Nanopores, megatonnes, and milliseconds: Exploring engineered peptides as antimicrobial, carbon-capture, and biocatalytic agents

by

Roberto Juan Barbero

A.B. Engineering modified with biology
Dartmouth College, 2001
B.E. Biochemical Engineering
Thayer School of Engineering at Dartmouth College, 2002

Submitted to the Department of Biological Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

February 2012

Signature of Author .....
Roberto Juan Barbero
Department of Biological Engineering
November 28, 2011

Certified by ....
.....
Angela Belcher
W.M. Keck Professor of Energy
Professor of Materials Science and Engineering
Professor of Biological Engineering

Accepted by...
.....
K. Dane Wittrup
C.P. Dubbs Professor of Chemical Engineering and Biological Engineering
Associate Director, Koch Institute for Integrative Cancer Research
Chair, Course XX Graduate Program Committee

Abstract

This work investigates the roles that peptides play in the fields of antimicrobials, surface functionalization, carbon capture, and biocatalysis. The results demonstrate that peptides, sometimes dismissed for their lack of complexity, can have a breadth of applications.

First, the killing kinetics of a pore-forming, engineered antimicrobial peptide (CM15) were imaged using a high-speed atomic force microscope (HS-AFM). The fast time resolution of the HS-AFM (13 seconds per image) enabled characterization of the initial stages of the killing of live *Escherichia coli* cells. The results suggested that the killing process of CM15 is a combination of a time-variable incubation phase and a more rapid execution phase, offering an interesting parallel between antimicrobial-peptide-induced death and mammalian cell apoptotic death.

As a follow-up, an engineered peptide (2K1) with high affinity toward oxide surfaces was used to functionalize a diverse set of materials, including titanium dioxide, zinc, and stainless steel. After demonstrating that 2K1 works as affinity tag for small molecules and fusion proteins, a 2K1-CM15 peptide was made in an attempt to develop a single-step, facile antimicrobial functionalization of oxide surfaces.

Second, motivated by the role of peptides in mineralization processes, the yeast *Saccharomyces cerevisae* was engineered to display peptides and proteins that enhanced the capture of CO_2. An industrial-scale CO_2 mineralization process was designed using this engineered yeast with an associated cost of \$52 per tonne of CO_2. The effect of the engineered yeast on the process was significant – the cost of CO_2 capture was decreased by 8.5–13.5%, as compared to a process with no biological components.

Finally, M13 bacteriophage (M13 phage) was established as a temperature stable, highly-multivalent biocatalytic scaffold through display of engineered histidine-biased peptides. A protocol for generating histidine-biased peptide libraries displayed on the major coat protein (pVIII) of M13 phage was developed. By analogy to known histidine-based active sites, seven sequences were chosen from amongst hundreds of sequenced histidine-biased pVIII peptides. Two demonstrated esterase activity with a $\frac{K_{cat}}{K_{uncat}} \approx 170$ that matches, and a $K_m \approx 4$ mM that is only 20-fold lower than, that reported for a computationally designed esterase.

Acknowledgements

"We struggle with, agonize over and bluster heroically about the great questions of life when the answers to most of these lie hidden in our attitude toward the thousand minor details of each day."

- Robert Grudin, *Time and the Art of Living*

When I graduated from college ten years ago, I had no plans to return to school for a graduate degree in engineering. For me, engineering skills were best learned while working on practical applications with a team of colleagues, and earning a Ph.D. meant that one must toil alone unnecessarily focused on the minor details of a project of limited utility to the wider world. However, my experiences working in industry showed me that it is actually those who have mastery over the nuances and minor details who have the capacity to affect real change.

In graduate school, I expected to sacrifice my preference for working with teams of colleagues in order to improve my technical foundation. Fortunately, I had the opportunity to become a part of the biological engineering community at MIT, which provided me with a collaborative and supportive environment. I would like to thank Doug Lauffenburger for establishing and inspiring this community and all of my B.E. colleagues for their dedication to Doug's vision.

I would especially like to thank my adviser, Angela Belcher, for her support and guidance. I am absolutely sure that there is no Ph.D. adviser at MIT who would have been better for me than Angie. She has the ability to strategically guide projects from a high-level, while also being able to suggest how to prioritize and manage the daily experiments. I learned a lot from her about how to write about and present my work, skills that are of critical importance for any scientist. In spite of all her success, she is a very fun and easy person to work with. She has established a wonderful working environment in her lab, allowing her graduate students and post-docs to work independently or collaboratively, and she has an uncanny ability to know whether someone will be a good fit for the collegial Belcher lab culture within minutes of meeting them. Finally, she has a positive and supportive way of motivating the members of her lab. I cannot thank her enough for her continued interest in and excitement over my work, regardless of how few and far between the results may have been at times.

I originally approached my thesis committee with the CO_2 capture project as the only one for my thesis. Over the years, though, I added projects, and every time I did so, I received encouragement and solid technical advice from my committee, Ed DeLong and Dane Wittrup. They are both very busy, world-class scientists and I thank them for their time. I was always impressed with how quickly they could immerse themselves in the details of a new project and suggest experiments or approaches that I had not yet considered.

I would like to acknowledge all of my Belcher lab colleagues for their help and friendship. Eric Krauland was the only B.E. student in the lab when I joined and he immediately helped me find a project and offered a lot of early advice. Rana Gosh, Hyunjung Yi, Yoon Sung Nam, and Ahmad Khalil have all been cubicle neighbors over the past five years, and I would like to thank them for conversations and support.

Georg Fanter and David Gray were a pleasure to work with on the HS-AFM project, and I thank them for their hard work despite how difficult it was to perform those experiments.

I would also like to thank Tanguy Chau, Mike Koeris, and Tim Lu for helpful conversations about AmP's.

The CO_2 capture project germinated in Angie's mind and she helped me to flesh out the details for the approach that we took. It was a complex project and I could not have completed it without the help of two amazing undergraduates, Anna Simon and Albert Kao. I would like to acknowledge Elizabeth Wood for her modeling and process design contributions and Jennifer Hsieh for help with x-ray diffraction. This project was funded through the generosity of Eni, and I would like to thank them for their support, both financially and intellectually. Nicola De Blasio was very adept in coordinating the project across two continents. Daniele Bianchi also helped manage the project, reviewed all of my work and helped to define the project from an industrial perspective. Lino Carnelli, Alessandra d'Arminio Monforte, and Moreno Riccò worked on the process design and the economic evaluation.

The biocatalytic phage project was challenging and I would like to thank Angie for encouraging me to start on it, for reassuring me when people said it could not be done, and for being so excited when we started to get results. I knew little about phage biology when I started and I would specifically like to thank Paul Widboom, Hyunjung Yi, Nimrod Heldman, and Rana Ghosh for teaching me how to make libraries and how to grow and purify phage. Paul's single His libraries were the starting point for my triple His libraries, Hyunjung's library protocol was a fantastic starting point for mine, and Nimrod's vast molecular biology knowledge (and his hard work in lab) made the triple His libraries a reality. It has been very exciting for me to have another graduate student demonstrate an interest in this project. I would like to thank John Casey for his help thus far and for continuing this project after I leave. I would also like to help Nurxat Nuraje, John Burpo, Noòemie-Manuelle Dorval Courchesne, and Vivian Wu for their help with aspects of this project.

In addition to Eni, I would like to acknowledge the MIT Biotechnology Training Program, the MIT Presidential Fellowship, the Medtronic Fellowship, and the Siebel Foundation for financial support. I would like to acknowledge the MIT Center for Materials Science and Engineering for help with x-ray diffraction and TGA, the MIT Biopolymers & Proteomics Core Facility for DNA sequencing services, the MIT Central Machine Shop for help with my CO_2 capture reactor, and Debby Pheasant in the MIT Biophysical Instrumentation Facility for help with MST and ITC. I would like to thank all of the BE staff, especially Jared Embelton, Dalia Fares, and Aran Parillo.

Finally, I would like to thank my wife, Megan, for being my best friend for the last fourteen years. She has been the source of inspiration and motivation for all my work and I am so fortunate to have her as part of my life. She also read every single page of this thesis and helped make it readable and understandable. Our daughter Tessa joined us in April 2010 and I want her to know how much I love her and that I look forward to every new day with her. My parents, Carol and Fred, gave me a childhood that was as close to ideal as is possible, and I thank them for teaching me to love to learn, to want to make the world a better place, and to recognize the importance of family. My brother, Cris, has been a part of my life for almost thirty one years and I can only say that if the next thirty one years together are that much fun, I will have no complaints.

Contents

List of Figures

List of Tables

Chapter 1

High-speed AFM imaging of the killing kinetics of an engineered antimicrobial peptide

1.1 Summary

Observations of real-time changes in living cells have contributed much to the field of cellular biology. Eluding the field thus far had been the ability to image whole, living cells with nanometer resolution on a time scale that is relevant to dynamic cellular processes. [1, 2] In this work the kinetics of individual bacterial cell death were investigated using a novel high-speed atomic force microscope (AFM) optimized for imaging live cells in real time. The increased time resolution (13 seconds per image) allowed the characterization of the initial stages of the action of an antimicrobial peptide (AmP) CM15 on individual *Escherichia coli* cells with nanometer resolution. The results suggested that the killing process of CM15 is a combination of a time-variable incubation phase (which takes seconds to minutes to complete) and a more rapid execution phase. The results also offered an interesting parallel between antimicrobial-peptide-induced death and mammalian cell apoptotic death.[1]

1.2 Introduction

High resolution AFM images of live cells and cell fragments have resolved the nanoscale structure of cell walls, the structural dynamics of single spores germinating, the structural changes of bacteria treated with bactericidal enzymes, and the location of specific binding sites. [4–9] One of the limitations that made AFM unsuitable for studying many dynamic processes in cell biology was the long image acquisition time of several minutes for one high-resolution AFM image. [10] Progress in AFM instrumentation had enabled imaging of single molecules at up to video rate, but this had been limited to small areas (hundreds of nanometers) and to flat samples. [11–13] The results in this work resulted from the use of a high-speed AFM specifically designed to bridge the gap between the improved AFM

[1]The work presented in this chapter was previously published in *Nature Nanotechnology*. [3]

technology and the biologically relevant but challenging experiments interesting to cellular biology. The key benefit of this instrument was the combination of very soft, but rapid imaging and a sufficient scan size to monitor multiple cells at once.

AmPs are a promising class of antimicrobials which have demonstrated activity against antibiotic resistant bacteria, parasites, viruses and fungi. [14–17] Moreover, resistance to AmPs appears to be much less probably than antibiotic resistance has proven to be. [18] There are several proposed targets of antimicrobial peptides, including inhibition of cell-wall synthesis, inhibition of nucleic-acid synthesis, and inhibition of protein synthesis, though the majority of AmP appear to kill their targets through the formation of transmembrane pore formation. [19] This class of AmPs have high killing activity against a wide variety of bacteria and much lower activity against mammalian cells. This is due to an interesting difference between bacterial and mammalian cell membranes. In bacterial membranes, the outermost layer exposed to the world is heavily populated with negatively charged phospholipids. In mammalian cells, though, the outermost layer of the membrane mostly is composed of lipids with no net charge. Most AmPs are positively charged with clusters of hydrophobic residues. They are attracted to the negatively-charged outer membrane of bacterial cells and, at critical concentrations, are able to displace the lipids and disrupt the membrane structure (because of the hydrophobic residues), sometimes even by forming pores in the membrane. [18]

In this work, a high-speed AFM was used to measure the kinetics of the pre-death activity of a pore-forming, membrane-disrupting AmP, called CM15, on individual live *E. coli* cells in an aqueous solution. CM15 is a well-studied, engineered AmP, comprised of the first seven amino acids of the from AmP cecropin-A and residues 2–9 from the bee venom AmP melittin. CM15 has higher potency relative to cecropin A and without the cytotoxic effects of melittin. Moreover, being only 15 residues long, it is in a size range that is therapeutically relevant. It has high activity against *E. coli* ($<$ 1–10 μm, depending on concentration of cells) and *Staphylococcus aureus* (2 μm). [20,21]

Prior to this work, electron microscopy and AFM experiments had demonstrated the endpoint surface morphological changes of a population of cells treated with AmPs. [22–24] Spectroscopic analyses of synthetic membranes or vesicles had provided insight into the sizes and structures of pores formed by AmPs. [25–27] However, the early stage kinetics of the membrane-disrupting activity of an AmP on live cells had not been reported with nanometer-spatial and seconds-temporal resolution.

1.3 Results

1.3.1 Small cantilevers reduce forces at high imaging speeds

Imaging live bacteria in an aqueous environment is one of the most challenging applications for AFM. [28] The force interaction between the cantilever tip and the sample is critical; if the force is too high, the bacteria are damaged or disconnected from the surface, whereas if it is too low, information about the sample is lost. Maintaining this balance at high imaging speeds was achieved using prototype AFM components and micro-fabricated small cantilevers ($\approx$ 1000 times smaller mass than conventional cantilevers) with integrated tips. [29] These cantilevers have resonance frequencies in liquid at $\approx$ 100 kHz and spring constants between

300 and 1000 $\frac{pN}{nm}$. The increased resonance frequency and the reduced mass, quality factor and spring constant enabled an increased imaging speed in tapping mode while retaining low imaging forces on the bacteria. [30] Figure 1.1a shows an SEM image of the small cantilevers used for this research. The inset shows a comparison with a conventional cantilever (NPS-B, Veeco Metrology). Figure 1.1b shows the thermal spectra of the two cantilevers. The resonance frequency of the small cantilever is 17 times higher than that of the conventional cantilever in air and is 30 times higher in fluid.

1.3.2 High-speed AFM imaging of AmP activity on live *E. coli* cells in liquid

In this work, bacteria were immobilized on poly-L-lysine coated cover slides. The bacteria were imaged in aqueous solution for at least 10 minutes to ensure that the cells were not altered or displaced by the AFM tip and to ensure that the poly-L-lysine did not change the cells in the timescales observed in our experiments. CM15 was added to the liquid droplet around the sample to a final concentration of 50 $\frac{\mu g}{mL}$ L or >5 times the minimum inhibitory concentration (MIC) at the concentration of cells being treated (see Table 1.1). Images were acquired every 13 seconds.

To determine the concentration of the cells on the slide, the average number of cells in a 100 μm square was determined by averaging the values for several images. This average "coverage" was then extrapolated across the surface area of the cover slide to determine a total number of cells. This total number of cells was divided by the volume of the imaging liquid to provide an estimated concentration of cells (see Table 1.1).

Description	Value
Area of disk (μm^2)	9.42 x 10^6
Number of cells per square 10 μm	10
Number of cells on disk	9.42 x 10^6
Volume of liquid (μL)	100
Concentration of cells ($\frac{cells}{mL}$)	9.42 x 10^7
MIC of CM15 at 10^8 $\frac{cells}{mL}$ ($\frac{\mu g}{mL}$)	8 [21]

Table 1.1: E. coli *concentration in high-speed AFM flow cell and MIC of CM15*

Figure 1.2a shows AFM phase images of the surfaces of two bacteria before and at several times after the addition of CM15. The most apparent effect of the CM15 was that the surfaces of the bacteria changed from smooth to corrugated. AFM phase data is shown due to increased clarity in the image. The same changes in surface morphology are present in AFM amplitude data (Figure 1.3). AFM height data was more difficult to interpret due to large background variation on the surface of the bacteria. The changes in the surfaces of the bacteria were consistent with published electron microscopy data that report ultra structural damage to the outside of peptide-treated bacterial cells. [22–24] Interestingly, there was a wide range in the time of onset of the change between individual bacteria. Bacterium 1 in

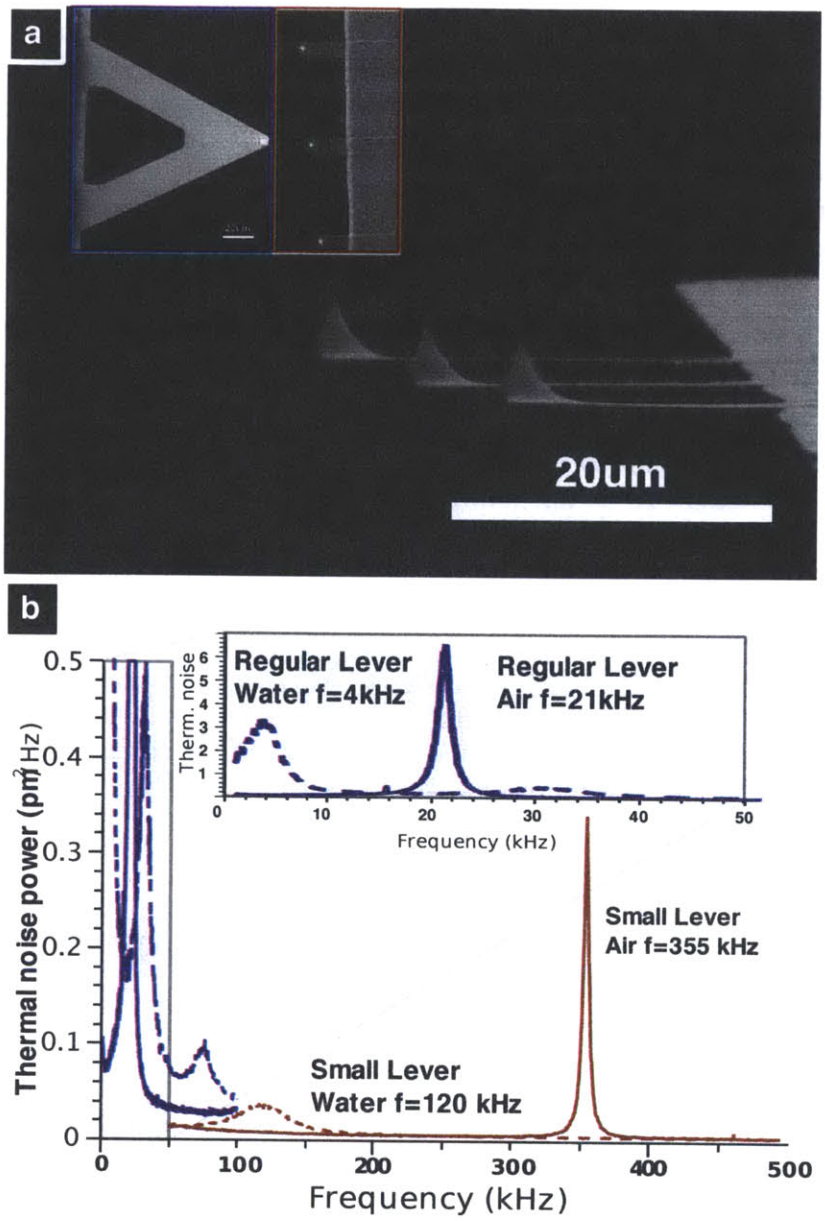


Figure 1.1: *Small AFM cantilevers for high-speed AFM imaging of bacterial cells in liquid* (a) SEM image of the small SiN cantilevers ($\approx$ 10 μm wide, 100–350 nm thick, and 20–30 μm long). The inset image compares the small cantilevers (on the right) to a conventional Veeco NPS-D cantilever used for imaging in fluid (on the left) at the same magnification. (b) Thermal noise power spectra of regular and small cantilevers. In air (red solid line), the first resonance frequency of the small cantilever is $\approx$ 350 kHz. In aqueous solution this drops to 100–120 kHz (red dashed line). The inset shows the thermal noise power spectra of an NP-S cantilever B with resonance frequencies of 21 kHz in air (blue solid line) and 4 kHz in aqueous solution (blue dashed line). Reprinted with permission from *Nature Nanotechnology*.

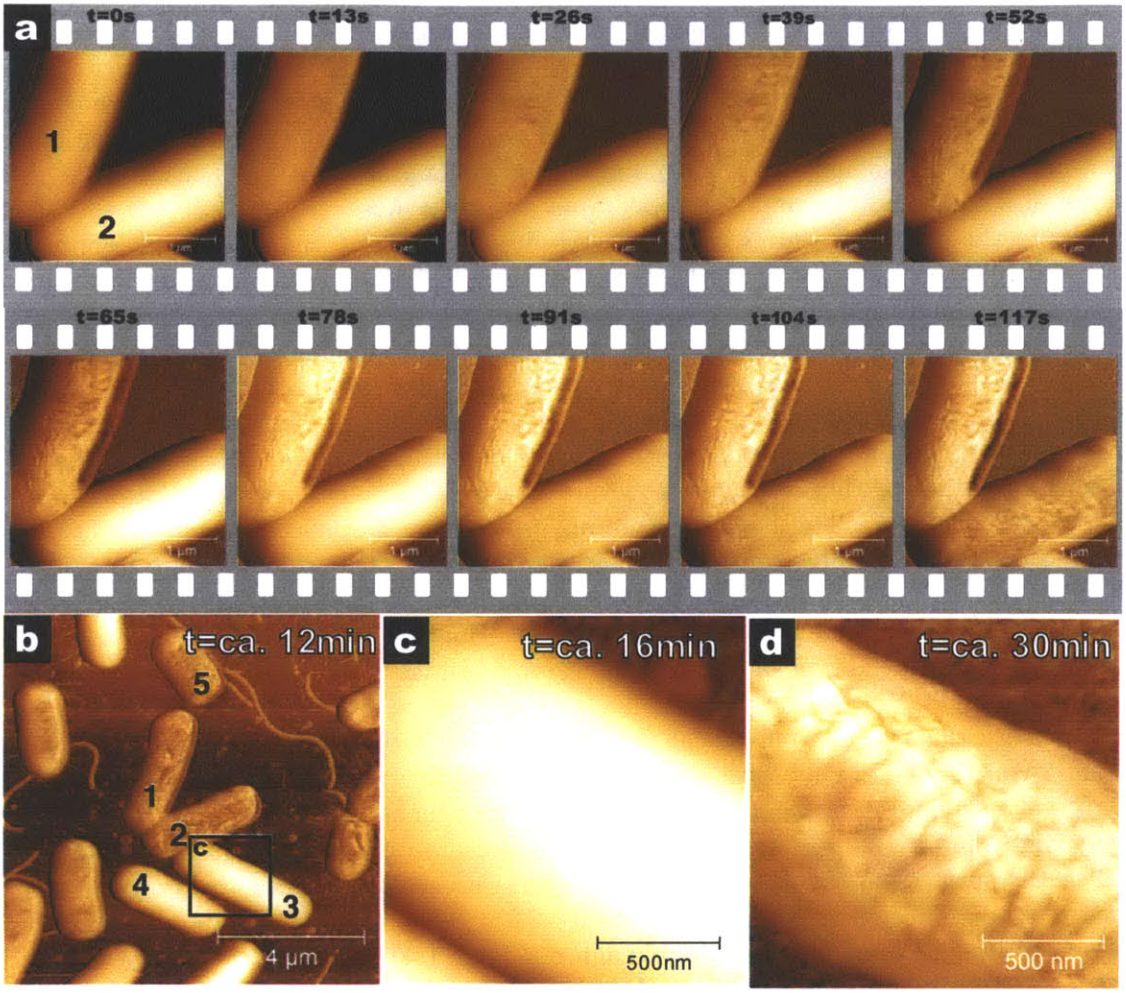


Figure 1.2: E. coli *membrane disruption induced by CM15 imaged with high-speed AFM* (a) Time series of CM15 antimicrobial action. CM15 was injected at $t = -6$ seconds and images were recorded every 13 seconds, with a resolution of 1024 x 256 pixels and rate of 20 $\frac{lines}{second}$. The upper bacterium's surface started changing within 13 seconds. The lower bacterium resisted changing for 78 seconds. (b) Larger area view recorded 12 minutes after addition of CM15. Most bacteria were corrugated but some were still smooth. (c) High-resolution image of bacterium 3 showed that this bacterium was still smooth at $t = 16$ minutes. (d) Image of the now corrugated bacterium 3 at $t = 30$ minutes. Eventually, all bacteria in the field of view were affected by CM15. Images were recorded in liquid in tapping mode with a tapping frequency of 110 kHz. Phase images are shown here for high contrast; amplitude data is shown in Figure 1.3. Images (b), (c), and (d) were recorded with 1024 x 256 pixels at 2 $\frac{lines}{second}$. Reprinted with permission from *Nature Nanotechnology*.

Figure 1.2a started changing within 13 seconds after addition of CM15 and the change was completed in ≈60 seconds. Bacterium 2 did not start changing until ≈80 seconds and the change was not complete until ≈120 seconds.

The phase data was used for the evaluation of the changes in the bacterial surface due to the increased contrast in these images. Since phase data can also include information about a change in materials properties it was not possible to absolutely rule out that the contrast was increased by such a change. However, a change in topography does have a significant effect on the phase signal, such that obtaining materials properties information from phase images of samples with significant topography is hardly possible. Also the fact that the changes in the phase data are somewhat scan-direction dependent (for example comparing trace and re-trace, data not shown) led to the conclusion that the majority of the features in the phase signal were due to a change in topography.

Figure 1.2b shows a larger area of the same sample 12 minutes after addition of CM15 bacteria 1 and 2 are in the center of Figure 2b. Some bacteria still had not changed (bacteria 3 and 4). Figure 1.2c is a higher resolution image of the smooth surface of bacterium 3, 16 minutes after addition of CM15. Eventually, this bacterium (Figure 1.2D) and all bacteria in the field of view became corrugated. Several reasons for the cell-to-cell variation in onset time were considered. The bacteria were all grown from a single clone, and, thus, were expected to be genetically identical. They were close to each other (within 10 μm) and therefore were exposed to the same CM15 concentration at the same time. The cells were likely not all in the same stage in the cell cycle. However, in all of the experiments, no correlation between the time of onset of change and cell length (an indicator of stage in cell cycle) was found. [31]

To ensure that the morphological change observed in this experiment was induced by CM15, a control experiment with a positively-charged peptide (2K1 from Chapter 2) with no known antimicrobial action was performed. As demonstrated in Chapter 2, 2K1 has a very high binding affinity to any negatively charged surface, and so it was expected to bind to the surfaces of the *E. coli* cells, but it has no antimicrobial activity on *E. coli* at concentrations lower than 200 μm (see Figure 2.8). No morphological change in the cells was observed, even after 45 minutes of imaging (Figure 1.4).

A second control experiment was performed with the conventional antibiotic ampicillin (Figure 1.5). Ampicillin inhibits cell wall synthesis, so it would not be expected to induce the kind of morphological changes caused by a cell-membrane disrupting AmP. As expected, even after almost 2 hours of imaging, there was still no significant morphological change in the cells.

1.3.3 AmP-induced surface morphology change correlates with cell death

To correlate the corrugation of the cell surface with cell death, combined AFM and fluorescence microscopy using a fluorescent indicator of cell viability (LIVE/DEAD BacLight) was performed. Figure 1.6a shows an AFM image of live cells. To minimize disruptions caused by the AFM tip, the imaging force was reduced until only the tops of the bacteria were imaged. Figure 1.6b shows a fluorescence image of the same cells recorded immediately after

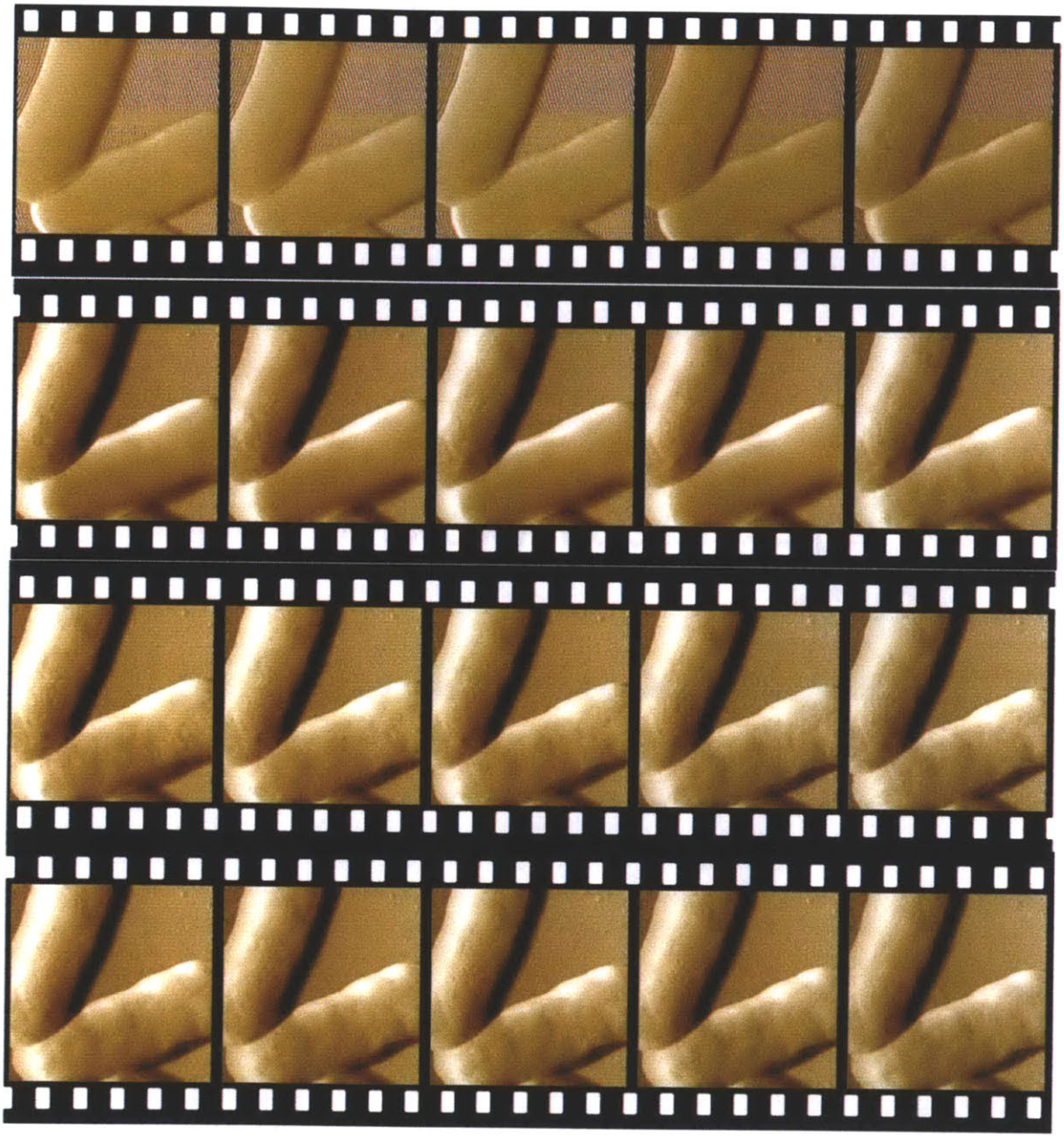


Figure 1.3: *Amplitude data of full time series from Figure 1.2* Controller gains were set aggressively to ensure proper tracking of the ROI (region of interest) on top of the bacteria, which resulted in oscillations in areas where no bacteria were present. Images are 3 μm x 3 μm. Reprinted with permission from *Nature Nanotechnology.*

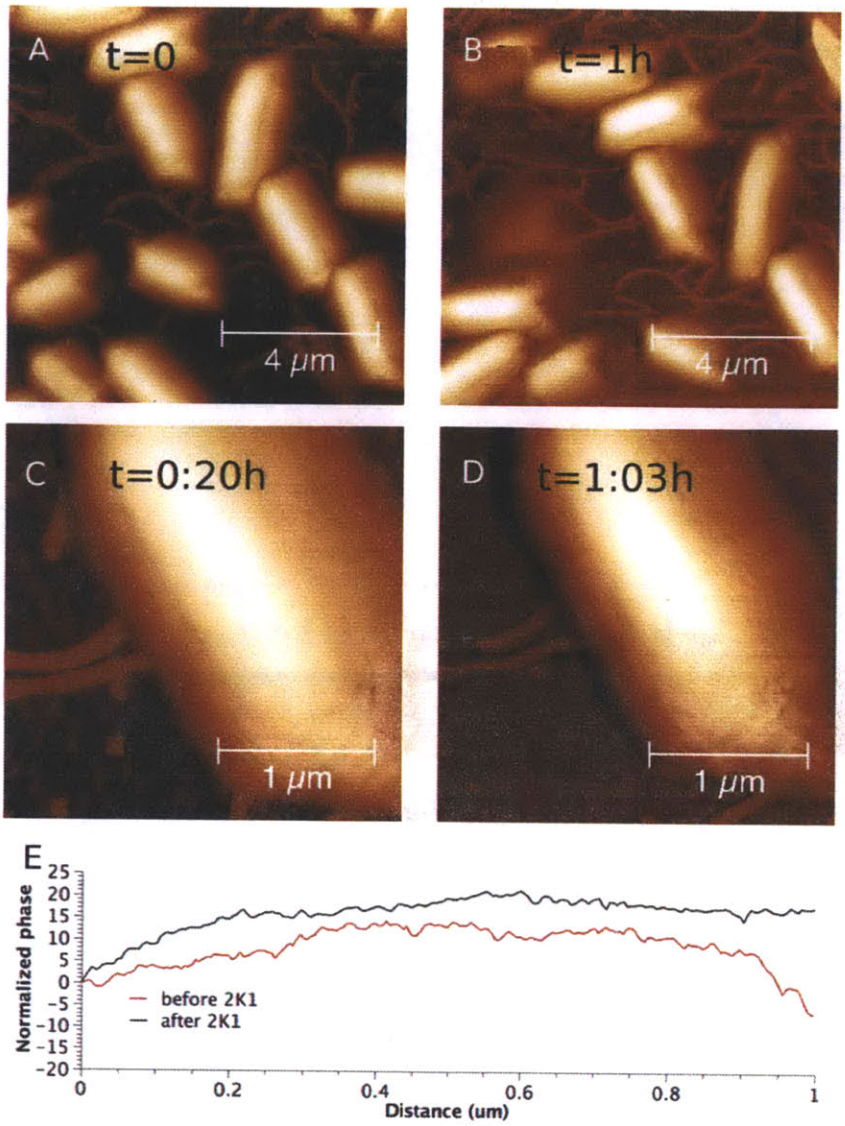


Figure 1.4: *AFM images of negative control with 2K1* In order to ensure that the observed morphology changes are in fact due to the antimicrobial action of CM15 and not just to the addition of a peptide, 2K1 was used as a control. 2K1 is positively charged, so one would expect it to bind to the surface of bacterial cells, yet it lacks the hydrophobic regions that antimicrobial peptides typically have and it has no measurable antimicrobial action at the concentration used (see Figure 2.8). Phase images before addition of 2K1 (a) and (c) and 45 minutes after addition of 2K1 (b) and (d). The cross-sections in panel (e) showed no significant difference in the surface variations before and after 2K1 addition. Reprinted with permission from *Nature Nanotechnology*.

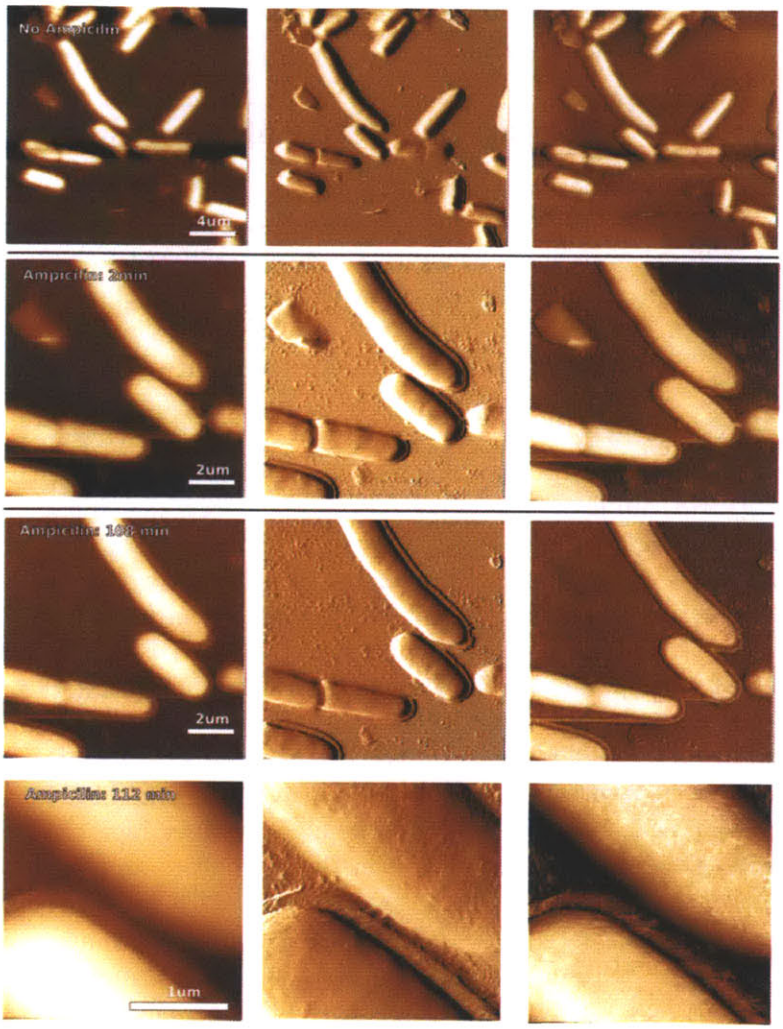


Figure 1.5: *AFM images of negative control with the antibiotic ampicillin* The effect of a common antibiotic ampicillin on the surface morphology of the *E. coli* cells was tested. The columns represent height, amplitude, and phase data, respectively. The rows represent different time points after addition of ampicillin (no ampicillin, 2 minutes, 108 minutes and 112 minutes after addition of ampicillin). After 112 minutes, minor changes were visible on the cell surface, but they were much less than that caused by CM15. Reprinted with permission from *Nature Nanotechnology*.

the AFM imaging. Figure 1.6c and 1.6d show the AFM and fluorescence images of the same bacteria 30 minutes after the cells were exposed to 2 times the MIC of CM15. Almost all of the bacteria have corrugated surfaces and there are no green (live) cells remaining. The apparent broadening of the cells is an imaging artifact due to the pyramidal shape of the tip. [32] There was a strong correlation between the fluorescence and surface variation of the cells (p-value < 0.004) (see Figure 1.6e).

1.3.4 Quantitative analysis of kinetics of AmP activity

The high time resolution of these measurements enabled the characterization of the initial stages of antimicrobial action on individual live bacteria. The high-speed AFM experiment was repeated with a larger scan area to investigate relative differences in the response times of adjacent cells. At $t = 0$ seconds, using a flow-through system the imaging solution was exchanged with solution of CM15 that was twice the MIC. One image was taken every 21 seconds, with every 5^{th} image in Figure 1.7 (full series in Figures 1.8 and 1.9). It is evident that bacteria 1, 7 and 8 responded more quickly to the addition of CM15 than the other bacteria.

To quantify the kinetics of the change for each bacterium separately, the change in the root mean square (RMS) value of the surface corrugation along the long axis of the bacteria in every frame of the image was measured. Figure 1.7b shows representative cross sections through bacterium 1 for each image. The change in surface corrugation was visible $\approx$80 seconds after the addition of CM15. Figure 1.7c shows the normalized RMS corrugation values for the bacteria in Figure 1.7a as a function of time after addition of CM15.

Time of onset of change was highly variable. In this experiment it took 40 seconds (bacterium 7) to more than 4 minutes (bacterium 2) average of 155 seconds $\pm$ 89 seconds (mean $\pm$ standard deviation). The time for each bacterium to complete the change (smooth to corrugated) was more consistent – 50 % of the damage was completed in 52 seconds $\pm$ 16 seconds. Once again, phase data was used for the analysis, since here the changes are most apparent. The same trends could be observed in the height and amplitude images (Figure 1.10), though they were more difficult to interpret because of the higher background noise in those images. The CM15-treated cells were monitored for up to 100 minutes after treatment and no additional changes were observed. Eventually, the cells disappeared, but it was not possible to say whether this was due to lysis or due to the detachment of cells. Figure 1.7d shows the results of a measurement of the kinetics of CM15 activity on a population of cells. The bulk killing rate of CM15 could be described reasonably well with a single exponential with a half-time of 4.6 minutes.

With this AFM technique, one might expect to see the pores that are formed in the outer membrane by CM15. The pores formed by CM15 are reported to have diameters of 2–4 nm. [21] Unfortunately, the pixel resolution of the high-speed images is 4–10 nm, scan-size dependent, and is thus insufficient to see features of this size. Using a slower acquisition speed ($2\ \frac{lines}{second}$), images were obtained where the pixel resolution was 1 nm for a 1 μm image (Figure 1.11). While there are features whose size match the reported pore size of CM15 pores in this higher resolution image, it is difficult to interpret much from them.

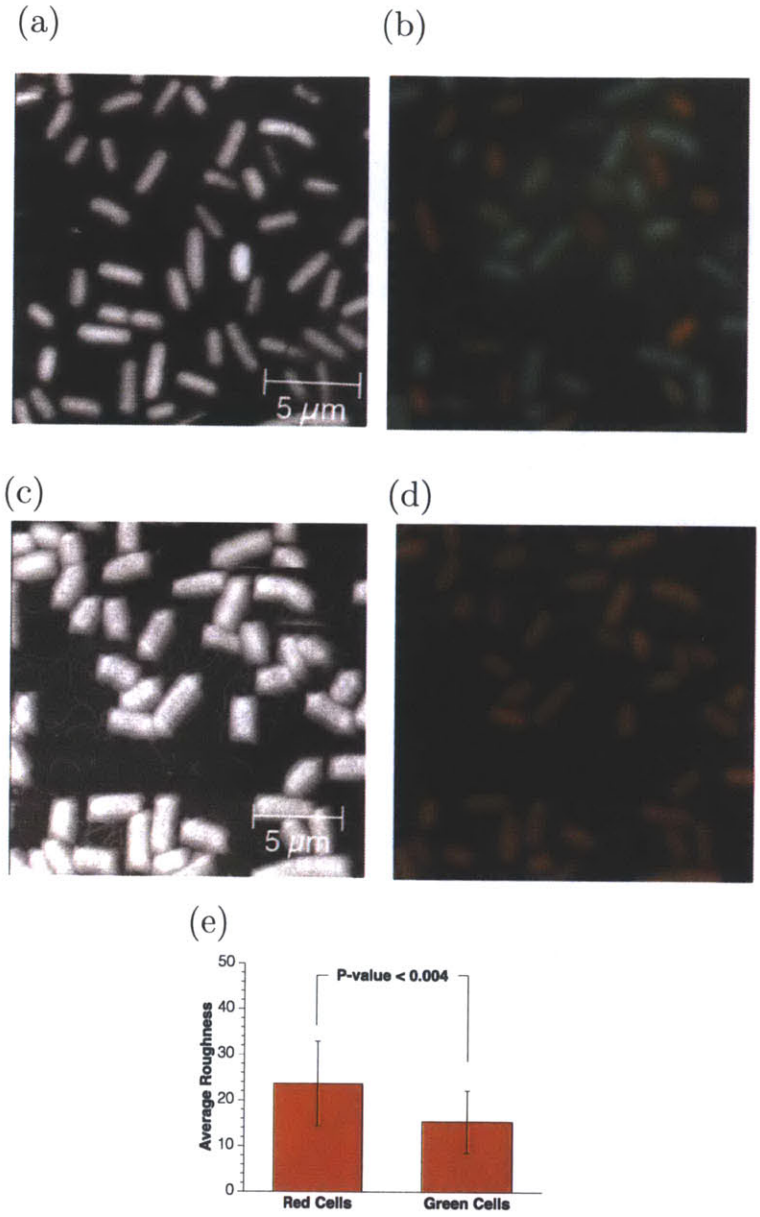


Figure 1.6: *AmP-induced surface morphology change correlates to cell death* Combined AFM and fluorescence microscopy images recorded on the same spot before and after addition of CM15. (a) Tapping mode image of bacteria before addition of CM15 (phase data). The surfaces of most bacteria were smooth. (b) Fluorescence image before addition of CM15. Green represents live bacteria; red represents dead bacteria (LIVE/DEAD stain). (c) AFM image 30 minutes after exposure to a solution of CM15 that was 2 times the MIC. Nearly all the bacteria exhibited a corrugated surface. (d) Fluorescence image after addition of CM15. All bacteria were red, indicating that they were dead. (e) Cells from (b) and (d) were divided into a "red" and "green" groups and the average roughness was calculated for each group. AFM images were taken with 512 x 256 pixels and a scan rate of 0.5 Hz. Reprinted with permission from *Nature Nanotechnology*.

