Solid Binding Peptides: Applications in Affinity Protein Purification and Bioprotonic Devices

Jessica Soto-Rodriguez

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

François Baneyx, Chair

Cole DeForest

Guozhong Cao

Program Authorized to Offer Degree:

Department of Chemical Engineering

University of Washington

Abstract

Solid Binding Peptides: Applications in Affinity Protein Purification and Bioprotonic Devices

Jessica Soto-Rodriguez

Chair of the Supervisory Committee:

François Baneyx

Department of Chemical Engineering

Solid binding peptides (SBPs) are short sequences of amino acids that can be fused to proteins to endow them with the ability to bind to inorganic or synthetic interfaces without affecting structure or function. Here, we describe the use and optimization of SBP usage in two important technological applications: the affinity purification of proteins, and the construction of a new class of bionanoelectronic devices. First, we use Car9, a silica binding peptide whose interaction with silica can be disrupted with lysine, to optimize the conditions for the purification of Car9-tagged proteins using underivatize silica. We demonstrate that capture of Car9 fusion proteins is enhanced on small particle and large pore size silica gels, that addition of 0.3% Tween 20 prevents non-specific protein binding and that efficient elution is achieved under alkaline conditions. Using the optimized conditions, we developed a small-scale purification kit that allows the inexpensive recovery of milligram-quantities of Car9-tagged proteins with purities higher than 90%. Then, we extended the functionality of SBPs to the integration of rhodopsins to electronic devices for the light activated conversion of protonic currents into electronic currents.

To this end we fused Pd4, a Pd-binding peptide, to the N-terminus of the green light absorbing *H.turkmenica* deltarhodopsin (HtdR) to bring the protein and protons transported by it to a closer proximity to the surface of Pd/PdHx-based devices. We demonstrate that addition of the Pd4 tag leads to a nearly one order of magnitude increase in electronic signal when compared to untagged protein. To enhance their functionality and develop multi-wavelength devices, we optimized the production of a blue absorbing proteorhodopsin (BPR) by exploring signal sequence requirements and export pathway. We show that BPR does not rely on the Sec pathway for inner membrane integration and that its signal sequence is not necessary to obtain a functional product. We also identified a mutation in the signal peptidase I recognition site that prevents cleavage of the signal sequence resulting in a product that exhibits all functional attributes of the wild type protein but produced at higher levels. After optimizing the production of BPR, we constructed a Pd binding BPR to fabricate HtdR- and BPR-based devices with light absorption maxima - and consequently photocurrent maxima - separated by 37 nm. These devices exhibit wavelength-dependent photocurrent production, opening up the possibility to develop a new type of biological camera. Together, this work demonstrates that SBPs are powerful tools for the development of cost effective purification schemes and more efficient devices.

Table of Contents

Table of Figures

Acknowledgements

I would like to acknowledge all the people that have supported me during my years as a graduate student. First, I want to start with my family: my mom Nilda, who always encouraged me to be independent; my brother Audeliz, that has been my biggest inspiration; my husband José Luis that has always motivated me to accomplish my goals even when it meant living in different cities; my sister-in-law Eunice who has been one of my best friends and my niece Mylana Victoria who was born while I was writing my dissertation and brought me joy and inspiration to keep writing.

I am also grateful for the support that I received in the Chemical Engineering Department. I want to thank my advisor François Baneyx for his guidance in my research project, for giving me the opportunity to become an independent researcher and for encouraging me to give back to the community by outreach and recruiting events. I also want to thank the members of my committee: Cole DeForest, Guozhong Cao and Daniel Ratner for their invaluable advice for the completion of my dissertation. Additionally, I want to thank Marco Rolandi at UCSC for allowing me to collaborate in his lab, as well as his former postdoc Zahra Hemmatian for her contributions to my projects. Also, a special thank to Lilo Pozzo for her support and advice for my research project and future career options. Finally to my co-workers: Brittney, Alex, Yundi, Lena, Amy, Matt, Brian, Brandon, James, Sonja, Wayne, Wenlan, Rosie, Kannan, Tong, Yatong and Weibin for your help in the lab and friendship during the past years. There are so many other people who have contributed to my success that I just do not have the space to thank you all individually, so this is my collective "gracias" to everyone I have had the pleasure of working with.

Chapter 1 Introduction

1.1 Solid binding peptides (SBP)

Solid binding peptides (SBPs) are short sequences of amino acids that bind to specific inorganic materials [1-6] with equilibrium dissociation constants that often falls in the nanomolar to sub- micromolar range. SBPs can be isolated by combinatorial techniques such as phage or cell surface display [1,7] for an ability to bind to palladium [8], titana [9], gold [10], silver [11], calcium phosphate [12] and many other inorganic and synthetic phases. All of these materials, especially at the nanoscale, hold promise for a broad range of applications including catalysis, electronics and therapeutics. To increase biocompatibility and usefulness of these materials, SBPs can be fused to proteins and enzymes to yield molecules capable of binding to a material while retaining their activity [13]. This makes SBPs powerful molecular tools for the immobilization of biomolecules to inorganic surfaces.

In addition to their ability to bind to inorganic surfaces, SBPs have been used as synthesizers for the mineralization of inorganic materials. Living systems have evolved to mineralize a variety of inorganic materials including silica [14,15], calcium carbonate [16] and magnetite [17] using proteins, peptides, and other biomolecules in aqueous and mild conditions. Finding inspiration in nature, SBPs have been isolated to bind and mineralize materials under ambient temperatures and neutral pH [8,9,11,18,19]. For example a silver binding peptide isolated by phage display and called Ag4 was found to precipitate silver nanoparticles of different morphologies when added to a silver nitrate solution [11]. Platinum nanocubes and nanotetrahedrons have been synthesized by selecting facet-specific platinum binding peptides by phage display [19]. Palladium [8], silver [11], titania [9,18] and other nanoparticles have also been synthesized by peptide-aided mineralization. Moreover, when fused to proteins, SBPs remain

capable of mineralizing inorganic materials, adding functionality to inorganic-protein nanostructure [20-22].

1.2 Silica binding peptide for affinity purification

1.2.1 Protein Purification

Proteins play crucial roles in living systems and are responsible for the proper functions of cells, tissues and organs. Advances in recombinant technologies have facilitated the study of proteins and made significant impact in the fields of biotechnology and therapeutics. Recombinant proteins have been successfully produced in a wide variety of hosts[23] such as bacteria, yeast, insects and mammalian cells. After expression in the selected host, the target protein must be separated from the rest of cellular proteins, in a process referred to as protein purification[24,25]. This is a crucial step for the production and characterization of recombinant proteins. There are various methods available for the purification of proteins based on electrostatic interaction with a charged resin (ion exchange chromatography)[26], hydrophobic-hydrophobic interactions (hydrophobic interaction chromatography)[27], size and shape (gel filtration or size exclusion chromatography)[28], and affinity for a ligand immobilized on a solid matrix (affinity chromatography)[29].

In ion exchange chromatography[26] the surface charge of a target protein in a buffer at defined pH mediates binding to a matrix with opposite charge. If the buffer pH is higher than the isoelectric point (pI) of the protein, it will carry a net negative charge and thus bind to a positively charged anion exchanger resin. When the pH is lower than the pI, the protein is positively charged and binds to a negatively charged cation exchangers. This reversible electrostatic interaction can be disrupted by changing the solution ionic strength with a salt or by using a pH gradient to recover the purified protein. In hydrophobic interaction chromatography[27],

proteins are separated based on their interaction with a hydrophobic matrix. A high salt concentration is used to facilitate binding to the matrix and the protein is recovered by lowering the salt concentration. In gel filtration or size exclusion chromatography[28], proteins are separated according to intrinsic differences in molecular mass and shape. The matrix or gel used for this type of separation consists of spherical beads with pores of defined size distribution. Proteins smaller than the size of the pores diffuse through them and require longer times (e.g., higher volumes of mobile phase) to elute. Larger proteins elute faster with the largest ones emerging in the void volume of the column. Finally, in affinity chromatography, a target protein is purified on the basis of its high affinity for an immobilized ligand and elution is preformed via addition of free ligand or a change in environmental conditions that weakens the interaction (e.g., pH). Due to the widespread use of the technology and relevance to this work, this method is further discussed in the next section.

1.2.2 Affinity chromatography

Affinity chromatography allows for the purification of a target protein based on its selective binding for a ligand immobilized on a matrix. Binding is reversible and allows for protein concentration on the matrix. After binding, washing and elution steps allow for the recovery of a highly concentrated and purified protein, often in a single step[30,31]. There are many available affinity tags that can be fused to a target protein to provide affinity to a matrix. Some examples are: 1) the maltose binding protein (MBP)[32] which has high affinity to amylose and is often used as a fusion partner to increase protein solubility[33] (however, the size of MBP is ~ 45kDa which can be undesirable for further applications; 2) gluthathione S-transferase (GST)[34] which can be purified using a glutathione resin but generally requires an additional treatment of the cell lysate to eliminate GroEL, a 70-kDa protein that co-purifies and elutes with GST fusion proteins[35]; 3)

the strep tag,[36] a 9 amino acid tag with affinity for streptavidin which allows elution of fusion proteins by competition with biotin; and 4) the His tag [37] composed of 6-10 histidine residues which confers affinity for transition metal ions such as Ni^{2+} or Co^{2+} immobilized on a solid support through a coordinated ligand such as nitrilotriacetate.

The His tag is one of the most commonly used affinity tags and the technology is generally referred to as immobilized metal affinity chromatography (IMAC)[38]. Proteins fused to a His tag are captured by the resin while untagged proteins flow through. The target protein is typically recovered by competitive elution with imidazole. One of the main advantages of the His tag is its small size which has little deleterious impact the target protein. The tag can also be easily removed with an endoprotease if an appropriate cleavage site, such as TEV recognition site, has been engineered between His tag and target protein. On the negative side, His-tagging has been reported to cause protein insolubility [39-41], alter structure and function [42], and lead to product inhomogeneity [43]. In addition, the stripping of protein-coordinated metal ions by imidazole can induce protein misfolding [44], and His-tagging remains expensive at the laboratory scale due to the cost of the resin [45]. Additionally, most manufactures recommend regeneration of the matrix after 5 rounds of purification, which adds additional labor time and costs. The costs associated with disposable Ni-NTA based columns for small scale and quick purification schemes are even higher (~ $10 per run). Thus, there is a need to find a more economical resin option for affinity purification.

1.2.3 Silica as a protein purification matrix

Silica (SiO_2) is one of the most abundant materials on earth making it attractive as a matrix for protein purification with the potential of decreasing purification costs. Due to the low cost of producing silica particles of controlled size and porosity and the high mechanical strength of the

material, manufacturers have widely use silica as a solid phase for chromatographic matrices. Silica[46] is composed of silicon atoms that are tetrahedrally coordinated to oxygen atoms through siloxane (Si-O-Si) bonds. Silica surfaces contain both hydrophobic siloxane sites and hydrophilic silanols (Si-OH). Silanols can be present in isolated (single), geminal or vicinal (H-bonded) forms. Ionization of Si-OH to SiO^- occurs at pH higher than 3, which results in an increasingly negatively charged surface [47-50]. Therefore, non-specific binding of proteins to silica is generally dominated by electrostatic and hydrophobic interactions, although H-bonding and van der Waals interaction may also play a role[51-53].

There has been growing interest in using underivatized silica as a stationary phase for preparative scale purification of proteins whose interactions with the matrix is well understood,[54] using the isoelectric point and surface hydrophobicity of the target protein to determine favorable elution conditions. However, approaches relying on the specific interaction of silica binding domains with the matrix can be more versatile and provide better results. For instance, Ikeda and coworkers have shown that fusion of the 273-residues long ribosomal protein L2 (*a.k.a.* Si-tag) to the C-terminus of *Staphylococcus aureus* protein A, allows for recovery of the chimera at better than 85% purity when it is released from matrix with 2M $MgCl_2$ [55]. Purification of Si-tag fusion proteins was also achieved under denaturing conditions[56]. Similarly, fusion of the 171 amino acids-long *Bacillus cereus* CotB protein to mCherry with an intervening SUMO recognition sequence supports batch-mode binding of the fusion protein on silica beads, and intact mCherry can be released with high purity after 3h incubation with SUMO protease [57]. The large size of L2 and CotB is a disadvantage of both approaches, as is the elution strategy (high concentrations of $MgCl_2$ for L2 and lengthy and costly SUMO protease processing for CotB). And while Abdelhamid and coworkers have found that a 14 residues-long peptide derived from

the C-terminus of CotB (CotBp, SGRARAQRQSSRGR) supports the binding of mCherry to silica, SUMO protease treatment remains required for target protein release [57].

1.2.4 Car9 – a silica binding peptide

Car9 (DSARGFKKPGKR) was initially identified by cell surface display for its ability to bind to carbonaceous substrates,[58] but it was later found to have high affinity for silica [59]. Binding of Car9 to silica is not yet fully understood but clues are available based on its composition and experimental observations. Basic residues (three Lys and two Arg) in the peptide sequence are likely to bind to silanol groups in silica via electrostatic interactions. However, the binding mechanism is not limited to electrostatics, since Car9 retains high affinity for silica under high salt concentrations [59]. Hydrophobic interactions and hydrogen bonding may be also involved in Car9 binding to silica as it has been shown for other proteins and peptides that bind to silica. Interestingly, the interaction of Car9 with silica can be disrupted with lysine or arginine[59]. Based on this observation, Car9 fusion proteins have successfully been entrapped in silica sol-gels with controlled protein release with arginine [60] and can be reversible immobilized and microcontact printed on glass surfaces [61]. Chapter 2 details a technology for the purification of Car9 fusion proteins on silica gel and its optimization for various proteins of interest.

1.3 Integrating rhodopsins in bioprotonic devices

1.3.1 Secretion and membrane protein insertion in *E. coli*

In order to perform their intended functions, about 30% of proteins synthesized by *E. coli* have to be exported to the inner membrane, the periplasm, the outer membrane, or even translocated to the extracellular space[62,63]. Proteins destined for export to the periplasm and outer membrane typically contain an N-terminal cleavable signal sequence (see **Figure 1.1 A**)

consisting of a net positively charged region, a hydrophobic core, and a cleavage site recognized by signal peptidase I (SPase I).[64] These proteins are usually exported by the Sec-dependent pathway **(Figure 1.1 B)**, in which signal sequence-containing precursor proteins, or pre-proteins, are captured post-translationally by secretory (e.g., SecB) or generic molecular chaperones, which maintain them in an unfolded conformation for transfer to SecA, a molecular motor that uses cycles of ATP binding and hydrolysis to thread pre-proteins across the membrane through the SecYEG pore. Once translocation is complete, signal sequences are cleaved by SPase I to yield a mature product that is released in the periplasm. [62,65].

By contrast, typical *E. coli* inner membrane proteins do not contain a signal sequence and most traffic to the membrane through the cotranslational signal recognition particle (SRP) pathway **(Figure 1.2)**. Here, the specialized chaperone SRP engages nascent proteins displaying highly hydrophobic α-helical TMS as soon as they emerge from the ribosomal tunnel. The resulting ribosome-nascent protein complexes (RNC) are delivered to the FtsY receptor at the membrane, and the cargo is transferred co-translationally to the SecYEG/YidC machinery. The opening of a lateral gate in the SecYEG channel allows for integration of membrane proteins within the lipid bilayer [66-68].

1.3.2 Prokaryotic Rhodopsin

Prokaryotic rhodopsins include proton pumps [69-72], chloride pumps [73] and sensory rhodopsins [74]. Although there is significant variation in the amino acid sequence of rhodopsins, they share a common structure consisting of 7-transmembrane segments (TMS). One of the most studied prokaryotic rhodopsin is bacteriorhodopsin (bR) [71] which support the light-activated, outwards transport of protons across membranes to power archaeal ATP synthases.[75] The prototypical *H. salinarum* bR is a homotrimer that packs into crystalline arrays to form the

purple membrane patches in the archaeon. Each bR subunit is organized as a seven helix bundle and contains a molecule of retinal bound to a lysine residue in the C-terminal helix G via a protonated Schiff base.[71] When retinal absorbs a photon of light, it isomerizes from an all-*trans* to a 13-*cis* conformation, which changes its molecular shape and induces a conformational change in the protein backbone. The excited chromophore also becomes positioned to transfer a proton to a neighboring aspartate residue (Asp-85) that itself shuttles the proton to the outside of the cell through a release group. The negatively charged retinal next takes up a proton from Asp-96, which is oriented towards the cytoplasm (see **Figure 1.3**). In doing so, the retinal returns to the ground state and Asp-96 replenishes its lost proton from the cytoplasm, completing a photocycle.[71]

1.3.3 Deltarhodopsin

Bacterorhodopsin can be obtained from the purple membrane of Halobacteria but its expression in common hosts such as *E. coli* has proven challenging. However, other archaeal rhodopsins such as *Haloterrigena turkmenica* deltarhodopsin (HtdR) have been expressed and purified from the *E coli* membrane [70]. Like bR, HtdR acts as a light activated outward proton pump as demonstrated by the light activated extrusion of ethidium bromide from *E. coli* cells co-expressing HtdR and Emr, a transporter that uses proton gradients to power toxicant extrusion across the bacterial inner membrane. Comparisons with the crystal structure of *Haloterrigena thermotolerans* deltarhodopsin (dR3)[72], indicates that the HtdR residues responsible for proton pumping are conserved with Asp-84 substituting for Asp-85) as proton a donor, and Asp-95 replacing Asp-96) as a proton acceptor.

The yields of HtdR can be further improved in *E. coli* by placing the *htdR* gene under transcriptional control of a mutant pBAD promoter that helps match HtdR production rates with

the ability of the cell to insert proteins within the inner membrane,[76] and by utilizing trigger factor deficient strains that are beneficial for the production of membrane proteins by reducing competition for SRP binding[77]. The maximum absorbance of HtdR is in the green region of the visible spectrum (~545 nm), which is useful for the fabrication of bioprotonic devices manipulated with green light.

1.3.4 Proteorhodopsin

A distinct group of light-activated proton pumps are bacterial proteorhodopsins (PR) [69,78]. Since the discovery of PR in uncultured samples from the Monterey Bay in California, this group of proteins has attracted considerable interest because they can be expressed at much higher yields in *E. coli* than typical membrane proteins. PRs are divided in two groups distributed at different depths in the ocean and absorbing light at different wavelengths. Green-absorbing proteorhodopsin (GPR) has an absorption maximum at 525 nm and is typically found in microorganisms dwelling at the surface. Blue-absorbing proteorhodopsin (BPR) absorb maximally at 490 nm and are found in deeper zones [79]. In addition, their spectral characteristics are pH dependent and spectral tuning from blue to green can be achieved by introducing a Gln to Leu mutation at position 105 in BPR or from green to blue by introducing a Leu to Gln mutation at position 105 in GPR[79,80]. Although PRs have less than 30% homology with bR, they share a common 7-TMS architecture, and the residues critical for proton-pumping are conserved with Asp-97 (equivalent to Asp-85 in bR) acting as proton acceptor to the retinal Schiff base and Glu-108 serving as proton donor (in place of Asp-96 in bR) [81,82].

The blue-absorbing HOT75m1 proteorhodopsin has been extensively studied and its structure[83] and function[84] are well understood. Although HOT75m1 can be expressed in *E. coli* at

fairly high yields, the role of its cleavable N-terminal signal sequence and how this extension impacts protein expression and membrane insertion has so far not been studied.

Signal sequences are typically present in secretory and outer membrane proteins that are exported across the inner membrane (see section 1.3.1) but not in inner membrane proteins. Chapter 4 investigates the role of the signal sequence, showing that it should be thought as a beneficial pro sequence rather than as secretory peptide, and demonstrating that a mutation preventing signal sequence processing allows for the accumulation of a fully functional protein at 50% higher yields.

1.3.5 Bioprotonic devices

The field of bioelectronics [85] focuses on the production of devices that mimics the function of biological components, and on the incorporation of biomolecules in these devices to transduce biological signals into electronic currents. Although there has been successes in the development of bioelectronics devices, and although some have seen important applications in the biomedical field (e.g., for blood glucose sensing[86]), challenges related to the compatibility of the inorganic component with biological systems have prevented further applications. In parallel, rapid advances in nanotechnology have provided new means for interfacing materials with appropriate size (e.g., nanotubes[87,88] and nanowires[89,90]) with biological component.

An additional challenge related to the biotic/abiotic interface is the difference in communications language. While most biological system use ions and other small molecules for function, devices mainly rely in electrons. Thus, efficient integration of biological components and electronics requires controlled conversion of biochemical ionic gradients into electrical signals that can be read or processed by conventional electronics. To this end, ionic[91] and mixed

conductivity in biological[92] and organic polymers[93] have been used to record and stimulate physiological functions and even assemble logic circuits,[94] most recently inside plants.[95]

In addition to ion and small molecules, protons (H^+) play crucial roles in biological energy generation and transduction processes[96] including ATP synthesis,[97] bacterial propulsion[98] and bioluminescence.[99] Pd/PdHx bioprotonic devices[100-107] that conduct H^+ rather than electrons have provided a new way to control and integrate these proton gradients. Palladium is a desirable material for contacts due to its ability to form palladium hydride (PdHx) when exposed to hydrogen[106]. H^+ adsorb to the surface of Pd and can be reduced to H when a negative voltage is applied to the surface, resulting in the formation of PdHx according to the following reaction:[108]

$$Pd + H^+ + e^- \rightleftharpoons PdH_x$$

PdHx is both proton conducting, which allows for direct control of protonic gradients, and electron conducting, which allows for the construction of interfaces with external electronics. One such example is a protonic field effect transistor (H^+-FET) [106], composed of a proton-conducting maleic-chitosan channel that bridges palladium hydride (PdHx) source and drain contacts on top of SiO_2 gate dielectric (**Figure 1.4**). The application of an electrostatic potential to the gate electrode allows control of the protonic current through the channel. More importantly, the functionality of these devices has been extended to aqueous solutions, the appropriate environment for biological systems.

An enzyme logic bioprotonic transducer[103] was used to demonstrate the functionality of these devices in aqueous solutions. In this device, PdHx formation (and the electronic current associated with interfacial proton reduction) is controlled by the decrease in pH caused by the glucose dehydrogenase catalyzed production of gluconic acid from glucose and nicotine adedine

dinucleotide (NAD^{+}). The result is the transduction of a biological signal into a measurable electronic current. More complex devices incorporating ion channels in Pd/PdHx devices have also been described[107]. In these systems Gramicidin A (gA) and alamethicin (ALM) embedded in supported lipid bilayers (SLB) allowed the regulation of proton flow across the SLB, with gA used to linearly control proton currents as a function of voltage and ALM acting as a voltage-gated channel equivalent to an ON-OFF switch.

The selectivity and binding activity of SBPs should prove useful for the integration of biomolecules in such devices. Indeed, Chapter 3 describes the integration of light activated proton pumping rhodopsins in Pd/PdHx devices using a Pd binding dodecapeptide (Pd4) and the efficient conversion of light generated proton currents into measurable electronic currents. Chapter 5 further explores the possibility of tuning device responses by integrating rhodopsins from archaeal and bacterial origin and selecting appropriate wavelengths for illumination.

1.3.6 Pd4 – a palladium binding-peptide

Pd4 (TSNAVHPTLRHL) is a palladium binding peptide that was isolated via phage display [8]. It has been shown to mineralize Pd nanoparticles ~ 2nm in size by acting as a capping agent during the reduction of a Pd precursor by sodium borohydryde [8]. Moreover, the particles exhibit catalytic activity for the formation of C-C bonds of the Stille coupling reaction[8]. Coarse grain Monte Carlo simulations has shown that Pd4 binds with strong affinity to palladium surfaces by anchoring its His residues (positions 6 and 11) to the surface [109].

Taking advantage of SBP versatility, this dissertation tackles: (i) the optimization of a silica based protein purification technique in which the silica-binding peptide Car9 (described in section 1.2.4), is fused to target proteins and elution is achieved by disrupting the Car9-silica interaction with the amino acid lysine (Chapter 2); (ii) the development of a novel strategy for

the integration of proteins in Pd/PdHx bioprotronic devices that relies on the fusion of the palladium binding-peptide Pd4 (described below in section 1.3.6) to deltarhodopsin to achieve green light-controlled conversion of protonic to electronic currents (Chapter 3); iii) the investigation of the role of the signal sequence of a blue absorbing proteorhodopsin with the goal of improving protein yields for incorporation in electronic devices (Chapter 4); and the development of a two channel Pd/PdHx photodetector that integrate palladium binding deltarhodopsin and proteorhodopsin that exhibit wavelength-dependent photocurrent production.

1.5 Figures

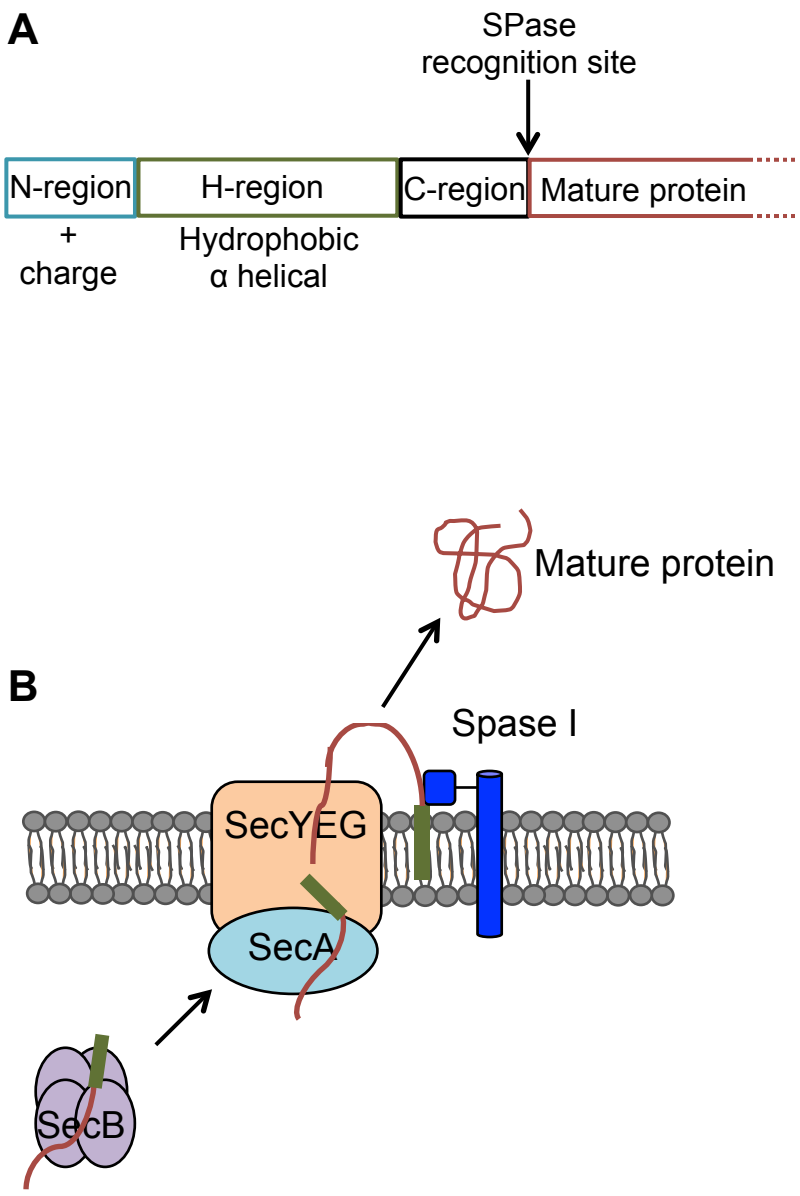

Figure 1.1 (**A**) Proteins destined for the Sec-dependent mechanism contain an N terminal signal sequence consisting of a positively charged N-region, an hydrophobic core with α-helical conformation and a cleavable region recognized by SPase (**B**) Molecular chaperones such as SecB maintain the pre-protein in unfolded conformation. The pre-protein is transferred to SecA, which drives the pre-protein through the SecYEG pore. SPase then removes the signal sequence (represented in green).

