# Transient complexes of haem proteins

Volkov, O.M.

# *Concluding Remarks*

## *Tools to Study Transient Protein-Protein Complexes*

### *NMR chemical shift perturbation analysis*

Interaction between two proteins is accompanied by changes in the chemical environment of the nuclei at the interface. If one of the proteins is $^{15}N$ labelled, these changes can be followed in its 2D [$^{1}H$, $^{15}N$] HSQC spectrum, which allows to map the interface of the protein complex[161] (Chapter II). This NMR approach has been used extensively in our group (ref. 4,5 and references therein; ref. 10,238,239). The size of the binding shifts has been noticed to vary widely among different complexes, and it was suggested that such variation correlates with the protein dynamics (ref. 5 and references therein).

In this study, we confirm this earlier suggestion by PRE analysis that allows to observe directly the dynamic state of the complex. We propose that the size of the binding-induced NMR chemical shift perturbations ($\Delta\delta$) can be used as a reliable diagnostic tool for the dynamics within protein-protein complexes (Chapters V-VI). Qualitative interpretation of the magnitude of the binding shifts allows to discriminate between highly dynamic (small $\Delta\delta$) and mostly single-orientation (large $\Delta\delta$) protein complexes. We show that the former group includes Cyt *c* – Cyt $b_5$ (Chapter II), R13A Cyt *c* – CcP (Chapter V), and R13K Cyt *c* – CcP (Chapter V) complexes, while the latter category contains complexes of wt or T12A Cyt *c* with CcP (Chapter VI).

### *Paramagnetic Relaxation Enhancement NMR spectroscopy*

Considered a nuisance for a long time, paramagnetism has pushed the boundaries of biomolecular NMR in the past decade. The paramagnetic NMR-based approaches for structure determination of individual proteins and protein complexes in solution have rapidly become standard[163,164]. In this work, we have successfully used intermolecular PREs – caused by a spin-label attached to the surface of CcP and detected on the Cyt *c* backbone amides – to solve the structures of several Cyt *c* – CcP complexes in solution (Chapters III,

V-VI). The time- and cost-efficiency of this method make it an attractive approach for rapid structure determination of other protein complexes in solution.

Recent, elegant work by Clore and co-workers[173,199] has extended the application of PRE NMR spectroscopy to the study of transient intermediates in biomolecular interactions. Employed in our research, the PRE analysis has allowed to directly observe the dynamic encounter state of the Cyt *c* – CcP complex and provided a semi-quantitative estimate for the fraction of time spent by the proteins in this state (Chapters IV-VI). In this way, we could assess the effects caused by mutations of interfacial Cyt *c* residues on dynamics within Cyt *c* – CcP complex (Chapters V-VI).

### *Isothermal Titration Calorimetry*

ITC, an established bioanalytical technique in its own right, has proven to be a valuable partner in the tandem with NMR spectroscopy for the investigation of Cyt *c* – CcP complex (Chapters V-VI). First of all, ITC has been used to ascertain the effect of mutations on the binding affinity, which was shown to decrease for R13A and R13K Cyt *c* (Chapter V) and, unexpectedly, increase for T12A variant (Chapter VI). Furthermore, separate analysis of $\Delta H$ and $\Delta S$ contributions to the mutation-induced changes in the binding energy has alleviated the interpretation of the differences observed in the NMR spectra. Substantial increase of $\Delta S$ term with concomitant decrease in $\Delta H$ for the binding of R13A or R13K Cyt *c* relative to that of the wt protein has suggested an increased mobility within the variant Cyt *c* – CcP complexes, which has been further confirmed by NMR spectroscopy (Chapter V). By the same token, the increase of $\Delta H$ with decrease in $\Delta S$ for the binding of T12A Cyt *c* has suggested that additional – most likely solvent-mediated – intermolecular contacts stabilize this complex, again consistent with the complementary NMR analysis (Chapter VI).

## *Macromolecular Recognition in Transient Protein Complexes*

In the protein-protein interactions, several factors determine the choice of a partner (specificity) and the strength of the interaction (affinity). Although the determinants of the

affinity and specificity of transient protein complexes remain poorly understood[4], it is believed that electrostatic and hydrophobic forces dominate the process of protein recognition.

Electrostatic attraction between oppositely-charged molecules accelerates bimolecular association rates[25,27,240] and often pre-orients the reacting proteins[241]. In addition to enhancing the binding kinetics, Coulombic attraction constitutes the first step in the recognition between partner proteins and frequently determines the overall specificity of the interaction. However, as illustrated by the disagreement between the theoretical models for Cyt *c* – Cyt $b_5$[89,99] or Cyt *c* – CcP[26,108] with the experimental data (Chapters II-III), consideration of intermolecular electrostatic forces alone is often not enough for the correct prediction of the binding geometry in transient protein-protein complexes. Therefore, additional factors must contribute to the specificity of the protein association.

The hydrophobic effect, one of the driving forces in protein-protein interactions[217], is associated with binding-induced water release from the interface, which provides most of the binding affinity in a protein complex[236]. In addition to that, desolvation provides an important contribution to the specificity of the protein association[23,24]. Due to the desolvation penalty of the charged groups, electrostatically favourable configurations should not be necessarily so thermodynamically[25]. Indeed, it has been shown that for a number of protein complexes the minimum of the electrostatic potential does not coincide with the minimum in the total free energy of the complex[23]. When the contribution of desolvation is taken into account, the simulations produce results that are close to those observed experimentally[23,24,27].

