Ligand migration and escape pathways in haem proteins

D. Lavalette1, C. Tétreau and L. Mouawad
Institut Curie, Centre Universitaire, 91405 Orsay, France

Abstract
Biophysical techniques developed during the last three decades have provided an increasingly detailed description of the internal processes associated with ligand capture and release by haem proteins. Myoglobin has long been the paradigm for these studies. More recently, cytochrome P450cam (the camphor-metabolizing cytochrome P450 from Pseudomonas putida) has also received considerable interest. In spite of sharing the same prosthetic group, the Fe(II)-haem, these proteins are structurally unrelated and they perform different functions. Recent works show that both proteins exhibit a common feature: a series of permanent or fluctuating, mostly hydrophobic, cavities of the protein matrix are providing transient docking sites as well as migration, escape and possibly entry pathways for the ligand. Remarkably, these systems of cavities connect the distal and the proximal regions of the haem, a disposition that may contribute to ligand capture enhancement.

Introduction
The elementary event of binding and releasing a small ligand is essential in a variety of protein functions such as transport (e.g. oxygen transport by haemoglobin), chemical transformation (e.g. enzymatic function of cytochromes P450) as well as signalling and molecular recognition processes. Yet, more than three decades have been necessary to achieve a reasonable description of the main processes accompanying this basic reaction. The specific examples mentioned above are two haemoproteins that played a unique role in these investigations thanks to the presence of a convenient spectroscopic marker, the Fe(II)-haem, the absorption spectrum of which changes when a ligand binds or dissociates.

The basic issues one would obviously like to address are very simple indeed: where and how does a ligand enter or leave the protein? How does it migrate towards the haem in the binding process or away from it in the dissociation process? How is the ligand–haem bond established? Whereas only quantum chemistry can answer the last question, the others are relevant to protein dynamics and amenable to experiment and modelling.

For the ‘general’ protein, in spite of the conceptual simplicity of the reaction scheme, some difficulties might have been anticipated from the beginning. For those who wish to study the forward (or ‘on’) reaction, the main problem is that mixing reactants is too slow a technique to permit a detailed investigation of the subsequent steps. The difficulty is no less for those who prefer to study the backward (or ‘off’) reaction: spontaneous dissociation is a rare stochastic event occurring incoherently. In fact, both approaches are plagued by the absence of any particular signal accompanying ligand entry, migration and escape. In other words, the processes one wishes to study are simply invisible in most proteins. In haemoproteins, they become, although indirectly, detectable as soon as the ligand binds to the haem or dissociates.

Materials and methods: a brief history
Photodissociation of the ligand–protein complex solves a lot of difficulties [1]. The absorption of one single photon breaks the haem–ligand bond. Using a flash, or more adequately nowadays, a laser pulse of short duration (picosecond to nanosecond), the dissociation of a large ensemble of protein–ligand complexes is achieved almost synchronously and gives rise to a large spectroscopic signal (absorbance change). It becomes possible to monitor the re-binding of the ligand from within the protein (‘geminate’ re-binding). In these experiments, carbon monoxide (CO) is generally used as a ligand for convenience.

Unfortunately, ligand and protein motions are generally so fast at ambient temperature that most important details escape detection. In general, the protein must be cooled down to very low temperature in order to stop all motions and, then, to watch the sequence of their reappearance by recording re-binding kinetics at successively higher temperatures. Of course this can be achieved only by embedding the protein in a cryoprotecting hydro-organic solvent (e.g. glycerol/water mixtures).

Reaction rates are predominantly determined by the structure of the proteins. Since these large polymers permanently undergo structural fluctuations, a protein sample is a heterogeneous ensemble in which an infinite number of ‘statistical’ substrates react at different rates. Therefore, instead of being
characterized by one unique, well-defined, rate parameter, such a system must be described by a distribution of reaction rates. At high temperature in a fluid solvent such as water, internal fluctuations are much faster than any other motions. As a consequence, differences are averaged out in the recording procedure and elementary reaction steps are apparently single exponentials. This effect, called kinetic averaging [1,2], is comparable with the ‘spectral narrowing’ known in NMR.

The apparent rate parameter is the ensemble average, ⟨k⟩, of the rate parameter statistical distribution. The gain in simplicity is paid by information loss about individual protein substrates.

