Exploring amino-acid radicals and quinone redox chemistry in model proteins

Kristina Westerlund
Tillägnad min familj
List of publications

This thesis is based on the following publications, which will be referred to by their Roman numerals (I-V) in the text:


‡These authors have contributed equally.

*Published papers are reproduced with permission from the publisher.*
Contents

1. Introduction to thesis ................................................................. 13
2. Amino-acid radical systems .................................................... 14
   2.1. Amino-acid radical enzymes ........................................... 14
   2.2. The ribonucleotide reductases – a well studied class of amino-acid radical enzymes .................................................. 15
   2.3. Glycyl radical enzymes – Important enzymes in anaerobic metabolism ................................................................. 17
   2.4. Amino-acid radicals and oxidative stress ................................ 17
   2.5. Studying amino-acid radicals in natural systems .................. 18
   2.6. Redox properties of amino-acid radicals in natural systems .......... 18
   2.7. Significance of thesis project, part I ....................................... 20
3. Quinones ..................................................................................... 21
   3.1. Quinones in biology ............................................................ 21
   3.2. Protein-derived cofactors present in quinoproteins ................. 21
   3.3. Quinones involved in biological energy transduction .............. 22
   3.4. Binding sites for non-covalently bound quinones .................... 23
   3.5. Redox chemistry of para-quinones in solution ...................... 23
   3.6. The photosynthetic reaction center of R. sphaeroides – a well studied quinone-containing system ................................. 24
   3.7. Electrochemical characterization of quinones in proteins ............ 27
   3.8. Significance of thesis project, part II ...................................... 27
4. De novo protein design ............................................................... 28
   4.1. What is de novo protein design? ........................................ 28
   4.2. Protein sequence space and the inverse folding problem .......... 28
   4.3. Design of coiled coils and helical bundles – an example of iterative rational design ......................................................... 29
   4.4. Template-assembled synthetic proteins .................................. 30
   4.5. Designed combinatorial libraries of de novo proteins ............. 30
   4.6. Computational protein design .............................................. 31
   4.7. Significance of thesis project, part III .................................... 32
5. Results and discussion ............................................................... 33
   Making a single-chain four-helix bundle for redox chemistry studies (Papers I-II) ................................................................ 33
   Moving a phenol hydroxyl group from the surface to the interior of a protein: Effects on the phenol potential and pKₐ (Paper III) ............ 35
Abbreviations

ATP  adenosine-5'-triphosphate
bCys  N-acetyl-L-cysteine methyl ester
CD  circular dichroism
CTQ  cysteine tryptophylquinone
CV  cyclic voltammetry
DMBQ  2,6-dimethyl-1,4-benzoquinone
DPV  differential pulse voltammetry
ENDOR  electron nuclear double resonance
EPR  electron paramagnetic resonance
FAD  flavin adenine dinucleotide
HF  high frequency
HPLC  high-performance liquid chromatography
HSQC  heteronuclear single quantum coherence
LTQ  lysine tyrosylquinone
MP  mercaptophenol
NAD  nicotinamide adenine dinucleotide
NAYA  N-acetyl-L-tyrosinamide
NHE  normal hydrogen electrode
NMR  nuclear magnetic resonance
RC  reaction center
RNR  ribonucleotide reductase
ROS  reactive oxygen species
SAM  S-adenosylmethionine
TPQ  topaquinone
TTQ  tryptophan tryptophylquinone
UV/Vis  ultraviolet-visible absorption spectroscopy

Amino acids are throughout the text referred to by their standard one-letter or three-letter codes.
Nomenclature

Tyr-122  tyrosine 122
Tyr-122*  tyrosine 122 as a neutral radical
Ile-M256  isoleucine 256 in subunit M
Abstract

Amino-acid radical enzymes have been studied extensively for 30 years but the experimental barriers to determine the thermodynamic properties of their key radical cofactors are so challenging that only a handful of reports exist in the literature. This is a major drawback when trying to understand the long-range radical transfer and/or catalytic mechanisms of this important family of enzymes. Here this issue is addressed by developing a library of well-structured model proteins specifically designed to study tyrosine and tryptophan radicals. The library is based on a 67-residue three-helix bundle (α<sub>3</sub>W) and a 117-residue four-helix bundle (α<sub>4</sub>W). α<sub>3</sub>W and α<sub>4</sub>W are single-chain and uniquely structured proteins. They are redox inert except for a single radical site (position 32 in α<sub>3</sub>W and 106 in α<sub>4</sub>W). Papers I and II describe the design process and the protein characteristics of α<sub>3</sub>W as well as a voltammetry study of its unique tryptophan. Paper III and V describe two projects based on α<sub>3</sub>C, which is a Trp-32 to Cys-32 variant of α<sub>3</sub>W. In Paper III we use α<sub>3</sub>C to investigate what effect the degree of solvent exposure of the phenolic OH group has on the redox characteristics of tyrosine analogs. We show that the potential of the PhO•/PhOH redox pair is dominated by interactions with the OH group and that the environment around the hydrophobic part of the phenol has no significant impact. In addition, we observe that interactions between the phenolic OH group and the protein matrix can raise the phenol potential by 0.11-0.12 V relative to solution values. The α<sub>3</sub>C system is extended in Paper V to study quinone redox chemistry. Papers III and V contain protocols to generate the cofactor-containing α<sub>3</sub>C systems and descriptions of their protein properties. Paper IV describes efforts to redesign α<sub>3</sub>Y (a Trp-32 to Tyr-32 variant of α<sub>3</sub>W) to contain an interacting Tyr-32/histidine pair. The aim is to engineer and study the effects of a redox-induced proton acceptor in the Tyr-32 site.
1. Introduction to thesis

Amino-acid radicals are involved in functional as well as harmful redox chemistry in living organisms and, yet, very little is known about their thermodynamic properties. The main experimental barrier arises from the typically highly oxidizing potentials of these species, which severely hamper electrochemical investigations. The central strategy of this thesis project is to use designed proteins that contain the redox-active species of interest as well as features that facilitate spectroscopic, structural and, most importantly, electrochemical characterization of the model proteins. The studies included in this thesis are based on the structurally characterized de novo \( \alpha_3 \)W three-helix bundle protein (Papers III-V) and the recently constructed \( \alpha_4 \)W four-helix bundle protein (Paper I-II). Paper I-II represents mainly protein-design oriented efforts while Papers III-IV describe functionally oriented studies on tyrosyl radicals. Paper V illustrates a successful extension of the three-helix bundle system to study protein quinone redox chemistry. To provide a background to Papers I-V, we briefly summarize the main characteristics of amino-acid radical systems and quinone-containing enzymes, and then describe the basics of de novo protein design.
2. Amino-acid radical systems

2.1. Amino-acid radical enzymes

Amino-acid radical enzymes catalyze a number of fundamental biochemical reactions including DNA and hormone synthesis, carbohydrate metabolism, cell detoxification reactions and energy transduction (Table 1) (1-5). Amino acids acting as redox cofactors involved in these processes are shown below.

Figure 1. Amino-acid radical enzymes contain redox-active (A) tyrosine, (B) tryptophan, (C) cysteine, and (D) glycine residues. Post-translationally modified tyrosyl radical cofactors including covalently cross-linked (E) Tyr-Cys and (F) Tyr-His species are also used. The amino acids serve as single electron redox cofactors involved in electron-transfer or proton coupled electron-transfer reactions.

Table 1 lists enzymes shown to contain amino-acids radicals. The ribonucleotide reductases (RNRs), photosystem II, galactose oxidase, pyruvate formate lyase, prostaglandin H synthase, and cytochrome c peroxidase are long-standing members of this class of enzymes and the essential function of the radicals in these systems is firmly established. For other enzymes, such as linoleate diol synthase, bovine catalase and the katG proteins, the role of the observed radicals need further investigations. Some exciting recent developments in the radical enzyme field include: i) The discovery of the Radical-SAM protein superfamily (6), which most likely will lead to a rapid increase of known glycyl radical proteins (5). ii) Reports of the possible involvement of tyrosyl radicals in the reaction mechanisms of human monoamide oxidase A (7) and cytochrome P450cam (8), both of which are of considerable medical and pharmaceutical importance. iii) That tryptophan- and tyrosine-based radical transfer chains may be involved in the action of bluelight-activated photoreceptors. These systems mediate developmental, growth and/or circadian responses in a variety of species (9-11).
Table 1. Radical detection methods, structural studies, and thermodynamic characteristic of amino-acid radical enzymes. a) redox-active residue(s), b) cofactor(s) involved in amino-acid radical generation, c) methods used in the initial detection of the radical(s), d) resolution of the first crystal structure reported for each protein, and e) references to all reported amino-acid $E_a$ values. x indicates no available data in the literature. References are given in parentheses and abbreviation stand for: absorption (abs.) spectroscopy (spec.), electron paramagnetic resonance (EPR), electron nuclear double resonance (ENDOR), high frequency (HF), substitutions (sub.), Clostridium difficile (Cd), Escherichia coli (Ec), Aspergillus nidulans (An), Arabidopsis thaliana (At), Rhodobacter sphaeroides (Rs), Streptomyces coelicolor (Sc), Bos taurus (Bt), Mycobacterium tuberculosis (Mt) and Synechocystis (Sy).

2.2. The ribonucleotide reductases – a well studied class of amino-acid radical enzymes

Tyrosine-122 in E. coli RNR was the first protein residue to be identified as a catalytically essential amino-acid radical. In 1972, Reichard & Ehrenberg reported the EPR spectrum of an organic radical in RNR purified from E. coli grown under oxygenic conditions (66). Isotopic labeling assigned the spectrum to a tyrosyl radical (16) and site-directed mutagenesis studies subsequently identified the redox-active species as Tyr-122 (17).

The RNR enzymes catalyze the conversion of ribonucleotides to deoxyribo-nucleotides and they are essential for the de novo DNA synthesis in all organisms known (67, 68). The catalytic mechanism involves the generation of a cysteiny radical at the active site. This radical initiates the catalytic cycle by abstracting a hydrogen atom from the ribose ring of the substrate.
The methods for generating the active-site cysteinyl radical are diverse and divide the RNR family into three major classes (1, 69).

![Proposed radical-transfer pathway in the E. coli ribonucleotide reductase enzyme. The spatial organization shown in Figure 2 is based on an in silico docking model of the R1 and R2 subunits (70). Reversible radical transport has been proposed to occur via a radical hopping mechanism along the following pathway: Tyr-122* ⇔ 2 Fe ⇔ Trp-48 ⇔ Tyr-356 ⇔ Tyr-731 ⇔ Tyr-730 ⇔ Cys-439 ⇔ substrate (68, 71). Tyr-356 resides within the unstructured C-terminal part of R2 and it is not visible in the crystal structure. The distance marked in red is based on the docking model. The figure is adapted from reference (72).]