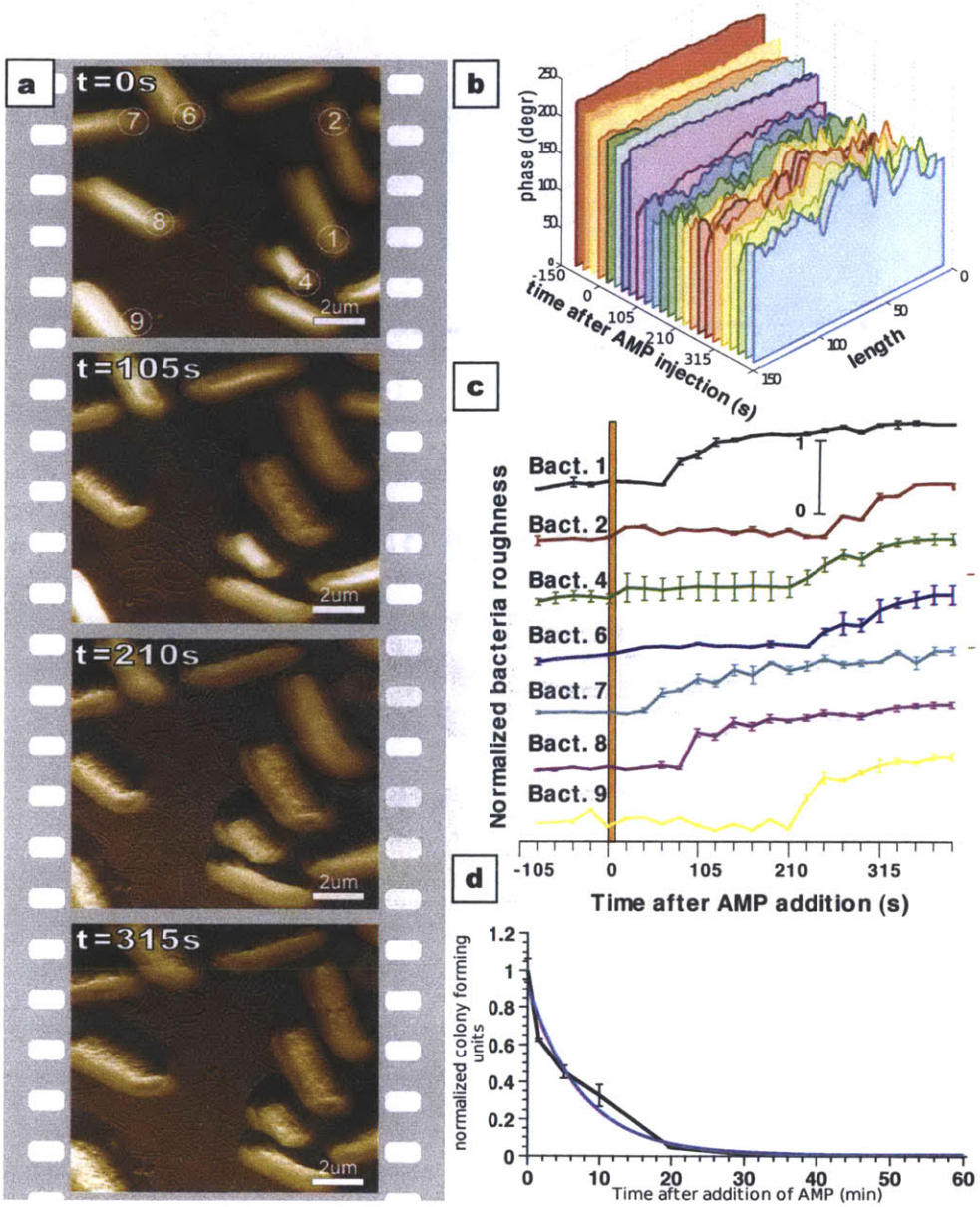


Figure 1.7: *Early stage kinetics of CM15 action measured by AFM correlates with bulk killing activity experiment* (a) Time series of bacteria after injection of CM15. Images recorded every 21 seconds (1024 x 256 pixels, 12.2 $\frac{lines}{second}$) with every 5^{th} image shown (full time series in Figure 1.8). (b) Cross sections along the long axis of bacterium 1 showing the time progression of the surface variation. Each slice represents data extracted from one image in the full time series. (c) Averaged surface variation of the bacteria as a function of time after injection of CM15 (bacteria numbers correspond to those in frame one of panel (a)). (d) Bulk measurement of CM15 antimicrobial activity. The interpolated behaviour between 0 and 5 minutes correlated well with the single cell measurements. Reprinted with permission from *Nature Nanotechnology*.

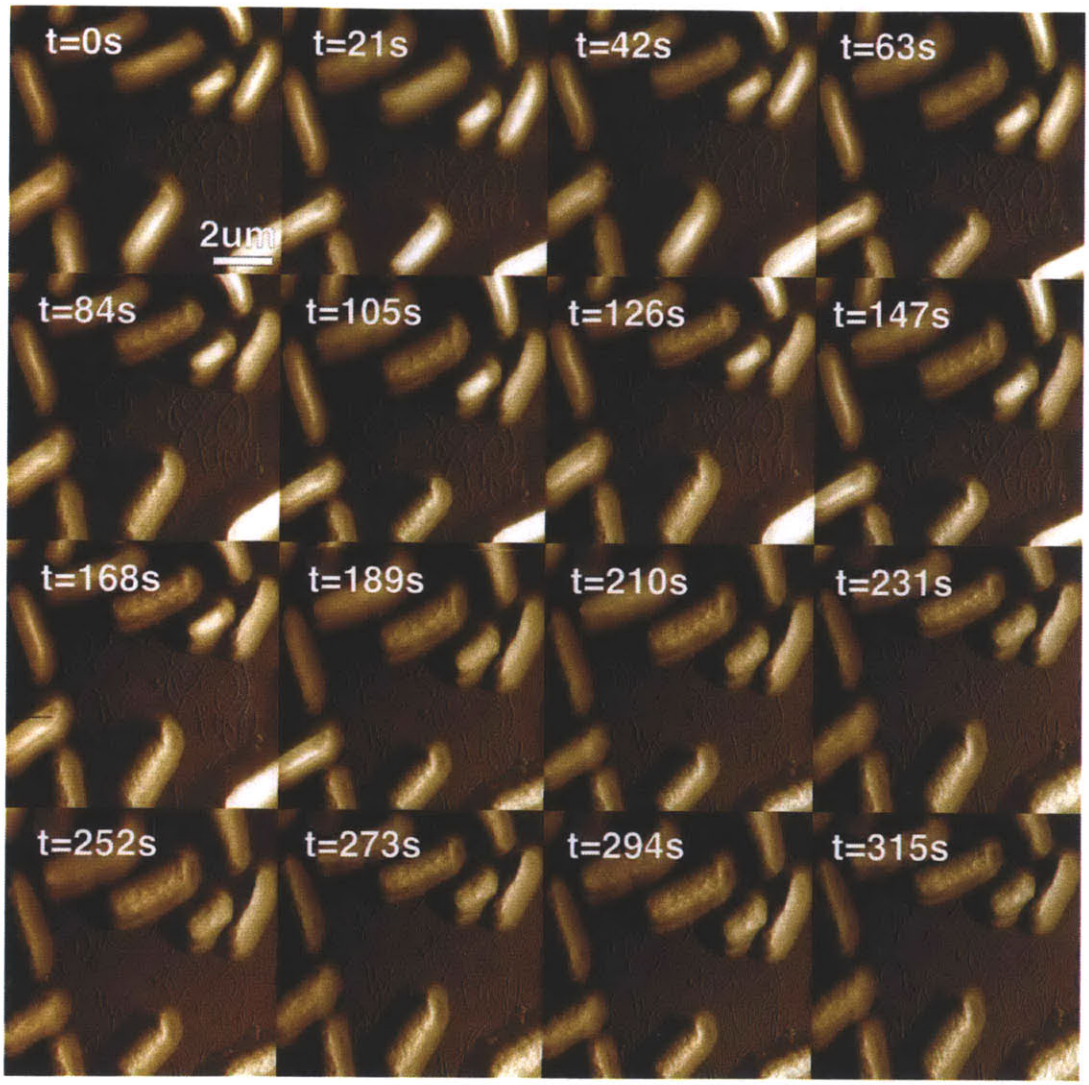


Figure 1.8: *Phase data of full time series from Figure 1.7* Rotated 90 degrees counter-clockwise from Figure 1.7(a). Reprinted with permission from *Nature Nanotechnology.*

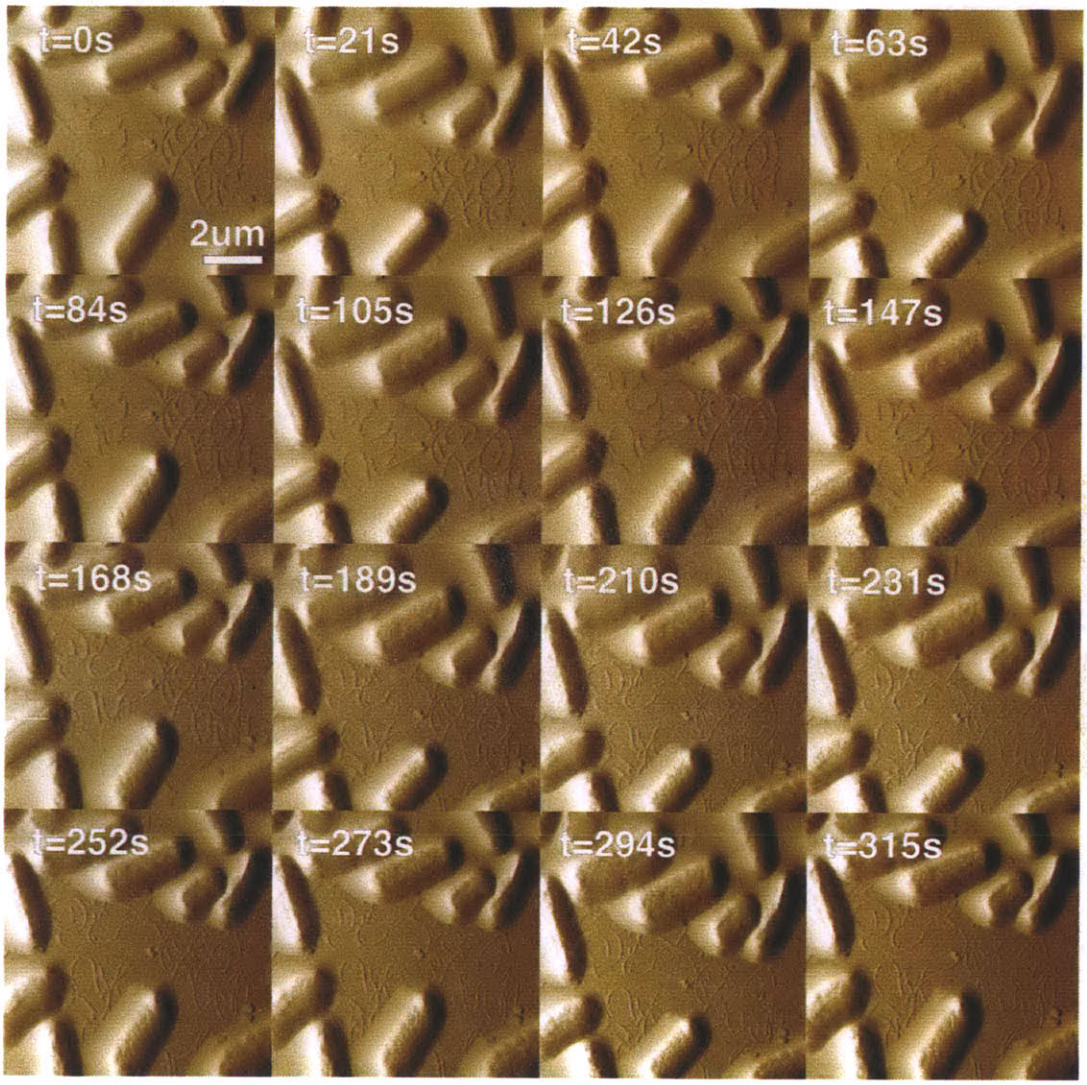


Figure 1.9: *Amplitude data of full time sequence from Figure 1.7* Rotated 90 degrees counter-clockwise from Figure 1.7(a). Reprinted with permission from *Nature Nanotechnology.*

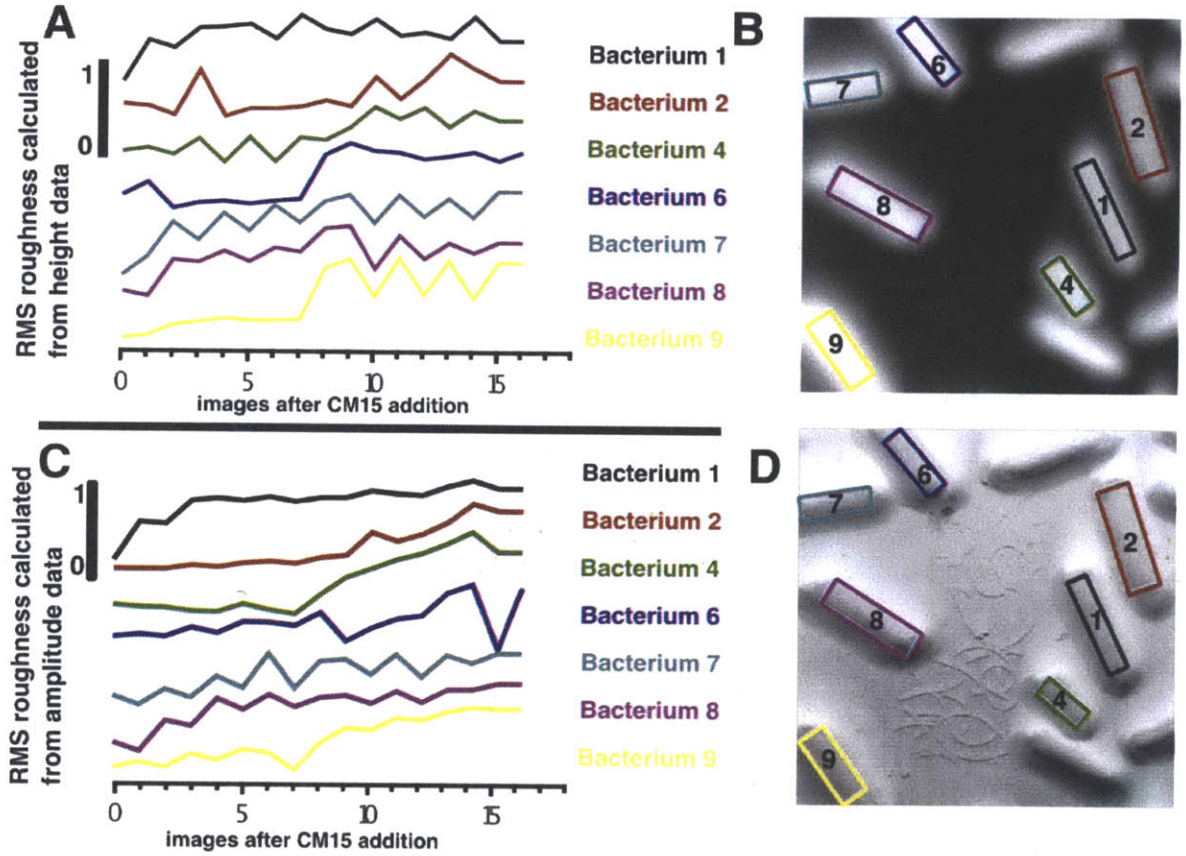


Figure 1.10: *Analysis of height and amplitude high-speed AFM images for CM15 killing of* E. coli RMS variation was calculated from height and amplitude data in the same manner as it was for the phase data. These data show the same trends as the values calculated from the phase images (Figure 1.7c), but the more pronounced features in the phase images result in better signal to noise ratio. (a) RMS variation calculated from height data as a function of images after addition of CM15 normalized to the maximum value. (b) Initial height image immediately after CM15 addition indicating the areas that were used for calculating the roughness. (c) Normalized RMS variation as calculated from the amplitude images. (d) Initial amplitude image immediately after addition of CM15. The absolute values of the final RMS roughness of the height data are 20–30 nm. The absolute values of the final RMS roughness of the amplitude data is 50–200 nm. The absolute values however varied between the individual cells which was likely a result of the different orientations of the cells with respect to the fast scan axis. Reprinted with permission from *Nature Nanotechnology.*

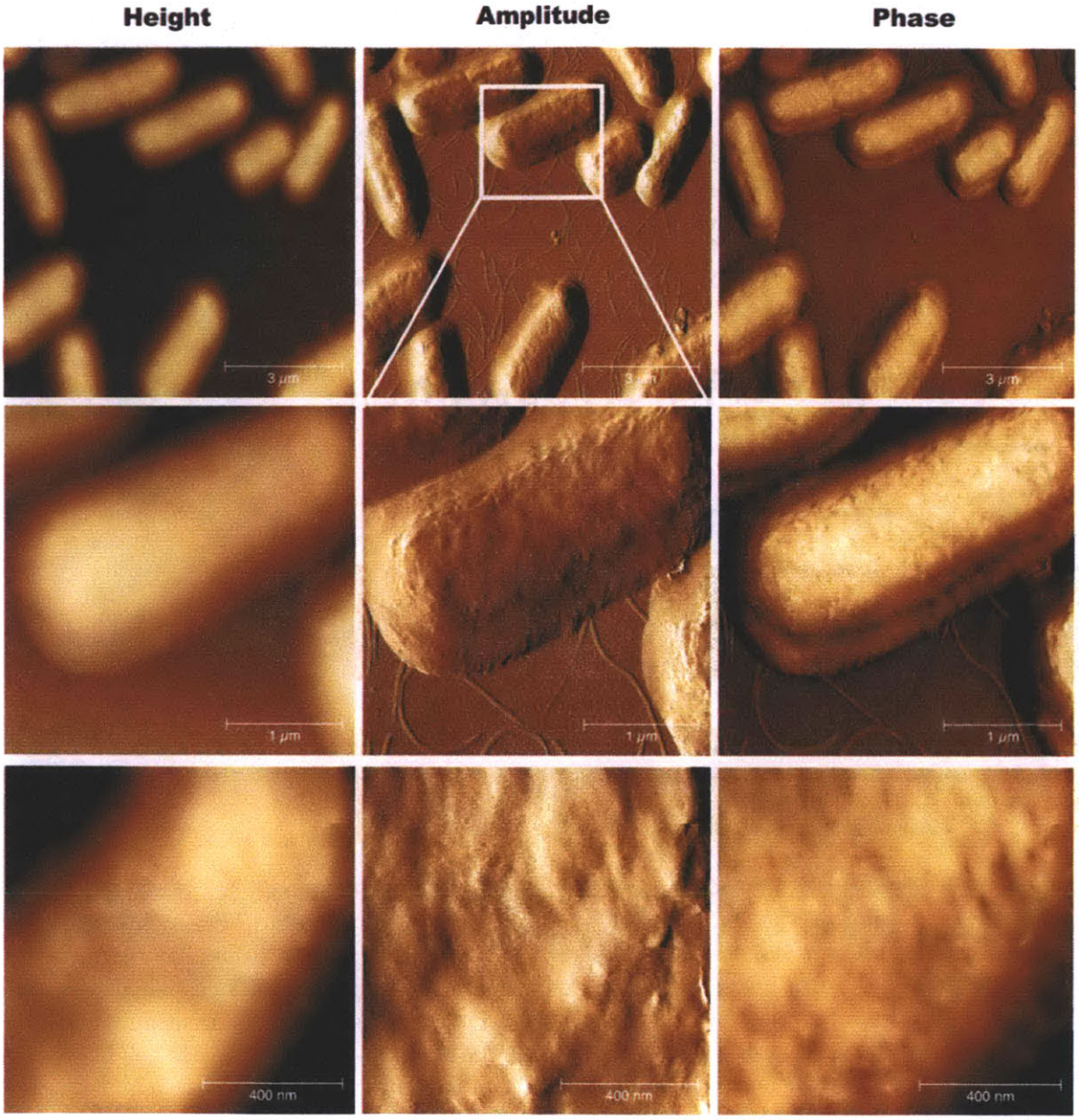


Figure 1.11: *High-resolution images of the CM15-treated bacteria* These images were captured after the high-speed AFM time series ended for Figure 1.7. The columns represent height, amplitude and phase data, respectively. The rows represent different magnifications. The small features have dimensions in the range of the reported CM15 pore sizes, but it cannot be unequivocally concluded that these features are in fact pores formed by the CM15. Reprinted with permission from *Nature Nanotechnology*.

1.4 Conclusions

These results led to the hypothesis that the killing of *E. coli* by CM15 is a two-stage process. The first stage is a time-variable, and slow to start, incubation phase, which can last seconds to minutes. The second stage is an execution phase, which is much less time-variable and which reaches completion more quickly. In the results shown in the work, for example, 50% of the damage was completed in less than one minute. The time to complete the incubation phase varies considerably more from cell to cell than the time to complete the execution phase. This result suggests that the bulk kill rate (e.g., the phenomenon measured in Figure 1.7d) is dominated by the time it takes to complete the incubation phase, rather than the time it takes to complete the execution phase. This also raises the question of whether the incubation and the execution phases are prolonged equally in bacteria that are more resistant to AmPs or if one particular phase is more prolonged than the other. Answering this could be important for understanding the mechanism by which bacteria can develop resistance to this class of peptide antibiotics.

The heterogeneity in the dynamics of AmP-induced bacterial cell death has an interesting parallel in eukaryotic cell biology. Mammalian cell apoptotic death recently has been demonstrated to be a two-stage process comprised of a time-variable and long-in-duration incubation phase which is followed by a more constant and relatively short execution phase. [33] Despite the apparent differences in the overall mechanisms between the two systems, protein/peptide-induced pore formation in a membrane is a critical component of both mammalian cell apoptotic death (mitochondrial outer membrane permeabilization) and AmP-induced bacterial cell death (outer membrane permeabilization).

The observation of this two-stage process was made possible by the use of a high-speed AFM capable of recording dynamic changes on the single cell level. These measurements demonstrate the enormous potential of high-speed AFM imaging for cellular biology. This technique can be applied to other cell types such as yeast or mammalian cells and even to eukaryotic cell organelles and future work with this instrument will surely be focused on those systems, as well as on additional characterization of AmP-induced killing of bacterial cells.

1.5 Experimental

1.5.1 Synthesis of the AmP CM15

A well-studied, pore-forming AmP called CM15 with the sequence KWKLFKKIGAVLKVL was used. The peptide was made by Genscript (Piscataway, NJ, USA) using standard solid phase peptide synthesis methods, was purified by desalting and was shipped lyophilized. The peptide was resuspended and stored in Millipore water and diluted to the appropriate concentrations in Millipore water.

1.5.2 Bacteria preparation

E. coli bacteria (ATCC 25922) were grown from a single colony overnight in LB growth medium at 37 degrees Celsius. Cells were diluted 1:100 in fresh LB medium and were grown

for 3 hours. Cells were spun down and washed three times with Millipore water.

To increase adherence of the bacteria to the glass substrate, glass cover slips were coated with poly-l-lysine before depositing the bacteria. Round glass cover slips were boiled in 2.5 M HCl solution for ten minutes to clean the glass. The cover slips were rinsed six times with Millipore water and were immersed for ten minutes in a pH 8.0 solution of 0.05 $\frac{mg}{mL}$ poly-L-lysine hydrobromide (Sigma part number: P1524) and 10 mM Tris. Afterwards, they were covered and dried vertically overnight at room temperature. The coated cover slips were stored at room temperature and were used within one week.

Concentrated bacteria suspended in un-buffered Millipore water were deposited on the coated cover slips and incubated for 30 minutes. Excess cells were rinsed off with 3 washes of 1 ml Millipore water. Images were taken in Millipore water.

1.5.3 High-speed AFM imaging

All high-speed AFM images were taken on a customized AFM instrument based on a Veeco Multimode with a Nanoscope 5 controller (Veeco Metrology, Santa Barbara, USA). The most enabling modifications are small cantilevers (SCL-Sensor.Tech. fabrication GesmbH, Vienna, Austria) and a new AFM head designed for use with the small cantilevers. A high-speed scanner was not used for these experiments, since turnaround ripples and other distortions were small at these scanning parameters. All measurements were performed in tapping mode in fluid with an open fluid cell. Flow-through fluid exchange was achieved using a dual syringe pump. Height, amplitude and phase signals were recorded for both trace and retrace. The data was processed using ImageJ, ImageSXM and Gwyddion, using standard modification commands applied over the whole sample.

1.5.4 Quantification of bacterial surface morphological changes

To quantify the change in the variations on the bacterial surface, line sections were extracted from the phase data on the top of each bacterium in the longitudinal direction. The exact number of line sections per bacterium per image depended on the size of the bacterium but was always more than 20. A 3^{rd} degree polynomial was fit through each line section and subtracted from the original data. The resulting data represented only the higher spatial frequencies from which the RMS value was calculated. The RMS values of each line were then averaged over all the lines along the bacterium. This number represents one data point in Figure 1.7. The error bars represent standard deviation of this calculation when varying the exact area on each bacterium and the number of longitudinal sections used for the calculation. This calculation was repeated for each frame in the time series and for each cell. The data processing was done with custom software written in Labview and NI-Vision (National Instruments, Austin, USA).

1.5.5 Combined AFM/fluorescence microscopy imaging

AFM images were recorded on a Digital Instruments Bioscope (Veeco Metrology Inc., Santa Barbara, USA) with an extended Nanoscope 3a controller, mounted on top of a Zeiss Axiovert 135 inverted microscope (Carl Zeiss Microimaging GmbH, Germany). AFM images were

taken with a Veeco NPS cantilever in tapping mode (Veeco Instruments Inc., Camarillo, USA). Fluorescence images were recorded using a Zeiss 100X Plan-Neofluar oil immersion objective and a Chroma Technology Corporation (Rockingham, USA) filter set 41012 . Cells were stained with LIVE/DEAD BacLight stain (Invitrogen PN: L13152). The LIVE/DEAD stain is a mixture of SYTO 9 green-fluorescent nucleic acid stain and propidium iodide red-fluorescent nucleic acid stain. These stains differ in their ability to penetrate healthy bacterial cells, with the green stain permeating healthy cells and the red stain permeating membrane-compromised cells.

Due to the loss in viability during the preparation process, some cells already are dead before the experiment starts (see Figure 1.6). During initial AFM imaging, some loosely connected cells detach from the surface and therefore not all the cells in Figure 1.6a are present in Figure 1.6b.

1.5.6 Analysis of LIVE/DEAD stain results and correlation with bacterial surface morphological changes

Fluorescence images were analyzed using ImageJ with nucleus counter plug-in (downloaded from MacBiophotonics). Cells were identified and numbered using the nucleus counter plug-in. The color images were split into red, green, and blue channels. For each cell, the mean red and the mean green values were measured and the green to red ratio for each cell was calculated. The cells were divided into two populations – those with green to red ratios less than one (dead) and those with green to red ratios greater than one (live). Using the axial RMS variation for each cell (calculation described below), a one-tailed students t-test was used to assess whether the live cell population and the dead cell population had statistically different RMS variation values.

1.5.7 Bulk cell killing assay

Cells were grown overnight in LB from a single colony. In the morning, they were diluted 1:100 in fresh LB and grown for three hours. After three hours, they were washed three times in Millipore water. The washed cells were resuspended in Millipore water to a concentration of $2\text{x}10^6$ $\frac{cells}{mL}$. The CM15 peptide was added to the cells to a final concentration of 20 $\frac{\mu g}{mL}$, which was the same peptide-to-cell ratio as that used in the AFM experiments. Aliquots were taken from the cells at various time points and were plated in triplicate on fresh, prewarmed LB agar plates, which were incubated overnight at 37°C. The number of colonies formed is an indicator of the number of viable cells remaining in the suspension at each time point. The number of colonies formed was counted and the average and standard deviation of each time point was calculated. All data was normalized to the zero time point (before addition of peptide).

Chapter 2

Functionalization of oxide surfaces with an engineered peptide

2.1 Summary

Peptide affinity tags are useful for facile surface functionalization of materials with proteins and other biomolecules. By directly interacting with the surfaces, peptide affinity tags obviate the need for multistep chemical surface functionalization. Affinity peptide tags are often also reversible, which allows for easier surface cleaning than covalent chemical functionalizations require, while still retaining high binding strength. In this work, applications for a novel peptide affinity tag (2K1) with high affinity for metal oxide surfaces were explored. The versatility of this peptide affinity tag was demonstrated by functionalization of a diverse set of relevant materials, including titanium dioxide, zinc, and stainless steel. The 2K1 peptide was used as an affinity tag for a small molecule (biotin) and for two larger fusion proteins (basic fibroblast growth factor and protein A). Finally, a 2K1-CM15 peptide was made in an attempt to develop a single-step, facile antimicrobial functionalization of oxide surfaces.

2.2 Introduction

The development of peptides with high affinities for medically and industrially relevant materials is of growing interest, especially as the field of synthetic biology promises to continue to bring down the costs of DNA and protein synthesis. [34, 35] Despite the advantages that peptide affinity tags offer (single-step, reversible fuctionalization in aqueous chemistry) there are still a limited number of peptide tags with high affinities toward relevant materials. One such set of materials with limited options for peptide-driven fuctionalization is metal oxides, which span the application areas from stainless steel industrial reactors to dental implants made of aluminum oxide.

In this work, the versatility and utility of a high-affinity sapphire-binding peptide (2K1) was explored. 2K1 had been shown previously to enable the immobilization of an MBP-fusion protein to sapphire, SiO_2, and tissue culture treated polystyrene (TCT-PS). [36] The range of substrates has now been expanded to three additional ubiquitous metal oxide

surfaces. A biotin-2K1 fusion was made and used to demonstrate the immobilization of a streptavidin-conjugated horseradish peroxidase (HRP) enzyme to TCT-PS, TiO_2, zinc metal, and stainless steel.

Additionally, 2K1's utility as affinity tag in two different recombinantly-produced fusion proteins was demonstrated, each with relevant biotechnology applications. The first was protein A (protA), which binds to the Fc region of IgG molecules, thus making it a useful alternative to biotin-streptavidin interactions for capturing or immobilizing antibodies. [37] The second was basic fibroblast growth factor (bFGF), which binds heparin. [38] Heparin acts as an anticoagulant and is given to patients during several types of medical treatments, including heart surgeries and hemofiltration. It is critical to monitor heparin levels in the blood, and one can imagine the utility of functionalizing a highly sensitive detector (such as the suspended microchannel resonator pioneered by the Manalis group at MIT) with a molecule that can recognize and bind heparin. [39]

Finally, fusions of 2K1 and the antimicrobial peptide investigated in Chapter 1, CM15, were made in an attempt to make antimicrobial surfaces. CM15 was chosen as the peptide of interest for these fusions because it is relatively short, thus keeping the synthesis costs within reason.

2.3 Results

2.3.1 Expression and purification of 2K1 fusion proteins

The 2K1 peptide was expressed on the C-terminus of two different MBP fusion proteins, one with the IgG-binding fragment of protA on the N-terminus and one with the heparin-binding fragment of bFGF. The fusion proteins were then expressed and purified in *E. coli* cells as described in Section 2.5.3. Figure 2.1 shows Coomasie stained gels with the different fractions from the purification. Since the expression vector that was used has a His6 tag, a Ni-NTA column was used to purify the fusion proteins (with an imidazole elution). The 2K1-MBP-protA and 2K1-MBP-bFGF fusion proteins have estimated molecular weights of $\approx$ 63 and 65 kDa, respectively. Proteins of these sizes would be expected to migrate between the second and third bands in the protein ladder (80 kDa and 58 kDa), which was confirmed by visual inspection of the gels.

To confirm that the bands in the Coomasie-stained gels were the fusion proteins, the two or three fractions that appeared the most pure were analyzed with Western blots using Anti-MBP and with PentaHis antibodies. Images of those gels can be seen in Figure 2.2. This analysis confirmed that the MBP and His6 portions of the fusion proteins were present. For further analysis, the highest concentration fractions at the highest purity were used. For 2K1-MBP-protA, the 200 mM imidazole fraction was used at an estimated concentration of $\approx$ 44.5 μM. For 2K1-MBP-bFGF, the 70 mM imidazole fraction was used at an estimated concentration of $\approx$ 5.8 μM (see Section 2.5.8 for details on calculating concentrations).

surfaces. A biotin-2K1 fusion was made and used to demonstrate the immobilization of a streptavidin-conjugated horseradish peroxidase (HRP) enzyme to TCT-PS, TiO_2, zinc metal, and stainless steel.

Additionally, 2K1's utility as affinity tag in two different recombinantly-produced fusion proteins was demonstrated, each with relevant biotechnology applications. The first was protein A (protA), which binds to the Fc region of IgG molecules, thus making it a useful alternative to biotin-streptavidin interactions for capturing or immobilizing antibodies. [37] The second was basic fibroblast growth factor (bFGF), which binds heparin. [38] Heparin acts as an anticoagulant and is given to patients during several types of medical treatments, including heart surgeries and hemofiltration. It is critical to monitor heparin levels in the blood, and one can imagine the utility of functionalizing a highly sensitive detector (such as the suspended microchannel resonator pioneered by the Manalis group at MIT) with a molecule that can recognize and bind heparin. [39]

Finally, fusions of 2K1 and the antimicrobial peptide investigated in Chapter 1, CM15, were made in an attempt to make antimicrobial surfaces. CM15 was chosen as the peptide of interest for these fusions because it is relatively short, thus keeping the synthesis costs within reason.

2.3 Results

2.3.1 Expression and purification of 2K1 fusion proteins

The 2K1 peptide was expressed on the C-terminus of two different MBP fusion proteins, one with the IgG-binding fragment of protA on the N-terminus and one with the heparin-binding fragment of bFGF. The fusion proteins were then expressed and purified in *E. coli* cells as described in Section 2.5.3. Figure 2.1 shows Coomasie stained gels with the different fractions from the purification. Since the expression vector that was used has a His6 tag, a Ni-NTA column was used to purify the fusion proteins (with an imidazole elution). The 2K1-MBP-protA and 2K1-MBP-bFGF fusion proteins have estimated molecular weights of $\approx$ 63 and 65 kDa, respectively. Proteins of these sizes would be expected to migrate between the second and third bands in the protein ladder (80 kDa and 58 kDa), which was confirmed by visual inspection of the gels.

To confirm that the bands in the Coomasie-stained gels were the fusion proteins, the two or three fractions that appeared the most pure were analyzed with Western blots using Anti-MBP and with PentaHis antibodies. Images of those gels can be seen in Figure 2.2. This analysis confirmed that the MBP and His6 portions of the fusion proteins were present. For further analysis, the highest concentration fractions at the highest purity were used. For 2K1-MBP-protA, the 200 mM imidazole fraction was used at an estimated concentration of $\approx$ 44.5 μM. For 2K1-MBP-bFGF, the 70 mM imidazole fraction was used at an estimated concentration of $\approx$ 5.8 μM (see Section 2.5.8 for details on calculating concentrations).

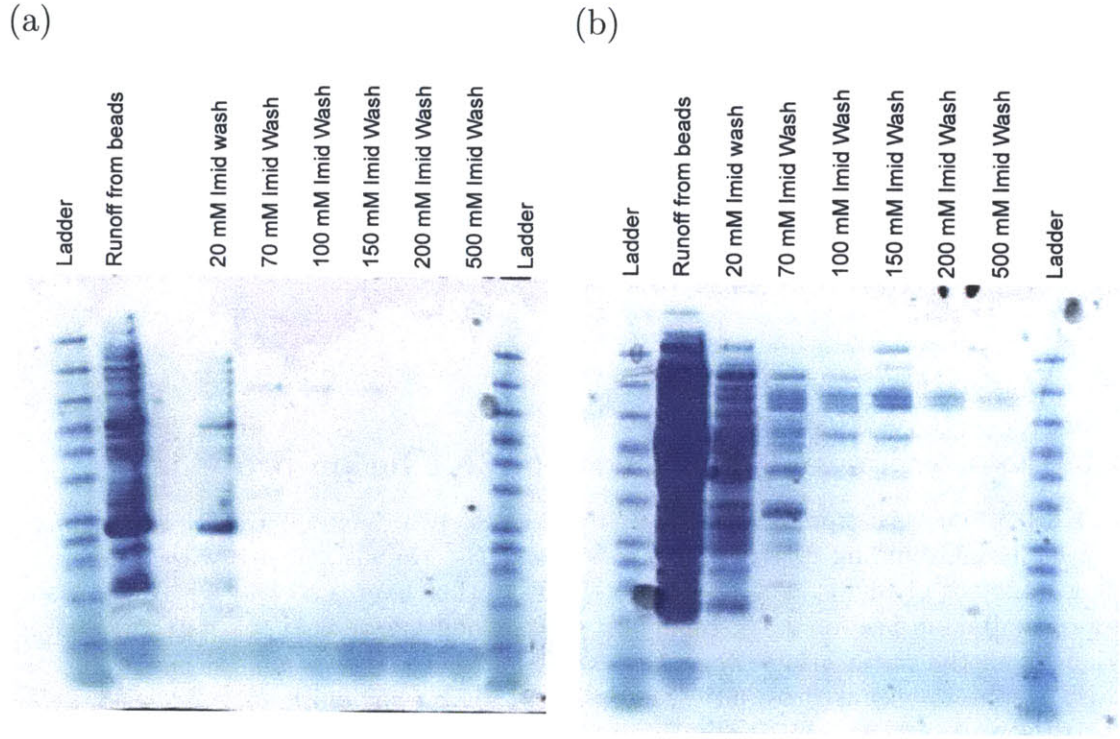


Figure 2.1: *SDS-PAGE of 2K1-MBP-bFGF-His6 fractions and 2K1-MBP-protA-His6 fractions* The gel on the left is a Coomasie stained gel of the 2K1-MBP-bFGF-His6 cell lysate with the different imidizole elution fractions listed from left to right. The gel on the right is a Coomasie-stained gel of the 2K1-MBP-protA-His6 cell lysate with the different imidizole elution fractions listed from left to right. The protein ladder is NEB Prestained Protein Marker, Broad Range (7-175 kDa).

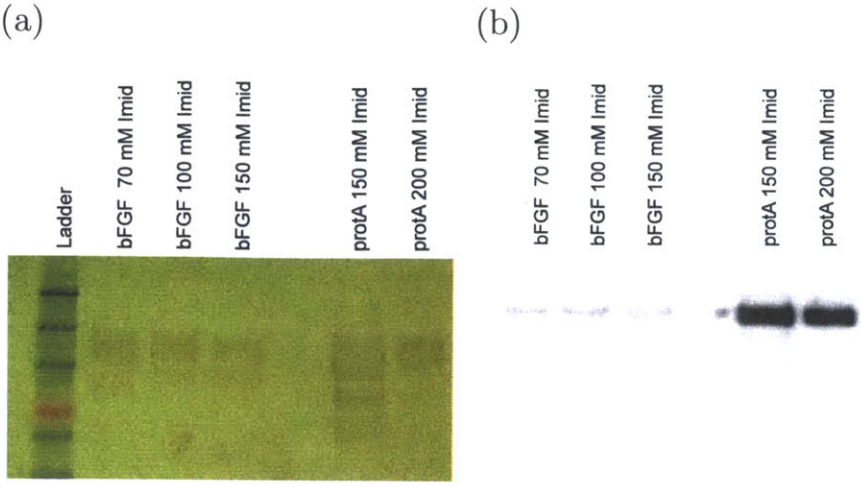


Figure 2.2: *Anti-MBP western blot of 2k1-MBP-bFGF-His6 fractions and PentaHis western blot of 2K1-MBP-bFGF-His6 fractions* The image on the left is a western blot using an anti-MBP antibod. The image on the right is a western blot using a PentaHis antibody. Both western blots confirm the presence of the 2K1-MBP fusion proteins in the fractions. The protein ladder is NEB Prestained Protein Marker, Broad Range (7-175 kDa).

2.3.2 2K1 peptide can be used as an affinity tag in fusion proteins

Once the expression and purification of 2K1 fusion proteins in an *E. coli* host was confirmed, the ability of the 2K1 peptide to immobilize and properly orient the fusion proteins on TCT-PS was tested. 2K1 had already been demonstrated to have a high affinity for the negatively charged oxide surface of plasma treated polystyrene, and this is an easy surface on which to perform binding assays. [36] Figures 2.3 and 2.4 show the results of these binding assays, as well as cartoons describing the layout of the binding assays. Both of the fusion proteins were able to capture and immobilize their target molecules (IgG and heparin) on the TCT-PS substrate with high affinities ($K_D << 100$ nM). The K_D's for these fusion proteins were not calculated, because of the variability introduced by the fusion protein concentration and purity calculations. A visual inspection of the data, however, suggests that the K_D is between 10 nM and 30 nM.

The His6 tag that is present on the 2K1 fusion proteins was used only as a purification tag, and it did not have a detectable effect on the immobilization of the fusion proteins to TCT-PS. If, however, these proteins were to be used in future applications, the expression vector encodes a thrombin cut site between the His6 tag and the rest of the fusion protein, which could be used to remove the His6 tag.

2.3.3 2K1 enables binding to a variety of oxide surfaces

Since 2K1 had already been demonstrated to have a high affinity to more than one type of oxide surface, including SiO_2, Al_2O_3, and TCT-PS, the promiscuity of the peptide's binding was tested using several additional materials. A panel of four ubiquitous metals were tested (iron, zinc, TiO_2, and stainless steel) for 2K1-MBP-protA binding. 2K1 was shown to enable

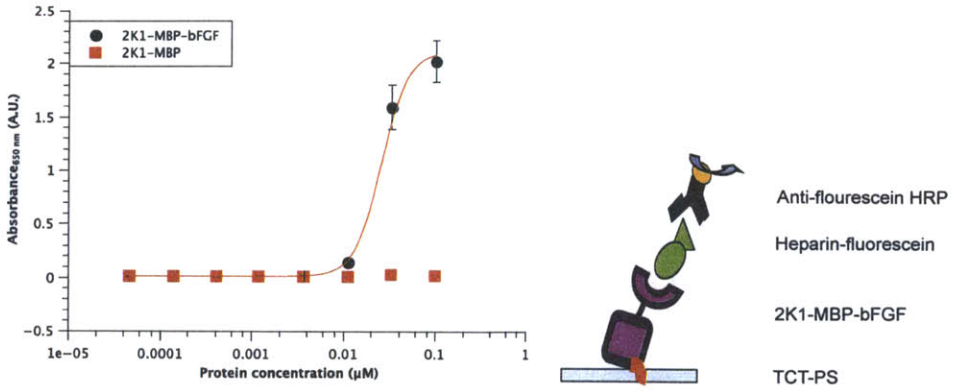


Figure 2.3: *Binding of 2K1-MBP-bFGF fusion protein to TCT-PS* To test the ability of a 2K1 affinity tag to immobilize and properly orient a bFGF moiety to an oxide surface, a modified ELISA was set-up. The 2K1 fusion protein was immobilized on the bottom of TCT-PS wells in a 96-well plate. The bFGF moiety was used to capture a fluorescein heparin conjugate, and the capture of this target was verified using an anti-fluorescein-HRP conjugate. After 5 minutes of exposure to TMB, the signal was read in a plate reader at 650 nm and the endpoint absorbance values are plotted against the concetrations of 2K1 fusion protein used.