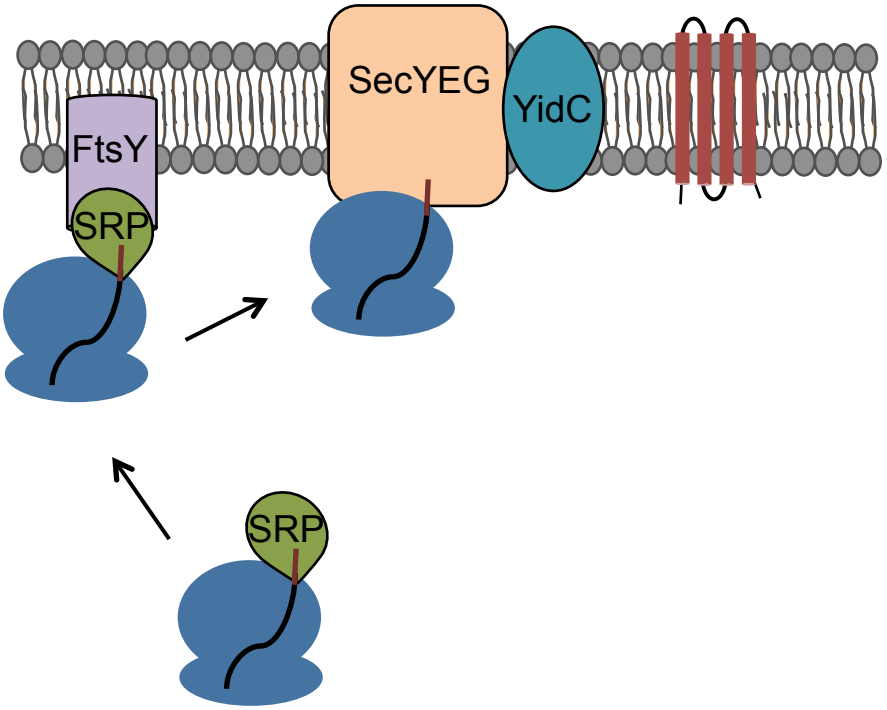

Figure 1.2 As proteins emerge from the ribosome tunnel, SRP bind to hydrophobic regions of transmembrane domains. The ribosome nascent chain complex (RNC) is delivered to FtsY that transfers it to the SecYEG for integration in the membrane.

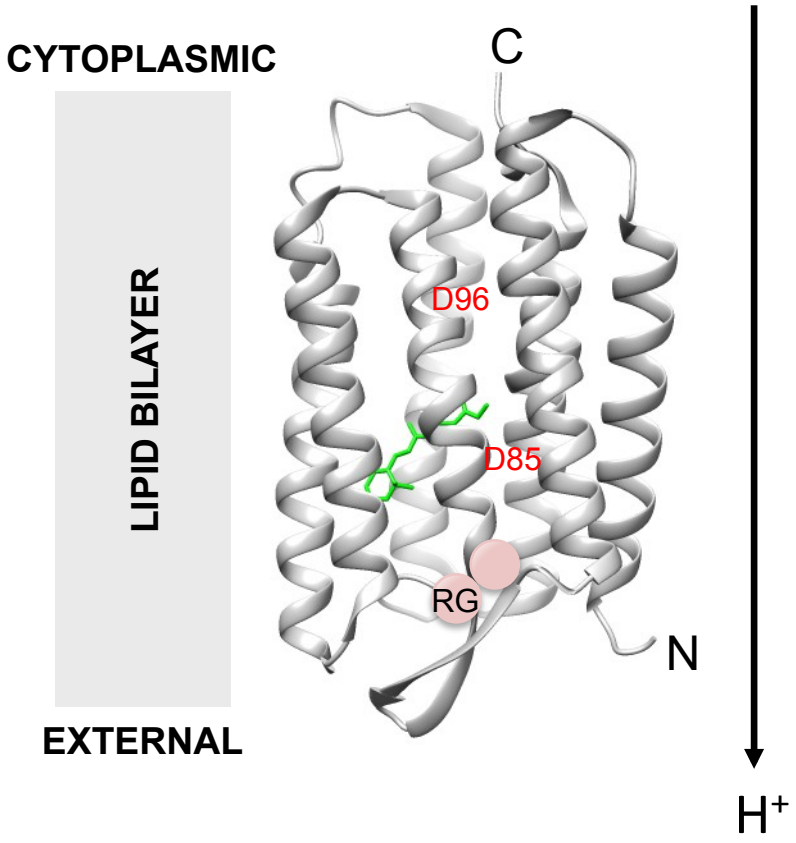

Figure 1.3 Bacteriorhodopsin structure. When retinal (green) absorbs light it changes conformation from all-trans to 13 cis. The excited chromophore can transfer a proton to Asp-85 that shuttles it to the outside through a release group. Retinal replenishes the proton from Asp-96.

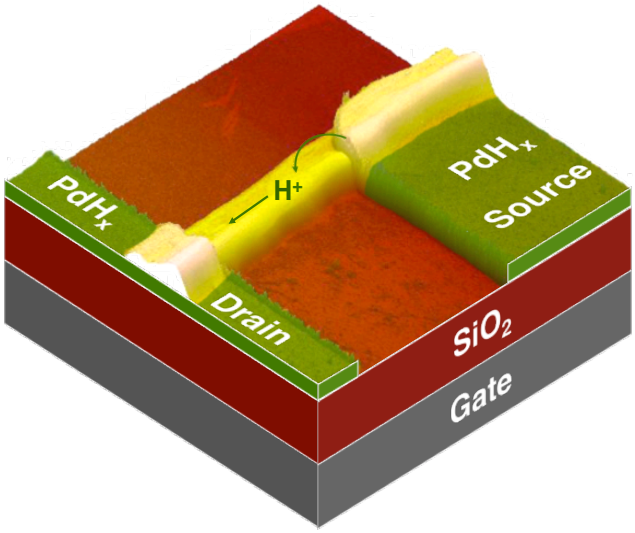

Figure 1.4 Bioprotonic field effect transistor composed of a maleic chitosan channel bridged between palladium hydride (PdHx) source and drain contacts on top of a SiO2 gate dielectric. Adapted from Ref. 106.

Chapter 2 Affinity Purification of Car9-Tagged Proteins on Silica Matrices: Optimization of a Rapid and Inexpensive Protein Purification Technology

2.1 Introduction

Affinity tags have transformed protein purification by converting time-intensive, multistep chromatography processes into single-step schemes that yield fairly pure polypeptides [30,31,45,110]. The primary function of a genetically (or chemically) appended affinity tag is to endow a protein of interest with the ability to reversibly bind to a ligand coupled to a stationary phase, or to directly interact with the stationary phase itself. Tags may also function as solubility enhancers, help facilitate the detection of proteins to which they are fused, and are often excisable if the fusion joint is engineered to contain elements supporting endoproteolytic or self-cleavage events [111].

Previously, we reported that Car9 (DSARGFKKPGKR), a dodecapeptide originally isolated for its ability to bind to the edge of carbonaceous materials [58] exhibits high affinity for silica [112]. Capitalizing on this finding and on the fact that Car9 can be competitively released from silica by addition of 1M lysine or arginine, we further demonstrated that C-terminal Car9 extensions support the microcontact printing of proteins on glass substrates [61], their self-immobilization and dynamic release from silica sol-gels [60], the assembly of functional multi-material architectures [113], and affinity protein purification on unmodified silica gels [112].

In this study, we show that Car9 is functional and tobacco etchv (TEV) protease-excisable when fused to the N-terminal of target proteins and that it remains capable of binding to silica under denaturing conditions, although it does so with lower affinity. Additionally, we harness the current understanding of how proteins and small molecules interact with SiO_2 to optimize a

simple kit that enables rapid and inexpensive affinity purification of Car9-tagged proteins at the laboratory scale on a disposable silica phase.

2.2 Materials and Methods

2.2.1 Plasmids and DNA manipulations

Plasmids pET24a(+)-sfGFP-Car9 and pET24a(+)-mCherry-Car9 which encode sfGFP-Car9 and mCherry-Car9 under transcriptional control of the IPTG-inducible T7 promoter, and plasmid pBLN200-MBP-Car9, which encodes a MBP-Car9 fusion protein under transcriptional control of the arabinose-inducible P_{BAD} promoter have been described [112,114]. Plasmid pNC9, a vector suitable for inserting gene products as *Nco*I-*Xho*I or *Nco*I-*Hind*III fragments downstream of a N-terminal Car9 tag followed by a flexible GGGS linker and TEV protease cleavage site was constructed by inserting an *Nde*I-*Nco*I cassette encoding tag, linker and TEV site in the same sites of pET-22b(+). An *Nco*I-*Xho*I cassette encoding TEM1-β-lactamase was amplified by PCR using primer pair 5'-CCGTAGCCATGGATCACCCAGAAACGCTGGTG-3' and 5'-GACCGGCTCGAG TTACCAATGCTTAATCAGTGAGGC-3' and plasmid pET-22b(+) as a template. An *Nco*I-*Hind*III cassette encoding sfGFP was amplified using primer pair 5'-GGGCCATGGGGCGTAAAGGCGAAGAGCTGTTC-3' and 5'- GCCAAGCTTTTAT TTGTACAGTTCATCCATACC-3' and plasmid pET24a(+)-sfGFP-Car9 [112] as a template. Both genes were inserted into the same sites of pNC9 to produce plasmids pET-22b(+)-Car9-β-lactamase and pET-22b(+)-Car9-sfGFP. The Car9-b-lactamase gene was transferred to pET-24a(+) by inserting the small *Nde*I-*Xho*I fragment of pET-22b(+)-Car9-β-lactamase into the same sites of that vector.

2.2.2 Optimizing elution pH and silica matrix

Seed cultures (25 mL) of BL21(DE3) cells (Novagen) harboring plasmid pET24a(+)-sfGFP-Car9 were used to inoculate 500 ml of LB medium supplemented with 50 µg/mL of kanamycin. Cells were grown to mid exponential phase (A_{600} ~ 0.5) at 37^{o}C and protein synthesis was induced by addition of 1 mM IPTG. After 3h of cultivation at 37^{o}C, cells were harvested by centrifugation at 7000*g* for 10 min, resuspended in 35 mL of 20 mM Tris-HCl, pH 7.5, 2 mM EDTA, and disrupted by six round of sonication for 3 min at 30% duty cycle using a Branson sonifier. Insoluble material was removed by centrifugation at 10,000*g* for 10 min. To test the effect of pH, clarified lysate (1.5 mL) were contacted with 1.5 mL of silica slurry (Sigma Merck Grade 7734, 6 nm pore size, 60-200 µm particle size) for 5 min on a rotary mixer operated at room temperature. The slurry was loaded on a 0.5 cm ID glass column (BioRad), and 10 mL of 20 mM Tris-HCl, pH 7.5 was added to the top of the bed. A 5 mL syringe was connected to the outlet of the column and the same volume of fluid was collected by aspiration to remove non-specifically bound proteins. Bound sfGFP was eluted by aspiration of successive 0.5 ml fractions using 20 mM Tris-HCl, pH 6.0-9.0 supplemented with 1 M L-lysine. The influence of silica particle and pore size was additionally tested using Davisil 646 (15 nm pore size, 250-500 µm particle size), Davisil 62 (15 nm pore size, 75-250 µm particle size) and Davisil 643 (15 nm pore size, 35-70 µm particle size) as above except that protein elution was only conducted with 20 mM Tris-HCl, pH 8.25 supplemented with 1 M L-lysine.

2.2.3 Optimizing small-scale purification

Seed cultures (5mL) were used to inoculate 25 ml of LB medium supplemented with 50 µg/mL kanamycin or 100 µg/ml carbenicillin. Mid-exponential phase cultures were treated with 1 mM IPTG (pET derivatives) or 0.2% arabinose (pBLN200 derivatives) and recombinant

proteins were allowed to accumulate for 3h at 37°C, or for 5h at 30°C in the case of Car9-β-lactamase due to its lower solubility. Culture samples (5ml) were subjected to centrifugation at 7000*g* for 10 min. The paste was resuspended in 3 mL of 20 mM Tris-HCl, pH 7.5, 2 mM EDTA, and cells were disrupted with a French press operated at 10000 psi. Lysates were clarified by centrifugation at 10000*g* for 10 min. To optimize operational parameters, 1 mL of silica slurry and 3 mL of clarified lysate from cells expressing sfGFP-Car9 were added to conical 0.8 x 4 cm polypropylene disposable columns (BioRad). After 10 min of mixing at 23°C on a rotary mixer, a stopcock and syringe were connected to the outlet of the column and contaminating proteins were removed by aspirating 15 mL of 20 mM Tris-HCl, pH 7.5 through the bed. Successive elution fractions (0.5 mL) were collected by aspiration after supplying the column with 5 mL of 20 mM Tris-HCl, pH 8.25 supplemented with 0.5 M or 1 M L-lysine. The effect of glycerol and NaCl was tested by adding 10% glycerol or 250 mM NaCl to the load, wash and elution buffer. For these experiments, elution was conducted with 20 mM Tris-HCl, pH 8.25 supplemented with 0.5 M lysine. The effect of Tween 20 was tested by preparing clarified extracts of sfGFP-Car9, Car9-β-lactamase, MBP-Car9 and mCherry-Car9 as above and adding 0.1 or 0.3% (v/v) of the detergent to the lysis buffer. Target proteins were eluted as above.

2.2.4 Removal of N-terminal Car9 tags with TEV protease

Car9-β-lactamase and Car9-sfGFP purified using the small-scale system under optimized conditions were dialyzed against 2 L of 20 mM Tris-HCl, pH 7.5, and concentrated with 10-kDa cutoff ultracentrifugal filters (Amicon). Protein concentration was determined using a BCA Assay (Sigma) with BSA as a standard. TEV protease with an N-terminal His tag [115] was purified by Ni-NTA affinity chromatography. The protease (1.5 μg) was mixed with 25 μg of

target proteins (0.5 μM TEV to ~8 μM protein) in the presence of 5 mM 2-mercaptoethanol. All digestions were conducted for 16 h at 23°C. For experiments performed in the presence of lysine, 94 μL of the second eluted fraction was mixed with 1 μL of 1.7 mg/mL TEV protease and 5 μL of 20 mM Tris-HCl, 0.5 M L-lysine and 5 μL of 100 mM 2-mercaptoethanol in the same buffer for a final concentration 5 mM.

2.2.5 Purification of Car9-β-lactamase under denaturing conditions

Car9-β-lactamase was expressed as above except that the temperature was raised to 37°C to promote aggregation (~65% insoluble). Culture samples (5ml) were subjected to centrifugation at 7000*g* for 10 min. The paste was resuspended in 3 mL of 20 mM Tris-HCl, pH 7.5, 2 mM EDTA, and cells were disrupted with a French press operated at 10000 psi. Lysates were centrifuged at 1500*g* for 10 min to remove unbroken cell and the supernatant was subjected to centrifugation at 10000*g* for 10 min. The pellet was washed in 3 mL of 20 mM Tris-HCl, pH 7.5, 2 M urea and 0.5% Tween and with 20 mM Tris-HCl, pH 7.5 with intervening cycles of centrifugation at 10000*g* for 10 min. Washed inclusion bodies were solubilized in 20 mM Tris-HCl, pH 7.5, 8 M urea and 0.3% Tween for 1 h at 23°C with gentle shaking. The solution was centrifuged at 10000*g* for 10 min to remove remaining aggregates. Unfolded proteins were contacted with silica slurry and subjected to small-scale purification as above, except that all buffers were supplemented with 8 M urea.

2.3 Results and Discussion

2.3.1 Car9 is functional as an N-terminal affinity tag

We previously reported that fusion of the Car9 sequence to the C-termini of sfGFP, MBP and mCherry enables efficient affinity purification of these proteins on silica gel when 1 M L-lysine

or L-arginine is used as an eluent [112]. To determine if the tag would be agnostic of placement, we fused the Car9 extension to the N-terminus of sfGFP, inserting a flexible GGGS linker and a TEV protease cleavage site between the two domains (**Figure 2.1 A**). Clarified extracts from cells overexpressing the Car9-sfGFP fusion protein were loaded on a Merck 7734 silica column equilibrated in Tris-HCl buffer, pH 7.5. The bound material was eluted with 1 M L-lysine after contaminating proteins had been removed by washing. This operation led to the recovery of 60-75% pure Car9-sfGFP (**Figure 2.1 B**), proving that Car9 is fully functional as an N-terminal affinity tag. Under the same experimental conditions, sfGFP-Car9 was recovered with a purity of 75-97% and with a slightly different pattern of contaminants (**Figure 2.1 C-D**). Considering that both fusion proteins are expressed at comparable levels, this result suggests that the orientation in which the Car9 sequence of amino acids is presented to silica, and/or the local structural or chemical context experienced by the tag, influences silica binding affinity. Of note, although both Car9-sfGFP and sfGFP-Car9 continued to elute after two column volumes of buffer had flowed through, it was possible to achieve quantitative release of sfGFP-Car9 by increasing the lysine concentration to 2 M (**Figure 2.2**). We therefore embarked on a campaign of buffer and stationary phase optimization to maximize the purity and yields of Car9-tagged proteins.

2.3.2 Alkaline conditions improve the efficiency of sfGFP-Car9 elution

Silica gels are amorphous materials with variable porosity in which silicon atoms are tetrahedrally coordinated to oxygen atoms through siloxane bonds (Si—O—Si). The surface of silica contains both hydrophobic siloxane sites [116], and unreacted, hydrophilic hydroxyls called silanols (Si—OH). These silanols can be found in isolated (single), geminal (silanediol) or vicinal (H-bonded) forms. The concentration and distribution of these species depends on the mode of silica preparation and influences adsorption properties [46]. In addition, deprotonation of

silanol groups at pH higher than ~3 (the point of zero charge for silica) leads to their ionization to SiO^-, and to the development of an increasingly negative surface charge as the solution becomes more basic [47-50]. For instance, the degree of ionization of 80 nm SiO_2 nanoparticles increases from 9% at pH 5, to 18% at pH 7.4, reaching a maximum of 50% at a pH somewhat above 8.5 [48,50].

The nonspecific binding of proteins to silica is dominated by a combination of electrostatic and hydrophobic interactions, although H-bonding, van der Waals interactions and structural flexibility also play significant roles [51-54,117]. Similar mechanisms appear to control the binding of combinatorially-selected silica-binding peptides to SiO_2 [48,118]. In these specific adsorption schemes, binding affinity can be influenced by the extent of surface ionization. As an example, molecular dynamics (MD) simulations have revealed that a silica-binding peptide of sequence KLPGWSG strongly adsorbs to SiO_2 through electrostatics and H-bonding under conditions of high silanol ionization, but exhibits much weaker affinity for less ionized surfaces [50]. Not surprisingly, the adsorption of L-lysine to silica also increases with the solution pH, and therefore with the extent of surface ionization [119]. On the basis of these reports and our empirical observation that the release of Car9-tagged proteins entrapped within silica sol-gels improves under basic conditions [60], we reasoned that lysine might be most effective at disrupting Car9-silica interactions under conditions of high pH where electrostatics dominate.

To test this hypothesis, clarified extracts from cells overexpressing sfGFP-Car9 were loaded onto Merck 7734 silica, and the matrix was washed to remove nonspecifically bound proteins as above. Lysine elution (1 M) was conducted on replicate samples over a range of 3 pH units. **Figure 2.3 A** shows that elution under neutral or acidic conditions significantly reduced the recovery yields of sfGFP-Car9, with low amounts of protein eluting over many fractions.

Conversely, increasing the pH of the elution buffer to 8.0-8.5 led to quantitative release of sfGFP-Car9 in a small number of highly concentrated and fairly pure fractions (**Figure 2.3 B**). To minimize potential problems with protein unfolding under alkaline conditions, we selected an elution pH of 8.25 for all follow-on experiments.

2.3.3 Small size, large pore silica particles are optimal for the purification of Car9-tagged proteins

A wide variety of inexpensive silica gels with different particle and pore sizes are commercially available. While the experiments of **Figures 2.1** - **2.3** were conducted with 63-200 μm particles with 6 nm pores (Merck 7734), we used three additional silica gels to further investigate the influence of size distribution and porosity on sfGFP-Car9 recovery.

Figure 2.4A shows that the elution profile of sfGFP-Car9 from Davisil 636 was similar to that observed with Merck 7734 silica (**Figure 2.3A**, pH 8), except that less protein was present in successive fractions. Considering that the two silica gels have the same porosity, this result is well explained by a decrease in surface area available for sfGFP-Car9 binding on the larger Davisil 636 particles. Conversely, increasing the pore size to 15 nm while maintaining particle size at 250-500 μm provided additional binding sites for the Car9 extension, and led to more concentrated fractions (**Figure 2.4 B**). Finally, and as expected from the comparison of Merck 7734 and Davisil 636 matrices, we observed a further improvement in yields with Davisil 643 whose particles have the same 15 nm pore size as Davisil 646 but are an order or magnitude smaller (**Figure 2.4 C**). Although the use of that small particle size, large pore silica was accompanied by an increase in nonspecific protein binding and a shift to higher elution volumes, the Davisil 643 matrix provided the best outcome (defined here as a high concentration of pure protein over a small number of fractions) and was selected for subsequent experiments.

2.3.4 Optimization of the small-scale Car9 purification technology

We previously demonstrated that the Car9 affinity purification technology can be implemented in a small scale, disposable format that enables rapid and inexpensive protein purification [112]. We improved on our original design as described in Materials and Methods, building a kit that uses 1 mL of Davisil 643 slurry. In an effort to minimize the amount of costly L-lysine used in the elution step, we first investigated how reducing the concentration of this amino acid from 1 M to 0.5 M would impact sfGFP-Car9 elution at pH 8.25. **Figure 2.5** shows that the use of the lower lysine concentration led to a 0.5 mL delay in the onset of sfGFP-Car9 elution and to a slight broadening of the elution peak. A likely explanation for this result is that too few lysine molecules are initially available to occupy all Car9 binding sites in the region of the bed that is first percolated by the eluent. As the mobile phase brings in additional lysine molecules, competition for site occupancy increases and Car9 fusion protein rebinding events are progressively eliminated in the local zone. The phenomenon repeats itself down the length of the bed, causing a lag in elution. Making use of smaller silica particles similarly delays sfGFP-Car9 breakthrough because it increases the number of Car9 binding sites that require lysine passivation (**Figure 2.4 B-C**). Bearing in mind that the economic benefits of a lower lysine concentration far outweigh the inconvenience of collecting more fractions, we conducted all subsequent optimization steps using an elution buffer supplemented with 0.5 M L-lysine.

We next turned our attention to the suppression of nonspecific binding events. As mentioned earlier, proteins adsorb to silica nonspecifically through a combination of electrostatic and hydrophobic interactions that are supplemented by H-bonding and van der Waals interactions [51,54,117]. We therefore explored how the presence of NaCl, a charge-screening agent that modulates protein-protein and protein-silica interactions [120,121], and that of glycerol, a co-solvent

that increases protein compactness [122] and interacts with large patches of contiguous surface hydrophobicity [123], would impact sfGFP-Car9 purity.

Supplementing the lysis, wash and elution buffers with 250 mM NaCl had little impact on sfGFP-Car9 purity. However, inclusion of 10% glycerol improved the purity of both sfGFP-Car9 (**Figure 2.6 A**) and MBP-Car9 (**Figure 2.6 B**). By contrast, while a Car9-b-lactamase fusion protein efficiently bound to silica (confirming that Car9 is functional as an N-terminal tag), glycerol did not appreciably suppress nonspecific protein binding and the protein eluted with very low purity (**Figure 2.6 C**). We attribute this outcome to the fact that sfGFP-Car9 and MBP-Car9 represent 30-40% of the total protein in clarified cell lysates while Car9-b-lactamase only accounts for ~15%. Because they are present at higher molarities, sfGFP-Car9 and MBP-Car9 will occupy 2-to-3 times as many Car9 binding sites as Car9-b-lactamase will, and possibly more since tethered moieties might block neighboring sites via steric hindrances. Under such conditions of high silica surface coverage, glycerol effectively suppresses residual interactions between host proteins and silica. The beneficial effect of the additive is however lost when fewer Car9-b-lactamase molecules bind to the surface as there are more opportunities for host proteins to adsorb/unfold at the silica interface. In agreement with this hypothesis, reducing the number of Car9 binding sites by making use of smaller pore, larger size Davisil 636 particles significantly improved the purity of Car9-b-lactamase (**Figure 2.7 A**). The use of this matrix was however not pursued since it led to an unacceptable reduction in binding capacity (about 50% for sfGFP-Car9 based on fluorescence; **Figure 2.7 B**).

Tween 20 is a nonionic polysorbate surfactant that is commonly used as a blocking agent in the immunological detection of proteins and to prevent non-specific binding in protein purification schemes [37,124]. Because it lacks a charged head group, Tween 20 has a much lower

critical micelle concentration that anionic detergents, and because it cannot reach a high enough monomer concentration to cooperatively bind to proteins, it does not promote their denaturation [125]. At concentrations lower than ~0.4% (v/v), Tween 20 strongly binds to silica through its polyol head, producing a hydrophobic surface studded with progressively denser alkyl chains as the concentration increases [126].

To determine if the surfactant would improve the purity of Car9-β-lactamase, we processed cell paste in the presence of 0.1 or 0.3% (v/v) Tween 20 and contacted clarified lysates with silica slurry as above. The surfactant was not included in the wash and elution buffers to minimize carry-over concentration in eluted fractions. **Figure 2.8 A** shows that whereas the use of 0.1% Tween 20 caused a slight improvement in Car9-β-lactamase purity, a significant amount of contaminating proteins still eluted with the target protein. By contrast, using 0.3% Tween 20 yielded an ~90% pure target. Importantly, the surfactant did not interfere with Car9-β-lactamase binding, as little protein was found in the flow-through (**Figure 2.8 A**). The observed improvement in purity may be related to the effective passivation of silica at high Tween concentrations through the formation of irreversible complexes between hydrophobic regions of host proteins and projecting surfactant brushes. The remaining contaminant at ~43-kDa is likely to be an RNA-binding protein (a class of polypeptides prone to nonspecific interactions with silica [53]) that is abundant enough to compete with both Tween molecules and Car9 fusion proteins for binding to silica.

