Approaches to single molecule studies of metalloprotein electron transfer using scanning probe-based techniques

Martin Elliott\textsuperscript{1} and D. Dafydd Jones\textsuperscript{2}

\textsuperscript{1}School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK
\textsuperscript{2}School of Biosciences, Cardiff University, Cardiff CF10 3AT, UK

Abstract

The single–molecule properties of metalloproteins have provided an intensely active research area in recent years. This brief review covers some of the techniques used to prepare, measure and analyse the electron transfer properties of metalloproteins, concentrating on scanning tunnelling microscopy–based techniques and advances in attachment of proteins to electrodes.

Introduction

Approximately a third of all proteins are metalloproteins, integrating one or more metal ions within their molecular structure. The ion(s) plays crucial structural and functional roles, one of the most important being in electron transfer (ET), which is associated with processes such as catalysis, respiration and photosynthesis. Relatively recently, the ability to study ET in individual protein molecules has become possible through technological advances centred around scanning probe microscopy (SPM) techniques – in particular scanning tunnelling microscopy (STM) \cite{1} and electrochemical STM (EC–STM) \cite{2}. The latter technique especially can probe the redox properties of single metalloprotein molecules, going beyond simple molecular imaging by extracting ‘spectroscopic’ current-voltage characteristics of molecules when electrochemically ‘gated’ by a third electrode.

Interfacing metalloproteins with electrodes

Because STM and EC–STM (outlined below) depend on current passing between an electrically conducting tip and a substrate, it is necessary to immobilise metalloproteins on a surface, often a noble metal such as gold,
in order to measure current flow between tip and substrate as mediated by the molecule. The nature of the protein molecule/substrate (and tip) interface is critical to making successful electrical measurements of the ‘intrinsic’ molecular conductance, as first demonstrated in experiments on small organic molecules [3, 4] which showed that without chemically bonded contacts a very high resistance was measured, arising from the large effective tunnelling barrier between molecule and tip or substrate.

For large molecules like proteins, there are other important considerations. In natural metalloprotein systems, the metal cofactor is organised and tuned within the protein environment. The distance and orientation of metal cofactors with respect to donor and acceptors within the same protein, and with different proteins or small molecules, plays a key role in function, including ET [5]. Thus, the protein-gold interface can be critical in determining orientation and distances between metal centres and electrodes, which may dictate the electronic paths through the protein. The ideal scenario for study (and practical applications) is of a homogenous system where all proteins orientate in a designed and desirable manner, with a protein acting as a single molecule bridge between tip and substrate.

Many proteins tend to bind to gold surfaces through passive adsorption. This is mediated primarily by non-covalent interactions, including hydrophobic and polar. However, such interactions are weak and non-specific, making measurements between the tip and substrate difficult. Adsorbed proteins may lose function through processes such as unfolding, and tips can quickly become fouled with poorly-bound proteins during measurement. To overcome this issue, proteins are normally covalently linked directly or indirectly to a gold substrate. A common and straightforward approach is to derivatise the gold surface to form a surface assembled monolayer (SAM), figure 1(a). These ‘linking’ molecules comprise a head group for attachment to gold at one end and a protein attachment site at the other end, separated by a chemical spacer – see figure 1 for an overview.

Gold attachment chemistry [7] includes amines (R − NH₂), phosphines, or thiol (R − SH), but others are becoming available such as N-heterocyclics carbenes [8]. By far the most commonly utilised is the gold–thiol complex [3] as it forms an interaction with a strength approaching that of a covalent bond (∼200 kJ/mol). The head group is then separated from the protein attachment site by an inert spacer, normally an alkane or ethylene chain. The protein attachment end is commonly a carboxyl group or a derivative thereof but can also be an amine or thiol. Usually the attachment site is activated to enable crosslinking to the inherent chemistry of a protein. For example, making the head group reactive to primary amines allows covalent attachment of protein via lysine residues. This can be achieved by the commonly-employed EDC-NHS crosslinking reaction (see [9] for a recent review), a method familiar to many biochemists in a variety of different contexts when there is a requirement to conjugate proteins. Other approaches
Figure 1: Approaches to attach proteins to gold via surface assembled monolayer. The protein shown in the picture is the haemoprotein myoglobin PDB 3rgk [6]. (a) general overview of SAM system. (b) Chemical features of various gold and protein attachment events.

include the use of maleimide chemistry (figure 1(b)), via a thiol head group and protein amine group or vice versa, and even Click Chemistry (azide-alkyne cycloaddition) [10–12] and Ni(2+)-His tag affinity approaches [13].

