Single Molecule Enzymology Using Femtoliter Arrays

A dissertation submitted by

Pratyusha Mogalisetti

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Chemistry

TUFTS UNIVERSITY

August 2015

Advisor: Professor David R. Walt

Abstract

This thesis demonstrates the application of optical fiber based femtoliter arrays for single molecule enzymology studies. Study of enzymes at single molecule level provides us with mechanistic insights that cannot be obtained using traditional bulk studies. Several approaches for isolating and studying single enzyme molecules have been reported in the literature. These methods are detailed in Chapter 1 with an emphasis on femtoliter array platforms for single molecule isolation.

Chapter 2 describes the application of optical fiber based arrays to study the stochastic binding and release of the inhibitors to single β-galactosidase molecules. The competition between the substrate and inhibitor for the enzyme active site is studied at single molecule level. The effect of inhibition on enzyme activity is also described. In Chapter 3, the α-complementation reaction of β-galactosidase was studied at single molecule resolution. In a novel application of the femtoliter arrays, the numbers of active molecules formed in the complementation reaction were directly counted and the stoichiometry of complementation reaction was determined.

Chapter 4 discusses the effect of temperature on single molecules of wild type and single surface cysteine variants of β-glucuronidase. The identification of the thermal switching property of a fraction of the enzyme population in some variants is reported. Finally, chapter 5 discusses the current progress towards identifying hetero-oligomers with unnatural backbones that can catalyze notoriously difficult reactions. The approach of screening large oligomer libraries for the presence of rare catalysts using the optical fiber femto-well array platform is also described.

To my beloved parents, Renuka and Suresh Mogalisetti, who have always supported me.

Acknowledgements

I would like to begin by thanking my PhD advisor, Prof. David Walt, for his constant support and encouragement throughout my graduate program. He has constantly provided me with guidance during my research and has also given me the intellectual freedom for my projects which has helped me develop into an independent researcher. I would like to also thank my committee members, Prof. Krishna Kumar and Prof. Joshua Kritzer, for providing me with critical feedback and helpful discussions throughout my program. Additionally, I am grateful to Prof. Jeff Gelles from Brandeis University for serving as an external member on my thesis committee.

This thesis would not have been possible without the contributions of several individuals outlined below. I would like to thank Prof. Volker Jäger from Stuttgart University for providing me with inhibitors used in Chapter 2 and past lab members, Dr. Marcin Rojek and Dr. Hans Gorris, who helped with the experimentation and data analysis for the project discussed in this chapter. I would also like to acknowledge DiscoveRx Corporation for generously providing me with the α-donor and α-acceptor samples used in Chapter 3 and thank Dr. Edwin Ullman for providing several helpful discussions for this project.

I acknowledge and thank our collaborators at University of Texas, Austin, Prof. Andrew Ellington and his students Dr. Christien Kluwe and Jimmy Gollihar, for designing and expressing the β-glucuronidase variants used in Chapter 4 and for their invaluable input and discussions for this project. I would also like to thank Dr. Marcin Rojek for contributing to experiment design and analysis in this project. I sincerely thank the collaborators at University of Wisconsin Madison - Prof. Samuel Gellman, Dr. Zvika Hayouka and Dr. Mellissa MacDonald –

for designing and synthesizing polymers used in Chapter 5 and for several helpful detailed discussions. I thank my colleagues, Dr. Shudan Bian, Dr. Barrett Duan and Dr. Yael Zilberman. for their contribution to this project. I would like to acknowledge Dr. Shonda Gaylord, Dr. Mark Hartman and Prarthana Khanna for their critical review of the chapters of this thesis.

I would also like to extend my deepest thanks to the past and current members of the Walt lab who have been an absolute pleasure to work with throughout my graduate school experience. I would like to acknowledge their valuable input in group meetings and their help with several qualification exams. I thank the members of the Kumar lab and the Kritzer lab for their enthusiastic discussions and input for my committee meetings. I thank my friends and colleagues Dr. Marcin Rojek, Dr. Shonda Gaylord, Dr. Elena Benito Pena, Dr. Milena Milutinovic, Trinh Dinh and Prarthana Khanna. Not only have they helped me with critical discussions on my projects, but they made graduate school fun both inside and outside of lab. I thank Ralph Hunnewell for his immense support and encouragement during the preparation of this thesis.

I would like to thank my undergraduate mentor, Prof. Anju Chadha, who initially got me interested in enzymology and encouraged me to pursue research. I thank my high school math teachers Koteswara Rao and Rajashekarra Reddy, who have taught me to be critical and inquisitive and who have always been an inspiration to me. I would not be able to succeed in my graduate studies without them.

Finally, I would like to extend my sincere and heartfelt thanks to my parents, Renuka and Suresh Mogalisetti, for their constant love and support.

Table of Contents

Appendix

CHAPTER 1

Single Molecule Enzymology in Femtoliter Arrays

1.1 Introduction

Enzymes are often referred to as workhorses of cells as they catalyze all the reactions required for a cell to survive and function. The remarkable and unrivalled performance of enzymes in terms of catalytic efficiency, substrate selectivity and substrate specificity has found use in several applications. Enzymes are used in industrial catalysis, recombinant DNA technology and diagnostic and therapeutic applications. In addition, enzymes are also drug targets for several diseases. Due to the omnipresent role of enzymes in biology and their commercial use, understanding the structure and function of enzymes as well as the basis of enzyme catalysis has always been a major research goal in biochemistry.

Traditional enzymology relies on the observation of activity of a population of enzymes. Typically, thousands of enzyme molecules are incubated with an excess of fluorogenic or chromogenic substrate and the turnover of the substrate into product is monitored over time. This approach to studying enzymes provides an average behavior of the whole population. The activity measured at any time point, is the average activity of the population. Any heterogeneity of activity within the population of enzymes or transient kinetic events are masked within the averaged activity. Observing the activity of an individual enzyme molecule makes it possible to identify variations in enzyme activity over time and correlate the transient kinetic events to the mechanism of catalysis. The advent of single molecule enzymology has led to several insightful discoveries about enzyme catalysis over the last 20 years. Single molecule enzymology methods have enabled researchers to study the mechanisms of motor proteins, DNA transcription and ATP hydrolysis[1-3]. Single molecule methods to identify active site conformational dynamics associated with enzyme action were demonstrated[4,5]. The application of single molecule techniques to study static and dynamic heterogeneities in the activity of single enzyme

molecules, interaction of substrates with enzymes and allosteric enzyme inhibition was demonstrated[6-9]. To date, single molecule methods have been applied to the study of DNA and RNA enzymes, motor proteins, oxido-reductases and hydrolases. Several reviews have been written summarizing these single molecule studies[10-14]. In this chapter, we will focus on single molecule studies performed in femtoliter arrays.

The first experiment of isolating single enzyme molecules and measuring its activity was reported by Rotman in 1961[15]. To isolate single enzymes, Rotman sprayed a solution containing ß-galactosidase and a fluorogenic substrate into an oil chamber. By using a dilute solution of enzyme, Rotman was able to generate water-in-oil droplets (ø = 14-15 μm) with single enzyme molecules in them (Figure 1). By conducting these experiments, Rotman was able to determine the molecular weight of the enzyme and the kinetics of the individual enzyme molecules. The general approach of encapsulating highly diluted enzyme solutions along with a fluorogenic substrate demonstrated in this innovative paper was applied in single molecule enzymology as well as in single cell papers published much later[16,17]. These studies are discussed in further detail in Section 1.2. At present, single molecule detection technologies have come a long way from the water-in-oil droplets technique, but this approach remains a simple and powerful tool to study the kinetic events of single enzymes at a relatively larger time scale. This approach is employed in the single molecule isolation technique used in this thesis.

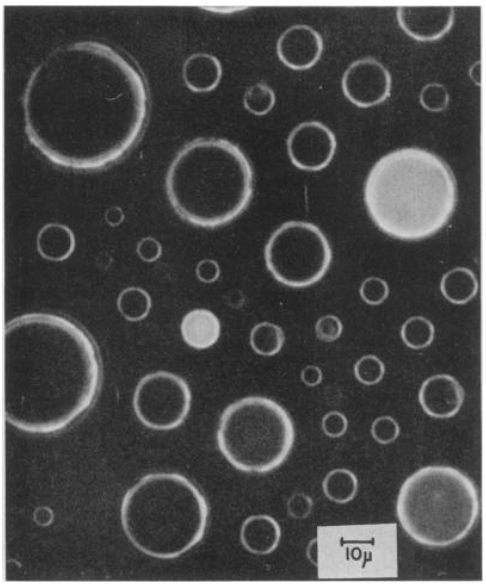


Figure 1.1 Oil in water droplets were generated by spraying an enzyme/substrate solution into oil. A very dilute solution of enzyme was used to ensure no more than one enzyme per bead. Most droplets did not have any enzyme in them and a few droplets had a single molecule of enzyme. The droplets containing the enzyme molecule could be detected due to the generation of fluorogenic product. Figure adapted from Reference 15.