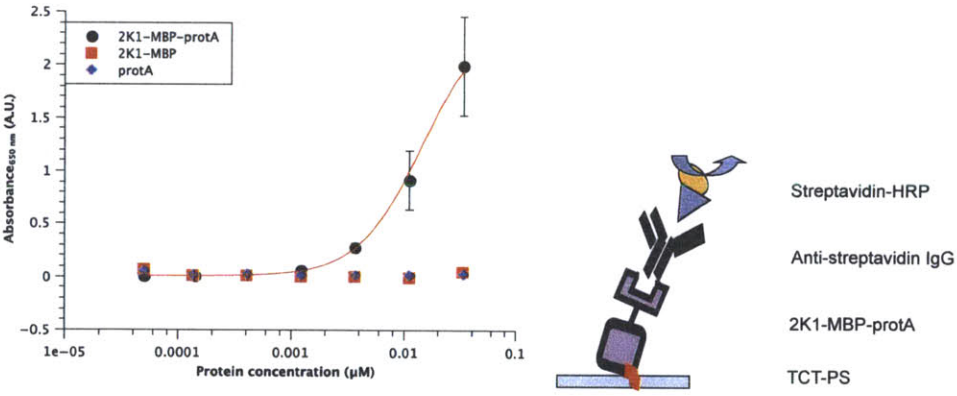


Figure 2.4: *Binding of 2K1-MBP-protA fusion protein to TCT-PS* To test the ability of a 2K1 affinity tag to immobilize and properly orient a protA moiety to an oxide surface, a modified ELISA was set-up. The 2K1 fusion protein was immobilized on the bottom of TCT-PS wells in a 96-well plate. The protA moiety was used to capture a rabbit anti-streptavidin IgG, and the capture of this target was verified using a streptavidin-HRP conjugate. After 5 minutes of exposure to TMB, the signal was read in a plate reader at 650 nm and the endpoint absorbance values are plotted against the concentrations of 2K1 fusion protein used.

binding of the protA fusion protein to TiO_2, stainless steel, and zinc (Figure 2.5). There was no detectable binding to the iron bar, though it should be noted that the iron bar appeared to be reacting with the binding solution and bits of iron were flaking off into the solution, which would have prevented the formation of a peptide layer on the surface of the bar. The results of this binding assay were not quantitative, because each of the materials used was a different size with different surface areas. Rather, this was meant to provide a qualitative assessment of whether 2K1 would bind to these materials. Further characterization of the interaction between 2K1 and stainless steel was performed and is discussed in Section 2.3.4.

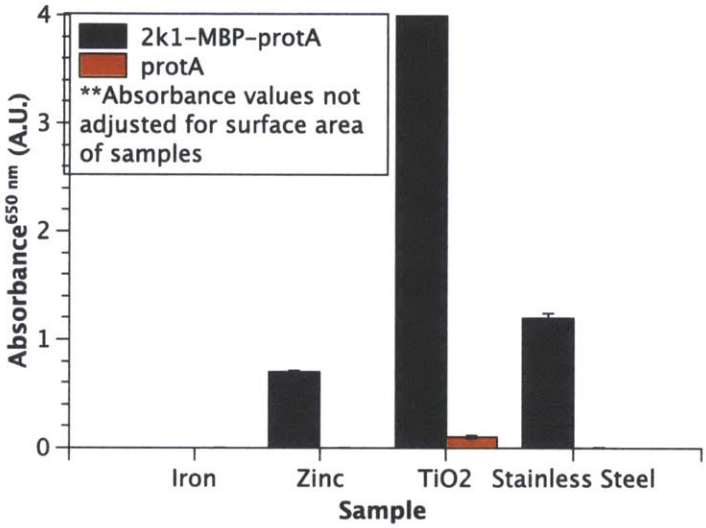


Figure 2.5: *Binding assay for 2K1 to other metals* Since 2K1 had already been shown to bind to SiO_2, sapphire, and TCT-PS, additional oxide surfaces were tested. As expected, 2K1 also showed a binding affinity for zinc, TiO_2, and stainless steel. Each of these materials forms an oxide layer on the surface in aqueous solutions, and it is to this oxide layer that 2K1 is likely binding. The lack of binding to zinc was likely due to the reaction of the zinc metal with the aqueous binding solution and the constant sluffing off of the outer layers during binding.

2.3.4 2K1-biotin enables single-step biotinylation of stainless steel

Adhesion of microorganisms to material surfaces occurs in a variety of contexts, often with deleterious effects. For example, biofilms form in industrial processes where they clog pipes or enhance corrosion. [40, 41] In the food processing industry, biofilms are associated with contamination of surfaces and transmission of infectious diseases, while biofilms on medical devices can provide safe harbor for antibiotic resistant pathogenic bacteria. [42, 43].

Stainless steel is a steel alloy that does not corrode or rust as easily as normal steel, due to the presence of $\approx$ 11% chromium by mass. It is the chromium that protects the bulk of the material through the formation of a thin oxide layer on the surface. There are

over one hundred grades of stainless steel and it is available in a plethora of forms: sheets, bars, wires, pipes, etc. Due to this availability and the inherent advantages over normal steel, it is used widely, including in the medical device and food processing industries. Thus, a facile means of functionalizing stainless steel is of obvious value. Several groups have reported on chemical approaches or adsorption to stainless steel functionalization to reduce biofouling. [44–46] Others have used polyelectrolyte layers to confer antimicrobial properties to stainless steel for medical implants. [47]

After demonstrating that the 2K1 peptide could bind to stainless steel, the next step was to assess the binding affinity. A 2K1 peptide with biotin attached to the C-terminus was synthesized using solid phase chemical synthesis (Genscript, Piscataway, NJ). The binding affinity of 2K1-biotin was measured as described in Section 2.5.7 and the results can be seen in Figure 2.6 and the relative binding affinity of 2K1 to a variety of oxide surfaces is summarized in Table 2.1.

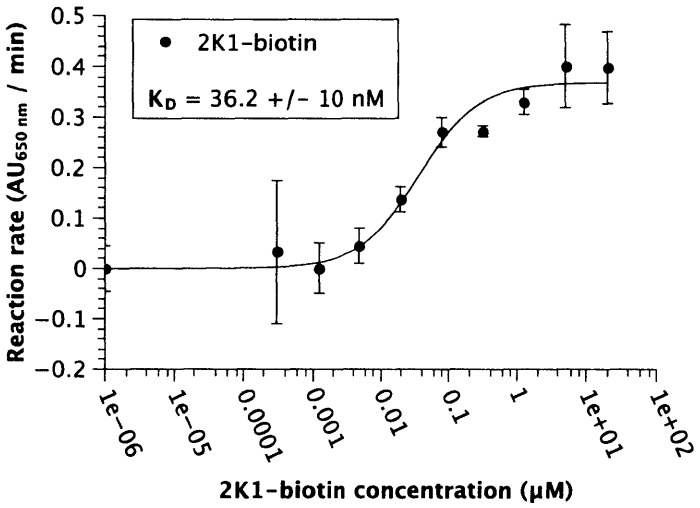


Figure 2.6: *Measurement of binding affinity of 2K1 to stainless steel* Stainless steel forms a chromium oxide layer on the surface in aqueous solutions. 2K1 has a K_D of $\approx$ 36 nM to stainless steel, which is the highest affinity for a small peptide to a stainless steel that has been reported.

The measured binding affinity of 2K1 to stainless steel is 36 nM $\pm$ 10 nM. Interestingly, this affinity is very comparable to one of the most widely used affinity tags available: His6 has an affinity for Ni-NTA plates of 30 $\pm$ 3 nM when measured using a similar assay. [36] Overall, it appears that 2K1 is a versatile affinity peptide for the functionalization of oxide materials. It can be used as an affinity tag in fusion proteins, or, if it has a biotin attached to it, it can be used for facile biotinylation of oxide surfaces. The binding affinities are in the low to mid nM ranges for Al_2O_3, TCT-PS and stainless steel. These results lead to the expectation that other highly negatively charged oxide or phosphate surfaces would also have K_D's << 100 nM, though obviously the binding to any new surface would need to be characterized.

Material	Binding Strength
A-face sapphire (Al_2O_3)	+++ (K_D = 1.3 ± 0.3 nM) [36]
Tissue culture treated polystyrene (TCT-PS)	+++ (K_D = 1.5 ± 0.1 nM) [36]
Z-cut quartz (single crystal SiO_2)	+++[a]
TiO_2	+++
Stainless steel	++ (K_D = 36 ± 10 nM)
Thermally grown 100 nm SiO_2 layer on Si wafer	++[a]
Zinc	+
Standard borosilicate glass	+[a]
Iron	-
His6-tag on Qiagen Ni-NTA HisSorb plate	++ (K_D = 30 ± 3 nM)[a]

Table 2.1: *2K1 surface binding versatility* The material binding promiscuity of 2K1 is summarized. 2K1 binds with low nM affinity to sapphire and TCT-PS and mid nM affinity to stainless steel. The affinities to SiO_2 and TiO_2 were not calculated, but 2K1 showed strong binding to each in a qualitative assay. The lack of binding to zinc was likely due to the reaction of the zinc metal with the aqueous binding solution and the constant sluffing off of the outer layers during binding.

2.3.5 Functionalization of TCT-PS with antimicrobial peptide using 2K1

Finally, with the goal of establishing a protocol for single-step and facile antimicrobial functionalization of oxide surfaces, a fusion peptide of 2K1 and the AmP CM15 was made. The fusion peptide was synthesized using solid phase chemical synthesis (Genscript, Piscataway, NJ). To test the ability of this fusion peptide to make antimicrobial surfaces, peptides were incubated with TCT-PS at four concentrations (3.1–200 μM) for one hour. To confirm that the concentrations of peptides were sufficient to achieve full surface functionalization, the K_D's were measured for CM15 (39 ± 15 nM) and 2K1 (3.8 ± 1.3 nM) to TCT-PS (see Figure 2.7. These results confirmed that even the lowest concentration used in the antimicrobial assay is still at least two orders of magnitude higher than the K_D's of both peptides. For the subsequent assays, four peptide conditions were tested: no peptide on surface (i.e., bare TCT-PS), 2K1 on surface, 2K1-CM15 fusion peptide on surface, and CM15 on surface.

There were two assays of antimicrobial activity that were performed with the functionalized TCT-PS. The first assay was a test of the ability of the funtionalized surfaces to inhibit growth of *E. coli* cells in a suspension culture. The concentration of mid-log phase growing *E. coli* cells was monitored over a four hour period. At the end of the time period the fold-growth for each concentration and type of peptide was calculated.

The results of this assay are shown in Figure 2.8a. There was no discernible difference between the wells with no peptide and the 2K1-biotin peptide, and they all had between 2- and 2.5-fold increases in *E. coli* concentration. The 2K1-CM15 functionalized surfaces appeared to keep the concentration of the cells at around the starting concentration for all four concentrations of 2K1-CM15. The results showed that the CM15 alone was able to inhibit growth as well as the 2K1-CM15 at the lower concentrations and performed even better than the 2K1-CM15 at the higher peptide concentrations. At 200 μM peptide, the

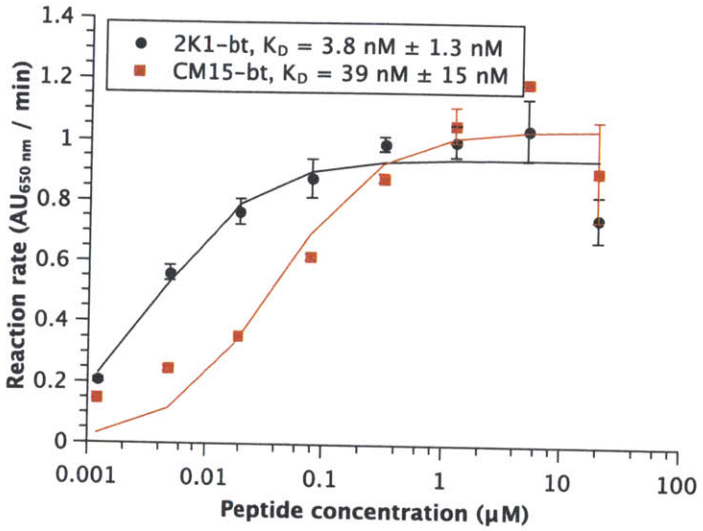
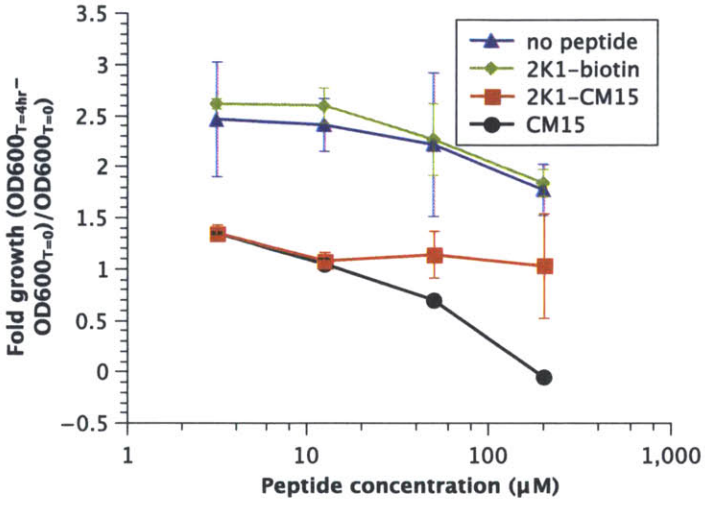


Figure 2.7: *Binding affinity of 2K1 and CM15 on TCT-PS* To confirm that the concentrations of peptides used in the antimicrobial assays would provide sufficient surface functionalization, the binding affinity of 2K1 and CM15 to TCT-PS was measured.

CM15 alone actually reduced the concentration of cells to below the detection limit of the assay, whereas the 200 μM 2K1-CM15 only succeeded in holding the concentration of cells constant. This difference is likely due to the fact that the CM15 has a higher K_D than the 2K1-CM15. A higher K_D would lead to faster off-rates and over a 4 hour period some of the peptide might have leached off the surface and into the solution, where it killed the cells.

The second assay was a measurement of what fraction of the cells that had adhered to the TCT-PS surface were alive. These cells are the likely progenitors of biofilms; they are the first colonizers that lead to downstream economic damages and health complications on medical devices, in industrial processes and in food processing plants. [40–43] If a surface can be functionalized such that fewer bacteria can colonize it, then these effects could be minimized. To measure this, the peptide-treated TCT-PS wells from the growth inhibition experiment were washed thoroughly to remove any unbound cells. The remaining cells were stained with a LIVE/DEAD stain. A fluorescence microscope was used to capture images of the stained bacteria. The images were processed as described in Section 2.5.9 and a green:red (LIVE:DEAD) ratio was calculated for each well. Multiple images per well and replicate wells were all averaged together and the results are presented in Figure 2.8b. Only the highest concentration of CM15 had a higher "dead" signal than "live" signal, and the 2K1-CM15 did not perform any better than the CM15 alone.

(a)



(b)

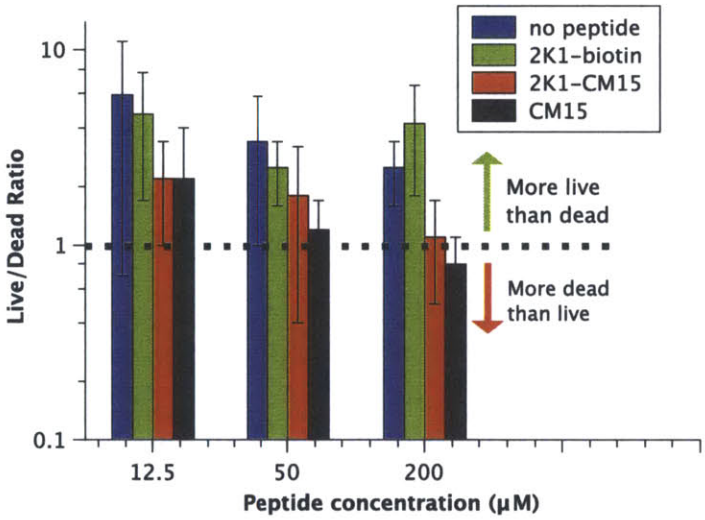


Figure 2.8: *Using 2K1 to functionalize TCT-PS with an antimicrobial peptide* (a) Peptide-functionalized TCT-PS was tested for its ability to inhibit growth of a suspension of mid-log phase *E. coli* cells over a four hour period in LB at 37°C. At lower peptide concentrations, the 2K1-CM15 fusion protein performed comparably to CM15 only, but at higher concentrations, the CM15 had a more dramatic impact on the *E. coli* cells. (b) For the three highest concentrations of peptide the wells from the assay in (a) were washed to remove unbound cells and then stained with a LIVE/DEAD stain to determine viability of adherent bacterial cells. For all three peptide concentrations, there were more live cells for the 2K1-CM15 surfaces than for the CM15 only surfaces and only at the highest concentration of peptide did CM15 wells have more dead than live cells.

2.4 Conclusions

The versatility and utility of an oxide surface binding peptide (2K1) was demonstrated. 2K1 was shown to bind to a wide range of materials, including Al_2O_3, SiO_2, TCT-PS, TiO_2, stainless steel, and zinc. [36] 2K1 can be used as an affinity tag in fusion proteins and it was shown to properly orient and immobilize two ubiquitous biomolecules (bFGF and protA). Alternatively, a biotinylated version of the peptide alone can be used to easily functionalize oxide surfaces, which opens the door to the biofuncationalization of a wide array of industrially relevant materials. The utility of 2K1-AmP fusion peptides in making antimicrobial TCT-PS surfaces and inhibiting bacterial colonization was investigated. Due to the fact that the AmP that was used in these assays already has an affinity for negatively-charged TCT-PS and because the mechanism of AmP action requires that the peptides be in solution in order to form pores on the cells surfaces, there were no perceived advantages in using 2K1-AmP fusions.

2.5 Experimental

2.5.1 Materials

Type 304 stainless steel and titanium (IV) oxide were purchased from AlfaAesar. Stainless Steel foil was purchased as 0.1 mm (0.004 in) thick, Type 304 100 x 100 mm, Fe:Cr:Ni 70:19:11 wt% (stock number: 41583) and was then cut into 0.14 inch squares (for the initial binding experiments) or punched into 3/16 inch circles using a Precision Brand Punch and Die Set (for the K_D and k_{off} measurements). Titanium dioxide was purchased from AlfaAesar as naturally occurring mineral grains, approximately 0.06-0.19 inches (stock number: 42648). Iron was used as 5 mm x 3 mm x 3 mm bars and zinc was used as 2.5 mm x 6 mm x 1 mm bars.

2.5.2 2K1 fusion proteins cloning and analysis

To build 2K1-MBP-protA and 2K1-MBP-bFGF expression constructs, 2K1-MBP-protA and 2K1-MBP-bFGF were cut from pMAL-c2x plasmid using NdeI and HindIII restriction enzymes. The purified insert DNA was then ligated into the pET-28A plasmid using NdeI and HindIII restriction sites such that the fusion protein was downstream of the His-Tag in the pET-28a plasmid (for example: $^{N-terminus}$ His6-bFGF-MBP-2K1$^{C-terminus}$). Plasmids were transformed into BL21(DE3) cells and plated on kanamycin plates. Proper sequences were confirmed through DNA sequencing of purified plasmid DNA from single colonies. Initial expression of fusion proteins was determined by growing cells at 37°C picked from a single colony for 4 hours in 25 mL of LB plus kanamycin. IPTG then was added to a final concentration of 1 mM to initiate induction of the fusion proteins and cells were grown for another 2.5 hours at 37°C. 200 μL aliquots were added to 140 μL of 1x PBS / 1 M NaCl and were frozen at -20°C overnight. These cell lysates (or more dilute version if still too thick too load on a gel) were then run directly on SDS-PAGE gels with LDS sample buffer and reducing agent. Expression was confirmed with coomasie staining and western blotting. Glycerol

stocks of cells with expression plasmids were made by adding 100 μL of 75% glycerol to 900 μL of exponential phase cells (OD_{600} = 0.5-0.75).

2.5.3 2K1 fusion proteins expression and purification

Cells containing the appropriate expression plasmids were seeded into 500 mL TB plus kanamycin at 9 OD_{600} / 500 mL. Once the OD_{600} reached 0.5 IPTG was added to a final concentration of 1 mM and the heater in the shaker was turned off to allow for overnight expression of proteins at room temperature. The next day the cells were pelleted by centrifugation (5000g for 10 min). Pellets were resuspended in 5 mL lysis buffer (20 mM imidazole, pH 7; 4 mM $MgCl_2$; 10 mM mercapto ethanol; 0.02 $\frac{mg}{mL}$ Peptstatin A; 0.2 $\frac{mg}{mL}$ TPCK; 0.2 $\frac{mg}{mL}$ TAME; 0.02 $\frac{mg}{mL}$ leupeptin; 0.2 $\frac{mg}{mL}$ soybeans; 20 mM PMSF in water). Cell suspension was incubated on ice for 30 minutes, frozen in liquid nitrogen and stored at -90°C overnight. The next day the frozen cells were thawed with alternating 2 minute exposures to a 37°C water bath with agitation and ice. The 37°C / ice procedure was repeated three times, then the cells were frozen in liquid nitrogen and the thawing cycle was repeated. This freeze / thaw cycle was performed three times. RNAse (0.5 mL of Sigma R5000) and DNAse (0.25 mL of Sigma D4527) were added to the lysates and the samples were incubated for 30 minutes on ice on a slowly tilting shaker. The solutions were then spun at 21,800g for 20 minutes and the pellets were discarded. The supernatent was spun at 180,000g for 30 minutes at 4°C.

The supernatent was then added to 2 mL of pre-equilibrated Ni-NTA resin (Qiagen Ni-NTA Superflow) and were incubated overnight on a rotating shaker. The next day the resin/supernatent mixture was added to a clean column. Resin was washed with 20 mM imidazole pH 7, 4 mM $MgCl_2$. Eluent was collected and protein content was monitored with Bradford reagent. Resin was washed until no protein eluted. Fractions were then eluted from column with increasing concentrations of imidazole in 4 mM $MgCl_2$, 10 mM βME (70 mM imidizole – 500 mM imidizole). Fractions were analyzed by SDS-PAGE and fractions with high purity of fusion proteins were concentrated in a Vivaspin Centricon 15 (MW cutoff 30 kDa) and were doubly buffer exchanged into 1x PBS, pH 7.2.

The Edlehoch model was used to calculate the extinction coefficients of the two fusion proteins. [48]. Extinction coefficients of 67280 and 79850 at 280 nm were calculated for 2k1-MBP-protA and 2K1-MBP-bFGF, respectively. The 200 mM imidazole fraction for 2K1-MBP-protA was used for further binding analysis, with a calculated MW of 63298 $\frac{grams}{mole}$. The 150 mM imidazole fraction for 2K1-MBP-bFGF was used for further binding analysis, with a calculated MW of 65149 $\frac{grams}{mole}$.

2.5.4 2K1-MBP-protA and 2K1-MBP-bFGF testing

TCT-PS wells (BD Falcon 35-3072 96-well plate) were incubated for 1 hour in 200 μL of protein dissolved in 1x PBST with 0.1% Tween (PBST). For 2K1-MBP-protA, the wells were washed three times each in 200 μL PBST and were then incubated for 1 hour with 200 μL of 500 $\frac{ng}{mL}$ rabbit anti-streptavidin in 1x PBS with 5 $\frac{mg}{mL}$ BSA (PBS/BSA). The wells were washed three times with 200 μL PBST and were incubated with 200 μL of a 1:1000 dilution of a 1 $\frac{mg}{mL}$ Streptavidin-HRP (SAv-HRP; Upstate 18-152) in PBST for 1 hour. This was followed by three washes with PBST. Finally, the wells were incubated with 150 μL of TMB

substrate. After 5 minutes, the absorbance at 650 nm was measured. Each concentration and protein combination was measured in duplicate.

For 2K1-MBP-bFGF, the wells were washed three times each in 200 μL PBST and were then incubated for 4 hours with 200 μL of 500 ng/mL heparin-fluorescein in 1x PBS with 5 $\frac{mg}{mL}$ BSA (PBS/BSA). The wells were washed three times with 200 μL PBST, and were incubated with 200 μL of a 1:1000 dilution of a 1 $\frac{mg}{mL}$ anti-fluorescein-HRP (PerkinElmer NEF710001EA) in PBS/BSA for 1 hour. This was followed by five washes with PBST. Finally, the wells were incubated with 150 μL of TMB substrate. After 5 minutes, the absorbance at 650 nm was measured. Each concentration and protein combination was measured in duplicate.

2.5.5 2K1 binding to other metals

Metal samples were incubated for 3 hours in 200 μL of protein dissolved in 1x PBST with 0.1% Tween (PBST) in a 96-well plate well. The samples were washed three times each in 500 μL PBST in a 48-well plate and were then incubated for 45 minutes with 500 $\frac{ng}{mL}$ rabbit anti-streptavidin in 1x PBS with 5 $\frac{mg}{mL}$ BSA (PBS/BSA). The samples were moved to a clean 96-well plate, were washed five times with 200 μL PBS/BSA, and were incubated with 200 μL of a 1:1000 dilution of Streptavidin-HRP (SAv-HRP) in PBS/BSA for 30 minutes. This was followed by five washes with PBST. Finally, the samples were transferred to clean wells and were incubated with 200 μL of TMB substrate. After 15 minutes, the samples were removed from the wells and the absorbance at 650 nm was measured. Each sample and protein combination was measured in duplicate.

2.5.6 2K1-biotin and 2K1-CM15 peptide synthesis

The peptides were made by Genscript (Piscataway, NJ, USA) using standard solid phase peptide synthesis methods, were purified by desalting and were shipped lyophilized. The peptide was resuspended and stored in Millipore water and diluted to the appropriate concentrations in Millipore water. Analysis of the HPLC traces included with the peptides showed that the 2K1-CM15 peptide was roughly 60% pure and the 2K1-biotin peptide had $>$ 95% purity. Peptide synthesis was verified by MS analysis by Genscript. 2K1-CM15 had a 4 glycine linker between the 2K1 and the CM15 moieties.

2.5.7 2K1-biotin binding to stainless steel

Stainless steel circles were incubated for 1 hour in 200 μL of protein dissolved in 1x PBST with 0.1% Tween (PBST) in a 96-well plate well. After addition of the protein solution, each stainless steel circle was checked to ensure that no air bubbles were blocking the surfaces. The samples were washed three times each in 200 μL PBST and were incubated with 200 μL of a 1:1000 dilution of Streptavidin-HRP (SAv-HRP) in PBS/BSA for 30 minutes. This was followed by five washes with PBST. Finally, the samples were transferred to clean wells in a 48-well, clear-bottomed plate and were incubated with 500 μL of TMB substrate. The plates were immediately placed in a plate reader and the absorbance was measured at 650 nm every 60 seconds for 20 minutes with a vigorous agitation (plate reader setting) between each read

to move the stainless steel disks to a new location in the well. This was done so that any blockages of the light beam by the disk would not last for more than one consecutive read. The slopes of the kinetic data were calculated, after removing any data points where it was obvious that the stainless steel disk had blocked the light path. Each protein combination was measured in triplicate and the background absorbance rates were subtracted from each value.

2.5.8 Calculating binding affinity

Enzymatic rates were plotted against protein/peptide concentrations. Using QTI Plot, the equation

$$y = \frac{(R_{max})(x)}{K_D + x} \tag{2.1}$$

was used to fit a sigmoidal curve through the values, to solve for R_{max} and K_D.

2.5.9 Functionalization of TCT-PS with 2K1-CM15

Tissue culture treated polystyrene wells were incubated with different concentrations of peptides ranging from 200 μM to 3.12 μM in PBST for one hour. Wells were then washed with PBST three times. *E. coli* cells (ATCC 25922) were grown overnight from a single colony and were then rediluted the next morning and were grown to mid-log phage (OD_{600} = 0.4). The cells were then diluted to OD_{600} = 0.0035 in fresh LB and were added to the wells. The plate was placed in a plate reader with shaking and temperature control. The cells were grown for 4 hours at 37°C. During the growth, the absorbance at 600 nm was measured and recorded on a regular basis. To calculate growth inhibition for each sample and concentration, the following equation was used:

$$Fold\ growth = \frac{Absorbance_{T=4\ hours} - Absorbance_{T=0}}{Absorbance_{T=0}} \tag{2.2}$$

The plates were then removed from the shaker, stored overnight at 4°C and were washed the next morning 2x in 1X PBS to remove unbound cells. To measure the inhibition of adhesion of live cells to the material surface, the wells were stained with LIVE/DEAD (Invitrogen) cell stain with a final concentration of 2X propodium iodide and 1X Syto 9 in 1X PBS. Three replicate images were taken in each well and there were two replicate wells for each peptide and concentration. Images were taken using a green filter (GFP) and a red filter (PI). The images were processed using ImageJ according to the following protocol. 1) Adjust brightness and contrast (using auto adjust when possible). 2) Adjust threshold (using auto adjust when possible). 3) Make a histogram of number of pixels above the threshold (this counts the area covered in the image by the number of green (live) or red (dead) cells). 4) Record the number reported in channel 256 (total number of pixels above threshold). The six values for each condition (three replicate images per well and two replicate wells per condition) were averaged.

Chapter 3

Engineered yeast for enhanced CO_2 mineralization

3.1 Summary

In this work, a biologically catalyzed CO_2 mineralization process for the capture of CO_2 from point sources was designed, constructed at a laboratory scale, and, using standard chemical process scale-up protocols, was evaluated at an industrial scale. Engineered *Saccharomyces cerevisae* was used to enhance the hydration of CO_2 and the mineralization rate of $CaCO_3$. The enhanced rates were verified using quantitative techniques in lab-scale measurements. In a variety of different tests, including pH and temperature stability and CO_2 hydration activity, yeast displayed bovine carbonic anhydrase II (bCA2-yeast) performed as well as soluble bovine carbonic anhydrase II (bCA2), which has a reported turnover rate of $1x10^6 \frac{1}{second}$. Several possible yeast-displayed mineralization peptides were tested for $CaCO_3$ mineralization rate enhancement. At concentrations below $2x10^{12} \frac{peptides}{mL}$, there was little to no effect on the mineralization rate for the yeast-displayed peptides. At concentrations around $2x10^{12} \frac{peptides}{mL}$, though, there was a roughly 50% increase in mineralization rate, which appeared to be independent of the peptide sequence displayed. Even the yeast with no peptide displayed was able to increase the mineralization rate by 20%-30% when tested at the same cell concentrations as the peptide-displaying yeast.

The effect that these enhanced CO_2 hydration and $CaCO_3$ mineralization rates have on the CO_2 capture cost in a CO_2 mineralization process using coal fly ash as the CaO source was then evaluated. The cost impact of the bCA2-yeast on the process is significant, decreasing the predicted cost of CO_2 capture by 8.5% to 13.5%, when compared to a CO_2 mineralization process with no biological components. This process model also predicts that the use of purified soluble, recombinant CA2 would not provide these same cost savings (only accounting for the difference in production and purification costs and ignoring the mineralization rate and settling rate effects).

While the total cost of CO_2 capture by this process is still higher than the costs of other techniques for CO_2 capture, mineralized CO_2 has the advantage of being permanently sequestered and also of being in a form that lends itself to use in downstream applications (especially in the building and construction industries).

3.2 Introduction

3.2.1 Climate change is caused by anthropogenic CO_2 emissions

> "Warming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice and rising global average sea level."
>
> – Intergovernmental Panel on Climate Change (IPCC)
> IPCC Fourth Assessment Report: Climate Change 2007 [49]

Over the last century and a half, atmospheric carbon dioxide (CO_2) levels have increased 35% and anthropogenic CO_2 sources are largely responsible for this increase, which is, in turn, a major contributor to climate change. [50] Though the full impact of continued climate change is still uncertain, first effects are already evident. According to the IPCC, there is "high confidence that recent regional changes in temperature have had discernible impacts on physical and biological systems." [49] The changes are having broad impacts, including on human health, travel, and recreation, as well as on the suitability of land for residential development, and agricultural, forestry and fishery management.

There is no simple solution for a problem of this magnitude. The human response to global warming will require a combination of adapting to a changing planet, altering the manner and amount of energy consumption, and reducing the release of greenhouse gases (like CO_2) during energy production. The largest energy-related emitters of global CO_2 are the world's 2100 coal-fired power plants. [51] The capture of CO_2 from these sources motivated this work.

The concentration of greenhouse gases in the earth's atmosphere has progressively increased as the human population has grown, especially since the onset of the Industrial Revolution in the 1850's. [49]. Atmospheric CO_2 levels have increased from 280 parts per million (ppm) in the nineteenth century to 380 ppm in 2006. [50] Over the same time period, the global temperature has increased by 0.6 ± 0.17°C. [49] The increase in CO_2 and other greenhouse gas levels is widely accepted as the cause of the increase in the average global temperature. [52] The consensus among climate scientists is that CO_2 levels should not be allowed to rise above 500 ppm – otherwise the result may be significant and deleterious alterations to the global climate. This means that the global community must avoid releasing 175 gigatonnes (Gt) of carbon into the atmosphere in the next 50 years. [50]

Emissions from the world's 2100 coal-fired power plants are responsible for about a third of the CO_2 released into the atmosphere. [53] Since 2003, coal has been the leading contributor to global energy-related CO_2 emissions. [51] Atmospheric CO_2 released by burning coal will continue to increase, due to the vast worldwide reserves of coal (almost 150 years' supply at current usage rates) and the increasing global demand for energy. [54] In fact, the overall capacity of the world's coal-fired power plants is projected to double by 2030, to 2200 gigawatts. [55] Additionally, the number of all large CO_2 emission sources (electric power and industrial manufacturing plants) is projected to continue increasing. [52]

To reduce the amount of CO_2 that is released into the atmosphere by anthropogenic sources, a multi-pronged strategy will be required. This strategy must include more conservation of energy via improved energy efficiency and reduced energy consumption, as well as

increased use of low- and near-zero carbon energy sources. However, given the vast reserves of fossil fuels still available and the large global infrastructure in place to create energy from fossil fuels, it seems very likely that most, if not all, of the fossil fuels will be consumed. [56] In order to mitigate the global harms, the CO_2 produced during the combustion of fossil fuels must be captured and stored rather than released into the atmosphere. Carbon capture and storage (CCS) is most easily applied to the large point sources of CO_2 (those that emit more than 0.1 megatonne (Mt) per year), such as fossil fuel power plants and industrial manufacturing plants. [52] This work focuses on capturing CO_2 from these sources.

3.2.2 Carbon capture and storage strategies

Several technologies for transporting and storing large volumes of CO_2 are already beyond the research stage, as a result of innovations piloted by the mining and energy industries. Additionally, several CO_2 capture technologies are already mature enough to be considered economically viable in certain situations. For example, transporting large volumes of liquid or gaseous CO_2 from a capture point to a storage point via a pipeline could be achieved using the same technologies that the oil industry already uses to move oil and natural gas. As part of a process called enhanced oil recovery (EOR), the CO_2 can then be pumped into an underground oil bed to help extract additional oil while simultaneously geologically storing the CO_2.

The existence of mature CO_2 capture and storage technologies should not preclude the development of additional CCS options. A reduction in anthropogenic CO_2 emissions will not be achieved with a single strategy and there are significant costs and risks associated with existing CCS technologies. For example, once the CO_2 has been captured, it must be transported and stored. If mature CCS technologies are to be relied upon, then the captured CO_2 will be in a gaseous or liquid form. The two most likely options for storing it are in deep underground geological formations or in the ocean. Both of these strategies carry legitimate risks of CO_2 leakage back into the atmosphere and these sites will require long-term monitoring. [57]

Additional drawbacks to some of the CO_2 storage options are the storage capacity and the storage time. [57] For example, EOR is only capable of storing tens of Gt of CO_2 – just barely more than several years' worth at current emissions rates. [57] In contrast to EOR and other mature CCS technologies, mineral carbonation, one of the least mature technologies, has a significant storage capacity and long storage time. [57] Mineral carbonation entails the conversion of CO_2 to solid carbonate minerals like calcium or magnesium carbonate.

CO_2 mineralization is an attractive CCS technology for several reasons. Mineral carbonates are theoretically capable of storing all of the twenty-first century's CO_2 emissions. [57] Also, it is unlikely that the CO_2 will be released into the atmosphere if it has been mineralized, because carbonate minerals are the most geologically stable form of carbon dioxide – being stable on the order of thousands of years. [57]

Carbonate minerals have a significantly lower energy state than CO_2 and, at least theoretically, the mineralization process could produce energy. The reactions of CO_2 with calcium and magnesium oxide generate 179 $\frac{kJ}{mole}$ and 118 $\frac{kJ}{mole}$, respectively. [58] In context, this is only two to three times less energy than is generated by the combustion of carbon, which generates 394 $\frac{kJ}{mole}$. It is important to clarify that calcium and magnesium are much more

abundant in mineral form than they are as simple oxides, a topic which any CO_2 mineralization strategy must address. Still, the carbonation reaction is exothermic even for these minerals, generating energy on the order of 50-100 $\frac{kJ}{mole}$. [58] Thus, at the very least, the mineralization reaction could be expected not to consume additional energy. A net energy requirement of zero is a desirable feature of any CCS technology, since consumption of energy during CCS equates to an increase in the cost of the energy produced by the power plant.

Another concern with mineral carbonation is that the molar weight of carbonate minerals is heavier than that of pure CO_2. A typical (300 MW) coal plant releases 7.0 x 10^6 kg of CO_2 into the atmosphere. Capturing this CO_2 in the form of calcium carbonate ($CaCO_3$) would produce 7.0 x 10^8 kg of $CaCO_3$, which would then need to be transported and stored. This is a staggering number. However, the mining and coal industries already function on this scale. A 300 MW coal plant burns 20% of this mass of coal every day and, thus, a nearly 20% reduction in CO_2 emissions using mineral carbonation could be met without the need to increase the plant's solids handling capacity. Moreover, the transportation and storage costs of carbonate minerals must be compared to those of liquid or gaseous CO_2, which will be substantial in their own right. If the mineralization process could create a material of economic value, then the transportation and storage costs might be born by the downstream user.

Mineral carbonation faces its own challenges for CO_2 sequestration because the mineralization process occurs slowly or requires extreme and costly operating conditions. [56] Some attempts have been made to improve the reaction with high temperature or pressure, but the use of these reaction pathways leads to cost estimates of $60-100 per tonne of CO_2 sequestered, which is well above the $1-5 range for ocean and geologic sequestration. [58,59]

3.2.3 Biologically catalyzed mineralization as a CCS solution

A biologically catalyzed approach might make mineral carbonation economically feasible by combining standard temperature and pressure (STP) conditions with the low cost of cell-surface-anchored enzymes and peptides (as compared to the costly amine-scrubbers and recapture processes that are required for other CCS technologies). Fortunately, there are many examples in nature of biologically catalyzed CO_2 mineralization. The growth of carbonate minerals in mollusk shells, sea urchin spicules, coral, avian eggshells, and oyster pearls is controlled and catalyzed by organic biomolecules. [60–64]

The goal of this work, then, was to design, build, model the scale-up of, and evaluate a biologically catalyzed system that can be used to capture CO_2 directly from the flue gas of a fossil fuel power plant. This would eliminate the necessity of pre-capturing the CO_2 from the flue gas, which is typically 15% CO_2 and 85% N_2. Integrating CO_2 capture and mineral sequestration has been proposed as one way to make mineralization more competitive as a CCS technology. [58]

Others have investigated biomineralization for CO_2 sequestration. In prior work, bacteria and cyanobacteria suspected of being capable of biomineralization were screened for the ability to remove CO_2 from a closed reactor. [65–67] The species that were identified took several days to have a detectable impact on the CO_2 levels in a small reactor – results that are promising, but ones that leave biomineralization out of contention as a serious large-scale CCS option.

The approach reported here differs in that it entailed starting with a well-known biological organism and engineering *in* the ability to catalyze the mineralization of CO_2, rather than testing organisms with known biomineralization abilities. Since CO_2 mineralization is a multi-step process, this allowed for the determination of which peptides and proteins participate in the mineralization process and the contribution of each to the overall reaction. This bottom up approach provided more exact control over the molecular components in the system, aiding in system optimization. Additionally, knowing the contributions of each component in the system facilitated the development of a scaled-up process model, which was used to simulate and evaluate the performance at a large scale.

The mineralization process of CO_2 can be broken down into three steps: the dissolution of CO_2 into an aqueous solution, the conversion of CO_2 to bicarbonate, and, finally, the mineralization of the bicarbonate by divalent cations like calcium or magnesium (see Table 3.1). The approach for this work was to address each of the three steps individually. The first step, the dissolution of CO_2, can be controlled by modifying the process equipment used to absorb the CO_2. The rate of the second step, the conversion of CO_2 to bicarbonate, can be enhanced by an enzyme, carbonic anhydrase. The hypothesis for this work was that the rate of the third step (the mineralization of bicarbonate) could be enhanced by the presence of mineralization peptides.

	Reaction	**Method for enhancing reaction rate**
0	$CO_{2_{(g)}} \rightleftharpoons CO_{2_{(aq)}}$	CO_2 absorber design affects dissolution rate
1	$CO_{2_{(aq)}} + H_2O \rightleftharpoons HCO_3^- + H^+$	Addition of carbonic anhydrase
2	$H^+ + HCO_3^- + Ca^{2+} + 2OH^- \rightharpoonup CaCO_3 + H_2O$	Addition of mineralization peptides

Table 3.1: *Reactions in the CO_2 mineralization process* Strategies for enhancing the rate of each reaction are listed in the column next to the reactions. Double arrows indicate that a reaction is reversible. The single arrow in the last reaction indicates that the reaction is essentially irreversible under the conditions proposed for this work.

Rather than attempt to utilize existing biomineralization organisms for CO_2 sequestration, a biomineralization system was engineered from the bottom up. This approach had two benefits. First, it allowed for the selection of a biological organism that is easy to cultivate, well-understood and relatively innocuous. Easy cultivation was desirable because the scale of the solution will require large amounts of the engineered organism. The objective also dictates that the organism be well-understood and innocuous. The carbonate mineral product or the waste materials from each CO_2 capture site could contain some of the biological material used to catalyze the reaction. Given the public's general cautiousness with new biological technologies, using an unknown organism for the CCS solution might lead to implementation problems due to (possibly unsubstantiated) safety concerns associated with the release of the organism.

Second, by choosing to engineer from the bottom up, there was greater control over the amounts and concentrations of the biomineralization peptides and proteins. This was

beneficial because the engineered organism needed to catalyze CO_2 mineralization under the specific reaction conditions defined by the final application – that is, in a industrial-scale reactor designed to capture CO_2 from the flue gas of a power plant.