Finally, we verified that supplementation of the lysis buffer with 0.3% Tween 20 did not interfere with the recovery of highly expressed proteins (sfGFP-Car9 and MBP-Car9), and demonstrated that it was beneficial to the purification of another low-expressing protein (mCherry-Car9). As shown in **Figure 2.8 B**, all targets could be recovered with purities ranging

from 93 to 99%, as judged by videodensitometric analysis of the gels. However, we found by measuring GFP fluorescence in the various fractions that inclusion of 0.3% Tween reduced the binding of the high-expressing sfGFP-Car9 to silica by 23% and that an additional 9% of the material was lost during column wash, reducing overall yields by ~35% (from about 1 mg to 0.63 mg for a 5 mL culture).

2.3.5 The Car9 tag is processed by the TEV protease in the absence and presence of lysine

Removal of affinity tags is necessary if a protein with (near) native sequence is required, or if the tag interferes with structure or function [30]. Previously, we have shown that it is possible to excise a C-terminal Car9 tag preceded by paired basic residues by incubating purified fusion protein with whole *E. coli* cells expressing the outer membrane protease OmpT on their surface [112]. Although the approach may reintroduce contaminating proteins in the preparation and could lead to nonspecific cleavage, it is effective, inexpensive, and only extends the target protein by one basic residue.

For the removal of N-terminal tags, the TEV protease is often preferred because of its stringent specificity, and the fact that in only adds a single serine residue to the N-terminus of the target [111]. To determine if TEV would process the Car9 extension, we purified Car9-sfGFP and Car9-β-lactamase (which both contain a TEV recognition sequence downstream of Car9; **Figure 2.1 A**) using small-scale purification under optimized conditions. **Figure 2.9** shows that the TEV protease processed about 80-90% of Car9-β-lactamase and 65% of Car9-sfGFP under our experimental conditions whether or not digestion was conducted in the presence of 0.5 M lysine.

2.3.6 The Car9 tag retains partial functionality under denaturing conditions

A relatively small number of affinity tags support protein purification under denaturing conditions either because the tag does not require an ordered conformation to bind to

immobilized ligands (e.g., polyhistidine [37], polyarginine [127] and polylysine [128] tags), or because it remains structured under conditions where most polypeptides unfold (e.g., cellulose-binding domain [129]). The Si-tag, which is believed to be disordered [130], has also been found to function under denaturing conditions [124], encouraging us to explore the possibility that Car9 might be useable in the presence of high concentrations of chaotropic agents.

To test this hypothesis, we unfolded Car9-b-lactamase inclusion bodies in 8M urea and applied the solubilized material to a small-scale silica column, conducting all steps in the presence of the same concentration of denaturant. Videodensitometric analysis of the various fractions revealed that about 70% of the protein was lost during the load and wash steps (**Figure 2.10**). However, the remainder of the material bound to the matrix and could be eluted free of contaminants with 0.5 M lysine (**Figure 2.10**). Because high urea concentrations do not prevent the Si-tag from binding to silica [124], the inefficient capture of Car9-b-lactamase under denaturing conditions is unlikely to be due to silica-chaotrope interactions. Rather, we believe that high-affinity binding of the Car9 dodecapeptide requires it adopts a specific conformation.

2.4 Conclusions

Building on a previous study [112], we have shown that Car9 is versatile affinity tag that supports protein purification on unmodified and inexpensive silica gels. The tag may be appended to the C- or N-termini of target proteins. It is excisable from the latter position via TEV protease digestion, and it retains partial silica-binding affinity under denaturing conditions. We have further demonstrated that capture of Car9-tagged proteins can be enhanced by appropriate selection of the silica matrix, and that efficient elution can be achieved at reduced lysine concentration and under alkaline conditions. Finally, we have determined that inclusion of 0.3% Tween 20 in the lysis buffer leads to product purities higher than 90%. An optimized

small-scale purification kit incorporating the above embellishments enables the recovery of high-purity Car9-tagged proteins in minutes and at very low cost. The Car9 affinity purification technology should prove particularly valuable for laboratory-scale purification from small culture samples (5 mL), and for commercial scale purification since both processes are sensitive to cost and time.

2.5 Figures

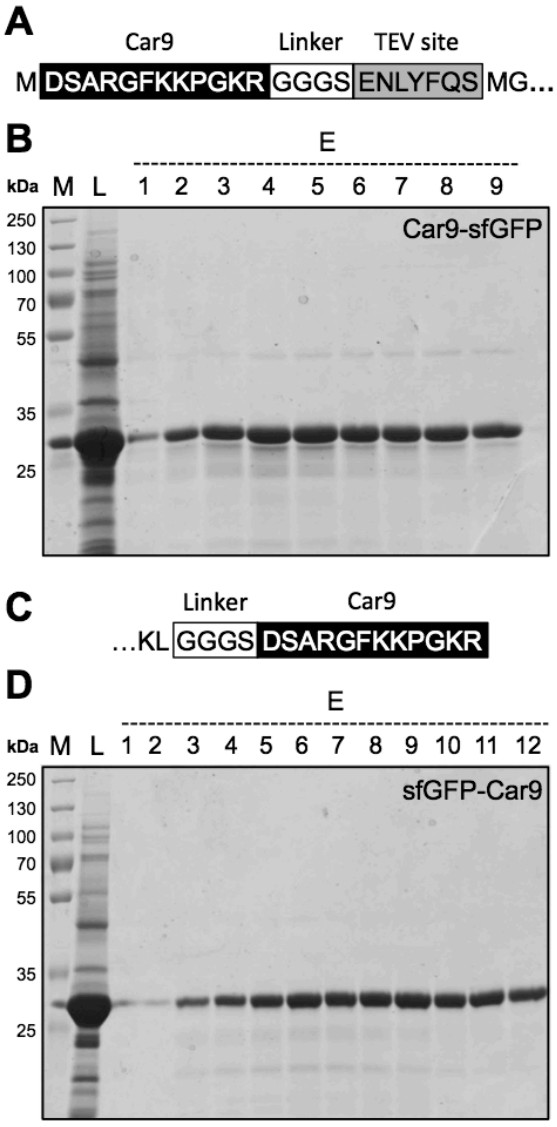

Figure 2.1 Influence of Car9 placement **(A)** Structure of the N-terminal Car9 extension. The sfGFP sequence starts with Met-Gly. **(B)** Purification of Car9-sfGFP on Merck 7734 silica with 1M lysine elution at pH 7.5. **(C)** Structure of the C-terminal Car9 fusion. The sfGFP sequence ends with Lys-Leu. **(D)** Purification of sfGFP-Car9 on Merck 7734 silica with 1M lysine elution at pH 7.5. Lanes: M, markers; L, load; E, successive elution fractions.

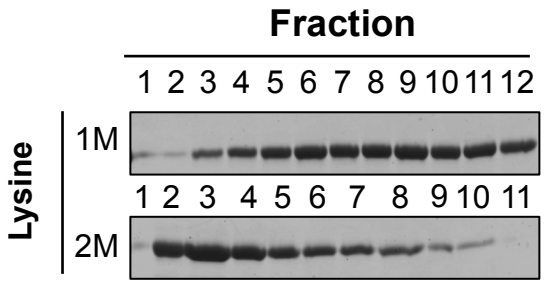

Figure 2.2 An increase in the elution buffer lysine concentration leads to quantitative elution of bound sfGFP-Car9. BL21(DE3) cells overexpressing sfGFP-Car9 were disrupted by sonication and clarified extracts were contacted for 5 min with 1.5 mL of Merck 7734 silica slurry and loaded in a 0.5 cm ID column. Nonspecifically bound proteins were removed by aspirating 10 mL of 20 mM Tris-HCl, pH 7.5 through the bed. The target protein was eluted in the same buffer supplemented with 1 or 2M Lysine. Successive fractions (~0.5 mL) were fractionated by SDS-PAGE.

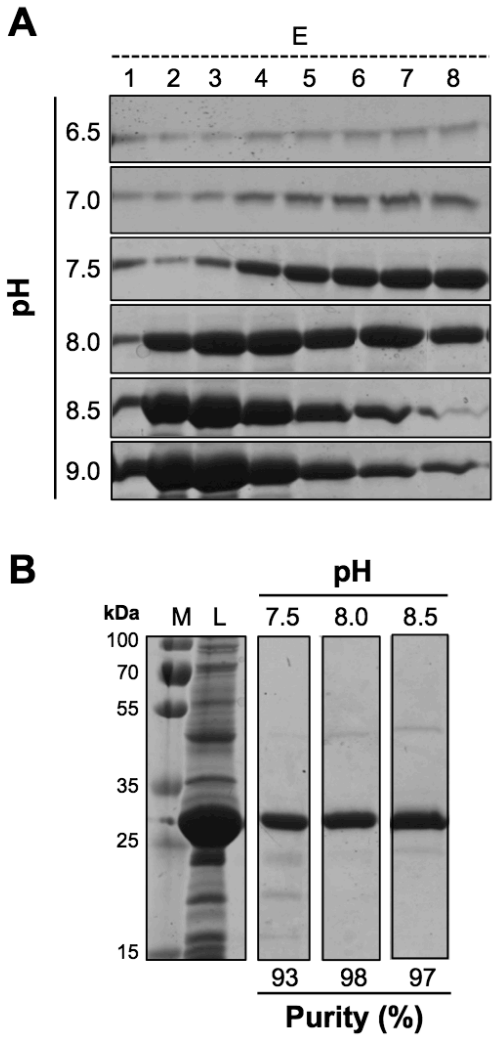

Figure 2.3 Influence of elution pH (**A**) SDS-PAGE analysis of sfGFP-Car9 fractions eluted with 1M lysine at the indicated pH. Successive elution fractions (E) are shown. (**B**) Comparison of the purity of eluted fractions at comparable sfGP-Car9 concentrations, as determined by videodensitometric analysis.

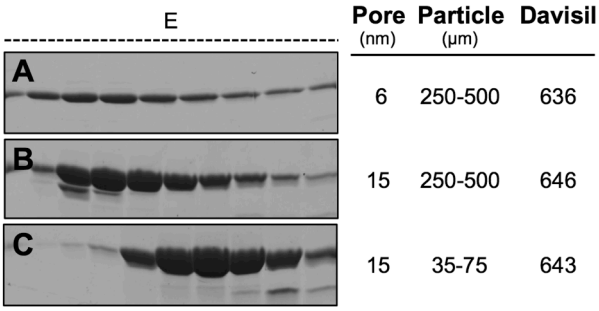

Figure 2.4. Influence of pore and particle size. Clarified extracts from cells overexpressing sfGFP-Car9 were loaded onto Davisil 636 (**A**), Davisil 646 (**B**) or Davisil 643 (**C**). The fusion protein was eluted with 1 M lysine in 20 mM Tris-HCl, pH 8.25. Successive elution fractions (E) are shown, as are the characteristics and commercial names of the silica gels.

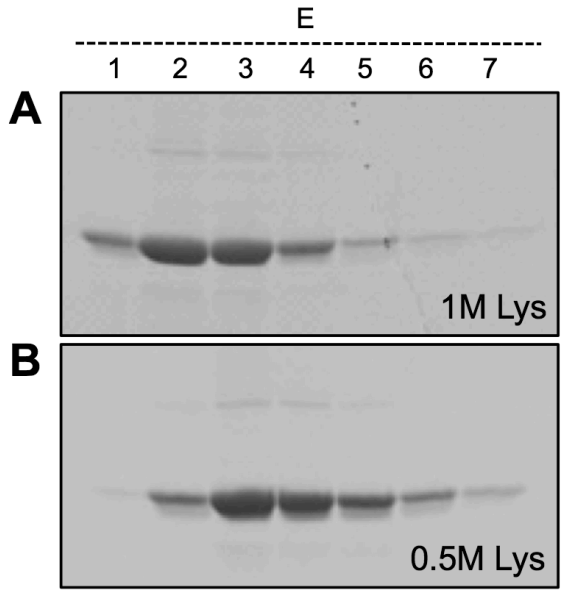

Figure 2.5 Influence of the lysine concentration on the elution of sfGFP-Car9 in a small-scale purification system. Clarified extracts from cells expressing sfGFP-Car9 were loaded onto 1 mL of Davisil 643 slurry and the protein was eluted with 1 M (**A**) or 0.5 M lysine (**B**) in 20 mM Tris-HCl buffer pH 8.25.

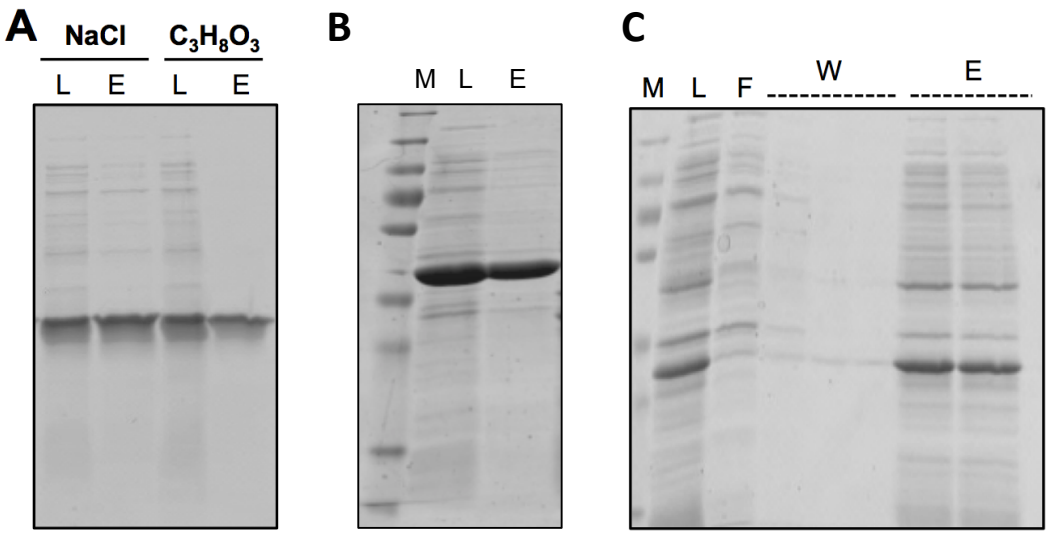

Figure 2.6 Influence of NaCl and glycerol on nonspecific protein binding. (**A**) Clarified extracts from cells expressing sfGFP-Car9 were loaded onto 1 mL of Davisil 643 slurry and the target protein was eluted in 20 mM Tris-HCl buffer pH 8.25 supplemented with 0.5 M lysine and 250 mM NaCl (NaCl) or 10% glycerol ($C_3H_8O_3$). (**B**) Rapid purification of MBP-Car9 in glycerol-containing buffers. Lanes: M, markers; L, load; E, material eluted with 0.5 M lysine. (**C**) Clarified extracts from cells expressing Car9-b-lactamase were loaded and eluted as above using buffers supplemented with 10% glycerol.

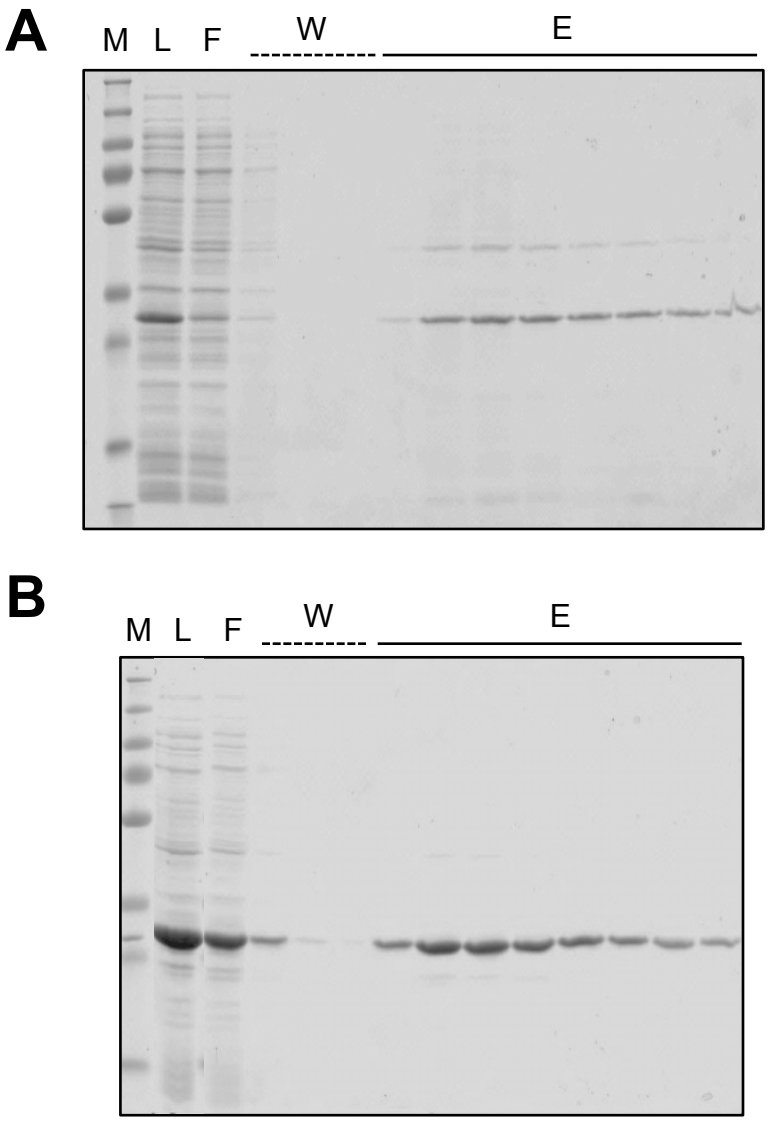

Figure 2.7 Rapid purification of Car9-β-lactamase (A) and sfGFP-Car9 (B) on Davisil 636 silica (250-500 μm particles, 6 nm pores).

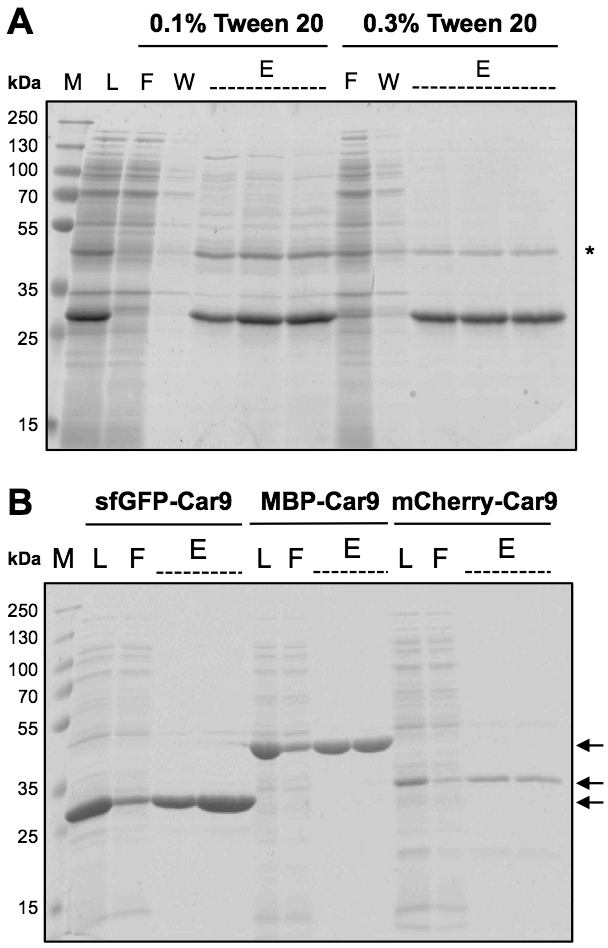

Figure 2.8 Influence of Tween 20 on nonspecific protein binding. (**A**) Clarified extracts from cells expressing Car9-b-lactamase were prepared in lysis buffer supplemented with 0.1% or 0.3% (v/v) Tween 20. Proteins were loaded onto 1 mL of Davisil 643 slurry and elution was conducted with 0.5 M Lysine in 20 mM Tris-HCl, pH 8.25. The star denotes the main contaminant. (**B**) Clarified extracts from cells expressing sfGFP-Car9, MBP-Car9 or mCherry-Car9 were prepared in lysis buffer supplemented with 0.3% Tweeen 20 and loaded and eluted as above. Arrows indicate the migration positions of the purified proteins.

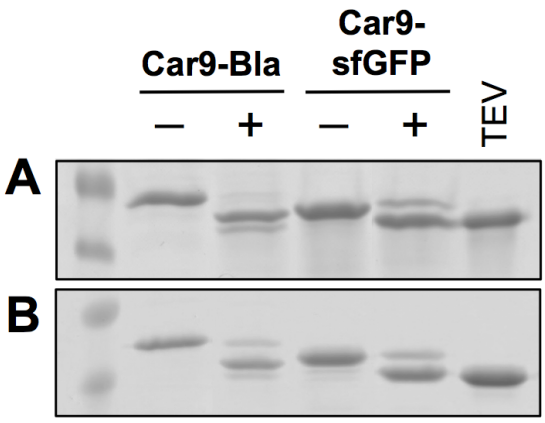

Figure 2.9 Excision of the Car9 tag by the TEV protease before (**A**) or after (**B**) removal of lysine by dialysis. Samples were supplemented (+) or not (-) with 0.5 µM of TEV protease and digested were conducted for 16h at room temperature. TEV lanes contain purified TEV protease.

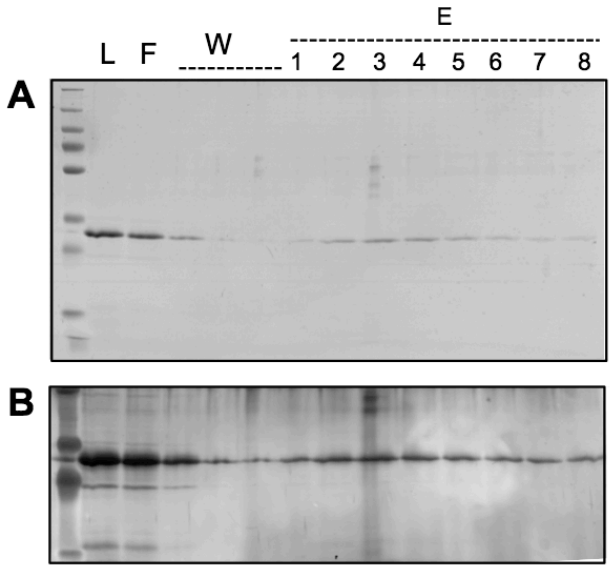

Figure 2.10 Coomassie blue- (**A**) and corresponding silver-stained (**B**) minigel of various fractions collected during the affinity purification of Car9-b-lactamase under denaturing conditions. Lanes: L, load; F, flow through; W, wash; E, successive elution fractions (0.5 mL). All steps were conducted in the presence of 8M urea.

Chapter 3 A palladium-binding deltarhodopsin for light-activated conversion of protonic to electronic currents

3.1 Introduction

Efficient integration of biology and electronics in the fast-paced field of bionanoelectronics[85,131,132] requires controlled conversion of biochemical ionic gradients into electrical signals that can be read or processed by conventional electronics. Here, we set out to integrate a genetically engineered archaeal bacteriorhodopsin with palladium hydride (PdHx) contacts to convert light-driven protonic currents into electronic signals. PdH_x[100-106] is used as a contact in protonic devices and converts a H^+ current into an e^- current according to:[108]

$$Pd + H^+ + e^- \rightleftharpoons PdH_x$$

In an effort to improve biotic-abiotic integration, we fuse a palladium (Pd) binding peptide to the N-terminus of *H. turkmenica* detarhodhopsin (HtdR), use genetic strategies to express the fusion protein at high level in the inner membrane of *E. coli* and reconstitute the protein into proteoliposomes that spontaneously fuse to form a continuous bilayer at the surface of a Pd electrode within a microfluidic device. By monitoring the electronic current that directly correlates with proton reduction and hydrogen diffusion into the Pd/PdHx "protode", we demonstrate that addition of the Pd-binding tag to HtdR leads to a nearly one order of magnitude increase in electronic signal with light-controlled "on" and "off" responses that are complete in seconds.

3.2 Materials and Methods

3.2.1 DNA manipulations

Plasmid pHtdR400 which encodes HtdR-His_6 under control of a mutant P_{BAD} promoter was described elsewhere.[76] To construct Pd4-HtdR-His_6, forward primer 5'- TAT GAC CAG CAA CGC GGT GCA TCC GAC CCT GCG TCA TCT GGG AGG CGG TAG CTG TTA CGC TGC TCT AGC ACC ACC - 3' and reverse primer 5'- CAT GGG TGG TGC TAG AGC AGC GTA ACA GCT ACC GCC TCC CAG ATG ACG CAG GGT CGG ATG CAC CGC GTT GCT GGT CA - 3' encoding the Pd4 dodecapeptide followed by a GGGS linker and flanked by *Nde*I and *Nco*I restrictions sites were annealed, digested with *Nde*I and *Nco*I and ligated at a 20:3 volume ratio with the large fragment of pHtdR400 that had been cut with the same enzymes. The structure of the fusion joint was confirmed by DNA sequencing and the resulting plasmid was named pPd4-HtdR.

3.2.2 Protein expression and purification

E. coli KTD101 cells [Δ(*araD-araB*)*567* Δ*lacZ4787* (::*rrnB-3*) l- *rph1* D(*rhaD-rhaB*)*568*, *hsdR514*, Δ*tig100*] harboring pPd4-HtdR or pHtdR400 were grown at 37°C in 2 L shake flasks containing 500 mL of LB media supplemented with 50 μg/mL kanamycin to mid-exponential phase (A_{600} ~ 0.45). Cultures were supplemented with 10 μM of all-*trans* retinal (Sigma) and protein synthesis was induced by addition of 0.2% *L*-arabinose. Cells were harvested 3 h post-induction growth at the same temperature (37° C) and disrupted with a French press operated at 10,000 psi. Aggregated material was removed by centrifugation at 14,000*g* for 10 min at 4°C. Supernatants were harvested and centrifuged at 150,000*g* for 1 h at 4°C to recover membrane fractions.[76] Pellets were resuspended in Buffer I (50 mM 2-(*N*-morpholino) ethanesulfonic acid [MES], pH 6.5, 300 mM NaCl, 5 mM imidazole, 1.0% n-dodecyl b-D-

maltoside [DDM]). Samples were incubated overnight at 4°C with gentle shaking. Solubilized membranes were loaded onto a 5 mL Ni-NTA chromatography column equilibrated in Buffer II (50 mM MES, pH 6.5, 300 mM NaCl, 50 mM imidazole, 0.1% DDM) and washed extensively with Buffer II to remove non-specifically bound proteins. Target proteins were eluted in Buffer III (50 mM Tris-HCl, pH 7.2, 300 mM NaCl, 250 mM imidazole, 0.1% DDM) [133]. Samples were diluted to a 50 mM imidazole concentration using Buffer III without imidazole and reloaded onto an Ni-NTA column equilibrated as above. Wash and elution steps were performed as above. Final purity was about 90 % and final yield about 5 mg per run. Proteins were stored at -80^{o}C at a final concentration of 0.2 mg/ml. Before use, proteins were concentrated to 0.55 mg/mL with a 10-kDa cutoff centrifugal concentrator. Protein concentrations were determined using the BCA assay (Thermo Fisher Scientific) with Bovine Serum Albumin as a standard.