The class I systems, and in particular the E. coli enzyme, has been extensively studied. The homodimeric enzyme contains a radical-generating domain (R2) and a catalytic domain (R1). R2 binds two Fe ions in close proximity to Tyr-122 (Figure 2). O₂-dependent redox chemistry at the di-iron center generates the oxidized form of Tyr-122. The tyrosyl radical is unusually stable, which in parts can be attributed to its highly sequestered location. Tyr-122* is involved in reversible long range radical transfer to the catalytically active Cys-439* in the active site of R1 (a distance of >35 Å based on a docking model of the E. coli R1R2 complex) (69, 70). Studies involving chemically modified amino acids have recently been conducted in order to probe the characteristics of the proposed radical transfer pathway (73-76).

Class II RNRs contains the vitamin B₁₂ derivative 5'-deoxyadenosylcobolamin. Homolytic cleavage of the C-Co bond is proposed to generate a deoxyladenosyl radical, which, in turn, abstracts a H-atom from the active-site cysteine (77). Class III RNRs are oxygen sensitive and are expressed in strict or facultative anaerobic bacteria and in some bacteriophages. These enzymes contain a redox-active glycine, which upon oxidation generates the active-site cysteinyl radical. Generation of the glycyl radical requires an activase, which is described in more detail below.
2.3. Glycyl radical enzymes – Important enzymes in anaerobic metabolism

Glycyl radical enzymes operate under anaerobic conditions. The peptide bond adjacent to the glycyl radical is cleaved upon oxygen exposure and the system is irreversible inactivated. The members of this enzyme family are involved in a broad range of metabolic pathways including anaerobic breakdown of pyruvate (24), toluene metabolism (26), fermentative production of para-cresol (27) and glycerol fermentation (78). Class III RNRs belong to this protein family. All glycyl radical enzymes are expressed in an inactive precursor form and require activation by a specific activase. The activases are members of the Radical-SAM superfamily and contain an iron-sulfur cluster and S-adenosylmethionine (SAM). Reductive cleavage of SAM by the iron-sulfur cluster produces a transient 5′-deoxyadenosyl radical, which abstracts a H-atom from the α-carbon of the strictly conserved glycyl residue. From sequence homology studies, putative new glycyl radical enzymes with yet unknown functions have been identified in E. coli and other facultative and strict anaerobic bacteria and archaea. Glycyl radical enzymes may play an important role in the anaerobic metabolism in various species and the number of identified enzymes is likely to increase in the near future (79).

2.4. Amino-acid radicals and oxidative stress

Oxidative stress results from exposure to high levels of reactive oxygen species (ROS) that have not been detoxified by the system. ROS are generally small, highly reactive molecules formed by an incomplete reduction of oxygen. ROS can be formed by ionizing radiation of biological molecules, by exogenous agents such as asbestos or ozone, or as biproducts in the respiratory electron transport (80). ROS can oxidize virtually all cellular components including DNA, lipids and proteins. Increased levels of ROS are associated with aging and many diseases such as cardiovascular diseases, diabetes mellitus (81) and rheumatoid arthritis (82). In proteins, ROS can oxidize sites at the backbone and/or the side chains and even cause cleavage of the peptide bond. Backbone oxidation in proteins occurs mainly by H-atom abstraction at the α-carbon and the resulting radical is stabilized by electron delocalization onto the peptide backbone. Glycine can adopt particularly favorable backbone configurations upon oxidation. All amino acid side-chains are susceptible to ROS oxidation but aromatic residues and sulfur-containing amino acids are the major sites of oxidations in proteins. Thus, the four residues used for functional redox chemistry in radical enzymes (Figure 1) represent particularly vulnerable positions under oxidative stress conditions. Most oxidative damage is non-repairable and may lead to deleterious consequences on protein structure and function.
2.5. Studying amino-acid radicals in natural systems

Dysfunctional radical chemistry typically occurs infrequently during the normal course of cellular processes. It is consequently difficult to obtain a high steady-state concentration of the reactive radical, which hampers experimental investigations. Studies on the experimentally more controllable amino-acid radical enzymes have been performed at a more molecular level. Crystal structures are available for the majority of these proteins, including several membrane-bound multisubunit enzymes. We note that only the resolution of the first crystal structure reported for each system is shown in Table 1 and that more refined structural data are available for many of the listed proteins. In addition, structures from different organisms and/or structures representing different redox states exist for several of the proteins listed in Table 1. An impressive array of spectroscopic techniques, including multi-frequency EPR-based methods, Fourier transform infrared and transient optical spectroscopy, has been used to identify the radical species and to derive detailed structural and kinetic information.

In sharp contrast to the extensive structural and spectroscopic work, virtually nothing is known about the reduction potentials of amino-acid radicals and their dependence on solution pH and the protein environment. The sheer size and complexity of amino-acid radical enzymes combined with the highly oxidizing nature of their radical cofactors severely hamper electrochemical measurements. Voltammetry and spectroelectrochemical measurements must typically be performed in the +1.0 V potential range, which for most enzymes will lead to oxidation of other cofactors, chromophores, and/or non-active amino acids with potential destruction of the protein as a consequence. Voltammetry experiments performed in this potential range in aqueous buffer will furthermore have a substantial background signal arising from water oxidation occurring at the surface of the working electrode. This complicates the analysis of the voltammogram, particularly if it reflects multiple redox events. Studies based on optical detection are further limited by the low extinction coefficients of amino-acid radicals. The weak spectral features of amino-acid radicals can easily be obscured by the spectral bands of other chromophores and/or added redox mediators.

2.6. Redox properties of amino-acid radicals in natural systems

The experimental limitations described above are so significant that no more than 10 published reports are available in the literature on the potentials of amino-acid redox cofactors (Table 1). In fact, the Cu-ligated, cysteine cross-linked Tyr-272 species in the active site of galactose oxidase is the only amino-acid radical system that has been studied systematically as a function of the protein microenvironment and solution pH (Figure 3; see legend for more details on this enzyme).
Galactose oxidase catalyzes the oxidation of primary alcohols to the corresponding aldehydes. The active site of galactose oxidase contains a single copper (II) ion (shown in gold), which is coordinated by Tyr-272, Tyr-495, His-496, His-581 and an exogenous ligand (in this crystal structure an acetate ion; PDB code 1GOF). Tyr-272 is the redox-active amino acid and it is covalently linked at its ortho position to the sulfur atom of Cys-228. Also shown in the Figure is Trp-290, of which the aromatic headgroup stacks up against the Tyr-272/Cys-288 bond. The \( E_m \) value of \( \text{Tyr-272}^\bullet/\text{Tyr-272}^- \) vs. NHE varies with pH from 0.38 V at pH 8.5 to 0.54 V at pH 5.5. Tyr-272 is deprotonated in its reduced state and the pH dependence arises from the pKa value of Tyr-495. The radical has a pH-independent potential of 0.425 V in Y495F (47). Other residues that influence the reduction potential of Tyr-272 include the cross-linked Cys-228 residues and the closely located Trp-290 residue. A single-point mutation raises \( E_m \) (pH 7.5) from 0.63 V in the wild type (45, 47) to 0.73 V in C228G (48). Trp-290 is believed to restrict entry to the active site and to help stabilizing the radical on Tyr-272 (46). Replacement of the bulky Trp with His destabilizes the radical and increases its potential to 0.83 V. Finally, the \( E_m \) (pH 7.5) value is 0.57 V in the apo enzyme (48). The combined effect of the copper ion and the Tyr-495, Cys-228 and Trp-290 residues is to significantly lower the reduction potential of Tyr-272 relative to aqueous tyrosine.

There are two protein film voltammetry studies characterizing the cooperative, two-electron Fe-O(IV)W+/Fe(III)W redox pair in cytochrome c peroxidase (14, 15), and three estimates for unmodified tyrosyl redox cofactors. Vass & Styring estimated \( E_m \) values of 0.74 and 0.97 V vs. NHE for the Y_D and Y_Z tyrosine residues in photosystem II, respectively (31). These numbers were derived by measuring equilibrium constants between several of the redox cofactors in the photosystem II electron-transfer chain and fixing the calculated \( \Delta E_m \) values to the potential obtained for QA/QA^- by redox potentiometry. The equilibrium constants and the QA/QA^- potential were collected from several different studies, and therefore, the experimental conditions, including the pH, to which the estimated Y_D and Y_Z potentials correspond, are not well defined. A similar strategy was used in a more recent study placing the Y_Z^\bullet/Y_Z^- potential at 1.21 V (32). Finally, Silva et al. made an attempt to measure the \( E_m \) of Tyr-122 in E. coli RNR by redox titration (19). The Tyr-122 site is sequestered and only six of the 16 redox mediators included in this study were reported to interact with Tyr-122. Of these six mediators, five (with \( E_m \) in the range of 0.341-0.724 V) reduced the radical to various degrees and one (\( E_m = 1.33 \) V) gave rise to a slight oxidation of Tyr-122. The poor contact between the bulk medium and the radical pre-
vented a true equilibrium potential to be determined and an apparent value of 1.0 V (pH 7.6) was given for the Y122*/Y122 pair.

2.7. Significance of thesis project, part I

Amino-acid radical enzymes have been studied for over 30 years and to date there is only one system, galactose oxidase (Figure 3), for which protein structural features have been systematically correlated with changes in the reduction potential of the active amino-acid radical. It appears safe to conclude that a complete chemical understanding of these biochemically important redox species cannot be obtained by concentrating experimentally solely on the naturally systems. The main motivation for the work described in this thesis was to develop an alternative approach to study the thermodynamic properties of amino-acid radicals. The goal was to gain insights into essential and unresolved issues such as if, and by how much, protein amino-acid radical potentials are modulated by solvent accessibility and the hydrogen-bonding environment. Papers III and IV describe such efforts and are discussed in more detail in Section 5.
3. Quinones

3.1. Quinones in biology

Quinones are aromatic diketone compounds with their two carbonyl groups placed on the same or separate rings. Naturally occurring quinones can be divided into subgroups of which benzo-, naphto-, and anthraquinones represent the major groups (Figure 4). Quinones can serve one- or two-electron acceptors and they participate in a wide range of biological redox processes such as NAD- and FAD-dependent reactions, electron and proton-transfers associated with photosynthetic and respiratory energy transduction, and in reactions catalyzed by a group of enzymes denoted quinoproteins.

![Figure 4. Quinones found in enzymes and redox protein include (A) 1,4-benzo-quinone, (B) 1,4-naphtoquinone, and (C) anthraquinone (9,10-dioxoanthracene).](image)

3.2. Protein-derived cofactors present in quinoproteins

Quinoproteins contain both covalently and non-covalently bound quinones, although the majority of these enzymes have amino-acid derived cofactors (Figure 5). Topaquinone (TPQ) is the tyrosine-derived prosthetic group of amine oxidases and this molecule was the first protein-derived quinone to be characterized in detail (83, 84). Mammalian lysyl oxidase contains lysine tyrosylquinone (LTQ), which is formed by a covalently cross-linked Tyr-Lys pair. (84). Tryptophan tryptophylquinone (TTQ) and cysteine tryptophylquinone (CTQ) are two examples of tryptophan-derived quinones. TTQ is found in bacterial methylamine dehydrogenase and aromatic amine dehydrogenase and is formed from two covalently linked tryptophan residues, while CTQ from quinohemoprotein amine dehydrogenase is derived from a cross-linked Trp-Cys pair. (85). Quinoproteins containing tyrosine-derived cofactors, TPQ and LTQ, oxidize primary amines to aldehydes while transferring two protons and two electrons to molecular oxygen and forming hydrogen peroxide. Proteins containing tryptophan-derived cofactors are also involved in the oxidation of primary amines but their cofactors are reoxidized by exogenous acceptors such as other proteins rather than by molecular oxygen (86).
Figure 5. The protein-derived redox cofactors found in quinoproteins include (A) topaquinone (TPQ), (B) lysine tyrosylquinone (LTQ), (C) tryptophan tryptophylquinone (TTQ), and (D) cysteine tryptophylquinone (CTQ).