With these constraints in mind, the yeast *S. cerevisae* was chosen. *S. cerevisae* has been used in the food and beverage industry for years and has "generally regarded as safe" (GRAS) status. This addresses two of the requirements established above. First, the protocols for cultivation and disposal of large quantities of *S. cerevisae* already exist. Second, utilizing *S. cerevisae* minimizes the chances of encountering safety concerns associated with the use a biological organism.

In addition to being relatively safe and easy to cultivate, there is a convenient and robust method for engineering *S. cerevisae* cells to express specific peptides and proteins on their outer cell walls called "yeast surface display" (see Figure 3.1). [68,69] This method allowed for the systematic testing of each component of the biomineralization process. In yeast display, the *aga1* gene is in the yeast genome under the control of a galactose-inducible cassette, whereas the *aga2* gene and the gene for the protein to be displayed (in the image, human carbonic anhdrase II, or hCA2) are under the control of galactose-inducible cassetteon a plasmid. Once induced, Aga2 is fused to Aga1 via disulfide bonds and the hCA2 enzyme is displayed on the N-terminal end of Aga2.

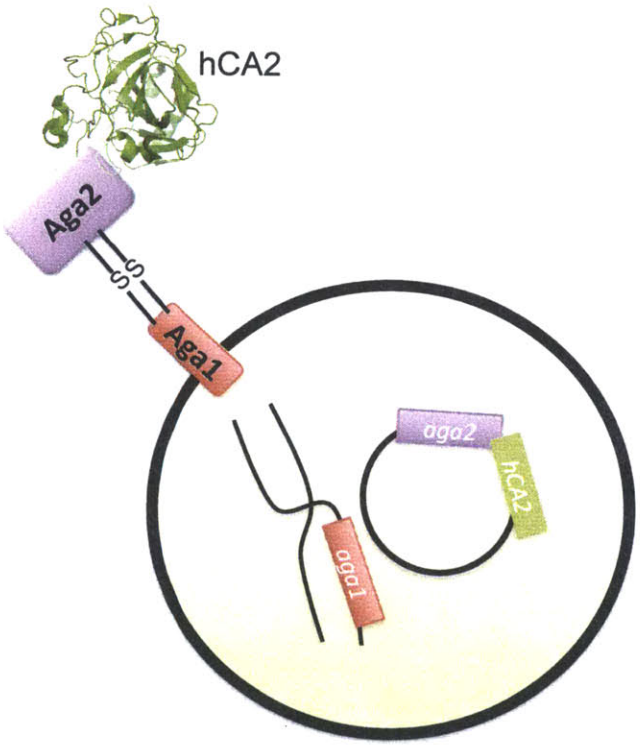


Figure 3.1: *Cartoon image of yeast-displayed hCA2* The *aga1* gene (red) is in the yeast genome under the control of a galactose-inducible cassette, whereas the *aga2* (purple) and *hCA2* (green) genes are under the control of galactose-inducible cassette on a plasmid. Once induced, Aga2 is fused to Aga1 via disulfide bonds and the hCA2 enzyme (in green) is displayed on the N-terminal end of Aga2.

3.3 Results

3.3.1 Enhanced conversion of CO_2 to HCO_3^-

The first part of the work was to accelerate the rate of hydration of CO_2 to bicarbonate (HCO_3^-) under the conditions necessary for CO_2 mineralization from the flue gas of a power plant. The enzyme carbonic anhydrase II (CA2) was selected as the model way of achieving this acceleration, since it is a well-studied, robust enzyme with high activity (see Table 3.2). In addition, CA2 is a cytosolic isoform, which means that it is compatible with the yeast-display technique.

Bovine carbonic anhydrase is commercially available in a lyophilized form, so it was used as a control against which to measure the activity of carbonic anhydrase displayed on the surface of *S. cerevisae* cells. Human and bovine carbonic anhydrase II genes were cloned into the yeast display plasmid as described below. Expression of genes from the pCTCON2 plasmid leads to proteins that are fused to the N-terminal end of the Aga2 protein – a yeast mating protein that is permanently anchored to the surface of the yeast cell. In addition, the fusion protein has two epitope tags, an HA tag in between Aga2 and the gene of interest (carbonic anhydrase, in this case) and a c-MYC tag on the C-terminal end of the gene of interest. See Figure 3.1 for a cartoon image of the yeast display system. By staining the yeast cells with fluorescently labeled antibodies against these epitope tags, one can confirm and quantify expression levels of the fusion protein and the protein of interest.

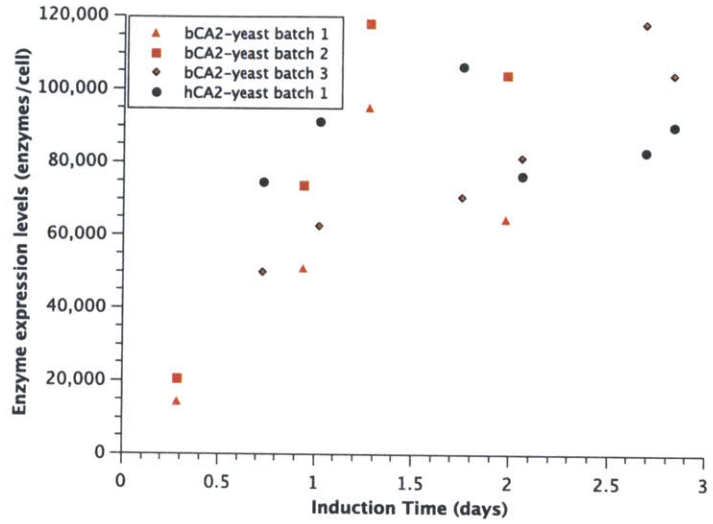
Initially, two isoforms of carbonic anhydrase II, the bovine (bCA2) and human (hCA2) were displayed on the surface of *S. cerevisae* cells. Since maximal expression of the enzymes is critical to the performance of the CO_2 mineralization process, the expression levels of the enzymes was measured at several induction time points over the course of three days of induction. Somewhere between 1 and 1.5 days, the cells reach the maximal average expression level, which is $\approx$ 100,000 $\frac{enzymes}{cell}$ (see Figure 3.2a). It takes another 12 to 24 hours of induction to reach the maximum numbers of enzymes in the fermentation (Figure 3.2) due to the fact that the cell concentration continues to increase. The important metrics from these data for the CO_2 capture process design are the average expression levels of CA2 per cell and the amount of time it takes to reach those cell densities.

After verifying that the *S. cerevisae* yeast display system can be used to express carbonic anhydrase enzymes on the surface of the cells, the next step was to determine whether the

Isozyme	Activity level	k_{cat} (s^{-1})	Cellular location
hCA I	moderate	2.0 x 10^5	cytosol
hCA II	very high	1.4 x 10^6	cytosol
bCA II	very high	1.0 x 10^6	cytosol
hCA III	very low	1.0 x 10^4	cytosol
hCA VA	low	2.9 x 10^5	mitochondria
hCA VI	moderate	3.4 x 10^5	secreted
hCA IX	high	3.8 x 10^5	transmembrane

Table 3.2: *Selected carbonic anhydrase isoforms* Human and bovine carbonic anhydrase isoforms, their turnover numbers, and their cellular location. [70, 71]

(a)



(b)

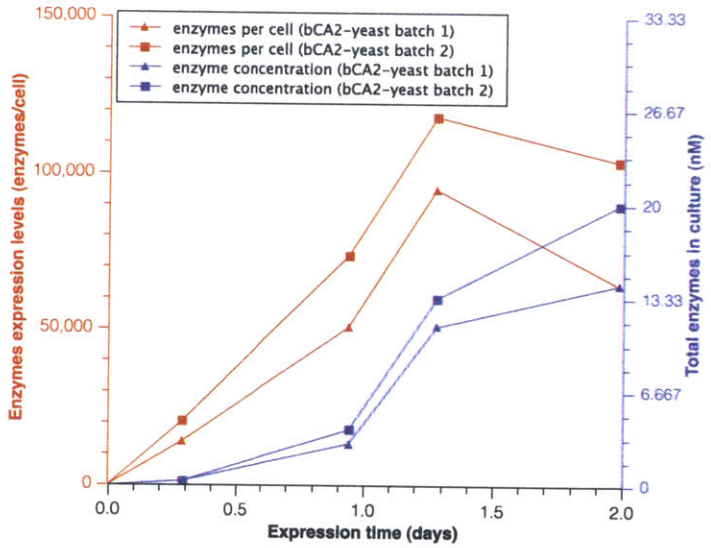


Figure 3.2: *Measurement of carbonic anhydrase expression levels at different induction times* (a) This figure shows the average number of enzymes displayed per cell for four different batches of carbonic anhydrase II expressing cells. (b) For two batches of bCA2-yeast, the levels of enzyme displayed per cell is plotted in addition to the total amounts of enzyme in the fermentation. While the average number of enzymes per cell reaches maximal levels sometime shortly after 1 day of induction, the total numbers of enzymes in the culture continues to increase up to around 2 days of induction, due to the fact that the cells continue to grow.

enzymes were active. To do this, a modified version of the Wilbur-Andersen method was used. [72] Soluble bovine carbonic anhydrase from Worthington Enzymes was used as a positive control in the experiments. Using the enzyme expression levels per cell that were measured previously, the amount of yeast-displayed enzyme added to the reactions could be determined by measuring the amount of cells added to the reaction. A comparison of reaction rates for the various yeast-displayed carbonic anhydrase isoforms to the soluble carbonic anhydrase could be made in order to determine the activity of each yeast-displayed isoform relative to the commercially available soluble bCA2.

In addition to hCA2 and bCA2, a third isoform of carbonic anhydrase (CAH) from the mildly thermophilic bacteria *Streptococcus thermophilus* was displayed on the surface of the yeast. The hypothesis was that CAH would be more stable at the elevated temperatures that might be present in the industrial scale CO_2 capture process. This isoform was chosen because *S. thermophilus* has a preference for growth at temperatures that are slightly higher than that preferred by humans and cows (42°C vs. 37°C). The use of a carbonic anhydrase isoform from a more severe thermophile was considered, but those isoforms all appeared to be γ-type with multiple transcriptional subunits combining to form the completed enzyme. This approach was not expected to be compatible with the yeast-display system in its current form.

As can be seen in Figure 3.3, all three isoforms were expressed in an active form on the surface of the *S. cerevisae* cells. Both the bCA2 and CAH isoforms had comparable activity to the soluble bCA2, whereas the hCA2 appeared to have lower activity. To more quantitatively assess the activity of the yeast displayed isoforms, a modified version of Wilbur-Andersen units was used. [72] Briefly, the amount of time that it took for the blank reaction to reach 50% completion was subtracted from the time for the sample reaction to reach 50% and the result was divided by the sample time to reach 50% completion. An example of these results for the bCA2-yeast sample are reported in Figure 3.3d, and the same analysis was performed for the other two isoforms displayed on yeast.

Using this assay, it is not possible to obtain an exact measurement of the turnover rate of the carbonic anhydrase enzymes displayed on the yeast cells, because the assay is performed at 4°C. This assay is suitable, however, for determining relative activity compared to a control sample with a known turnover rate. Both the CAH-yeast and the bCA2-yeast had activities that were comparable or better than the soluble bCA2, so further assessment of the utility of the yeast-displayed enzyme in a large-scale CO_2 mineralization process was performed with these two isoforms.

The next step in assessing the utility of yeast-displayed carbonic anhydrase in a CO_2 mineralization process is to measure the stability of the enzymes under conditions to which they might be exposed in the industrial process. The industrial process is being designed with a specific CO_2 generation source in mind, and it is onerous to cover every possible condition that could exist in an industrial CO_2 capture process. However, the most likely source of degradation for the enzymes is exposure to elevated temperatures in the capture process (flue gas temperatures often exceed 100°C and proteins are well-known to be susceptible to degradation at elevated temperatures).

To assess the temperature stability of bCA2-yeast and CAH-yeast, a large batch of each sample was made and expression levels were measured to determine the amount of enzyme per unit of cells. The cells were then rinsed, resuspended in a lightly buffered solution at

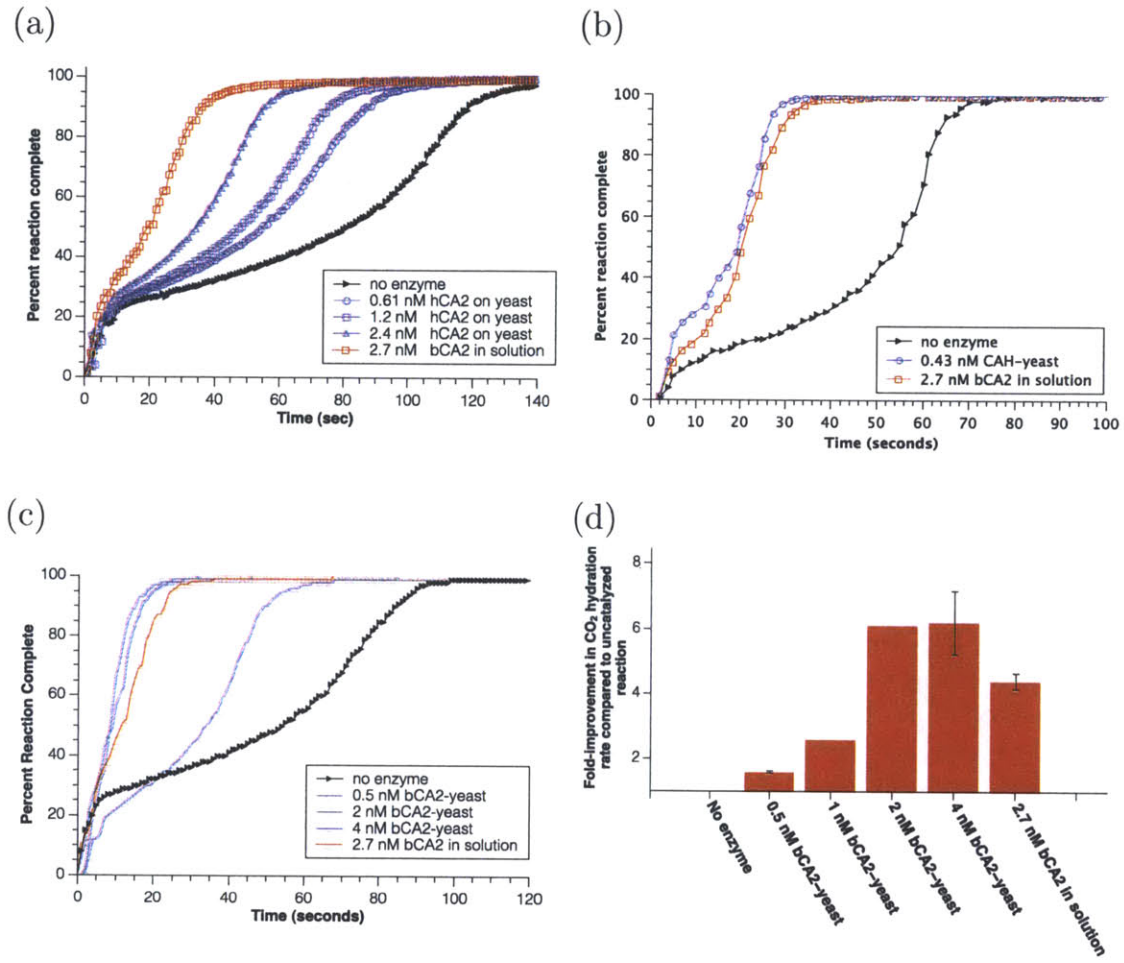


Figure 3.3: *Activity of yeast displayed carbonic anhydrase isoforms compared to commercially available bovine carbonic anhydrase* (a) hCA2-yeast catalyzed hydration of CO_2 measured at 4°C (b) CAH-yeast hydration of CO_2 measured at 4°C (c) bCA2-yeast hydration of CO_2 measured at 4°C (d) Results of bCA2-yeast titration normalized to the no enzyme sample. Activity was calculated by the amount of time to reach 50% reaction completion. Each concentration was measured in duplicate and values are reported as averages with 1 standard deviation error bars.

pH 8.8 and were split into aliquots that were stored at 50°C or at 4°C. The samples were returned to 4°C to measure the activity. The starting activity levels of each was measured using a modified version of Wilbur-Andersen units (see Section 3.5.6 wherein each sample's activity in Wilbur-Andersen units are reported per nmole of enzyme in each reaction). By converting the results of the activity assay into activity units per numbers of enzyme, the dynamic range of the assay can be extended and samples of varying concentrations (or varying activities) can be compared to each other. The results of the 50°C stability study are reported in Figure 3.4a. An examination of these results shows that the bCA2-yeast and the CAH-yeast had roughly the same starting activity (which was about 4 times higher than that of the soluble bCA2). It is unlikely that the yeast-displayed enzymes are really more active than pure bCA2. This difference could be due to a slight loss of activity in the bCA2 sample from the preparation, shipping, or storage conditions of the soluble bCA2 (purified from bovine erythrocytes, lyophilized and shipped from Worthington Enzymes and stored in a -20°C freezer in lab).

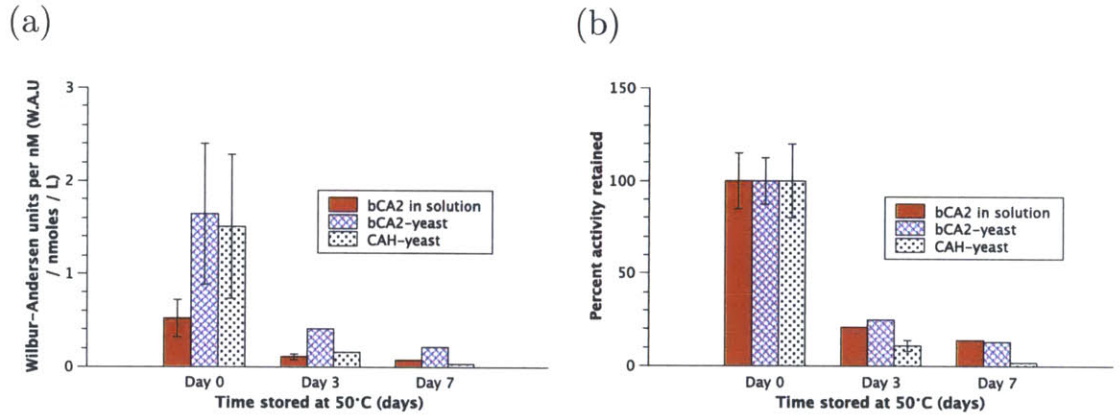


Figure 3.4: *Temperature and pH stability of yeast displayed carbonic anhydrase isoforms* (a) Modified Wilbur-Andersen units (amount of time for reaction to reach 50% completion compared to blank reaction) per nM are reported for each enzyme sample. Day 0 samples are freshly prepared enzyme solutions. The samples were stored at 50°C for 3 days and 7 days and activity was measured for each sample using the assay described in Section 3.5.5. (b) The results for each sample were normalized to the starting activity levels so that the stability of each sample could be compared to that of the other samples. bCA2-yeast and bCA2 in solution both retain more than 20% of their original activity even after 3 days at 50°C.

In order to compare the temperature stabilities of the yeast-displayed carbonic anhydrase isoforms to each other and to the soluble bCA2 more easily, the results from Figure 3.4a were normalized to the Day 0 activity levels for each sample (e.g., $\frac{bCA2_{Day\ i}}{bCA2_{Day\ 0}}$). These results (reported in Figure 3.4b) show that the bCA2 in solution and bCA2-yeast have very similar temperature stability profiles. Even after three days at 50°C, they retain more than 20% of their original activity, and after seven days at 50°C, they retain more than 10% of their original activity. Interestingly, the carbonic anhydrase isoform from the thermophilic *S. thermophilus* did not have superior stability compared to bCA2 at 50°C. For this reason,

the use of CAH-yeast was not further pursued in this work.

Finally, the pH stability at pH 7.0, pH 6.0, pH 5.0, and pH 3.0 of the bCA2-yeast was calculated and compared to soluble bCA2, using the same assay as that used for the temperature stability. The samples were stored for sixteen hours in the different buffered solutions, then were rinsed and resuspended into (for bCA2-yeast) or diluted into (for bCA2 in solution) the assay buffer. The activity was measured at 4°C and the result was normalized to the activity of freshly prepared sample. The bCA2-yeast and the bCA2 in solution have comparable stabilities, with both retaining more than 60% of the original activity down to pH 5.0, but losing 90% of the original activity at pH 3.0.

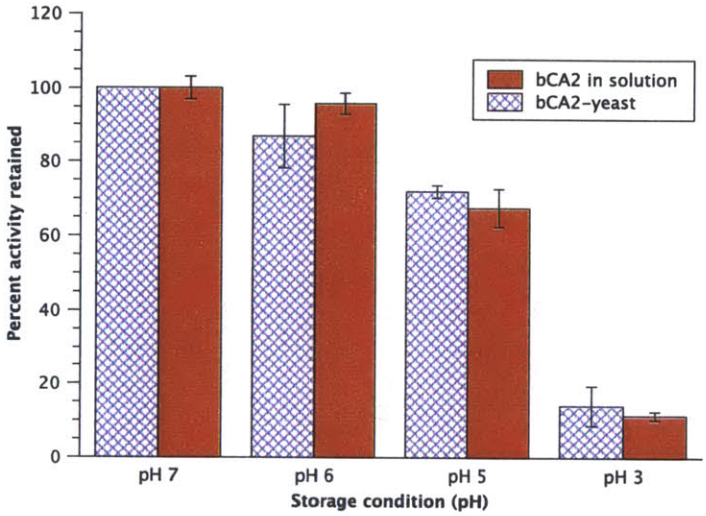


Figure 3.5: *pH stability of bCA2-yeast and bCA2 in solution* Samples were stored for 16 hours at 4°C from pH 7.0 to pH 3.0. Samples were then transferred to the CA2 assay buffer, activity was measured and compared to the starting activity.

These results demonstrated that bCA2 displayed on the surface of yeast cells has comparable activity to soluble bCA2. The reported turnover rate of bCA2 is 1 x 10^6 $\frac{1}{second}$ (see Table 3.2). A conservative turnover rate of 6 x 10^5 $\frac{1}{second}$ was chosen for the value to be used in the scaled-up process design (see Section 3.3.7) to account for unforseen process effects that might decrease the efficiency of bCA2-yeast.

3.3.2 Enhanced mineralization of $CaCO_3$

The second part of this work was to identify biomolecules displayed on the yeast *S. cerevisae* that could accelerate the mineralization rate of $CaCO_3$. To do this, a mineralization assay that could measure the kinetics of mineralization had to be developed. Figure 3.6 shows a cartoon image and a photo of this assay configuration. Briefly, a calcium-sensitive electrode was used to measure the rate of decrease of Ca^{2+} ions in a solution as sodium bicarbonate was added. As the bicarbonate and calcium ions combined to form calcium carbonate crystals (releasing a hydrogen ion in the process), the concentration of Ca^{2+} decreased. The

rate of addition of the bicarbonate solution was set to be faster than the expected rate of mineralization, so as not to be rate-limiting. The change in Ca^{2+} concentration was recorded via an attached computer and the slope was converted to moles $CaCO_3$ formed per second, adjusting for the change in concentration of the solution as the bicarbonate was added.

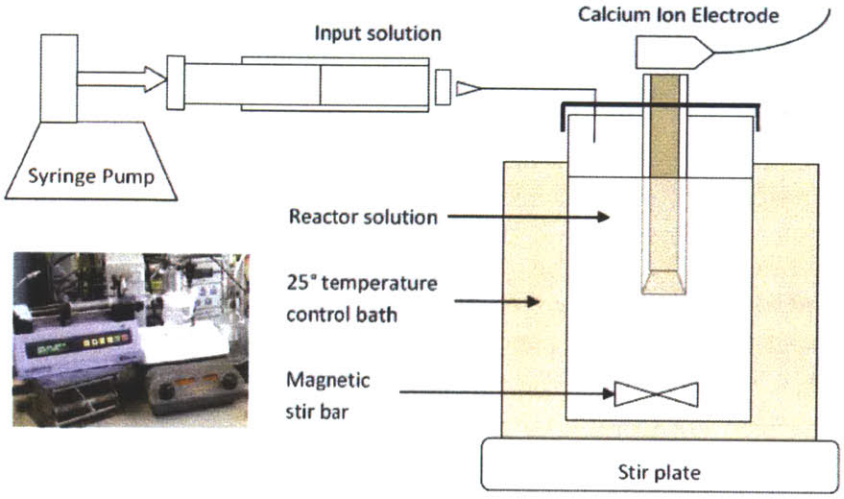


Figure 3.6: *Mineralization rate measurement* A beaker (or 50 mL Falcon tube) with a magnetic stir bar was set on top of a stir plate. The beaker was filled with a solution of Ca^{2+} ions (with chloride as the counter ion) and the peptide displaying yeast samples. A calibrated calcium sensitive electrode was placed in the solution. A syringe pump was used to inject, at a consistent rate, a solution of HCO_3^{-1} ions (with sodium as the counter ion) into the beaker. The temperature of the solution was verified to be 25°C. The electrode was connected to an ISE meter and the data was pulled into a spreadsheet on an attached computer via the RS232 port. (Inset: A photo of the mineralization rate measurement system on the lab bench.)

There are several natural examples of biologically catalyzed mineralization of carbonate. [60–64] The mineralized tissues of many organisms have been analyzed and found to contain peptides rich in acidic amino acids and phosphorylated amino acids, though they occasionally also contain acidic sulfated polysaccharides or glycoproteins. [73]

With respect to the mineralization of calcium carbonate, the acidic peptides appear to interact electrostatically with calcium in order to initiate the nucleation and, in some circumstances, to control the growth of the mineral. These peptides are rich in aspartate and, to a lesser degree, glutamate, and they often appear in repeated motifs. For example, in the scallop shell protein MSP-1, the aspartate residues are arranged "more or less regularly ... exhibiting some repeated motifs such as Asp-Gly-Ser-Asp and Asp-Ser-Asp ... . [These] domains indicate remarkable overall sequence similarities to each other. These observations suggest that the regular arrangements of COO^- groups" are important for the growth of calcium carbonate. [74] In the protein nacrein, which assists in the mineralization of calcium

previously. All of the peptides were displayed at high levels on the surfaces of the *S. cerevisae* cells $\approx$ 100,000 $\frac{peptides}{cell}$. The total number of peptides per 2 x 10^7 cells that was used in the mineralization assay was calculated and is reported in Table 3.3. Each yeast-peptide strain and a non-expressing strain of yeast was then tested in the mineralization assay at two concentrations: 2 x 10^8 $\frac{cells}{mL}$ and 2 x 10^7 $\frac{cells}{mL}$ (1 OD_{600} and 10 OD_{600}, respectively). The results are shown in Figure 3.7. At 2 x 10^7 $\frac{cells}{mL}$, there is little effect on the mineralization rate, but at 2 x 10^8 $\frac{cells}{mL}$, some of the yeast-displayed peptides increase the mineralization rate 50% or more. GPA, $(D_3G)_6$, $(D_3S)_6$, and FLK have the greatest effect on the mineralization rate, though these are also the peptides that are expressed at the highest concentrations.

The cells that are not displaying any peptide also increase the mineralization rate of $CaCO_3$ and there is a dose-dependent response (i.e., more yeast cells leads to faster mineralization). This is not surprising given that *S. cerevisae* cells are covered in negatively charged biomolecules and it has been previously reported that $CaCO_3$ can be mineralized on the surface of *S. cerevisae* cells. [78] Note that the majority of the peptide-yeast strains that are expressing negatively charged peptide do not have significantly higher mineralization rates than the yeast cell with no peptide. One explanation for this phenomenon could be that the Aga1-Aga2 fusion protein that is expressed with every mineralization peptide could be large enough to be sterically blocking the growth of $CaCO_3$ crystals on the negatively charged native surface of the *S. cerevisae* cells. And only once the peptide concentration reaches a minimum level ($\approx$ 1 x 10^{13} $\frac{peptides}{mL}$) does the presence of the mineralization peptide overcome this effect.

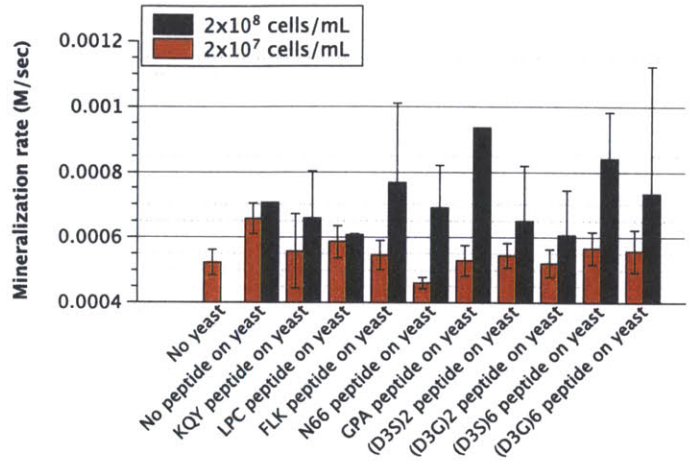
The important parameters that are necessary for the development of a large-scale CO_2 mineralization process are what concentration of cells to use and what the mineralization rate will be for the system in the presence of those cells. Thus, two possible scenarios were evaluated for the large-scale process. The first scenario was one in which just one strain of yeast (the bCA2-yeast) would be used to catalyze the hydration of CO_2 and also to serve as a mineralization template (with the assumption that the mineralization rate would be on the lower end of the peptide-yeast sample mineralization rates $\approx$ 6.3 x 10^{-4} $\frac{moles}{(L)(sec)}$). The second would be to have two strains of yeast, bCA2-yeast to catalyze the hydration of CO_2 and then a second strain to promote the mineralization of $CaCO_3$. The second strain was proposed to be GPA-yeast at a concentration of 2 x 10^8 $\frac{cells}{mL}$, which increased the mineralization rate to $\approx$ 9.3 x 10^{-4} $\frac{moles}{(L)(sec)}$.

3.3.3 $CaCO_3$ is mineralized on the surface of the yeast cells

The formation of $CaCO_3$ on the surface of the N66-yeast cells was verified by analyzing the cells with a light microscope. The samples were prepared as described in the experimental section. Representative images from the light microscope with and without the cross-polarization filters in place are shown in Figure 3.8a and b. It is apparent that the crystals are bound to the surface of the cells and that they are capable of changing the polarization of light (since only the objects that are "attached" to the cell surface are illuminated in the cross-polarization microscopy (CPM) images).

The bulk sample from the N66-yeast mineralization reaction was then analyzed by x-ray diffraction (XRD) to determine what type and morphology of crystal was being formed. The XRD analysis confirmed that $CaCO_3$ was the only crystal being formed. The resulting

(a)



(b)

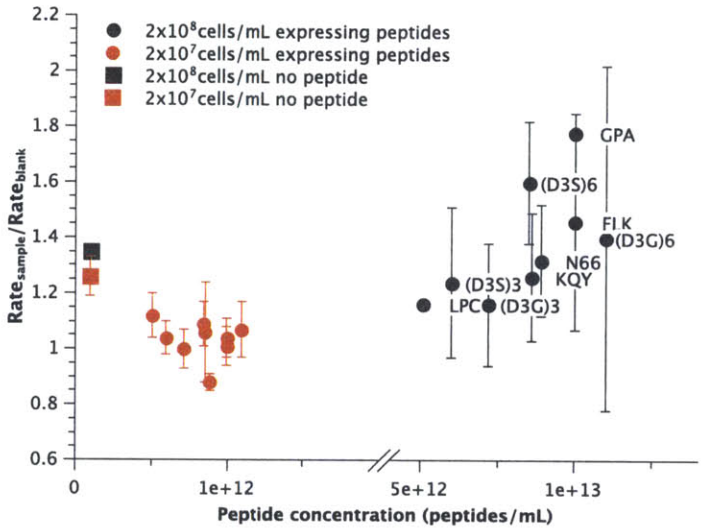


Figure 3.7: *Measurement of mineralization rate enhancement by yeast displayed peptides* (a) The mineralization rate of $CaCO_3$ in the mineralization reactor was measured for each of the ten strains of peptide-expressing yeast, as well as for a no peptide strain of yeast, at two concentrations of cells 2 x 10^8 $\frac{cells}{mL}$ and 2 x 10^7 $\frac{cells}{mL}$. The baseline mineralization rate was measured without yeast in the mineralization reactor. Each sample was measured in duplicate and the averages are reported. Error bars represent one standard deviation. (b) The mineralization rates for each sample were normalized to the rate of the blank (no yeast) sample and are plotted against the concentration of peptide in the sample. The yeast with no peptide is plotted at an arbitrary concentration of 1 x 10^{11} $\frac{peptides}{mL}$ so that it can be viewed on the same chart as the peptide-yeast strains.

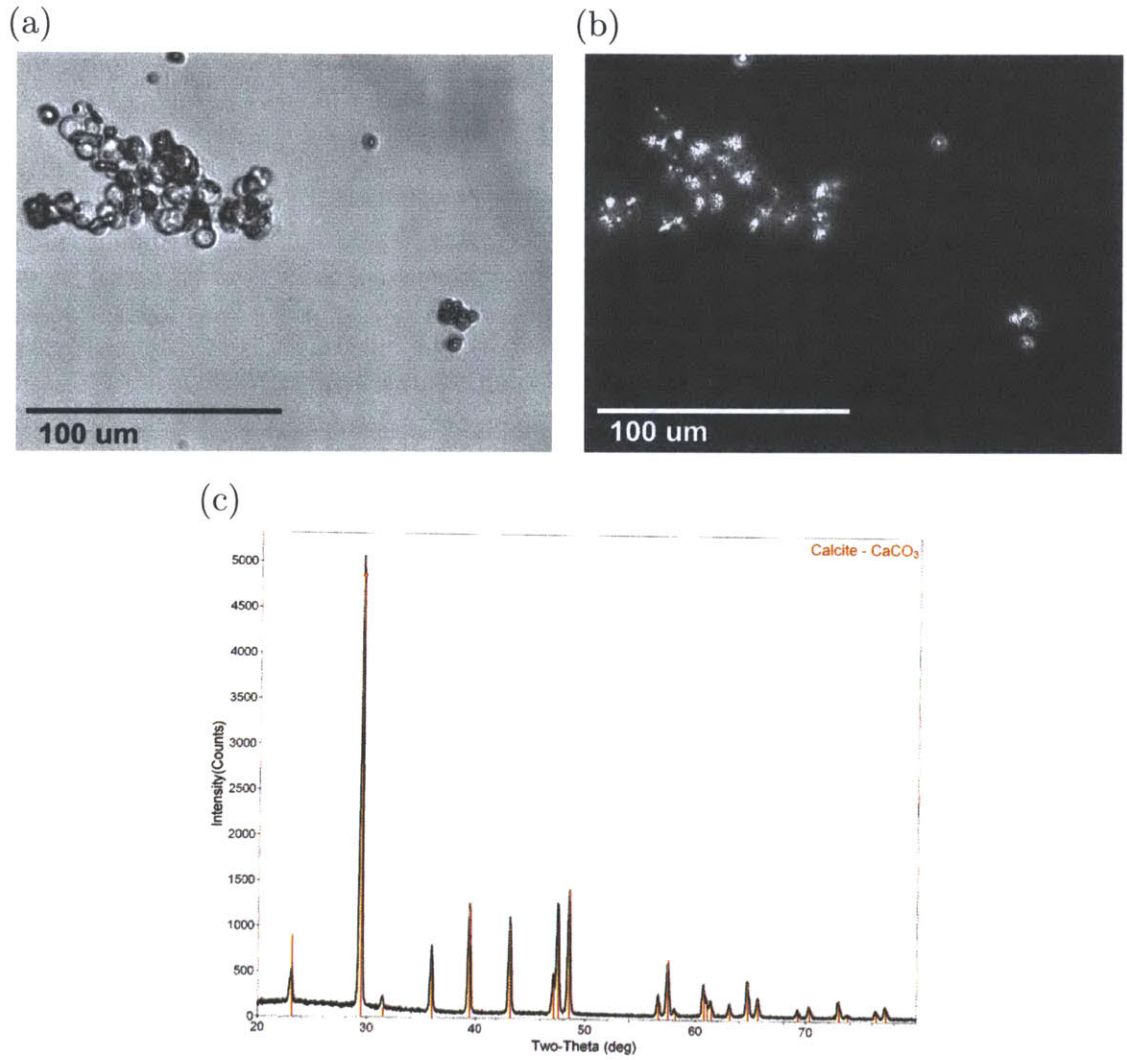


Figure 3.8: *Cross-polarized microscopy and X-ray diffraction confirmation of $CaCO_3$ mineralization* (a) Representative image of the N66-yeast from the mineralization reactor. The dark objects that are attached to the yeast cells are $CaCO_3$. For scale, the average yeast cell is 5 μm in diameter. (b) Cross-polarized microscopy image of the same sample from (a). In this image, only objects that can change the polarization of light (i.e., $CaCO_3$ crystals) are illuminated (c) XRD chromatogram confirms that calcite is mineralized in the presence of N66-yeast.

chromatogram is in Figure 3.8c. Moreover, the peaks in the XRD chromatogram match the peaks from the calcite reference card, with no discernible matching to vaterite or aragonite reference cards (not shown).

XRD analysis was performed on a mineralization reaction with 10X lower concentrations of $CaCl_2$ and $NaHCO_3$ (10 mM instead of 100 mM) using the "no peptide on yeast" sample and the presence of calcite only was confirmed. Interestingly, under those same mineralization conditions, if there was no yeast present in the reaction, the $CaCO_3$ was composed of calcite and vaterite (data not shown). This result is consistent with the well-established role of biomolecules in the stabilization of specific crystal structures in biomineralization. [61]

Using these methods, it was not possible to confirm absolutely whether the nucleation was occurring on the yeast cells, or whether it was occurring in solution and then the peptides were binding the nascent crystals. However, the evidence suggests that it is the former, rather than the latter that is occurring, especially in light of the XRD data from the yeast and the no yeast mineralization reactions. Either way, the critical focus of this project was the rates and the costs of mineralization, rather than the mechanism.

3.3.4 Development of a lab-scale CO_2 mineralization reactor

The next step of this work was to verify that the engineered yeast could be used to enhance the rate of the entire process (i.e., the couple reactions of gaseous CO_2 to mineralized $CaCO_3$). To do this, a lab-scale CO_2 mineralization reactor was designed and built. This reactor was composed of a standard glass laboratory bottle (the volume of the bottle was dependent on the volume of the reaction) with a threaded cap. A custom-made stainless steel cap for the bottle was made that fit the threading on the bottle and that had three ports in the top. Two of the ports accommodated stainless steel tubing that could slide into and out of the port (to adjust the delivery height of the mock flue gas and the capture height of the off gas). Each port had a compression-tightened o-ring to ensure an airtight seal when the stainless steel tubing was in place. The third port accommodated the Ca^{2+}-sensitive electrode. This electrode could slide in and out of the port to adjust the detection height. The port had a compression-tightened o-ring to ensure an airtight seal once the electrode was in place. The metal to glass connection between the cap and the bottle was not airtight, and so petroleum jelly was applied to the threads every time the cap was put in place. The result of this set-up was that all of the flue gas that was delivered to the reactor was collected by the off-gas port and was analyzed by the in-line CO_2 detector.

The temperature of the reactor was controlled using a temperature-controlled stir plate, wherein the thermocouple was placed into a solution of water surrounding the CO_2 reactor. The temperature of the liquid inside the reactor was thus controllable from 25°C-60°C $\pm$ 2°C and was verified to stay at the set temperature for at least 20 hours.

The mock flue gas was a custom-made mixture of 15% CO_2 and 85% N_2. The flow rate of the gas was controlled by an Alicat Scientific mass flow controller that was calibrated for that gas mixture. The % CO_2 in the off-gas was measured using an Alpha Omega Instruments 9510 analyzer with a built-in temperature controller to keep the temperature in the detector higher than the temperature in the reactor (so that condensation of water vapor would not occur in the detector).

For each reaction, the reactor was filled with the appropriate Ca^{2+} source, OH^- source,

and yeast strain, the cap was placed on the bottle, and the system was purged with 100% N_2. The signals from both the Ca^{2+}-sensitive electrode and the CO_2 detector were recorded using an attached computer. At time, $t = 0$, a solenoid valve was used to switch the gas source from the N_2 to the 15% CO_2.

3.3.5 Mineralization in the CO_2 capture reactor

The lab-scale CO_2 reactor was then used to test the effectiveness of the engineered *S. cerevisae* cells in enhancing the mineralization rate of $CaCO_3$ using a mock flue gas as the starting source of carbon. For these experiments, three conditions were tested: 1) no biomolecules in the reactor; 2) soluble bCA2 in the reactor; and 3) bCA2-yeast in the reactor. The results of these experiments are shown in Figure 3.9 and are summarized in Table 3.4. Note that the bCA2-yeast and the soluble bCA2 are both added in excess of the expected required amount so as to not be rate-limiting in the reaction. Both the bCA2-yeast and the soluble bCA2 enhance the mineralization rate of $CaCO_3$ in the reactor by 20%-30% (see Table 3.4). The rates that are reported from this reaction are roughly one order of magnitude slower than those measured in the mineralization rate experiments from Section 3.3.2. This discrepancy is likely due to the fact that the absorption of the gaseous CO_2 in the mineralization reactor is limiting the amount of substrate available for mineralization, which would be avoided in the scaled-up process by the use of a more efficient absorption column.

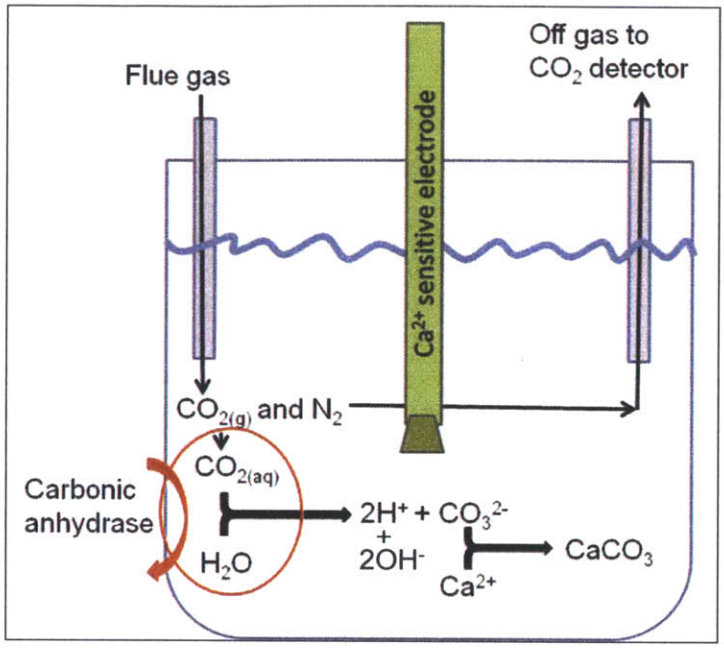
Sample	**Concentration (nM)**	**Mineralization rate ($\frac{M}{sec}$)**	**Temperature (°C)**
No carbonic anhydrase	0	$1.07\text{x}10^{-5} \pm 4.12\text{x}10^{-7}$ (S.E.)	25
soluble bCA2	67	$1.45\text{x}10^{-5} \pm 4.12\text{x}10^{-7}$ (S.E.)	25
bCA2-yeast	67	$1.35\text{x}10^{-5} \pm 4.12\text{x}10^{-7}$ (S.E.)	25

Table 3.4: *Conversion of gaseous CO_2 to $CaCO_3$ in CO_2 capture reactor* The bCA2-yeast and the soluble bCA2 are both added in excess of the expected required amount so as to not be rate-limiting in the reaction. Both the bCA2-yeast and the soluble bCA2 enhance the mineralization rate of $CaCO_3$ in the reactor by 20%-30%.