Current single molecule detection techniques can observe single turnovers of enzymes. Funatsu *et al.* first demonstrated the detection of single fluorophore using total internal reflection microscopy (TIRF). Detection of single fluorophores enables observation of dynamic heterogeneity in enzyme activity and the conformational dynamics of enzymes. Using TIRF, Funatsu *et al.* were able to observe individual turnovers of ATPase[18]. To perform single turnover resolution experiments, enzymes were immobilized on a surface or in a matrix. Funatsu et al. used a fluorescent substrate so that a fluorescence signal could be detected when the substrate was bound to the enzyme. When the product was generated due to the enzyme, it diffused away from the surface resulting in a loss of fluorescence. (Figure 2a) The fluctuation in fluorescence could therefore be correlated to individual enzyme turnover. Alternately, the conversion of a single fluorogenic substrate into a fluorescent product molecule at the surface could be observed

using TIRFM or confocal microscopy (Figure 2)[6]. In these single molecule measurements, the conversion of substrate into product was observed as a burst in fluorescence signal. The turnover events are stochastic and analysis of the time between two turnover events (waiting time) provides information about the reaction intermediates in the catalytic cycle. In addition, these measurements enable measurement of the variation of reaction velocity over time (dynamic heterogeneity of enzyme activity or dynamic disorder). This dynamic disorder has been attributed to the fluctuation of the enzyme between different conformations. Conformational dynamics of the enzyme can be studied by attaching fluorescent labels to different domains of the enzyme and monitoring the fluorescence resonance energy transfer (FRET) between the fluorophores and correlating it with the enzyme turnover.

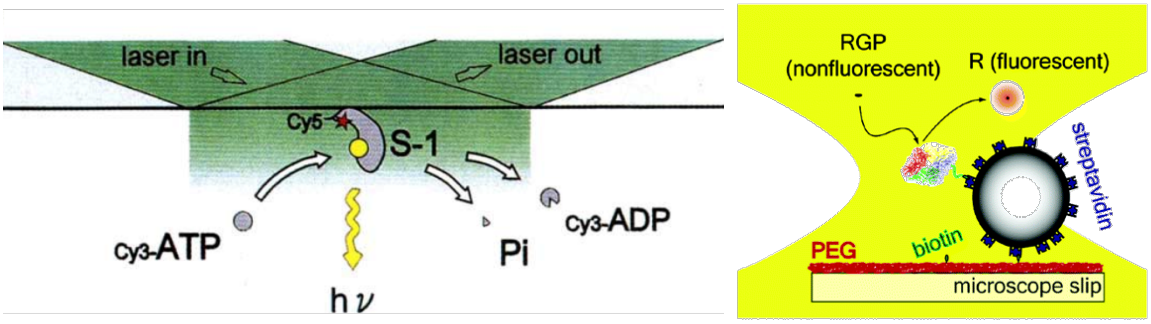


Figure 1.2 Schematics showing single turnover detection systems. (a) TIRFM was used to monitor the turnovers of single ATPase molecule immobilized on a cover slip. A laser beam incident on the coverslip undergoes total internal reflection, generating an evanescent wave at the surface of the coverslip (shown in green). The evanescent wave decays exponentially from the surface and therefore only excites fluorophores close to the surface. As a result, only fluorophores close to the interface will be detected. In this experiment, the ATPase molecules immobilized on the surface are labeled with Cy5 fluorophore so that their location on the coverslip can be pinpointed with the signal observed in Cy5 channel. The substrate molecules are labeled with Cy3 fluorophore. When the substrate is bound to the enzyme at the interface, Cy3 fluorescence can be detected. When the substrate is converted

into the product, it diffuses away resulting in a loss of fluorescence. Therefore individual turnover events can be detected by monitoring the fluorescence near the labeled enzyme. (Figure adapted from Reference 18). (b) Confocal microscopy was used to monitor turnovers of single molecules of β-galactosidase immobilized on a microsphere. Confocal microscopy uses a pinhole to limit the excitation to the focal plane as well as a pinhole to collect emission only from the focal plane. The resulting focal volume of detection is very small. In the schematic here, the focal volume is positioned near the microsphere. When the enzyme converts the fluorogenic substrate into a product, a burst in fluorescence can be detected before the product diffuses away from the focal volume. The bursts in fluorescence, therefore, indicate a turnover event. (Figure adapted from Reference 6).

1.2 Single molecule isolation methods

Isolation of single enzyme molecules can be achieved either through immobilization of the enzyme on a surface or encapsulation within a reaction chamber. Immobilization of the enzyme on a surface enables measurements at the millisecond level time scale using sensitive optical methods like TIRFM and confocal microscopy, as discussed previously. However, immobilization can potentially alter the activity of the enzyme. In addition, the use of laser in TIRFM and confocal microscopy can result in photobleaching of the fluorophores which limits the time for which single enzymes can be observed. Encapsulation of single enzyme molecules within a reaction chamber allows the enzyme to be free in solution, resulting in measurement of unaltered activity. However, since the enzyme is free to move in solution, optical methods such as TIRFM or confocal microscopy cannot be employed. The enzyme is no longer restricted to the small focal plane or the focal volume that is essential for these methods. Enzyme activity is usually measured by an accumulation of fluorescent product and, therefore, reduces the time resolution to the order of seconds. Single turnover events or conformational dynamics cannot be

detected; however, static disorder and perturbations to enzyme activity over longer time scales can be studied as photobleaching is not a limiting problem.

1.3 Isolation of single enzyme molecule in containers

1.3.1 Poisson Statistics

Poisson statistics are employed to ensure encapsulation of single enzyme molecules within the containers in all the methods discussed below. Unlike immobilization methods discussed above, where the focal plane or focal volume ensures observation of single enzyme molecules, the methods discussed below, use wide field fluorescence microscopy to monitor the reaction chambers. Therefore, reducing the probability of encapsulating more than one enzyme molecule is critical. In order to achieve this, a highly diluted enzyme solution is used to fill the reaction chambers. Poisson statistics can be used to determine the probabilities of having *n* molecules within a reaction chamber when a specific enzyme concentration is used. The enzyme concentration is chosen such that the probability at *n = 2* is negligible. First the ratio, μ, of enzymes to wells at the enzyme concentration is determined.

$$\mu = \frac{concentration\ x\ volume\ of\ a\ well\ x\ number\ of\ wells\ x\ Avagadro\ number}{number\ of\ wells}$$

$$\mu = concentration\ x\ volume\ of\ a\ well\ x\ Avagadro\ number \qquad \text{Equation 1}$$

At the ratio μ, the probability of having n enzymes in a reaction chamber can be calculated using the poisson equation:

$$P_{\mu}(n) = \frac{e^{-\mu}\mu^{n}}{n!} \qquad \text{Equation 2}$$

The calculated poisson probabilities P(n) for different ratios of enzymes to wells are shown in Table 1.1. At low ratios of enzymes to wells (0.01, 0.05 and 0.1), the probability of having 2 or more enzymes per well is very low. At these ratios, more than 90 % of wells are expected to have zero enzymes in them. Therefore, most of the reaction chambers are not expected to show any signal. A small fraction of the wells (1 – 10%) are expected to have one enzyme in the well. These wells provide kinetic data of single enzyme molecules. At enzyme to well ratios of 1 or more, the number of wells expected to have 2 enzymes is over 18% and is not negligible anymore. The ratio of enzymes to wells must be chosen such that there are a statistically significant number of wells containing 1 enzyme and negligible wells containing 2 enzymes. For example, if an array has a total of 100,000 wells, using a low ratio of 0.01, results in a 1000 wells containing single enzymes and less than 5 wells containing 2 enzymes. On the other hand, if the array only has a 1000 wells, using a ratio of 0.01 results in only 10 wells with single enzymes, which do not provide enough statistics for single enzyme measurements. A higher ratio of 0.1, results in about 100 wells with single enzymes and less than 5 wells with 2 enzymes.

Ratio	P(0)	P(1)	P(2)	P(3)
0.01	0.99	0.01	4.95×10^{-05}	1.65×10^{-07}
0.05	0.95	0.05	1.19×10^{-03}	1.98×10^{-05}
0.1	0.90	0.09	4.5×10^{-03}	1.51×10^{-04}
1	0.37	0.37	0.18	0.06
2	0.14	0.27	0.27	0.18

Table 1. Probability of finding 0, 1, 2 or 3 enzyme molecules in a well at different ratios of enzymes to wells. The probabilities were calculated using Equation 2.