3.3. Quinones involved in biological energy transduction

Due to the ability of quinones to carry both electrons and protons (see Section 3.5 below) they play an important biochemical role in coupling electron transfer to proton movement. Thus quinones are essential for the formation of electrochemical proton gradients in photosynthetic and respiratory systems. Ubiquinone (Figure 6A) has been given its name because of its ubiquitous presence in nature and this molecule is found, for example, in the mitochondrial respiratory chain. Both ubi- (Figure 6A) and menaquinones (Figure 6B) are present in reaction centers (RCs) isolated from photosynthetic bacteria while plastoquinone (Figure 6C) is involved in electron and proton transfers in plant chloroplasts. An essential feature of the quinones involved in energy transduction is their long isoprenoid chain. The hydrophobic tail makes the quinones more soluble in the lipid bilayer of biological membranes and allow the oxidized and reduced cofactors to diffuse between membrane-bound protein complexes.

Figure 6. Quinones involved in biological energy transduction include (A) ubiquinone, (B) menaquinone, and (C) plastoquinone. The length of the isoprenoid chain depends on the species.
3.4. Binding sites for non-covalently bound quinones

Quinones can be covalently bound, as is the case with the quinoproteins described above, or, more commonly, associated to proteins by non-covalent forces including hydrogen bonding to the carbonyl/hydroxyl groups, and van der Waals interactions with the aromatic ring and the aliphatic isoprenoid chain. Interactions at the binding sites can have a profound effect on the redox characteristics of the quinones and different types of binding sites are associated with different functional properties of quinones (87). A well-studied example is the RC from the purple bacterium *R. sphaeroides* in which the two electron acceptor-side ubiquinones have very different redox characteristics although they are chemically identical. In order to discuss this system in more detail, the redox properties of *para*-quinones in solution will first be briefly described.

3.5. Redox chemistry of *para*-quinones in solution

Quinones have nine possible redox states (Figure 7) due to their ability to bind two electrons and two protons (88, 89). In aprotic media a quinone is reduced in two steps, the first forming the anionic semiquinone species, Q\(^-\), and the second forming the dianionic quinol Q\(^{2-}\). Due to electrostatic effects, the formation of Q\(^{2-}\) requires more energy than the formation of Q\(^-\) and, consequently, the redox behavior of quinones in aprotic media is relatively simple with two reversible and typically well-resolved \(n = 1\) reactions (90).

![Figure 7. Different para-quinone states resulting from electron and proton binding. Horizontal arrows indicate uptake of an electron while vertical arrows indicate protonation. States marked in bold are those available at pH > 0 (91).](image-url)
In a protic solution the redox behavior of quinones becomes complicated by the protonic reactions associated with the hydroxyl groups. Most semi-quinones have $pK_a$ values around 3-5, which is lower than the $pK_a$ values for the corresponding quinol (91). The quinol states are greatly stabilized relative to the semiquinone states and, as a result, the reduction of quinone in a protic media give rise to an $n = 2$ behavior. This overall $n = 2$ conversion of quinone to quinol can be treated as two $n = 1$ reactions, rate-limited by the first energetically unfavorable one-electron transfer (90).

3.6. The photosynthetic reaction center of R. sphaeroides – a well studied quinone-containing system

Reaction centers are integral membrane protein complexes found in photosynthetic organisms such as the purple bacterium R. sphaeroides. In this system, the RC is part of a light-driven “redox loop” that gives rise to a transmembrane proton gradient subsequently used for ATP synthesis. Central to this mechanism are the QA and QB ubiquinones, which are both buried within the RC. The secondary quinone QB is reduced in two one-electron transfer steps via the primary quinone QA and is released as a protonated quinol, QBH2, to the membrane bilayer (Figure 8; see legend for more details). For each QBH2 produced, the RC goes through two light-induced turnovers and two protons are taken up from the cytosolic side of the membrane. After dissociation from the RC, QBH2 is reoxidized by cytochrome bc1, where the release of two protons to the periplasm completes a “proton translocation redox loop” (92). Although QA and QB are chemically identical molecules they serve two very different roles in the RC quinone redox cycle. Clearly the protein environment plays a major role in tuning the redox properties of the two quinones.

In the QA site (Figure 9) only the fully oxidized QA and anionic semiquinone QA− are found. QA is bound 20 times tighter than QB to the RC (93) and does not dissociate from the protein during the reduction cycle. The QA site is highly hydrophobic and no proton path is associated with the reduction of the quinone. Protonation is needed to stabilize the fully reduced quinol state and inaccessibility to protons may be the reason why QA only functions as a one-electron acceptor under normal conditions (94). The side chain of Ile-M256 is in van der Waal contact with several atoms of the QA head group and changing this residue to a polar Ser or Thr lowers the midpoint potential for the QA/QA− pair by as much as 60 and 115 mV (pH 8), respectively (95). As a contrast, a replacement of Ile-M256 with a Val has a negligible effect on the midpoint potential of the ubiquinone (95, 96). A lowered QA midpoint potential is consistent with a destabilization of the anionic semiquinone state relative to the oxidized quinone. The destabilization may arise from a change in the polarity of the binding site, which will affect the negatively charged QA− more than the neutral quinone, and/or a change in the hydrogen-bonding pattern to the quinone head group (95, 96). Similar effects on the midpoint potential were observed when replacing the native ubiquinone with anthraquinone (Figure 4C) indicating that this effect does
not have a large steric component (96). Substitution of other residues in the vicinity of QA (e.g. Met-M218 and Trp-M252) have a similar effect by increasing the equilibrium constant for electron transfer to QB consistent with a negative shift in the midpoint potential of QA (95).

**Figure 8.** The ubiquinone reduction cycle in the photosynthetic reaction center of *R. sphaeroides*. Different stages in the reduction cycle are numbered by Roman numerals. Figure adapted from references (92, 94). Stage I represents the light-induced electron transfer from the primary electron donor P, a bacteriochlorophyll dimer, to the primary electron acceptor quinone QA. The electron transfer occurs through a series of intermediates before reducing QA to an anionic semiquinone forming the \( P^+Q_A^-Q_B^- \) state. \( P^+ \) is re-reduced to P by cytochrome c2 after each photooxidation event. Stage II involves the first electron transfer step from QA to QB forming the \( PQ_AQ_B^- \) state and a sub-stoichiometric (\( H^+/e^- <1 \)) proton uptake predominantly to Glu-L212 near QB. After a second light-induced electron transfer from P (stage III), QA is reduced to QA forming the \( PQ_A^-Q_B^- \) state. Stage IV involves a fast, super-stoichiometric proton uptake and direct protonation of QB via Asp-L213 and Ser-L223 forming the \( PQ_A^-Q_BH^- \) state. This is followed by stage V in which a second, slow, electron transfer from QA occurs forming the \( PQ_A^-Q_BH^- \) state. In stage VI the doubly reduced \( Q_BH^- \) is protonated via Glu-L212 forming the \( PQ_A^-Q_BH_2 \) state. The fully reduced and protonated quinol \( QBH_2 \) can now leave the binding pocket and be replaced by a fully oxidized ubiquinone molecule from the membrane pool in stage VII (92, 94).
Figure 9. Binding sites for the $Q_A$ and $Q_B$ ubiquinones in the photosynthetic reaction center from R. sphaeroides. $Q_A$ and $Q_B$ is linked by hydrogen bonds through a His-Fe$^{2+}$-His complex. The non-heme Fe$^{2+}$ ion shown in orange is not redox active during electron transfer from $Q_A^-$ to $Q_B$ (97). Also shown are residues in the proposed proton transfer pathway to $Q_B$. His-H126 and His-H128 are situated on the surface of the reaction center and may function as a local proton buffer to keep proton transfer to $Q_B$ non-rate limiting (92). Protonation of reduced $Q_B$ occur in the first step via Asp-L213 and Ser-L223 and in the second step via Glu-L212 and perhaps via one or more water molecules (92). Side-chain carbons of hydrophobic residues are shown in yellow, acidic residues in blue and basic and polar residues are shown in grey. Coordinates from 1DV3.pdb.
3.7. Electrochemical characterization of quinones in proteins

Since quinones are involved in many biological redox processes, the solvated quinone/quinol redox couple is a very well-studied system electrochemically (91). The para-quinone/para-quinol pair (Figure 6) is possibly the most studied of all organic redox couples by electrochemical methods. In contrast to the numerous articles on the redox behavior of quinone cofactors free in solution (98-101) or covalently attached to the working electrode (102, 103), reports on the redox behavior of quinones in proteins are scarce. Quinone-containing enzymes are often large, multi-subunit and multi-cofactor complexes and, in addition, several of them are membrane bound. These factors have hampered electrochemical characterization and only a few systems have been studied by other electrochemical methods than simple redox potentiometry. A weak electrochemical response from the CTQ cofactor (Figure 5D) in P. denitrificans amine dehydrogenase was detected when using standard cyclic voltammetry (CV). Reversible CV and square wave voltammograms exhibiting resolved waveforms of two bacteriochlorophyll species and QA cofactors in the R. sphaeroides RC were recorded by Kong et al. (104). Two types of protein film samples (i.e. protein samples deposited directly on the surface of the working electrode) were used to achieve rapid electron transfer between the electrode and the protein cofactors. The RC complex was either incorporated into multiple lipid bilayers or positioned between layers of polyanion molecules (104). Gray and coworkers synthesized a diethylaniline-terminated oligo-(phenyl-ethynyl)-thiol “electron wire” with the aim to electronically connect to the sequestered TPQ cofactor (Figure 5A) in the active site of A. globiformis amine oxidase. The electron wire molecule was covalently attached to the surface of the working electrode and then bound to the protein via its 20 Å hydrophobic substrate channel. Rapid electron tunneling between the buried cofactor and the working electrode was successfully obtained by this method (105). Haehnel and coworkers covalently attached ubi- and menaquinones to a template-assembled four-helix bundle. Characterization of the de novo quinone proteins included determination of reduction potentials by cyclic voltammetry (106).

3.8. Significance of thesis project, part II

Quinones are important cofactors in enzymes and redox proteins and voltammetry-based methods represent a key tool to study the properties of these cofactors as a function of the protein environment. As described above, the size and complexity of most quinone proteins preclude however voltammetry characterization. The second motivation for this thesis project was to construct a model quinone protein which redox properties could be studied by voltammetry. These efforts are described in Paper V and discussed in Section 5.
4. De novo protein design

4.1. What is de novo protein design?
The construction of proteins from scratch, de novo protein design (107), represents a model system approach to explore protein folding, protein/cofactor interactions, and protein-based catalysis. Various methods including iterative rational design, template-assembled strategies, combinatorial libraries, and computational methods have been used to create α-helical, β-sheet and mixed α/β-protein folds, cofactor- and metal-containing designs, and systems exhibiting some degree of catalytic activity.

