. Keizers and M. Ubbink, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**(1-2), 88.

4. J. Koehler and J. Meiler, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **59**(4), 360.

5. M. Ubbink, *FEBS Lett.*, 2009, **583**(7), 1060.

6. G. M. Clore and J. Iwahara, *Chem. Rev.*, 2009, **109**(9), 4108.

7. G. Schreiber, G. Haran and H. X. Zhou, *Chem. Rev.*, 2009, **109**(3), 839.

8. H. M. McConnell, *J. Chem. Phys.*, 1958, **28**(3), 430.

9. D. M. Yu, A. N. Volkov and C. Tang, *J. Am. Chem. Soc.*, 2009, **131**(47), 17291.

10. J. Iwahara and G. M. Clore, *Nature*, 2006, **440**(7088), 1227.

11. P. H. J. Keizers, B. Mersinli, W. Reinle, J. Donauer, Y. Hiruma, F. Hannemann, M. Overhand, R. Bernhardt and M. Ubbink, *Biochemistry*, 2010, **49**(32), 6846.

12. S. Arumugam, C. L. Hemme, N. Yoshida, K. Suzuki, H. Nagase, M. Bejanskii, B. Wu and S. R. Van Doren, *Biochemistry*, 1998, **37**(27), 9650.

13. H. G. Hocking, K. Zangger and T. Madl, *ChemPhysChem*, 2013, **14**(13), 3082.

14. R. M. Almeida, C. Geraldes, S. R. Pauleta and J. J. G. Moura, *Inorg. Chem.*, 2011, **50**(21), 10600.

15. L. de la Cruz, W. N. Chen, B. Graham and G. Otting, *Febs J.*, 2014, **281**(6), 1517.

16. G. Pintacuda, A. Y. Park, M. A. Keniry, N. E. Dixon and G. Otting, *J. Am. Chem. Soc.*, 2006, **128**(11), 3696.

17. H. van Ingen and A. Bonvin, *J. Magn. Reson.*, 2014, **241**, 103.

18. N. P. Cowieson, B. Kobe and J. L. Martin, *Curr. Opin. Struct. Biol.*, 2008, **18**(5), 617.

19. N. J. Anthis and G. M. Clore, *Quart. Rev. Biophys.*, 2015, **48**(1), 35.

20. I. M. A. Nooren and J. M. Thornton, *Embo J.*, 2003, **22**(14), 3486.

21. W. Leupin, G. Otting, H. Amacker and K. Wuthrich, *FEBS Lett.*, 1990, **263**(2), 313.

22. M. Nilges, *Proteins*, 1993, **17**(3), 297.

23. V. Gaponenko, A. S. Altieri, J. Li and R. A. Byrd, *J. Biomol. NMR*, 2002, **24**(2), 143.

24. S. Rumpel, S. Becker and M. Zweckstetter, *J. Biomol. NMR*, 2008, **40**(1), 1.

25. M. A. S. Hass and M. Ubbink, *Curr. Opin. Struct. Biol.*, 2014, **24**, 45.

26. M. Ubbink, M. Ejdeback, B. G. Karlsson and D. S. Bendall, *Structure*, 1998, **6**(3), 323.

27. T. Saio, K. Ogura, M. Yokochi, Y. Kobashigawa and F. Inagaki, *J. Biomol. NMR*, 2009, **44**(3), 157.

28. G. Bjorkoy, T. Lamark, A. Brech, H. Outzen, M. Perander, A. Overvatn, H. Stenmark and T. Johansen, *J. Cell Biol.*, 2005, **171**(4), 603.