1.3.2 Droplet methods

Several methods of isolating single enzyme molecules in chambers have been used since Rotman's oil-in-water droplet technique. An improved droplet method was used for detection of single molecules of α-chymotrypsin[19]. In this experiment, the droplets were generated using a blender to create an emulsion of the enzyme/substrate solution in silicone oil followed by loading the droplets into microfabricated PDMS wells. In both experiments, the micro droplets generated had a distribution in size, which had to be taken into account for analysis. A relatively more mono disperse and smaller droplets (ø = 2.5-3 μm) were generated using a microfluidic chip to isolate single molecules of β-glucosidase[20]. Droplets were created using lipid bilayers (liposomes) to encapsulate enzymes. Liposomes containing single labeled molecule of alkaline phosphatase and substrate, respectively, were generated using the rotary evaporation method, followed by electrofusing enzyme and substrate liposomes to observe the activity of individual enzymes[21]. Unlike the water-in oil droplets, liposomes could be generated with either positive or negative charge and, therefore, could be manipulated by voltage. The ability to fuse the liposome containing enzyme and substrate respectively enabled the researchers to introduce different reaction components. Liposomes as small as 50 nm diameter (small unilamellar vesicles) were generated to isolate alkaline phosphatase enzyme and study its activity[22]. Large unilamellar vesicles of ~125 nm diameter were used to trap single molecules of HRP and study its allosteric inhibition[23]. Droplet encapsulation strategies enable observation of single enzymes for long periods of time. Since the droplets have to be freshly generated with each experiment, care must be taken to characterize the droplets each time. Comellas-Aragones et al., overcame this drawback by using a virus capsid for encapsulation of horseradish peroxidase (HRP)[24]. The cowpea chlorotic mottle virus (CCMV) capsid has a well-defined dimension (inner ø = 18 nm,

Figure 3) and can be disassembled and reassembled by switching the pH of the solution. Disassembling the capsid and reassembling in the presence of a highly diluted solution of enzyme resulted in single HRP molecules enclosed within the capsids. Owing to the very small volume of the capsids, the millisecond time scale measurement was achieved using a confocal microscope, although, it was not possible to observe individual enzymatic turnovers due to product accumulation within the capsids.

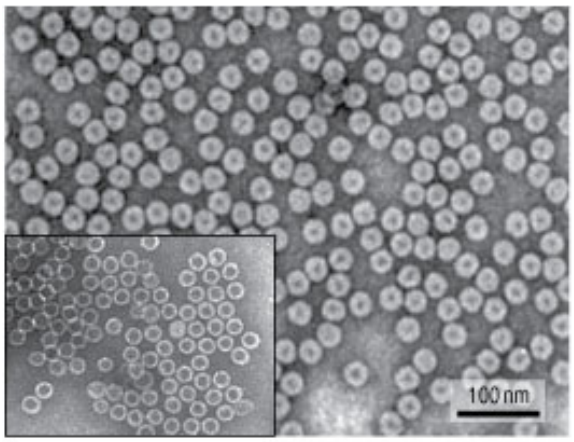


Figure 1.3 Transmission electron microscopy staining of CCMV virus. The inset shows empty capsids. Compared to the water-in-oil droplets the virus capsids are uniform in size. Figure adapted from Reference 24.

1.3.3 Femtoliter arrays

A relatively simpler method of encapsulating enzymes within containers is the use of pre-fabricated wells. Arrays of femtoliter sized reaction chambers have been generated on quartz[25], PDMS[26-28] and silica[29,30] surfaces using standard lithography techniques. In this thesis, femtoliter arrays generated on optical fiber bundles are used for single molecule encapsulation (Section 1.4). Generating femtoliter arrays on optical fiber bundles is relatively fast and does not require lithography techniques[8]. In all the femtoliter array techniques, the reaction chambers are filled with the enzyme/substrate solution and then sealed so that the fluorescent product generated is

contained within the wells. The fluorescence intensity of the wells containing the enzyme increases with time due to product accumulation. Uniform size of the reaction wells and the ability to simultaneously monitor up to thousands of enzyme molecules are the main advantages of the femtoliter arrays. Femtoliter arrays have been used to encapsulate single molecules of β-galactosidase, HRP and β-glucuronidase and α-chymotrypsin. As the enzyme is not immobilized in solution, single turnover resolution imaging is not possible. However, static disorder and transient kinetic events occurring in the time scale of seconds can be studied. Imaging can be performed for long periods of time, as photobleaching is not a limiting factor. Finally, very low concentrations of enzymes can be detected, making these techniques adaptable to ultrasensitive detection of biomolecules using an enzymatic label[31].

1.3.4 Capillary electrophoresis:

Isolation of single enzyme molecules via capillary electrophoresis is another method that does not require enzyme immobilization. The capillary was filled with very dilute solution of enzymes to ensure that individual enzymes are spatially separated from each other along the capillary length. In the presence of a fluorogenic substrate, detectable fluorescent product accumulated in the vicinity of the enzyme after a certain incubation time[17,32,33]. Therefore, discrete zones of product molecules were generated along the length of the capillary. The accumulated product was then separated from the enzyme with electrophoresis. A detector positioned downstream of the capillary tube was used to quantify the amount of product generated. Alternatively, the single molecule experiment were performed in a continuous flow format to observe the activity of enzyme over time[34]. In this case, due to the difference in electrophoretic mobility of the product and the enzyme, either the product or the enzyme moved faster along the capillary when compared to the enzyme. Therefore, the enzyme was constantly

separated from the product. The concentration of the product, which is proportional to enzyme activity, could be measured as it flowed through the detector. As a result a time profile of enzyme activity could be obtained. Capillary electrophoresis is an excellent technique to observe the activity of a single enzyme molecule in different solvent conditions as the solvent around the enzyme can be changed via electrophoresis. This method was applied to study the activity of single molecules of alkaline phosphatase and β-galactosidase at different temperatures[32,34]. The main drawback of this technique is that relatively fewer enzyme molecules can be studied when compared to femtoliter arrays.

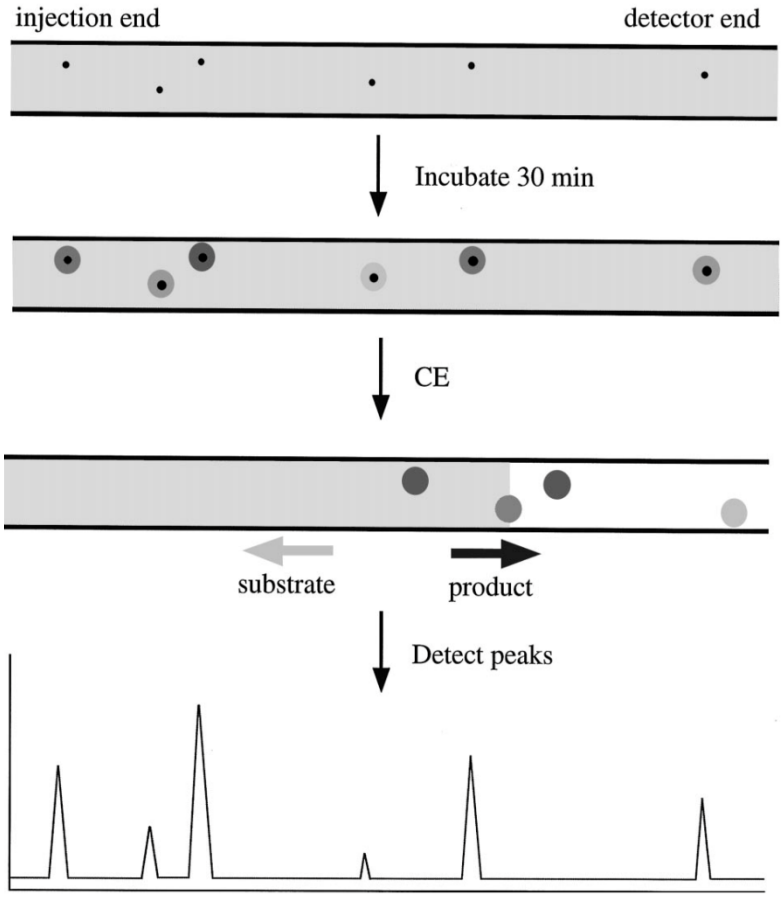


Figure 1.4 Schematic showing capillary electrophoresis based measurement of single enzyme activity[35]. The enzyme was highly diluted into a solution of substrate and injected into the capillary (a). Due to the low concentration of the enzymes, they were spatially

separated along the length of the capillary. The enzyme/substrate solution was incubated in the capillary till fluorescent product accumulated in the vicinity of each enzyme (b). The product was then pulled towards the detector using electrophoresis (c). Each product zone appeared as a peak in the detector (d). The area under the peak indicates the activity of individual fiber. Figure adapted from Reference 35.

1.4 Single molecule studies using optical fiber platform

Optical fiber bundle femtoliter arrays have been used in the past for sensitive bead based detection of analytes such as nucleic acids and proteins. More recently, the femtoliter arrays were used to isolate and analyze single molecules of β-galactosidase[36]. Since then, the optical fiber bundle femtoliter array has been used to study the static heterogeneities of single molecules of β-galactosidase[8], single molecule inhibition of β-galactosidase[9], mechanistic studies of HRP and thermodynamic studies of β-galactosidase. The single enzyme molecule detection in the femtoliter array was used to develop a digital format of the widely used enzyme linked immunosorbent assays (ELISA) [31]. The digital ELISA has been successfully applied for detection of sub-femtomolar concentrations of disease biomarkers. In this thesis, the optical fiber bundle femtoliter arrays are used to study enzyme inhibition at single molecule level, thermodynamic stability of β-glucuronidase variants and for the determination of stoichiometry of the α-complementation reaction of β-galactosidase. In this section, the optical fiber bundle platform used in the rest of the chapters is described in detail.