Study of the conserved R13K mutation of Cyt *c* on its binding to CcP (Chapter V) suggests that intermolecular electrostatic and hydrophobic interactions are not sufficient for a specific complex formation. It appears that while electrostatics pre-orient and guide protein molecules to form an encounter complex and hydrophobics provide the affinity and further enhance the specificity of the interaction, particular intermolecular contacts further stabilize the complex and fine-tune the specificity. As exemplified by the mutational analysis of Cyt *c* – CcP complex, affinity and specificity of transient interactions – though clearly connected – are not linearly related: it is possible to greatly reduce the specificity without substantially decreasing the affinity (Chapter V) and, vice versa, largely increase the affinity without affecting the specificity (Chapter VI).

## *Encounter State of Transient Protein Complexes*

The concept of the encounter state is central to the understanding of transient protein interactions (Chapter I). In the case of Cyt *c* – CcP complex, extensively studied in this work, it helps to resolve a seeming contradiction: while some reports propose only one protein-protein orientation in this complex[107,127,128], other data suggest that it exists in multiple forms[26,120-122,124,125] or is even unstructured[123]. Furthermore, the two lines of thought appear to be mutually exclusive. In this work we show that, in fact, both are right – for most of the lifetime of the complex the proteins are in a single relative orientation (Chapters III-IV), with the rest of the time spent in the dynamic encounter state (Chapter IV). As it appears now, previous reports addressed only one aspect of Cyt *c* – CcP complex formation, either the dominant form or the encounter state, but not both at the same time; hence, the long-time controversy.

A persistent confusion of the terms used to describe the encounter state has greatly contributed to the controversy around this single-orientation, yet dynamic complex. For a long time, various values for the lifetime of a complex obtained by FRET (5 ns)[122], used in Brownian dynamics simulations (100 ns)[26], or determined by NMR (1 – 4 ms)[126], have been hard to reconcile[87]. Here we show that – using the definitions given in Chapter I (pp. 10-11) – the seemingly disparate time values reported for Cyt *c* – CcP, together with those determined in this work (Chapters IV-V), can be explained within the concept of the encounter state.

Using our terminology (for the definition of the words in italic see pp. 10-11), the FRET rates provide the lifetime of a protein-protein *collision* (5 ns), the values used in the simulations (100 ns) correspond to the lifetime of an *encounter complex*, those determined by NMR (1 – 4 ms) are the lifetime of the *protein complex*, and, finally, the values estimated in this work ( < 0.3 – 1.2 ms for wt Cyt *c* – CcP, Chapter IV, and *ca.* 0.75 – 3 ms for R13A Cyt *c* – CcP, Chapter V) represent the lifetime of the *encounter state*. The hitherto irreconcilable values reflect a mere linguistic confusion, rather than conceptual incompatibility, of the terms referring to different stages in the life of the Cyt *c* – CcP complex.

Until now the encounter state of Cyt *c* – CcP complex could only be visualized by theoretical modelling[26,27]. A classical Brownian dynamics study showed that there is a broad negatively-charged area on the surface of CcP available for the interaction with Cyt *c*[26]. For a long time, several local minima of the electrostatic energy observed in the simulations were thought to represent separate protein complexes likely co-existing in solution (see ref. 87,106 and references therein). Using PRE NMR approach, we could directly observe both the dominant, most favourable form of the Cyt *c* – CcP complex and its encounter state (Chapters III, IV). As the conformational space of the encounter state is similar to the one revealed by the Brownian dynamics, we propose that the local minima of the electrostatic energy observed in the simulations should not be seen as separate protein complexes, but rather a path taken by the molecules to arrive to the most favourable orientation.

## *Implications for Protein Electron Transfer*

Most of the current knowledge of the protein ET comes from the study of paradigmatic Cyt *c* complexes with Cyt $b_5$[28,90,91] and CcP[28,87,90,105,106]. The present report, concerned with structural aspects of these protein complexes, does not address the ET event itself. Nevertheless, the results obtained in this study could aid understanding of the intermolecular ET process in these systems.

Our results suggest that the interaction between Cyt $b_5$ and Cyt *c* is highly dynamic (Chapter II), which is consistent with the findings of a very recent study that supported a dynamic docking paradigm for this complex[145]. The paradigm, initially proposed for Cyt $b_5$ – myoglobin[11], suggests that a multitude of weakly bound conformations of the docked complex contribute to the overall binding of Cyt *c* to Cyt $b_5$. However, only few of these conformations are ET active, and they need not be the most favourable for binding. If one assumes that only the conformations with the two proteins in close haem-to-haem contact are ET favourable[145], the docked structures with the "head-on" orientation (Figure 2.9 B, p. 35) might represent the ET productive complexes.

As we show in this work, the dominant form of the Cyt *c* – CcP complex in solution is the same as the protein-protein orientation observed in the crystal structure

(Chapter III). If the dominant form of the complex is also the most ET reactive, as indeed was suggested by Poulos and co-workers[118], this could explain similar ET rates observed in the crystal[242] and in solution[243]. However, the forward ET between Cyt *c* and CcP is likely to be conformationally gated, implying that the most favourable binding geometry is not necessarily the most ET active[242,243]. In fact, there might be a multitude of reactive ET configurations as suggested by the crystallographic study of variant Cyt *c* – CcP complexes[198]. In this case the ET may be rationalized by the protein-protein orientations, some of which could be ET active, sampled in the dynamic encounter state of the complex. The conformational space of the encounter state delineated in this work (Chapter IV) could, therefore, represent the extent of conformational fluctuations in Cyt *c* – CcP complex that modulate the ET in solution[243].