4.2. Protein sequence space and the inverse folding problem
De novo protein design is ultimately about finding amino-acid sequences that fold into structures with desired properties, a problem sometimes referred to as “the inverse folding problem” (108). The sequence space, the number of different possible sequence combinations, is enormous even for a very small protein. For example, a protein of 50 residues, in which all 20 naturally occurring amino acids are permitted at every position, have \(20^{50} \approx 10^{65}\) different amino-acid sequence combinations. These sequences would weigh in at \(10^{42}\) kilograms which is about \(10^{17}\) times the mass of the Earth (109). Given the vastness of protein sequence space it is easy to see that the optimal sequence for a given application will never be found using a totally random search. A number of complementary methods exists to explore the sequence space more efficiently and to enhance the probability of a successful de novo design.

Rational protein design is based on selecting specific folding patterns and/or known stabilizing features, e.g. propensities of the amino acids to stabilize or destabilize specific secondary structures, construct typically a small number of variants containing these features, and then examine their physical properties to connect the design with the product. The strategy of template-assembled methods is to reduce the complexity of the system by connecting peptide chains to a template, e.g. a metal or a small circular peptide. In combinatorial protein design, randomness and elements of rational design are combined to make libraries of de novo proteins. The libraries are then screened to select molecules with the desired properties. Finally, computational protein design algorithms search through sequence space to find low energy sequences that satisfy the target fold. Advances using these different methods are briefly summarized below.
4.3. Design of coiled coils and helical bundles – an example of iterative rational design

The coiled coil is a widespread structural motif found in a variety of proteins including transcription factors, chaperones, and structural proteins such as keratin (110). It is a strikingly simple motif consisting of α-helices wrapped around each other in a highly organized fashion. The oligomeric state of coiled coils can vary from 2 to 5 and the helices can pack in either a parallel or an anti-parallel manner. The contributing helices may come from a single peptide chain or from separate chains in an oligomeric protein. Most coiled coils consist of interacting amphipathic helices of which the amino acids are organized in a characteristic heptad (7-residue) repeat pattern (Figure 10). The heptad repeat pattern was first discovered in tropomycin (111, 112). Tropomycin is a parallel two-stranded coiled coil involved in the contraction and relaxation of skeletal muscles.

![Figure 10. Interactions between residues in a dimeric coiled coil. The positions of the amino acids in a heptad segment are denoted abcd for (111). Residues a and d are normally non-polar and their side chains pack against each other forming the main part of the hydrophobic core in the coiled-coil structure. Residues e and g are often charged and may be involved in inter-helical salt bridges. Residues in the remaining positions are usually polar and form the surface of the coiled coil.](image)

The first de novo designed coiled coils were constructed in 1981 by Hodges and coworkers (113). Since then, many groups have used the heptad-repeat pattern to construct coiled coils and helical bundles (114). The construction of the α3D three-helical bundle by DeGrado and coworkers (115) provides an excellent example of iterative rational design and highlights several important features for the successful design of a helical bundle including: I) the importance of loop regions to promote a monomeric structure, II) the introduction of helix capping interactions to stabilize helix formation, III) the addition of features of “negative design” to destabilize unwanted topologies, and IV) the importance of a well-defined hydrophobic core (115). An early variant of α3D suffered from an ill-defined hydrophobic core made up of all Leu residues (115). Leucine is a big and bulky residue and has many possible rotamers close in energy and an overabundance of Leu may result in a dynamic and/or “over-packed” core (116, 117). The introduction of more sterically constrained aromatic and β-branched amino acids at
core positions tend to make *de novo* proteins less dynamic and more “native-like” (118, 119). The NMR solution structure of the repacked α3D protein was found to be in good agreement with the original design (120) although, interestingly, substitutional (121) and NMR side-chain dynamics (122) studies on α3D showed a more dynamic behavior at core positions relative to natural proteins studied at this level of detail.

### 4.4. Template-assembled synthetic proteins

The unfavorable loss of main-chain and side-chain entropy occurring upon protein folding increases with the number of possible unfolded conformations. Consequently, the loss in peptide entropy upon folding will be less if the number of unfolded conformations for a given peptide chain is reduced. In template-assembled protein design, peptides are attached to a rigid molecular scaffold with the aim to reduce the size of the ensemble of unfolded conformations. The peptides are pre-ordered on the template prior to folding and the entropic cost of folding the peptide chain is thereby lessened. The templates varies and include peptides, carbohydrates, porphyrin derivatives, steroids, transition metals, and cavitands (rigid organic molecules with an internal cavity) (123, 124). The vast majority of template-assembled *de novo* proteins are α-helical with 3-6 interacting helices. The helices can have identical or different sequences and they can be oriented in a parallel or anti-parallel fashion. Rational, combinatorial and computational methods have been used to generate this type of synthetic proteins (123, 124). Haehnel and coworkers have used the template-assembled synthetic protein methodology to design, synthesize and characterize a number of *de novo* proteins including heme proteins (125), heme proteins with catalytic activity (126), chlorophyll binding proteins (127), quinoproteins (106), and proteins with copper-binding sites (128).

### 4.5. Designed combinatorial libraries of *de novo* proteins

Hecht and coworkers have constructed combinatorial libraries of *de novo* four-helix bundles using a simple binary pattern. The binary pattern approach is based on the assumption that the ability of a peptide chain to form amphipatic secondary structure elements is sufficient to promote the chain to fold into a globular structure. The binary pattern specifies the order of hydrophilic and hydrophobic residues in the peptide sequence but the identity of the amino acid at each site is allowed to vary combinatorially (129). Libraries of synthetic genes were constructed based on the binary pattern and the organization of the genetic code that allowed for the combinatorial diversity. Protein expression tests conducted on the first-generation library showed that about 60 % of randomly chosen clones generated soluble proteins. Some of the expressed proteins showed native-like characteristics, but most of them displayed properties consistent with fluctuating structures (130, 131). A second-generation library was created by redesigning one of the non-native like proteins characterized from the original library. The heli-
ces were elongated and several elements of rational design were allowed to influence the redesign (132). The NMR solution structure of a well-folded de novo four-helix bundle protein from this second-generation library has been presented (133).

4.6. Computational protein design

With more powerful computers and efficient algorithms, computational design strategies have gained in popularity. In silico de novo protein-design methods have the potential to explore the protein sequence space more extensively than experimentally based methods. However, a full de novo design of even a small 100-residue protein remains a considerable challenge (134). Even if only the most preferred amino-acid rotamers are taken into account, the complexity of the computational problem quickly grows into overwhelming proportions. For example, a 100-residue protein for which the ~100 most preferred side-chain conformations of the 20 naturally occurring amino acids are considered at each position, have in total $100^{100}=10^{200}$ possible conformations (134). In addition, natural proteins typically have a structural flexibility that allows the back-bone to relax upon mutations. Freezing the back-bone position during the calculations reduces the computational complexity greatly, but this approach has its limitations in de novo design since the starting structure by definition is not well defined (134).

The ORBIT program by Mayo and coworkers is constructed to optimize the sequence that will stabilize a fixed template structure (135). The ORBIT methodology has been used for the full sequence design of a 28-residue zinc finger motif (135), redesign of calmodulin to improve the binding of one of its target peptides, redesign of 24 core positions in T4 lysozyme, and in an attempt to design a α/β-barrel with an “idealized” artificial back-bone fold (134).

The RosettaDesign software, developed by Baker, Kuhlman and coworkers (136), uses an iterative design approach and alternates between atomic-resolution structure prediction and Monte Carlo-based sequence optimization. Back-bone flexibility is incorporated by considering a set of initially fixed back-bone templates (134). The Rosetta design methodology has been used for protein structure prediction (137), to redesign and stabilize natural proteins (138, 139), for the design of a peptide sequence that can switch between a trimeric coiled coil and a zinc finger-like fold (140), and for the de novo design of Top7 a protein with a novel fold (136). Recently, Rosetta was used to design novel catalysts for a retro-aldol reaction in which a carbon bond is broken in an unnatural substrate. 44% of the 72 experimentally characterized de novo proteins displayed retro-aldolase activity with rate enhancement up to four orders of magnitude as compared to the uncatalyzed reaction (141). Two of the active enzymes, which were designed to display different protein folds and different active-site architectures, were crystallized. In both cases the crystal structures were in good agreement with the design models (141).
4.7. Significance of thesis project, part III

Our group has previously utilized *de novo* design to construct a stable and well-structured single-chain three-helical bundle. This system has been used to study amino-acid redox chemistry (Papers III and IV) and quinone redox chemistry (Paper V) occurring within a protein environment. This thesis describes the characterization of a single-chain four-helix bundle protein that was designed using the methods outlined in Section 4.3. The goal was to construct a larger protein scaffold that will be used in future studies similar to those performed with the three-helix bundle system. These efforts are described in Papers I and II and discussed in Section 5.
5. Results and discussion

Making a single-chain four-helix bundle for redox chemistry studies (Papers I-II)

The three-helix bundle system used in Papers III-V exhibits some key features. The protein scaffold is well structured and, in addition, uncomplicated with respect to its optical properties. Earlier work and data presented here demonstrate that NMR as well as UV/Vis, fluorescence and CD spectroscopy can be used to derive detailed structural information of the protein as a whole (142, 143) and of the radical/quinone site specifically (e.g. Paper IV). The 67-residue, single-chain protein scaffold is “redox inert” except for a dedicated site, position 32, in the core of the protein (Figure 6, Paper I). We define “redox inert” as tolerating voltammetry measurements in the oxidative range of up to +1.3 V vs. NHE without triggering uncontrolled redox reactions. The inert nature of the scaffold is essential as it allows us to probe the redox properties of highly oxidizing molecules without triggering unwanted side reactions. Thus far three main variants have been developed in which position 32 is either occupied by a tryptophan (\(\alpha_3W\)), a tyrosine (\(\alpha_3Y\)) or a cysteine (\(\alpha_3C\)). In the latter, Cys-32 is used as a tag to ligate phenol (Paper III) or quinone (Paper V) molecules to the protein.

A valid criticism of the three-helix bundle scaffold is that it is small in size relative to naturally occurring radical and quinone-containing enzymes. In addition, its structural architecture is simple which can provide limitations (see Discussion Paper IV). The goal with this project was to create a protein of roughly twice the molecular mass of \(\alpha_3W\) (7.5 kDa) but otherwise exhibiting the same overall properties as the smaller system. That is, the larger protein should be well structured. It should have a redox inert sequence except for a single site containing a redox-active molecule, which, in turn, should give rise to a Faradaic response in voltammetry investigations. The initial design of this larger protein, a single chain four-helix bundle with a molecule mass of about 13 kDa, was first introduced in Paper I. The refined design and the characteristics of the final protein are described in Paper II.