1.4.1 Optical fibers bundles

Optical fibers are made up of a core and a cladding, which are composed of two different types of glass. The refractive index of the core is higher than the refractive index of the cladding. When light travelling through the core is incident on the cladding, at an angle greater than the

critical angle ϕ_c, it cannot pass through the cladding and is reflected into the core (Figure 1.5a). This phenomenon is called total internal reflection. As a result, light entering one end of an optical fiber is transmitted with minimal loss to the other end of the optical fiber, as long as it is incident on the cladding at an angle greater than ϕ_c. Optical fibers are able to efficiently transmit light of different wavelengths through long distances with minimal loss. Due to these features, optical fibers are widely used for data transmission applications such as telecommunications and the internet.

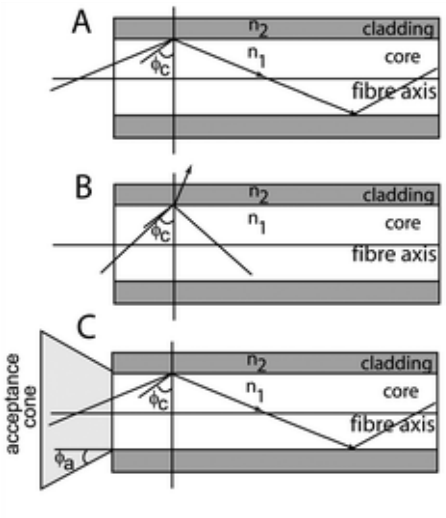


Figure 1.5 (a) Schematic showing an optical fiber[37]. The refractive index of the core, n_1, is greater than the refractive index of the cladding, n_2. Light travelling through the core incident on the cladding at an angle greater than the critical angle ϕ_c undergoes total internal reflection. (b) Light incident at an angle smaller than the critical angle is transmitted through the cladding. (c) The acceptance cone is defined by the entire light incident on the cladding at an angle greater than ϕ_c. Only light from the acceptance cone is transmitted along the length of the fiber. Figure adapted from Reference 37.

Optical fibers can have core diameters ranging from a couple of μm to hundreds of μm. In our laboratory, sensing and detection applications fibers as small as 1-15 micron core diameters are used. Thousands of fibers are fused together to form an array of optical fibers, commonly

referred to as an optical fiber bundle, are obtained from Schott. Apart from the optical transmission properties discussed above, the fiber bundles have two unique features that make them particularly suitable for single molecule applications.

(a) Differential etching of core and cladding: The core and cladding of the optical fibers are made of differently doped glass. The etching rate of core in acid is faster than the etching rate of cladding. As a result, when the fiber bundle is submerged in acid, faster etching of the core results in the formation of wells[38]. The size of the wells formed depends on the concentration of acid used for etching and core diameter. In this thesis, the optical fiber bundles consisting of fibers with 4.5 μm core diameter are used. Etching in 0.025 N HCl for 1'55'' results in 2.9 μm deep wells with a well volume of 46 fL[36]. Therefore, a micro or nano well can be fabricated without any lithography techniques by a simple wet etching process. In addition, the etched end of a fiber bundle can be easily polished with a lapping film to remove the wells and re-etched to generate fresh wells. This is particularly useful when different surface modifications of the fiber wells are employed for different experiments.

(b) High-density arrays. Fiber bundles with very high density of optical fiber can be manufactured. For example, typical fiber bundles used in this thesis have ~ 50,000 individual fibers fused together within a hexagonal cross section of 2mm side length. The whole bundle can be imaged with a 5x objective. Therefore, when single enzymes are encapsulated in the etched femtowells such that there average number of enzymes per well, μ is 0.1, up to 5000 single enzymes molecules can be imaged simultaneously. This feature is extremely important in single molecule enzymology studies.

1.5 Fundamental enzymology studies using optical fiber platform

Recently, the use of optical fiber bundles to study the behavior of single enzymes has been demonstrated with different model systems[8,9,39,40]. Rissin *et al.* used optical fiber bundles to study the activity of single molecules of β-galactosidase[36]. They demonstrated that single molecules of β-galactosidase had a large variation in the activity. This variation in activity could be attributed to the fact that single enzyme molecules within a population can have different conformations resulting in different activities. The inhibition of β-galactosidase was studied at single molecule level[9]. D-galactal was shown to be a cooperative inhibitor of β-galactosidase. Inhibitor binding and release events were shown to be accompanied by conformational changes in the enzyme active site. The optical fiber bundle was used to study the mechanism of HRP. Finally, thermodynamic properties of single molecules of β-galactosidase were studied[40]. The enzyme molecules were shown to interconvert between different conformations when repeatedly exposed to short heat pulses. The enzymes were shown to randomly change their activity after every heat pulse. The average activity of the population remained the same.

In this thesis, the optical fiber bundle platform is used to study several features of fundamental enzymology that could not be obtained from traditional bulk studies. The inhibition of β-galactosidase with two reversible slow binding inhibitors is elucidated at the single molecule level. The competition between inhibitor and substrate for the enzyme active site is demonstrated at the single molecule level. In addition, the competition between two substrates for the enzyme active site is also reported. Using the fiber bundles, the stoichiometry of α-complementation of β-galactosidase is deduced. The mechanism behind the thermostability of several variants of β-glucuronidase is studied. The inhibition of β-glucuronidase by a non-competitive inhibitor is observed at the single molecule level. Finally, the application of fiber

bundle arrays to screen oligomer libraries for the presence of rare biomimetic catalysts is discussed.

1.7 References

(1) Smiley, R. D.; Hammes, G. G. *Chemical reviews* **2006**, *106*, 3080.

(2) Ishii, Y.; Yanagida, T. *Single Molecules* **2000**, *1*, 5.

(3) Puchner, E. M.; Gaub, H. E. *Annual review of biophysics* **2012**, *41*, 497.

(4) Lu, H. P. *Chemical Society reviews* **2014**, *43*, 1118.

(5) Min, W.; English, B. P.; Luo, G.; Cherayil, B. J.; Kou, S. C.; Xie, X. S. *Accounts of Chemical Research* **2005**, *38*, 923.

(6) Xie, S. *Single Molecules* **2001**, *2*, 229.

(7) Lu, H. P.; Xun, L.; Xie, X. S. *Science* **1998**, *282*, 1877.

(8) Rissin, D. M.; Gorris, H. H.; Walt, D. R. *Journal of the American Chemical Society* **2008**, *130*, 5349.

(9) Gorris, H. H.; Rissin, D. M.; Walt, D. R. *Proceedings of the National Academy of Sciences* **2007**, *104*, 17680.

(10) Gershenson, A. *Current Opinion in Chemical Biology* **2009**, *13*, 436.

(11) Deniz, A. A.; Mukhopadhyay, S.; Lemke, E. A. *Journal of The Royal Society Interface* **2008**, *5*, 15.

(12) Ishii, Y.; Kitamura, K.; Tanaka, H.; Yanagida, T. *Methods in enzymology* **2003**, *361*, 228.

(13) Weiss, S. *Nature structural biology* **2000**, *7*, 724.

(14) Jorgensen, S. K.; Hatzakis, N. S. *Biophysical Reviews and Letters* **2013**, *08*, 137.

(15) Rotman, B. *Proceedings of the National Academy of Sciences of the United States of America* **1961**, *47*, 1981.

(16) Cai, L.; Friedman, N.; Xie, X. S. *Nature* **2006**, *440*, 358.

(17) Xue, Q.; Yeung, E. S. *Nature* **1995**, *373*, 681.

(18) Funatsu, T.; Harada, Y.; Tokunaga, M.; Saito, K.; Yanagida, T. *Nature* **1995**, *374*, 555.

(19) Lee, A. I.; Brody, J. P. *Biophysical Journal* **2005**, *88*, 4303.

(20) Arayanarakool, R.; Shui, L.; Kengen, S. W.; van den Berg, A.; Eijkel, J. C. *Lab on a Chip* **2013**, *13*, 1955.

(21) Hsin, T.-M.; Yeung, E. S. *Angewandte Chemie International Edition* **2007**, *46*, 8032.

(22) Christensen, S. M.; Bolinger, P.-Y.; Hatzakis, N. S.; Mortensen, M. W.; Stamou, D. *Nat Nano* **2012**, *7*, 51.

(23) Piwonski, H. M.; Goomanovsky, M.; Bensimon, D.; Horovitz, A.; Haran, G. *Proceedings of the National Academy of Sciences* **2012**, *109*, E1437.

(24) Comellas-Aragones, M.; Engelkamp, H.; Claessen, V. I.; Sommerdijk, N. A. J. M.; Rowan, A. E.; Christianen, P. C. M.; Maan, J. C.; Verduin, B. J. M.; Cornelissen, J. J. L. M.; Nolte, R. J. M. *Nat Nano* **2007**, *2*, 635.