3.2.3 Spheroplasting experiments

Samples (2 mL) from KTD101 (pHtdR400) and KTD101 (pPd4-HtdR400) cultures grown and induced as above were incubated with 20 µL of chloramphenicol (50 µg/mL) on ice for 10 min. Cells were centrifuged at 5000*g* for 5 min and pellets were resuspended in 200 µL of Buffer IV (100 mM Tris-HCl, pH 8.0, 500 mM sucrose, 0.5 mM EDTA). Samples were supplemented with 10 µL of a freshly prepared solution of 2mg/mL lysozyme followed by 400 µL of Buffer IV and 400 µL of ddH_2O. After 20 min incubation at room temperature, the tubes were centrifuge at 12800*g* for 30 sec. The supernatant was discarded and spheroplasts were resuspended in Buffer V (100 mM Tris-HCl, pH 8.0, 300 mM sucrose, 10 mM $MgCl_2$). Aliquots of vesicles (50 µL) were stained with 1 µL of a freshly prepared 50 µM stock solution of Nile Red (Sigma). Droplets (35 µL) were contacted for 30 min with ~50 nm thick Pd films deposited onto silicon wafers coated with a Cr adhesion layer. Substrates (~2x2 cm) were washed three

times with 15 mL of Buffer V with gentle shaking to remove nonspecifically bound spheroplasts. Samples were imaged in reflectance mode using a GE Typhoon FLA 9000 Gel Imaging Scanner with excitation at 532 nm.

3.2.4 Proteoliposomes preparation

Multilamellar vesicles (MLVs) were produced by dissolving 10 mg/mL of 1,2-Dioleoyl-*sn*-glycero-3-phosphocholine (DOPC, Avanti Polar Lipids) in a solution of chloroform/methanol at a 2:1 (v/v) ratio and by evaporating the solvents under a constant nitrogen stream. The dry lipidic film was then hydrated for 1h at room temperature and under vigorous shaking in 10 mM HEPES, pH 7.4, 150 mM NaCl to a final lipid concentration of 4 mg/mL (lipids). The hydrated solution was pre-filtered through a 0.2 µm cellulose acetate filter (VWR) and extruded through 200 nm polycarbonate membranes using an Avestin LiposoFast device to generate large unilamellar vesicles (LUVs).[134,135] Purified HtdR-His_6 and Pd4-Htdr-His_6 were reconstituted into liposomes using the three step detergent method[136,137]. Briefly, LUVs (4 mg/ml lipids) were solubilized with a saturating concentration (R_{sat}) of DDM, which corresponds to a detergent to lipid ratio of 1 mol/mol. After 2 h, membrane proteins were added at 1:40 (w/w) protein to lipid ratio (about 75 µL of a 0.55 mg/mL protein solution for a final volume 500 µL) and samples were incubated for 2 additional hours. The detergent was removed by contacting the samples with SM2 Bio-Beads (BioRad) for 2h. The hydrodynamic diameters of liposomes and proteoliposomes were determined by dynamic light scattering (DLS) on a Malvern Zetasizer instrument.

3.2.5 Device Fabrication and current measurements

Conventional photolithography procedures were used to fabricate devices consisting of 50nm-thick Pd contacts evaporated onto a Si wafer substrate with a 100 nm silicon oxide layer

and a 5nm Cr binding layer. A top layer of SU-8 leaves a microfluidic channel across multiple Pd micro-contacts. PDMS side walls were placed on top of microfluidic channel to provide enough room for an Ag/AgCl reference electrode to complete the circuit. Aliquots (200 μL) of control DOPC liposomes and proteoliposomes synthesized as above were deposited in the central microfluidic channel of the device that had been subjected to plasma cleaning to eliminate contamination and activate the Pd surface. (The hydrophilic activated Pd surface is thermodynamically attractive to the hydrophilic heads of DOPC vesicles and helps vesicles fuse to the Pd surface.)[138] The device was agitated during vesicle fusion which was allowed to proceed overnight at 4^{o}C. Current measurements were performed with a grounded semiconductor parameter analyzer (Agilent 4155C). Electrical contact between analyzer and device were made with a home-built probe station with micro probes contacting Pd pads out of the microfluidic channel. For illumination (typically 20 cycles at 10 s intervals), a 523 nm green LED (10 W) was held at 11V and 0.3A and 10 cm away from the device. The LED was allowed to warm up and reach constant output for 15-20 min before measurements were taken.

3.3 Results and Discussion

3.3.1 Construction of a Pd binding deltarhodopsin

HtdR is a light-activated outwards proton pump whose photocycle is similar to that of *H. salinarum* bR.[70] However, it can be expressed at much higher levels than bR in *E. coli*[70] and we previously described genetic strategies that further improve yields to values as high as 20 mg/L in shake flask cultures.[139,140] Starting with a variant of HtdR fitted with a C-terminal hexahistidine tag residing in the cytoplasm (HtdR-His_6), we used standard molecular biology to fuse the Pd4 palladium-binding peptide[141] of amino acid sequence Thr-Ser-Asn-Ala-Val-His-Pro-Thr-Leu-Arg-His-Leu to the N-terminus of HtdR through a flexible Gly-Gly-Gly-Ser linker.

If properly folded within the inner membrane of *E. coli*, the resulting tripartite fusion protein (Pd4-HtdR-His_6) should adopt the topology depicted in **Figure 3.1 A** with an N-terminal Pd-binding segment projecting into the periplasm.

Genes encoding HtdR-His_6 and Pd4-HtdR-His_6 were placed under transcriptional control of a mutant P_{BAD} promoter that improves the yields of integral membrane proteins by reducing the flux of unfolded species to the SecYEG translocon.[139] The resulting plasmids (pHtdR400 and pPd4-HtdR400, respectively) were introduced into KTD101, an *E. coli* strain that is beneficial to the production of polytopic membrane proteins by virtue of lacking the molecular chaperone trigger factor.[140] Bands corresponding to the expected migration positions of HtdR-His_6 and Pd4-HtdR-His_6 could be readily detected by SDS-PAGE analysis of whole cells harvested 3h post-induction of recombinant protein production with L-arabinose (**Figure 3.1 B**). Furthermore, when grown in the presence of all-*trans* retinal, both transformants exhibited the purple color characteristic of *E. coli* expressing biologically active rhodopsin (**Figure 3.1 C**). To determine how much material localized to the inner membrane, cultures were further processed into membrane and insoluble fractions. As expected,[140] most of the HtdR-His_6 and Pd4-HtdR-His_6 produced misfolded and accumulated as inclusion bodies in the insoluble fractions of the cell. However, in both cases, significant (and comparable) amounts of protein localized to the inner membrane (**Figure 3.1 D**). Thus, fusion of the Pd4 extension to the N-terminus of HtdR does not affect its ability to traffic to, and properly fold within, the inner membrane of *E. coli*.

3.3.2 Pd4 tagged HtdR binds to palladium

To test the functionality of the Pd4 extension, cultures producing HtdR-His_6 or Pd4-HtdR-His_6 were subjected to spheroplasting,[142] a procedure that strips the cells of most of their outer membrane and peptidoglycan and leads to the formation of spherical vesicles consisting of

cytoplasmic material bounded by the inner membrane (**Figure 3.2 A-B**). Spheroplasts were stained with the lipophilic dye Nile Red, imaged by fluorescence microscopy (**Figure 3.2 C-D**). Aliquots corresponding to identical amounts of fluorescence were next contacted with ~50 nm thick Pd films that had been deposited onto silicon and subjected to UV-ozone-cleaning. The slides were washed extensively with gentle shaking to remove non-specifically bound material and samples were imaged on a Typhoon scanner (GE Healthcare). **Figure 3.2 E-F** shows that while nearly all HtdR-His_6 spheroplasts were removed upon washing, Pd4-HtdR-His_6 spheroplasts were efficiently retained on the Pd film. We conclude that the Pd4 extension is properly translocated to the periplasmic side of the *E. coli* inner membrane where it remains functional for Pd binding.

3.3.3 Purification and reconstitution of HtdR and Pd4-HtdR

To purify Pd4-HtdR-His_6 and HtdR-His_6, induced cultures were lysed with a French press, aggregated species removed by centrifugation and membrane fractions collected by ultracentrifugation. Membranes were solubilized in buffer supplemented with 1% (wt/v) of the non-ionic detergent n-dodecyl β-*D*-maltoside (DDM). The detergent-extracted proteins were loaded onto a Ni-NTA chromatography column equilibrated in DDM-containing buffer and contaminants were removed by extensive washing. Target proteins were eluted with 250 mM imidazole. To further improve purity, samples were subjected a second round of Ni-NTA purification (**Figure 3.3 A**). The UV/visible spectra of these preparations exhibited the characteristic 550 nm absorption peak of functional HtdR, indicating that the detergent-extracted material was properly folded (**Figure 3.3 B**).

The purified membrane proteins were integrated within liposomes using detergent-mediated reconstitution,[143,144] a process that involves solubilizing preformed liposomes by detergent

addition, incubating the resulting vesicles with detergent-solubilized protein and, finally, removing the detergent to drive protein insertion within the liposomal membrane. Here, we used 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) and the thin film hydration method followed by extrusion through polycarbonate membranes to produce liposomes 210 ± 90 nm in hydrodynamic diameter. The vesicles were solubilized at saturating DDM concentration and purified HtdR-His_6 and Pd4-HtdR-His_6 in DDM-containing buffer were added at a 1:40 wt/wt ratio of proteins to lipids and incubated for 2h. The detergent was removed by 2h incubation with SM2-BioBeads (Bio-Rad) to produce pink proteoliposomes with hydrodynamic diameters comparable to those of the original vesicles (**Figure 3.4**).

3.3.4 Fabrication of a bioprotonic device

To compare the performance of HtdR-His_6 and Pd4-HtdR-His_6 in delivering protons to a Pd/PdHx electrode, we used conventional photo- and soft lithography to fabricate a microfluidic protonic device that will be fully described elsewhere. Each device consists of 50 nm-thick Pd contacts evaporated onto a Si wafer substrate with a 100 nm layer of silicon oxide and a 5 nm Cr adhesion layer. A top layer of SU-8 defines a microfluidic channel with PDMS side walls across multiple Pd micro-contacts (**Figure 3.5 A**). Devices were plasma-cleaned to remove contaminants and make the contacts hydrophilic before 200 µL of proteoliposomes (or control DOPC liposomes) were dispensed in the center channel. The vesicles were allowed to fuse overnight on the hydrophilic surface with gentle agitation to prevent aggregation. Finally, a reference Ag/AgCl electrode was introduced in the well to complete the circuit. The device operating principle is schematically illustrated in **Figure 3.5 B**. Illumination with green light (HtdR absorbs maximally at 550 nm)[70,140] excites the retinal molecules of HtdR, which initiates proton pumping across the bilayer. Translocated protons diffuse to the Pd/PdH_x surface. There,

incoming electrons reduce them to hydrogen, which inserts itself within the palladium lattice to form PdH_x. As a result, for every proton reduced at the Pd/PdH_x surface an electron is drawn from the leads and results in a measurable electronic current.

3.3.5 Protonic current measurements

Current measurements were taken in the dark while the Pd contact was kept under a constant applied voltage of -50 mV (vs. Ag/AgCl). This applied voltage prevents changes in ionic potential across the membrane due to electrochemical potential differences between the Pd and Ag/AgCl electrodes. After about 20 seconds, the current reached a constant value ranging between -2 and -6 pA. This variability is typical in such devices and likely reflects small differences in geometry and protein density. Devices were next illuminated with a 523 nm green LED at consecutive 10 s intervals. **Figure 3.5 C** (black trace) shows that the recorded current was constant in devices prepared with protein-free liposomes under both dark and light conditions. Thus, a continuous and electrically insulating bilayer is formed under our experimental conditions that shows low proton permeability and results in no current change coupled with illumination. When the experiment was repeated with a device fabricated with HtdR-His_6 liposomes, we measured a weak, illumination-correlated current on several contacts (**Figure 3.5 C**, blue trace). In sharp contrast, devices made with a Pd4-HtdR-His_6 bilayer yielded clear signals that were highly reproducible in shape and magnitude over repeated cycles of illumination (**Figure 3.5 C**, purple trace).

How does one explain the 125% improvement in maximum photocurrent between the two devices? Kinetic Monte Carlo[109] and Replica Exchange with Solvent Tempering (REST)[145] molecular dynamics simulations suggest that the Pd4 dodecamer lies in close proximity to metallic surfaces where it is anchored through its 6^{th}, and, to a lesser extent,[145] 11^{th} histidine

(His) residue. Thus, protons transported through Pd4-HtdR-His_6 should experience a significantly reduced diffusion path length before reaching the Pd/PdHx contact compared to those released in the interstitial space formed between the electrode and the lipid bilayer containing unmodified HtdR-His_6. We note that the Pd4 sequence does not contain aspartic (Asp) or glutamic (Glu) acid, two residues commonly used by natural proton channels because of their ability to effectively move protons along a transfer path through the rotation of their carboxyl groups about the terminal C-C bond of their side chains.[146] On the other hand, Pd4 specifies two threonines (Thr) and one serine (Ser) whose side chain hydroxyls might help proton transport by being part of a hydrogen-bonded chain involving water molecules. In short, we attribute the dramatic photocurrent improvement in devices built with Pd4-HtdR-His_6 to the Pd4-mediated formation of an improved contact between rhodopsin and Pd/PdHx electrode.

In these devices, illumination triggered a rapid current increase and maximum response was reached within 2s. The response to light being switched off was equally rapid (**Figure 3.5 C**). Interestingly, the photocurrent decreased to a steady state value corresponding to ~30% of the maximum signal after about 10 seconds of illumination (**Figure 3.5 D**; the smaller photoresponse in this device is within experimental sample variability and likely reflects higher device resistance or lower protein density). Furthermore, progress towards the development of a steady state could be entirely eliminated by cycling the device under shorter (1s) illumination times (**Figure 3.6**). Considering that bacteriorhodopsin photocyles are on a millisecond timescale,[71] the observed behavior is consistent with the progressive saturation of Pd sites available for H^+ reduction on the electrode.[147] Preliminary calculations suggest that as many as 80% of the available Pd sites in a single contact may become occupied by H in the 2s spike (see Appendix A). The rate of occupancy is expected to be significantly higher in Pd regions vicinal to Pd4-

HtdR-His_6 due to a high local proton concentration. Under this scenario, the steady state photocurrent would be determined by the rate-limiting diffusion of H into bulk Pd. Indeed, we have previously observed diffusion limited H^+ currents in PdH_x contacts with Nafion devices.[102] Further work is under way to more quantitatively understand the dynamics and efficiency of the photoresponse.

3.4 Conclusions

Our work shows that solid binding peptides, which can be isolated on virtually any material[148] and used to control adhesion, orientation, assembly and morphogenesis[1-3,5,6,149] also hold promise in bioelectronics. Indeed, by bringing HtdR in intimate contact with a Pd/PdHx protode, the Pd4 domain allowed for the prototyping of a responsive device that efficiently couples light-activated proton transport to an electrical signal. Although optimization of the photoresponse beyond a twofold increase in current will be desirable, such an approach to interface light harvesting, ionic transport and electronics should prove valuable for the construction of more complex devices including biotic-abiotic photodetectors.

3.5 Figures

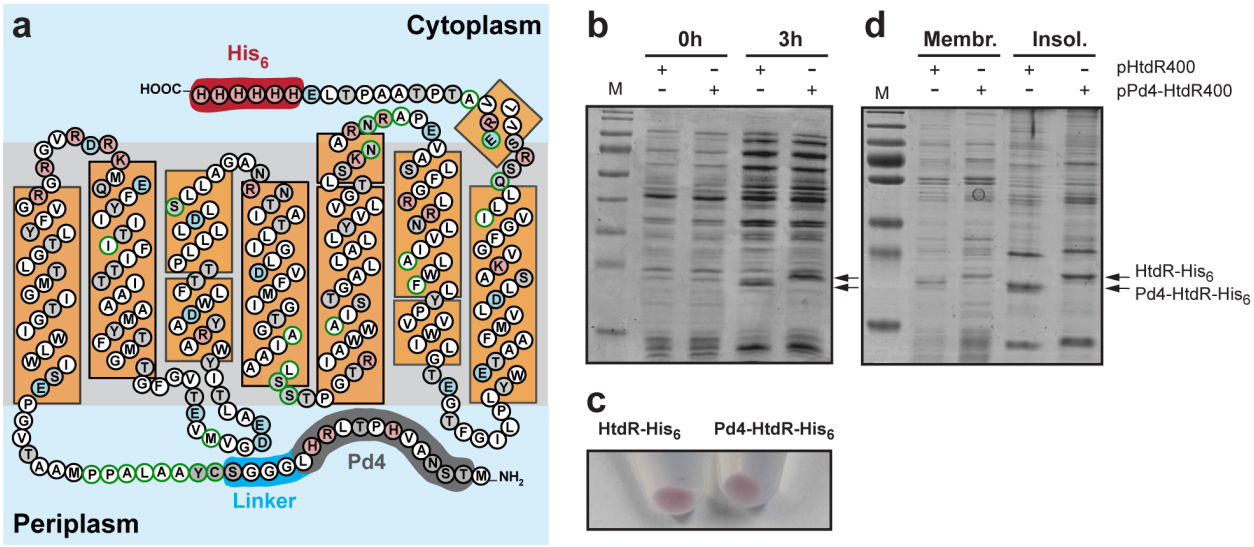

Figure 3.1 Structure and purification of Pd4-HtdR-His_6. **(a)** Predicted topology of Pd4-HtdR-His_6 in the *E. coli* inner membrane based on the crystal structure of its homolog dR3 (Ref. 72) The Pd4 palladium-binding sequence is identified in dark gray, the flexible linker in blue and the hexahistidine tag in red. Transmembrane segments are boxed and polar, acidic and basic amino acids are color-coded in light gray, blue and pink, respectively. Residues circled in green differ in HtdR and dR3. **(b)** KTD101 cells harboring plasmid pHtdR400 or pPd4-HtdR400 were grown to mid-exponential at 37°C in LB medium supplemented with 10 µM of all-*trans* retinal. Cultures were induced with 0.2% L-arabinose and recombinant proteins were allowed to accumulate for 3h at the same temperature. Whole cell samples are shown immediately before induction (0h) and 3h post-induction. Top and bottom arrows denote the migration positions of Pd4-HtdR-His_6 and HtdR-His_6, respectively. **(c)** The pink color of induced cells indicates the production of retinal-containing HtdR-His_6 and Pd4-HtdR-His_6. **(d)** Cultures collected 3h post-induction were fractionated into membrane (Membr.) and insoluble (Insol.) fractions.

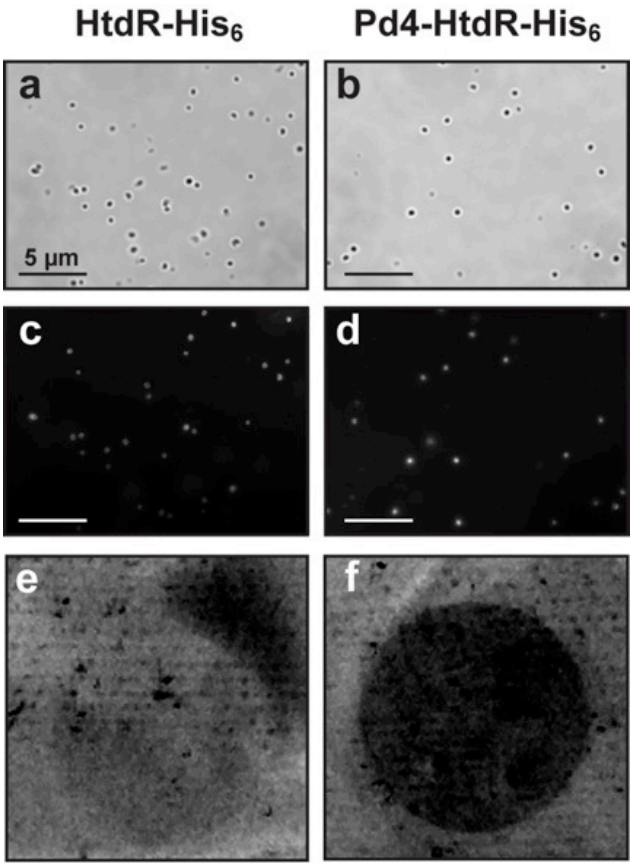

Figure 3.2 Pd4-HtdR-His_6 is functional for Pd-binding. KTD101 cells expressing Pd4-HtdR-His_6 or HtdR-His_6 were subjected to spheroplasting and examined by phase contrast (**a,b**) or fluorescence microscopy (**c,d**) following staining with Nile Red. Samples (35 µL) corresponding to identical amount of fluorescence were contacted with Pd for 30 min and washed to removed nonspecifically bound material. The contacts were examined in reflectance mode on a Typhoon scanner with excitation at 532 nm (**e,f**).

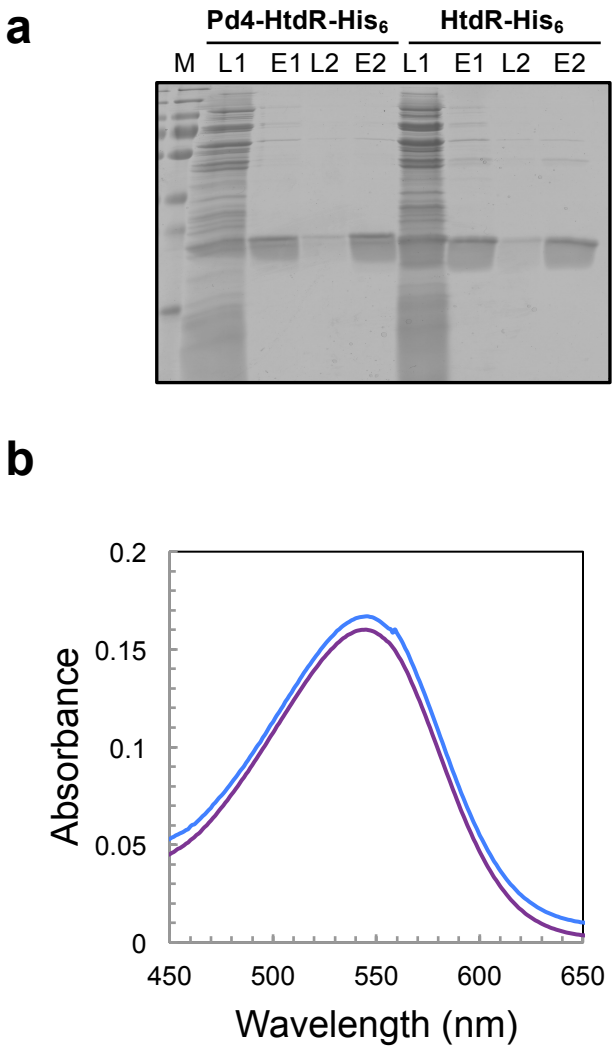

Figure 3.3 (**a**) Purification of Pd4-HtdR-His$_6$ and HtdR-His$_6$. Membrane proteins solubilized in DDM were loaded (lane L1) on a Ni-NTA column equilibrated as described in section 3.2. Proteins were eluted (E) with 250 mM imidazole. Eluted samples were diluted to reduce the imidazole concentration to 50 mM, loaded (L2) a second time onto Ni-NTA, and eluted with 250 mM imidazole (E2) to increase protein purity. Lane M contain molecular mass markers. The smearing is characteristic of membrane proteins fractionated on SDS-PAGE gels. (**b**) UV/Visible absorption spectra of DDM-solubilized and Ni-NTA purified HtdR-His$_6$ (blue) and Pd4-HtR-His$_6$ (purple).

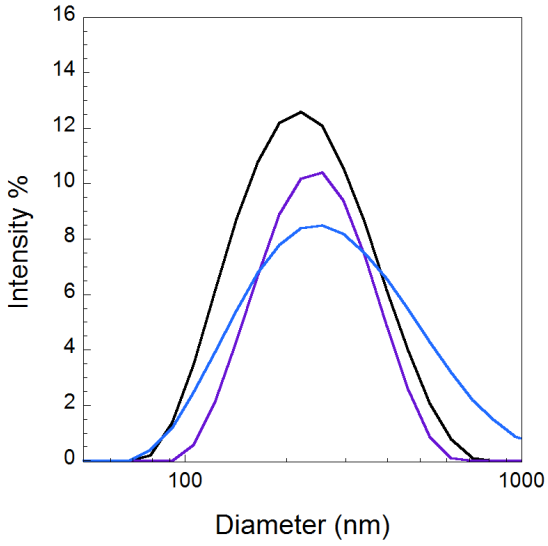

Figure 3.4. Dynamic light scattering (DLS) analysis of control DOPC liposomes (black trace), and of HtdR-His_6 (blue) and Pd4-HtR-His_6 (purple) proteoliposomes prepared as described.

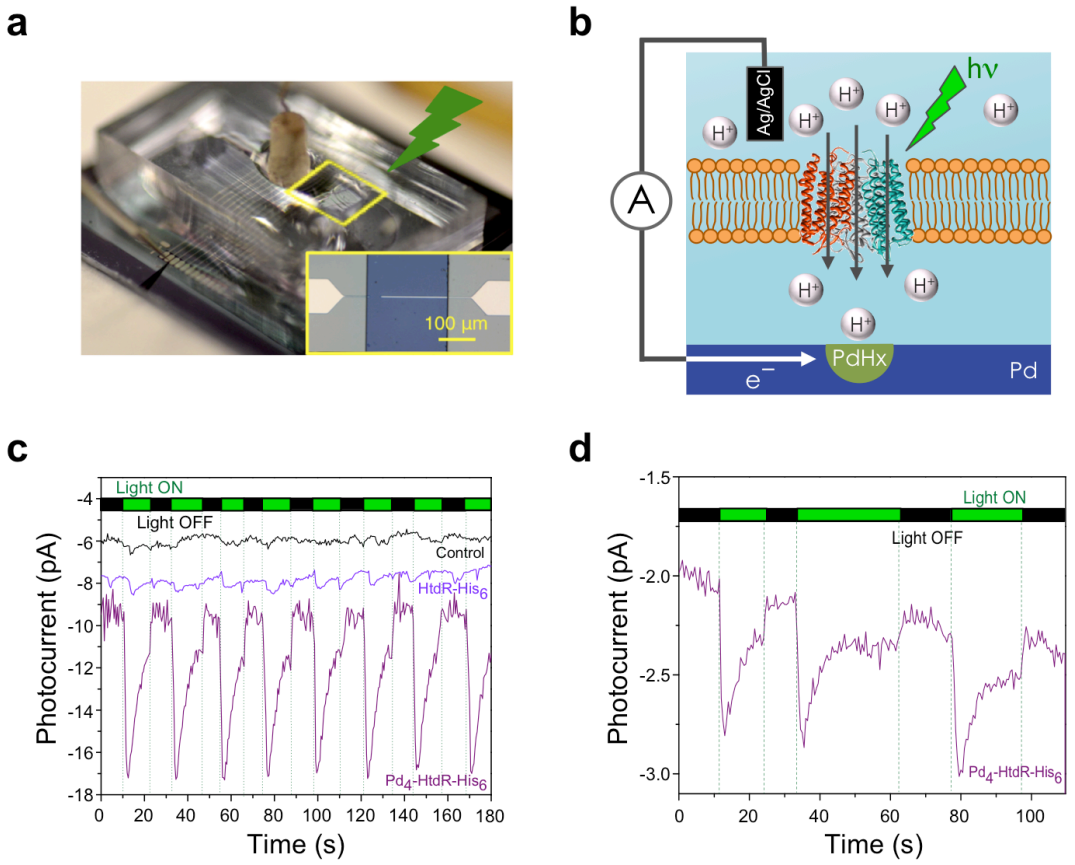

Figure 3.5 Light-activated conversion of protonic to electronic current. **(a)** Photograph of the bioprotonic microfluidic device. Each chip consists of multiple, lithographically fabricated Pd micro-contact inside a microfluidic channel (inset). The PDMS side walls provide enough room for an Ag/AgCl reference electrode that is immersed in the buffer to complete the circuit. **(b)** Operating principle of the device. **(c)** The photocurrent response of devices prepared with DOPC control liposomes (black trace), HtdR-His6 (blue) or Pd4-HtdR-His6 proteoliposomes (purple) was recorded while the Pd contact was kept at an applied voltage of -50 mV vs. Ag/AgCl and exposed to consecutive 10 sec illumination cycles with a 523 nm LED. Traces are offset by -2 pA for clarity. **(d)** A device with Pd4-HtdR-His6 proteoliposomes was illuminated for 10, 30 and 20 sec to better show the onset of the steady state. The lower photocurrent is likely due to higher device resistance and/or the fact that less Pd4-HtdR-His6 is present on the Pd microcontact.