In contrast to the de novo \(\alpha_3W\) protein, a natural protein was used as a starting point for the four-helix bundle design. The sequence of the E. coli DNA-binding Rop protein was chosen as the template. Thus this project does not represent de novo design, but rather redesign of a natural protein. The redesign of Rop followed classical iterative design principles and involved multiple changes in the primary sequence with large-scale changes in inter-helical interactions and bundle topology as a result.

In solution, Rop is a homo dimeric four-helix bundle with its two helix-turn-helix domains arranged in an anti-parallel manner. The sequences of the helical regions in Rop follow the heptad repeat pattern described in Section 4.3. Wild-type Rop has eight hydrophobic packing layers with small residues at heptad \(a\) positions and large residues at \(d\) positions with the exception of...
the equivalent layers 2 and 2' (see Figure 1, Paper II) in which the small versus large residue packing pattern is reversed.

The anti-parallel topology of Rop places the C- and N-termini of each monomer on opposite sides of the protein. In our redesign, 20 out of the 32 heptad a and d core residues were changed following a small versus large packing pattern to induce and stabilize a \( \sim 180^\circ \) flip of one of the subunits (illustrated schematically in Supplementary Figure S1, Paper II). The C- and N-termini of the two subunits were linked by the introduction of a five-residue glycine loop and multiple inter-helical sites were changed to avoid potentially unfavorable electrostatic interactions. The single-chain sequence was made to contain no tyrosine, tryptophan or cysteine residues except for a unique tryptophan at heptad a position 106 (Figure 1, Paper II).

Characterization of the first protein product revealed a predominately helical structure (71.0 ± 0.6% helical, pH 5.5-10.0) with a poor global stability (\( \sim –2.7 ± 0.5 \) kcal mol\(^{-1}\)). The protein appears monomeric at low protein concentrations (\( \leq 100 \) \( \mu \)M) but broadened NMR spectral lines suggest that aggregation occurs at higher concentrations (Figure 4A, Paper II). The subsequent redesign process was guided by two main goals: To improve the protein stability (which hopefully would resolve the aggregation problem) and to make the protein more suitable for NMR structural studies by increasing the overall chemical-shift dispersion. Two Thr (T19 and T80) were replaced by apolar amino acids and two Phe were placed in predicted core positions (T80F and L109F). The introduction of aromatic side chains was predicted to provide ring current shifts and influence the chemical shifts of nearby atoms. The changes were made in a stepwise fashion and the physical properties of each variant were investigated (Table I, Paper II).

The final protein is a three-site variant (T19I, T80F, L109F) of the initial design. This protein forms a stable helical structure in water. Its free energy of unfolding is \(-4.7 \) kcal/mol (25° C, pH 7.2) and the protein was shown to be thermostable with an estimated midpoint temperature of unfolding above 355 K. The helical content is high and displays little variance over the 5.5 – 10.0 pH range (69.8 ± 0.8 %). Analytical ultracentrifugation and NMR data showed that the protein is monomeric up to at least 0.5 mM. 1D proton and 2D \(^{15}\)N-HSQC data display narrow spectral lines consistent with a uniquely structured protein. The spectral resolution of the 2D \(^{15}\)N-HSQC spectrum (Figure 4C, Paper II) is excellent considering the all-helical nature of the protein and suggests that resonance assignment of the protein backbone should be straightforward. The blue-shifted fluorescence spectrum of Trp-106 suggests that the aromatic residue is buried, as indented. Finally, a differential pulse voltammetry investigation showed a Faradaic response of a single proton-coupled oxidation reaction consistent with the unique Trp-106.

In conclusion, the efforts described in Papers I and II resulted in the extension of our model protein library with a stable and well-structured 13 kDa radical protein. This protein is denoted \( \alpha_4W \) to reflect that it is a single chain four-helix bundle with a unique tryptophan. It is our hope that the \( \alpha_4W \) protein can be explored in a manner similar to the \( \alpha_3W \) system and that new information can be gained by comparing the two.
Moving a phenol hydroxyl group from the surface to the interior of a protein: Effects on the phenol potential and $pK_A$ (Paper III)

Despite the number and biochemical importance of enzymes utilizing tyrosine redox chemistry, experimental studies characterizing the thermodynamic properties of tyrosyl radicals are scarce, essentially nonexistent, in the literature (see Table 1 and discussion in Section 2.6.). Since very little is known, virtually any aspect of “tyrosine redox properties vs. the protein environment” is important to strive to understand. The work described in Paper III aims to address a very basic issue: How does the degree of solvent exposure of the phenol hydroxyl group influences the redox characteristics of a tyrosine? Experimentally we wanted to gradually alter the solvent exposure of the phenol hydroxyl group but without introducing large-scale changes in the overall environment of the aromatic head group. A “phenol rotation strategy” was developed that consisted of ligating phenols containing a SH tag in the *para*, *meta* and *ortho* position (site 4, 3 and 2 relative to the phenol hydroxyl site 1) to the unique cysteine in $\alpha_3C$.

![Figure 11. Chemical structures of (A) 4-mercaptophenol, (B) 3-mercaptophenol, and (C) 2-mercaptophenol attached via an S–S bond to a cysteine. (D) Tyrosine is also included to provide a structural comparison. R’ and R” represents the peptide back-bone in the $\alpha_3C$ systems. For the solution reference compounds, R’ and R” represents –COCH$_3$ and –OCH$_3$, respectively.](image)

Trp-32 in $\alpha_3W$ has an accessible surface area of < 3% (143) and the smaller side chain of the corresponding Cys-32 residue in $\alpha_3C$ is therefore predicted to be buried within the hydrophobic core. The tyrosine analogues 4-mercaptophenol (4MP), 3-mercaptophenol (3MP) and 2-mercaptophenol (2MP) were covalently attached to $\alpha_3C$ via a disulfide bond. Chemical representations of the three Cys-reacted phenols are shown in Figure 11. The simple heptad design of the three-helix bundle, which was confirmed by the structural characterization of $\alpha_3W$ (Figure 6, Paper I), predicts that the protein scaffold will force the geometry of the bound mercaptophenols rather than the other way around. That is, we anticipated that binding a mercaptophenols to Cys-32 would not give rise to large-scale shifts in the position of the cysteine and that the phenol OH group of 4MP-$\alpha_3C$ would be more exposed to the bulk solvent relative to the phenol OH group in the 3MP-$\alpha_3C$ and 2MP-$\alpha_3C$ proteins. The aim of the “phenol rotation strategy” is shown in the second Figure in Paper III. We note that the purpose of Figure 2 is only to illustrate the design and not to make detailed structural predictions.
Paper III describes the synthesis and purification of the three MP-α₃C proteins as well as characterization of their properties relative to α₃C and α₃W. It was shown that homogenous MP-α₃C protein samples could be obtained. UV/Vis spectra and extinction coefficients of the protein-bound phenols were measured and described. The MP-α₃C proteins were shown to be stable and highly helical. These results are summarized in Table 1 in Paper III. One-dimensional NMR spectra of the MP-α₃C proteins display chemical shift dispersion in the amide region that is similar to the α₃W 1D NMR spectrum. This indicates that no catastrophic changes, i.e. conversion from a structured scaffold to a molten globule, occur upon binding of the phenols. This conclusion has more recently been confirmed by multi-dimensional NMR data, which show characteristics fully consistent with well-structured proteins (Tommos et al. unpublished).

Two data sets suggest that the phenol OH group of 4MP-α₃C is significantly more solvent exposed relative to the OH group of the phenols in the 3MP-α₃C and 2MP-α₃C proteins. Firstly, the global stability of 3MP-α₃C and 2MP-α₃C (–3.1 kcal mol⁻¹) is lower than the stability of 4MP-α₃C (–3.7 kcal mol⁻¹) and α₃C (–4.1 kcal mol⁻¹). The core of α₃C is made of Leu, Ile and Val residues and solvating a hydrophilic OH group within this highly hydrophobic environment is expected to have a destabilizing effect on the protein. Secondly, the CD spectrum of 4MP-α₃C is virtually identical at pH 7.6 and at pH 10 (Figure 8, Paper III). The phenol pKₐ of 4MP-α₃C is about 9.5, which means that at pH 7.6 the phenol is predominately protonated and at pH 10 it is predominately deprotonated. To form a charge inside the protein is expected to have a significant impact on the secondary structures of the protein and the fact that there is no change at all strongly suggests that the OH group resides at or near the surface of the protein.

The last part of Paper III describes an electrochemical characterization of the MP-α₃C proteins. In order to isolate the effects of the protein matrix on the redox properties of the bound phenols, all electrochemical data derived from the MP-α₃C proteins were compared to data derived from aqueous reference compounds. Thus, we are not interested in the absolute phenol potentials but rather on the shifts introduced as the phenols are bound to the protein relative to being freely solvated in aqueous buffer. The reference compounds were made by reacting the three mercaptophenols with the blocked, neutral Cys derivate N-acetyl-L-cysteine methyl ester (bCys). The preparation and purification of these compounds are described in Paper III.

Differential pulse voltammetry (DPV) measurements have been used previously to estimate the reduction potentials and pKₐ values of tyrosine and tryptophan in solution and in α₃W and α₃Y (142). The average width of the DPV peaks representing the MP-α₃C proteins and the MP-bCys compounds were broader than expected for an n = 1 reversible reaction (144), suggesting irreversible redox chemistry. Cyclic voltammetry traces of 2MP-α₃C and 2MP-bCys confirmed this conclusion (not shown). Irreversible electrochemistry is typical for organic radicals, since these species usually have life times that are short relative to the time scale of the experiment. For a fully reversible redox reaction with no kinetic complications, the average DPV...
peak potential approximates the formal reduction potential \( (E^0') \) (145). \( E_{\text{peak}} \) values derived by voltammetry from aqueous tyrosine and tryptophan are very close to \( E^0' \) values derived from pulse radiolysis equilibrium studies. We estimate that \( E_{\text{peak}} = E^0' \pm 0.02 \) V for the freely solvated species (142). For the protein samples it is more difficult to estimate a direct relationship between \( E_{\text{peak}} \) and the true reduction potential, although we note that scan rates were deliberately kept low in order to avoid any kinetic effects on \( E_{\text{peak}} \). At this point, we assume that the model proteins behave similarly to the solvated compounds and that their \( E_{\text{peak}} \) values can be compared directly.

Figure 12. Redox states of phenol (PhOH) in aqueous solution as a function of pH. (A) At a solution pH below the \( pK_A \) for the oxidized species \( (pK_{\text{OX}}) \) the potential becomes pH-independent and \( \text{PhOH}^+/\text{PhOH} \) is the predominant redox couple. (B) Between \( pK_{\text{OX}} \) and \( pK_{\text{RED}} \) the redox reaction is pH dependent with a slope of \(-\ln(10)RT/nF\) (59 mV per pH unit for a one electron, one proton redox event at 25 °C). The major redox couple in this region is \( \text{PhO}^-/\text{PhOH} \). (C) At a pH higher than the \( pK_A \) of the reduced species \( (pK_{\text{RED}}) \) the redox reaction becomes pH independent again with the \( \text{PhO}^+/\text{PhO}^- \) as the predominant redox couple.