(25) Tan, W.; Yeung, E. S. *Analytical Chemistry* **1997**, *69*, 4242.

(26) Chen, A. Y.; Jani, A. S.; Zheng, L.; Burke, P. J.; Brody, J. P. *Biotechnology Progress* **2009**, *25*, 929.

(27) Rondelez, Y.; Tresset, G.; Tabata, K. V.; Arata, H.; Fujita, H.; Takeuchi, S.; Noji, H. *Nat Biotech* **2005**, *23*, 361.

(28) Ota, S.; Kitagawa, H.; Takeuchi, S. *Analytical Chemistry* **2012**, *84*, 6346.

(29) Ehrl, B. N.; Liebherr, R. B.; Gorris, H. H. *Analyst* **2013**, *138*, 4260.

(30) Liebherr, R. B.; Renner, M.; Gorris, H. H. *Journal of the American Chemical Society* **2014**, *136*, 5949.

(31) Rissin, D. M.; Kan, C. W.; Campbell, T. G.; Howes, S. C.; Fournier, D. R.; Song, L.; Piech, T.; Patel, P. P.; Chang, L.; Rivnak, A. J.; Ferrell, E. P.; Randall, J. D.; Provuncher, G. K.; Walt, D. R.; Duffy, D. C. *Nat Biotech* **2010**, *28*, 595.

(32) Craig, D. B.; Arriaga, E. A.; Wong, J. C. Y.; Lu, H.; Dovichi, N. J. *Journal of the American Chemical Society* **1996**, *118*, 5245.

(33) Shoemaker, G. K.; Juers, D. H.; Coombs, J. M.; Matthews, B. W.; Craig, D. B. *Biochemistry* **2003**, *42*, 1707.

(34) Craig, D. B.; Nichols, E. R. *ELECTROPHORESIS* **2008**, *29*, 4298.

(35) Craig, D. B.; Dovichi, N. J. *Canadian Journal of Chemistry* **1998**, *76*, 623.

(36) Rissin, D. M.; Walt, D. R. *Nano Letters* **2006**, *6*, 520.

(37) Walt, D. R. *Chemical Society Reviews* **2010**, *39*, 38.

(38) Pantano, P.; Walt, D. R. *Chemistry of Materials* **1996**, *8*, 2832.

(39) Gorris, H. H.; Walt, D. R. *Journal of the American Chemical Society* **2009**, *131*, 6277.

(40) Rojek, M. J.; Walt, D. R. *PLoS ONE* **2014**, *9*, e86224.

CHAPTER 2

Single Molecule Inhibition of β-Galactosidase

2.1 Introduction

Competitive enzyme inhibition is a fundamental mechanism in biochemistry. Bulk kinetic analyses, however, only give a limited picture of enzyme reaction mechanisms, as they often assume the same behavior for all enzyme molecules in a population. Single molecule studies of enzymes provide a high resolution picture of the underlying behavior of individual molecules that gives rise to their bulk properties. While bulk studies only provide ensemble averaged properties, single molecule studies shed light on the inherent heterogeneity between different molecules[1-7] and within one molecule over time[3, 8-15]. Such studies also enable the detection of transient kinetic and dynamic events[16-23] that cannot be detected in bulk systems.

In this paper, the competition between inhibitor and substrate molecules for the active sites of hundreds of individual enzyme molecules is studied in arrays of femtoliter-sized wells etched into the end of an optical fiber bundle[5, 7, 21]. Specifically, we observe and compare stochastic binding and release of two inhibitors, D-galactal and NpBHC, to the tetrameric enzyme β-galactosidase to determine if the inhibitors bind or release in a sequential manner. We compare the enzyme activity before and after inhibitor binding and release to investigate if inhibitor binding changes the enzyme conformation. Finally, using autocorrelation analysis the inhibitor exchange rates of the enzymes are correlated to the substrate turnover rates.

The inhibition of β-galactosidase by D-galactal has been extensively studied both in bulk and at the single molecule level[21, 24-26]. D-galactal binds to β-galactosidase to form a covalent glycosyl-enzyme adduct, which dissociates slowly to regenerate the free enzyme and release the product 2-deoxygalactose, with an off-rate of ~2×10^{-3} s^{-1} [21, 24]. The

inhibition has characteristic slow binding kinetics with a reported on-rate constant of ~$4x10^2$ $M^{-1}s^{-1}$ [24]. Using an optical fiber bundle set up it was demonstrated previously that the release of D-galactal from the tetrameric enzyme β-galactosidase occurs in a single step[21]. It was found that the subunits of β-galactosidase release D-galactal in a cooperative manner. In addition, the binding and release of D-galactal was demonstrated to induce a conformational change of the enzyme resulting in a change in the enzyme activity. Therefore, the enzyme does not always revert to the activity it had before inhibitor binding. The activity profile of the enzyme in the presence of inhibitor thus exhibits a fluctuation between distinct activity levels.

NpBHC is another potent inhibitor of β-galactosidase modelled after mannostatin, which is a strong inhibitor of mannosidase[27]. It is a reversible transition state inhibitor with an inhibition constant of 0.5 nM (calculated from bulk experiments). Bulk studies show that NpBHC is a tight binding inhibitor with an on-rate constant of $4.5x10^5$ $M^{-1}s^{-1}$ and an off-rate of $2x10^{-3}$ s^{-1}, estimated by fitting the steady state progress curves to the integrated equation described in[28]. Therefore, under steady state conditions, NpBHC has a low dissociation and association rate on the order of 10^{-3} s^{-1}, which makes it ideal for observing the binding and release from β-galactosidase because inhibitor binding events can be readily distinguished from substrate turnover events. In this paper, we study the single molecule inhibition of β-galactosidase by NpBHC and compare it to inhibition with D-galactal.

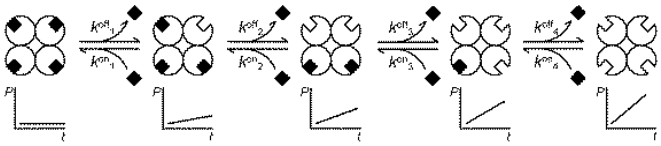
2.2 Results and discussion

2.2.1 Pre-steady state experiments

First, in analogy to the single molecule studies on inhibition by D-galactal[21], the inhibition of β-galactosidase by NpBHC was investigated to determine if β-galactosidase

subunits release NpBHC in an independent or cooperative manner. In a pre-steady state experiment, all active sites of the enzyme are first blocked by pre-incubating with the inhibitor. The enzyme-inhibitor complex is then diluted into a substrate solution such that the final concentration of the inhibitor is orders of magnitude lower than the inhibition constant K_i. Under these pre-steady state conditions, re-binding of the inhibitor to the enzyme is highly unlikely and therefore release of the inhibitor from the enzyme subunits can be exclusively monitored by measuring the onset of substrate turnover. If NpBHC release from the subunits is cooperative, the enzyme is expected to jump from a state of no activity to its highest activity state without any intermediate activity levels. On the other hand, if the subunits release NpBHC independent from each other, intermediate activity levels are expected before the enzyme reaches its maximum activity (Scheme 2.1).

We conducted pre-steady state experiments where β-galactosidase is pre-incubated with NpBHC for 15 minutes and then diluted 1000 fold into substrate solution. The results are shown in Figure 2.1. Inhibitor dissociation is indicated by the onset of an increase in fluorescence as shown in Figure 2.1a. A histogram of the off-times (see Section 6.1 for a description of how the off-times were determined) before inhibitor release shows an exponential decay (Figure 2.1b), with a dissociation rate of $2.8 \times 10^{-3} \pm 0.8 \times 10^{-3}$ s^{-1}, which is consistent with the dissociation rate computed from bulk experiments (2.6×10^{-3} s^{-1}, 2.6.1). Notably, the dissociation rate of NpBHC is also similar to the dissociation rate of D-galactal[21].



Scheme 2.1. Schematic showing model for sequential inhibitor release from a tetrameric enzyme[21]. In this model, the inhibitors are released from the tetrameric enzyme in a sequential manner, and the enzyme can exist in five different states as shown, depending on the number of active sites blocked by the inhibitor. Each of the five forms have distinct activities as depicted by the product (P) vs time (t) profiles under the respective enzyme state. For example, the enzyme state that has all the active sites free will have higher activity than the enzyme that has only three free active sites and one blocked active site and so on. To go from the state of no activity (i.e. all active sites blocked) to the state of highest activity (i.e. all active sites free), the enzyme has to go through all the intermediate states.