29. T. Saio, M. Yokochi, H. Kumeta and F. Inagaki, *J. Biomol. NMR*, 2010, **46**(4), 271.

30. P. H. J. Keizers, J. F. Desreux, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2007, **129**(30), 9292.

31. P. H. J. Keizers, A. Saragliadis, Y. Hiruma, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**(44), 14802.

32. J. J. Muller, A. Lapko, G. Bourenkov, K. Ruckpaul and U. Heinemann, *J. Biol. Chem*, 2001, **276**(4), 2786.

33. A. Spaar, C. Dammer, R. R. Gabdoulline, R. C. Wade and V. Helms, *Biophys. J.*, 2006, **90**(6), 1913.

34. G. Adam and M. Delbrück, in *Structural Chemistry and Molecular Biology*, ed. N. D. A. Rich, W. H. Freeman and Co., San Francisco, 1968, p. 198.

35. G. McLendon, in *Long-Range Electron Transfer in Biology*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1991, p. 159.

36. Y. C. Kim, C. Tang, G. M. Clore and G. Hummer, *P. Natl. Acad Sci. USA*, 2008, **105**(35), 12855.

37. A. N. Volkov, J. A. R. Worrall, E. Holtzmann and M. Ubbink, *P. Natl. Acad Sci. USA*, 2006, **103**(50), 18945.

38. Q. Bashir, A. N. Volkov, G. M. Ullmann and M. Ubbink, *J. Am. Chem. Soc.*, 2010, **132**(1), 241.

39. C. Tang, J. Iwahara and G. M. Clore, *Nature*, 2006, **444**(7117), 383.

40. N. L. Fawzi, M. Doucleff, J. Y. Suh and G. M. Clore, *P. Natl. Acad Sci. USA*, 2010, **107**(4), 1379.

41. N. L. Fawzi, J. F. Ying, D. A. Torchia and G. M. Clore, *J. Am. Chem. Soc.*, 2010, **132**(29), 9948.

42. L. Zhao, G. Ji, X. Le, Z. Luo, C. Wang, M. Feng, L. Xu, Y. Zhang, W. B. Lau, B. Lau, Y. Yang, L. Lei, H. Yang, Y. Xuan, Y. Chen, X. Deng, T. Yi, S. Yao, X. Zhao, Y. Wei and S. Zhou, *Oncogene*, 2017, **36**(24), 3384.

43. K. D. Brewer, T. Bacaj, A. Cavalli, C. Camilloni, J. D. Swarbrick, J. Liu, A. Zhou, P. Zhou, N. Barlow, J. J. Xu, A. B. Seven, E. A. Prinslow, R. Voleti, D. Haussinger, A. Bonvin, D. R. Tomchick, M. Vendruscolo, B. Graham, T. C. Sudhof and J. Rizo, *Nat. Struct. Mol. Biol.*, 2015, **22**(7), 555.

44. L. de la Cruz, H. D. N. Thi, K. Ozawa, J. Shin, B. Graham, T. Huber and G. Otting, *J. Am. Chem. Soc.*, 2011, **133**(47), 19205.

45. R. A. Marcus and N. Sutin, *Biochim. Biophys. Acta*, 1985, **811**(3), 265.

46. C. C. Moser, J. M. Keske, K. Warncke, R. S. Farid and P. L. Dutton, *Nature*, 1992, **355**(6363), 796.

47. J. M. Guss, P. R. Harrowell, M. Murata, V. A. Norris and H. C. Freeman, *J. Mol. Biol.*, 1986, **192**(2), 361.

48. S. E. Martinez, D. Huang, A. Szczepaniak, W. A. Cramer and J. L. Smith, *Structure*, 1994, **2**(2), 95.

49. A. Kannt, S. Young and D. S. Bendall, *Biochim. Biophys. Acta, Bioenergetics*, 1996, **1277**(1), 115.

50. P. B. Crowley, G. Otting, B. G. Schlarb-Ridley, G. W. Canters and M. Ubbink, *J. Am. Chem. Soc.*, 2001, **123**(43), 10444.

51. I. Diaz-Moreno, A. Diaz-Quintana, M. A. De la Rosa and M. Ubbink, *J. Biol. Chem*, 2005, **280**(19), 18908.

52. R. Hulsker, M. V. Baranova, G. S. Bullerjahn and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**(6), 1985.

53. S. Scanu, J. Forster, M. G. Finiguerra, M. H. Shabestari, M. Huber and M. Ubbink, *ChemBioChem*, 2012, **13**(9), 1312.