3.3.6 Analysis of cation and alkalinity sources

The aqueous mineralization of CO_2 requires two material inputs. The first is a divalent cation such as calcium or magnesium and the second is a source of alkalinity (see Table 3.1 and Figure 3.9). There are several possible sources of alkalinity, including reactive mineral deposits, industrial byproducts (such as coal fly ash or paper mill ash), electrochemically generated alkalinity from brines or seawater, or non-reactive mineral deposits. A full review of these sources has been discussed elsewhere. [79–81] To summarize these results, however, it is sufficient to say that there are limited options for the mineralization of significant fraction

(a)



(b)

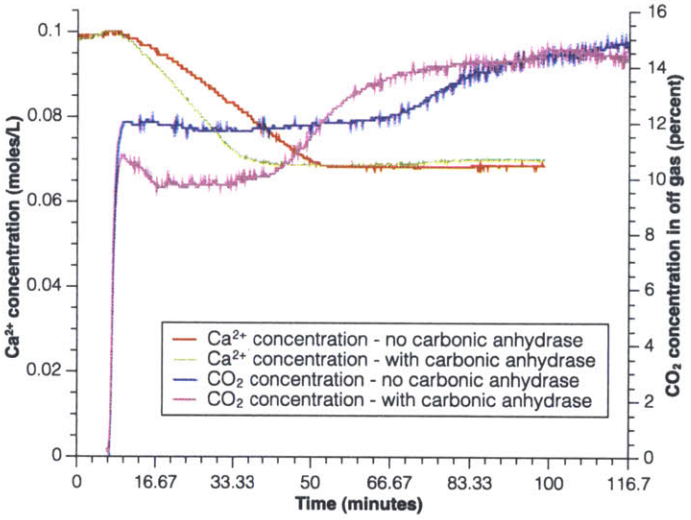


Figure 3.9: *Lab-scale CO_2 capture reactor* (a) A lab-scale CO_2 capture reactor was built that was used to measure the effectiveness of the engineered *S. cerevisae* cells in enhancing the mineralization rate of CO_2 from a mock flue gas. A Ca^{2+}-sensitive electrode was used to monitor Ca^{2+}concentration in the reactor and an in-line CO_2 detector was used to measure the CO_2 concentration in the off-gas. (b) The results of a mineralization reaction with real-time monitoring of the Ca^{2+}concentration and the CO_2 concentration in the off-gas for no enzyme and a soluble bCA2 solution (67 nM bCA2). The measured mineralization rates are reported in Table 3.4

of the global CO_2 emissions. The first of these is injection of CO_2 into underground reservoirs that contain the necessary cations and alkalinity. Another is to produce alkalinity through electrochemical weathering of water, to neutralize the acid generated as part of the process with mineral sources of alkalinity (using non-reactive mineral sources like serpentinite), and then to use the hydroxide to neutralize the acid generated by the CO_2 mineralization process. [80] This is the approach that is being pursued by the company Calera. [81]

While the electrochemical production of alkalinity allows for the capture of CO_2 on a global scale, a concern with this approach is that it adds additional energy demand to the CO_2 capture process, thus increasing the cost of the process. Since process cost is a primary metric for the approach that was designed and evaluated in this work, it was decided to use a less expensive source of alkalinity – alkaline industrial byproducts like fly ash or paper mill ash, even though this means that it must be implemented on a smaller scale.

It was estimated that the fly ash produced from a single 200 megawatt coal fired power plant is approximately 40 $\frac{ktonne}{year}$. Assuming a calcium oxide content of 17% by weight, this amount of fly ash could sequester about 6.5 $\frac{ktonne}{year}$ of CO_2, or roughly 0.5% of the CO_2 generated by the power plant. Of course, more of the CO_2 generated by that power plant could be mineralized if other CaO sources were utilized, but that would require the transport of those materials to the power plant, which would add additional cost to the process. Thus, in order to identify the cost feasibility of the most simple process as designed, fly ash from the power plant is used as the source for both the divalent cations and the alkalinity.

3.3.7 Development of an industrial-scale process model for CO_2 mineralization

Using the results from the laboratory scale experiments, an industrial scale process was designed and evaluated for biologically catalyzed CO_2 mineralization. The mineralization rates and the CO_2 hydration rate parameters that were used for the dimensioning of the mineralization reactor and the CO_2 absorption column were taken from the lab-scale measurements and are summarized in Table 3.5. In order to compare the effects of various process parameters, a base case was designed and the cost of the process was determined including full capital equipment, associate labor, maintenance, etc. The annual cost of the process was then divided by the annual CO_2 capture capacity or the annual solids production ($CaCO_3$ plus other solids, including yeast and inactive materials in the CaO source). These values allowed for the comparison of different processes in terms of dollars per unit of CO_2 captures or dollars per unit of material produced.

The following assumptions were made in order to develop the base-case process model:

1. CO_2 flux $= 6.7 \frac{ktonnes}{year} * 95\%(capacity\ factor) = 5.09 \frac{moles\ CO_2}{second}$

2. The flue gas has a rate of 3000 $\frac{(N)(m^3)}{hour}$.

3. The flue gas has a composition, in percent volume of $N_2 = 73.8$, $O_2 = 3.4$, $CO_2 = 13.7$, $H_2O = 9.1$ and it comes from the flue stack at a temperature of about 85°C and at atmospheric pressure.

	Reaction	Catalyzed by	Uncatalyzed rate parameter	Enhanced rate parameter
1	$CO_{2(aq)} + H_2O \rightharpoonup HCO_3^- + H^+$	bCA2-yeast	$0.045 \frac{1}{second}$ [82]	$6x10^5 \frac{1}{second}$ [a]

	Reaction	Catalyzed by	Uncatalyzed mineralization rate	Enhanced mineralization rate
2	$H^+ + HCO_3^- + Ca^{2+} + 2OH^- \rightharpoonup CaCO_3 + H_2O$	GPA-yeast	$5.2x10^{-4} \frac{moles\ CaCO_3}{(L)(second)}$ [b]	$9.3x10^{-4} \frac{moles\ CaCO_3}{(L)(second)}$ [b]
		No peptide yeast	$5.2x10^{-4} \frac{moles\ CaCO_3}{(L)(second)}$ [b]	$6.3x10^{-4} \frac{moles\ CaCO_3}{(L)(second)}$ [b]

[a] Conservative estimate based on activity level comparisons to soluble bCA2.

[b] From Table 3.7.

Table 3.5: *CO_2 mineralization reactions and associated rates.* Reactions in the CO_2 mineralization process with the biological catalysts for each reaction and the uncatalyzed and catalyzed reaction rates measured in the lab. The single arrow in reactions indicates that the reaction are treated as irreversible under the conditions proposed for this work.

4. A heat exchanger cools the flue gas to 40 °C to reach the optimal temperature for enzyme stability and a fan blows it into the absorber.

5. Coal fly ash is used as the source of CaO.

6. The calcium oxide has an 85% efficiency of CO_2 capture.

7. The fly ash has a composition, in weight percent of $SiO_2 = 50$, $Al_2O_3 = 25$, $Fe_2O_3 = 8$, $CaO = 17$, which corresponds to class C fly ash.

8. The rate of fly ash addition to the process is roughly $5 \frac{ton}{hour}$.

9. Only one strain of yeast is to be used: bCA2-yeast, with a CO_2 hydration rate parameter of $6 \times 10^5 \frac{1}{second}$ and a $CaCO_3$ mineralization rate of $6.3 \times 10^{-4} \frac{moles\ CaCO_3}{(L)(second)}$.

10. The concentration of the bCA2-yeast will be $\approx 1.1 \times 10^8 \frac{cells}{mL}$ in the reactions.

11. Water is added to the system such that the slurry in the absorption column is maintained at a concentration of 7.5% (by weight) solids. The water will have to be added to compensate for evaporation in the absorption column and to compensate for the liquid that exits the process with the solids at the end.

12. The water exiting the process at the end is 15% by weight.

13. Water is added at a rate of 2.5 $\frac{tonnes}{hour}$ with a circulating rate of roughly 70 $\frac{tonnes}{hour}$.

14. After mineralization, there are two solid separation stages: the first is a gravity settling stage that leads to a solid content of 25% (from 8%) and the second is a press filtration

stage that leads to a solid content of up to 85%. Two filter presses are utilized, working in alternate operation. Finally, a dryer is used to increase the solids content up to 95.5%.

15. The bCA2-yeast are not separated from the solids (because the calcium carbonate mineralizes on them), but the water from the settling stage and the filter press stage can be recirculated and re-used. The water from the gravity settling stage is cooled to 35°C before being recirculated to the absorber to prevent overheating of the bCA2-yeast.

16. The spray column absorber allows for a gas flow rate of $\approx 2\ \frac{m}{second}$ with a pressure drop of just a few millibars. The diameter of the spray column (1.2 m) is determined by the speed of the gas flow inside, which is set to a standard rate of $0.75\ \frac{m}{second}$. With a liquid hold-up of 5%, a slurry concentration of 7.5%, and a concentration of 1.1 x $10^{11}\ \frac{cells}{L}$, a 10 m column can absorb the CO_2 at the assumed CO_2 flux.

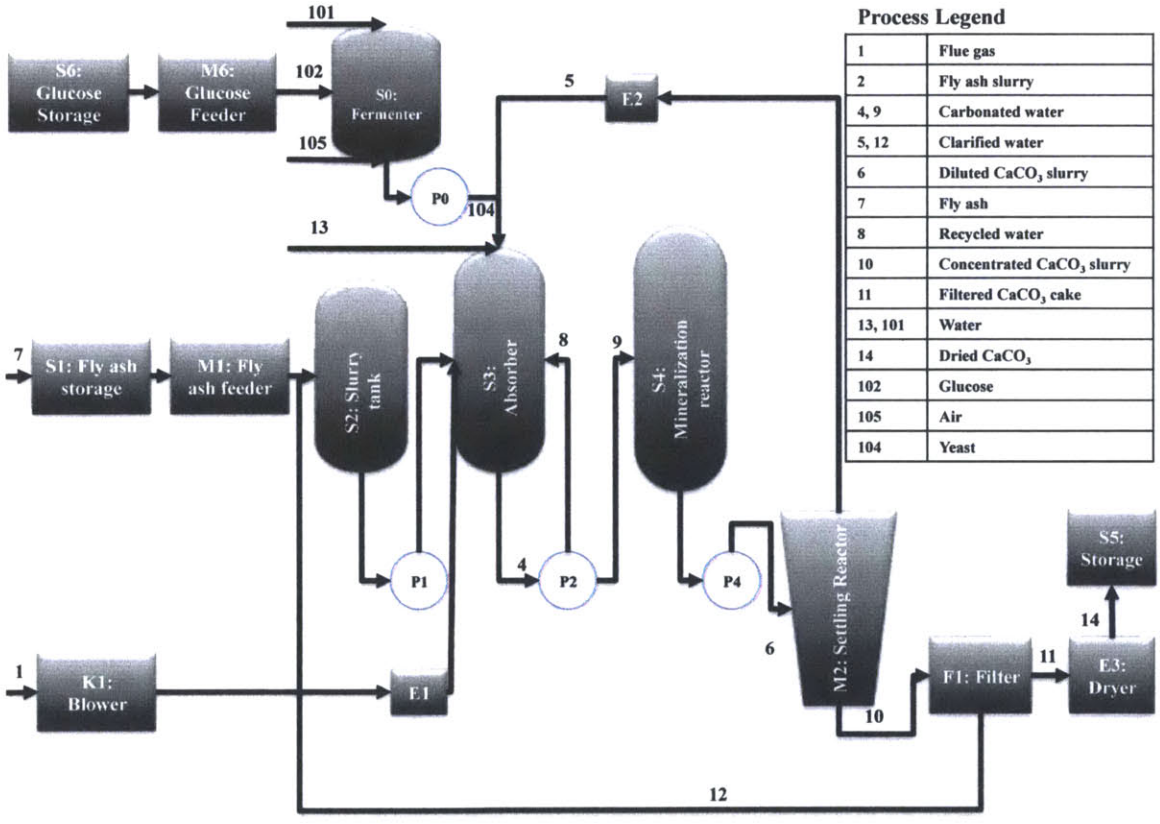


Figure 3.10: *Process diagram for biologically catalyzed mineralization of CO_2 using fly ash as a Ca^{2+} and alkalinity source*

3.3.8 Impact of carbonic anhydrase on absorber column size

Reaction 2 in Table 3.5 can be thought of as a enzymatic reaction, and the Michaelis-Menten rate law can be applied to this reaction. By making a quasi-steady-state approximation for the concentration of enzyme-substrate complex, assuming that the enzyme is not in great excess over substrate ([E] $\approx$ nM, [S] $\approx$ mM), and assuming that [S] $\equiv$ $[S]_0$ (since new substrate is constantly added), the following equation can be derived:

$$v_{catalyzed} = \frac{v_{max}[S]}{K_m + [S]} \tag{3.1}$$

where $v_{max} = (k_{cat})$(moles enzyme) is the maximum reaction velocity, K_m = 8 mM, k_{cat} = 6 x $10^5 \frac{1}{second}$, and [S] = solubility of CO_2 at 40°C = 22 mM. [83, 84] The desired $v_{catalyzed}$ is the same as the CO_2 flux = 5.09 $\frac{moles\ CO_2}{second}$. A rearrangement of equation 3.1 allows one to solve for the moles of enzyme needed to catalyze the reaction in the absorber.

$$Moles\ enzyme = \frac{(v_{catalyzed})(K_m + [S])}{[S](k_{cat})} \tag{3.2}$$

This leads to a value of 1.16 x 10^{-5} moles of enzyme. Next, to determine the volume of the reactor, the concentration of enzyme in the reactor must determined. It was assumed that it was best to work at as high of a concentration of bCA2-yeast as possible, to minimize the reactor volume and the associated captial equipment cost. The assays performed in earlier experiments (e.g., Figures 3.3, 3.4, 3.7) routinely used concentrations of yeast between 2 x 10^6 and 2 x 10^8 $\frac{cells}{mL}$. Thus, a concentration of cells on the high end of this range was selected: 1.1 x 10^8 $\frac{cells}{mL}$ $\approx$ 8 OD_{600}). At this concentration, the yeast solution still behaves like a liquid, but the amount of bCA2 is high enough to have an impact on the process. Assuming 100,000 enzymes per cell, this means that the final concentration of enzyme will be around 27 nM.

The volume of the reactor then can be calculated:

$$Vol_{reactor} = \frac{moles\ enzyme}{concentration\ enzyme} \tag{3.3}$$

which leads to a catalyzed reaction liquid volume of 420 L. This liquid volume necessitates an absorber reactor volume of 1130 L (see Table 3.6).

It is worthwhile to perform the same calculation for the uncatalyzed reaction, to determine the impact that the enzyme has on the reactor volume. To do this, the following equation is used:

$$v_{uncat} = k_1[S] \tag{3.4}$$

Using the uncatalyzed reaction rate parameter value of 0.045 $\frac{1}{sec}$ for k_1, [82] and a value for [S] that is equal to the solubility of CO_2 at 40°C = 0.022 moles per L, [84] gives an uncatalyzed reaction rate of 0.001 $\frac{moles}{(L)(sec)}$. Dividing the CO_2 flux by the uncatalyzed reaction rate gives a required liquid volume of 5090 L, more than 10 times the volume required if using the carbonic anhydrase yeast.

3.3.9 Impact of mineralization rate on mineralization reactor

The calculations to determine the size of the mineralization reactor are based on the mineralization rate of $CaCO_3$ and and the amount of CO_2 to be mineralized. The volume of the reactor can be calculated using equation 3.5

$$Vol_{reactor} = \frac{CO_2 \ capture \ rate}{rate_{mineralization}} \tag{3.5}$$

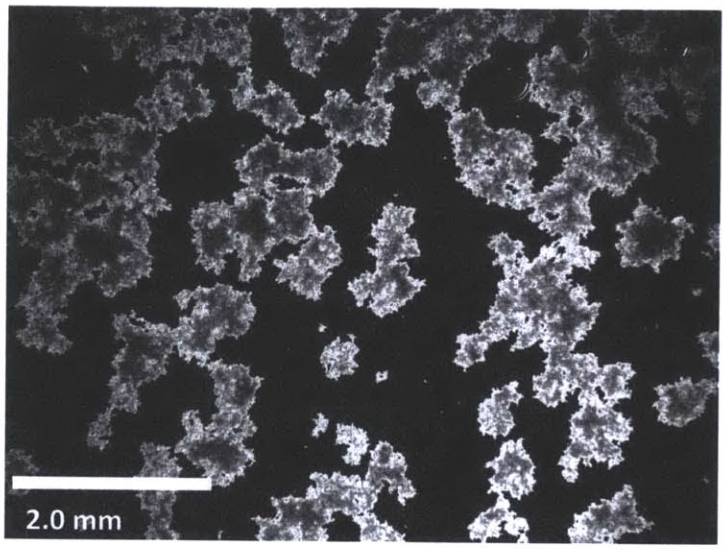
where CO_2 capture rate = (CO_2 flux)(CO_2 capture efficiency of CaO) = 4.33 $\frac{moles}{second}$ and $rate_{mineralization}$ is the baseline rate measured in the laboratory experiments for yeast with no mineralization peptide = 6.3 x 10^{-4} $\frac{moles \ CaCO_3}{(L)(second)}$. Substituting these values into equation 3.5 gives an estimated mineralization reaction volume of $\approx$ 6900 L. If the mineralization rate could be increased by 3x, the reactor volume could be reduced to $\approx$ 2300 L, and if the mineralization rate could be increased by 10x, then the mineralization reactor could be eliminated and the reaction could occur in the bottom of the absorption reactor. It is interesting to note, though, that the mineralization part of the process does not have a significant impact on the process cost (see Figure 3.8). It is also worth considering what the mineralization rate would be in the absence of the mineralization template provided by the yeast cells. To do this, the uncatalyzed mineralization rate from the mineralization rate measurement experiments was used = 5.2 x 10^{-4} $\frac{moles \ CaCO_3}{(L)(second)}$ (see Table 3.5 and Figure 3.7). Substituting this value into equation 3.5 gives an estimated reaction volume of $\approx$ 8300 L or 20% larger than the catalyzed reactor volume.

3.3.10 Impact of yeast cells on settling reactor

The presence of yeast cells in the mineralization reactor results in much larger particles than when the CO_2 is mineralized without the yeast cells (see Figure 3.11). Dozens or hundreds of yeast cells interconnect with dozens or hundreds of calcium carbonate crystals (see the top image in Figure 3.11) whereas the carbonate minerals formed in the absence of yeast cells stay as individual crystals (see the bottom image in Figure 3.11). The average particle sizes for the two conditions were estimated to be 1 mm for the yeast-$CaCO_3$ particles and 25 μm for the individual $CaCO_3$ particles.

While theoretical ways of predicting settling time exist, they are difficult to apply to complex materials of irregular shape, and it is best to measure settling times empirically. The settling time was measured in the lab-scale mineralization reactor by turning off the stirring and then monitoring the amount of time that it took for the materials to settle to the bottom of the reactor. The GPA-yeast in Figure 3.11 settled a distance of 7 cm in approximately 30 seconds (settling time = 14 $\frac{cm}{sec}$) whereas the no yeast sample took more than 4 hours to fully settle (settling time < 0.029 $\frac{cm}{min}$), which leads to settling rate enhancement of almost five hundred fold. Even if it is assumed that the inactive materials in the fly ash are able to increase the settling rate of the $CaCO_3$ in the settling reactor by a factor of 100 (to $\approx$ 0.3 $\frac{cm}{min}$), the difference is still almost fifty fold. This is not an unreasonable assumption, since the particle sizes in fly ash are < 100 μm. This slower settling rate could have a significant impact on the residence time in the settling reactor (and thus the total number of settling reactors necessary to keep the process running continuously).

(a)



(b)

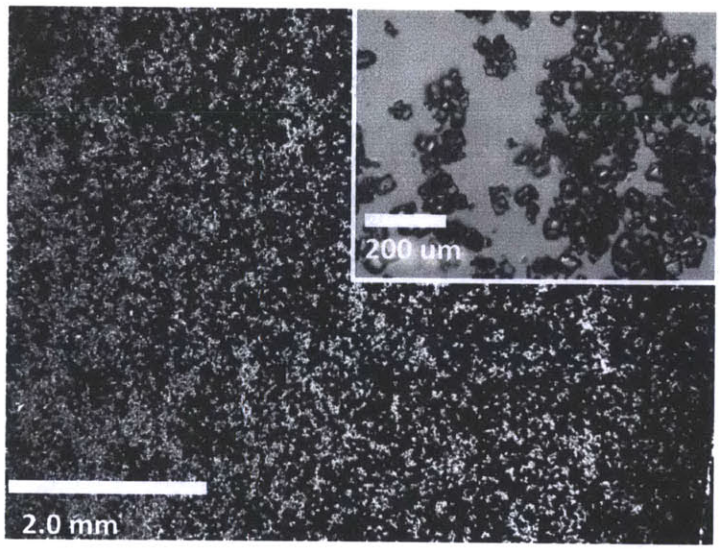


Figure 3.11: *Size of calcium carbonate particles mineralized in the presence and absence of engineered yeast cells* (a) This image shows the typical particle sizes of the yeast-$CaCO_3$ particles when CO_2 is mineralized in the presence of GPA-yeast cells with a final concentration of 6 x 10^{11} $\frac{peptides}{mL}$ or 2 x 10^7 $\frac{cells}{mL}$. (b) This image shows the typical particle sizes of the $CaCO_3$ particles when CO_2 is mineralized without yeast cells. Both reactions took place in the CO_2 capture reactor described in Figure 3.9 with a starting concentration of 0.1 M Tris-base and 0.1 M $CaCl_2$ and with a flow rate of 0.1 $\frac{L}{minute}$ of 15% CO_2.

The height of the settling reactor in the base case is 15 m, which would require a settling time of $\approx$ 110 minutes. If the settling rate is five times slower, then four additional settling tanks, staggered, but running parallel would be required to keep the process running. This would add roughly 2% to the cost per unit of CO_2 captured to the base case.

3.3.11 Production of yeast

In order to determine the amount of bCA2-yeast required for the procees, the following equation is used:

$$yeast\ flux = (Volum.\ rate_{absor.})(\frac{bCA2 - yeast\ cells}{L})(MW_{cells}) \approx 7.5\frac{kg\ DCW}{hour} \tag{3.6}$$

where the volumetric rate in the absorber is $\approx$ 6.98 x 10^4 $\frac{L}{hour}$, the concentration of cells in the absorber 1.1 x 10^8 $\frac{cells}{mL}$ from subsection 3.3.8, and the molecular weight of the cells is 1 x 10^{-12} $\frac{g\ DCW}{cell}$. The bCA2-yeast are grown in fermenters with an assumed yield of 50% dry cell weight yeast per gram of sugar (glucose) and with an assumed growth rate of 0.05 $\frac{kg\ DCW}{(L)(day)}$. The calculated yeast fermentor volume is thus 4.5 m^3.

There are several assumptions underlying the calculations in this section, which are worth a brief discussion. These assumptions deal with some technical nuances relevant to the use of the yeast-display system. The yeast-display system implemented in this work uses plasmid (i.e., non-genomic) DNA to host the Aga2 and carbonic anhydrase genes, while the Aga1 gene is chromosomally integrated. The plasmid is retained in successive generations of cells with auxotrophic pressure. Expression of the Aga1 and Aga2-CA2 genes is under the control of a galactose inducible expression cassette. In practice, this means that minimal media must be used to retain the plasmid. Even with this selection pressure, however, only 40%-60% of the cells in a population retain the plasmid, presumably because the missing nutrient is secreted into the media at a sufficient concentration by the cells with the plasmid so as to support the cells without it. In the implementation proposed for the final process, however, the Aga2-CA2 gene could also be genomically integrated, thus avoiding the problems associated with the use of the plasmid DNA. [85]

3.3.12 Comparison of alternative sources of CA2

There are alternative ways to procure a suitable quantity of carbonic anhydrase to catalyze the CO_2 to HCO_3^- reaction in this process and it is thus important to compare the costs of those approaches to the cost of the approach proposed in this work. The amount of carbonic anhydrase that is required can be calculated using the following equation:

$$CA2\ flux = (Volum.\ rate_{absor.})(\frac{cells}{L})(\frac{CA2}{cell})(\frac{mole}{CA2})(\frac{grams\ CA2}{mole}) \approx 38\frac{g}{hour} \tag{3.7}$$

where the volumetric rate in the absorber is $\approx$ 6.98 x 10^4 $\frac{L}{hour}$, the concentration of cells in the absorber 1.1 x 10^8 $\frac{cells}{mL}$ from subsection 3.3.8, the number of enzymes per cell is 100,000 (see Figure 3.2), and the molecular weight of CA2 = 30 kDa.

The bCA2 that was used during the development of this process was purified from bovine red blood cells and is available for purchase at a cost of $\approx \$1.25\ \frac{1}{mg}$. Assuming that economies of scale could lower this cost by 100X (and ignoring the number of cows that would need to be regularly bled to provide 38 grams of enzyme per hour), this would still result in over $4 million per year.

Another way to produce the necessary amount of enzyme would be to recombinantly produce CA2 in *E. coli.* [86] A previous report describes a process for the recombinant production and purification of 3 $\frac{kg}{year}$ of heparinase in *E. coli* for a cost of $\approx \$1.5$ million. [87] Assuming that this process could produce 100X the amount of carbonic anhydrase as it could heparinase without an increase in the capital equipment costs (300 kg vs. 3 kg per year), then the direct effect of the capital equipment costs for the *E. coli* CA2 vs. the bCA2-yeast can be made by substituting the equipment costs from the *E. coli* process for the equipment costs for the *S. cerevisae* process. The equipment list and associated costs (in 1998 dollars) were placed into the capital equipment costs spreadsheet for the CO_2 biofixation process, which resulted in an increase of 13.8% in the CO_2 capture cost compared to the bCA2-yeast. The resulting CO_2 capture costs are reported in Table 3.8.

3.3.13 Evaluation of process costs for base case and alternatives

The ultimate goal of this work is to determine the cost to capture CO_2, and to determine the cost to produce $CaCO_3$ for downstream applications. Using the process designed in Figure 3.10 as the base case, Aspen Process Economic Analyzer V7.2.1 was used to determine the capital equipment costs. For special items not listed in Aspen Process Economic Analyzer, quotes were requested from suitable vendors. The purchase costs of the necessary equipment for the base case are listed in Table 3.6. Additional infrastructure and set-up costs (including piping, steel, instrumentation, electrical, insulation, paint, and installation labor) were added to determine the direct costs associated with this process. Indirect costs, such as general and overhead costs, engineering costs, contract fees and contigencies were included in the final cost. Finally, an additional 20% of the total direct and indirect costs was added for the integration of this process into an existing power plants infrastructure, which brings the total capital cost to $9.04 million.

The operating costs were calculated based on the amount of electrical energy, the cooling water, the well cooling water, air, and glucose necessary to operate the process. The operating costs are reported in Table 3.7. It was estimated that four additional workers would be needed to operate the system, that annual maintenance would be 3% of the capital costs, and that an additional 25% of labor and maintenance costs would be required for indirect costs. The capital equipment costs for a plant depreciation of 15 years and a conservative estimate of return on capital employed of 9% were calculated. The costs per tonne of CO_2 captured and per tonne of $CaCO_3$ were then calculated by adding all of these costs. For the base case, the vast majority of the capture costs are due to the cost of capital (almost 65%), while maintenance and labor contribute 14% and 12%, respectively. Indirect costs and operating expenses are 7% and 2%, respectively. The results are reported in Table 3.8. Several parameters in the process were varied in order to evaluate the impact that those parameters would have on the final costs.

To see what effect an enhanced mineralization rate would have on the process costs,

Item ID	Equipment	Size	Purchase cost (thousand $)	Direct cost (thousand $)
S0	Fermenter	3.83 m^3	48	157
S1	Fly ash storage	884 m^3	188	275
S2	Slurry tank	28.3 m^3	89	203
S3	Absorber	11.3 m^3	31	164
S4	Mineralization reactor	17.7 m^3	58	161
S5	Product storage	180 m^3	0	166
S6	Glucose storage	5 m^3	36	224
M1	Fly ash feeder (conveyer belt)	81.7 $\frac{tonne}{hour}$	45	128
M2	Settling reactor	577 m^3	86	276
M6	Glucose feeder (rotary feeder)	25 $\frac{kg}{hour}$	4	7
P0	Yeast pump	0.1 $\frac{L}{second}$	6	25
P1	Circulation pump	7 $\frac{L}{second}$	7	39
P2	Recirculation pump	40 $\frac{L}{second}$	23	110
P4	Diluted $CaCO_3$ slurry pump	20 $\frac{L}{second}$	9	44
E1	Flue gas cooler (heat exchanger)	3.54 x 10^4 $\frac{kcal}{hour}$	14	72
E2	Clarified water cooler (heat exchanger)	2.39 x 10^5 $\frac{kcal}{hour}$	13	69
E3	Plate dryer	4700 $\frac{L}{hour}$	1875	1875
F1	Filter press	25 $\frac{m^3}{hour}$	852	852
K1	Flue gas blower	3000 $\frac{m^3}{hour}$	5	15

Table 3.6: *Equipment sizes and costs for base case* Equipment costs were determined using Aspen Process Economic Analyzer V7.2.1 where applicable. For equipment not available in Aspen Process Economic Analyzer, estimated costs were provided from equipment vendors. Costs are in US dollars for a US location in 2011.

	Units	$\frac{Units}{kg\ CaCO_3}$	$\frac{\$}{Unit}$
Electricity	Kwh	0.0138	0.068
Cooling water	m^3	0.0006	0.0005
Well cooling water	m^3	0.0058	0.0007
Air	Nm^3	0.0037	0
Glucose	kg	0.0025	0.112

Table 3.7: *Operating costs for base case*

mineralization rates of 3X = 1.9 x 10^{-3} $\frac{moles}{(L)(second)}$ and 10X = 6.3 x 10^{-3} $\frac{moles}{(L)(second)}$ were evaluated in the process design. A 3X mineralization rate allows for the reduction of the mineralization reactor size and a 10X mineralization rate allows for the complete elimination of the mineralization reactor. Interestingly, though, these improvements have little impact on the cost of the CO_2 captured (reducing it from \$399 per tonne to \$386 per tonne for the 10X rate). Next, the impact of culturing the yeast on the CO_2 capture cost was evaluated. The equipment and associated costs for culturing the yeast were removed from the process costs and it was assumed that the engineered yeast could be purchased from a third-party vendor for a price of \$2.5 $\frac{1}{kg}$, which is a reasonable price for wild type *S. cerevisae* . This change had about the same effect as the 10X mineralization rate enhancement, reducing the cost of CO_2 capture to \$384 $\frac{1}{tonne}$.

Since capital equipment expenses are the major driver of process cost, the most expensive piece of equipment (the plate dryer, E3; Table 3.6) was removed from the process model. The removal of this piece of equipment results in an end product that has 15% water, but if there is no end-user specifying that the mineralized CO_2 must be dried, then this would be a reasonable change to make to the process in order to reduce the CO_2 capture costs. Removal of the plate dryer reduces the cost of CO_2 capture to \$290 $\frac{1}{tonne}$. Pursuing this approach even further by removing the filter press, which is the second most expensive piece of equipment, provides an additional savings of \$50 $\frac{1}{tonne}$ (data not shown). It should be noted, though, that this would result in significantly more circulation water lost with the product (the filter press increases the solids percentage from 25% to 85%) and it would thus require a larger product holding tank and more replacement water added at the beginning of the process.

The impact that the engineered yeast has on the process costs was evaluated by adjusting the process equipment sizes to handle the slower absorption and mineralization times. This results in an estimated 15% increase in the capital equipment costs and an estimated 50% increase in the electricity consumption for a final CO_2 capture cost of \$433 $\frac{1}{tonne}$.

Finally, two additional sources of CaO were evaluated: paper mill ash from a plant that uses the Kraft process (with higher CaO concentrations than fly ash) and CaO that has been leached from fly ash prior to the addition of the flue gas CO_2. These two processes would hypothetically result in higher concentrations of $CaCO_3$ in the final product, and, of the three sources of CaO, the paper mill ash is most intriguing because of the reduced cost of CO_2 capture. It is important to note that this ash would have to be transported to the CO_2 capture point, and transportation costs have been ignored in these calculations.

	CaO source	**$CaCO_3$ in product (weight %)**	**Cost per tonne of product ($\frac{\$}{tonne}$)**	**Cost per tonne of CO_2 ($\frac{\$}{tonne}$)**	**Cost decrease (increase) per tonne of CO_2 ($\frac{\$}{tonne}$)**	**Percent cost decrease (increase) compared to base case**
Base case (bCA2-yeast)	Coal fly ash	26.7	47.0	399		
3x mineralization rate[a]	Coal fly ash	26.7	46.8	398	1	0.3%
10x mineralization rate[a]	Coal fly ash	26.7	45.5	387	12	3.0%
Externally purchased bCA2-yeast	Coal fly ash	26.7	45.2	384	15	3.8%
No dryer in process[b]	Coal fly ash	21.7	34.2	290	119	27.3%
No biological components (no CA2, no yeast)	Coal fly ash	26.7	51.1	433	(34)	(8.5%)
No biological components (no CA2, no yeast), extra settling tanks	Coal fly ash	26.7	53.4	453	(54)	(13.5%)
Recombinantly produced CA2	Coal fly ash	26.7	53.4	454	(55)	(13.8%)
Base case	Paper mill ash	62.4	79.0	275	124	31.1%
Base case	Pre-carbonation CaO leaching	99.6	170	387	12	3.0%

[a]Hypothetical rate enhancement

[b]Removal of the plate dryer results in 15% water in the final product

Table 3.8: *Process costs for different CO_2 mineralization process designs*

3.4 Conclusions

In this work, engineered *S. cerevisae* was used to enhance the hydration of CO_2 and the mineralization rate of $CaCO_3$. The enhanced rates were verified using quantitative techniques in lab-scale measurements. In a variety of different tests, including pH and temperature stability and CO_2 hydration activity, the bCA2-yeast performed as well as soluble bCA2, which has a reported turnover rate of 1 x 10^6 $\frac{1}{second}$. Several possible mineralization peptides were tested for $CaCO_3$ mineralization rate enhancement. At concentrations below 2 x 10^{12} $\frac{peptides}{mL}$, there was little to no effect on the mineralization rate for the yeast-displayed peptides. At concentrations around 2 x 10^{12} $\frac{peptides}{mL}$, though, there was a roughly 50% increase in mineralization rate, which appeared to be independent of the peptide sequence displayed. Interestingly, even the yeast with no peptide displayed was able to increase the mineralization rate by 20%-30% when tested at the same cell concentrations as the peptide-displaying yeast.

The effect that these enhanced CO_2 hydration and $CaCO_3$ mineralization rates would have on the CO_2 capture cost in a CO_2 mineralization process using coal fly ash as the CaO source was then evaluated. The cost impact of the bCA2-yeast on the process is significant, decreasing the predicted cost of CO_2 capture by 8.5% to 13.5%, when compared to a CO_2 mineralization process with no biological components. Interestingly, this process model also predicts that the use of purified soluble, recombinant CA2 would not provide these same cost savings (only accounting for the difference in production and purification costs and ignoring the mineralization rate and settling rate effects).

If a downstream use can be identified for the product (the mineralized CO_2 and the other solid contaminants introduced with the CaO), then the cost of CO_2 capture in this process would be partially offset by the income received for the sale of the product. For example, for the base case in Table 3.8, 1 tonne of captured CO_2 would generate 8.5 tonnes of product (with 2.27 tonnes of $CaCO_3$ in the product). Even if the product is only sold for \$10 $\frac{1}{tonne}$, the cost of CO_2 capture would be reduced by as much as \$85 $\frac{1}{tonne}$ of CO_2 captured. Moreover, the use of the product in downstream applications would obviate the need to dispose of the coal fly ash in a landfill. Assuming a landfill disposal charge of \$0.035 $\frac{1}{kg}$, this would further reduce the cost of CO_2 capture by \$262 $\frac{1}{tonne\ CO_2}$ if coal fly ash is used as the CaO source, by \$72 $\frac{1}{tonne\ CO_2}$ if paper mill ash is used as the CaO source, and by \$6 $\frac{1}{tonne\ CO_2}$ if pre-carbonation leached CaO is used. Thus, for the base case, the cost of CO_2 capture could be as low as \$52 $\frac{1}{tonne\ CO_2}$ (\$47 $\frac{1}{ton\ CO_2}$) if the landfill non-use credit is applied and if the product is sold for \$10 $\frac{1}{tonne}$. While the total cost of CO_2 capture by this process might be higher than the costs of other techniques for CO_2 capture, it is important to keep in mind that mineralized CO_2 has the advantage of being permanently sequestered in solid form.

3.5 Experimental

3.5.1 Yeast display of enzymes

The cDNA's for bCA2 and hCA2 were cloned into the yeast surface display plasmid pCTCON2 using standard molecular biology techniques. All cloning steps were performed in *E. coli.* Bovine CA2 cDNA in the pCMV-SPORT6 plasmid was ordered from Open Biosystems (clone ID: 7985245; Accession number: BC103260). Human CA2 cDNA in the pDONR221 plasmid was ordered from the Dana Farber / Harvard Cancer Center DNA Resource Core (plasmid ID: HsCD00005312; Refseq ID: NM_000067). The pCTCON2 plasmid was a generous gift from the Wittrup lab. Both CA2 genes contained internal BamHI restriction sites, which were removed using a Stratagene Quikchange Lightning Site Directed Mutagenesis Kit to make them compatible with the yeast display vector, pCTCON2. The genes were PCR amplified off of the plasmids and an upstream NheI restriction site and a downstream BamHI restriction site were added to make them compatible with the pCTCON2 plasmid. The yeast display vector pCTCON2 and the bCA2 and hCA2 PCR products were digested with the appropriate restriction enzymes and the digestion products were ligated into the vector. The CAH gene (from *S. thermophilus* LMG 18311 (Accession number: YP_140076) was codon optimized for *S. cerevisae*, then synthesized and cloned into the pCTCON2 plasmid by DNA 2.0. Deletion of the first 35 amino acids of the CAH protein was accomplished through site-directed mutagenesis using the protocol described previously. [88] Correct insertion of the genes of interest into the yeast-display plasmid was confirmed by DNA sequencing reactions prior to transformation of the plasmids into competent EBY100 *S. cerevisae* cells. Transformed cells were propagated in SD-CAA media. Expression of the hCA2, bCA2, and CAH enzymes was induced by transferring the cells to fresh SG-CAA media at a concentration of 0.5 $\frac{OD_{600}}{mL}$ and growing them for at least 24 hours at 22°C.

3.5.2 Yeast display of mineralization peptides

Mineralization peptides were codon optimized for *S. cerevisae* and were designed with the appropriate restriction enzyme sites for cloning into the pCTCON2 plasmid using the Gene Designer program from DNA 2.0. Oligonucleotides were synthesized using solid phase synthesis by IDT. The oligonucleotides were annealed, phosphorylated and inserted into the pCTON2 plasmid. Correct insertion of the genes of interest into the yeast-display plasmid was confirmed by DNA sequencing reactions prior to transformation of the plasmids into competent EBY100 *S. cerevisae* cells. Transformed cells were propagated in SD-CAA media. Expression of the hCA2, bCA2, and CAH enzymes was induced by transferring the cells to fresh SG-CAA media at a concentration of 0.5 $\frac{OD_{600}}{mL}$ and growing them for at least 24 hours at 22°C.

3.5.3 Measuring yeast display expression levels

Expression levels were measured as described previously. [89] After expression, 0.2 OD_{600} cells were washed three times with 700 μL cold 1X PBS, 5 $\frac{mg}{mL}$ bovine serum albumim, sterile filtered (PBS-BSA). Washing consisted of spinning the cells, removing supernatent,

and resuspending in wash solution. The cells were then resuspended in 100 μL of a 1 to 200 dilution of AlexaFluor 488 conjugated antibody (mouse anti-cmyc 9B11 or mouse anti-HA 16B12) diluted in PBS-BSA. The cells were incubated for 30 to 60 minutes on ice. The cells were washed twice with PBS-BSA and were resuspended in 750 μL PBS-BSA. Cells were analyzed on a BD FACScan flow cytometer.

3.5.4 Calculating yeast display expression levels

A Quantum Simply Cellular anti-Mouse IgG kit was used to quantitate the levels of expressed c-myc or HA epitopes on the yeast cells according to the directions in the kit. Briefly, four populations of beads, each with a different level of antigen binding on it were stained with the same AlexaFluor 488 conjugated mouse antibody that was used to label the yeast cells. A fifth, unstained population of beads was added after staining and the beads were analyzed on a BD FACScan flow cytometer using the same settings and on the same day as the yeast cells. A calibration curve was established by plotting the antibody binding capacity versus the peak channel for each population of stained beads. The Bang's Lab calibration file was used to provide a calibration curve.

The auxotrophic marker that is used to maintain selective pressure on the cells to retain the pCTCON2 yeast-display plasmid is not 100% effective. Thus, in any population of EBY100 cells that have been transformed with the pCTCON2 plasmid, some fraction of the cells do not express and display the protein of interest. Moreover, within the positive population of cells, there is a range of expression levels. This, in turn, affects the total numbers of proteins for the entire population of cells. To account for this for this, the flow cytometry data was processed by first calculating the percentage of cells that was expressing the full-length protein or peptide. The geometric mean of the positive population of cells was used to determine the average expression level per cell using the calibration curve from the Bang's lab beads. The "number of proteins per population of cells" was then calculated by multiplying the "percent expressing" by the "average expression level." If a positive population had a high and a low expression population, they were treated as separate populations and the sum of these was the number of proteins per population of cells.

3.5.5 Carbonic anhydrase activity assay

To test the activity of carbonic anhydrase isoforms, a modified version of the Wilbur-Anderson method was used. [72] First the cells expressing the carbonic anhydrase were washed three times in distilled water. The cells were then resuspended in fifteen mL of the chilled assay buffer. Carbon dioxide gas was bubbled through a flask of stirred distilled water that was packed in ice for at least one hour to reach CO_2 saturation in the liquid (roughly 67 mM CO_2 at 3°C). The reaction buffer, Tris-HCl, pH 8.8 (0.02 M) was chilled on ice for at least 1 hour. A pH electrode was 3-point calibrated at 3°C. For the blank, fifteen mL of the Tris-HCl was added to a clean glass beaker with a small stir bar sitting in an ice bath. For the soluble carbonic anhydrase experiments, the enzyme was added to the buffer. For the carbonic anhydrase displaying yeast cells, fifteen mL of the cells in assay buffer was added to the clean glass beaker with a small stir bar sitting in an ice bath. The pH electrode

was added and allowed to reach equilibrium. Using the RS232 port, the data was collected and saved onto an attached computer. Once equilibrium was reached, 10 mL of the CO_2 saturated water was added. The pH was recorded every 1.5 seconds by the pH meter and saved to the computer. The raw pH values were converted to delta pH values (max pH - pH at time = t). The delta pH values were converted to percent reaction complete values (e.g., 0% for $t = 0$ to 100% when the pH reached equilibrium).