To investigate whether the inhibitor is released in a cooperative or sequential manner, the fluorescence intensity trajectories in Figure 2.1a are converted to substrate turnover rates by taking the derivative of the trajectories and multiplying by a calibration constant followed by correction for photobleaching[21]. Turnover rates of a few representative trajectories from the pre-steady state experiment are shown in Figure 2.1c. The turnover rate increases in a stepwise manner for many enzyme trajectories. That is, multiple levels of activity are observed after inhibitor release before maximum activity is reached for many enzyme trajectories. Since the final concentration of NpBHC is orders of magnitude lower than the inhibition constant, it is highly unlikely that NpBHC rebinds β-galactosidase during the observation time. The different levels of activity therefore may be attributed to sequential inhibitor release from the enzyme. This behavior is unlike that of D-galactal where the enzymes were observed to go from a state of no activity to the highest activity level in a single step[21].

In Figure 2.1c, an inhibitor release event, characterized by a sudden change in turnover rate, can be clearly distinguished from the background variation in the turnover rate. The turnover rate between two subsequent inhibitor release events is stable and is defined as an 'activity level'. Since β-galactosidase is a tetramer, each enzyme molecule is expected to display 5 different activity levels (including the inactive state) corresponding to the stoichiometry of inhibitor binding, before it reaches maximum activity (Scheme 2.1). While several enzymes show different levels of activity, demonstrating sequential inhibitor release, the number of activity levels observed is not the same for all enzymes as can be seen in Figure 2.1c. The number of activity levels for each trajectory was determined by using a method based on the student t-test[29] (Section 6.1). The number of activity levels observed ranged from 2-4, unlike the expected five levels of activity for each enzyme based on scheme 2.1.

The average turnover for each activity level observed in Figure 2.1c was computed (Section 6.1). A histogram of these turnovers is shown in Figure 2.1d along with the best fit obtained for a normal distribution. The turnover distribution does not fit well to a normal distribution. Multiple peaks can be identified in the histogram in Figure 2.1d and may be generated from multiple populations of the enzyme, rather than a single population of enzyme. These multiple populations of the enzyme could be a result of differences in the number of active sites blocked by the inhibitor. For example, the population of enzyme with all active sites blocked has zero activity; the population of enzyme with three of the active sites blocked has the lowest level of activity and so on. If the activity of each enzyme state in scheme 2.1 is well quantized, then the histogram in Figure 2.1d is expected to contain five peaks. However the five peaks are not clearly

distinguishable in the histogram. The reason could be that each of the enzyme-inhibitor complexes shown in Scheme 2.1 may exhibit static heterogeneity resulting in a broad distribution of activity for each complex. The activities of different complexes can potentially overlap resulting in the histogram shown in Figure 2.1d, from which it is not possible to quantize the activity of the five different states.

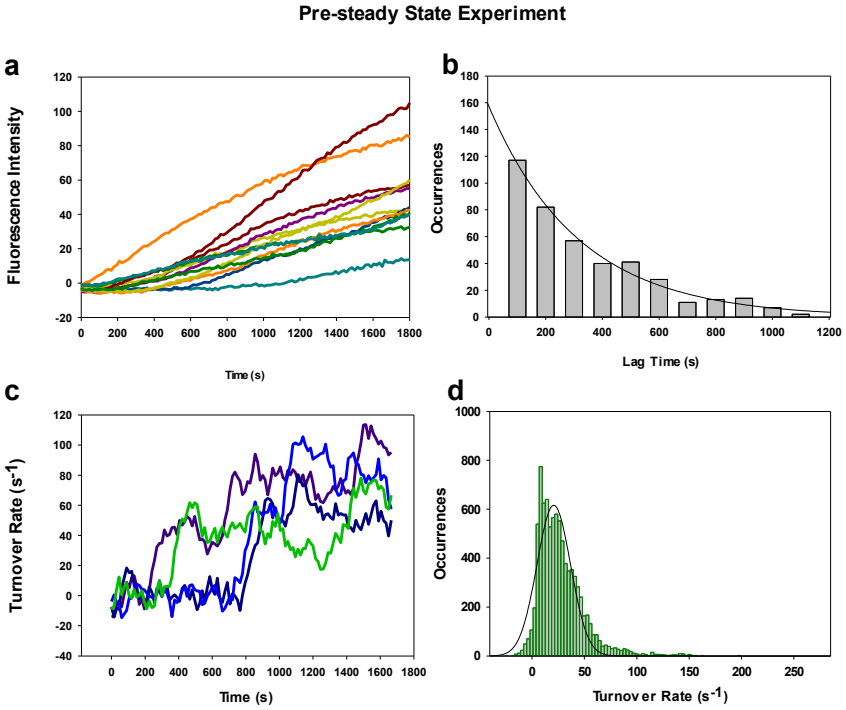


Figure 2.1. (a) Pre-steady state inhibition experiment with the inhibitor NpBHC performed by incubating 3.6 nM of β-galactosidase with 10 nM inhibitor followed by a 1000-fold dilution into the substrate RGD (100 μM) and sealing in optical fiber wells. The figure shows background-corrected and smoothed (moving average filter with span = 3) fluorescence time traces obtained by imaging the fiber. Initially all the active sites of the tetrameric enzyme molecule are bound by the inhibitor resulting in no build up of fluorescence. After the inhibitor is released from the enzyme, the fluorescence intensity increases due to the enzyme regaining activity. Different lag times can be observed in the trajectories for different molecules due to in the stochastic nature of the dissociation times[21]. In the absence of inhibitor no such lag time is observed[21] (b) Typical histogram of the lag times from a pre-steady

state experiment with NpBHC used to determine the k_{off}. (c) Turnover rates calculated for a few trajectories in (a). The slopes at every point in the fluorescence intensity trajectories were computed to give the change in fluorescence intensity per unit time. The slopes were then converted into numbers of product molecules per unit time using a calibration curve obtained by sealing various concentrations of resorufin into optical fiber wells and recording the raw fluorescence intensity value. (d). Histogram of average turnover rates calculated for each activity level distinguishable from traces in (c).

The conclusion for these results indicates that D-galactal and NpBHC differ in the kinetics of inhibitor release from β-galactosidase subunits. While D-galactal release seems to be strictly cooperative, NpBHC seems to be released sequentially from the four enzyme subunits. Sequential release of inhibitor was reported by Craig *et al.*[26] in the case of D-galactal for a minor fraction of the total molecules observed. In our studies, we were only able to observe cooperative release for D-galactal both in this study and in Reference 21. In contrast, for NpBHC, we observe sequential release for approximately 30% of the enzyme molecules that show an initial lag in activity. To determine the number of enzymes that display a lag in activity, the turnover rate profiles of the enzymes were first analyzed to identify inhibitor release events as described in Section 6.1. If the enzyme exhibits more than one inhibitor release event, then the enzyme is determined to release the inhibitor in a sequential manner.

2.2.2 Steady state experiments

The variations in enzyme turnover rates due to inhibitor binding and release can also be monitored under steady state conditions ($[I] = K_i$). Beginning with an un-inhibited free enzyme, we can observe fluctuations in enzyme activity. Steady state experiments of NpBHC and D-galactal were performed at an inhibitor concentration equal to the

inhibition constants (K_i = 0.5 nM for NpBHC and K_i = 20 μM for D-galactal). The steady state trajectories are shown in Figure 2.2a. During the initial stages of the trajectories, the enzymes are active resulting in an increase in fluorescence. A pause in fluorescence build up indicates the duration during which the inhibitor is bound to the enzyme. Figure 2.2b shows the turnover rate variations during a steady state experiment with NpBHC and D-galactal respectively.

Significant fluctuations in the enzyme turnover can be observed in the presence of NpBHC as well as D-galactal. The enzyme fluctuates between distinct activity states, indicating an inhibitor binding or release event. A histogram of the turnover rates of the enzyme in the presence of the inhibitors is shown in Figure 2.2c (only the durations in which the enzyme was active were considered). While the activity of the enzyme is lower in the presence of both inhibitors, in comparison to NpBHC, the activities are lower for D-galactal as is evident from the distributions in Figure 2.2c. Even though NpBHC is a more potent inhibitor of β-galactosidase than D-galactal, the average enzyme activity is lower in the presence of D-galactal than NpBHC when both inhibitors are present at a concentration equal to their K_i. D-galactal is a reversible inhibitor where an inhibition cycle (binding and release of inhibitor) is accompanied by conversion of D-galactal to 2-deoxygalactose[24]. Since the enzymatic reaction of β-galactosidase with D-galactal proceeds through the transition state to the production of the product, inhibition occurs with an activation barrier that is much higher than the barrier for substrate hydrolysis. Crossing this activation barrier can possibly result in significant alteration of the enzyme conformation[25], resulting in lower activities. NpBHC, on the other hand, is a classical reversible inhibitor, where release of inhibitor from the enzyme does not accompany the

conversion of inhibitor into a product. As a result, the enzyme may not undergo major conformational changes due to inhibition, leading to relatively unaltered and higher activity compared to D-galactal.