Upon cooling, protein fluctuations eventually come to a stop. The frozen protein ensemble now displays the whole heterogeneity of reaction parameters, and the re-binding kinetics is given by

\[ N(t) = \int P(k)e^{-kt}dk \]  

(1)

Generally such kinetics extend over several orders of magnitude in time and the width of the rate distributions is considerably blurring important kinetic features.

The useful information in eqn (1) is P(k), the statistical distribution of rate parameters in the protein ensemble. It can be recovered by numerical inversion of the Laplace transform eqn (1) that describes the observed kinetic trace. This problem was solved by using the Maximum Entropy Method, based on Information Theory principles [2,3]. Because they are connected to each other by eqn (1), the kinetics N(t) and the ‘rate spectrum’ P(k) are entirely equivalent descriptions, in the same way as its frequency spectrum describes a complicated electric signal. The difference between Fourier and Laplace transforms is that the latter is based on the exponential rate parameter (k) as a variable rather than on frequency. Far from being an additional complication, the description of complex (heterochromatic) kinetics in terms of their rate spectrum P(k) is very powerful because details that are hardly visible in the kinetic trace give rise to bands or groups of bands that can be further analysed (Figure 1).

Once P(k) is known, it is possible to label and to quantify the individual processes present in the heterogeneous, low temperature kinetics. The problem is to interpret each process in terms of structural or dynamical events. Kinetics alone are unable to provide a unique answer unless some perturbation is applied to the system. Site-directed mutagenesis can be, and has been, used but it is a costly and lengthy procedure [6]. Long ago, crystallographers discovered that proteins are able to bind heavy xenon atoms inside a small number of matrix cavities [7]. Kinetic experiments in which ligand (e.g. CO) and xenon (∼0–20 atm; 1 atm = 101.325 kPa) compete for the cytochrome P450cam spectrum into four bands was firmly established by fitting globally the rate spectra over a range of temperatures where the amplitude of individual processes varied continuously [4,5]. The kinetic processes shown are all geminate recombinations, i.e. correspond to re-binding of the ligand that has remained within the protein after photodissociation. Such processes are independent of the external ligand concentration. One exception is process S in Mb at 250 K. Its rate depends on the [CO]. Process S is the bimolecular re-binding of ligands present in the solvent with proteins whose dissociated ligand has escaped into the solvent instead of re-binding. S processes are never observed below approx. 200 K because protein surface fluctuations are so much damped that geminate recombination occurs before the ligand can escape. Note that the area of rate spectra is normalized to unity.

Figure 2 | A simplified view of CO docking sites (full circles) and protein cavities in Mb (upper panel) and cytochrome P450cam (lower panel)

Upper panel: The cavities were identified as xenon-binding sites by crystallography. CO docking sites D1 and P1 as well as distal and proximal escape pathways (wavy arrows \( E_d \) and \( E_p \)) have been determined experimentally. Lower panel: The four CO docking sites were inferred from kinetic experiments and competition with xenon. Their location, their correspondence to protein cavities and the escape pathways have been suggested by MD simulations. The positions of the camphor substrate before and after relaxation are indicated by the pentagonal figures. In both schemes, wide and thin arrows give the order of magnitude of transition rates.

Results and discussion

Figure 2 summarizes the dominant features of CO binding and escape in Mb [sperm whale (\textit{Physeter macrocephalus}) myoglobin] and cytochrome P450cam (the camphor-metabolizing cytochrome P450 from \textit{Pseudomonas putida}) established in the last few years using various combinations of the experimental approaches outlined above. To compare proteins, it will be convenient to refer to Figure 1 that displays the rate spectra at a temperature where all geminate processes are seen simultaneously. This occurs at approx. 250 K in Mb and at approx. 200 K in cytochrome P450cam.

Myoglobin at 250 K

Immediately after photodissociation, CO is found in a primary distal docking site, D1, slightly off axis and above the haem plane. From this site, CO may either re-bind directly (fast geminate process \( G^1 \)) or escape the protein as indicated by the subsequent, slower, bimolecular process \( S \) corresponding to ligands originating from the outside. As a third possibility, the ligand may start migrating through a series of xenon-binding cavities that were identified long ago in the crystal [7].