3.5.6 Calculating activity of carbonic anhydrase

Using the data collected from the carbonic anhydrase activity assay, carbonic anhydrase activity was calculated using a modified version of Wilbur-Anderson units. [72] The units of activity per nM of enzyme were calculated normalizing to the blank reactions

$$(T_{blank\ average} - T_{sample\ average})/(T_{sample\ average}) \tag{3.8}$$

where $T_{blank\ average}$ is the average of the times that it took the blank reactions to reach 50% completion and the $T_{sample\ average}$ is the average time it took the sample reaction to reach 50% completion. Fifty percent completion was chosen because this was the point at which some of the reactions reached an inflection point (see Figure 3.3). This inflection point was probably due to the weaker buffering capacity of tris buffer below pH 7.5.

3.5.7 Carbonic anhydrase stability studies

For the temperature stability studies, carbonic anhydrase displaying yeast or soluble carbonic anhydrase was resuspended in assay buffer either at the concentration to be used in the final assay (in the case of yeast displayed enzyme) or to a concentration of 0.1 $\frac{mg}{mL}$ (for the soluble carbonic anhydrase). The samples were then stored at 4°C or in a 50°C oven in sealed containers (to avoid evaporation of the buffer). For the pH stability studies, the samples were resuspended in buffers of the appropriate pH and were stored at 4°C in sealed containers. Prior to use in the assay, the samples were then washed and resuspended in cold assay buffer. The buffers used for the pH studies were: 65 mM Phosphate-Citrate buffer, pH 3.0; 72 mM phosphate-citrate buffer, pH 5.0; distilled water, pH 6.0; 81 mM phosphate-citrate buffer, pH 7.0.

3.5.8 Mineralization rate measurement

The mineralization rate of calcium carbonate was measured using calcium sensitive electrodes. The probes were calibrated using three standard solutions of 0.01 M, 0.05 M, and 0.1 M calcium chloride. Then 7.5 mL of 0.2 M $CaCl_2$ was placed in a 50 mL falcon tube with a magnetic stir bar and the tube was placed on a magnetic stir plate. 0.3 mL of calcium ionic strength adjuster (Orion 932011) was added to the solution and the stirring was set to 850 RPM. The yeast cells were prepped by washing several times with distilled water and then either 180 OD_{600} or 18 OD_{600} of cells were resuspended to a final volume of 7.7 mL (for the 2 x 10^8 or the 2 x 10^7 $\frac{cells}{mL}$ samples, respectively). The calcium ion electrode was added to the sample and allowed to equilibrate. Then 2.5 mL of 0.1 M $NaHCO_3$ was added using a syringe

pump set to inject at 5 $\frac{mL}{min}$ and the change in calcium ion concentration was measured. Using the RS232 port, the data was collected and saved onto an attached computer. The slope of the change in calcium ion concentration was calculated and converted to a value with units of $\frac{moles}{(L)(sec)}$. Each peptide and concentration was measured in duplicate and values reported in Figure 3.7 are averages of these measurements. The sample was then saved and analyzed by light microscopy and by x-ray diffraction, to confirm presence of calcium carbonate.

3.5.9 Cross-polarized microscopy and X-ray diffraction analysis of crystals

Samples for microscopy were prepared by pipetting a small aliquot directly out of the mineralization reaction onto a glass slide and covering with a glass coverslip. Cross-polarized microscopy was performed using orthogonal polarizing filters before and after the sample. XRD samples were prepared by taking $\approx$ 50 mL out of a reaction, centrifuging for 10 minutes it to pellet the sample, and removing the supernatent by decanting. If necessary, the sample was resuspended in a small amount of distilled water. The sample was pipetted into an x-ray diffraction (XRD) sample holder. The sample was placed at 80°C to dry and was analyzed on a Rigaku Powder Diffractometer using a 10 minute scan. Crystal structure was verifed by comparing to appropriate standard reference cards.

3.5.10 Process modeling

Microsoft Excel was used to calculate the mass balances at each process stage for the various species involved in the CO_2 capture process, including nitrogen, oxygen, carbon dioxide, the components of fly ash (or paper mill ash) water, calcium oxide, yeast, glucose, and the mineralized carbonate product. These mass balances were used to determine the minimum process equipment sizes. Those dimensions were used to determine the appropriate equipment dimensions, the associated capital costs, and the associated direct, indirect, and operating costs for the process in Aspen Process Economic Analyzer V7.2.1. These costs were then aggregated and converted to a cost per tonne of CO_2 captured or per tonne of product produced based on the CO_2 flux and the process production rates.

Chapter 4

Optimization of M13 bacteriophage pVIII library construction process for histidine-biased libraries

4.1 Summary

Biased peptide libraries are useful for focusing library screening efforts on peptides with predefined characteristics. In this work, the aim was to develop histidine-biased libraries on the major coat protein (pVIII) of M13 bacteriophage (M13 phage, or phage), as histidines are critical in many biological processes. Histidines can effectively coordinate first row transition metals like zinc, copper, and nickel. Histidines have a pKa that is very close to physiological pH and they play a critical role in the active sites of many enzymes. [90] The imidazole ring on histidines can also form π-stacking interactions. [91] To this end, a vastly improved M13 phage pVIII library construction protocol was developed, which led to the construction of high-quality, histidine-biased libraries. The improvements included optimization of several cloning steps and, after a quantitative analysis of the impact that amplification time had on library quality, the selection of a 3-hour amplification time. This protocol was then used to create libraries with histidines in positions 3 and 5 or 4 and 6 in the N-terminal peptide insert of pVIII as well as a histidine mutation at the twenty-third residue in the pVIII backbone.

4.2 Introduction

The pVIII protein is 55 amino acids long and forms an alpha helix (see Figure 4.1a). There are 2700 copies of the pVIII protein arranged in a helical array around the M13 genome. Figure 4.1b shows a reconstruction of the likely arrangement of 60 of the pVIII proteins based on TEM images of full M13 phage, with the pVIII from Figure 4.1a still highlighted in red. [92] Each pVIII protein comes into contact with three other pVIII proteins, and the overall arrangement is very regular. The first 28 amino acids of pVIII (those that are most solvent accessible) are listed in Figure 4.1c. To make a pVIII library, residues 2–4 (EGD) are replaced with a library of 8 amino acid peptides. Those 8 amino acids are referred to hereinafter as 1 through 8. After those 8 amino acids, the wild-type pVIII sequence resumes

with an aspartic acid (D_5) and the subsequent residues in the pVIII protein are referenced using their position in the wild-type sequence (e.g., Q_{15}). A single phage sequence with three histidines in it (two in the peptide insert and one in the backbone) would be called xxHxHxxx-23H, where the x's represent any amino acid and the 23H denotes the presence of a G23H mutation in the backbone. A pVIII library with the third and fifth positions locked as histidines and with a G23H mutation in the backbone would then be called a His3,5,23 library.

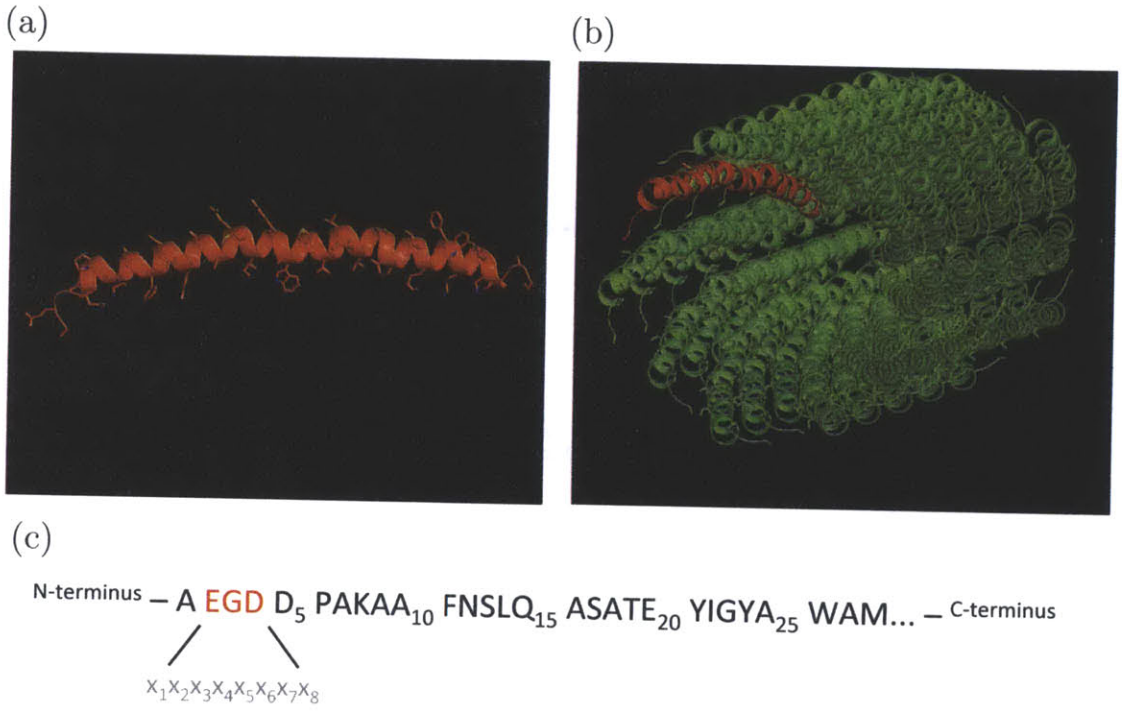


Figure 4.1: *M13 phage pVIII protein structure and sequence* (a) The crystal structure of a single pVIII protein (from PDB file 1IFJ) with the N-terminus on the left and the C-terminus on the right. The pVIII protein forms an alpha helix. [92] (b) A reconstruction of 60 pVIII proteins of M13 viral capsid based on TEM images of M13 phage. [92] The pVIII protein from (a) is highlighted in red. (c) The first 28 amino acids of wild-type pVIII protein. The three residues that are replaced with the 8 amino acid peptide library are highlighted in red. The numbering convention in this scheme assigns 1 through 8 to the peptide library and 5 to 28 to the rest of the pVIII residues.

4.3 Results

4.3.1 Improvement on pVIII library construction process

Previous work in the Belcher lab led to a greatly improved pVIII library construction process. [93] Improvements included the use of longer oligonucleotides for easier agarose gel

purification and some improved DNA purification steps, which led to pVIII libraries that were sufficient for identification of a carbon-nanotube-binding pVIII sequence. [94] Despite these improvements, there were several characteristics of this improved process that limited its use for the construction of histidine-biased libraries. Wild-type M13 phage prefers to have negatively charged amino acids on the N-terminus of the pVIII protein and positively charged residues on the C-terminus to interact with the negatively charged DNA enclosed in the viral capsid. The replacement of negatively charged residues on the N-terminus with positively charged residues affects viral capsid assembly and growth rate. M13 pVIII libraries with large numbers of histidines are slow growers and are thus prone to contamination by fast growing M13 sequences.

The cloning techniques used in library construction also make it very difficult to eliminate completely the presence of the parental M13 sequence. If the parental sequence has a faster growth rate than the library sequences (for example, if the parental has more negative residues on the N-terminus than the library), then the parental sequence will overgrow the library sequences and reduce the library diversity. Finally, M13 phage resists modifications made to the pVIII sequence such that it will make mutations or truncations to the 8-mer peptides added on the N-terminus in order to return to a sequence that is closer to wild-type. The more time that a population of engineered-pVIII M13 phage spends in a growth environment (i.e., during *E. coli* amplification), the more likely those phage are to return to a wild-type-like pVIII sequence.

With those limitations in mind, additional improvement to the pVIII library construction protocol were introduced to increase the library size and quality. These improvements are listed and discussed below. A general overview of the library construction process is shown in Figure 4.2. The full protocol is written out in Section 4.5.5. Table 4.1 summarizes the average yields of each step in the process for four libraries made using the final protocol.

- **Full randomization of position 1 provides greater sequence diversity** The previous library construction protocol resulted in amino acid 1 of the peptide insert being limited to five residues. This was an artifact of the restriction enzymes that were used in earlier protocols. However, digestion with BamHI and BspHI results in restriction sites that are far removed from the library insert position, meaning that the first position can be fully randomized, instead of being restricted to only five amino acids. This increased the theoretical diversity of the library by a factor of four (from 5 possible residues to 20 in position 1).

- **Selection of a parental sequence with a growth rate that matches that of the library sequences improves library diversity** The results in this work identified the relative growth rates of sequences as a critical parameter. pVIII sequences that led to faster assembly into a full phage ended up dominating a library within just a few hours of amplification. This phenomenon has been identified by others and was described in a mathematical model by Derda, et al. [95] In that model, a 10 minute difference in secretion time between two phage led to a 5 to 1 difference in abundance after just 5 hours of amplification.

  Since it is very difficult to completely eliminate parental sequence from a library, it is important to use a parental sequence that grows more slowly (or at least not faster)

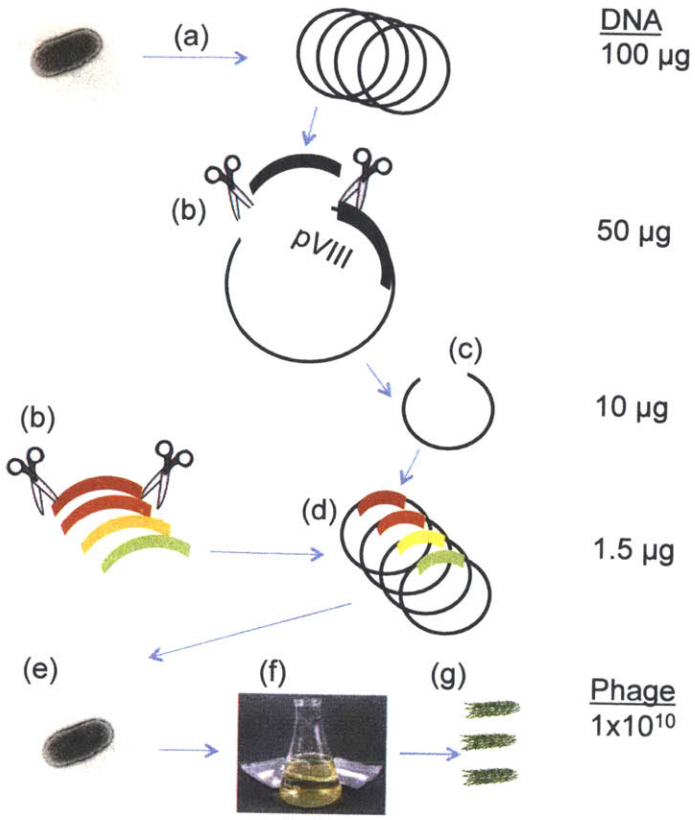


Figure 4.2: *Outline for M13 pVIII library construction* (a) The first step in making a pVIII peptide library is to purify a large quantity of parental M13 DNA from an infected batch of *E. coli*. One hundred μg of DNA is a reasonable amount in order to ensure a sufficient quantity for downstream use. The numbers on the right of the diagram show roughly the quantity of DNA remaining after each step. Table 4.1 shows the actual amounts remaining after each step for the final protocol (average results from four libraries). (b) After isolating the DNA, the parental DNA is digested to remove the pVIII protein region where the peptide will be inserted. The oligonucleotide library that encodes the peptide inserts is digested in parallel with the parental DNA. (c) The parental DNA is purified and dephosphorylated to minimize self-ligation, and (d) the digested oligonucleotide library is ligated into 2 μg parental vector. (e) The ligation reaction is cleaned, concentrated, and transformed into competent *E. coli* cells. After a 45-minute recovery period, a small quantity of infected cells is titered to assess phage concentration (Figure 4.4) and the diversity of peptide sequences (Figure 4.3). At this stage, only one copy of each sequence is present. (f) The final step is to amplify these phage so that at least 10^3 copies of each sequence are present. The bulk of the cells are grown for 3 hours in 400 mL of LB. (g) The amplified phage library is then purified, concentrated and assessed for the total phage amount (Figure 4.5a) and diversity (Figure 4.5b). The full protocol is available in Section 4.5.5.

Step	**Average yield for this step**	**Average yield from start**	**Mass left from 100 μg starting parental vector**
After digestions	56%	56%	56 μg
After gel extraction	30%	15%	15 μg
After dephosphorylation	68%	10%	10 μg
Into ligation (only use 2 μg)	27 %	2%	2 μg
After ligation, purification, and concentration	52%	1.3%	1.3 μg

Table 4.1: *Typical yields for each step in the pVIII library construction process* These averages are from the libraries that were made using the final protocol that is listed in full in Section 4.5.5 (Libraries 6–9). Note that the average yield after dephosphorylation (10 μg) is several-fold higher than the amount of DNA used in the ligation reaction (2 μg). This means that a larger ligation reaction can be performed, which would yield more DNA for the subsequent transformation reactions. Alternatively, the digested and dephosphorylated parental DNA can be used for more than one ligation reaction (i.e., for the construction of multiple libraries in parallel).

than the sequences in the library. Thus, the libraries constructed for this work were built using a triple-His sequence (ESHTHGNE-23H), which is based on M13SK, as the parental sequence. [96]

- **Starting library construction process with a larger amount of DNA (>100 μg) increases library size** As shown in Figure 4.4, the amount of DNA used for transformation affects the size of the library. By starting with 100 μg of parental DNA, the protocol developed in this work delivered at least 10 μg of digested, dephosphorylated DNA ready for ligation, which was 5 times more than needed for the ligation reaction (2 μg). A 2 μg ligation reaction typically yielded around 1.3 μg of DNA for the transformation reactions (see Table 4.1). The results from this work clearly show that the amount of DNA used in the transformation reaction has an impact on the total library size, with libraries of 1 x 10^7 sequences (measured by counting the plaque forming units or PFUs) coming from transformations with $\approx$ 1 μg (see Figure 4.4). These libraries were close to the theoretical maximum for the triple-His libraries (20^5 x 13 = 3.8 x 10^7), which have two His locked in the 8 amino acid peptide insert (see Figure 5.2) and the last residue limited to codons ending with the nucleotide G. Future protocol optimization will include investigation of the impact of doubling the ligation size (to 4 μg), which would yield even more DNA for the subsequent transformations.

- **Optimization of cloning steps, including using a double digest with BamHI-HF and BspHI instead of sequential digestions with BamHI and BspHI, enhances DNA yields** The largest improvement in yield was achieved by combining the BspHI and BamHI digestions in one double digestion reaction instead of performing them sequentially. This was made possible with the introduction of the enzyme BamHI-HF, which has high activity in the same buffer as BspHI (NEB Buffer 4) and the same restriction site specificity of BamHI. This also had the advantage of reducing the digestion time from a total of 7 hours (3 hours for the first digestion, an hour for cleaning and preparing the next digestion, and 3 hours for the second digestion) to 3 hours (3 hours for the double digestion). This reduction in digestion time made it possible to condense all of the DNA manipulation steps (annealing, extending, parental purification from *E. coli*, digestion, dephosphorylation, and ligation) into one day of work (see Section 4.5.5).

- **Using fresh (not frozen) DNA increases DNA yields and library size** The quality of the libraries was improved by emphasizing not freezing any DNA during the library construction process. Based on anecdotal evidence that freezing DNA at any point in the cloning process could reduce the efficiency of subsequent steps (digestions, ligations, transformations), the entire library construction process was coordinated so that naked DNA was never frozen. In other words, the times of each step and the overall process flow were set so that the only overnight stop point during the DNA manipulation steps occurred during an overnight ligation.

  The impact that frozen ligated DNA potentially can have on library size can be seen by looking at the results for Library 9, which has much lower diversity than Libraries 6–8 (see Figure 4.4). Library 9 was made using the exact same process as Libraries

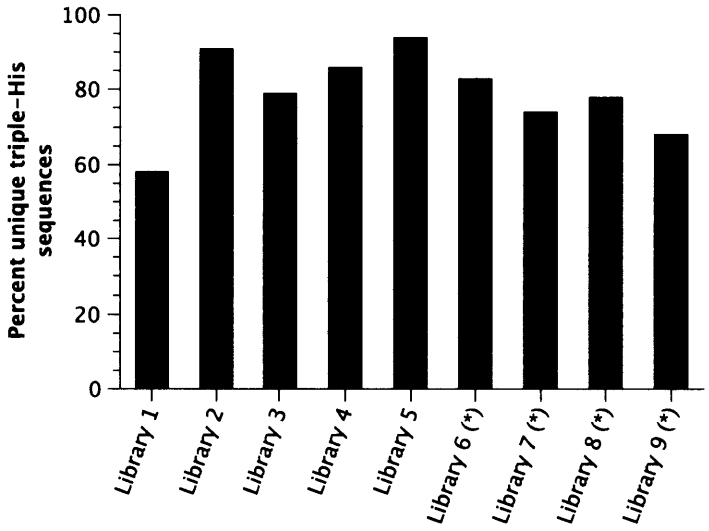


Figure 4.3: *Pre-amplification library quality* This data represents the results of the library quality after step (e) in Figure 4.2. The libraries are listed in chronological order. The percent unique sequences was calculated by dividing the number of unique sequences by the total number of sequences analyzed (15 to 45 sequences analyzed per library). The majority of the non-unique sequences were parental sequences, with the occasional mutation or truncation. The libraries marked with a (*) were all made with the process listed in Section 4.5.5, with the exception that the ligation reaction was stored frozen for 1 week prior to transformation for Library 9. Library process improvements do not seem to have affected the percent unique sequences.

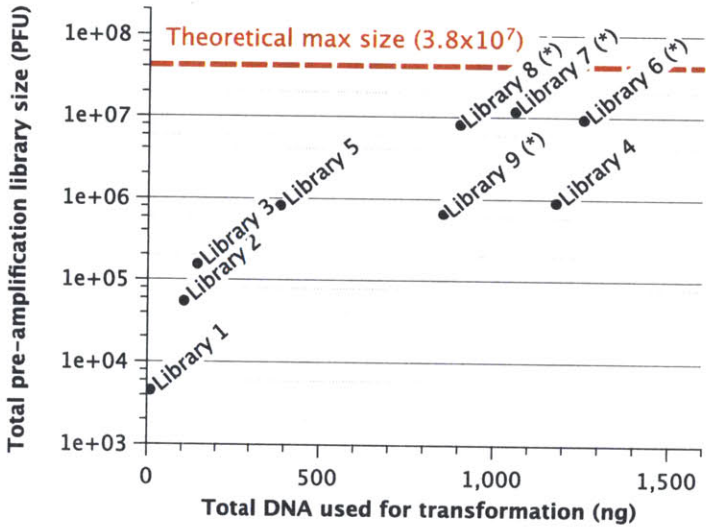


Figure 4.4: *Pre-amplification library size* This data represents the results of the library size after step (e) in Figure 4.2. Total library size was calculated by plating 10 μL of the transformation reaction, calculating $\frac{PFU}{\mu L}$ and multiplying by total transformation volume. The red dotted line indicates the theoretical maximum diversity for a biased pVIII peptide library with 5 fully randomized and 1 partially randomized position (20^5 x 13). The libraries marked with a (*) were all made with the process listed in Section 4.5.5, with the exception that the ligation reaction was stored frozen for 1 week prior to transformation for Library 9. There is a clear relationship between total amount of DNA used in transformation and the total size of the library.

6–8, with the exception that the ligated DNA was stored frozen for 1 week prior to being transformed into *E. coli* for amplification.

- **Reducing amplification time to 3 hours achieves optimal balance of library quality and fold amplification** As discussed above, the relative growth rates of different phage sequences in a population of M13 phage can lead to a faster growing sequence taking over a population. The final step in making an M13 pVIII library is to amplify the library. Prior to amplification, only one copy of each phage sequence is present. Screening of libraries typically calls for $\approx$ 600-fold oversampling in order to account for a 95% chance of sampling each sequence and for the differences in abundances caused by differences in the genetic code. [97] Moreover, given that the time and material resources that are required to create a library are significant, it is desirable to make enough of the library so that it can be screened more than one time. It is reasonable to target 10 screens per library batch made. Using these values in the equation below:

$$\begin{aligned} \textit{Fold amplification} &= (\textit{fold oversampling})(\textit{screens per batch}) \\ &\approx 6\text{x}10^3 \end{aligned}$$

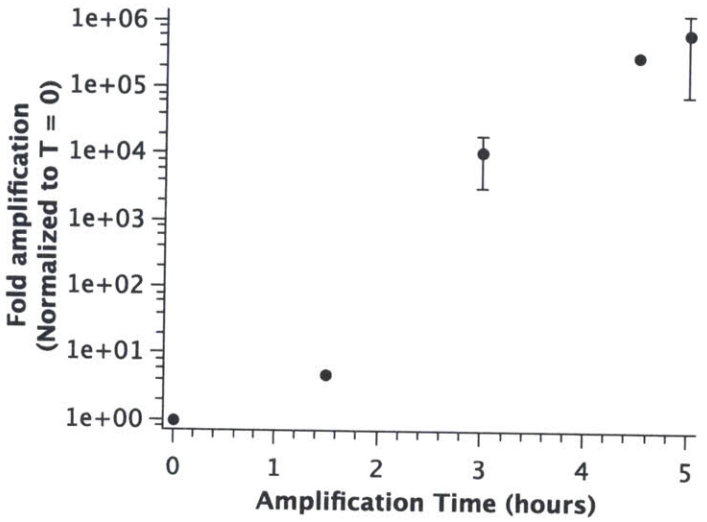
gives a value for the desired fold amplification. Fold amplification is only one component of the final library quality, and any analysis of a library protocol must incorporate the other component – the percent unique sequences.

With these two metrics in mind, nine M13 triple-His pVIII libraries were amplified (step (f) in Figure 4.2). The fold amplification and the percent unique triple-His sequences were calculated at several amplification time points. The results are summarized in Figure 4.5. The original pVIII library protocol called for an amplification time of 5–7 hours. It was obvious for the triple-His pVIII libraries, however, that even a five-hour amplification would be too long, because at five hours the percent unique triple-His sequences was reduced to less than 30%. Even more problematic was that more than 50% of the sequences were truncations and mutations (sequences without three histidines). These sequences grow much more quickly than triple-His sequences and completely take over a phage population in just a few more hours of amplification (data not shown). However, this work demonstrated that a 3-hour amplification provides sufficient fold amplification (average fold amplification$_{3\ hours} = 10^4$) while minimizing the truncations and mutations ($< 10\%$). The increase in the parental sequence at 3 hours is concerning in that it reduces the total diversity of the libraries, but since the parental sequence is a triple-His, it is less concerning than the presence of mutation and truncation sequences.

4.3.2 Histidine-biased libraries with three histidines in every copy of pVIII

As discussed above, there are limitations on the diversity of a pVIII library that are imposed by the biology of M13 phage. Expression of positively charged residues near

(a)



(b)

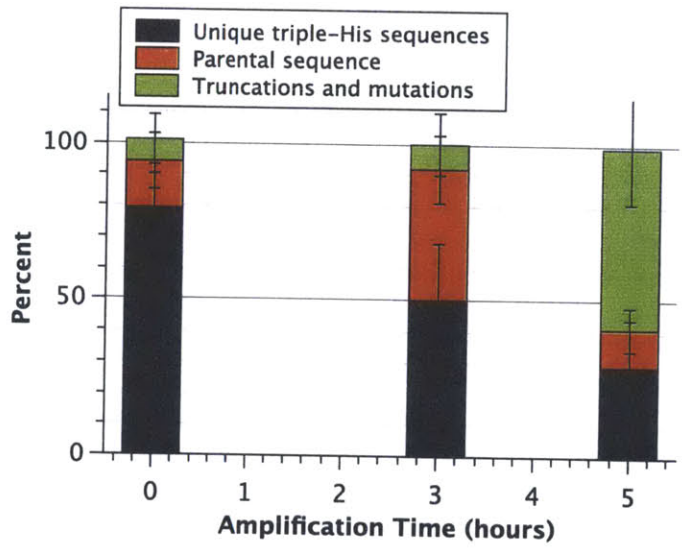
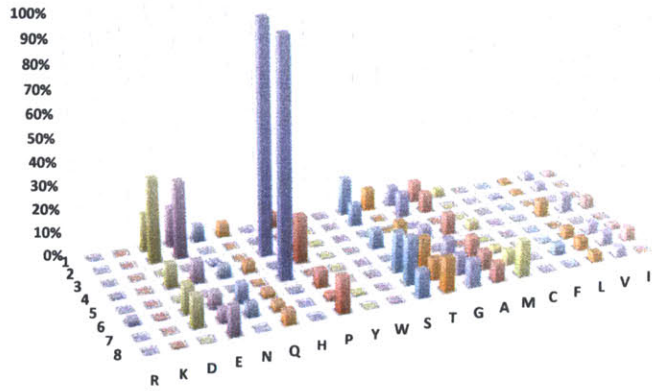


Figure 4.5: *Library amplification time and its influence on fold-amplification and library quality* (a) Fold-amplification was calculated by dividing the library size at each time point by the pre-amplification library size (time = 0). (b) At each time point, phage plaques were sequenced and the quality of the library was assessed by calculating percent unique triple-His sequences, parental sequence, and truncation and mutation sequences. All nine libraries were assessed at 0 hours. Three libraries were assessed at 5 hours, one was assessed at 4.5 hours, five were assessed at 3 hours, and one was assessed at 1.5 hours.

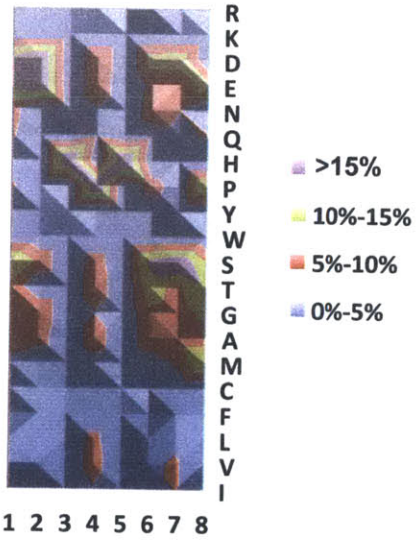
the N-terminus of pVIII is difficult, though this work has shown that it is not impossible. Not surprisingly, nearly every pVIII peptide insert that was sequenced had a net neutral or negative charge. In other words, for every positively charged histidine inserted into the pVIII, there was a charge-balancing aspartate or glutamate. There were also very few positively charged lysine or arginine residues expressed in the histidine-biased libraries. The libraries were also essentially devoid of bulky aromatic residues and cysteines. Figure 4.6a shows the prevalence of each amino acid at each position in the peptide insert (1–8) for 212 clones that were sequenced from several builds of a His3,5,23 library (i.e., this data is a compilation of all His3,5,23 sequences, not just those from a single library).

In addition to limitations on the expression of certain residues in the N-terminal pVIII insert, there also seem to be preferred locations for certain residues. Figure 4.6b is a contour plot view of the same data from Figure 4.6a and Figure 4.6c shows the same analysis for all of the His4,6,23 phage that were sequenced. It appears that aspartates and glutamates are preferentially expressed on the N-terminal end of the insert. This is not altogether surprising, given that an aspartate and glutamate are present in positions 2 and 4 in the wild-type sequence, though one might have expected them to prefer to surround the histidines for charge balancing.

(a)



(b)



(c)

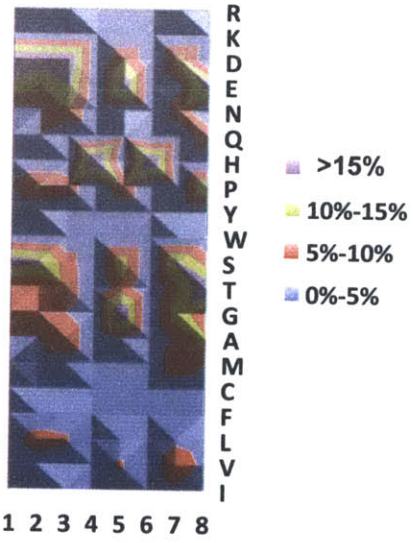


Figure 4.6: *Positional amino acid diversity of triple-His libraries* These triple-His libraries have two His residues locked in either positions 3 and 5 (a) and (b) or 4 and 6 (c) and the third His is located in the pVIII backbone at position 23. The percentages shown in these charts represent the percent of sequences expressing that amino acid at that position. 212 sequences were analyzed for His3,5,23 and 101 for His4,6,23.

4.4 Conclusion

The protocol that was developed in this work was optimized for the construction of histidine-biased libraries. Some of the parameters might not be optimal for non-biased pVIII libraries (such as shorter amplification time), while others are likely to be beneficial for any pVIII library (using a parental sequence that matches in growth rate with the rest of the library). Any new type of pVIII library will need to be evaluated for the diversity and the fold amplification, much as the histidine-biased libraries in this work were evaluated, and all of the factors discussed in this work will need to be considered for future library protocols. It is anticipated that the lessons learned as part of the histidine-biased library optimization will prove useful and informative for future library construction.

Finally, it is worthwhile to discuss the implications of the differential growth rates on downstream use of pVIII libraries. For most applications, the completion of a high diversity library with $> 10^4$ copies of each sequence is only the beginning. The next step for most projects is to perform a biopanning experiment and to identify M13 phage that are displaying pVIII proteins of novel functionality. However, even if the percentage mutant and truncation sequences is low prior to biopanning, they are often still present (or, possibly, they are likely to form during amplification by utilizing *E. coli*'s DNA recombination pathways). After the selection round in biopanning, which might reduce the total number of phage from 10^8 to fewer than 10^4, for example, one must amplify the phage to achieve a sufficient number of each sequence for the next round of selection (see discussion about fold amplification above). For the histidine-biased libraries, this amplification always resulted in a complete overgrowth of mutant and truncated sequences (data not shown) after only one round of selection. Biopanning often requires three to five rounds of selection (with amplification in each round), and this was not possible using normal amplification protocols (even if the amplification time was reduced to three hours). As such, alternative amplification protocols will need to be investigated. Fortunately, one such protocol already exists: the use of microgels for isolation and amplification of individual phage sequences in a massively parallel fashion. [98] Future work with the histidine-biased libraries for biopanning applications will benefit from the use of microgel amplification.

4.5 Experimental

4.5.1 Library primers

The oligonucleotide for the library extension primer has the sequence: 5-TTAAT GGAAA CTTCC TCATG AAAAA GTCTT TAGTC CTCAA AGCCT CTGTA GCCGT TGCTA CCCTC GTTCC GATGC TGTCT TTCGC TGCA -3. The oligonucleotides for the biased libraries are variations on the following sequence: 5'-AAG GCC GCT TTT GCG GGA TC CNN MNN MNN MNN GTG MNN GTG MNN TGC AGC GAA AGA CAG CAT CGG AAC GAG GGT AGC AAC GGC TAC AGA GGC TTT-3'. In this sequence, codons 2 and 4 were locked as histidines, but similar primers were made with codons 3 and 5 or 4 and 6 locked.

4.5.2 Measuring phage concentration with a UV-Vis spectrophotometer

The absorbance of a phage solution was measured on a Nanodrop UV-Vis spectrophotometer (path length = 1 mm). The following equation was used to calculate the concentration:

$$\frac{(A_{269} - A_{320})(6\ x\ 10^{17})}{7222\ bp} = \frac{PFU}{mL} \tag{4.1}$$

for each phage sample.

4.5.3 Media and solutions for phage handling

The recipes for making agarose top, LB medium, IPTG/Xgal/Tet plates, SOC medium, Tetracycline stock, TBS, and PEG/NaCl are available in the NEB Ph.D. Phage Display Libraries Instruction Manual (New England Biolabs).

4.5.4 Measuring phage concentration by titering

To 150 μL of an overnight culture of XL-1 Blue *E. coli* cells, 10 μL of phage sample (or diluted phage sample) was added. Four to five mL of melted agarose top (cooled to < 50°C) was added, swirled vigorously to mix, and the mixture was poured onto a warmed IPTG/Xgal/Tet LB agar plate. Agarose top was allowed to cool, then the plate was placed at 37°C for overnight. After overnight incubation, the number of plaques was counted (each representing a single phage infection event) to determine the number of PFUs (plaque forming units). To calculate concentration, the dilution factor was multiplied by the concentration number of PFUs per μL of phage solution plated.

4.5.5 Library construction protocol

Make 400 mL autoclaved LB in a 2 L flask for day 2.

1. Preparation of pVIII library inserts

   Perform these steps in parallel with miniprepping of the vector DNA.

   (a) Annealing

   - 44 μL TE with 100 mM NaCl (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 100 mM NaCl)
   - 3 μL 100 μM pVIII library extension primer
   - 3 μL 100 μM library oligonucleotides (300 pmoles, 5 μg oligo)

   - *50 μL Total volume*

   Place in the thermocycler with the following conditions. It should take about 30 minutes.

   - 95°C for 8 minutes

- Ramp down to 25°C at a rate of 0.1 $\frac{degrees}{second}$
- Hold at 25°C

(b) Extension

- 50 μL annealed reaction
- 20 μL (10X) Klenow buffer (NEB Buffer 2)
- 8 μL Klenow fragment (5 $\frac{units}{\mu L}$) = 40 units
- 8 μL of 10 mM dNTP's
- 114 μL ddH_2O

- *200 μL Total volume*

Aliquot 50 μL into each of 4 PCR tubes and place in thermocycler with the following conditions.

- 37°C for 20 minutes
- 65°C for 20 minutes

(c) Recombine aliquots and purify DNA.

Add 3 volumes buffer QG (600 μL) and 1 volume isopropyl alcohol (200 μL) to the recombined DNA. Using the DNA Gel Extraction Kit from Qiagen (page 25 in the QIAquick Spin manual), purify the annealed and extended oligonucleotides. Elute the purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

Measure the concentration and calculate the total mass of DNA:

$$48\ \mu L\ x\ \_\_\_\_\_\ \frac{ng}{\mu L} = \_\_\_\_\_\ ng \tag{4.2}$$

(d) Double digest of pVIII library oligonucleotides using BamHI-HF and BspHI

Perform this digest in parallel with the vector digestion.

i. Total reaction volume = 200 μL; 20 μL per 0.5 μg of DNA; only add 5 μg of oligo DNA

ii. Save 4 μL for the negative control (uncut) sample

- _____ μL extended reaction (5 μg DNA)
- 20 μL (10X) NEB buffer 4
- 2.5 μL BamHI-HF
- 5 μL BspHI
- _____ μL ddH_2O

- *200 μL Total volume*

iii. Digest in 1.5 mL tube at 37°C for at least 3 hours. No heat kill necessary.

iv. Using the DNA Gel Extraction Kit from Qiagen (page 25 in the QIAquick Spin manual), purify the digested oligonucleotides. Elute the purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.
Measure the concentration and calculate the total mass of DNA:

$$48\ \mu L\ x\ \_\_\_\_\_\ \frac{ng}{\mu L} = \_\_\_\_\_\ ng \tag{4.3}$$

(e) Agarose gel purification of digested pVIII library oligonucleotides
Perform this in parallel with the gel purification of the digested vector.

i. Make 3% agarose gel by adding 3 grams of agarose to 100 mL of TBE. Add 1.0 μL SYBR Safe (10,000X) after the gel is partially cooled and pour into gel holder ($\approx$ 70 mL for 1 gel). Use the 8 well comb to create wells.

ii. Prepare DNA samples
- Double digested DNA = 44 μL of digested DNA and 9 μL NEB orange loading dye
- Uncut DNA = 4 μL uncut DNA, 6 μL ddH_2O, and 2 μL orange loading dye
- DNA ladder = 1 μL 100 bp NEB ladder, 11 μL ddH_2O, and 2 μL orange loading dye

iii. Load $\approx$ 55 μL per well.

iv. Set voltage to 90V and run for $\approx$ 60 minutes, checking the gel frequently.

v. While gel is running, preweigh 2.0 mL eppendorf tubes for gel extraction.

vi. Use blue light imager to verify that DNA is cut before extracting the digested DNA bands.

vii. Using the DNA Gel Extraction Kit from Qiagen (page 25 in the QIAquick Spin manual), purify the DNA. Do not put more than 2400 μL of dissolved sample on one column. Use Table 4.3 to calculate volumes of QG and isopropanol. Elute the purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

2. Preparation of M13SK vector

Perform these steps in parallel with preparation of library oligo DNA.

(a) Extraction of M13SK vector DNA from *E. coli* cell pellets
Always prepare the M13SK vector DNA on the same day as the digestions. Fresh parental DNA from frozen cell pellets provides better yields in subsequent steps. The frozen cell pellets should have been saved in $\approx$ 200 μL volumes in 1.5 mL tubes from a previous phage amplification (8 hour to 16 hour growth) and the sequence should be previously verified by DNA sequencing (from one tube). One liter of amplified phage should provide between 8 and 16 of these pellets.

i. Prep three tubes of parental DNA. This should provide more than 100 μg of DNA, using QIAprep Miniprep kit (page 22).

ii. Elute the purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

iii. Mix all the DNA samples together, measure concentration, and calculate total mass.

$$\_\_\_\_\_\mu L \; x \; \_\_\_\_\_\frac{ng}{\mu L} = \; \_\_\_\_\_ng \quad (4.4)$$

(b) Double digest of parental vector DNA using BamHI-HF and BspHI

Perform this digest in parallel with the library oligonucleotide digestion.

i. Total reaction volume = 2000 μL; 200 μL per 5 μg of DNA; total of 50 μg of vector DNA

ii. Save 4 μL for the negative control (uncut) sample.

iii. Use Table 4.2 to set up digestion reactions.

Sample	**Volume for 1 reaction (μL)**	**Number of parallel reactions**	**Volume for master mix (μL)**	**Volume for one-half master mix (μL)**
Vector DNA		10		
NEB Buffer 4 (10X)		10		
BamHI-HF	2.5	10	25	12.5
BspHI	5	10	50	25
ddH_2O		10		
Total	200	10	2000	1000

Table 4.2: *Table for calculating pVIII library vector digestion volumes*

iv. Digest in 1.5 mL tube at 37°C for at least 3 hours (preferable 4 hours). No heat kill necessary.

v. Using the DNA PCR Purification Kit from Qiagen (page 19 in the QIAquick Spin manual), purify the digested oligonucleotides. Purify using 5 columns (10 μg DNA per column). Elute the purified DNA from each column in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

Measure the concentration and calculate the total mass of DNA:

$$240 \; \mu L \; x \; \_\_\_\_\_\frac{ng}{\mu L} = \; \_\_\_\_\_ng \quad (4.5)$$

(c) Agarose gel purification of digested parental vector

Perform this in parallel with the gel purification of the digested library oligonucleotides.

i. Make two 0.8% agarose gels by adding 1.2 grams of agarose to 150 mL of TAE. Add 1.5 μL SYBR Safe after the gel is partially cooled and pour into gel holder ($\approx$ 70 mL for 1 gel). Use the 8 well comb to create wells.

ii. Prepare DNA samples
   - Double digested DNA = 240 μL of digested DNA and 48 μL NEB blue loading dye
   - Uncut DNA = 4 μL uncut DNA, 6 μL ddH_2O, and 2 μL blue loading dye
   - DNA ladder = 1 μL 1 kbp NEB ladder, 11 μL ddH_2O, and 2 μL blue loading dye

iii. Load $\approx$ 55 μL per well. (Will probably need 6 wells for parental vector digestion reaction.)

iv. Set voltage to 90V and run for $\approx$ 60 minutes, checking the gel frequently.

v. While gel is running, preweigh 2.0 mL eppendorf tubes for gel extraction.

vi. Use blue light imager to verify that DNA is cut before extracting the digested DNA bands.

vii. Using the DNA Gel Extraction Kit from Qiagen (page 19 in the manual), purify the DNA. Use Table 4.3 to calculate volumes of QG and isopropanol. Do not put more than 2400 μL of dissolved sample on one column. Elute the purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

Sample	**Gel weight (mg)**	**Volume QG (μL)**	**Volume isoprop. (μL)**
Insert			
Parental 1			
Parental 2			
Parental 3			
Parental 4			
Parental 5			
Parental 6			
Parental 7			
Parental 8			

Table 4.3: *Table for calculating pVIII library gel extraction volumes*

3. Calculation of DNA concentrations after digestion and purification

Mix all the parental DNA together and then measure the concentration and calculate the total mass of DNA for the parental vector and the insert:

$$\textit{Digested parental vector}: \_\_\_\_\_ \mu L \; x \; \_\_\_\_\_ \frac{ng}{\mu L} = \_\_\_\_\_ ng \quad (4.6)$$

$$\textit{Digested oligonucleotide insert}: \_\_\_\_\_ \mu L \; x \; \_\_\_\_\_ \frac{ng}{\mu L} = \_\_\_\_\_ ng \quad (4.7)$$

4. Dephosphorylation of vector to prevent self-ligation

   (a) Prepare reaction.