The main problems with the use of chemical intermediaries are lack of intimate and specific contact. In terms of intimacy, the additional physical spacing between the metalloprotein and gold due to an intermediary will in general provide a tunnelling barrier for ET, an effect well studied for small molecules. In terms of specificity, while conjugation to protein via lysine residues is relatively straightforward, their amine groups are relatively common on a protein’s surface so that multiple attachment sites are available. This in turn gives rise to the unwanted problem of a heterogeneous population of orientations and distances between the protein’s metal centre and the gold substrate, which results in variation in measured electronic properties.
that is particularly problematic for single molecule analysis. Furthermore a
defined and strong tip-protein interaction is lacking.

One way to overcome some of these challenges is to utilise the inher-
ent chemistry in proteins, most notably in the form of cysteine, one of the
natural twenty amino acids, which contains a thiol group $R - CH_2SH$ (fig-
ure 2). However, cysteine is one of the least common amino acids found in
proteins and is rarely present as the free thiol on a protein surface, which is
necessary to bind gold. Nevertheless, naturally placed cysteines have been
used to good effect with regards to metalloprotein gold interactions. For
example, Artés et al [14] utilised the native cysteines to attach azurin di-
rectly to a gold surface [15]. The main drawbacks of such an approach are
(i) the point of protein attachment to gold is fixed, so preventing different
relative orientations (and thus pathways) from being studied, and (ii) the
tip–protein interaction is variable and relatively weak.

Ideally, one would like to place the gold substrate interfacing site at
any desired position in the protein and introduce a secondary ‘tip’ inter-
acting site (figure 2b). Using a simple twelve amino acid peptide, Nichols
and colleagues generated linked substrate-peptide-tip bridges via N-terminal
histidine (imidazole) and C-terminal cysteine (thiol) [16]. Although chemi-
cally synthesising short peptides with desired amino acid sequence features
is straightforward, it is not so easy for functional metalloproteins, whose
sizes range from $\sim 100$ amino acid single domain proteins (e.g. cyt $b_{562}$,
cyt c, azurin, ferredoxin) to large proteins complexes (e.g. photosystem I)
that generally have to be produced through recombinant cellular methods.
Instead, protein engineering can be employed together with recombinant
protein production.

Protein engineering is the process by which a protein’s sequence is altered
in a designed manner; defined residues are mutated to those with properties
the researcher requires. Mutagenesis is implemented at the level of the gene
using oligonucleotide-directed mutagenesis approaches. The power of the
approach is that any amino acid residue in any protein is potentially open
to change as a single event, or as multiple events (more than one amino acid
mutated per sequence), although in practice the choice is limited. With
respect to substrate and tip interfacing, residues on the surface and not
involved in function are the main targets. Amino acid substitutions to cyste-
ine residues are usually employed to facilitate a strong protein-metal inter-
action. For example, we have engineered both single [18] and pairs [19, 20]
of cysteines into cyt $b_{562}$ to sample how different attachment sites relative
to the haem centre affect conductance. The mutations in the double cyst-
eine cyt $b_{562}$ variants were designed so that only one thiol would interface
directly with substrate and the second with the tip, and opened up direct tip-
protein-substrate conductance experiments; these demonstrated remarkably
enhanced conduction, interpreted as suggesting the possibility of coherent
transport [19]. Interestingly, attachment position and thus orientation and
Figure 2: Direct attachment of protein to gold substrate and tip through engineered cysteines residues. The protein shown in the figure is cyt \(b_{562}\) (PDB 1qpu [17]). The haem is shown as green sticks with central Fe atom in orange. Different attachment points and thus different orientations and distances can be measured using approaches such as EC-STM. (a) A single cysteine generates a defined interface with the gold substrate. (b) By introducing a second cysteine residue an interaction with the tip (right hand side) is also possible. See references [18, 19] for details.