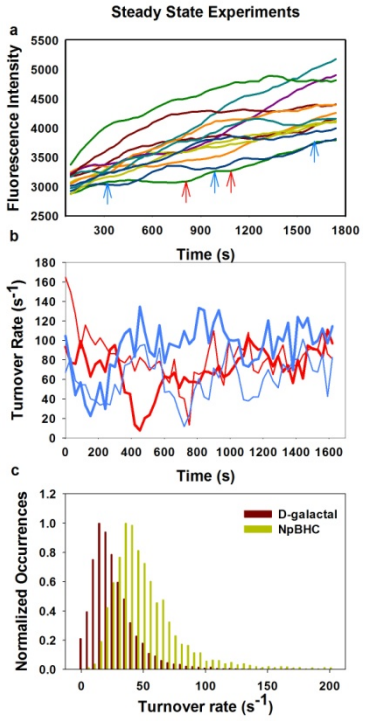


Figure 2.2. (a) Representative trajectories of raw fluorescence intensity in the active wells under steady state reaction conditions. NpBHC binding and release events are indicated for one representative trajectory (green). The blue arrows indicate when inhibitor binds to the active site and resulting in lack of substrate turnover and a stable fluorescent intensity, while the red arrows indicate a later time point when the inhibitor is released from the active site resulting in substrate turnover and an increase in fluorescence intensity. (b) An example of substrate turnover rates calculated from fluorescence intensity trajectories for D-galactal (blue traces) and NpBHC (red traces). Distinct levels of activity within the duration of the experiment can be identified indicating different stochiometries of binding of the inhibitor to the subunits of the tetrameric β-galactosidase. (c) Histogram of the average turnover rates under steady state reaction conditions. To compute the average turnover rate

of an enzyme, only the times of the trajectory during which the enzyme is active are taken into account.

The fluctuations in single molecule turnover rates are stochastic in nature and indicate the enzyme switches between several active states and an inactive state. An analysis of the on-times (duration in which enzyme is active) and off-times (duration in which the enzyme is inactive) was performed (Section 6.2). The histograms for D-galactal can be fitted to an exponential function to obtain an association rate constant of 10^2 $M^{-1}s^{-1}$ (R^2 = 0.98). The dissociation rates calculated from the exponential fits are $(7\pm0.5)x10^{-3}$ s^{-1} for D-galactal and $(1\pm0.09)x10^{-2}$ s^{-1} for NpBHC (R^2 = 0.98 and 0.97 respectively) and compare well with the pre-steady state experiments.

Inhibitor induced changes in activity can be easily identified in the steady state experiments. In a steady state experiment, for a sequential binding and release model the enzyme switches between any two adjacent states depicted in scheme 2.1 upon an inhibitor binding or release event. On the other hand for a cooperative binding and release model, the enzyme switches between the first and last states in scheme 2.1. To identify any relationship between the activity of an enzyme before and after inhibitor binding, scatter plots of the pre and post activities are plotted. The first binding or unbinding event was considered for each trajectory in the steady state experiment. To identify the event of inhibitor binding or release, we looked for the first time point at which the enzyme switches between two activity levels. T(i) is the average turnover of the enzyme before inhibitor binding and T(i+1) is the average turnover of the enzyme after inhibitor binding and before a second binding event. A cross-correlation of T(i+1) with T(i) is shown in Figures 2.3a and b for D-galactal and NpBHC, respectively. In the case of D-galactal, the turnovers are only weakly correlated (spearman correlation coefficient = 0.20, $p<0.05$).

Therefore the new activity state induced by D-galactal is relatively random. However, in the case of NpBHC, a significant positive correlation can be seen (spearman correlation coefficient = 0.674, $p<0.05$). Therefore, an enzyme with low activity before NpBHC binding is more likely to have low activity after NpBHC release and vice versa. In other words, the enzyme seems to have a memory of its activity before inhibitor binding. This result supports the hypothesis that D-galactal binds to β-galactosidase in a cooperative manner and alters its conformation significantly compared to NpBHC, which binds in a sequential manner.

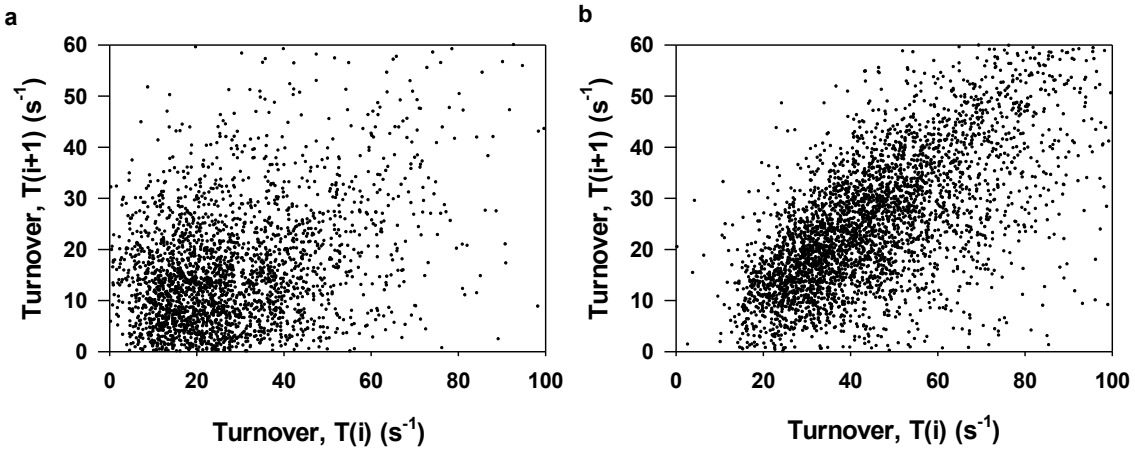


Figure 2.3. (a) and (b) Cross-correlation plots of substrate turnover rates of the adjacent activity levels of individual β-galactosidase molecules. T(i) is the activity before inhibitor induced change in activity and T(i+1) is the activity after the inhibitor induced change. The scatter plot for D-galactal shows that there is no relation between the activities (a), while that for NpBHC shows a positive correlation (b).

2.2.3 Autocorrelation analysis

Static heterogeneity is a well established feature in the activity of β-galactosidase[5]. Considering that static heterogeneity originates either from translational errors, post-

translational modifications, oxidative damage or different conformational states of the enzyme, it is reasonable to anticipate that the extent of interaction of individual enzymes with the inhibitor will also have a broad distribution. To quantify the inhibitor interaction with the enzyme an autocorrelation analysis can be performed as described in Reference 21 to extract the exchange rates of the inhibitor with the enzyme (Section 6.3). Exchange rate of the inhibitor is the sum of the on and off rates and is therefore a better indicator of the rate of the number of inhibition cycles per unit time rather than counting the number of binding and release events for each trajectory.

To study the static heterogeneity of inhibitor interaction with the enzyme an autocorrelation analysis is performed on the steady state turnover rate trajectories (Section 6.3). The autocorrelation curves for both D-galactal and NpBHC fit well to a mono-exponential decay (Figure 6.4a). For a single step binding and release, the autocorrelation is expected to have a mono-exponential decay. However, for a multi-step binding and release scheme, the autocorrelation of the turnover rates is expected to be a multi-exponential decay and should fit to a stretched exponential decay equation $C_m(t) = C_m(0)\exp[-(k_c \times t)^{\beta}]$. From the sequential release kinetics evident in pre-steady state experiments, the autocorrelation curves for activity traces of steady state NpBHC experiments were expected to fit to a stretched exponential decay equation. When plotted, the majority of them could be fit better to a mono exponential decay curve ($\beta=1$ in the stretched exponential decay equation). It is possible that we do not have the necessary dynamic range of lag times to be able to observe a stretched exponential decay profile. To obtain the average exchange rate, the autocorrelation curves of thousands of enzymes (4999 for D-galactal and 2993 for NpBHC) were averaged and fit to a mono-exponential

decay (Figure 6.4b). The exchange rates were computed to be $6.6x10^{-3}$ s^{-1} for D-galactal and $4.7x10^{-3}$ s^{-1} for NpBHC. The inhibitor exchange rates of individual β-galactosidase molecules show a broad distribution (Figure 6.4c) as expected.

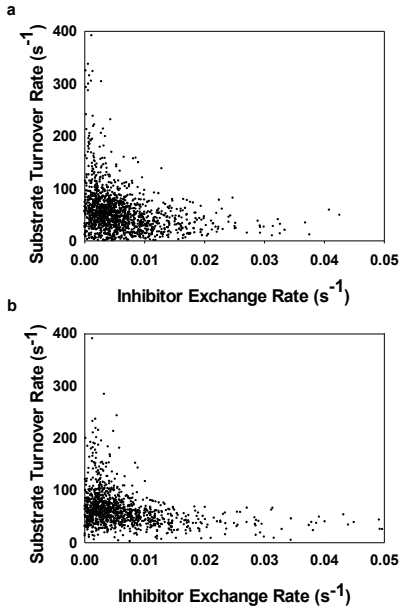


Figure 2.4. **Cross correlation plot between the average substrate turnover rate and the inhibitor exchange rate computed for each β-galactosidase molecule under steady state conditions with two tight binding inhibitors (a) D-galactal (n = 1476) and (b) N-p-bromobenzylamino-hydroxymethyl-cyclopentanetriol (n = 1109).**

Thus static heterogeneity can be observed for the inhibitor exchange rate as well as the substrate turnover rate. By measuring both parameters in the same experiment, the relative variation of these parameters across an enzyme population can be measured. The average substrate turnover rate computed for each enzyme molecule in the presence of the inhibitors was plotted against the inhibitor exchange rate calculated for the same molecule in a steady state experiment (Figure 2.4a and b). A negative correlation can be observed for both inhibitors – D-galactal ($p<10^{-25}$, Pearson's product moment correlation = -0.34) and NpBHC (negative correlation, $p<10^{-24}$, pearson's product moment correlation = -0.31).Therefore, enzyme molecules that have a high substrate turnover rate have a low

inhibitor exchange rate and enzyme molecules that have high exchange rate have low turnover rate.