54. S. Scanu, J. M. Foerster, G. M. Ullmann and M. Ubbink, *J. Am. Chem. Soc.*, 2013, **135**(20), 7681.

55. S. Scanu, J. M. Foerster, M. Timmer, G. M. Ullmann and M. Ubbink, *Biochemistry*, 2013, **52**(38), 6615.

56. C. J. Camacho, Z. Weng, S. Vajda and C. DeLisi, *Biophys. J.*, 1999, **76**(3), 1166.

57. C. J. Camacho, S. R. Kimura, C. DeLisi and S. Vajda, *Biophys. J.*, 2000, **78**(3), 1094.

58. T. Z. Grove, G. M. Ullmann and N. M. Kostic, *J. Inorg. Biochem.*, 2012, **106**(1), 143.

59. T. Z. Grove and N. M. Kostic, *J. Am. Chem. Soc.*, 2003, **125**(35), 10598.

60. I. Diaz-Moreno, R. Hulsker, P. Skubak, J. M. Foerster, D. Cavazzini, M. G. Finiguerra, A. Diaz-Quintana, B. Moreno-Beltran, G. L. Rossi, G. M. Ullmann, N. S.

Pannu, M. A. De la Rosa and M. Ubbink, *Biochim. Biophys. Acta, Bioenergetics*, 2014, **1837**(8), 1305.

61. X. F. Xu, W. G. Reinle, F. Hannemann, P. V. Konarev, D. I. Svergun, R. Bernhardt and M. Ubbink, *J. Am. Chem. Soc.*, 2008, **130**(20), 6395.

62. X. F. Xu, P. H. J. Keizers, W. Reinle, F. Hannemann, R. Bernhardt and M. Ubbink, *J. Biomol. NMR*, 2009, **43**(4), 247.

63. Y. Hiruma, M. A. S. Hass, Y. Kikui, W. M. Liu, B. Olmez, S. P. Skinner, A. Blok, A. Kloosterman, H. Koteishi, F. Lohr, H. Schwalbe, M. Nojiri and M. Ubbink, *J. Mol. Biol.*, 2013, **425**(22), 4353.

64. W. M. Liu, P. H. J. Keizers, M. A. S. Hass, A. Blok, M. Tirnmer, A. J. C. Sarris, M. Overhand and M. Ubbink, *J. Am. Chem. Soc.*, 2012, **134**(41), 17306.

65. T. A. Lyons, G. Ratnaswamy and T. C. Pochapsky, *Protein Sci.*, 1996, **5**(4), 627.

66. B. OuYang, S. S. Pochapsky, M. Dang and T. C. Pochapsky, *Structure*, 2008, **16**(6), 916.

67. S. Tripathi, H. Y. Li and T. L. Poulos, *Science*, 2013, **340**(6137), 1227.

68. W. Andralojc, Y. Hiruma, W. M. Liu, E. Ravera, M. Nojiri, G. Parigi, C. Luchinat and M. Ubbink, *P. Natl. Acad Sci. USA*, 2017, **114**(10), E1840.

69. W. K. Myers, Y. T. Lee, R. D. Britt and D. B. Goodin, *J. Am. Chem. Soc.*, 2013, **135**(32), 11732.

70. S. H. Liou, M. Mahomed, Y. T. Lee and D. B. Goodin, *J. Am. Chem. Soc.*, 2016, **138**(32), 10163.

71. S.-H. Liou, W. K. Myers, J. D. Oswald, R. D. Britt and D. B. Goodin, *Biochemistry*, 2017, **56**(33), 4371.

72. S. P. Skinner, W. M. Liu, Y. Hiruma, M. Timmer, A. Blok, M. A. S. Hass and M. Ubbink, *P. Natl. Acad Sci. USA*, 2015, **112**(29), 9022.