   - 5 μL Antarctic phosphatase (concentration = 5 $\frac{units}{\mu L}$; need 5 units per μg vector
   - 20 μL 10X Antactic phosphatase buffer
   - 175 μL purified vector = _____μL vector

     _____
   - *200 μL Total volume*

   (b) Dephosphorylate for 1.5 hours at 37°C. No heat kill necessary.

   (c) Using the DNA Gel Extraction Kit from Qiagen (page 19 in the QIAquick Spin manual), purify the dephosphorylated vector. Elute the purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

   (d) Measure the concentration and calculate the total mass of DNA for the parental vector and the insert:

$$\textit{Dephosph. parental vector}: \_\_\_\_\_ \mu L \; x \; \_\_\_\_\_ \frac{ng}{\mu L} = \_\_\_\_\_ ng \quad (4.8)$$

5. Ligation of library oligonucleotide insert and vector

   A molar ratio of 1:10 for vector to insert is used and the mass of the library oligonucleotides is 29495 $\frac{g}{moles}$. For 100 ng of vector, 13.2 ng of insert is needed:

$$\frac{(100\ ng\ vector)(95\ bp\ insert)(10\ moles\ insert)}{(7200\ bp\ vector)(1\ mole\ vector)} = 13.2\ ng\ insert \quad (4.9)$$

   (a) Dilute insert 1:10 to reach a workable concentration (90 μL EB and 10 μL insert).

$$\textit{Insert conc.} = \textit{Result from Equation (4.7)}\ x\ 0.1 = \_\_\_\_\_ \frac{ng}{\mu L} \quad (4.10)$$

$$\textit{Volume insert needed} = \frac{13.2\ ng}{\textit{insert concentration}} = \_\_\_\_\_ \mu L \quad (4.11)$$

Sample	Volume for 1 reaction (μL)	Number of parallel reactions	Volume for master mix (μL)
Vector DNA		20	
Insert DNA		20	
T4 DNA Ligase Buffer (10X)	4	20	80
T4 DNA Ligase	2	20	40
ddH_2O		20	
Total	40	20	800

Table 4.4: *Table for calculating pVIII library ligation reaction volumes*

(b) To ligate 2 μg vector use Table 4.4.

(c) Split the reaction into separate PCR tubes with 40 μLin each tube. Incubate the reaction overnight at 16°C.

Start preparation of an overnight of XL1-Blue cells in 25 mL LB-Tet for Day 2.

END OF DAY 1

(d) Using the DNA Gel Extraction Kit from Qiagen (page 19 in the QIAquick Spin manual), purify the ligation reaction. Use 4 columns for the purification. Elute each column of purified DNA in 50 μL of prewarmed EB (microwave for $\approx$ 10 seconds to warm EB to around 50°C). Let the EB sit on the column for three minutes before eluting. These two steps increase the yield of DNA.

(e) Concentrate the vector using Amicon Ultra 0.5 mL 30K spin column.

(f) Mix all four purifications together = 200 μL and add to spin column.

(g) Centrifuge at 14,000 RCF for 10 minutes. This will reduce the volume to $\approx$ 23 μL.

(h) Add 200 μL EB and centrifuge at 14,000 RCF for 10 minutes.

(i) Flip filter over into a new tube and spin for 2 minutes at 1000 RCF. This should provide $\approx$ 23 μL concentrated DNA.

(j) Measure the concentration and calculate the total mass of DNA:

$$\textit{Ligated vector}: \_\_\_\_\_ \mu L \; x \; \_\_\_\_\_ \frac{ng}{\mu L} = \_\_\_\_\_ ng \tag{4.12}$$

6. Transformation of library DNA into electrocompetent *E. coli* and amplification of library
This part takes about 7 hours from start to finish.

(a) The ligation reaction will be transformed in 12 separate transformations.The following is needed:

- 12 14 mL round bottom tubes
- 12 1.5 mL tubes, chilled
- 12 electroporation cuvette, chilled
- 1 kit XL1-Blue electrocompetent *E. coli* (Agilent 200228)
- 60 mL melted agarose top
- 12 mL warm SOC media
- 2.5 mL of XL1-blue cells (grown overnight from previous day)
- 9 warmed IPTG/Xgal/Tet plates

(b) Thaw cells on ice and aliquot 40 μL cells into each of 12 chilled 1.5 mL tubes. Keep on ice.

(c) Add 1/12 of the total ligated DNA (between 1.5 and 2.0 μL, depending on final volume) to each cells and calculate mass of DNA transformed.

$$\frac{Total\ volume = \_\_\_\_\_\ \mu L}{12} = \_\_\_\_\_\ \mu L = \_\_\_\_\_\ ng \tag{4.13}$$

(d) Transform, according to protocol for cells.

(e) Incubate transformed cells for 45 minutes (from the time the last batch of cells went into incubator).

(f) Remove cells from shaker and mix all of them together $\approx$ 10.5 mL.

(g) Take 100 μL from pre-amplification cells and set aside for titering. Finish preparing amplification reaction before returning to titer using Table 4.5 for calculations. Titering consists of adding 10 μL of phage mixture with 150-200 μLof overnight XL1-blue cells and 4-5 mL of $\approx$ 50°C agarose top (if it is too hot, it will kill the cells).

(h) Add all remaining transformed *E. coli* cells to 400 mL LB with 0.4 mL tetracycline. Amplify at 37°C for 3 hours.

(i) Take a 100 μL aliquot and use for titering. Titering consists of adding 10 μL of phage mixture with 150-200 μL of overnight XL1-blue cells and 4-5 mL of $\approx$ 50°C agarose top (if it is too hot, it will kill the cells). Use Table 4.5 for calculations. Place plates in 37°C incubator overnight.

(j) To purify the amplified phage, centrifuge the mixture at 12200 RCF for 30 minutes.

(k) Transfer the supernatent into a new tube and add 20% (w/v) polyethylene glycol8000, 2.5 M NaCl (PEG/NaCl) (1/6 volume = 75 mL PEG/NaCl). Mix by inversion and store at 4°C overnight.

END OF DAY 2

(l) Centrifuge the phage mixture at 15900 RCF for 1 hour.

(m) Gently pour off the supernatent.

(n) Resuspend the pellet in 30 mL of 1X TBS and transfer to a 40 mL centrifuge tube.

(o) Centrifuge at 3500 RCF for 10 minutes to remove residual cell material.

(p) Transfer the supernatent to a clean 40 mL centrifuge tube.

(q) Add 1/6 volume PEG/NaCl to the supernatent (5.5 mL) and mix by inversion.

(r) Store at 4°C overnight.

(s) Count plaques from Day 2 titering and calculate amplification ratio.

Sample	**Dilution factor**	$\frac{PFU}{10\ \mu L}$	$\frac{PFU}{\mu L}$	**Volume (mL)**	**Total PFU**	**Average Total PFU**
Before amplification	0					
	10					
	10^2					
	10^3					
After 3 hour amplification	10^2					
	10^3					
	10^4					
	10^5					
	10^6					
Final concentrated library	10^5					
	10^6					
	10^7					
	10^8					
	10^9					

Table 4.5: *Table for calculating pVIII library titering results*

END OF DAY 3

(t) Centrifuge phage mixture at 21000 RCF for 45 minutes.

(u) Remove supernatent by decanting, spin again for 2 minutes to pull pellet and residual liquid to bottom.

(v) Remove residual liquid by pipette and Kimwipe (if necessary).

(w) Resuspend the pellet in 1 mL of 1X TBS. Transfer to 1.5 mL tube and centrifuge at 3500 RCF for 5 minutes to remove last cellular material. Store at 4°C in short term.

(x) For long term storage, add 1:1 volume of glycerol such that the final concentration is 50% glycerol. Store at -20°C.

(y) Titer the final library to determine concentration. Titering consists of adding 10 μL of phage mixture with 150-200 μL of overnight XL1-blue cells and 4-5 mL of $\approx$ 50°C agarose top (if it is too hot, it will kill the cells). Use Table 4.5 for calculations. Place plates in 37°C incubator overnight.

Chapter 5

Engineered M13 bacteriophage as highly multivalent, temperature-stable biocatalysts

5.1 Summary

In this work, M13 bacteriophage (M13 phage, or phage) was established as a temperature stable, highly-multivalent biocatalytic scaffold. A panel of histidine-biased M13 phage was screened for esterase activity and 2 out of 7 histidine phage enhance the hydrolysis of the substrate pNPA. The $\frac{K_{cat}}{K_{uncat}} \approx 170$ matched that reported for a computationally designed pNPA hydrolysis enzyme (PZD2) and the $K_m \approx 4$ mM was only 20-fold lower than PZD2. [99] The esterase phage retained almost 50% of their pNPA hydrolysis activity after a 4-hour, 80°C heat treatment, whereas bCA2, which also catalyzes this reaction, lost 97% of its activity in just 10 minutes at 80°C. The mechanism of action for esterase activity appears to be metal-ion-independent, though the histidine phage were shown to bind specifically to Zn^{2+}, Ni^{2+}, and Cu^{2+}, but not to other divalent cations. Future work will be focused on improving the K_m and K_{cat}, on identifying other industrially relevant substrates, and on incorporating metallic centers into the pVIII active sites.

5.2 Introduction

The use of enzymes and other biomolecules as catalysts has been an established tool for industrial chemical synthesis for over two decades. [100] Recently, biocatalysis has received renewed attention as the chemical and pharmaceutical industries emphasize "green chemistry" for sustainable manufacturing of their products. [101] There is a strong conviction that, due to innovations in computational optimization algorithms, rational design, directed evolution, and more efficient screening of higher quality biomolecular libraries, many new biological catalysts are on the horizon. [102]

There are two general strategies for the developmnet of new biological catalysts, each with its own tradeoffs. In one, *de novo* biomolecules are designed and built from scratch, allowing for complete control over the size, shape, and subcomponents of the biomolecule. [103–109]

However, there is vast complexity in biomolecules and nearly infinite possible combinations of amino acids; appropriately incorporating every desired feature when starting from scratch can be overwhelming. The other seeks to harness Nature's optimization of biomolecules and to bestow new functionalities on those biomolecules. [99, 110, 111] This allows one to start with a known entity with a pre-established structure and function (and with a built-in existence proof), and to focus effort on the desired novel functionalities.

This work takes the second approach, using M13 phage as the starting point. M13 phage is a rod-shaped virus with a highly ordered viral capsid that can be culture concentrations of greater than 20 $\frac{mg}{L}$. The major coat protein (pVIII) has roughly 2700 copies that stack in a five-fold symmetry down the length of the virus. This stacking enables regular 3D-ordering of fusion proteins that are added to the N-terminus of pVIII. [112] As demonstrated in Chapter 4, it is possible to make biased pVIII libraries, which are naturally endowed with the repetitive and consistent arrangement of the viral capsid. This appropriately reduces the sequence space for a library to a subset that is more likely to possess enzymatic activity, since well folded structures are a prerequisite for such activity. [113, 114] Every phage carries its own genome, meaning that there is a tight genotype-phenotype linkage, which simplifies the identification of the sequences that are causing activity. With genetically modifiable proteins on both ends and along the length of the capsid, if a catalytic site could be engineered into pVIII, it would be possible to make a highly multivalent, multi-functional biocatalytic molecule. [115, 116] Finally, M13 phage is more stable than many enzymes and proteins under non-standard conditions, such as high temperatures and low pH – a desirable feature for industrial catalytic reactions.

Enzymatic hydrolase reactions, including those catalyzed by esterases, lipases, and proteases, constitute a significant fraction of industrial biotransformations. [117] The proof-of-concept reaction chosen to demonstrate that the pVIII protein on M13 phage can act as a biocatalyst was the nucleophilic hydrolysis of *p*-nitrophenyl acetate (pNPA) (see Figure 5.1 for a scheme of the reaction). This reaction was chosen because it proceeds relatively slowly in neutral aqueous solutions, it has been well-characterized, and it generates a product, *p*-nitrophenol (pNP), that can be measured with a spectrophotometer. Also, since carbonic anhydrase has been shown to catalyze this reaction, there was a readily available positive control against which any M13 phage results could be benchmarked.

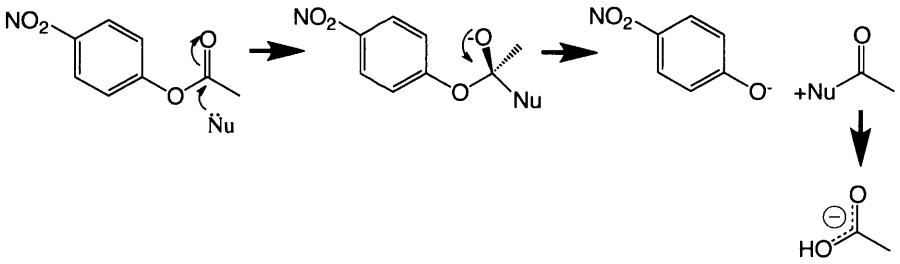
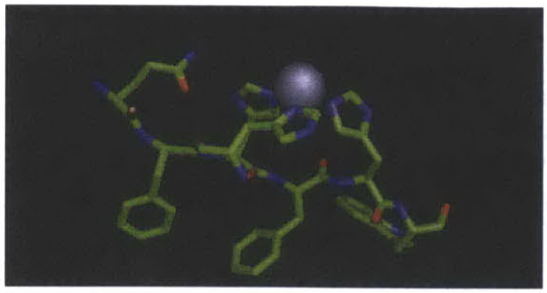


Figure 5.1: *Nucleophilic hydrolysis of pNPA* Image adapted from [99].

5.2.1 Active site of carbonic anhydrase and structural analogies to pVIII protein

In carbonic anhydrase, a coordinated zinc ion acts as a nucleophile to catalyze the hydrolysis of pNPA. [118] The structure of the active site of hCA2 was used as a model for the design of an active site on the pVIII protein. In the active site of hCA2, there are three histidine residues that coordinate the Zn^{2+} ion (see Figure 5.2a). Two of the histidine residues are on the same strand and are separated by one residue and the third histidine is on a separate strand but is in close proximity to the other two (see Figure 5.2b). The structure of the M13 viral capsid is similar to the carbonic anhydrase in that one pVIII protein is arranged in close proximity with three other pVIII proteins (see Figure 5.2c). The goal of this work was to build an active site on every copy of pVIII on an assembled M13 phage using the carbonic anhydrase active site as a model. The histidine-biased libraries built as part of the work described in Chapter 4 generated hundreds of histidine-biased sequences and it was this set of sequences that formed the starting point for generating a panel of histidine sequences to test for pNPA hydrolysis activity.

(a)



(b)

$Q_{92}F_{93}H_{94}F_{93}H_{96}W_{97}$ ... H_{119} and a Zn^{2+} ion

(c)

(d)

N-terminus – A EGD DPAKAAFN S_{13} L Q_{15} A S_{17} A T_{19} EYI G_{23} YAW A_{27} M... – C-terminus

xxHxHxxx

Figure 5.2: *hCA2 active site compared to four aligned pVIII proteins* (a) In the active site of hCA2, three histidine residues coordinate a zinc ion (PDB 3HS4). This active site can catalyze the hydrolysis of pNPA. (b) The sequence of the active site of hCA2. Two of the histidines are on the same strand and the third is on an adjacent strand. (c) Four aligned strands of pVIII on the surface of M13 virus (from PDB file 1IFJ). [92] The residues that are highlighted show the six residues that were identified to be tolerant of a histidine mutation. (d) The first 28 amino acids of pVIII. The three N-terminal amino acids that are replaced with the 8 amino acid library are highlighted, as are the other 6 amino acids that tolerate a histidine mutation. Only one backbone mutation was introduced in a design at a time (i.e., no sequences exist with multiple backbone histidine mutations).

5.3 Results

5.3.1 Engineered histidine phage with enzymatic activity

In the histidine-biased library work in Chapter 4, pVIII libraries with two histidines in the peptide insert at the N-terminus and a third histidine at the twenty-third residue (G23H) in the backbone were made. To test the versatility of the pVIII backbone in accepting histidines at locations other than G23, additional backbone residues were mutated. These mutations included S13H, Q15H, S17H, G23H, and A27H (N12 was also tested but would not tolerate a histidine mutation). See Figures 5.2c and 5.2d for locations of the histidine mutations in the pVIII backbone.

After identifying which backbone residues would tolerate histidine mutations, small-scale triple-His libraries were made with the 13H, 15H, and 23H backbone mutations using a scaled-down but similar version of the protocol described in Section 4.5.5. Phage from each of these libraries were amplified and tested for their ability to catalyze the hydrolysis of pNPA. The reaction rate was monitored across a range of pNPA concentrations, with each concentration measured in duplicate for each sample. An example of the raw rate data for the duplicate samples for a DDAHVHWE-23H phage, a bCA2 positive control, and a blank reaction (no enzyme) with 4 mM pNPA is shown in Figure 5.3a. For each sample and each concentration of pNPA, the reaction rate was calculated over the first 30 minutes of the reaction by fitting a linear regression through the data. Linearity was confirmed by analysis of the R^2 value and by visual inspection of the raw kinetic data. Rates from replicate wells were averaged and were plotted against substrate concentration (Figure 5.3b). Rates were converted to $\frac{nM}{second}$ as discussed in Section 5.5.3. The blank reaction rate was subtracted to obtain background-subtracted rates for each sample at each concentration of substrate. To compare the reaction rate enhancement for a variety of phage sequences, the average percent rate acceleration was calculated by dividing the rate for each sample at each substrate concentration by the blank reaction rate at the same substrate concentration. This value was converted to percent rate acceleration using the following equation:

$$Percent\ rate\ acceleration = (\frac{Rate_{sample}}{Rate_{blank}} - 1)(100\%) \tag{5.1}$$

The results of this calculation for all the phage sequences tested are shown in Figure 5.4 along with a bCA2 positive control. Nine different phage sequences were tested in total; one sequence had four histidines (ADHDHHSV-15H), five had three histidines, (DDAHVHWE-23H, EGHVHATM-23H, ATHDHGDQ-23H, DVHSHGLP-15H, and DMHEHNGE-13H), one had two histidines (TENMHGTM-23H), one had one histidine (DSPHTELP), and a control sequence had no histidines (EEE). Only two of the sequences demonstrated activity over background (DDAHVHWE-23H and TENMHGTM-23H). To assess the biological reproducibility of these results, the DDAHVHWE-23H phage was amplified and tested three separate times. The three biological replicates showed similar activities. Note that the pVIII peptides were present at $\approx$ 500X higher concentrations than bCA2 (22.5 μM vs. 45 nM).

In order to better assess the results of the pNPA hydrolysis assay, the enzyme parameters were calculated for the sequences that demonstrated activity. For two batches of DDAHVHWE-23H and for the one batch of TENMHGTP-23H, the K_m, K_{cat}, $\frac{K_{cat}}{K_{uncat}}$, and

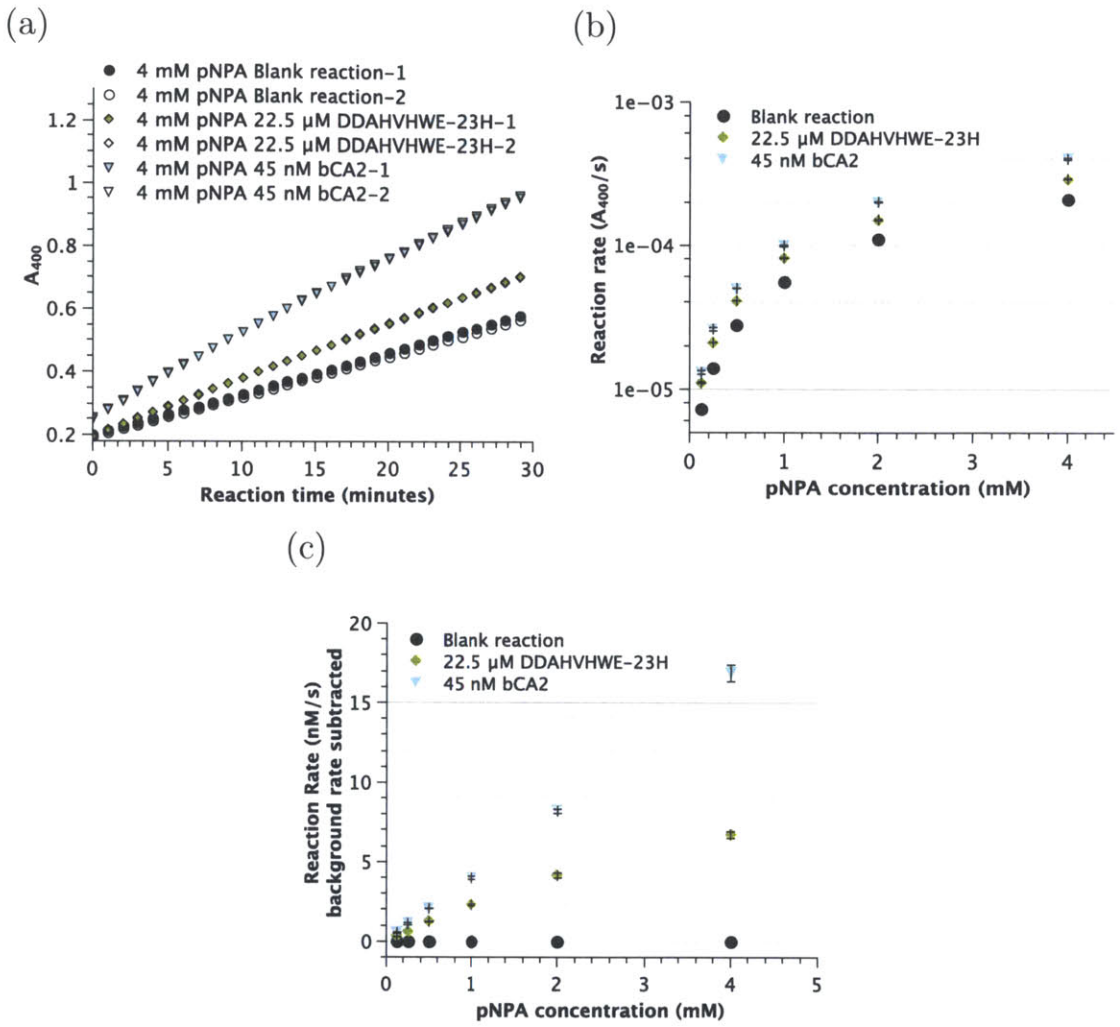


Figure 5.3: *Hydrolysis of pNPA by DDAHVHWE-23H phage* (a) The formation of pNP was measured by monitoring the absorbance at 400 nM. Data from the replicate wells for a blank reaction, a bCA2 postive control, and a DDAHVHWE-23H phage are shown. (b) For each concentration of pNPA, the slopes from the raw rate data were calculated for the first 30 minutes of the reaction. The values from the replicate wells were averaged and the reaction rate with standard deviations was plotted against the substrate concentration for each sample. To calculate the fold-improvement over background, the rate for each sample at each concentration of substrate was divided by the blank reaction rate at the same substrate concentration. (c) Rates were converted to $\frac{nM}{second}$ as discussed in Section 5.5.3. The blank reaction rate was subtracted to obtain background-subtracted rates.

(a)

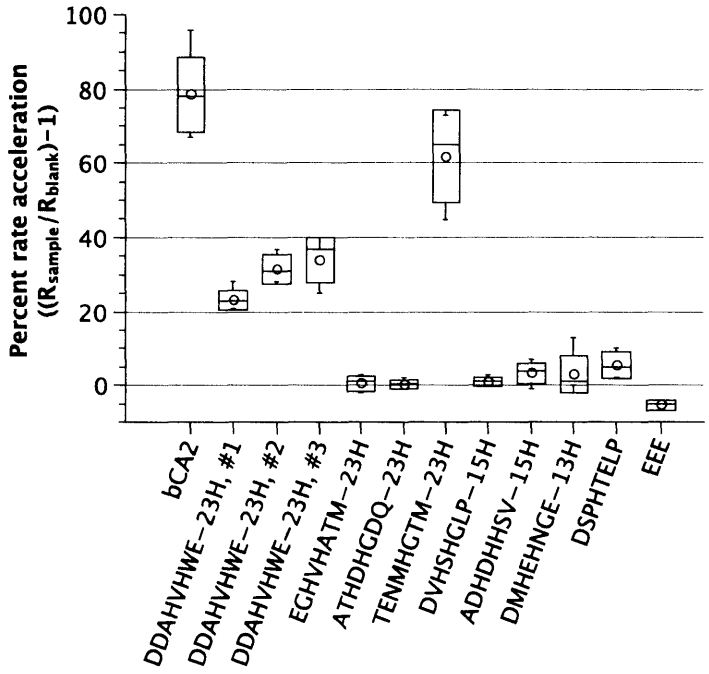


Figure 5.4: *pNPA hydrolysis assay results for a panel of engineered histidine phage* This figure shows the results of the pNPA hydrolysis assay for nine different sequences of engineered M13 phage and for a bovine carbonic anhydrase (bCA2) control. All phage were tested at a concentration of 5 x 10^{12} $\frac{PFU}{mL}$, or 22.5 μM pVIII, and the bCA2 was tested at a concentration of 45 nM. Six 23H sequences were tested, two 15H sequences were tested, one 13H sequence was tested, and two sequences with no backbone histidines were tested (DSPHTELP and EEE). Two of the histidine phage increased the reaction rate, DDAHVHWE-23H and TENMHGTM-23H. Three batches of DDAHVHWE-23H were tested. Each box and whiskers data point represents rate data from six concentrations of substrate. The ellipse represents the mean, the boxes the standard deviations and the whiskers the full range of the data. Each concentration of substrate was measured in duplicate and the rates were averaged and normalized to the blank reaction rates at that concentration of substrate.

$\frac{K_{cat}}{K_m}$ were calculated. These results are reported in Table 5.1, where they are benchmarked against the enzyme parameters of other enzymes that can catalyze pNPA hydrolysis. In 2001, Bolon, et al., engineered a pNPA hydrolysis site into a thioredoxin scaffold (PZD2) and their results are presented in Table 5.1 along with the engineered phage results. Also, the bCA2 positive control results from this work and from previous reports are included in Table 5.1. The bCA2 parameters calculated in this work match well with those reported in the literature, validating the 96-well pNPA plate assay developed as part of this work.

Sample	**K_m (mM)**	**K_{cat} ($\frac{1}{sec}$)**	$\frac{K_{cat}}{K_{uncat}}$	$\frac{K_{cat}}{K_m}$ ($\frac{1}{(M)(sec)}$)	Source
DDAHVHWE-23H (batch 1)	7.1 ± 0.3	8.3 ± 0.25 x 10^{-4}	168 ± 5	0.12 ± 0.004	this study
DDAHVHWE-23H (batch 3)	4.0 ± 1.7	4.7 ± 1.1 x 10^{-4}	98 ± 24	0.12 ± 0.03	this study
TENMHGTM-23H	4.1 ± 0.4	8.4 ± 0.45 x 10^{-4}	170 ± 9	0.21 ± 0.01	this study
Engineered thioredoxin (PZD2)	0.17 ± 0.02	4.6 ± 0.2 x 10^{-4}	180	2.71	[99]
bCA2	31.4 ± 7.26	7.8 ± 1.7	1.6 ± 0.34 x 10^6	248 ± 54	this study
literature value	*9.2*	*2.45*	*6 x 10^5*	*266*	[119]

Table 5.1: *Enzyme parameters for pNPA hydrolysis by engineered phage* All values reported for engineered phage are reported per pVIII protein, not per phage molecule. The engineered phage have K_{cat} values that are comparable with PZD2, a thioredoxin engineered to have pNPA hydrolysis activity. The K_m and $\frac{K_{cat}}{K_m}$ values for the engineered phage are only $\approx$ one order of magnitude worse than that of PZD2.

The engineered phage have K_{cat} values that are comparable with PZD2 and K_m values that are only 10X higher - impressive results considering that PZD2 was rationally designed to catalyze this reaction, while the engineered phage were identified in a screen of just a few sequences. All values reported for engineered phage are reported per pVIII protein, not per phage molecule. If the same parameters were calculated per phage molecule the values are very close to that of bCA2 (M13 K_{cat} = 1.3–2.3 $\frac{1}{sec}$ and M13 $\frac{K_{cat}}{K_{uncat}}$ = 2.7–4.6 x 10^5).

Interestingly, the active phage sequences did not have any Zn^{2+} added to them (beyond what is present in the bacterial growth media), meaning that the mechanism of hydrolysis may be different from that of bCA2. While further experiments will clarify the exact mechanism, it appears likely that the histidines themselves are the active site (see the discussion in Section 5.4).

5.3.2 High temperature stability of enzymatic phage

After identifying two pVIII sequences that could catalyze the hydrolysis of pNPA, the temperature stability of the enzymatic phage was tested. First, DDAHVHWE-23H and bCA2

were subjected to a 10-minute treatment at 80°C in a PCR thermocycler, much like one might treat an enzyme to perform a heat kill during a cloning experiment. The pNPA hydrolysis activities of bCA2 and DDAHVHWE were compared before and after the 80°C heat treatment. The bCA2 sample lost nearly all of its activity, while the DDAHVHWE-23H sample retained most of its activity. Figure 5.5a shows these results.

After demonstrating that the DDAHVHWE-23H phage retained enzymatic activity after a 10 minutes heat treatment at 80°C, DDAHVHWE-23H and TENMHGTM-23H phage were exposed to 1 hour and 4 hour heat treatments at 80°C. The results of this experiment are shown in Figure 5.5b and are summarized in Table 5.2. Both phage sequences retained at least 45% of their pNPA hydrolysis activity even after 4 hours of exposure to 80°C, while the bCA2 sample lost 97% of its pNPA hydrolysis activity after just 10 minutes at 80°C.

Sample	Time at 80°C (hours)	Percent activity retained
DDAHVHWE-23H	0.17	86% **
	2	47% **
	4	46% **
TENMHGTP-23H	2	85%
	4	80%
bCA2	0.17	3%

**0.17 hour data and 2 hour, 4 hour data from separate experiments

Table 5.2: *Stability at 80°C for enzymatic phage compared to bCA2* Percent activity retained was calculated using the average percent rate acceleration value (the ellipses in Figure 5.5) normalized to the no-heat-treatment value for each sample.

5.3.3 Transition metal ion binding by engineered histidine phage

The capacity to engineer a histidine-based active site onto every copy of the surface coat protein pVIII and the subsequent resistance to high temperature treatment establishes M13 phage as a scaffold for enzymatic reactions. However, a larger suite of catalytic reactions would be accessible if bioinorganic active sites with coordinated metal ions could be engineered onto pVIII. [90] Metal ions play several roles in biomolecules and the coordination chemistry is a determining factor for which role they play. For example, Zn^{2+} has four coordination sites. If all four of those coordination sites are occupied by organic ligands (four amino acids), then the Zn^{2+} is playing a structural role in the protein. If, however, three of those sites are occupied by the imidazole ligand of histidine residues and the fourth is occupied by a water molecule, then the Zn^{2+} site can act as a catalytic site. [120]

The metal ion binding capacity of the histidine phage was explored using a Nanotember Monolith microscale thermophoresis (MST) instrument. MST is based on the movement of molecules in microscopic temperature gradients, and it provides information about molecule size, charge and hydration shell. MST has been used to measure a variety of biomolecular

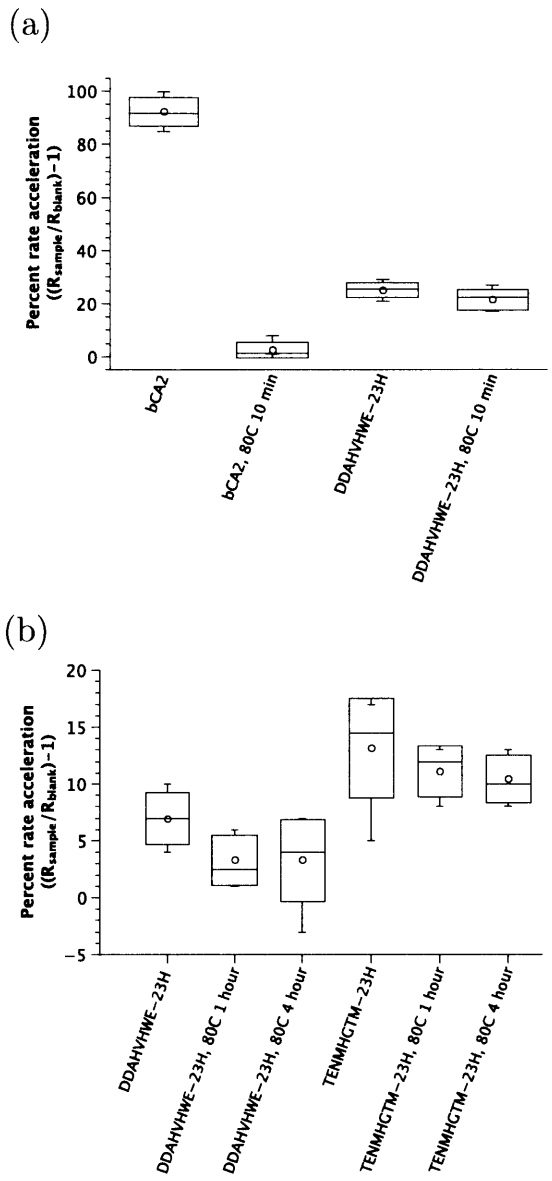


Figure 5.5: *Stability of enzymatic phage at 80°C* (a) The temperature stability of DDAHVHWE-23H phage was compared to bCA2 assessing the pNPA hydrolysis activity of each biomolecule before and after a ten minute 80°C heat treatment. Activity was measured using pNPA assay with phage at a concentration of 22.5 μM pVIII. (b) The temperature stability of DDAHVHWE-23H and TENMHGTM-23H phage was tested by measuring the activity of the phage before and after 80°C heat treatments. Activity was measured using pNPA assay with phage at a concentration of 4.5 μM pVIII. Each box and whiskers data point represents rate data from six concentrations of substrate. The ellipse represents the mean, the boxes the standard deviations and the whiskers the full range of the data. Each concentration of substrate was measured in duplicate, averaged, and normalized to the blank reaction at that concentration of substrate.

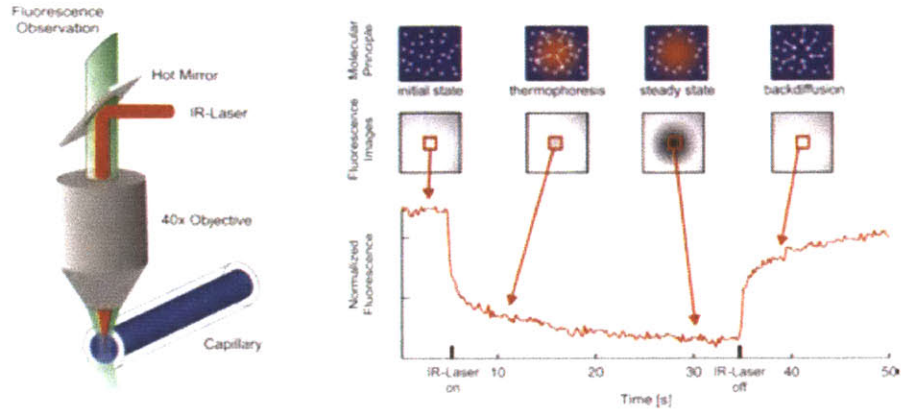
binding events, including protein-protein, protein-small molecule, and protein-ion interactions and a full description of the technique is available elsewhere. [121, 122] Briefly, a fluorescently labelled protein (in this case, M13 phage) is loaded into a glass capillary. An infrared laser is used to heat a small section of the capillary (Figure 5.6a) and the mobility of the protein is measured by monitoring the change in fluorescence. A set of 16 protein samples (all at the same concentration) are mixed with a serially-diluted binding partner (in this case, metal ions) in 16 different glass capillaries. In the samples where the concentration of the binding partner is higher than the K_D of the binding interaction, the thermophoretic mobility of the protein changes. Plotting the thermophoretic mobility of the protein against the concentration of the binding partner for all 16 samples provides data like that shown in Figure 5.7a, which approximates a characteristic binding curve.

The binding of four different engineered M13 phage to Zn^{2+} was tested using MST. Zn^{2+} was titrated from 250 μM to 0.01 μM. For all three of the triple-His phage (DDAHVHWE-23H, ESHTHGNE-23H, and EAHAHGDE-23H) there was an interaction with the Zn^{2+} at around 7–20 μM Zn^{2+}(see Figure 5.7a). The negatively charged EEE phage did not exhibit any change in thermophoretic mobility across the range of Zn^{2+} concentrations tested. Histidine residues are well know for their ability to interact with Zn^{2+}, as well as with other first row transition metals, whereas histidines do not typically interact with the divalent cations Ca^{2+} or Mg^{2+}. The binding of DDAHVHWE-23H phage to five divalent cations was measured using MST. The DDAHVHWE-23H phage showed ion-specific binding with the three transition metals, Zn^{2+}, Cu^{2+}, and Ni^{2+}, but not with Ca^{2+} or Mg^{2+}.

In all of the experiments where the phage was interacting with the metal ions (i.e., for all binding experiments between three triple-His sequences and Zn^{2+}, Cu^{2+}, or Ni^{2+}), the MST traces looked like the one shown in Figure 5.6b. According to the Nanotemper product literature, the irregularity of the shapes of MST curves (compare the typical curve from the Nanotemper product literature in Figure 5.6a to the curves in Figure 5.6b) indicate that the Zn^{2+} was inducing aggregation of the phage. To further investigate this, the pNPA hydrolysis activity of DDAHVHWE-23H phage was tested in the presence of 0 mM Zn^{2+}, 0.1 mM Zn^{2+}, and 1 mM Zn^{2+}(see Figure 5.7c). At 100 μM Zn^{2+} (well above the measured K_D), the pNPA hydrolysis activity was severely quenched. At 1 mM Zn^{2+}, the raw absorbance data became very bumpy, leading to the very high standard deviation for the processed rate acceleration data shown in Figure 5.7c.

The DDAHVHWE-23H, 1 mM Zn^{2+} and 0 mM Zn^{2+} samples were also examined under AFM. Ten μL of each sample (aliquots taken prior to pNPA hydrolysis assay) was deposited onto an AFM disk and allowed to dry. The AFM images for the 0 mM Zn^{2+} samples are shown in Figures 5.8a and b. The features are typical AFM images of M13 phage, with some small clumps of virus (left side in a and bottom right in b), but with a significant number of phage (rod-like structures) spread out across the surface. The 1 mM Zn^{2+} samples in Figure 5.8c and d, however, are unusual for AFM images of M13 phage. The entire AFM substrate surface is covered in large clumps of material, which, when examined at a higher magnification (in d), seem to be composed of structures that have the same dimensions as M13 phage.

(a)



(b)

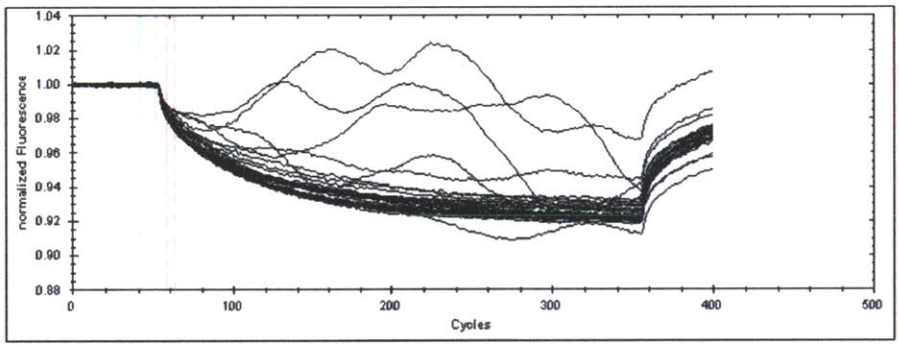


Figure 5.6: *Explanation of Microscale Thermophoresis* (a) Before temperature change, the molecules are homogeneously distributed and a constant fluorescent signal is measured. When the laser is turned on, a fast temperature jump is observed, followed by thermophoretic molecular motion. The fluorescence decrease is measured for 30 seconds. When the laser is turned off, an inverse temperature jump is observed, followed by the diffusion of molecules, which is purely driven by mass diffusion. The curve on the right is a typical MST curve from the Nanotemper product literature. Figure and explanation from [122]. (b) An example of MST data from a DDAHVHWE-23H and Zn^{2+} binding interaction experiment. Each trace represents data from a different capillary (Zn^{2+} was serially diluted across 16 capillaries). At a critical concentration of Zn^{2+} the traces stop looking like a "normal" trace, like the example on the right in (a), and start to exhibit bumpiness, which is indicative of sample aggregation. To identify this critical concentration of Zn^{2+}, the change in fluorescence was calculated using temperature jump data for each sample (*MST Temp Jump* = (*Ave. fluor. signal between vertical blue lines*) − (*Ave. fluor. signal between vertical red lines*)).