Distances between the protein’s metal centre and the substrate/tip influenced the conductance characteristics and provided insight into potential ET pathways through a protein. It has also been shown [14, 19, 21] that redox active metalloproteins can act as a transistor, with conductance being redox-gated; with respect to cyt \(b_{562}\) maximal conductance was observed at the redox potential for cyt \(b_{562}\) at which the Fe\(^{2+}\)/Fe\(^{3+}\) are at equilibrium. Our combined mutagenesis and EC-SDTM approach is a powerful one and should be extended to a broader range of proteins, linkers, and substrates in future. Indeed, it has been applied in a recent study of azurin molecules [22], the results again being suggestive of coherent tunnelling and supported by examination of temperature dependence.

At the other end of the protein size scale, Gerster et al [23] impressively engineered the photosystem I complex to form a bridge between tip and substrate to undertake single molecule scanning near field optical microscopy. Again using engineered cysteines, one face of the complex was interfaced with the gold substrate (via two cysteines) with an additional cysteine at the opposite end for tip interactions. They were then able to analyse the photo-induced current properties of the complex, providing insights into the dynamics and pathways of ET in this multi-metal centre complex.
Although protein engineering through designed placements of cysteines provides an attractive and powerful route to defining protein-gold interactions, there are some potential problems. The first is that the protein may have inherent cysteine residues. As cysteine is a rarely-used amino acid in proteins, this can be overcome by mutating out non-essential ones. Others (e.g. iron-sulphur and zinc fingers) may be coordinated to the metal centre of the protein or form disulphide bridges or buried away from the surface. When the protein is to act as a bridge between substrate and tip, at least one additional cysteine is required. The main downside is that it is difficult to control which of the two cysteines residues interfaces with the gold substrate and thus orientation can be one of two ways. One way to avoid this is to use another amino acid type such as histidine [16]), although binding is weaker. Another approach could be combine thiol chemistry with abiological chemistry through the use of expanded genetic codes [12, 24–28]. The thiol acts as an initial tether to the gold substrate, with bioorthogonal modification used to attach a tip-binding group to the abiological chemical handle.

**STM and EC–STM methodology**

STM and EC–STM have greatly advanced our ability to investigate single molecule conduction and have been widely applied to a range of small single organic molecules and more recently to proteins. From small model proteins such as cyt \(b_{602} \) [18, 29], cyt \(c \) [30, 31] and azurin [21, 32] to multiprotein complexes such as photosystems involved in light harvesting [23], a variety of metalloprotein systems have proved amendable to these approaches. STM-based measurements thus are a powerful approach for investigating protein function, especially ET events, of metalloproteins at the single molecule level [33,34]. Although proteins which are not redox active can be examined with STM [35] much work has been performed on metalloproteins such as the blue single-copper redox active protein azurin [21,36,37] which attaches naturally to Au surfaces through cysteine residues.

There are many excellent reviews [38–43] on STM and EC–STM, and we outline only the salient points here. EC–STM combines a simple STM with an electrochemical control circuit, figure 3(a). In STM, an electrical bias \(V_t\) is applied between the STM tip and sample (both typically gold). A quantum mechanical tunnel current \(I\) flows across the gap between tip and sample, which is decorated with attached protein molecules. By scanning the tip across the surface, variations in the controlled height \(z\) of the tip above the sample required to maintain a set feedback current (the ‘set point’ of \(\lesssim\) a few nA) are used to build up an image. Much more than just a molecule’s location and apparent height can be determined. By placing the tip above a molecule, with feedback turned off, \(I\) can be measured as a function of \(V_t\) to yield spectroscopic information which relates to the molecular electronic
Figure 3: (a) Schematic of EC-STM. A bipotentiostat controls the potentials on both the STM tip (typically 10-100 mV) and the surrounding solution with respect to the grounded substrate (or working electrode). The tip body is insulated to minimise ionic current flow. The tunnel current is measured with an I-V converter and depends on tip-sample distance $z$ as $I(z) \sim e^{-\beta z}$ (except when above the molecule). Not show is the computer-interfaced piezoelectric feedback system by which the tip’s position across, and distance above, the surface is controlled. (b) EC–STM apparent height images of cyt $b_562$ protein molecules [18] depend on the applied potential $V$ (overpotential top to bottom -0.21, -0.01 and 0.13V) which affects the molecular conductance.