The redox reactions of a phenol are dependent on the solution pH, as illustrated in Figure 12. DPV measurements were performed in the pH 4 to 10 range to investigate what effect the protein matrix has on the potential and \( pK_A \) values of the bound mercaptophenols. \( E_{\text{peak}} \) vs. pH plots for 2MP, 3MP, and 4MP ligated to \( \alpha_3 \)C and to bCys are shown in Figure 7, Paper III, and the results from these data sets are summarized in Table 2 below. The \( pK_A \) values of the phenolate/phenol couples were determined electrochemically (\( pK_{\text{RED}} \)) and optically via the absorption of the phenolate (\( pK_A \)). The values derived by these two methods were found to be in good agreement.

<table>
<thead>
<tr>
<th>Species</th>
<th>( pK_A )</th>
<th>( pK_{\text{RED}} )</th>
<th>-( \ln(10)RT/nF )</th>
<th>( E_{\text{peak},7} )</th>
<th>( \Delta E_{\text{peak},7} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4MP-bCys</td>
<td>9.3</td>
<td>9.7</td>
<td>0.059</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>4MP-( \alpha_3 )C</td>
<td>9.5</td>
<td>9.4</td>
<td>0.064</td>
<td>0.79</td>
<td>-0.01</td>
</tr>
<tr>
<td>3MP-bCys</td>
<td>9.6</td>
<td>9.8</td>
<td>0.059</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>3MP-( \alpha_3 )C</td>
<td>&gt;10</td>
<td>&gt;10</td>
<td>0.058</td>
<td>1.07</td>
<td>0.11</td>
</tr>
<tr>
<td>2MP-bCys</td>
<td>9.0</td>
<td>8.8</td>
<td>0.060</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>2MP-( \alpha_3 )C</td>
<td>&gt;10</td>
<td>&gt;10</td>
<td>0.057</td>
<td>0.91</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 2. Electrochemical properties of 4MP, 3MP, and 2MP bound to \( \alpha_3 \)C and bCys. All potentials are in the unit of V and are given vs. NHE.
Our study of the MP-α₃C proteins and their corresponding solution reference compounds provided two main results. As described above, the characterization of 4MP-α₃C suggests that the protein folds around the hydrophobic part of the ligated phenol and that the phenol OH group resides at or near the protein surface. Clearly the protein-ligated phenol resides in a very different environment relative to the solvated species (schematically illustrated in Figure 13) and yet their redox properties are remarkably similar (Figure 7, Paper III; Table 2). The $E_{\text{peak}}$ (pH 7.0) values of 4MP-α₃C and 4MP-bCys are within the error margin of the experiment. The average p$K_A$ values derived from the electrochemical and phenolate titrations are 9.45 and 9.50 ± 0.2 for 4MP-α₃C and 4MP-bCys, respectively. In all its simplicity, this is a surprising and important finding that represents the first experimental demonstration that the interactions with the OH group dominate the redox characteristics of the PhO•/PhOH redox couple and that interactions with the remaining part of the aromatic molecule have no significant impact.

**Figure 13.** Schematic representation of the environment for 4MP when (A) solvated in water and (B) attached to the α₃C protein. The aqueous compound experiences a homogenous environment with a high dielectric constant. The protein-bound species, in contrast, resides in a highly heterogeneous environment as it extends from the hydrophobic core (marked in dark blue), through the charged surface of the protein (red) and into the bulk phase (light blue). That the redox properties of 4MP in Figures A and B are essentially identical, indicate that interactions to the phenol OH group are the key parameters tuning the phenol reduction potential.

In contrast to 4MP, the electrochemical characterization of 3MP and 2MP show distinct differences when the phenols are ligated to α₃C relative to bCys. The average p$K_A$ values for 3MP-bCys and 2MP-bCys are 9.7 ± 0.1 and 8.9 ± 0.1, respectively. These values increase to >10 when the phenols are ligated to α₃C. The increase in the p$K_A$ values observed for 3MP-α₃C and 2MP-α₃C relative to the solution compound suggests that the phenol OH group is more shielded in the protein. $E_{\text{peak}}$ (pH 7.0) is elevated by 0.11 and 0.12 V for 3MP-α₃C and 2MP-α₃C, respectively, relative to the solution systems. The $E_{\text{peak}}$ vs. pH plots for 3MP-α₃C and 2MP-α₃C display slopes close to 59 mV per pH unit over the whole pH 4 to 10 pH range. This is also true for 3MP-and 2MP-bCys at a pH < p$K_A$ and show that PhO*/PhOH is the relevant redox couple in this region. Factors that may influence the potential of the PhO*/PhOH couple are hydrogen-bonding interactions of the phenol and/or the phenoxyl radical, the fate of the phenolic proton and its associated charge upon forming the neutral radical, and potential redox induced structural changes. The ~59 mV/pH slopes show that the MP-α₃C proteins remain
overall charge neutral upon oxidation suggesting that charge is not involved in raising the potentials of 3MP-α3C and 2MP-α3C. As discussed at some length on page 11899, Paper III, it is doubtful that structural changes are involved as well. Most likely differences in hydrogen bonding between the aqueous and protein-bound 3MP and 2MP molecules is the main parameter influencing the potential of the PhO+/PhOH redox couple.

To summarize, Paper III describes two interesting observations. Firstly, the potential of the PhO+/PhOH redox pair is dominated by interactions with the phenol OH group while interactions with the hydrophobic part of the aromatic ring appears to have little or no impact. Secondly, we propose that hydrogen-bonding interactions between the 3MP and 2MP species and the protein matrix either stabilize the reduced phenol state and/or destabilize the radical state, relative to the solvated compounds. Although not part of this thesis project, NMR structural characterization of the three MP-α3C proteins is in progress to investigate this further (Tommos et al. unpublished).

Probing proton-coupled electron transfer in model tyrosyl radical proteins (Paper IV)

Oxidation of a protonated tyrosine will generate the neutral radical state in an aqueous milieu. This is the result of the very low pK_{OX} value of tyrosine (<0; (146)). Here we aim to understand how the chemical nature of the acceptor(s) of the phenolic proton influences the thermodynamic properties of tyrosine. The project is based experimentally on the α3Y model protein, which sequence contains a single tyrosine at position 32 and originally no histidine residues. The goal is to engineer a histidine side chain in close vicinity of the Tyr-32 phenolic oxygen so that the imidazole ring acts as a proton acceptor or part of a proton-acceptor network involved in the proton-transfer reactions coupled to the electrochemical oxidation of Tyr-32. Histidine was chosen as the primary target for the α3Y re-engineering project since work on natural systems has shown that tyrosine/histidine interactions can have significant influences on tyrosine oxidation/reduction rates and radical yield ((147); see also page 108 in Paper I).

Overall the project involves an initial modeling step to identify a set of single-site α3Y/histidine variants in which the imidazole side chain is predicted to reside close to the phenolic oxygen of Tyr-32, the generation and purification of these proteins followed by spectroscopic and structural screening studies to probe for Tyr-32/histidine interactions. The proteins that pass these screening steps are studied electrochemically to further characterize the nature of the Tyr-32/histidine interactions and to probe for potential effects on the tyrosine redox properties. The major part of this work, including a detailed spectroscopic and structural characterization of α3Y and histidine variants has been completed. These results are shown in the manuscript labeled Paper IV in this thesis. The electrochemical characterization has more recently been conducted and detailed analyses of the obtained data sets have yet to be finalized. A summary of the results obtained thus far and a preliminary statement of the electrochemical studies is provided below.
To identify potential histidine sites in $\alpha_3Y$, a model of this protein was made from the $\alpha_3W$ NMR structure (Figure 1 Paper IV). Tyr-32 rotamers were visually inspected to identify residues within 5 Å of the phenolic oxygen. Histidine was modeled into each identified site and the Tyr-32 rotamers were paired with all available histidine rotamers to get a rough idea of relative ring geometries and distances. This study suggested the following nine sites for histidine incorporation: 9, 12, 13, 29, 33, 36, 58, 61 and 62.

The modeling was coarse and not intended to make detailed predictions but rather our re-engineering approach relied on a subsequent spectroscopic screening step to identify $\alpha_3Y$/His variants of interest. The $\alpha_3Y$-V9H, $\alpha_3Y$-L12H, $\alpha_3Y$-E13H, $\alpha_3Y$-K29H, $\alpha_3Y$-E33H, $\alpha_3Y$-K36H, $\alpha_3Y$-L58H, $\alpha_3Y$-E61H and $\alpha_3Y$-I62H proteins were made and purified. We take advantage of the sensitivity in the electronic absorption and emission properties of tyrosine as a function of its environment in the characterization of these proteins. Absorption and fluorescence spectra were obtained from the nine $\alpha_3Y$/His variants and compared to corresponding spectra obtained from $\alpha_3Y$. It is well known that the absorption energy of phenolic compounds shifts in response to the dielectric medium and hydrogen-bonding interactions ($^{142, 148}$). This effect is evident when comparing the blueshifted absorption spectrum of aqueous N-acetyl-L-tyrosinamide (NAYA; $\lambda_{\text{max}}$ 275.3 nm) relative to the $\alpha_3Y$ spectrum ($\lambda_{\text{max}}$ 277.8 nm) (Figure 2A Paper IV). Shifts in absorption maximum, interestingly both to the blue and to the red relative to the $\alpha_3Y$ absorption maximum, were detected for a main fraction of the $\alpha_3Y$/His proteins (e.g. Figure 2B Paper IV) indicating changes in the microenvironment of Tyr-32 in these systems. A more sensitive structural probe is provided by the fluorescence spectrum of tyrosine. For example, aqueous NAYA has an emission maximum at 302 nm, which redshifts to 319 and 335 nm upon hydrogen bonding to acetate and imidazole, respectively ($^{149}$). The excitation spectra of the hydrogen-bonded complexes are ~2 nm redshifted relative to the excitation spectrum of aqueous NAYA ($^{149}$). In contrast, upon forming a hydrogen-bond to an amide C=O group the NAYA emission remains unchanged while the excitation spectrum shifts up to 10 nm ($^{148}$). These results are particularly relevant since the available hydrogen-bonding partners to reduced Tyr-32 are water, the NH and C=O groups of the protein backbone, glutamate, and the introduced histidine. The fluorescence of $\alpha_3Y$-E13H, $\alpha_3Y$-E33H, $\alpha_3Y$-E58H, and $\alpha_3Y$-E61H look essentially identical to the fluorescence of $\alpha_3Y$ (e.g. Figure 2D). In contrast, the excitation and emission spectra of $\alpha_3Y$-V9H, $\alpha_3Y$-L12H, $\alpha_3Y$-K29H, $\alpha_3Y$-K36H and $\alpha_3Y$-I62H are shifted relative to the fluorescence of $\alpha_3Y$ (e.g. Figures 2C & E), indicating changes in the environment of Tyr-32 in these proteins.