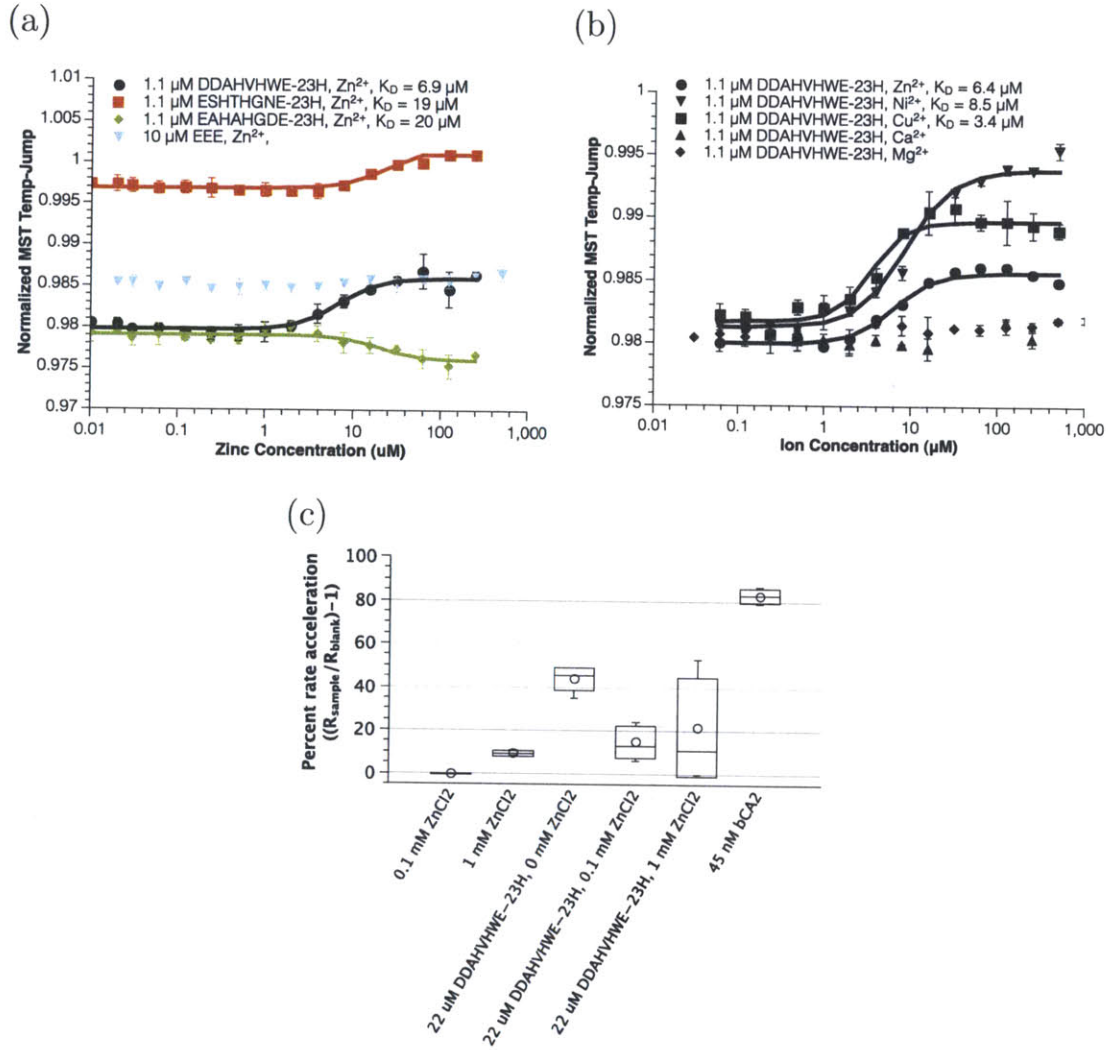


Figure 5.7: *Engineered phage bind transition metals, aggregate, and lose pNPA hydrolysis activity* (a) The thermophoretic mobility of four strains of engineered phage were tested across a range of Zn^{2+} concentrations. All three triple-His phage exhibited binding to Zn^{2+} with K_D's between 6 and 20 μM. (b) DDAHVHWE-23H binds the first-row transition metals Zn^{2+}, Ni^{2+}, and Cu^{2+}, but not Ca^{2+} or Mg^{2+}. (c) Zn^{2+}-induced aggregation of DDAHVHWE-23H leads to a loss of pNPA hydrolysis activity. Each box and whiskers data point represents rate data from six concentrations of substrate. The ellipse represents the mean, the boxes the standard deviations and the whiskers the full range of the data. Each concentration of substrate was measured in duplicate, averaged, and normalized to the blank reaction at that concentration of substrate.

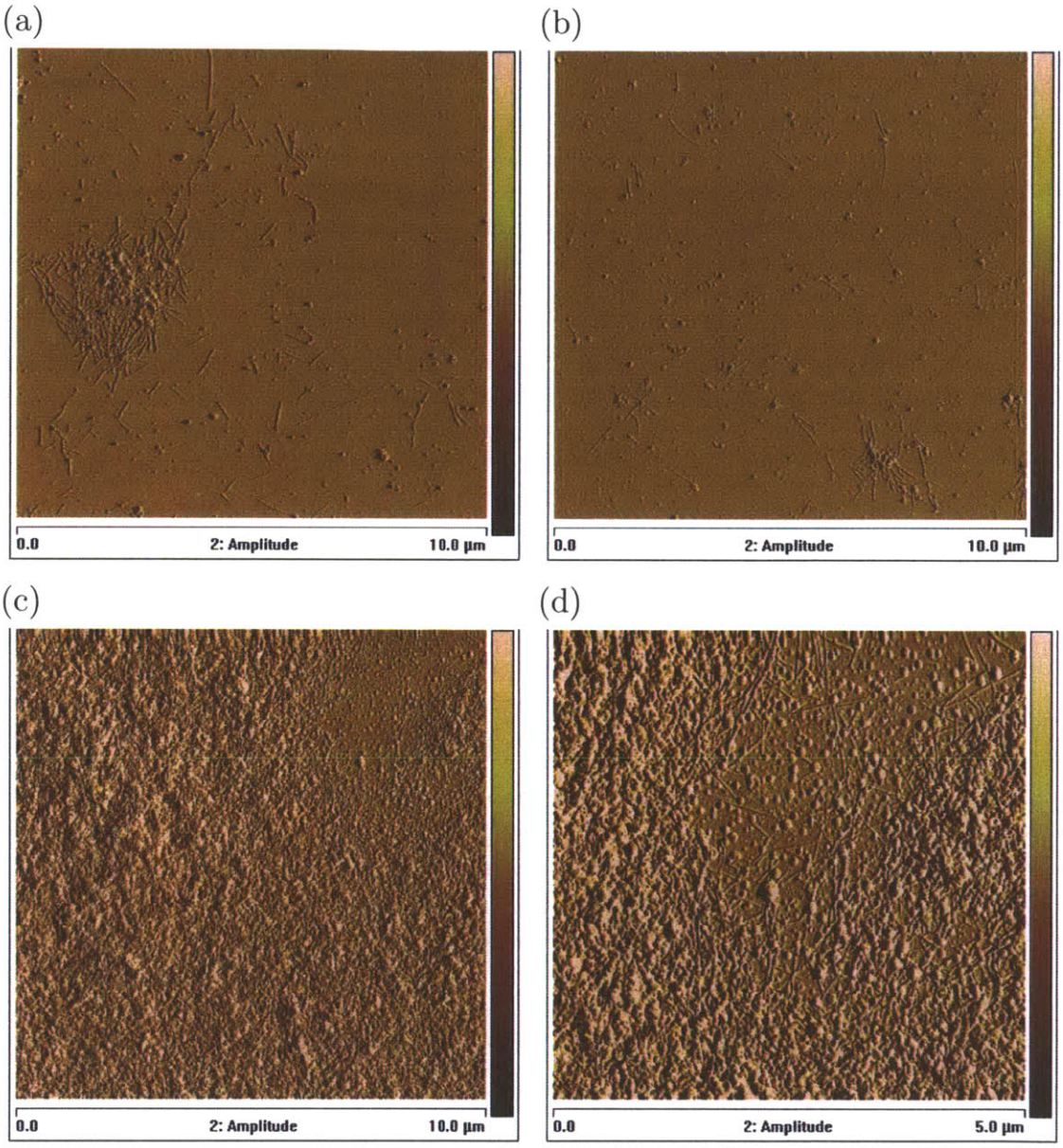


Figure 5.8: *AFM images of Zn^{2+} induced aggregation of DDAHVHWE-23H* (a) and (b) AFM amplitude images of DDHAVHWE-23H phage with no $ZnCl_2$. There are some clumps of material (which are often present in M13 AFM images, due to drying effects), but a lot of individual rod-like shapes (M13 phage). (c) and (d) AFM amplitude images of DDHAVHWE-23H phage with 1 mM $ZnCl_2$. Note that the AFM substrate surface is almost entirely covered with material. The image in (d) is a higher zoom image of the top right corner from (c) and shows some individual phage, but the majority appear to be aggregated together. Both M13 phage samples came from the same stock solution and were handled exactly the same, with the only exception being the amount of Zn^{2+} added to the samples.

5.4 Conclusions

The results of this work firmly establish M13 phage as a highly multivalent, temperature-stable scaffold for biocatalysis. While the results indicated that the esterase activity is independent of zinc, which would suggest that the histidines are themselves the active site, this mechanism needs to be confirmed. Bifunctional catalysis of ester hydrolysis by two histidines has been previously reported, with one "acting in the free base form as the nucleophilic or general-base catalyst and another acting in the protonated form, for example, in stabilizing the oxyanion intermediate" [123]. In a *de novo* designed 42-amino acid polypeptide that folded into a four-helix bundle, it was demonstrated that a two-histidine active site, with one His residue protonated and the other unprotonated, functioned through cooperative nucleophilic and general-acid catalysis. [124] The distance between the histidines and the nature of the surrounding residues were critical parameters in tuning the activity of the peptide, which is consistent with the fact that esterase activity was only measured in two of the M13 phage tested in this work. In fact, it turns out that a two-histidine active site is a common motif that has been used to validate other proteins as scaffolds for engineered enzymatic activity. Bolon and Mayo used computational design to identify mutations in a thioredoxin protein that would enable zinc-independent, histidine-mediated nucleophilic hydrolysis of pNPA and Host, et al., created a novel esterase by adding a two-histidine reactive site into a well-defined substrate binding site. [99, 125] Bolon and Mayo's results are presented in this work as a benchmark for engineered esterase activity in Table 5.1.

In this work two pVIII sequences that had esterase activity were identified from a panel of just seven sequences. A success rate of $> 20\%$ is unusual in library screening projects. However, it should be made clear that the M13 phage that were chosen for analysis were culled from the hundreds of phage that were sequenced based on some underlying rules-of-thumb for biocatalytic activity. First, the active site sequence of hCA2 was used as a model and sequences that looked similar, especially with regard to the histidine-adjacent residues were chosen. For example, in hCA2, the three residues around H_{94} and H_{96} are FHFHW and it has been shown previously there is a relationship between the hydrophobicity of these surrounding residues and the esterase activity of the enzyme. [126] It also has been noted that metal ions often bind in hydrophobic cores. [127] In other words, the results of this work were not the results from a random panel of sequences; the M13 phage sequences that were chosen in this work were chosen because of their perceived adherence to these rules (i.e., at an amino acid level DDAHVHWE looks like FHFHW).

Even with this preselection bias in the panel of tested phage, the hit rate is still high. Others have previously reported, though, on the catalytic abilities of *de novo* proteins that were designed only for structure, with no function in mind and that were not exposed to evolutionary pressure. [128] In that work, 30% of the tested proteins displayed esterase activity and 20% displayed lipase activity. In fact, it has been hypothesized that this feature (i.e., promiscuous, low-level catalytic activity) could have served as a central node for evolution in nature. [129]

Further work will need to be performed to verify that there is truly no metallic center in the active sites of the pVIII proteins. Also, an alanine scan of all of the residues in the 8 peptide insert and of the histidine in the backbone should offer more insight into the roles of each amino acid. Future work will use 4-methylimidazole as a control for the histidine

residue without the surrounding phage sequence and the active pVIII insert sequences will be tested in solution to determine the contribution of the viral capsid structure to activity. Finally, future work should focus on examining activity in other assays (like lipases and heme binding) and on improving the K_m and K_{cat} of the enzymatic phage to make them more appealing for industrial chemical synthesis.

5.5 Experimental

5.5.1 Phage preparation

Phage clones were picked from individual plaques off of a titering plate and were amplified for 6 hours at 37°C in a culture of XL1-Blue *E. coli* cells. The culture was made by adding 40 μL of an overnight culture of XL1-Blue cells and 4 μL of 20 $\frac{mg}{mL}$ tetracycline to 4 mL of LB. After amplification, the cells were pelleted. Double-stranded DNA was purified from the cells using a Qiagen plasmid prep kit and was sequenced with a neg96gIII primer (New England Biolabs) to confirm that the correct pVIII sequence. The 4 mL supernatent was retained and used to seed a 1 L XL1-Blue *E. coli* culture. The phage were amplified for 8 hours to overnight. After amplification, the culture was centrifuged at 12200 RCF for 30 minutes. A small part of the pellet was retained and double-stranded DNA was purified and sequenced as before to confirm the correct sequence. To the supernatent a 1/6 (volume to volume) amount of 20% (w/v) polyethylene glycol8000, 2.5 M NaCl (PEG/NaCl) was added. The sample was mixed by inversion and stored overnight at 4°C. The next morning the sample was centrifuged at 15900 RCF for 1 hour. The supernatent was removed by decanting, the pellet was resuspended in 30 mL of 1X TBS and was transferred to a clean 40 mL centrifuge tube. The tube was centrifuged at 3500 RCF for 10 minutes to remove residual cell material and was then transferred to another clean 40 mL centrifuge tube. A 1/6 volume of PEG/NaCl was added. The sample was mixed by inversion and stored overnight at 4°C. The next day, the samples were centrifuged at 21000 RCF for 45 minutes and the supernatent was removed by decanting. The tube was centrifuged again for 2 minutes to pull the pellet and residual liquid to the bottom. The residual liquid was removed and the pellet was resuspended in 1 mL of 50 mL Tris-sulfate, pH 7.6. The concentration was measured by UV-Vis absorbance.

5.5.2 pNPA assay

Two-fold serial dilutions of pNPA (Sigma number: N8130) were made in DMSO from 80 mM to 2.5 mM and were stored at 4°C in sealed tubes. Bovine carbonic anhydrase was purchased from Worthington Enzymes, was resuspended in ultrapure water and was stored in 0.1 $\frac{mg}{mL}$ aliquots at -20°C. Just prior to testing, bCA2 was diluted to the appropriate concentration in Tris-sulfate, pH 7.6. To set up the assay, 10 μL each of the different concentrations of pNPA was added to separate wells in a 96-well plate. For each sample to be tested (blank, bCA2, phage), two replicate concentrations of pNPA were prepared. To initiate the assay, a multichannel pipette was used to add 190 μL of phage, bCA2, or buffer only to each well. Each well was mixed by aspirating and dispensing gently. The absorbance at 400 nm was

read in kinetic mode in a plate reader at 25°C. The slopes of the A_{400} change were calculated for 30 minutes of reaction (during which the reaction remained in the linear phase).

5.5.3 Calculating enzyme parameters

To convert rate units from A_{400} per second to nM per second, the following equation was used

$$\frac{(\frac{A_{400}}{sec})(10^9\ \frac{nM}{M})}{(pNPA\ extinction\ coefficient)(Path\ length)} \tag{5.2}$$

where the extinction coefficient is 18,400 $\frac{1}{(M)(second)}$ and the path length is estimated to be 0.6 cm for 200 μL of liquid in a 96-well plate well. To calculate the k_{uncat}, the following equation was used

$$k_{uncat} = \frac{Rate_{uncatalyzed}}{pNPA\ concentration} \tag{5.3}$$

where the uncatalyzed reaction rate was the rate of the buffer only samples. The average k_{uncat} was calculated across all substrate concentrations. To determine V_{max} and K_D, the data was fit using QTI plot to the Michealis-Menten equation:

$$y = \frac{(V_{max})(x)}{(K_m + x)} \tag{5.4}$$

where $V_{max} = (k_{cat})([E])$, $x = [S]$, and K_m is equal to the substrate concentration at which the reaction rate is half of the V_{max}.

5.5.4 Labeling triple-His phage with AF488 for MST experiments

To 200 μL of phage in 1X PBS, 300 μL of PBS, and 20 μL of 10 $\frac{mg}{mL}$ AF488-5-SDP (Invitrogen) was added. The reaction was allowed to proceed for 1.5 hours at room temperature under slight agitation. Then the samples were dialyzed to remove the unreacted dye against 5 L of 1X PBS overnight. The PBS was exchanged twice the next day for two more 3 hour dialyses. Finally, a 1/6 volume of PEG/NaCl was added to the solution in the dialysis tubing and it was transferred to a 1.5 mL tube and stored at 4°C overnight. The next day, the samples were spun at 18,000 RCF for 20 minutes. The supernatent was removed and the samples were resuspended in 1X TBS. The concentrations were measured and the removal of unconjugated dye was confirmed by comparing the absorbances at 269 nm, 320 nm, and 495 nm (absorption peak for AF488).

5.5.5 MST experiments

Samples were prepared according to protocol for MST instrument. Briefly, the fluorescently labeled phage samples were diluted to roughly 1.1 μM pVIII (for histidine phage) or 10 μM pVIII (for EEE phage) in sample buffer. To 10 μL of phage was added 10 μL of metal ion in the sample buffer. Samples were read with 100% blue LED power and 40% laser power.

Data was analyzed using Nanotemper Analysis software 1.2.095. Since the histidine phage aggregated in the presence of the first row transition metals because aggregation causes large variations in the fluorescence signal traces (see Figure 5.6c for an example), it was not possible to use the thermophoretic mobility to calculate binding affinity. The change in the fluorescence after the temperature jump (the difference between the average fluorescence between the red lines and the blue lines in Figure 5.6c) was used to calculate the binding affinity for the phage and the metal ions; this approach is a valid alternative for analyzing aggregated samples, per the Nanotemper User Manual. To calculate K_D, data was fitted to a Boltzman curve in QTI Plot using the following equation:

$$y = \frac{(Bottom) - (Top)}{1 + e^{((x - K_D)(Slope) + (Top))}} \tag{5.5}$$

Bibliography

[1] Y. F. Dufrene. Using nanotechniques to explore microbial surfaces. *Nature Reviews Microbiology*, 2(6):451–460, 2004.

[2] Y. F. Dufrene. Towards nanomicrobiology using atomic force microscopy. *Nature Reviews Microbiology*, 6(9):674–680, 2008.

[3] G. E. Fantner, R. J. Barbero, D. S. Gray, and A. M. Belcher. Kinetics of antimicrobial peptide activity measured on individual bacterial cells using high-speed atomic force microscopy. *Nature Nanotechnology*, 5(4):280–285, 2010.

[4] R. Matzke, K. Jacobson, and M. Radmacher. Direct, high-resolution measurement of furrow stiffening during division of adherent cells. *Nature Cell Biology*, 3(6):607–610, 2001.

[5] D. J. Muller, W. Baumeister, and A. Engel. Controlled unzipping of a bacterial surface layer with atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 96(23):13170–13174, 1999.

[6] D. J. Muller, D. Fotiadis, S. Scheuring, S. A. Muller, and A. Engel. Electrostatically balanced subnanometer imaging of biological specimens by atomic force microscope. *Biophysical Journal*, 76(2):1101–1111, 1999.

[7] M. Plomp, T. J. Leighton, K. E. Wheeler, H. D. Hill, and A. J. Malkin. In vitro high-resolution structural dynamics of single germinating bacterial spores. *Proceedings of the National Academy of Sciences of the United States of America*, 104(23):9644–9649, 2007.

[8] G. Francius, O. Domenech, M. P. Mingeot-Leclercq, and Y. F. Dufrene. Direct observation of staphylococcus aureus cell wall digestion by lysostaphin. *Journal of Bacteriology*, 190(24):7904–7909, 2008.

[9] P. Hinterdorfer, W. Baumgartner, H. J. Gruber, K. Schilcher, and H. Schindler. Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 93(8):3477–3481, 1996.

[10] D. J. Muller and Y. F. Dufrene. Atomic force microscopy as a multifunctional molecular toolbox in nanobiotechnology. *Nature Nanotechnology*, 3(5):261–269, 2008.

[11] T. Ando, N. Kodera, E. Takai, D. Maruyama, K. Saito, and A. Toda. A high-speed atomic force microscope for studying biological macromolecules. *Proceedings of the National Academy of Sciences of the United States of America*, 98(22):12468–12472, 2001.

[12] M. Kobayashi, K. Sumitomo, and K. Torimitsu. Real-time imaging of dna-streptavidin complex formation in solution using a high-speed atomic force microscope. *Ultramicroscopy*, 107(2-3):184–190, 2007.

[13] P. K. Hansma, G. Schitter, G. E. Fantner, and C. Prater. Applied physics - high-speed atomic force microscopy. *Science*, 314(5799):601–602, 2006.

[14] E. Tiozzo, G. Rocco, A. Tossi, and D. Romeo. Wide-spectrum antibiotic activity of synthetic, amphipathic peptides. *Biochemical and Biophysical Research Communications*, 249(1):202–206, 1998.

[15] C. T. Gottlieb, L. E. Thomsen, H. Ingmer, P. H. Mygind, H. H. Kristensen, and L. Gram. Antimicrobial peptides effectively kill a broad spectrum of listeria monocytogenes and staphylococcus aureus strains independently of origin, sub-type, or virulence factor expression. *BMC Microbiology*, 8:205, 2008.

[16] R. E. W. Hancock and H. G. Sahl. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. *Nature Biotechnology*, 24(12):1551–1557, 2006.

[17] C. Loose, K. Jensen, I. Rigoutsos, and G. Stephanopoulos. A linguistic model for the rational design of antimicrobial peptides. *Nature*, 443(7113):867–869, 2006.

[18] M. Zasloff. Antimicrobial peptides of multicellular organisms. *Nature*, 415(6870):389–395, 2002.

[19] K. A. Brogden. Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? *Nature Reviews Microbiology*, 3(3):238–250, 2005.

[20] D. Andreu, J. Ubach, A. Boman, B. Wahlin, D. Wade, R. B. Merrifield, and H. G. Boman. Shortened cecropin-a melittin hybrids - significant size-reduction retains potent antibiotic-activity. *FEBS Letters*, 296(2):190–194, 1992.

[21] H. Sato and J. B. Feix. Osmoprotection of bacterial cells from toxicity caused by antimicrobial hybrid peptide CM15. *Biochemistry*, 45(33):9997–10007, 2006.

[22] V. C. Kalfa, H. P. Jia, R. A. Kunkle, P. B. McCray, B. F. Tack, and K. A. Brogden. Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. *Antimicrobial Agents and Chemotherapy*, 45(11):3256–3261, 2001.

[23] M. Meincken, D. L. Holroyd, and M. Rautenbach. Atomic force microscopy study of the effect of antimicrobial peptides on the cell envelope of *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*, 49(10):4085–4092, 2005.

[24] M. L. Mangoni, N. Papo, D. Barra, M. Simmaco, A. Bozzi, A. di Giulio, and A. C. Rinaldi. Effects of the antimicrobial peptide Temporin L on cell morphology, membrane permeability, and viability of *Escherichia Coli*. *Journal of Peptide Science*, 10:859–865, 2004.

[25] B. Bechinger. The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. *Biochimica Et Biophysica Acta-Biomembranes*, 1462(1-2):157–183, 1999.

[26] A. S. Ladokhin, M. E. Selsted, and S. H. White. Sizing membrane pores in lipid vesicles by leakage of co-encapsulated markers: Pore formation by melittin. *Biophysical Journal*, 72(4):1762–1766, 1997.

[27] M. T. Lee, W. C. Hung, F. Y. Chen, and H. W. Huang. Mechanism and kinetics of pore formation in membranes by water-soluble amphipathic peptides. *Proceedings of the National Academy of Sciences of the United States of America*, 105(13):5087–5092, 2008.

[28] Y. F. Dufrene. Atomic force microscopy and chemical force microscopy of microbial cells. *Nature Protocols*, 3(7):1132–1138, 2008.

[29] G. E. Fantner, G. Schitter, J. H. Kindt, T. Ivanov, K. Ivanova, R. Patel, N. Holten-Andersen, J. Adams, P. J. Thurner, I. W. Rangelow, and P. K. Hansma. Components for high speed atomic force microscopy. *Ultramicroscopy*, 106(8-9):881–887, 2006.

[30] M. B. Viani, T. E. Schaffer, G. T. Paloczi, L. I. Pietrasanta, B. L. Smith, J. B. Thompson, M. Richter, M. Rief, H. E. Gaub, K. W. Plaxco, A. N. Cleland, H. G. Hansma, and P. K. Hansma. Fast imaging and fast force spectroscopy of single biopolymers with a new atomic force microscope designed for small cantilevers. *Review of Scientific Instruments*, 70(11):4300–4303, 1999.

[31] D. J. Ferullo, D. L. Cooper, H. R. Moore, and S. T. Lovett. Cell cycle synchronization of escherichia coli using the stringent response, with fluorescence labeling assays for dna content and replication. *Methods*, 48(1):8–13, 2009.

[32] D. Alsteens, C. Verbelen, E. Dague, D. Raze, A. R. Baulard, and Y. F. Dufrene. Organization of the mycobacterial cell wall: a nanoscale view. *Pflugers Archiv-European Journal of Physiology*, 456(1):117–125, 2008.

[33] J. G. Albeck, J. M. Burke, B. B. Aldridge, M. S. Zhang, D. A. Lauffenburger, and P. K. Sorger. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Molecular Cell*, 30(1):11–25, 2008.

[34] M. Sarikaya, C. Tamerler, A. K. Y. Jen, K. Schulten, and F. Baneyx. Molecular biomimetics: nanotechnology through biology. *Nature Materials*, 2(9):577–585, 2003.

[35] R. Carlson. The changing economics of DNA synthesis. *Nature Biotechnology*, 27(12):1091–1094, 2009.

[36] E. M. Krauland. *Towards rational design of peptides for selective interaction with inorganic materials.* PhD thesis, M.I.T., 2007.

[37] F. Rusmini, Z. Y. Zhong, and J. Feijen. Protein immobilization strategies for protein biochips. *Biomacromolecules*, 8(6):1775–1789, 2007.

[38] P. J. Yu, G. Ferrari, A. C. Galloway, P. Mignatti, and G. Pintucci. Basic fibroblast growth factor (FGF-2): The high molecular weight forms come of age. *Journal of Cellular Biochemistry*, 100(5):1100–1108, 2007.

[39] T. P. Burg, M. Godin, S. M. Knudsen, W. Shen, G. Carlson, J. S. Foster, K. Babcock, and S. R. Manalis. Weighing of biomolecules, single cells and single nanoparticles in fluid. *Nature*, 446(7139):1066–1069, 2007.

[40] C. Potera. Biofilms invade microbiology. *Science*, 273(5283):1795–1797, 1996.

[41] T. E. Cloete. Biofouling control in industrial water systems: What we know and what we need to know. *Materials and Corrosion*, 54(7):520–526, 2003.

[42] M. Sharma and S.K. Anand. Characterization of constitutive microflora of biofilms in dairy processing lines. *Food Microbiology*, 19(6):627 – 636, 2002.

[43] B. Schachter. Slimy business – the biotechnology of biofilms. *Nature Biotechnology*, 21(4):361–365, 2003.

[44] A. Caro, V. Humblot, C. Methivier, M. Minier, M. Salmain, and C. Pradier. Grafting of lysozyme and/or poly(ethylene glycol) to prevent biofilm growth on stainless steel surfaces. *The Journal of Physical Chemistry B*, 113(7):2101–2109, 2009.

[45] M. Minier, M. Salmain, N. Yacoubi, L. Barbes, C. Methivier, S. Zanna, and C. Pradier. Covalent immobilization of lysozyme on stainless steel. Interface spectroscopic characterization and measurement of enzymatic activity. *Langmuir*, 21(13):5957–5965, 2005.

[46] N. P. Guerra, A. B. Araujo, A. M. Barrera, A. T. Agrasar, C. L. Macias, J. Carballo, and L. Pastrana. Antimicrobial activity of nisin adsorbed to surfaces commonly used in the food industry. *Journal of Food Protection*, 68(5):1012–1019, 2005.

[47] Z. Shi, K. G. Neoh, S. P. Zhong, L. Y. L. Yung, E. T. Kang, and W. Wang. In vitro antibacterial and cytotoxicity assay of multilayered polyelectrolyte-functionalized stainless steel. *Journal of Biomedical Materials Research Part A*, 76A(4):826–834, 2006.

[48] S. C. Gill and P. H. Vonhippel. Calculation of protein extinction coefficients from amino-acid sequence data. *Analytical Biochemistry*, 182(2):319–326, 1989.

[49] IPCC fourth assessment report: Climate change 2007 synthesis report. Technical report, International Panel on Climate Change, 2007.

[50] Q. Schiermeier. Putting the carbon back: The hundred billion tonne challenge. *Nature*, 442(7103):620–623, 2006.

[51] World energy outlook 2006. Technical report, International Energy Association – Organization for Economic Co-operation and Development, 2006.

[52] IPCC special report: Carbon dioxide capture and storage. Technical report, International Panel on Climate Change, 2005.

[53] E. Kintisch. Power generation - making dirty coal plants cleaner. *Science*, 317(5835):184–186, 2007.

[54] World Coal Institute website. Accessed on October 23, 2007; www.worldcoal.org.

[55] World energy outlook 2004. Technical report, International Energy Association – Organization for Economic Co-operation and Development, 2004.

[56] The future of coal. Technical report, Massachusetts Institute of Technology, Cambridge, MA (United States), 2007.

[57] K. S. Lackner. A guide to CO_2 sequestration. *Science*, 300(5626):1677–1678, 2003.

[58] H. Herzog. Carbon sequestration via mineral carbonation: overview and assessment. Technical report, Massachusetts Institute of Technology, March 14, 2002 2002.

[59] P. Goldberg, Z. Chen, W. O'Connor, R. Walters, and H. Ziock. CO_2 mineralization sequestration studies in U.S. Technical report, National Energy Technology Laboratory, 2000.

[60] G. Falini, S. Albeck, S. Weiner, and L. Addadi. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science*, 271(5245):67–69, 1996.

[61] A. M. Belcher, X. H. Wu, R. J. Christensen, P. K. Hansma, G. D. Stucky, and D. E. Morse. Control of crystal phase switching and orientation by soluble mollusc-shell proteins. *Nature*, 381(6577):56–58, 1996.

[62] Y. Helman, F. Natale, R. M. Sherrell, M. LaVigne, V. Starovoytov, M. Y. Gorbunov, and P. G. Falkowski. Extracellular matrix production and calcium carbonate precipitation by coral cells in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 105(1):54–58, 2008.

[63] H. Tong, P. Wan, W. T. Ma, G. R. Zhong, L. X. Cao, and J. M. Hu. Yolk spherocrystal: The structure, composition and liquid crystal template. *Journal of Structural Biology*, 163(1):1–9, 2008.

[64] H. Miyamoto, T. Miyashita, M. Okushima, S. Nakano, T. Morita, and A. Matsushiro. A carbonic anhydrase from the nacreous layer in oyster pearls. *Proceedings of the National Academy of Sciences of the United States of America*, 93(18):9657–9660, 1996.

[65] B. D. Lee, W. A. Apel, and M. R. Walton. Screening of cyanobacterial species for calcification. *Biotechnology Progress*, 20(5):1345–1351, 2004.

[66] T.J. Phelps, R.J. Lauf, and Y. Roh. Biomineralization for carbon sequestration. Technical report, Oak Ridge National Laboratory, 2003.

[67] Y. Roh, T.J. Phelps, A.D. McMillan, and R.J. Lauf. Utilization of biomineralization processes with fly ash for carbon sequestration. Technical report, National Engergy Technology Laboratory, 2000.

[68] E. T. Boder and K. D. Wittrup. Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology*, 15(6):553–557, 1997.

[69] E. T. Boder and K. D. Wittrup. Yeast surface display for directed evolution of protein expression, affinity, and stability. *Applications of Chimeric Genes and Hybrid Proteins, Pt C*, 328:430–444, 2000.

[70] I. Nishimori, T. Minakuchi, S. Onishi, D. Vullo, A. Scozzafava, and C. T. Supuran. Carbonic anhydrase inhibitors. DNA cloning, characterization, and inhibition studies of the human secretory isoform VI, a new target for sulfonamide and sulfamate inhibitors. *Journal of Medicinal Chemistry*, 50(2):381–388, 2007.

[71] J. C. Kernohan. Ph-activity curve of bovine carbonic anhydrase and its relationship to inhibition of enzyme by anions. *Biochimica Et Biophysica Acta*, 96(2):304, 1965.

[72] K. M. Wilbur and N. G. Anderson. Electrometric and colorimetric determination of carbonic anhydrase. *J. Biol. Chem.*, 176(1):147–154, 1948.

[73] L. Addadi and S. Weiner. Control and design principles in biological mineralization. *Angewandte Chemie-International Edition in English*, 31(2):153–169, 1992.

[74] I. Sarashina and K. Endo. The complete primary structure of molluscan shell protein 1 (MSP-1), an acidic glycoprotein in the shell matrix of the scallop *Patinopecten yessoensis*. *Marine Biotechnology*, 3(4):362–369, 2001.

[75] E. M. Krauland, B. R. Peelle, K. D. Wittrup, and A. M. Belcher. Peptide tags for enhanced cellular and protein adhesion to single-crystal line sapphire. *Biotechnology and Bioengineering*, 97(5):1009–1020, 2007.

[76] K. T. Nam, Y. J. Lee, E. M. Krauland, S. T. Kottmann, and A. M. Belcher. Peptide-mediated reduction of silver ions on engineered biological scaffolds. *ACS Nano*, 2(7):1480–1486, 2008.

[77] S. Elhadj, J. J. De Yoreo, J. R. Hoyer, and P. M. Dove. Role of molecular charge and hydrophilicity in regulating the kinetics of crystal growth. *Proceedings of the National Academy of Sciences of the United States of America*, 103(51):19237–19242, 2006.

[78] R. F. Fakhrullin and R. T. Minullina. Hybrid cellular-inorganic core-shell microparticles: Encapsulation of individual living cells in calcium carbonate microshells. *Langmuir*, 25(12):6617–6621, 2009.

[79] K.Z. House. Invited talk at MIT. 2010.

[80] K. Z. House, C. H. House, D. P. Schrag, and M. J. Aziz. Electrochemical acceleration of chemical weathering as an energetically feasible approach to mitigating anthropogenic climate change. *Environmental Science & Technology*, 41(24):8464–8470, 2007.

[81] D. Zaelke, O. Youn, and S. Andersen. Scientific synthesis of calera carbon sequestration and carbonaceous by-product applications. Technical report, Donald Bren School of Environmental Science and Management, University of California, Santa Barbera, 2011.

[82] T. H. Maren. Carbonic anhydrase kinetics and inhibition at 37 degrees - an approach to reaction rates in vivo. *Journal of Pharmacology and Experimental Therapeutics*, 139(2):129–139, 1963.

[83] A. Radzicka and R. Wolfenden. A proficient enzyme. *Science*, 267(5194):90–3, 1995.

[84] J. A. Dean. *Lange's Handook of Chemistry.* McGraw-Hill, New York, 10th edition, 1973.

[85] J. A. Rakestraw, A. R. Baskaran, and K. D. Wittrup. A flow cytometric assay for screening improved heterologous protein secretion in yeast. *Biotechnology Progress*, 22(4):1200–1208, 2006.

[86] C. Forsman, G. Behravan, A. Osterman, and B. H. Jonsson. Production of active human carbonic anhydrase-II in *Escherichia coli*. *Acta Chemica Scandinavica Series B-Organic Chemistry and Biochemistry*, 42(5):314–318, 1988.

[87] S. Ernst, O. A. Garro, S. Winkler, G. Venkataraman, R. Langer, C. L. Cooney, and R. Sasisekharan. Process simulation for recombinant protein production: Cost estimation and sensitivity analysis for heparinase I expressed in *Escherichia coli*. *Biotechnology and Bioengineering*, 53(6):575–582, 1997.

[88] O. Makarova, E. Kamberov, and B. Margolis. Generation of deletion and point mutations with one primer in a single cloning step. *Biotechniques*, 29(5):970–2, 2000.

[89] G. Chao, W. L. Lau, B. J. Hackel, S. L. Sazinsky, S. M. Lippow, and K. D. Wittrup. Isolating and engineering human antibodies using yeast surface display. *Nature Protocols*, 1(2):755–768, 2006.

[90] Berg J. M. Lippard S. J. *Principles of bioinorganic chemistry.* University Science Books, 1994.

[91] L. J. Wang, N. Sun, S. Terzyan, X. J. Zhang, and D. R. Benson. Histidine/tryptophan pi-stacking interaction stabilizes the heme-independent folding core of microsomal apocytochrome b(5) relative to that of mitochondrial apocytochrome b(5). *Biochemistry*, 45(46):13750–13759, 2006.

[92] D. A. Marvin, R. D. Hale, C. Nave, and M. H. Citterich. Molecular-models and structural comparisons of native and mutant class-I filamentous bacteriophages Ff (Fd, F1, M13), If1 and Ike. *Journal of Molecular Biology*, 235(1):260–286, 1994.

[93] H. Yi. *M13 virus/single-walled carbon nanotubes as a materials platform for energy devices and biomedical applications.* PhD thesis, M.I.T., 2011.

[94] X. N. Dang, H. J. Yi, M. H. Ham, J. F. Qi, D. S. Yun, R. Ladewski, M. S. Strano, P. T. Hammond, and A. M. Belcher. Virus-templated self-assembled single-walled carbon nanotubes for highly efficient electron collection in photovoltaic devices. *Nature Nanotechnology*, 6(6):377–384, 2011.

[95] R. Derda, S. K. Y. Tang, S. C. Li, S. Ng, W. Matochko, and M. R. Jafari. Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules*, 16(2):1776–1803, 2011.

[96] S. K. Lee, D. S. Yun, and A. M. Belcher. Cobalt ion mediated self-assembly of genetically engineered bacteriophage for biomimetic Co-Pt hybrid material. *Biomacromolecules*, 7(1):14–17, 2006.

[97] DNA 2.0 website. Accessed on November 9, 2011; www.dna20.com.

[98] R. Derda, S. K. Y. Tang, and G. M. Whitesides. Uniform amplification of phage with different growth characteristics in individual compartments consisting of monodisperse droplets. *Angewandte Chemie-International Edition*, 49(31):5301–5304, 2010.

[99] D. N. Bolon and S. L. Mayo. Enzyme-like proteins by computational design. *Proceedings of the National Academy of Sciences of the United States of America*, 98(25):14274–14279, 2001.

[100] A. Schmid, J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts, and B. Witholt. Industrial biocatalysis today and tomorrow. *Nature*, 409(6817):258–268, 2001.

[101] H. E. Schoemaker, D. Mink, and M. G. Wubbolts. Dispelling the myths - biocatalysis in industrial synthesis. *Science*, 299(5613):1694–1697, 2003.

[102] B. A. Smith and M. H. Hecht. Novel proteins: from fold to function. *Current Opinion in Chemical Biology*, 15(3):421–426, 2011.

[103] M. A. Fisher, K. L. McKinley, L. H. Bradley, S. R. Viola, and M. H. Hecht. De novo designed proteins from a library of artificial sequences function in escherichia coli and enable cell growth. *Plos One*, 6(1), 2011.

[104] W. F. DeGrado and J. Kaplan. De novo design of catalytic proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 101(32):11566–11570, 2004.

[105] Y. N. Wei and M. H. Hecht. Enzyme-like proteins from an unselected library of designed amino acid sequences. *Protein Engineering Design & Selection*, 17(1):67–75, 2004.

[106] M. H. Hecht, J. S. Richardson, D. C. Richardson, and R. C. Ogden. De novo design, expression, and characterization of felix - a 4-helix bundle protein of native-like sequence. *Science*, 249(4971):884–891, 1990.

[107] B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker. Design of a novel globular protein fold with atomic-level accuracy. *Science*, 302(5649):1364–1368, 2003.

[108] V. Nanda and R. L. Koder. Designing artificial enzymes by intuition and computation. *Nature Chemistry*, 2(1):15–24, 2010.

[109] L. Regan and W. F. Degrado. Characterization of a helical protein designed from 1st principles. *Science*, 241(4868):976–978, 1988.

[110] A. Tramontano, K. D. Janda, and R. A. Lerner. Catalytic antibodies. *Science*, 234(4783):1566–1570, 1986.

[111] K. Q. Chen and F. H. Arnold. Tuning the activity of an enzyme for unusual environments - sequential random mutagenesis of subtilisin-E for catalysis in dimethylformamide. *Proceedings of the National Academy of Sciences of the United States of America*, 90(12):5618–5622, 1993.

[112] C. B. Mao, D. J. Solis, B. D. Reiss, S. T. Kottmann, R. Y. Sweeney, A. Hayhurst, G. Georgiou, B. Iverson, and A. M. Belcher. Virus-based toolkit for the directed synthesis of magnetic and semiconducting nanowires. *Science*, 303(5655):213–217, 2004.

[113] T. E. Creighton. *Proteins: Structures and Molecular Properties.* W. H. Freeman, New York, 2nd edition edition, 1992.

[114] A. Fersht. *Structure and Mechanism in Protein Science.* W. H. Freeman, New York, 1999.

[115] A. S. Khalil, J. M. Ferrer, R. R. Brau, S. T. Kottmann, C. J. Noren, M. J. Lang, and A. M. Belcher. Single M13 bacteriophage tethering and stretching. *Proceedings of the National Academy of Sciences of the United States of America*, 104(12):4892–4897, 2007.

[116] Y. J. Lee, H. Yi, W. J. Kim, K. Kang, D. S. Yun, M. S. Strano, G. Ceder, and A. M. Belcher. Fabricating genetically engineered high-power lithium-ion batteries using multiple virus genes. *Science*, 324(5930):1051–1055, 2009.

[117] A. J. J. Straathof, S. Panke, and A. Schmid. The production of fine chemicals by biotransformations. *Current Opinion in Biotechnology*, 13(6):548–556, 2002.

[118] Y. Pocker and J. T. Stone. Catalytic versatility of erythrocyte carbonic anhydrase. enzyme-catalyzed hydrolysis of p-nitrophenyl acetate. *Journal of the American Chemical Society*, 87(23):5497–5498, 1965.

[119] Y. Pocker and J. T. Stone. Catalytic versatility of erythrocyte carbonic anhydrase. kinetic studies of enzyme-catalyzed hydrolysis of p-nitrophenyl acetate. *Biochemistry*, 6(3):668–678, 1967.

[120] D. S. Auld. Zinc coordination sphere in biochemical zinc sites. *Biometals*, 14(3-4):271–313, 2001.

[121] C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, and S. Duhr. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature Communications*, 1, 2010.

[122] M. Jerabek-Willemsen, C. J. Wienken, D. Braun, P. Baaske, and S. Duhr. Molecular interaction studies using microscale thermophoresis. *Assay and Drug Development Technologies*, 9(4):342–353, 2011.

[123] T. Darbre and J. L. Reymond. Peptide dendrimers as artificial enzymes, receptors, and drug-delivery agents. *Accounts of Chemical Research*, 39(12):925–934, 2006.

[124] K. S. Broo, H. Nilsson, J. Nilsson, A. Flodberg, and L. Baltzer. Cooperative nucleophilic and general-acid catalysis by the HisH(+)-His pair and arginine transition state binding in catalysis of ester hydrolysis reactions by designed helix-loop-helix motifs. *Journal of the American Chemical Society*, 120(17):4063–4068, 1998.

[125] G. E. Host, J. Razkin, L. Baltzer, and B. H. Jonsson. Combined enzyme and substrate design: Grafting of a cooperative two-histidine catalytic motif into a protein targeted at the scissile bond in a designed ester substrate. *Chembiochem*, 8(13):1570–1576, 2007.

[126] J. A. Hunt and C. A. Fierke. Selection of carbonic anhydrase variants displayed on phage - aromatic residues in zinc binding site enhance metal affinity and equilibration kinetics. *Journal of Biological Chemistry*, 272(33):20364–20372, 1997.

[127] M. M. Yamashita, L. Wesson, G. Eisenman, and D. Eisenberg. Where metal-ions bind in proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 87(15):5648–5652, 1990.

[128] S. C. Patel, L. H. Bradley, S. P. Jinadasa, and M. H. Hecht. Cofactor binding and enzymatic activity in an unevolved superfamily of de novo designed 4-helix bundle proteins. *Protein Science*, 18(7):1388–1400, 2009.

[129] R. A. Jensen. Enzyme recruitment in evolution of new function. *Annual Review of Microbiology*, 30:409–425, 1976.