properties - particularly the location of the highest occupied and lowest unoccupied molecular orbital (HOMO and LUMO) levels. A simple way to view this (although not capturing all the complications of a redox-active protein) is to picture the energetic line up of the tip and substrate Fermi levels with the molecular levels, figure 4(a, b).

In a picture [44] of coherent tunnelling (no inelastic scattering to destroy phase coherence) the molecular levels, which rigidly shift up or down with electrochemical bias $V$, can be aligned in the ‘Fermi window’ – i.e. the energy range between $\mu_1$ and $\mu_2$. If a HOMO or LUMO level lies in this window (‘on resonance’) then an electron can tunnel onto the level from the left electrode and off to the right electrode in a single quantum step. The tunnel time $\tau$, the major factor determining the overall current, depends on
Figure 4: Electron tunnelling from substrate (left, grounded) to STM tip (right, positive bias $V_t$) mediated by a molecular energy level. The vertical scale is energy in eV. The horizontal scale suggests the contact positions as well as the density of states (DOS). The DOS is shown as constant for the metals. The illustrated delta function DOS of the molecular levels actually has an energy width $\Gamma$ due to coupling with the metallic states. Electrochemical potentials $\mu_1$ and $\mu_2$ are separated by an energy $V_t$ due to the bias. At zero bias, the metals’ Fermi levels lie within the HOMO-LUMO gap. Tip bias lowers the molecular levels by $V_t/2$ (with respect to the substrate) if symmetrically coupled to the contacts. Applied bias $V$ affects just the molecular levels. Open (closed) circles indicate unoccupied (occupied) molecular states. (a) Even if the closest molecular level (LUMO here) is above $\mu_1$, a current can flow through its broadened tail states. (b) If the LUMO is between $\mu_1$ and $\mu_2$ an enhanced current flows - see figure 3(b). (c) Illustration of a cyt $b_{562}$ molecule with central haem cofactor responsible for redox activity and electron transfer.

the degree of wave function overlap between the relevant molecular orbital (MO) and metals. An equivalent description is an uncertainty-principle broadening $\Gamma = h/\tau$ of the MO. The MO is therefore partially occupied. Even when a MO is moved out of alignment (‘off resonance’) a smaller current can flow through the tail region of the broadened orbital. It interesting that, as suggested by a number of authors [19, 45, 46], if a protein molecule can be intimately (covalently) linked to both the STM tip and substrate, strong electronic coupling of molecular levels to the continuum of metallic
levels might lead to a degree of coherent transport. Depending on relevant time scales, it is also possible for a two-step process to occur, where the electron tunnels onto the MO, undergoes inelastic scattering with the molecule, and then tunnels off again. In this case, it is further possible that there is a physical reorganisation of the molecule (and orbital energies) during the process, as known from classical electrochemistry studies of ET in redox-active proteins. Indeed, most EC-STM data is analysed through a model to allow for this effect (see for example references [47,48]). However, Băldea [49] has considered an alternative approach, perhaps valid for strong tip and substrate coupling, applying instead a Newns-Anderson framework to detailed $I-V$ measurements on azurin [50]. In the opposite limit of weak electronic coupling and strong scattering an incoherent hopping charge transport may be expected [51]. To distinguish experimentally which, if any, is the dominant mechanism requires more information than simple $I-V$ characteristics. In some simple molecular systems this may be provided for example by examining the length dependence of conductance of a homologous series of molecules such as alkanedithiols [52,53], or (appropriate to proteins) by measuring temperature dependence, or by examining gate bias dependence.

It might be worth noting in this context, without being too pedantic, that the terms electron (or charge) ‘transfer’ and ‘transport’ are often used interchangeably in the wider literature. Transport is perhaps best reserved to describe the continuous flow of electrons across a protein contacted by two electrodes maintained at different electrochemical potential. See for example the discussion in reference [54] which also summarises the variety of physical mechanisms involved.