The ligand then reaches the proximal docking site, P1, where it has been located in the crystal as a trapped intermediate [8–12]. Kinetic experiments in the presence of xenon indicate that once in P1, the ligand may again either escape or migrate back to the primary site D1, giving rise to the slower delayed geminate re-binding \( G^1 \) [13].

Cytochrome P450cam at 200 K

The initial kinetics are somewhat more complicated due to the presence of the camphor substrate. Comparison of the crystal structures of liganded and unliganded cytochrome P450cam suggests that, when the ligand is photodissociated, camphor tends to relax some steric constraints by moving slightly towards the iron. There are therefore two distal re-binding paths respectively before (\( G^1 \)) and after (\( G^1_r \)) camphor relaxation (Figure 1) and two distal docking sites, D1 and D2, in close vicinity (Figure 2) [4]. The subsequent delayed geminate re-binding processes are due again to ligand migration. Experiments with xenon suggest two additional docking sites in series. In the most remote one, the ligand competes with xenon for occupancy. Xenon sites have been documented [14] and MD (Molecular Dynamics) was used to simulate ligand migration and to delineate protein cavities (L. Mouawad, C. Tétéreau, F. Abdel-Azeim, D. Perahia and D. Lavalette, unpublished work). The statistics of 47, 1-ns-long trajectories of cytochrome P450cam (at temperature \( T \) between 200 and 320 K) showed many cavities, most of them being transient or fluctuating. On the average, two of them are consistently found to coincide with the locations where ligand occupancy is high. In particular, one encompasses the lateral and proximal sites, which were found to be strongly populated by the ligand (docking sites L1 and P1). Experimental data and MD simulations are consistent with a high-affinity xenon site in the proximal cavity. In the present scheme, one long and one short migration paths give rise to a faster and a slower migration-delayed geminate re-binding process \( G_{Mf} \) and \( G_{Ms} \) respectively (Figure 1).

In cytochrome P450cam, escape is observed experimentally only at higher temperature but the amplitude of the geminate processes is then too small for quantification. MD simulations at \( T \geq 300 \) K indicated, however, that three escape pathways are possible.

Differences and common features

Structurally, myoglobin and cytochrome P450cam have nothing in common but the haem. Yet, Figure 2 shows that both proteins are equipped with a series of internal cavities with similar role. Some of them coincide with the distal docking site(s) of the ligand, another one with a highly populated proximal docking site. Intermediate cavities connect the distal

©2006 Biochemical Society
and the proximal space of the haem. They may be populated or not, depending on the protein. The ligand moves around from one cavity to another in both proteins. Return from the proximal site gives rise to delayed geminate re-binding. In Mb, the alternative proximal escape pathway was discovered experimentally [13]. In cytochrome P450cam, it is suggested by MD simulations (L. Mouawad, C. Tétreau, F. Abdel-Azeim, D. Perahia and D. Lavalette, unpublished work).

Because laser flash photolysis and MD are dealing with ligand dissociation only, we have no direct information about ligand entry paths and rates. However, to a first approximation, protein cavities, structural fluctuations and potential docking sites are always available, ignoring whether the ligand is on its way towards the haem or leaving the protein. If the protein matrix offers a way out, it is likely to provide a way in as well (although entry and escape rates may be quite different). What may be important, as found in the Mb case, is that the escape efficiency from the proximal site P1 is at most $N_{esc} \approx 0.2$ [13]. This means that, with a high probability, a ligand entering the proximal pathway would become trapped by the system of cavities and rapidly directed to the distal side of the haem where it would be available for binding. If it were not so, ligand collisions occurring in the proximal protein ‘space’ would remain unproductive.

In conclusion, whatever the vast differences in details between Mb and cytochrome P450cam and in spite of being disposed differently with respect to the haem symmetry axes, the arrangement of cavities and of ligand docking sites is strikingly similar in both proteins and is likely to enhance the efficiency of ligand capture. At present, we only have two proteins to compare. Future investigations of other haemoproteins may tell us whether this feature has occurred by pure chance or whether it might have resulted from some sort of ‘convergent’ functional evolution.

References


Received 23 June 2006

©2006 Biochemical Society