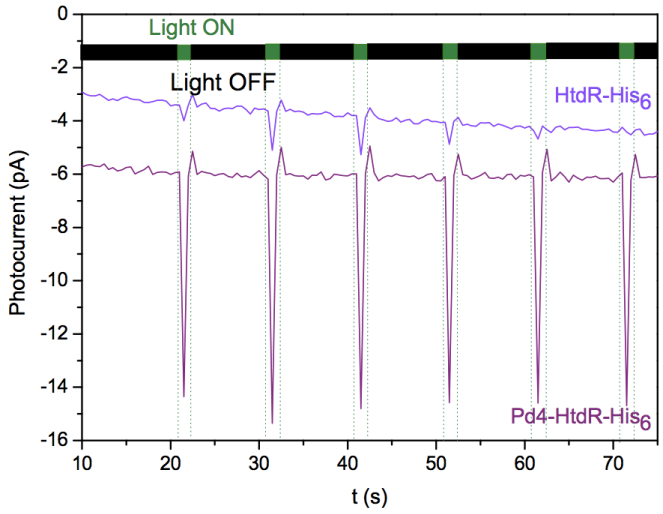

Figure 3.6 The photocurrent response of devices prepared with HtdR-His_6 (blue trace) and Pd4-HtdR-His_6 proteoliposomes (purple) was recorded while the device was kept under an applied voltage of 50 mV vs. Ag/AgCl and exposed to consecutive cycles of 1s illumination with a 523 nm LED and 9 s in the dark.

Chapter 4 Role of the Signal Sequence in Proteorhodopsin Biogenesis in *E. coli*

4.1 Introduction

Proteorhodopsins (PRs) are retinal-containing transmembrane proteins that function as light-activated proton pumps in marine organisms [78,150]. The blue-absorbing HOT75m1 proteorhodopsin, hereafter referred to as BPR, is one of the best characterized PRs in terms of structure [83] and function [84]. In contrast to most heterologous α-helical membrane proteins, and to the structurally related bR and G protein-coupled receptors (GPCRs) [151,152], proteorhodopsins can be expressed at appreciable levels in recombinant *E. coli* [78,153].

While optimizing the production of BPR in *E coli* to facilitate the construction of bioprotonic devices [101,154,155], we found surprisingly high levels of pre-protein accumulation in the inner membrane which led us to explore the role of the signal sequence in BPR biogenesis. Here, we show that BPR does not depend on the Sec pathway for inner membrane integration, and that although it greatly enhances yields, the signal sequence is not necessary to obtain a functional product. We further show that a full-length version of BPR obtained by precluding signal sequence cleavage through mutagenesis of the SPase I site is produced at higher level and exhibits all functional attributes of the wild type protein. Our results are discussed within the context of bacterial secretory pathways and the construction of bioprotonic devices.

4.2 Materials and Methods

4.2.1 DNA manipulations

A DNA cassette encoding a codon-optimized version of BPR (Hot75m1) followed by a C-terminal His_6 tag was commercially synthesized by Invitrogen. Both an *Nde*I and an *Nco*I site were built at the 5' end, leading to the addition of two additional amino acids (Met-Thr) at the N-

terminus of the native signal sequence. The corresponding gene was excised on an *Nde*I-*Xho*I fragment from plasmid pMAT-BPR-His_6 and inserted into the same sites of plasmid pBLN400 [76] to yield plasmid pBPR400.

A variant of BPR-His_6 lacking its signal sequence was constructed by amplifying the mature BPR gene from pMAT-BPR-His_6 with primers 5' – TATATACCATGGCAGCCGGTGGTGATCTG – 3' and 5' – CTCGAGTTAATGGTGATGGTGATG – 3', and ligating the *Nco*I-*Xho*I fragment into the same sites of pPBR400 to produce plasmid pPBR(Δss)400. Alanine 20 (using a numbering system that includes the two additional amino acids) was converted to a leucine by subjecting pBPR400 to site-directed mutagenesis with primers 5'-CAATTGCACTGCCGAGCTTTCTGGCAGCCGGTGGTGATCTGG - 3' and 5' – CCAGATCACCACCGGCTGCCAGAAAGCTCGGCAGTGCAATTG - 3' to produce pBPR(A20L)400.

4.2.2 Protein expression and characterization

E. coli KTD101 cells (Δ(*araD-araB*)*567* Δ*lacZ4787* (::*rrnB-3*) l-*rph1* Δ(*rhaD-rhaB*)*568 hsdR514* Δ*tig100*) [77] harboring plasmid pBPR400, pBPR(A20L)400, or pPBR(Δss)400 were grown to mid-exponential phase (A_{600} ~ 0.45) at 37°C in shake flasks containing 25 mL of LB medium supplemented with 50 μg/mL kanamycin. Protein synthesis was induced by addition of 0.2% L-arabinose and 10 μM all-*trans* retinal was added at the same time except where noted. For SecA inhibition experiments, 1.5 mM of freshly prepared sodium azide (NaN_3) was added to the cultures at the time of induction. Cells were harvested after 3 h of incubation at 37°C, disrupted with a French press operated at 10,000 psi, and the aggregated material was removed by centrifugation at 14,000*g* for 10 min at 4°C. Supernatants were harvested and centrifuged at

100,000*g* for 1 h at 4°C. Pellets corresponding to membrane fractions were resuspended in SDS sample buffer. The A_{600} of samples were adjusted to the same value and aliquots were subjected to SDS-PAGE and Western analysis using mouse anti-His antibody (Biolegend) at a 1:1500 dilution, and alkaline phosphatase-conjugated goat anti-mouse IgG (Sigma) at a 1:3000 dilution. Colorimetric detection was performed using 5-bromo-4-chloro-3-indolyl phosphate and nitrotetrazolium blue. Growth curves were constructed by measuring the A_{600} of triplicate cultures prepared from the same inoculum at 30 min intervals. For the experiments of Figure 4.2, cells were grown to mid-exponential phase at 37°C and protein expression was conducted at 25, 30 or 37°C.

4.2.3 Proton pumping activity

To measure *in vivo* proton pumping activity, KTD101 cells harboring pBPR(A20L)400 were grown in 250 mL of supplemented LB medium and induced with -arabinose with or without addition of *all-trans* retinal. After 3 hr of growth at 37°C, cells were sedimented by centrifugation at 4000*g* for 10 min, washed twice in an unbuffered solution consisting of 10 mM NaCl, 10 mM $MgSO_4 \cdot 7H_2O$ and 100 μM $CaCl_2$, and resuspended in 5 mL of the same solution. Samples were transferred to 50 mL centrifuge tubes cut to the 15 mL mark, and a glass pH electrode (Jenco 600P) was immersed in the solution. Samples were allowed to equilibrate in the dark until a constant pH value of ~6.6 was reached (about 30 min to 1 hr). Samples were then illuminated with a WheeLED Wavelength-Switchable LED Source at 505 nm positioned ~1.5 cm from the surface of the solution. Control experiments were conducted as above with KTD101 cells harboring pBPR400.

4.2.4 Spectral Characteristics

For protein purification, cultures (500 mL) were grown and induced as above, and membrane fractions were harvested by ultracentrifugation as described. Pellets were resuspended in Buffer I (50 mM 2-(*N*-morpholino) ethanesulfonic acid [MES], pH 6.5, 300 mM NaCl, 5 mM imidazole, 1.0% n-dodecyl β-D-maltoside [DDM]) and incubated overnight at 4°C with gentle shaking. Solubilized BPR(A20L) was purified by two rounds of Ni-NTA chromatography [154] and stored at -80°C until further use. For the experiments of Figure 4.7C, purified BPR(A20L) was diluted in Buffer II (50 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, 0.1% DDM) at the indicated pH and samples were concentrated using 10-kDa microconcentrators. Protein concentrations were determined by BCA assay using BSA as a standard. UV-visible spectra were collected at ~3.5 μM protein concentrations.

4.3 Results and Discussion

4.3.1 Significant amounts of precursor BPR accumulates in the inner membrane of *E. coli* under optimized conditions for polytopic membrane protein expression

In an effort to maximize the yields of the blue-absorbing proteorhodopsin BPR HOT75m1, a protein that has been successfully expressed and purified in *E. coli*,[69,84,153] we first employed strategies that have proven useful in improving the accumulation of polytopic (i.e., multi-TMS) inner membrane proteins in *E. coli* [76,77,156]. Briefly, a synthetic gene encoding a codon-optimized version of BPR HOT75m1 together with its predicted signal sequence and a C-terminal hexahistidine (His) tag was placed under transcriptional control of a downregulated P_{BAD} promoter advantageous for IMP expression, and the resulting plasmid was introduced into trigger factor deficient (Δ*tig*) cells which improve the yields of polytopic IMPs in the inner membrane [76,77,156]. Cultures grown in the presence of all-*trans* retinal exhibited the pink-orange

color that is characteristic of functional BPR expression (**Figure 4.1 A**). However, in contrast to studies reporting full processing of the BPR signal sequence [153,157], we found that the signal peptide was only removed in about half of the material associated with, or purified from, the plasma membrane (**Figure 4.1B-C**).

Incomplete signal sequence removal was not completely unexpected since the presence of a signal sequence-containing form of BPR (that we will refer to thereafter as pre-BPR) was identified in purified protein samples with sensitive techniques such as MALDI-TOF [158] and amino acid-specific isotope labeling [159]. What was surprising in our experiments was that such a large amount of membrane-integrated pre-BPR remained unprocessed.

4.3.2 Neither the presence, nor the processing of the signal sequence is required for BPR biogenesis which occurs in Sec-independent fashion

To confirm that inefficient signal sequence processing was related to enhanced BPR synthesis, we conducted expression experiments at 25, 30 and 37^{o}C and immunoblotted aliquots corresponding to identical amounts of cells with anti-His antibodies. **Figure 4.2 A-B** show that the amount of mature —and therefore membrane-integrated BPR— increased with a reduction in the expression temperature, as would be expected from a better match between the lower flux of pre-BPR to the plasma membrane and the availability of cellular Signal Peptidase I (SPaseI). However, even when expression was conducted at 25^{o}C, the fraction of unprocessed protein remained high with pre-BPR accounting for nearly 40% of the immunoreactive material produced by the cell (**Figure 4.2 B**).

Considering that most proteins residing in the bacterial inner membrane are inserted in the lipid bilayer through the signal recognition particle (SRP) dependent pathway and do not

contain a Sec-dependent signal sequence [160], we further explored the need for this feature by constructing BPR(Δss)-His, a BPR variant lacking residues Met-1 to Ala-20 which span the predicted signal sequence. When cells expressing the truncated protein were grown in the presence of retinal, a faint pink-orange color was observed (**Figure 4.3 A**), suggesting low-level expression of functional BPR. Indeed, Western analysis of membrane fractions with anti-His antibodies revealed the presence of a faint immunoreactive band at the migration position of mature BPR (**Figure 4.3 C**). Thus, although the BPR signal sequence of BPR is dispensable it significantly improves protein production.

We next asked if processing of the signal sequence was important to obtain active BPR. Informed by the facts that SPase I recognizes the Ser-Phe-Ala motif present at the C-terminus of its signal sequence [158,161](**Figure 4.4**), and that conversion of the highly conserved P1 alanine to leucine in the SPase I site of the alkaline phosphatase (PhoA) signal sequence completely inhibits its processing [162], we changed Ala-20 in position P1 of the BPR signal sequence to Leu to produce BPR(A20L). Cells expressing the resulting variant exhibited the pink–orange color of high producers of bioactive BPR (**Figure 4.3 A**) although none of the material present in membrane fractions was processed into mature BPR (**Figure 4.3 B-C**, lane 5). Overexpression of BPR(A20L) had a more deleterious effect on cell growth relative to that of BPR or BPR(Δss), possibly due to the toxicity of membrane protein overexpression or to the fact that the uncleavable signal sequence of BPR(A20L) inhibits SPase I and causes the accumulation of toxic preproteins (**Figure 4.5**). However, cells overproducing BPR(A20L) accumulated about 50% more immunoreactive material in their inner membranes compared to cells expressing the wild type protein. As a result, we achieved BPR(A20L) yields of ~ 2-3 mg of protein per L of culture after Ni-NTA purification from detergent solubilized membranes (**Figure 4.6**).

With the presence and processing of the BPR signal sequence seemingly unnecessary for the production of functional protein, we asked if the biogenesis of BPR in *E. coli* was dependent on the Sec pathway. To this end, we repeated expression experiments in the presence of sodium azide (NaN_3), a compound that inhibits the ATPase activity of SecA and blocks Sec-dependent protein translocation [163]. **Figure 4.3 B-C** shows that azide addition had no significant impact on the accumulation of wild type BPR, BPR(Δss) or BPR(A20L) in the inner membrane, indicating that trafficking to and insertion of these proteins within the inner membrane is occurring through a Sec-independent process.

4.3.3 BPR(A20L) is fully functional as a pH-dependent, light-activated proton transporter

While a change in cell color following addition of all-*trans* retinal to cultures of *E. coli* producing BPR(A20L) is consistent with the acquisition of a correct seven transmembrane α-helical architecture [70,78], we sought to confirm that the protein's proton pumping activity was unaffected by the presence of an uncleaved signal sequence. To do so, we expressed BPR(A20L) in the presence or absence of retinal, transferred the cells to unbuffered saline, and measured the change in solution pH upon illumination with a 505 nm light emitting diode (LED). **Figure 4.7 A** shows that exposure of retinal-treated cells to blue light caused immediate acidification, and that the pH started to rise again as soon as the light was turned off (black circles). The kinetics of acidification were well captured by an exponential decay model with rate constants of 14.9×10^{-4} $pH.min^{-1}A_{600}^{-1}$ for BPR(A20L) and 6.2×10^{-4} $pH.min^{-1}A_{600}^{-1}$ for the wild type, with differences well explained by the lower expression level of wild type BPR species relative to BPR(A20L) (**Figure 4.3 C**). Negative control, retinal-free culture of BPR(A20L) which produce a nonfunctional protein experienced a linear decrease in pH that was unaffected by the presence or absence of light (**Figure 4.7 A**, red circles). Of note, it was possible to achieve multiple cycles of

light-induced acidification after the cell suspension had returned to a more neutral value (**Figure 4.7 B**), a critical property for the construction of rhodopsin-based protonic devices [154].

Retinal-bound BPRs have characteristic pH-dependent absorption spectra in which maximum absorption wavelengths can vary by as much as 50 nm [80]. To further confirm that the purified signal-sequence-containing BPR(A20L) variant was functionally similar to the wild type, we recorded its UV-visible spectrum at four different pHs. Maximum absorption was at ~524 nm at pH 7.2 and blue shifted linearly to ~ 500 nm at pH 9.0 (**Figure 4.7 C** and **Figure 4.8**). These values are comparable to those of Hot75m1 except for a 5-7nm shift to longer wavelengths [84,164], possibly caused by variations in sample preparation and protein concentrations. In short, the presence of an intact signal sequence at the N-terminus of BPR does not affect light-induced proton transport or significantly alter pH-dependent spectral characteristics.

4.4 Discussion

The discovery of proton-pumping proteorhodopsins in planktonic marine bacteria [78] was a first in many respects: it established the power of metagenomics in identifying genes or functions that one was not particularly "looking for"; it ignited research that established the ubiquity of PRs in bacterioplankton; and it revealed that α-helical membrane proteins functionally or structurally related to difficult to express targets (e.g., bacteriorhodopsin and GPCRs), can be produced at reasonable yields in recombinant *E. coli* [69,78,153]. While significant effort was subsequently dedicated to the study of the biochemistry, physiology and genetic diversity of marine PRs [165], little attention has been paid to the details of PR biogenesis in *E. coli*.

In their original studies, Béja and coworkers predicted the existence of typical Sec-

dependent signal sequences at the N-terminus of green (and blue) proteorhodopsins amplified from environmental marine samples [69,78]. While the pathway responsible for plasma membrane integration was not studied, it was reported that these signal sequences were efficiently cleaved to yield mature products that are functional [78,157] and adopt the 7-TMS architecture common to bR and GPCRs [78,83]. Taken together with a recent report demonstrating that SecA supports the co-translational insertion of the single TMS membrane protein RodZ [166], the presence of a signal sequence raised the possibility that SecA could be involved in BPR biogenesis. However, our combined observations that a signal sequence-less version of the proteorhodopsin localizes to the plasma membrane where it folds into a functional protein (albeit with severely reduced yields), and that treatment of cultures with the SecA inhibitor sodium azide does not alter the distribution of BPR precursor and mature species (**Figure 4.3**) strongly argue against this possibility.

The majority of autologous inner membrane proteins containing multiple membrane-spanning segments rely on the cotranslational SRP pathway to integrate within the plasma membrane of *E. coli* [67]. In current models, SRP samples the hydrophobicity of nascent chains emerging from the ribosome exit tunnel and preferentially binds to regions exhibiting the high hydrophobicity that is typical of α-helical TMS [167]. Indeed, a recent analysis of the nascent *E. coli* proteome has revealed that over half of the sites bound by SRP contain a stretch of at least four consecutive phenylalanines, isoleucines, leucines or valines [68]. Such features only occur about 8% of the time in Sec-dependent signal sequence, yet a signature Leu-Leu-Leu-Ile-Leu motif is present in the signal sequences of both BPR and GPR (**Figure 4.4**). It is therefore possible that SRP recognizes the BPR signal sequence before it is engaged by SecA, and that it directs the ribosome nascent chain complex to its FtsY receptor for subsequent cotranslational membrane integration of the opsin through the lateral gate of the SecYEG pore. Further work

will be needed to explore this possibility.

Our finding that conversion of the Ser-Phe-Ala motif at the C-terminus of the BPR signal sequence to Ser-Phe-Leu completely abolishes leader peptide excision (**Figure 4.3**) firmly establishes SPase I as the peptidase responsible for BPR signal sequence processing. This is an unusual aspect of BPR biogenesis in *E. coli* since SPase I normally cleaves the signal sequences of periplasmic and outer membrane proteins that are exported post-translationally via the Sec-dependent pathway. It should be noted that the presence of a cleavable signal sequence in an inner membrane protein is not unheard of: the bacteriophage M13 procoat protein sports such a feature. However, there are significant differences between this short (79-residues long) single TMS membrane protein which directly integrates into the membrane through a "YidC-only" pathway [168] and BPR a 27 kDa polytopic membrane protein. The presence of a signal sequence has a clear beneficial effect on the yields of membrane-integrated BPR (**Figure 4.3**) and it may have evolved for this very reason.

Remarkably, we found that the "unprocessable" A20L pre-BPR variant was not only fully functional for light-activated proton transport (**Figure 4.7**), but that it was also produced at higher levels than the wild type protein (**Figures 4.3**). At present, we do not know if the accompanying increase in cellular toxicity (**Figure 4.5**) is the reason why evolution has favored a cleavable version in proteobacteria or if higher levels of expression affect *E. coli* growth rates.

4.6 Conclusions

This chapter focused in optimizing the production of BPR in *E coli* to facilitate the construction of bioprotonic devices. First, we demonstrated that cleavage of the signal sequence is inefficient resulting in the accumulation of significant amounts of pre-BPR. We demonstrated that BPR does not rely on the Sec pathway for inner membrane integration, and that although it

greatly enhances yields, its signal sequence is not necessary to obtain a functional product. We further showed that an unprocessable version of pre-BPR obtained by introducing the A20L mutation of the signal peptidase I site exhibits all functional attributes of the wild type protein and that is has the advantage of being produced at higher levels. BPR(A20L) yields sizable amounts of a homogeneous gene product that exhibit light cycling and pH-dependent characteristics similar to the wild type protein and is therefore very well suited for the construction of bioprotonic devices.

4.6 Figures

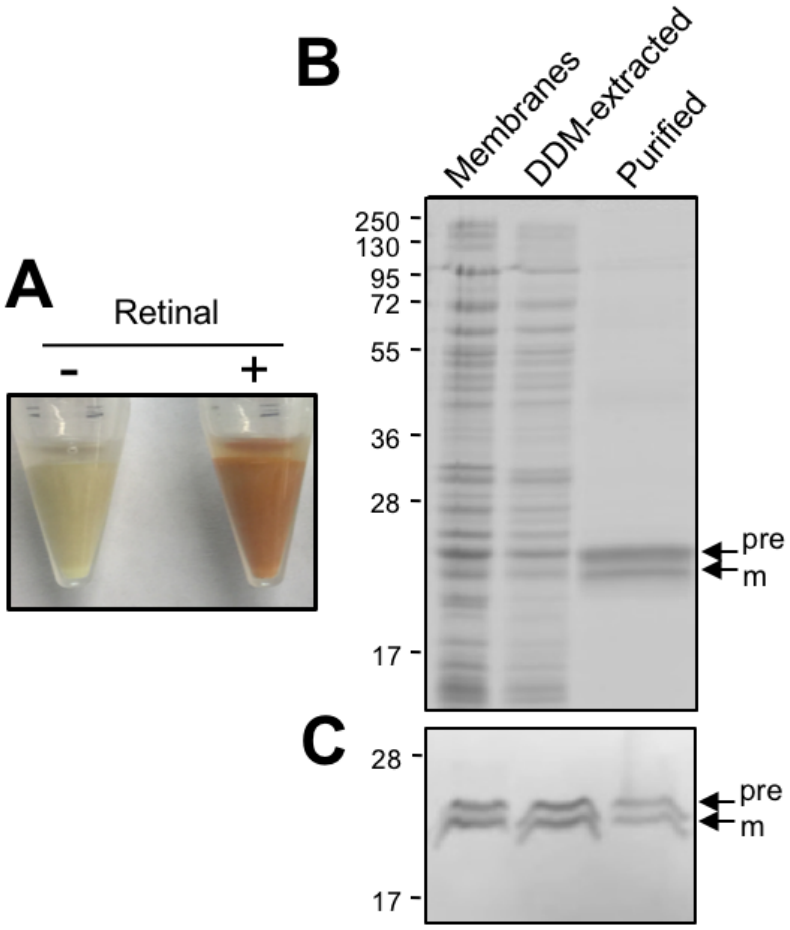

Figure 4.1 Expression and purification of wild type BPR. **(A)** KTD101 cells expressing BPR exhibit a pink orange color characteristic of a functional protein when retinal is added at the time of induction. **(B)** Bands corresponding to precursor (pre) and mature (m) BPR are present in membrane fractions and following Ni-NTA purification. **(C)** Western analysis with anti-His antiserum confirms that both products contain an intact C-terminal His tag, and thus that signal sequence processing is incomplete.

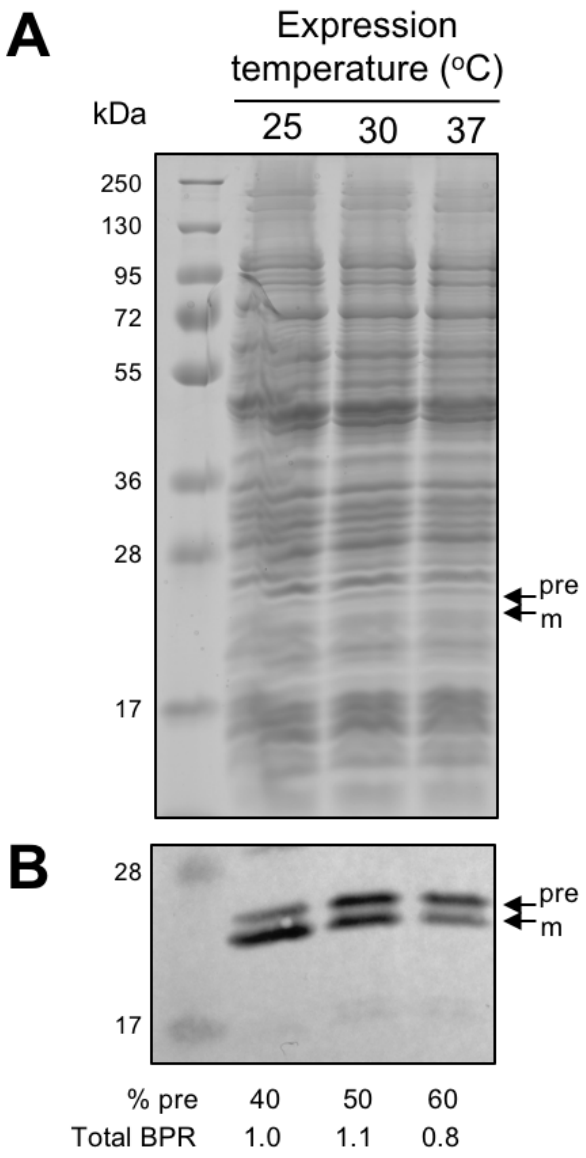

Figure 4.2 Lower cultivation temperatures improve BPR signal sequence processing. SDS-PAGE fractionation **(A)** and corresponding immunoblot analysis **(B)** of KTD101 cultures expressing BPR at the indicated temperatures. Samples correspond to identical amounts of cells collected 3h post-induction. Top and bottom arrows identify the migration positions of unprocessed (pre) and mature (m) BPR, respectively. Numbers below the blot indicate the percent of unprocessed BPR (pre) and the total (precursor and mature) BPR species in each sample, as determined by videodensitometric analysis. The percent of pre-BPR in cells grown at 37°C was found to be 58±3% in 5 independent experiments.