Individual protein molecules cannot always be clearly distinguished, for example on a rough surface or for weakly attached molecule. The so-called $I(z)$ or $I(s)$ method [55,56] nevertheless can provide conductance data. Here, $I$ is measured as the STM tip is withdrawn from the surface; occasional clear signature of molecular attachment is seen. The ‘break junction’ method [52] goes further and actually drives the tip onto the surface before withdrawal. The related $I(t)$ method [57,58] simply locates the STM tip near the surface until a molecule randomly bridges the gap. These techniques require a great deal of experimental care and careful interpretation.

Electrochemical control of course opens more powerful spectroscopic abilities over STM. In EC-STM a potentiostat controls or ‘gates’ the potential $V$ of the electrolyte solution with respect to the substrate and measures the resulting current $I_m$ flowing to a counter electrode. The potential is defined with respect to a reference electrode which is maintained at equilibrium by the circuitry. Thus, not only can the tip-substrate bias $V_t$ be chosen (shifting the MOs by $\sim V_t/2$ as a consequence) but also the potential $V$ in the solution around the molecule can be chosen (to shift the MOs independently). This is analogous to gating in the operation of a field effect
More recent work includes the detection of single-molecule redox events switching events [32] for azurin using time-resolved measurement. It has also been reported [59] that applied voltage pulses can control the conductance of single azurin molecules. Sub-molecular features of single streptavidin proteins in solution have been resolved with STM [60] and it is interesting to extend this to a metalloprotein to observe the redox centre under gating. (Most, though not all [45] EC-STM imaging experiments on proteins generally reveal rather featureless structures.) In fact, submolecular resolution in single azurin molecules, using a modified EC-STM setup, has very recently been reported [61]. There is also growing understanding of the important role of the complex surrounding medium [62]. For example, local water molecules and ions can both influence the average local electrostatics (shifting molecular levels) and have important dynamical effects. A protein’s environment is of course intimately linked to its overall physical properties.

There is also a need to understand at a detailed molecular level the process of ET between a protein and contacting electrodes. The two-step [63] and competing models [46] of charge transfer have been employed to describe the essential mechanisms phenomenologically, but a nanoscale description will involve numerical computation of molecular dynamics as well as the electronic properties of the protein hybrid structure.

Conclusions

Single molecule studies of electron transfer in redox active metalloproteins have reached a fairly mature state, with a few select proteins now used routinely as model systems. However, there are several areas that are ripe for further development. It is particularly interesting to see how far studies of dynamic behaviour can be advanced, potentially providing deeper insight into the nanoscale electron transfer mechanism. Theoretical understanding promises to develop, examining interactions with the detailed protein environment (water, ions, electrostatics, docking with other proteins) and the influence of strong electronic coupling between tip/molecule/substrate. Biosensor applications will doubtless be a strong motivation here. Protein engineering has already proved to be a powerful tool to define how proteins link to metallic electrodes. The advent of reprogrammed genetic code approaches allows new chemistry not found in nature to be incorporated by design, which will further our ability to couple proteins in defined orientations on technologically important substrates and facilitate single protein molecule to electrode bridging. Single molecule studies of metalloproteins should continue to be a fruitful research area for many more years.
Abbreviations

AFM, atomic force microscopy/microscope; STM, scanning tunnelling microscopy/microscope; EC–STM, electrochemical STM; MO, molecular orbital; HOMO, LUMO; highest occupied and lowest unoccupied molecular orbital.

Acknowledgements

We are indebted to many research colleagues and collaborators over recent years, with special thanks due to Jens Ulstrup and Qijin Chi at the Technical University of Denmark and Emyr Macdonald at Cardiff University. We should also like to acknowledge the several grants provided by EPSRC, BB-SRC and MRC as well as Cardiff University scholarships. Thanks are due to the anonymous referee who highlighted the distinction between electron transfer and electron transport, and pointed out references [22,61] to us.

Competing Interests

The Authors declare that there are no competing interests associated with the manuscript.

References