Five of the nine proteins passed the initial spectroscopic screening. The next step was to determine whether the observed spectral shifts are due to a global perturbation of the protein scaffold rather than to a local change in the vicinity of Tyr-32. The results from the structural screening are displayed in Figures 3 and 4, and summarized in Table I in Paper IV. Briefly, it was determined that the $\alpha$-helical contents and the global stability of $\alpha_3Y$-L12H and $\alpha_3Y$-I62H (Figure 3B) were significantly lower relative to $\alpha_3Y$, $\alpha_3Y$-
V9H, α3Y-K29H and α3Y-K36H (Figure 3A). The pK\textsubscript{app} values of Tyr-32 and the introduced histidine were determined by pH titrations monitored optically (tyrosinate absorbance at 293 nm) or by NMR (chemical shift of the imidazole C2 proton). Significantly broadened NMR spectral lines were observed for α3Y-L12H and α3Y-I62H (not shown) consistent with poorly structured proteins. The NMR pH titration of α3Y-V9H showed that the pK\textsubscript{app} of His-9 is significantly lower (< 5) relative to aqueous histidine (6.0) and that the protonation of this residue is tightly coupled to the unfolding of the protein (not shown). A CD study monitoring the degree of α-helical content as a function of pH confirmed this conclusion (Figure 4B). The importance of this observation is that the initial selection of α3Y-V9H was based on optical data obtained at pH 5. However, the spectral shifts observed are most likely due to global unfolding rather than to a specific change around Tyr-32. On the basis of these results, α3Y-V9H, α3Y-L12H and α3Y-I62H were excluded from further characterizations.

A more refined fluorescence study could be conducted on α3Y-K29H and α3Y-K36H following the characterization of their protein properties. Fluorescence spectra of Tyr-32 were obtained ± 2 pH units around the pK\textsubscript{app} of the introduced histidine in these two proteins (Figures 5B & C Paper IV) and corresponding control spectra were obtained from α3Y (Figure 5A). The intensity of the Tyr-32 emission spectrum is different in the two histidine-containing proteins relative to α3Y and, in addition, pH-induced spectral changes are significantly more pronounced for the histidine variants. The low and high pH spectra obtained from α3Y-K29H are also significantly different from those obtained from α3Y-K36H both with respect to intensity (Figure 5B vs. 5C) and emission maximum (Figure 5D). These results show that the local environment of Tyr-32 is different in the three proteins.

A few clarifications should be made to strengthen this statement. Firstly, the fluorescence spectra were obtained ≥ 1.7 pH units below the pK\textsubscript{app} of Tyr-32 (Figure 4 & Table I) to avoid spectral contributions from tyrosinate fluorescence. For α3Y-K29H, spectra were obtained both a pH 8.7 (2 pH units below pK\textsubscript{app} of Tyr-32) and at pH 9.0 (2 pH units above pK\textsubscript{app} of His-29) to check that there were essentially identical. Secondly, changes in the α-helical content were measured by CD spectroscopy in the pH region ± 2 pH units around the pK\textsubscript{app} value of the histidine in α3Y-K29H and α3Y-K36H. α3Y-K29H and α3Y-K36H are 66.8 ± 2.5 and 64.7 ± 0.8 percent helical between pH 5.1 and 9.1, respectively. From these results we conclude that pH-induced spectral changes arising from global unfolding are minimal in this pH range.

Fluorescence pH titrations were conducted to probe for more direct evidence of Tyr-32/His interactions in α3Y-K29H and α3Y-K36H. Figure 6A shows the emission center of mass as a function of pH for Tyr-32 in α3Y. A small, 1.0 nm shift was observed between pH 5.1 and 9.1 and fitting the α3Y fluorescence titration plot to a single pK yields a value 8.1 ± 0.1. A significantly more substantial shift of 4.3 nm, and titrating with a pK\textsubscript{app} of 7.1 ± 0.1, was observed for α3Y-K36H. This value is exactly consistent with the 7.1 pK\textsubscript{app} of His-36 suggesting an interaction between Tyr-32 and His-36.
Finally, the fluorescence pH plot of $\alpha_3$Y-K29H titrates with a $pK_{app}$ of 7.4 ± 0.1. This suggests that another residue or residues in the addition to His-29 influence the fluorescence properties of Tyr-32 in $\alpha_3$Y-K29H.

To briefly mention the recently obtained electrochemical data, there is no major change in the potential of Tyr-32 in the three proteins but there are differences with respect to the protonic reactions coupled to the oxidation event (the change in $E_{peak}$ vs. pH is 47 mV/pH unit for $\alpha_3$Y & 53 mV/pH unit for $\alpha_3$Y-K36H). The $E_{peak}$ vs. pH plot of $\alpha_3$Y-K29H is not linear and most likely reflects redox-induced shifts in the $pK_A$ values of nearby residues. Interestingly, the pH dependence of the DPV peak height (reflecting the number of molecules oxidized) is significantly different in the three proteins.

In conclusion, the data obtained thus far show that $\alpha_3$Y-K29H and $\alpha_3$Y-K36H are stable and well-folded proteins. The protonation state of their single histidine residue influences the fluorescence of Tyr-32 suggesting an interaction between the His and Tyr side chains in these proteins. NMR NOE experiments in the aromatic region combined with protein modeling studies will be conducted to try to get a sense of the relative geometry of the two aromatic side chains in $\alpha_3$Y-K29H and $\alpha_3$Y-K36H. This information will be coupled to detailed analyses of the electrochemical data (see reference (150) for an example).

As a final remark, ideally we wish to make a stable Tyr/His complex that involves a direct phenolic oxygen – imidazole nitrogen hydrogen bond. The architecture of our three-helix bundle model protein might be too limited to be able to place the two residues in a geometry that allows a direct Tyr/His hydrogen bond to be formed. The four-helix bundle system described above, or potentially a natural protein such as e.g. ubiquitin, will be explored in future efforts along these lines. The practical lessons learned from this study will definitely help in the construction of such systems.
Redox characteristics of a *de novo* quinone protein (Paper V)

There are only a handful of voltammetry studies on quinone-containing proteins present in the literature, which induced us to try to extend the $\alpha_3C$ system and study quinone redox chemistry in addition to radical reactions. These efforts are described in Paper V and summarized below.

**Figure 14.** Chemical structures of (A) free 2,6-dimethyl-1,4-benzoquinone, and (B) 2,6-dimethyl-1,4-benzoquinone ligated to a Cys residue. $R'$ and $R''$ represents the peptide back-bone in the $\alpha_3C$ protein. For the solution reference compounds, $R'$ and $R''$ represents –COCH₃ and –OCH₃ groups, respectively.

With respect to sample preparation, the ligation chemistry used in this project was different relative to that used earlier. In Paper III, the mercapto-phenols were ligated to $\alpha_3C$ via a disulfide linkage (Figure 12) while in Paper V 2,6-dimethyl-1,4-benzoquinone (DBMQ) was bound to $\alpha_3C$ via a 1,4-Michael-type thiol addition reaction (Figure 14). The former reaction was very efficient and, when using a 1:5 protein/mercaptophenol ratio at alkaline pH, the binding reaction went to completion and all of the apo-$\alpha_3C$ was consumed. The latter reaction was less efficient and gave rise to side products. An additional redox dialysis step was introduced in the purification to maximize the amount of oxidized DMBQ-$\alpha_3C$ in the reaction mixture. This simplified the subsequent HPLC purification step significantly. The final yield of DMBQ-$\alpha_3C$ was about 25% relative to the starting amount of $\alpha_3C$.

The protein characterization of DMBQ-$\alpha_3C$ provided results similar to those obtained for the MP-$\alpha_3C$ proteins. It was shown that homogenous protein samples could be generated. UV/Vis spectra of the oxidized quinone and reduced hydroquinone form of DMBQ-$\alpha_3C$ were obtained and compared to the spectra representing aqueous DMBQ and DMBQ-bCys. As shown in Figure 1, Paper V, the absorption characteristics of DMBQ-$\alpha_3C$ are distinct relative to those of freely solvated DMBQ and DMBQ-bCys. Clearly, the protein-bound DMBQ experiences a significantly different milieu relative to the solvated species. A 273 nm extinction coefficient of 10600 ± 900 M⁻¹ cm⁻¹ was estimated for the oxidized form of DMBQ-$\alpha_3C$. As shown in Figure 1C, Paper V, DMBQ-$\alpha_3C$ displays a classic CD spectrum of a helical structure with double minima at 222 and 208 nm. The DMBQ-reacted $\alpha_3C$ protein was found to be 65% helical at pH 5.5 by comparing its 222 nm CD intensity with that of the structurally-characterized $\alpha_3W$ protein (143). This value is close to the 72% $\alpha$-helical content of $\alpha_3C$ and suggests that the binding of DMBQ does not significantly alter the $\alpha$-helical content of the $\alpha_3C$ scaffold. The global stability of $\alpha_3C$ and DMBQ-$\alpha_3C$ are -4.0 ± 0.1 and -2.7
± 0.1 kcal mol⁻¹, respectively. The decrease in stability upon DMBQ binding most likely reflects the presence of a quinone oxygen atom in the hydrophobic core of α₃C. The conclusion from the characterization of DMBQ-α₃C is that the protein remains folded and stable in aqueous solution and that no large-scale structural reorganizations occur upon DMBQ binding.

There are similarities as well as several distinct differences between the electrochemical characterization of DMBQ-α₃C relative to the MP-α₃C proteins: i) In both studies the properties of the protein-bound redox cofactor are compared to those of the corresponding aqueous species. The non-zwitterionic N-acetyl-L-cysteine methyl ester compound was used to mimic Cys-32 in the phenol and quinone aqueous systems. As noted earlier, we are interested in relative “protein vs. water” potentials and not absolute potentials. ii) The potential range probed in the mercaptophenol study was about 0.4 to 1.3 V vs. NHE. The higher number represents a practical upper limit for voltammetry studies on aqueous sample since the baseline arising from bulk water oxidation becomes increasing severe at higher potentials. The range used in the quinone study, 0.2 to about 0.9 V, was considerably more modest and solvent oxidation did not cause any significant problems. ii) The potential ranges that we had to use for the two different cofactors determined the type of working electrode. For the mercaptophenol study, we used a glassy carbon electrode (water oxidation starts at ~0.7 V) while a chemically modified gold electrode (water oxidation starts at ~0.5 V) was utilized for the quinone study. Naturally, a bare gold electrode cannot be used since the protein would be denatured on the electrode surface. All electrochemical measurements described in Paper V were preformed using a gold electrode modified with a self-assembled layer of 3-mercaptopropionic acid. iii) The radical cofactors studied in Paper III exhibit one-electron, irreversible electrochemistry. For this study, DPV was our main method of choice since this technique is more sensitive than cyclic voltammetry (CV) and, in addition, provides more distinct waveforms (i.e. easier to determine Eₚₑᵃᵏ values). In contrast, two-electron, quasi-reversible electrochemistry was observed for the DMBQ-based systems. For the quinone study we used CV since this method has the potential to provide more information of the Faradaic system studied (e.g. kinetic information, see below). iv) The small-molecule aqueous systems and the model radical proteins display diffusion control electrochemistry (Tommos, 1999; unpublished data). However, for DMBQ-α₃C we could obtain diffusion controlled as well as protein film data. The main results from the electrochemical characterization of DMBQ, DMBQ-bCys and DMBQ-α₃C are described below.