73. S. D. Dai, R. Friemann, D. A. Glauser, F. Bourquin, W. Manieri, P. Schurmann and H. Eklund, *Nature*, 2007, **448**(7149), 92.

74. X. F. Xu, P. Schurmann, J. S. Chung, M. A. S. Hass, S. K. Kim, M. Hirasawa, J. N. Tripathy, D. B. Knaff and M. Ubbink, *J. Am. Chem. Soc.*, 2009, **131**(48), 17576.

75. H. Pelletier and J. Kraut, *Science*, 1992, **258**(5089), 1748.

76. A. N. Volkov, Q. Bashir, J. A. R. Worrall, G. M. Ullmann and M. Ubbink, *J. Am. Chem. Soc.*, 2010, **132**(33), 11487.

77. J. A. R. Worrall, U. Kolczak, G. W. Canters and M. Ubbink, *Biochemistry*, 2001, **40**(24), 7069.

78. K. Van de Water, Y. G. J. Sterckx and A. N. Volkov, *Nat. Commun.*, 2015, **6**, 7073.

79. C. Tang, R. Ghirlando and G. M. Clore, *J. Am. Chem. Soc.*, 2008, **130**(12), 4048.

80. J. Y. Suh, C. Tang and G. M. Clore, *J. Am. Chem. Soc.*, 2007, **129**(43), 12954.

81. P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan and U. Hommel, *Nat. Struct. Mol. Biol.*, 2006, **13**(4), 372.

82. Y. M. Kim, S. Gayen, C. B. Kang, J. Joy, Q. W. Huang, A. S. Chen, J. L. K. Wee, M. J. Y. Ang, H. A. Lim, A. W. Hung, R. Li, C. G. Noble, L. T. Lee, A. Yip, Q. Y. Wang, C. S. B. Chia, J. Hill, P. Y. Shi and T. H. Keller, *J. Biol. Chem*, 2013, **288**(18), 12891.

83. C. G. Noble, C. C. Seh, A. T. Chao and P. Y. Shi, *J. Virol.*, 2012, **86**(1), 438.

84. W. N. Chen, K. V. Loscha, C. Nitsche, B. Graham and G. Otting, *FEBS Lett.*, 2014, **588**(14), 2206.

85. B. Graham, C. T. Loh, J. D. Swarbrick, P. Ung, J. Shin, H. Yagi, X. Jia, S. Chhabra, N. Barlow, G. Pintacuda, T. Huber and G. Otting, *Bioconjugate Chem.*, 2011, **22**(10), 2118.

86. M. D. Vlasie, R. Fernandez-Busnadiego, M. Prudencio and M. Ubbink, *J. Mol. Biol.*, 2008, **375**(5), 1405.

87. F. Peters, M. Maestre-Martinez, A. Leonov, L. Kovacic, S. Becker, R. Boelens and C. Griesinger, *J. Biomol. NMR*, 2011, **51**(3), 329.

88. A. Velyvis, H. K. Schachman and L. E. Kay, *J. Am. Chem. Soc.*, 2009, **131**(45), 16534.

89. M. A. S. Hass, W. M. Liu, R. V. Agafonov, R. Otten, L. A. Phung, J. T. Schilder, D. Kern and M. Ubbink, *J. Biomol. NMR*, 2015, **61**(2), 123.

90. M. A. S. Hass, P. H. J. Keizers, A. Blok, Y. Hiruma and M. Ubbink, *J. Am. Chem. Soc.*, 2010, **132**(29), 9952.

91. Y. Xue, M. Okvist, O. Hansson and S. Young, *Protein Sci*, 1998, **7**(10), 2099.

92. C. R. Babu, B. F. Volkman and G. S. Bullerjahn, *Biochemistry*, 1999, **38**(16), 4988.

93. U. Badsberg, A. M. Jorgensen, H. Gesmar, J. J. Led, J. M. Hammerstad, L. L. Jespersen and J. Ulstrup, *Biochemistry*, 1996, **35**(22), 7021.