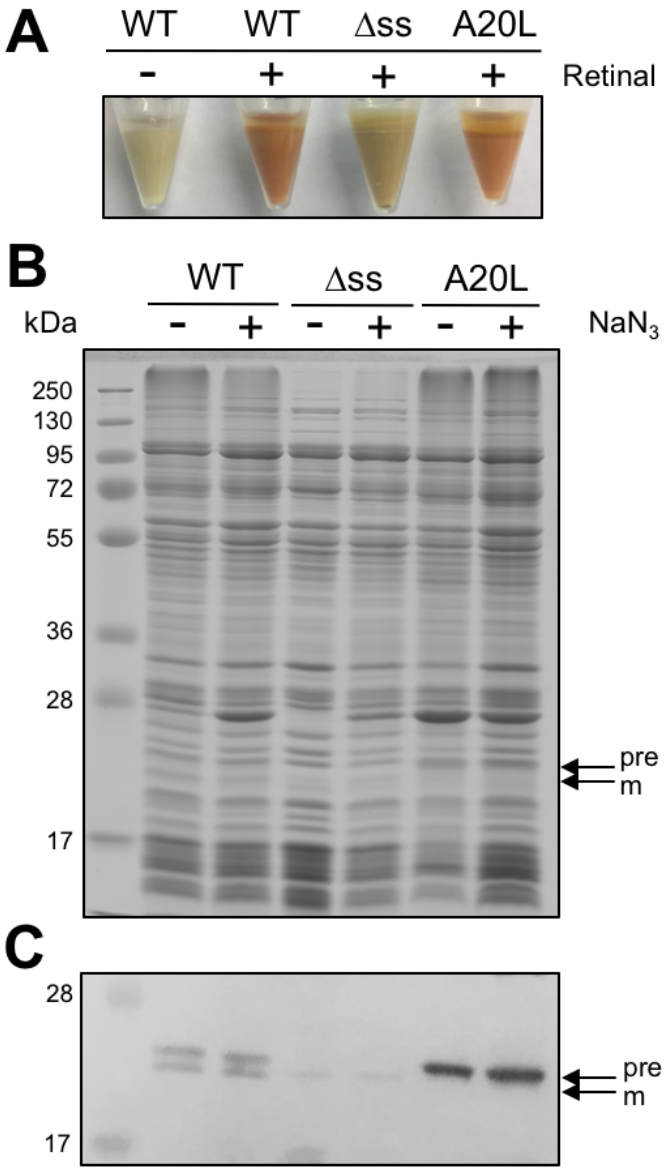

Figure 4.3 Role of the BPR signal sequence in membrane targeting and functional protein production. **(A)** Appearance of KTD101 cells expressing wild type BPR (WT), or its Δss or A20L variants after 3h of cultivation in the presence of retinal (+). A retinal-free control (-) is included for the wild type protein. SDS-PAGE fractionation **(B)** and corresponding immunoblot **(C)** of membrane fractions extracted from *E. coli* cells expressing BPR (WT), or its Δss or A20L variants following treatment (+) or not (-) with NaN_3. Top and bottom arrows identify the migration positions of unprocessed (pre) and mature (m) BPR, respectively.

Figure 4.4 Amino acid composition of the BPR (Hot75) and GPR (Ebac3108) signal sequences. Two extra amino acids were added to the N-terminus of the native BPR signal sequence as a result of construction constraints. Numbering starts with Met-1. SPase I recognition sequences are in blue and cleavage sites identified with arrows. The Ala-20 converted to Leu to prevent BPR signal sequence cleavage is in white. The motif preferentially recognized by SRP is in orange.

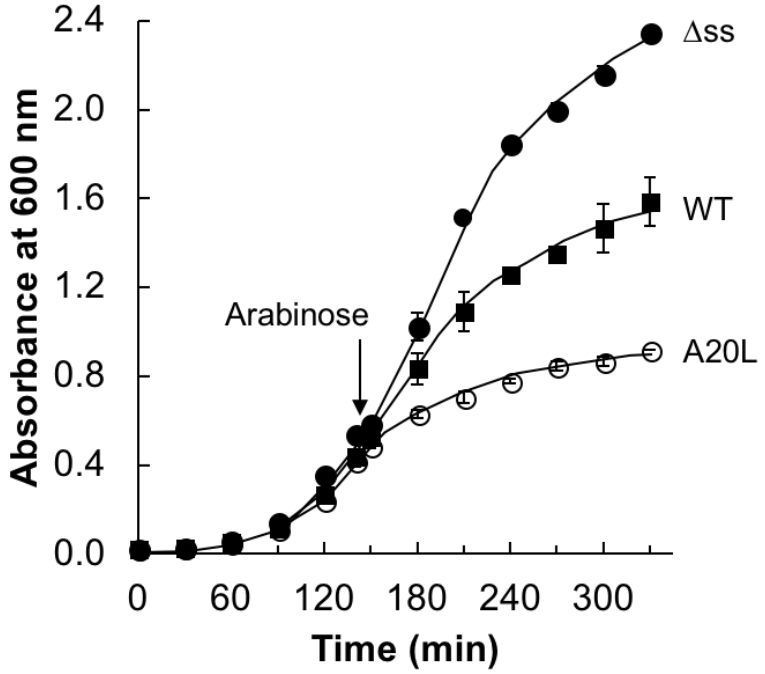

Figure 4.5 KTD101 cells harboring plasmids encoding wild type BPR (WT), or its Δss and A20L variants were grown at 37^{o}C in LB medium to mid-exponential phase, induced with arabinose and proteins were allowed to accumulate for the indicated times. Error bars correspond to triplicate experiments.

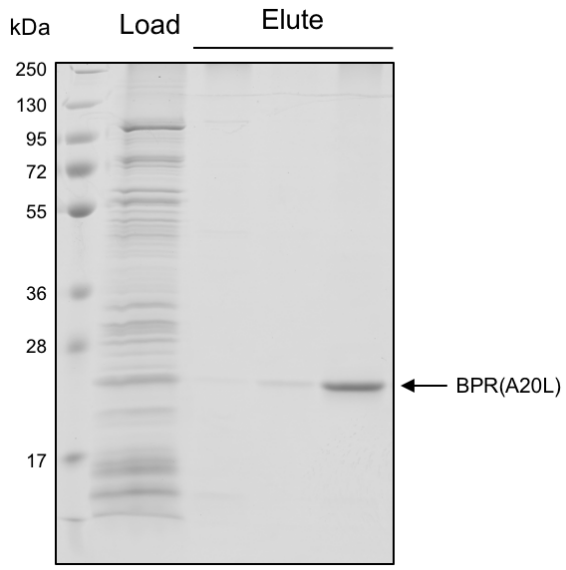

Figure 4.6 Purification of DDM-solubilized BPR(A20L) by Ni-NTA chromatography. Load and successive elution fractions are shown.

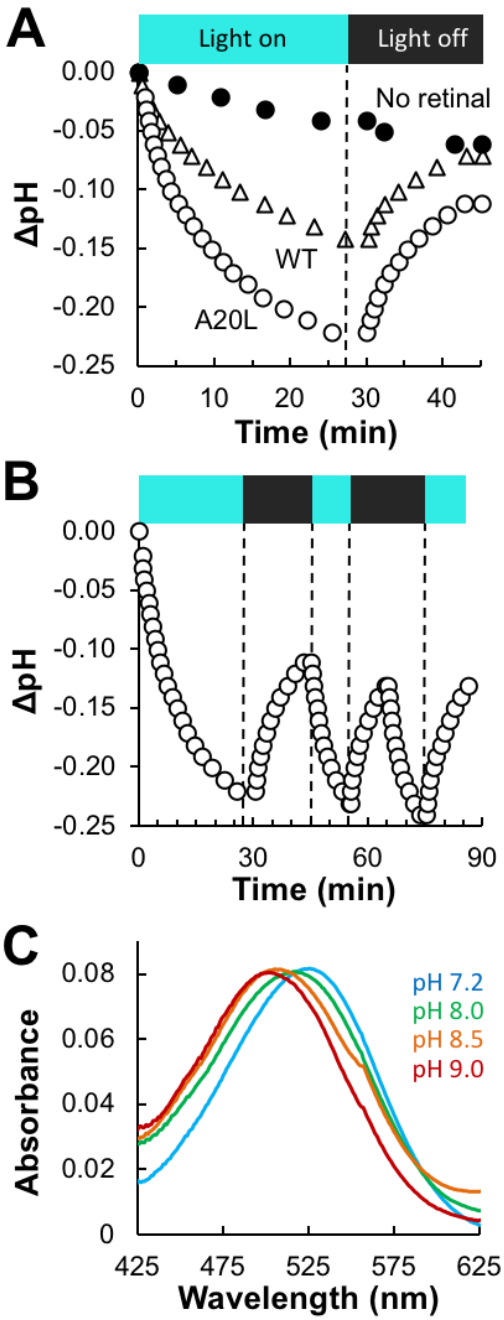

Figure 4.7 BPR(A20L) supports light-induced outwards proton transport and exhibits pH-dependent spectral characteristics. **(A)** Resuspended *E. coli* cells expressing wild type BPR (open triangles) or its A20L variant (open circles) were illuminated with blue light for 30 min and the light was turned off. The evolution of solution pH was monitored over time. A retinal-free culture of BPR(A20L) provides the negative control (filled circles). **(B)** Repeated cycles of illumination and darkness demonstrate that the observed changes in pH are associated with the photocycle of BPR(A20L). **(C)** The absorption spectra of purified BPR(A20L) was recorded at the indicated pHs.

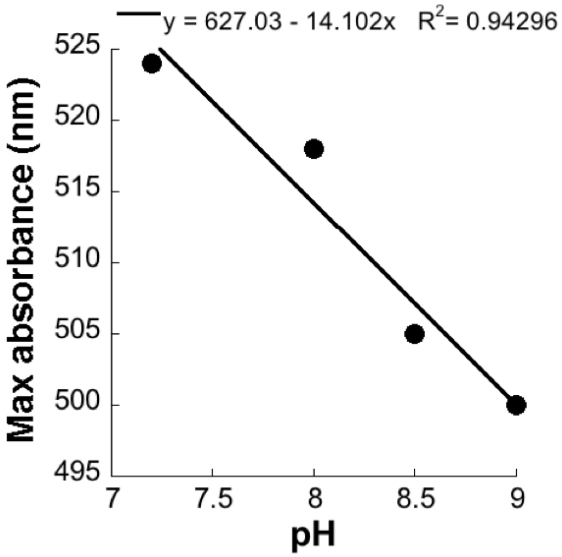

Figure 4.8 Relationship between solution pH and the maximum absorption intensity

Chapter 5 A Two-Channel Bioprotonic Photodetector

5.1 Introduction

In Chapter 3, we harnessed the proton pumping mechanism of a deltarhodopsin to build a new type of bioprotonic device that can be manipulated with green light.[154] A unique property of microbial rhodopsins is that the bound retinal chromophore that controls proton transport by reversibly shifting from a *trans* to a *cis* conformation upon the absorption of a photon of light[169], experiences large variations in absorption maximum depending on its microenvironment.[170] For example, the well-studied bacteriorhodopsin (bR) has an absorption maximum (λ^{abs}_{max}) of 568 nm, while that of HtdR is ~545 nm. In addition, in the case of proteorhodopsins (PR), a family of rhodopsins from marine bacteria that can be produced at high levels in the workhorse organism *E. coli,*[78,153] spectral and proton-pumping characteristics are pH-dependent,[79,80,158] and can be tuned by the introduction of a single amino acid mutation.[79] With a broad division into blue absorbing proteorhodopsin (BPR, λ^{abs}_{max} ~ 495 nm)[84,164] and green absorbing proteorhodopsin (GPR, λ^{abs}_{max} ~ 525nm),[78] PRs offer a unique opportunity to increase the functionality of light-activity bioprotonic devices by allowing for the possibility of tuning their response with specific wavelengths of light. This chapter focuses on the construction and characterization of such a multiwavelength photodetector.

5.2 Materials and Methods

5.2.1 DNA Manipulations and Protein Expression and Purification

Plasmid pPd4-HtdR encodes a Pd4-HtdR fusion protein with a C-terminal hexahistidine tag under transcriptional control of a modified arabinose-inducible P_{BAD} promoter.[76] Plasmid pBPR(A20L) encodes a codon-optimized version of BPR with an uncleavable pro sequence that

is homogeneous, fully functional for proton transport, and accumulates at high yields.[171] A DNA cassette encoding codon-optimized BPR(A20L) modified with a N-terminal Pd4 solid binding peptide (TSNAVHPTLRHL)[8] and a C-terminal His_6 tag was synthesized commercially. The gene was excised on a *Nde*I-*Xho*I fragment and inserted into the same sites of plasmid pBLN400 to yield pPd4-BPR(A20L). This plasmid was introduced into *E. coli* KTD101 cells[77] which improve the yields of polytopic membrane proteins. Cultures of KTD101 harboring plasmid pPd4-BPR(A20L) or pPd4-HtdR were grown to mid-exponential phase (A_{600} ~ 0.45) at 37°C in shake flasks containing 500 mL of LB media supplemented with 50 μg/mL kanamycin. Protein synthesis was induced with 0.2% L-arabinose, the medium was supplemented with 10 μM all-*trans* retinal, and proteins were allowed to accumulate for 3 hr. Cells were lysed with a French Press, and n-dodecyl β-D-maltoside (DDM) solubilized membrane proteins were prepared and purified by Ni-NTA chromatography as previously described.[154] Protein concentrations were determined by BCA assay using BSA as a standard. UV-visible spectra were collected at ~3.5 μM protein concentrations.

5.2.2 Partial Proteolysis Experiments

Aliquots (10 mL) of KTD101 cells expressing Pd4-BPR(A20L) or Pd4-HtdR were spheroplasted as described.[154] Spheroplasts were resuspended in 3 mL of 100mM Tris-HCl, pH 8.0, 0.3M sucrose, 10mM $MgCl_2$ and treated with 0.1mg/ml of a freshly prepared solution of trypsin (from bovine pancreas). After 1 hr incubation at room temperature, samples were supplied with 10 mM PMSF and held on ice for 10 min to stop the reaction. Spheroplasts were recovered by centrifugation and resuspended in sample buffer for SDS-PAGE and Western analysis. Gels were incubated in 25 mM Tris-HCl, pH 8.3, 20% methanol, 200 mM glycine and 3mM SDS for 1 h and transferred to a nitrocellulose membrane for 14 hr at 12 V and 4^{o} C.

Nitrocellulose membranes were probed with mouse anti-His (Biolegend) at 1:1500 dilution, incubated with goat antimouse IgG Alkaline Phosphatase (Sigma) at 1:3000 dilution, and detected by colorimetric reaction with 5-bromo-4-chloro-3-indolyl phosphate and nitrotetrazolium blue.

5.2.3 Proton Pumping Activity

Cultures (250mL) expressing Pd4-BPR(A20L) or BPR(A20L) as a control were grown and induced as above in the presence or absence of *all-trans* retinal. Cells were harvested by centrifugation at 4000*g* for 10 min, washed twice with 10 mM NaCl, 10 mM $MgSO_4 \cdot 7H_2O$, 100 μM $CaCl_2$ and resuspended in 5 mL of the same solution. Samples were allowed to reach a stable pH in the dark for 30 min to 1hr and illuminated at 505 nm with a WheeLED Wavelength-Switchable LED positioned ~1.5 cm from the surface. The evolution of pH was recorded with a glass electrode.

5.2.4 Proteoliposomes preparation

Proteoliposomes were prepared as previously described. Briefly, thin films from multilamellar DOPC vesicles were hydrated in 10 mM HEPES, pH 7.4 or 8.5, 150 mM NaCl for 1h with shaking to a lipid concentration of 4 mg/mL. Large unilamellar vesicles (LUVs) were generated by extrusion. Purified Pd4-Htdr-His_6 and Pd4-BPR-His_6 were diluted in 4mL of 50 mM Tris-HCl, pH 7.2, 300 mM NaCl, 0.1% DDM, and spun in a 10-kDa microconcentrator for imidazole removal. Protein concentrations were determined before proteoliposomes were prepared by the three-step detergent method,[136,137] as previously described.

5.2.5 Device Fabrication and Current Measurements

The bioprotonic device and its fabrication have been described in detail previously.[107,154]

Aliquots (150 μL) of control DOPC liposomes, Pd4-HtdR and Pd4-BPR(A20L) proteoliposomes synthesized as above were deposited in the microfluidic channel that confines the flow of liquid to the top on Pd contacts. Prior to aliquots deposition, devices were plasma cleaned to remove contaminations and also help vesicles fuse to the Pd surface.[138] The device was agitated during vesicle fusion, which was allowed to proceed overnight at 4^{o}C. Prior to electrical measurements, the device was rinsed with fresh HEPES buffer (pH=7.4) to remove unincorporated vesicles. Electrical measurements were performed with a grounded semiconductor parameter analyzer (Agilent 4155C). Electrical contact between analyzer and device were made with a Signatone S probe station with a custom built environmental chamber with microprobes. Each Pd contact device was connecting to the analyzer by means of an electrical probe pad out of the microfluidic channel and the microprobes. A PDMS well was sealed on top of the microfluidic channel provides space to insert the counter and reference electrodes (Ag/AgCl) and complete the electrical circuit. To determine wavelength dependent electrical characteristics, a tunable laser source was employed. Wavelength tunability was achieved with an acousto-optic tunable filter and a supercontinuum source (NKT Photonics SuperK EXTREME). The laser output was coupled to a single mode optical fiber (Newport F-SV) which was collimated and aligned. The fiber collimator was mounted above the device and electrical measurements collected as the wavelength was scanned. Data were normalized to the wavelength-dependent output intensity.

5.3 Results and Discussion

5.3.1 Construction of a Palladium-Binding Variant of Blue Proteorhodopsin

In an effort to broaden the usefulness of our HtdR-based photodetector,[154] we selected the Hot75m1 blue proteorhodospin (hereafter called BPR) to construct a bioprotonic device that would respond to multiple wavelengths of light. BPR, whose proton transport function is

activated by blue light and modulated by pH,[84] is synthesized with a 20 residues-long N-terminal extension that facilitates protein trafficking and insertion in the bacterial inner membrane.[171] This "pro" sequence, which has also been referred to as a signal sequence,[78] is cleaved by the membrane-embedded Signal Peptidase I to yield a mature opsin that spans the lipid bilayer 7 times (**Figure 5.1A**). [83,171] However, high-level expression of BPR in *E. coli* leads to inefficient processing of the pro sequence and causes both full-length and cleaved products to accumulate in the inner membrane.[171] Remarkably, when cleavage of the extension is prevented by introducing an Ala-20 to Leu (A20L) mutation in the Signal Peptidase I recognition sequence, a homogeneous protein with proton-pumping characteristics virtually identical to those of wild type BPR is obtained at 50% higher yields.[171] This BPR(A20L) variant was used in subsequent experiments.

To endow BPR(A20L) with the ability to bind to Pd/PdHx, we genetically fused the 12 residues-long Pd4 palladium-binding peptide[8] to the N-terminus of BPR(A20L) through a flexible Gly-Gly-Gly-Ser (GGGS) linker (**Figure 5.1A**). When *E. coli* cells expressing the resulting Pd4-BPR(A20L) fusion were grown in the presence of all-*trans* retinal, they exhibited a pink-orange color comparable to that of BPR(A20L) control cultures (**Figure 5.1B**). Because the Schiff-based mediated attachment of retinal to the 7th transmembrane segment of the opsin requires a properly folded protein,[70,78] color acquisition indicates that the Pd4 extension does not preclude BPR(A20L) from reaching the inner membrane and from acquiring a proper conformation once embedded in the lipid bilayer.

We purified Pd4-BPR(A20L) from detergent-solubilized membranes through two cycles of Ni-NTA chromatography, taking advantage of the presence of a hexahistidine affinity tag at the C-terminus of the protein (**Figure 5.1A**). Yields (~ 2-3 mg/L of culture) and purity (>85%) were

comparable to those obtained with BPR(A20L) (**Figure 5.1C**). More importantly, the two proteins had virtually identical UV-visible spectra suggesting that the retinal chromophore experiences similar environments and that Pd4-BPR(A20L) is properly folded (**Figure 5.1D**). Thus, fusion of the Pd4 sequence to the N-terminus of BPR(A20L) does not have significantly affect protein expression, trafficking, insertion and folding into the inner membrane, binding of retinal, purification efficiency, or recovery yields.

5.3.2 The Pd4 Extension Faces the Periplasm

For Pd4-BPR(A20L) to effectively function within the context of our bioprotonic device, the Pd4 extension must bring leaving protons close to the palladium surface when it binds to Pd/PdHx contacts. Thus, Pd4 should lie on the periplasmic side of the inner membrane, as depicted in **Figures 5.1A** and **5.2A**. However, the intervening pro sequence contains a hydrophobic region that has potential for forming an 8th transmembrane segment.[159] This topology would place Pd4 in an undesirable location: the cytoplasmic and proton donating side of the inner membrane (**Figure 5.2B**).

To determine if this was the case, we stripped *E. coli* cells expressing Pd4-BPR(A20L) of their outer membrane, producing spherical vesicles consisting of cytoplasmic contents bounded by the inner membrane. The resulting spheroplasts were incubated with trypsin, a protease that cleaves on the carboxyl side of arginine (R) and lysine (K) residues.[172] Because the transmembrane segments and cytoplasmic loops of membrane proteins are shielded from proteolytic attack by the lipid bilayer, only loops and termini facing the periplasm (and protruding on the surface of spheroplasts) are susceptible to digestion. As a result, trypsin should cleave within the Pd4 domain (Arg-11) and at the beginning of the pro sequence (Lys-21) if Pd4-BPR(A20L) adopts the 7 transmembrane segments fold that is common to all rhodopsins (**Figure**

5.2A, left panel; red arrows in **Figure 5.1A**). However, no cleavage should occur if Pd4-BPR(A20L) contains an 8th transmembrane segment because this structural feature would place Arg-11 and Lys-21 in the cytoplasm (**Figure 5.2B**).

The gel and immunoblot of **Figure 5.2C** unambiguously show that incubation of spheroplasts with trypsin leads to partial digestion of Pd4-BPR(A20L) and to the appearance of a truncated product that cross-reacts with anti-His_6 antibodies (which implies N-terminally processing). Using calibrated videodensitometric analysis of migration positions, we found that the degradation product was 1.2-kDa lighter than the full-length protein. This is in excellent agreement with a digestion event at Arg-11 which should excise a 1.21-kDa peptide from the N-terminus of Pd4-BPR(A20L). The reasons for the absence of a second degradation band corresponding to trypsin attack at Lys-21 is unclear but it is possible that local structure formation in the pro domain may impede protease access.

To validate the above results, we prepared Pd4-HtdR spheroplasts and subjected them to trypsin digestion. Because Pd4-HtdR adopts a 7 transmembrane segments structure and contains a single trypsin-accessible site at Arg-11 of its periplasmic Pd4 domain (**Figure 3.1 A**) a 1.21 kDa fragment should also be excised from the protein. **Figure 5.2 D** shows that this prediction is indeed correct. We conclude that Pd4-BPR(A20L) exists as a 7 transmembrane segments protein in the inner membrane of *E. coli* with a Pd4 domain that projects into the periplasm. As a result, residues His-6 and His-11 which have been identified as important for Pd binding[173] should be able to make contact with the Pd/PdHx electrode.

5.3.3 Pd4-BPR(A20L) Supports Proton Pumping

Previously, we have shown that the BPR(A20L) variant is functional as an outwards proton pump despite the presence of an uncleaved pro sequence.[171] To determine if addition of the Pd4

domain would interfere with proton transport, we resuspended *E. coli* cells expressing Pd4-BPR(A20L) in unbuffered medium and measured the change in solution pH upon illumination with a 505 nm LED. Cultures lacking retinal were used as a negative control while cells producing BPR(A20L) served as a positive control. **Figure 5.3 A** shows that retinal-supplemented cultures of Pd4-BPR(A20L) or BPR(A20L) rapidly acidified the medium upon blue light illumination while cultures of Pd4-BPR(A20L) grown in retinal-free medium experienced a slow linear drift in pH. With a recovery of the solution pH observed in the dark phase (**Figure 5.3 A**), it was possible to achieve successive cycles of acidification by turning the light on and off (**Figure 5.3 B**). These results indicate that Pd4-BPR(A20L) is fully functional as an outwards proton pump.

5.3.4 Wavelength Tuning of Bioprotonic Device Response

Unlike "typical" bacteriorhodopsins, the UV-visible absorption spectrum of BPR is affected by the environmental pH and its absorption maximum can shift by as much as 50 nm when recorded under acidic or alkaline conditions.[80,164] In Chapter 4, we have shown that the A20L mutation – and therefore the presence of an undigested pro sequence – has little effect on the pH-dependency of BPR spectral characteristics.[171] The data of **Figure 5.1 D** demonstrates that addition of a Pd4 extension does not alter λ^{abs}_{max} at pH 7.4 (524 nm) whereas **Figure 5.4 A** further shows that Pd4-BPR(A20L) remains pH-responsive with a maximum absorption blue-shifting to 508 nm at pH 8.5. By contrast, under the same conditions of pH, the λ^{abs}_{max} of Pd4-HtdR remains unchanged at 545 nm (**Figure 5.4 A**). Thus, a pH of 8.5 can be used to introduce a nearly 40 nm gap in the absorption maxima of the two rhodopsins, which is well suited for the development of the multi-wavelength bioprotonic device depicted in **Figure 5.4 B**. [107,154] The bioprotonic devices are essentially the same as previously described (**Figure 5.4 B)**. [107,154]

In this work, the lipid bilayer insulating the Pd contact from the interstitial layer (IL) contains either of Pd4-BPR(A20L) or Pd4-HtdR. [107,154] To create Pd4-BPR(A20L)-based devices, proteoliposomes were prepared and allowed to fuse on the Pd/PdHx microcontacts in pH 8.5 buffer. Thus, the liquid isolation layer (IL) that separates electrode from lipid bilayer also experiences a pH of 8.5, as do the periplasmic domains of proteorhodopsin molecules that are brought in proximity to Pd/PdHx by their Pd4 palladium-binding domains (**Figure 5.4B**). This configuration should confer a 508 nm absorption maximum to Pd4-BPR(A20L) molecules embedded within the supported lipid bilayer.

To demonstrate bioprotonic photodetection over a continuous portion of the visible spectrum, we deposit Pd4-BPR(A20L) and Pd4-HtdR incorporated vesicles on the Pd contacts and measured steady state photocurrents under laser illumination at nine discrete wavelengths. To increase photocurrent detection, we set V=- 0.05 V on the Pd contact so that the H^+ that are pumped into the interstitial layer can transfer into the Pd to form PdH_x. Consistent with the expectation that proton pumping activity should be highest at the rhodopsin absorption maxima, the largest currents (I) were observed at 508 nm for Pd4-BPR(A20L) and at 545 nm for Pd4-HtdR devices, respectively (**Figure 5.5 A**). This photocurrent is the measured current difference prior and after laser illumination. Photocurrents progressively decreased as illumination wavelengths diverged from absorption maxima, and there was a good agreement between current intensity and the UV-visible spectra of Pd4-BPR(A20L) and Pd4-HtdR over the 450-600 nm region of the spectrum (**Figure 5.4 A**). A plot of the photocurrent normalized to the baseline intensity at 450 nm (**Figure 5.5 B**), shows that maximum device response was ~1.6-fold for Pd4-BPR(A20L) and twofold for Pd4-HtdR, and that variations in intensities could be used to track changes in illumination wavelengths.