**DMBQ and DMBQ-bCys.** The CV characteristics of aqueous DMBQ and DMBQ-bCys were investigated as a function of pH. The obtained CV traces display broad and widely separated anodic and cathodic wave forms (Figure 2A, Paper V). The half-height width of the anodic and cathodic peaks, the difference between the anodic and cathodic peak potentials (ΔEₚₑᵃᵏ), and the ratio between the anodic and cathodic peak currents (iₚᵃ/iₚᶜ) were all pH-dependent and their values consistent with the typical quasi-reversible, n = 2 electrochemical behavior of solvated quinones (90).
Figure 15. Redox states of DMBQ-bCys (Q) in aqueous buffer. (A) At a solution pH below \( pK_{(1)} \) the potential has a pH dependence of 59 mV per pH unit for a two electron, two proton redox event at 25 °C and \( Q/QH_2 \) is the predominant redox couple. (B) Between \( pK_{(1)} \) and \( pK_{(2)} \) the redox reaction is pH dependent with 29 mV per pH unit slope for a two electron, one proton redox event at 25 °C. The major redox couple in this region is \( Q/QH^- \). (C) At a pH above the \( pK_{(2)} \) of the redox reaction becomes pH independent with the \( Q/Q^2- \) as the predominant redox couple.

DMBQ is a well-studied system, which allowed us to check the values we derived from our CV measurements against the literature. For DMBQ, a half-wave potential at pH 7.0, \( E_{1/2} \) (pH 7.0), of 179 mM was measured. This value is within the 160-180 mV range reported for aqueous DMBQ compounds. The 10.4 ± 0.3 (p\( K_{(1)} \)) and 12.6 ± 0.6 (p\( K_{(2)} \)) values derived from the DMBQ \( E_{1/2} \) vs. pH plot (Figure 2C, Paper V) are fully consistent with the p\( K_A \) values of 10.35 and 12.4 derived for DMBQ by optical methods (151). p\( K_{(1)} \) and p\( K_{(2)} \) were assigned to the hydroquinone \( QH^- + H^+ \rightleftharpoons QH_2 \) and \( Q^2- + H^+ \rightleftharpoons QH^- \) equilibria, respectively (see Figure 7 for quinone redox states).

Only a small change in \( E_{1/2} \) (pH 7.0) was observed for DMBQ-bCys relative to DMBQ (190 mV vs. 179 mV; Table 2 Paper V) but a substantial effect of the ring sulfur substitution was noted on the acidity of the quinone. Linear fits of the DMBQ-bCys \( E_{1/2} \) vs. pH data show a pH dependence of about 59 mV per pH unit below pH 7.0 and 29 mV per pH unit above pH 7.0. This indicates that the predominant redox couple is different in these two pH regions. A 59 mV per pH unit dependence at 25 °C is indicative of a two electron/two proton redox reaction (\( Q + 2e^- + 2H^+ \leftrightarrow QH_2 \)); reaction A in Figure 15) while a 29 mV per pH unit dependence is indicative of a two electron/one proton redox reaction (\( Q + 2e^- + H^+ \leftrightarrow QH^- \); reaction B in Figure 15). The p\( K_A \) for the \( QH^- + H^+ \rightleftharpoons QH_2 \) equilibrium of DMBQ-bCys was determined to 6.84 ± 0.09, which is significantly lower than the 10.4 ± 0.4 values derived for DMBQ. This shift may arise from the stronger electronegativity of sulfur relative to hydrogen, which predicts a stabilization of \( QH^- \) relative to \( QH_2 \).

**DMBQ-α3C.** The CV traces of DMBQ-α3C are significantly different relative to those of the two DMBQ solution systems (Figure 2B, Paper V). \( \Delta E_p \) is reduced by one order of magnitude in the diffusion controlled DMBQ-α3C data relative to the two DMBQ solution systems and is close to the expected value of 30 mV for a two-electron redox process. Non-unity
$i_{pa}/i_{pc}$ values and broadened peak widths were observed for DMBQ-$\alpha_3$C traces at all pH values measured. These quasi-reversible characteristics are common when other reactions, such as protonic reactions, are coupled to the electron transfer. It is also possible, and even likely, that protein movements are coupled to the quinone $n = 2$ redox change studied here.

The pH dependence of $E_{1/2}$ was investigated between pH 4.8 and 8.6 for DMBQ-$\alpha_3$C. In this case, a clear increase in $E_{1/2}$ (pH 7.0) was detected for DMBQ-$\alpha_3$C relative to DMBQ-bCys (297 mV vs. 190 mV; Table 2 Paper V) while there was no significant change in the $pK_A$ ($QH^- + H^+ \leftrightarrow QH_2$) value (6.84 ± 0.09 vs. 6.98 ± 0.08). The interaction between the surface of the electrode and DMBQ-$\alpha_3$C was found to be pH dependent and both diffusion-controlled and thin layer data could be obtained (for details, see page 3493, Paper V). The $E_{1/2}$ was investigated as a function of scan rate for the DMBQ-$\alpha_3$C monolayer at pH 7 and found to remain constant at 286 ± 4 mV in the range of scan rates analyzed. The measured $E_{1/2}$ value appears thus to be a true equilibrium value not governed by slow electron tunneling between the surface of the electrode and the protein bound quinone. We conclude that the protein matrix does influence the potential while the sulfur substitution of the ring is the main parameter controlling the acidity of the quinone.

To summarize, in Paper V we show that the $\alpha_3$C system could be extended also to study protein quinone redox chemistry. The cyclic voltammograms of DMBQ-$\alpha_3$C display overall characteristics similar to the few natural systems studied thus far. It appears likely that DMBQ-$\alpha_3$C will serve as a useful model protein and provide insights into the parameters that govern protein redox chemistry. This is particularly true once detailed structural characterization can be connected to the redox properties.
Summary and Future Perspectives

Although amino-acid radicals represent key components in processes ranging from energy transduction in plants to carcinogenesis in humans, virtually nothing is known about the thermodynamic properties of these species. The overall aim of this thesis was to develop a library of well-structured model proteins in which to study how specific protein features influence the electrochemical and thermodynamic properties of tyrosine and tryptophan radicals (Figure 16). The original members of this protein library are the $\alpha_3Y$ and $\alpha_3W$ de novo designed three-helix bundles (142). The primary sequences of $\alpha_3Y$ and $\alpha_3W$ are identical except for position 32, which is occupied by a tyrosine in $\alpha_3Y$ and by a tryptophan in $\alpha_3W$. Both $\alpha_3Y$ and $\alpha_3W$ are uniquely structured proteins (142) and a high-resolution NMR structural model has been obtained of the latter (143). These proteins are “redox inert” (i.e. solvent oxidation ultimately limits the electrochemical measurements and not spurious oxidation of the protein scaffold) and robust against changes in the solution pH (i.e. $\alpha_3W$ is $76 \pm 1\% \alpha$-helical between pH 4-10 (142, 143)). These are key features in order to probe the properties of the radicals generated in these systems in a meaningful way.

There are many published studies in which designed proteins are used to study redox reactions, but an overwhelming majority of the generated model proteins are not well structured. Detailed structural information can thus not be obtained and correlated with the redox characteristics observed for these systems. A major experimental effort of this thesis has been to strive for structural uniqueness so that both structural and electrochemical data can be obtained on the model radical systems. The $\alpha_4W$ project hopefully conveys this stance clearly. Here our aim was to make a single chain $\alpha$-helical protein about twice the size of the three-helix bundle system and containing an electrochemically active tryptophan. The initial protein design is described in the experimental part of Paper I and the refined design in Paper II. The data described in Paper II strongly suggest that the $\alpha_4W$ system can be used in the future to correlate radical characteristics with structural features, i.e., we demonstrate that electrochemical data can be obtained and that a structural NMR study is feasible. We anticipate that the larger scaffold of $\alpha_4W$ will broaden our possibilities to vary the properties of the radical site.

Paper III describes an extension of the three-helix bundle system with a third main variant denoted $\alpha_3C$. This protein was ligated with three different mercaptophenols (Figure 16) with the aim to study how the redox properties of a phenol are influenced by the solvent exposure of its hydroxyl group. We could show that the potential of the PhO'/PhOH redox pair is dominated by interactions to the phenolic OH group and that these interactions can elevate the protein potentials by 0.11-12 V relative to the phenol solution potentials. To investigate this further, magnetic resonance including both NMR and EPR techniques will be applied to structurally characterize the MP-$\alpha_3C$ proteins and their radical sites. To gain detailed structural information of the MP-$\alpha_3C$ proteins is likely to benefit a related project. In Paper V, Cys-32 is ligated to a quinone and a detailed cyclic voltammetry study was conducted on the resulting DMBQ-$\alpha_3C$ protein. This started as a bit of the side project
but the considering the well-behaved and interesting electrochemical properties of DMBQ-α₃C, this system is likely to be expanded in the future.

Perhaps the overall most challenging project in this thesis is to make histidine variants of α₃Y in which the imidazole side chain is involve in protonic reactions induced by the oxidation of Tyr-32 (Paper IV). As described earlier, spectroscopic and structural screening narrowed initially nine α₃Y/His variants to two that displayed properties suggestion a connection between the aromatic side chains of Tyr-32 and the introduced histidine. Electrochemical data have been obtained on these two proteins and α₃Y although more analysis is required in order to draw firm conclusions from these data sets. To make a Tyr/His pair, and most ideally a directly hydrogen-bonded Tyr/His pair, represents an important goal and further attempts using either the α₄W system or a naturally occurring well-characterized protein will be made along these lines.

Figure 16. Model proteins made to study various aspects of tryptophan and tyrosine radicals and quinone redox chemistry.
6. Acknowledgements

Jag vill tacka


Astrid Gräslund och Stefan Nordlund som får institutionen att fungera på ett föredömligt sätt.

Bruce W. Berry – for being a great colleague and friend. Room A350a has been very empty this last year. I’ll see you at your dissertation!

Sam Hay – for fruitful collaborations and for being so patient when teaching me the mysteries of electrochemistry.

Margaret Elvekrog – thanks for all good times in and out of the lab and for being such a great hostess in New York. I’ll be back to take advantage of your hospitality again soon.


Josh Wand and all the people at the Wand lab, UPenn.

Eeva-Liisa Karjalainen – tack för att du är en fantastisk vän som alltid ställer upp när jag behöver dig.


Sofia Unnerståle – för våra diskussioner om livets väsentligheter.

Ulrika Flock, Kristina Faxén och Lina Salomonsson – tack för en rolig vecka (med mycket shoppande) i Kalifornien.

Haidi och Torbjörn Astlind – som får allt att fungera på biofysik.

Alla vänner utanför jobbet! Ni är fantastiska.

Och så förstås tack till min familj,

Kenneth och Mikael – ni är de bästa bröder en (något bortskämd) lillasyster kan tänka sig! Vad gjorde jag utan er?

Mamma och Pappa – tack för allt stöd under alla år. Jag älskar er.

Slutligen tack till alla på biofysik som har gjort avdelningen till en trevlig plats att jobba på.
7. References