5.4 Conclusion

By genetically modifying a green-absorbing deltarhodopsin and a blue-absorbing proteorhodopsin with a palladium binding peptide that brings transported protons in close contact with a Pd/PdHx electrode, we have shown that it is possible to detect discrete wavelengths of light by measuring photocurrents associated with interfacial proton reduction over the 450-600 nm range of the visible spectrum. These results illustrate the power of the bioprotonic approach and are a first step towards the construction of wavelength-tuned photodetectors and more complex biological cameras that could be considered for artificial vision or prostheses.

5.5 Figures

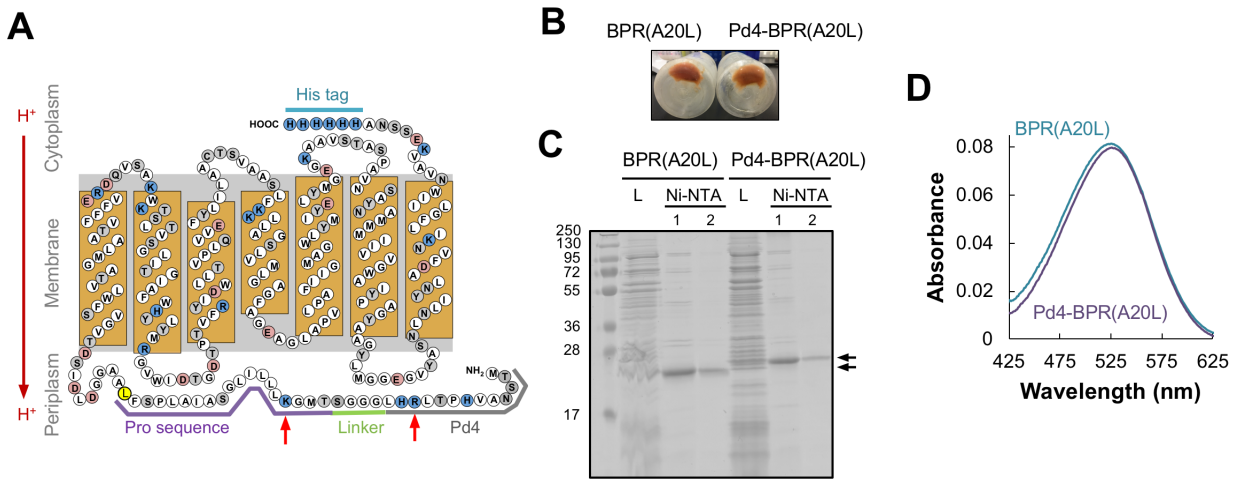

Figure 5.1 Structure, expression, purification and characterization of Pd4-BPR(A20L). **(A)** Cartoon structure of Pd4-BPR(A20L). Amino acids are represented using the one letter code. Transmembrane segments are boxed in orange. The N-terminal pro sequence, linker region, and Pd4 palladium-binding peptide, and the C-terminal hexahistidine tag are labeled. The A20L mutation is identified in yellow. The direction of proton pumping is indicated by a maroon arrow. Red arrows correspond to trypsin cleavage sites. **(B)** Cells expressing BPR(A20L) or Pd4-BPR(A20L) have a pink orange color when grown in the presence of all-*trans* retinal. **(C)** Both proteins can be recovered with a purity exceeding 85% after DDM-solubilized membrane fractions (lanes L) are subjected to two rounds of Ni-NTA chromatography (lanes 1 and 2). Top and bottom arrows identify the migration positions of Pd4-BPR(A20L) and BPR(A20L), respectively. **(D)** UV-Vis spectra of purified protein samples at 3.5 µM concentration and pH 7.4.

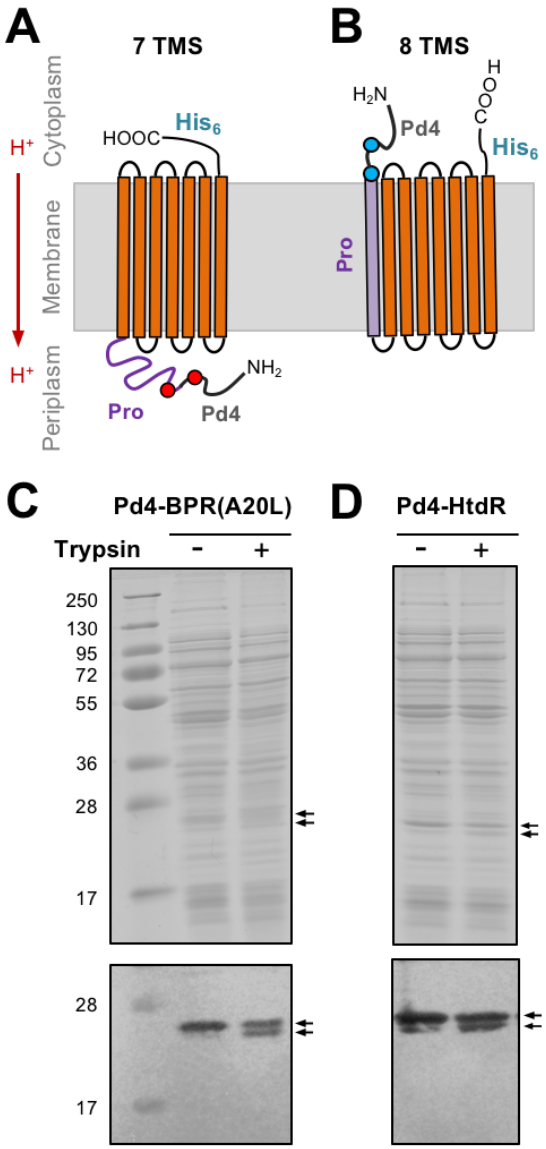

Figure 5.2 The Pd4 domain of Pd4-BPR(A20L) resides in the periplasm. Topologies of 7 transmembrane segments (TMS) version **(A)** and of the 8 TMS version **(B)** of Pd4-BPR(A20L). The direction of proton pumping is indicated by a maroon arrow. Trypsin cleavage sites are identified with red dots if accessible in spheroplasts, and blue dots if not. SDS PAGE fractionation (top panels) and corresponding immuno- analysis with anti-His_6 antibodies (bottom panels) of trypsin-treated spheroplasts prepared from Pd4-BPR(A20L) expressing cells **(C)** or Pd4-HtdR expressing cells **(D)**. The migration positions of intact proteins and trypsin degradation products are indicated by the top and bottom arrows, respectively.

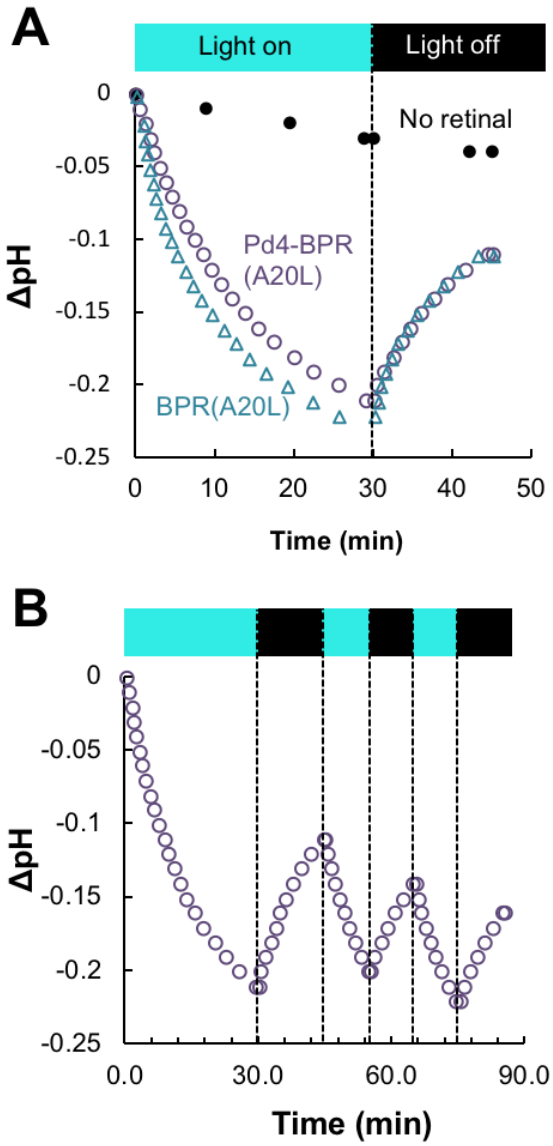

Figure 5.3 Pd4-BPR(A20L) is functional for proton pumping. **(A)** Cells expressing BPR(A20L) (open triangles) or Pd4-BPR(A20L) (open circles) were harvested 3h post-induction in retinal supplemented medium, transferred to unbuffered solutions and the change in pH was monitored with blue illumination on or off. Retinal-free cultures of Pd4-BPR(A20L) were used as a negative control (filled circles). **(B)** Pd4-BPR(A20L) cultures support repeated cycles of light-activated acidification.

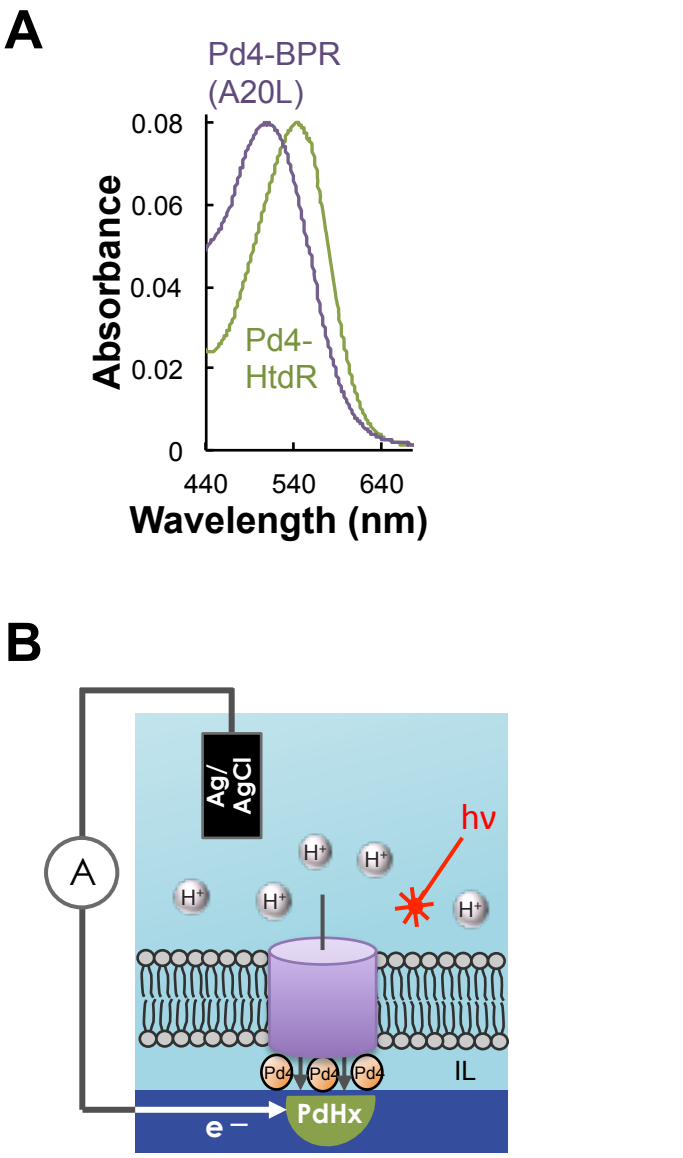

Figure 5.4 Construction and validation of a Pd4-BPR(A20L)-based bioprotonic device. **(A)** UV-visible spectra of detergent-stabilized Pd4-Htdr and Pd4-BPR(A20L) at pH 8.5. **(B)** Schematic representation of the bioprotonic devices. The isolation layer (IL) is labeled.

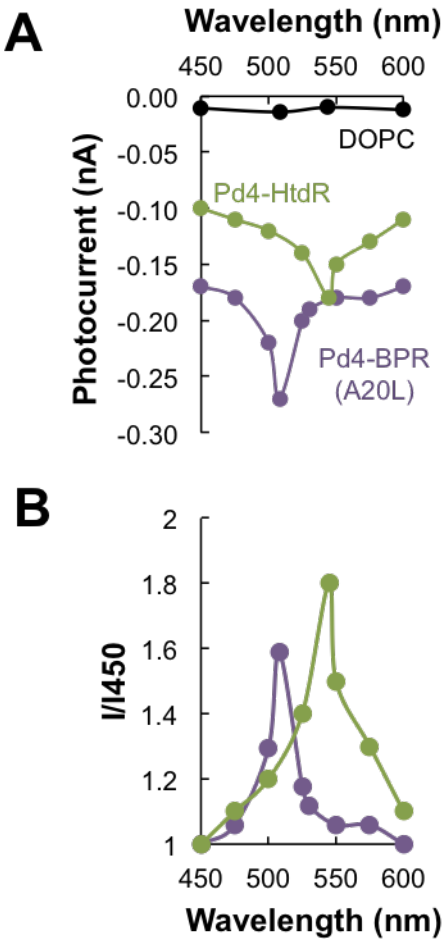

Figure 5.5 Photoresponse of Pd4-HtdR and Pd4-BPR(A20L) devices in response to laser illumination at discrete wavelengths. **(A)** Steady state currents of control (DOPC) and Pd4-HtdR and Pd4-BPR(A20L) based devices. **(B)** Photocurrent intensities were normalized to their value at 450 nm.

Chapter 6 Conclusions

This work exploited the use of SBP for the development of more efficient techniques for protein purification and the fabrication of bioprotonic devices. Chapter 2 focused on optimizing the conditions of binding, washing and elution of a silica based affinity protein purification technique that takes advantage of the reversible binding between proteins tagged with a silica binding peptide (Car9) and silica gels. Disruption of the Car9/silica binding is achieved by competition with lysine. It was demonstrated that binding to silica could be improved by selecting small particle size and large pore size. Additionally, a small concentration of Tween 20 (0.3%) reduces non-specific binding and optimal elution can be achieved under alkaline conditions. Using Car9 as an affinity tag and the optimized conditions, allowed the recovery of fusion proteins with purities higher than 90%. It was also demonstrated that binding and recovery of Car9 tagged proteins remains possible even under denaturing conditions. Finally, it was demonstrated that the Car9 tag is excisable by TEV protease digestion when appended to the N-terminus.

The functionality of SBP was extended to the integration of proteins to bioprotonic devices. In Chapter 3, a palladium binding peptide (Pd4) was fused to the N-terminus of a green light activated deltarhodopsin to efficiently integrate the protein with Pd/PdHx bioprotonic devices for the light controlled conversion of protonic currents into a measurable electronic current. It was demonstrated that the Pd4 tag does not affects protein expression and is capable of binding to palladium surfaces when appended to *H. turkmenica* deltarhodopsin. It was demonstrated that the Pd4 binder brings the protein into a more intimate contact with the devices, decreasing the diffusion path of protons to the surface of the Pd/PdHx protode. This resulted in an 8 to 10 fold increase in current response when compared to the untagged deltarhodopsin.

While optimizing the production of a blue absorbing proteorhodopsin in *E. coli* for the development of dual wavelength bioprotonic devices, it was evident that cleavage of the native signal sequence of BPR was inefficient, resulting in the production two products: a full length pre-BPR and a mature BPR. Since typically, inner membrane proteins in *E. coli* do not contain signal sequences, it was necessary to explore the role of this feature in the biogenesis of BPR in *E. coli.* In Chapter 4, it was demonstrated that a signal sequence-less version of BPR can fold within the inner membrane of *E. coli* but it is produced at significantly lower levels. Introducing an A20L mutation that prevents processing of the signal sequence by SPase I, resulted in the production of 50% more BPR(A20L) when compared to the wild type. It was also demonstrated that processing of the signal sequence occurs in a Sec-independent fashion and that although the signal sequence enhances expression it is not necessary to obtain a functional product. Moreover, the unprocessable BPR(A20L), remains functional for light activated proton transport in *E. coli* cells.

After optimizing the expression of BPR(A20L), a palladium binding version was constructed to fabricate Pd/PdHx bioprotonic devices. Chapter 4 focused on demonstrating that fusion of the Pd4 binder to BPR(A20L) does not affects protein expression and insertion into the *E. coli* membrane, that the Pd4 extension is accessible for binding to palladium surfaces and that the fusion protein (Pd4-BPR(A20L)) remains capable of outward proton transport upon light illumination. Devices were fabricated using Pd4-HtdR as a green light absorbing rhodopsin and Pd4-BPR(A20L) as a blue light absorbing rhodopsin. These devices exhibited wavelength-dependent photocurrent production when illuminated between 450 and 600 nm with higher current response at the proteins absorption maxima.

Together, this work demonstrates the versatility of solid binding peptides for biomimetic approaches. Using Car9 and optimizing the conditions for binding and elution of Car9 tagged proteins on silica gels we were able to develop a rapid and cost effective disposable technology for the purification of proteins. We were also able to use Pd4 to control the biotic/abiotic interface between rhodopsins and Pd/PdHx devices increasing the response of these devices. We demonstrated the wavelength dependency of devices fabricated with green and blue light absorbing rhodopsins, which opens up the possibility of fabricating more complex devices such as biological cameras. Additionally, we optimized the production of proteorhodopsin in *E. coli*, which should prove useful in many applications such as optogenetics and electronic devices.

References

1 Sarikaya, M., Tamerler, C., Jen, A. K., Schulten, K. & Baneyx, F. Molecular biomimetics: nanotechnology through biology. *Nat Mater* **2**, 577-585 (2003).

2 Sarikaya, M., Tamerler, C., Schwartz, D. T. & Baneyx, F. Materials assembly and formation using engineered polypeptides. *Annu Rev Mater Res* **34**, 373-408 (2004).

3 Tamerler, C. & Sarikaya, M. Genetically designed peptide-based molecular materials. *ACS Nano* **3**, 1606-1615 (2009).

4 Care, A., Bergquist, P. L. & Sunna, A. Solid-binding peptides: smart tools for nanobiotechnology. *Trends Biotechnol.* **33**, 259-268 (2015).

5 Coyle, B. L., Zhou, W. & Baneyx, F. in *Bionanotechnology: biological self-assembly and its applications* (ed B. H. A. Rehm) (Caister Academic Press, 2013).

6 Dickerson, M. B., Sandhage, K. H. & Naik, R. R. Protein- and peptide directed synthesis of inorganic materials. *Chem. Rev.* **108**, 4935-4978 (2008).

7 Baneyx, F. & Schwartz, D. T. Selection and analysis of solid-binding peptides. *Curr. Opin. Biotechnol.* **18**, 312-317, doi:10.1016/j.copbio.2007.04.008 (2007).

8 Pacardo, D. B., Sethi, M., Jones, S. E., Naik, R. R. & Knecht, M. R. Biomimetic synthesis of Pd nanocatalysts for the Stille coupling reaction. *ACS nano* **3**, 1288-1296 (2009).

9 Dickerson, M. B. *et al.* Identification and design of peptides for the rapid, high-yield formation of nanoparticulate TiO2 from aqueous solutions at room temperature. *Chem. Mater.* **20**, 1578-1584 (2008).

10 Tamerler, C., Oren, E. E., Duman, M., Venkatasubramanian, E. & Sarikaya, M. Adsorption kinetics of an engineered gold binding peptide by surface plasmon resonance spectroscopy and a quartz crystal microbalance. *Langmuir* **22**, 7712-7718 (2006).

11 Naik, R. R., Stringer, S. J., Agarwal, G., Jones, S. E. & Stone, M. O. Biomimetic synthesis and patterning of silver nanoparticles. *Nature materials* **1**, 169-172 (2002).

12 Gungormus, M. *et al.* Regulation of in vitro calcium phosphate mineralization by combinatorially selected hydroxyapatite-binding peptides. *Biomacromolecules* **9**, 966-973 (2008).

13 Care, A., Petroll, K., Gibson, E. S., Bergquist, P. L. & Sunna, A. Solid-binding peptides for immobilisation of thermostable enzymes to hydrolyse biomass polysaccharides. *Biotechnology for Biofuels* **10**, 29 (2017).

14 Kröger, N., Deutzmann, R. & Sumper, M. Polycationic peptides from diatom biosilica that direct silica nanosphere formation. *Science* **286**, 1129-1132 (1999).

15 Poulsen, N., Sumper, M. & Kröger, N. Biosilica formation in diatoms: characterization of native silaffin-2 and its role in silica morphogenesis. *Proceedings of the National Academy of Sciences* **100**, 12075-12080 (2003).

16 Hernández-Hernández, A. *et al.* Influence of eggshell matrix proteins on the precipitation of calcium carbonate (CaCO 3). *J. Cryst. Growth* **310**, 1754-1759 (2008).

17 Yan, L. *et al.* Magnetotactic bacteria, magnetosomes and their application. *Microbiol. Res.* **167**, 507-519 (2012).

18 Kröger, N. *et al.* Bioenabled synthesis of rutile (TiO2) at ambient temperature and neutral pH. *Angew. Chem. Int. Ed.* **45**, 7239-7243 (2006).

19 Chiu, C.-Y. *et al.* Platinum nanocrystals selectively shaped using facet-specific peptide sequences. *Nature chemistry* **3**, 393-399 (2011).

20 Dai, H. *et al.* Nonequilibrium Synthesis and Assembly of Hybrid Inorganic− Protein Nanostructures Using an Engineered DNA Binding Protein. *J. Am. Chem. Soc.* **127**, 15637-15643 (2005).
21 Chiu, D. *et al.* Biomineralization and Size Control of Stable Calcium Phosphate Core Protein Shell Nanoparticles: Potential for Vaccine Applications. *Bioconj. Chem.* **23**, 610 (2012).
22 Zhou, W. & Baneyx, F. Aqueous, protein-driven synthesis of transition metal-doped ZnS immuno-quantum dots. *ACS nano* **5**, 8013 (2011).
23 Brondyk, W. H. in *Methods Enzymol.* Vol. 463 131-147 (Elsevier, 2009).
24 Saraswat, M. *et al.* Preparative purification of recombinant proteins: current status and future trends. *BioMed research international* **2013** (2013).
25 Wingfield, P. T. Overview of the purification of recombinant proteins. *Current protocols in protein science* **80**, 6.1. 1-6.1. 35 (2015).
26 Selkirk, C. Ion-exchange chromatography. *Protein Purification Protocols*, 125-131 (2004).
27 Queiroz, J., Tomaz, C. & Cabral, J. Hydrophobic interaction chromatography of proteins. *J. Biotechnol.* **87**, 143-159 (2001).
28 Hong, P., Koza, S. & Bouvier, E. S. A review size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. *J. Liq. Chromatogr. Rel. Technol.* **35**, 2923-2950 (2012).
29 Urh, M., Simpson, D. & Zhao, K. Affinity chromatography: general methods. *Methods Enzymol.* **463**, 417-438 (2009).
30 Waugh, D. S. Making the most of affinity tags. *Trends Biotechnol.* **23**, 316-320 (2005).
31 Young, C. L., Britton, Z. T. & Robinson, A. S. Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnol J* **7**, 620-634, doi:10.1002/biot.201100155 (2012).
32 Lebendiker, M. & Danieli, T. in *Protein Chromatography* 281-293 (Springer, 2011).
33 Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8**, 1668-1674 (1999).
34 Harper, S. & Speicher, D. W. in *Protein Chromatography* 259-280 (Springer, 2011).
35 Thain, A., Gaston, K., Jenkins, O. & Clarke, A. R. A method for the separation of GST fusion proteins from co-purifying GroEL. *Trends Genet.* **12**, 209-210 (1996).
36 Korndörfer, I. P. & Skerra, A. Improved affinity of engineered streptavidin for the Strep‐tag II peptide is due to a fixed open conformation of the lid‐like loop at the binding site. *Protein Sci.* **11**, 883-893 (2002).
37 Bornhorst, J. A. & Falke, J. J. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol.* **326**, 245-254 (2000).
38 Kuo, W.-H. K. & Chase, H. A. Exploiting the interactions between poly-histidine fusion tags and immobilized metal ions. *Biotechnol. Lett.* **33**, 1075-1084 (2011).
39 Braun, P. *et al.* Proteome-scale purification of human proteins from bacteria. *Proc Natl Acad Sci U S A* **99**, 2654-2659, doi:10.1073/pnas.042684199 (2002).
40 Edwards, A. M. *et al.* Protein production: feeding the crystallographers and NMR spectroscopists. *Nat. Struct. Biol.* **7 Suppl**, 970-972, doi:10.1038/80751 (2000).
41 Shih, Y. P. *et al.* High-throughput screening of soluble recombinant proteins. *Protein Sci.* **11**, 1714-1719, doi:10.1110/ps.0205202 (2002).

42 Sastry, M. S., Zhou, W. & Baneyx, F. Integrity of N- and C-termini is important for E. coli Hsp31 chaperone activity. *Protein Sci.* **18**, 1439-1447, doi:10.1002/pro.158 (2009).
43 Geoghegan, K. F. *et al.* Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **267**, 169-184, doi:10.1006/abio.1998.2990 (1999).
44 Huang, A., de Jong, R. N., Folkers, G. E. & Boelens, R. NMR characterization of foldedness for the production of E3 RING domains. *Journal of structural biology* **172**, 120-127, doi:10.1016/j.jsb.2010.07.014 (2010).
45 Lichty, J. J., Malecki, J. L., Michelson-Horowitz, D. J. & Tan, S. Comparison of affinity tags for protein purification. *Protein Expr Purif* **41**, 98-105 (2005).
46 Iler, R. K. *The chemistry of silica: Solubility, polymerization, colloid and surface properites, and biochemistry*. (Wiley Intersciences, 1979).
47 Hertl, W. & Hair, M. L. Reaction of Hexamethyldisilazane with Silica. *J. Phys. Chem.* **75**, 2181-& (1971).
48 Puddu, V. & Perry, C. C. Peptide adsorption on silica nanoparticles: evidence of hydrophobic interactions. *ACS Nano* **6**, 6356-6363, doi:10.1021/nn301866q (2012).
49 Puddu, V. & Perry, C. C. Interactions at the silica-peptide interface: the influence of particle size and surface functionality. *Langmuir* **30**, 227-233, doi:10.1021/la403242f (2014).
50 Emami, F. S. *et al.* Force field and a surface model database for silica to simulate interfacial properties in atomic resolution. *Chem. Mater.* **26**, 2647-2658 (2014).
51 Sulkowski, E. in *Protein purification: micro to macro* (ed R. Burgess) (Alan R. Liss, 1987).
52 Mathe, C. *et al.* Structural determinants for protein adsorption/non-adsorption to silica surface. *PLoS One* **8**, e81346, doi:10.1371/journal.pone.0081346 (2013).
53 Klein, G. *et al.* RNA-binding proteins are a major target of silica nanoparticles in cell extracts. *Nanotoxicology* **10**, 1555-1564, doi:10.1080/17435390.2016.1244299 (2016).
54 Ghose, S., McNerney, T. M. & Hubbard, B. Preparative protein purification on underivatized silica. *Biotechnol. Bioeng.* **87**, 413-423, doi:10.1002/bit.20125 (2004).
55 Ikeda, T., Ninomiya, K., Hirota, R. & Kuroda, A. Single-step affinity purification of recombinant proteins using the silica-binding Si-tag as a fusion partner. *Protein Expr Purif* **71**, 91-95, doi:10.1016/j.pep.2009.12.009 (2010).
56 Ikeda, T. *et al.* The silica-binding Si-tag functions as an affinity tag even under denaturing conditions. *Protein Expression Purif.* **77**, 173-177 (2011).
57 Abdelhamid, M. A. A. *et al.* Affinity purification of recombinant proteins using a novel silica-binding peptide as a fusion tag. *Appl. Microbiol. Biotechnol.* **98**, 5677-5684, doi:10.1007/s00253-014-5754-z (2014).
58 Coyle, B. L., Rolandi, M. & Baneyx, F. Carbon-binding designer proteins that discriminate between sp2- and sp3-hybridized carbon surfaces. *Langmuir* **29**, 4839-4846 (2013).
59 Coyle, B. L. & Baneyx, F. A cleavable silica‐binding affinity tag for rapid and inexpensive protein purification. *Biotechnol. Bioeng.* **111**, 2019-2026 (2014).
60 Yang, W., Hellner, B. & Baneyx, F. Self-Immobilization of Car9 Fusion Proteins within High Surface Area Silica Sol-Gels and Dynamic Control of Protein Release. *Bioconjug Chem* **27**, 2450-2459, doi:10.1021/acs.bioconjchem.6b00406 (2016).

61 Coyle, B. L. & Baneyx, F. Direct and reversible immobilization and microcontact printing of functional proteins on glass using a genetically appended silica-binding tag. *Chem. Commun.* **52**, 7001-7004, doi:10.1039/c6cc02660e (2016).
62 Tsirigotaki, A., De Geyter, J., Sostaric, N., Economou, A. & Karamanou, S. Protein export through the bacterial Sec pathway. *Nat Rev Microbiol* **15**, 21-36, doi:10.1038/nrmicro.2016.161 (2017).
63 Economou, A. Following the leader: bacterial protein export through the Sec pathway. *Trends Microbiol.* **7**, 315-320 (1999).
64 Paetzel, M. Structure and mechanism of Escherichia coli type I signal peptidase. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1843**, 1497-1508 (2014).
65 Natale, P., Brüser, T. & Driessen, A. J. Sec-and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1778**, 1735-1756 (2008).
66 Beck, K., Wu, L. F., Brunner, J. & Müller, M. Discrimination between SRP‐and SecA/SecB‐dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *The EMBO journal* **19**, 134-143 (2000).
67 Luirink, J., Yu, Z., Wagner, S. & de Gier, J.-W. Biogenesis of inner membrane proteins in Escherichia coli. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1817**, 965-976 (2012).
68 Schibich, D. *et al.* Global profiling of SRP interaction with nascent polypeptides. *Nature* **536**, 219 (2016).
69 Beja, O., Spudich, E. N., Spudich, J. L., Leclerc, M. & DeLong, E. F. Proteorhodopsin phototrophy in the ocean. *Nature* **411**, 786-789 (2001).
70 Kamo, N. *et al.* A light-driven proton pump from Haloterrigena turkmenica: functional expression in Escherichia coli membrane and coupling with a H+ co-transporter. *Biochem. Biophys. Res. Commun.* **341**, 285-290, doi:10.1016/j.bbrc.2005.12.181 (2006).
71 Lanyi, J. K. & Luecke, H. Bacteriorhodopsin. *Curr. Opin. Struct. Biol.* **11**, 415-419 (2001).
72 Zhang, J. *et al.* Crystal structure of deltarhodopsin-3 from Haloterrigena thermotolerans. *Proteins* **81**, 1585-1592, doi:10.1002/prot.24316 (2013).
73 Schobert, B. & Lanyi, J. K. Halorhodopsin is a light-driven chloride pump. *J. Biol. Chem.* **257**, 10306-10313 (1982).
74 Gordeliy, V. I. *et al.* Molecular basis of transmembrane signalling by sensory rhodopsin II–transducer complex. *Nature* **419**, 484-487 (2002).
75 Schafer, G., Engelhard, M. & Muller, V. Bioenergetics of the Archaea. *Microbiol. Mol. Biol. Rev.* **63**, 570-620 (1999).
76 Nannenga, B. L. & Baneyx, F. Enhanced expression of membrane proteins in E. coli with a P(BAD) promoter mutant: synergies with chaperone pathway engineering strategies. *Microb Cell Fact* **10**, 105, doi:10.1186/1475-2859-10-105 (2011).
77 Nannenga, B. L. & Baneyx, F. Reprogramming chaperone pathways to improve membrane protein expression in Escherichia coli. *Protein Sci.* **20**, 1411-1420, doi:10.1002/pro.669 (2011).
78 Béja, O. *et al.* Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902-1906 (2000).
79 Man, D. *et al.* Diversification and spectral tuning in marine proteorhodopsins. *The EMBO journal* **22**, 1725-1731 (2003).

80 Kelemen, B. R., Du, M. & Jensen, R. B. Proteorhodopsin in living color: diversity of spectral properties within living bacterial cells. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1618**, 25-32 (2003).
81 Dioumaev, A. K. *et al.* Proton transfers in the photochemical reaction cycle of proteorhodopsin. *Biochemistry* **41**, 5348-5358 (2002).
82 Lanyi, J. K. Proton transfers in the bacteriorhodopsin photocycle. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1757**, 1012-1018 (2006).
83 Ran, T. *et al.* Cross-protomer interaction with the photoactive site in oligomeric proteorhodopsin complexes. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **69**, 1965-1980 (2013).
84 Wang, W.-W., Sineshchekov, O. A., Spudich, E. N. & Spudich, J. L. Spectroscopic and photochemical characterization of a deep ocean proteorhodopsin. *J. Biol. Chem.* **278**, 33985-33991 (2003).
85 Noy, A. Bionanoelectronics. *Adv. Mater.* **23**, 807-820, doi:10.1002/adma.201003751 (2011).
86 Yoo, E.-H. & Lee, S.-Y. Glucose biosensors: an overview of use in clinical practice. *Sensors* **10**, 4558-4576 (2010).
87 Yang, W., Thordarson, P., Gooding, J. J., Ringer, S. P. & Braet, F. Carbon nanotubes for biological and biomedical applications. *Nanotechnology* **18**, 412001 (2007).
88 Huang, S. C. *et al.* Carbon nanotube transistor controlled by a biological ion pump gate. *Nano Lett.* **10**, 1812-1816, doi:10.1021/nl100499x (2010).
89 Misra, N. *et al.* Bioelectronic silicon nanowire devices using functional membrane proteins. *Proceedings of the National Academy of Sciences* **106**, 13780-13784 (2009).
90 Zhou, W., Dai, X. & Lieber, C. M. Advances in nanowire bioelectronics. *Rep. Prog. Phys.* **80**, 016701 (2016).
91 Tybrandt, K., Larsson, K. C., Richter-Dahlfors, A. & Berggren, M. Ion bipolar junction transistors. *Proc Natl Acad Sci U S A* **107**, 9929-9932, doi:10.1073/pnas.0913911107 (2010).
92 Mostert, A. B. *et al.* Role of semiconductivity and ion transport in the electrical conduction of melanin. *Proc Natl Acad Sci U S A* **109**, 8943-8947, doi:10.1073/pnas.1119948109 (2012).
93 Campana, A., Cramer, T., Simon, D. T., Berggren, M. & Biscarini, F. Electrocardiographic recording with conformable organic electrochemical transistor fabricated on resorbable bioscaffold. *Adv. Mater.* **26**, 3874-3878, doi:10.1002/adma.201400263 (2014).
94 Tybrandt, K., Forchheimer, R. & Berggren, M. Logic gates based on ion transistors. *Nat Commun* **3**, 871, doi:10.1038/ncomms1869 (2012).
95 Stavrinidou, E. *et al.* Electronic plants. *Science Advances* **1**, e1501136, doi:10.1126/sciadv.1501136 (2015).
96 Morowitz, H. J. Proton semiconductors and energy transduction in biological systems. *Am J Physiol* **235**, R99-114 (1978).
97 Mitchell, P. Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. 1966. *Biochim. Biophys. Acta* **1807**, 1507-1538, doi:10.1016/j.bbabio.2011.09.018 (2011).

98 Walz, D. & Caplan, S. R. Bacterial flagellar motor and H(+)/ATP synthase: two proton-driven rotary molecular devices with different functions. *Bioelectrochemistry* **55**, 89-92 (2002).
99 Smith, S. M. *et al.* Voltage-gated proton channel in a dinoflagellate. *Proc Natl Acad Sci U S A* **108**, 18162-18167, doi:10.1073/pnas.1115405108 (2011).
100 Deng, Y. *et al.* H+-type and OH- -type biological protonic semiconductors and complementary devices. *Sci Rep* **3**, 2481, doi:10.1038/srep02481 (2013).
101 Hemmatian, Z. *et al.* Taking electrons out of bioelectronics: bioprotonic memories, transistors, and enzyme logic. *Journal of Materials Chemistry C* **3**, 6407-6412, doi:10.1039/c5tc00502g (2015).
102 Josberger, E. E., Deng, Y., Sun, W., Kautz, R. & Rolandi, M. Two-terminal protonic devices with synaptic-like short-term depression and device memory. *Adv. Mater.* **26**, 4986-4990, doi:10.1002/adma.201400320 (2014).
103 Miyake, T., Josberger, E. E., Keene, S., Deng, Y. X. & Rolandi, M. An enzyme logic bioprotonic transducer. *APL Mater* **3**, doi:Artn 014906
10.1063/1.4900886 (2015).
104 Ordinario, D. D., Phan, L., Jocson, J.-M., Nguyen, T. & Gorodetsky, A. A. Protonic transistors from thin reflectin films. *APL Mater* **3**, 014907 (2015).
105 Wunsche, J. *et al.* Protonic and Electronic Transport in Hydrated Thin Films of the Pigment Eumelanin. *Chem. Mater.* **27**, 436-442, doi:10.1021/cm502939r (2015).
106 Zhong, C. *et al.* A polysaccharide bioprotonic field-effect transistor. *Nat Commun* **2**, 476, doi:10.1038/ncomms1489 (2011).
107 Hemmatian, Z. *et al.* Electronic control of H+ current in a bioprotonic device with Gramicidin A and Alamethicin. *Nat Commun* **7** (2016).
108 Glasser, L. Proton conduction and injection in solids. *Chem. Rev.* **75**, 21-65 (1975).
109 Pandey, R. B. *et al.* Adsorption of peptides (A3, Flg, Pd2, Pd4) on gold and palladium surfaces by a coarse-grained Monte Carlo simulation. *Phys. Chem. Chem. Phys.* **11**, 1989-2001, doi:10.1039/b816187a (2009).
110 Bell, M. R., Engleka, M. J., Malik, A. & Strickler, J. E. To fuse or not to fuse: what is your purpose? *Protein Sci.* **22**, 1466-1477, doi:10.1002/pro.2356 (2013).
111 Waugh, D. S. An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr. Purif.* **80**, 283-293 (2011).
112 Coyle, B. L. & Baneyx, F. A cleavable silica-binding affinity tag for rapid and inexpensive protein purification. *Biotechnol. Bioeng.* **111**, 2019-2026, doi:10.1002/bit.25257 (2014).
113 Swift, B. J., Shadish, J. A., DeForest, C. A. & Baneyx, F. Streamlined Synthesis and Assembly of a Hybrid Sensing Architecture with Solid Binding Proteins and Click Chemistry. *J. Am. Chem. Soc.* **139**, 3958-3961, doi:10.1021/jacs.7b00519 (2017).
114 Nannenga, B. L. & Baneyx, F. Reprogramming chaperone pathways to improve membrane protein expression in Escherichia coli. *Protein Sci.* **20**, 1411-1420, doi:10.1002/pro.669 (2011).
115 Tropea, J. E., Cherry, S. & Waugh, D. S. Expression and purification of soluble His(6)-tagged TEV protease. *Methods Mol Biol* **498**, 297-307, doi:10.1007/978-1-59745-196-3_19 (2009).
116 Ung, K. K. *Porous silica*. Vol. 16 (Elsevier Science, 1979).

117 Stutz, H. Protein attachment onto silica surfaces--a survey of molecular fundamentals, resulting effects and novel preventive strategies in CE. *Electrophoresis* **30**, 2032-2061, doi:10.1002/elps.200900015 (2009).
118 Patwardhan, S. V. *et al.* Chemistry of aqueous silica nanoparticle surfaces and the mechanism of selective peptide adsorption. *J. Am. Chem. Soc.* **134**, 6244-6256, doi:10.1021/ja211307u (2012).
119 Vlasova, N. N. & Golovkova, L. P. The adsorption of amino acids on the surface of highly dispersed silica. *Colloid J.* **66**, 657-662, doi:DOI 10.1007/s10595-005-0042-3 (2004).
120 Dumetz, A. C., Snellinger-O'brien A, M., Kaler, E. W. & Lenhoff, A. M. Patterns of protein protein interactions in salt solutions and implications for protein crystallization. *Protein Sci.* **16**, 1867-1877, doi:10.1110/ps.072957907 (2007).
121 Tsumoto, K., Ejima, D., Senczuk, A. M., Kita, Y. & Arakawa, T. Effects of salts on protein-surface interactions: applications for column chromatography. *J. Pharm. Sci.* **96**, 1677-1690, doi:10.1002/jps.20821 (2007).
122 Priev, A., Almagor, A., Yedgar, S. & Gavish, B. Glycerol decreases the volume and compressibility of protein interior. *Biochemistry* **35**, 2061-2066, doi:10.1021/bi951842r (1996).
123 Vagenende, V., Yap, M. G. & Trout, B. L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry* **48**, 11084-11096, doi:10.1021/bi900649t (2009).
124 Ikeda, T. *et al.* The silica-binding Si-tag functions as an affinity tag even under denaturing conditions. *Protein Expr. Purif.* **77**, 173-177, doi:10.1016/j.pep.2011.01.012 (2011).
125 Makino, S., Reynolds, J. A. & Tanford, C. The binding of deoxycholate and Triton X-100 to proteins. *J. Biol. Chem.* **248**, 4926-4932 (1973).
126 Ghaemi, Y. & Wall, R. A. Hydrophobic Chromatography with Dynamically Coated Stationary Phases .3. Non-Ionic Surfactant Effects. *J. Chromatogr.* **198**, 397-405, doi:Doi 10.1016/S0021-9673(00)80508-7 (1980).
127 Stempfer, G., Holl-Neugebauer, B. & Rudolph, R. Improved refolding of an immobilized fusion protein. *Nat. Biotechnol.* **14**, 329-334, doi:10.1038/nbt0396-329 (1996).
128 Kweon, D. H., Lee, D. H., Han, N. S. & Seo, J. H. Solid-phase refolding of cyclodextrin glycosyltransferase adsorbed on cation-exchange resin. *Biotechnol. Prog.* **20**, 277-283, doi:10.1021/bp0341895 (2004).
129 Berdichevsky, Y. *et al.* Matrix-assisted refolding of single-chain Fv- cellulose binding domain fusion proteins. *Protein Expr Purif* **17**, 249-259, doi:10.1006/prep.1999.1125 (1999).
130 Ikeda, T. & Kuroda, A. Why does the silica-binding protein "Si-tag" bind strongly to silica surfaces? Implications of conformational adaptation of the intrinsically disordered polypeptide to solid surfaces. *Colloids Surf. B. Biointerfaces* **86**, 359-363, doi:10.1016/j.colsurfb.2011.04.020 (2011).
131 Meredith, P., Bettinger, C. J., Irimia-Vladu, M., Mostert, A. B. & Schwenn, P. E. Electronic and optoelectronic materials and devices inspired by nature. *Rep. Prog. Phys.* **76**, 034501, doi:10.1088/0034-4885/76/3/034501 (2013).

132 Svennersten, K., Larsson, K. C., Berggren, M. & Richter-Dahlfors, A. Organic bioelectronics in nanomedicine. *Biochim. Biophys. Acta* **1810**, 276-285, doi:10.1016/j.bbagen.2010.10.001 (2011).
133 Kandori, H. *et al.* Structural Changes of pharaonis Phoborhodopsin upon Photoisomerization of the Retinal Chromofore: Infrared Spectral Comparison with Bacteriorhodopsin. *Biochemistry* **40**, 9, doi:10.1021/bi0103819 (2001).
134 Liang, X., Mao, G. & Ng, K. Y. Mechanical properties and stability measurement of cholesterol-containing liposome on mica by atomic force microscopy. *J. Colloid Interface Sci.* **278**, 53-62, doi:10.1016/j.jcis.2004.05.042 (2004).
135 Laouini, A. *et al.* Preparation, Characterization and Applications of Liposomes: State of the Art. *Journal of Colloid Science and Biotechnology* **1**, 147-168, doi:10.1166/jcsb.2012.1020 (2012).
136 Rigaud, J. L., Paternostre, M. T. & Bluzat, A. Mechanisms of membrane protein insertion into liposomes during reconstitution procedures involving the use of detergents. 2. Incorporation of the light-driven proton pump bacteriorhodopsin. *Biochemistry* **27**, 2677-2688 (1988).
137 Lambert, O., Levy, D., Ranck, J.-L., Leblanc, G. & Rigaud, J.-L. A new "gel-like" phase in dodecyl maltoside–lipid mixtures: implications in solubilization and reconstitution studies. *Biophys. J.* **74**, 918-930 (1998).
138 Weber, M. J. *et al.* Atomic layer deposition of high-purity palladium films from Pd (hfac) 2 and H2 and O2 plasmas. *The Journal of Physical Chemistry C* **118**, 8702-8711 (2014).
139 Nannenga, B. L. & Baneyx, F. Enhanced expression of membrane proteins in E. coli with a PBAD promoter mutant: synergies with chaperone pathway engineering strategies. *Microb Cell Fact* **10**, 105, doi:10.1186/1475-2859-10-105 (2011).
140 Nannenga, B. L. & Baneyx, F. Reprogramming chaperone pathways to improve membrane protein expression in Escherichia coli. *Protein Sci.*, doi:10.1002/pro.669 (2011).
141 Pacardo, D. B., Sethi, M., Jones, S. E., Naik, R. R. & Knecht, M. R. Biomimetic synthesis of Pd nanocatalysts for the Stille coupling reaction. *ACS Nano* **3**, 1288-1296, doi:10.1021/nn9002709 (2009).
142 Minsky, A., Summers, R. G. & Knowles, J. R. Secretion of beta-lactamase into the periplasm of Escherichia coli: evidence for a distinct release step associated with a conformational change. *Proc Natl Acad Sci U S A* **83**, 4180-4184 (1986).
143 Knol, J., Sjollema, K. & Poolman, B. Detergent-mediated reconstitution of membrane proteins. *Biochemistry* **37**, 16410-16415, doi:10.1021/bi981596u (1998).
144 Rigaud, J. L., Pitard, B. & Levy, D. Reconstitution of membrane proteins into liposomes: application to energy-transducing membrane proteins. *Biochim. Biophys. Acta* **1231**, 223-246 (1995).
145 Tang, Z. *et al.* Biomolecular recognition principles for bionanocombinatorics: an integrated approach to elucidate enthalpic and entropic factors. *ACS Nano* **7**, 9632-9646, doi:10.1021/nn404427y (2013).
146 DeCoursey, T. E. & Hosler, J. Philosophy of voltage-gated proton channels. *J R Soc Interface* **11**, 20130799, doi:10.1098/rsif.2013.0799 (2014).
147 Palczewska, W. Catalytic reactivity of hydrogen on palladium and nickel hydride phases. *Adv. Catal* **24**, 235-291 (1975).

148 Baneyx, F. & Schwartz, D. T. Selection and analysis of solid-binding peptides. *Curr. Opin. Biotechnol.* **18**, 312-317 (2007).
149 Care, A., Bergquist, P. L. & Sunna, A. Solid-binding peptides: smart tools for nanobiotechnology. *Trends Biotechnol.* **33**, 259-268, doi:10.1016/j.tibtech.2015.02.005 (2015).
150 Bamann, C., Bamberg, E., Wachtveitl, J. & Glaubitz, C. Proteorhodopsin. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1837**, 614-625 (2014).
151 Costanzi, S., Siegel, J., Tikhonova, I. G. & Jacobson, K. A. Rhodopsin and the others: a historical perspective on structural studies of G protein-coupled receptors. *Curr. Pharm. Des.* **15**, 3994-4002 (2009).
152 McCusker, E. C., Bane, S. E., O'Malley, M. A. & Robinson, A. S. Heterologous GPCR expression: a bottleneck to obtaining crystal structures. *Biotechnol. Prog.* **23**, 540-547, doi:10.1021/bp060349b (2007).
153 Gourdon, P. *et al.* Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump. *Protein Expression Purif.* **58**, 103-113 (2008).
154 Soto‐Rodríguez, J., Hemmatian, Z., Josberger, E. E., Rolandi, M. & Baneyx, F. A Palladium‐Binding Deltarhodopsin for Light‐Activated Conversion of Protonic to Electronic Currents. *Adv. Mater.* **28**, 6581-6585 (2016).
155 Hemmatian, Z. *et al.* Electronic control of H+ current in a bioprotonic device with Gramicidin A and Alamethicin. *Nat Commun* **7**, 12981, doi:10.1038/ncomms12981 (2016).
156 Nannenga, B. L. & Baneyx, F. Folding engineering strategies for efficient membrane protein production in E. coli. *Methods Mol. Biol.* **899**, 187-202, doi:10.1007/978-1-61779-921-1_12 (2012).
157 Bischoff, L., Wickles, S., Berninghausen, O., Van Der Sluis, E. O. & Beckmann, R. Visualization of a polytopic membrane protein during SecY-mediated membrane insertion. *Nat Commun* **5**, 4103 (2014).
158 Friedrich, T. *et al.* Proteorhodopsin is a light-driven proton pump with variable vectoriality. *J. Mol. Biol.* **321**, 821-838 (2002).
159 Pfleger, N. *et al.* Solid-state NMR and functional studies on proteorhodopsin. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1787**, 697-705 (2009).
160 Ulbrandt, N. D., Newitt, J. A. & Bernstein, H. D. The E. coli signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell* **88**, 187-196 (1997).
161 Shi, L. *et al.* Three-dimensional solid-state NMR study of a seven-helical integral membrane proton pump—structural insights. *J. Mol. Biol.* **386**, 1078-1093 (2009).
162 Karamyshev, A. L., Karamysheva, Z. N., Kajava, A. V., Ksenzenko, V. N. & Nesmeyanova, M. A. Processing of Escherichia coli alkaline phosphatase: role of the primary structure of the signal peptide cleavage region. *J. Mol. Biol.* **277**, 859-870 (1998).
163 Chun, S.-Y. & Randall, L. L. In vivo studies of the role of SecA during protein export in Escherichia coli. *Journal of bacteriology* **176**, 4197-4203 (1994).
164 Hillebrecht, J. R. *et al.* Structure, function, and wavelength selection in blue-absorbing proteorhodopsin. *Biochemistry* **45**, 1579-1590 (2006).

165 Pinhassi, J., DeLong, E. F., Béjà, O., González, J. M. & Pedrós-Alió, C. Marine bacterial and archaeal ion-pumping rhodopsins: genetic diversity, physiology, and ecology. *Microbiol. Mol. Biol. Rev.* **80**, 929-954 (2016).
166 Wang, S., Yang, C.-I. & Shan, S.-o. SecA mediates cotranslational targeting and translocation of an inner membrane protein. *Journal of Cell Biology* **216**, 3639-3653 (2017).
167 Lee, H. C. & Bernstein, H. D. The targeting pathway of Escherichia coli presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proceedings of the National Academy of Sciences* **98**, 3471-3476 (2001).
168 Samuelson, J. C. *et al.* YidC mediates membrane protein insertion in bacteria. *Nature* **406**, 637 (2000).
169 Kühlbrandt, W. Bacteriorhodopsin—the movie. *Nature* **406**, 569-570 (2000).
170 Yan, B. *et al.* Spectral tuning in bacteriorhodopsin in the absence of counterion and coplanarization effects. *Journal of Biological Chemistry* **270**, 29668-29670 (1995).
171 Soto-Rodríguez, J. & Baneyx, F. Role of the Signal Sequence in Proteorhodopsin Biogenesis in E. coli *Biotechnol. Bioeng.* (2018).
172 Olsen, J. V., Ong, S.-E. & Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics* **3**, 608-614 (2004).
173 Pandey, R. B. *et al.* Adsorption of peptides (A3, Flg, Pd2, Pd4) on gold and palladium surfaces by a coarse-grained Monte Carlo simulation. *PCCP* **11**, 1989-2001 (2009).
174 Johansson, M. *et al.* Hydrogen adsorption on palladium and palladium hydride at 1 bar. *Surface Science* **604**, 718-729, doi:10.1016/j.susc.2010.01.023 (2010).

Appendix A: Supplementary Information for Chapter 3

Quantitative estimate for photocurrent time dependence

We suggest that the photocurrent (i_p) decrease that starts ~2s after the onset of illumination results from a decrease in the number of Pd sites available for the reduction of H^+ to H. This decrease is caused by Pd turning into PdH from the absorption of the H product of the H^+ reduction. To verify this hypothesis, we estimate the occupancy ratio (Θ) at t = 2s by calculating the total amount of H produced through the reduction of H^+ and dividing this number by the total number of Pd sites available on the Pd contact. Assuming a 1:1 correlation between the number of electrons in the reduction current and the number of protons reduced to hydrogen (N_{H+}) we use the data of **Figure 3.5 c** to obtain:

$$N_{H+} = \frac{\int_{t=0}^{t=2s} i_p \mathrm{d}t}{e} \qquad (1)$$

where e is the elemental charge, 1.6×10^{-19} C

Averaging values for the 7 current spikes of **Fig. 3**c yields $N_{H+} = 2.3 \times 10^8$.

The number of available Pd sites on a 10μm x 2μm contact is the product of the surface area by the number of Pd sites per unit area $N_s = 1.5 \times 10^{19}$ m^{-2}.[174]

$$N_{Pd} = 2 \times 10^{-11}\ m^2 \times 1.5 \times 10^{19}\ m^{-2} = 3 \times 10^8 \qquad (2)$$

Thus, the Pd site occupancy after a 2s current spike is

$$\Theta = N_{H+} / N_{Pd} \sim 0.8 \qquad (3)$$

It is likely that Θ is higher than 80% in the vicinity of each Pd4-HdtR-His_6 trimer leading to Pd site saturation and a gradual decrease in photocurrent until steady state is reached.