NpBHC is thought to mimic the structure of transition state intermediate in the hydrolysis of galactopyranosides by β-galactosidase. For the transition state mimic, the inverse correlation between turnover rate and exchange rate can be explained as follows: Enzymes that have high inhibitor exchange rates stabilize the transition state excessively resulting in a lower turnover rate with the substrate. Conversely enzymes that have higher turnover rates stabilize the transition state to a lesser extent and therefore have a lower inhibitor exchange rate. D-galactal is a non classical inhibitor. It forms a glycosyl enzyme intermediate which is slowly hydrolysed to release 2-deoxygalactose. Thus the inhibitor exchange rate of D-galactal is a measure of the rate of hydration of D-galactal. In this case the cross correlation plot in Figure 2.5a compares two different activities of the enzyme – hydration of D-galactal versus the hydrolysis of the substrate resorufin β-D-galactopyranoside. The enzyme employs a different mechanism for D-galactal hydration[24] when compared to the substrate hydrolysis. As a result, the individual enzyme molecules that are more efficient at catalyzing the substrate hydrolysis are less efficient at catalyzing D-galactal hydration and vice versa.

Therefore, for both NpBHC and Dgalactal, the enzyme active site is selective towards the inhibitor or the substrate. In the case of D-galactal, selectivity results from the heterogeneity in catalytic efficiencies of the enzyme molecules. In the case of NpBHC the selectivity arises from the heterogeneity in the transition state stabilization. The heterogeneities in the enzyme active site may be due to conformational differences or due

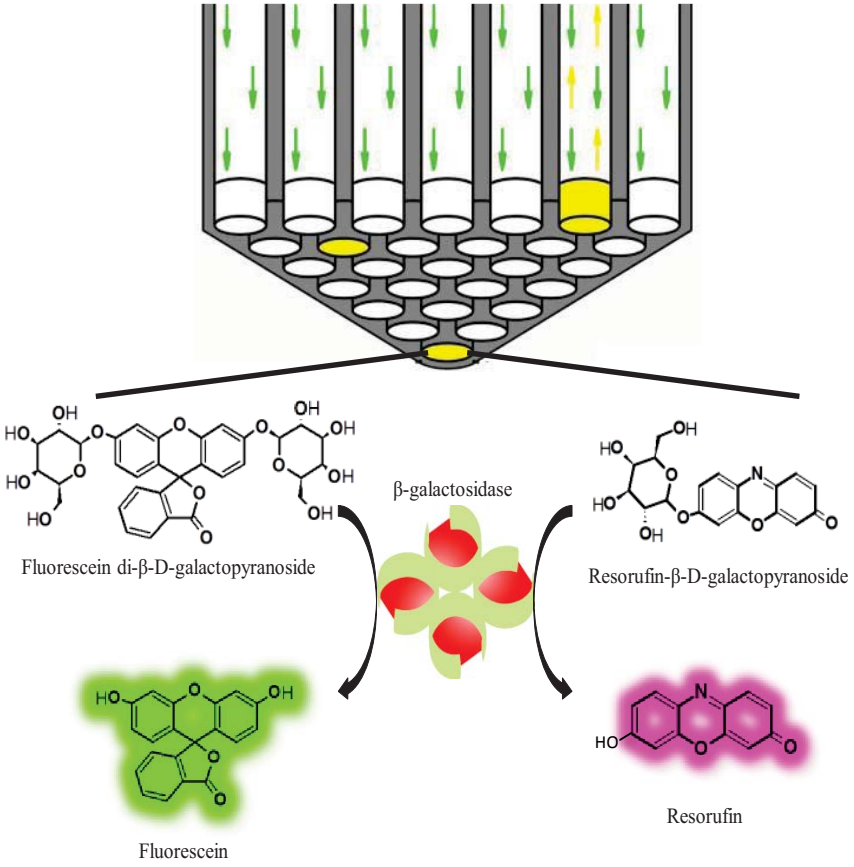
to differences in the amino acids of the active site arising from transcriptional or translational errors or a combination of these factors.

2.2.4 Multiple substrates

To ensure that the results from the competitive inhibition studies were not experimental artefacts, we performed a similar experiment and studied the competition for the active site between substrate and a non-effector species. To this end, we compared the relative variation of activities of the single β-galactosidase molecules in the presence of two fluorogenic substrates.

The substrates also compete for the active site but are not expected to affect each other's turnover rate[30]. Single molecules of β-galactosidase were trapped in optical fiber wells in the presence of an excess of two different substrates—FDG and RDG— (Figure 2.5a) and the competition between the substrates for the four active sites of β-galactosidase was studied by monitoring their relative turnover rates. Controls for determining signal leakage between the filter cubes were performed by omitting one of the substrates but imaging in both the cubes. No signal leakage between the filter cubes could be observed. Figure 2.5b shows a cross correlation plot between the net fluorescence generated from the two substrates, demonstrating a perfectly positive correlation between the two substrates. The positive cross correlation shown in Figure 2.5b indicates that the enzyme activity state affects the turnover rates of both substrates similarly. Therefore, a molecule that has high activity for one substrate also has a high activity for the other substrate, providing strong evidence that the results we obtained with substrate and inhibitors are valid. Additionally this plot demonstrates that in the competition between the substrates the active site is not selective for either substrate. On

the contrary, the competition between substrate and inhibitor is dictated by the active site, i.e. if the active site is more selective for the inhibitor then it is less selective for the substrate and vice versa.



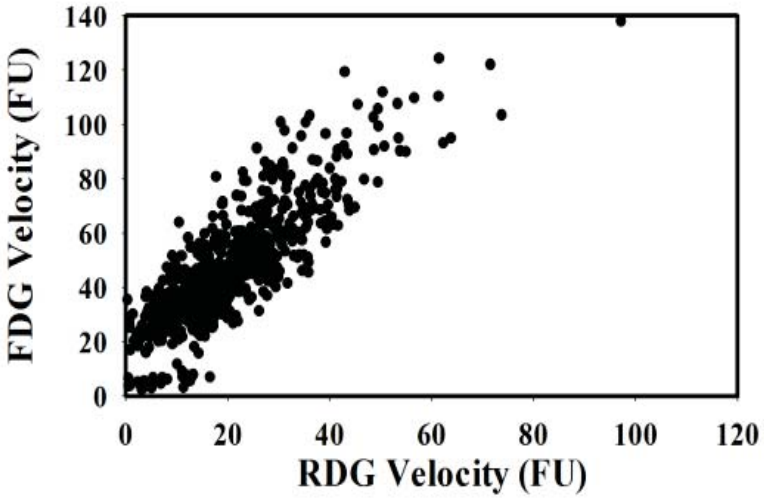


Figure 2.5. (a) Scheme for studying the relative kinetics of single β-galactosidase molecules in the presence of the two substrates FDG and RDG. Each well contains either 1 or 0 enzyme molecules and 10 μM of both FDG and RDG. The product generation was monitored by recording an image by alternating between the fluorescein and resorufin channels. The velocity for each reaction was computed by measuring the average change in fluorescence intensity per second recorded in the respective channel. (b) A cross correlation plot between the reaction velocities to compare the relative activities of individual enzyme molecules with respect to each substrate.

2.3 Conclusions

A single molecule technique was used to trap a large population of individual molecules of E. coli β-galactosidase and to study their activities in the presence of tight binding inhibitors. We studied and compared the stochastic binding and release of two different inhibitors of β-galactosidase —D-galactal and NpBHC. The inhibitor release kinetics of the two inhibitors was shown to be different. We were able to compare the effect of inhibitor binding on the enzyme conformation for the two inhibitors, which is impossible to differentiate in a bulk experiment. We demonstrated that D-galactal, a less potent inhibitor has a more drastic effect on the enzyme conformation than NpBHC, a more potent inhibitor. Furthermore, by using autocorrelation analysis on the observed turnover rates, the interaction of individual enzymes with the inhibitors and the substrate could be quantified. We demonstrated that the substrate turnover rate and inhibitor exchange rate are inversely correlated for both D-galactal and NpBHC.

While inhibitor binding and substrate turnover are mutually exclusive processes in competitive inhibition, we show that they are not independent of one another. The exchange rate of the inhibitor and the turnover rate of the substrate, for example, may be

linked to each other via the enzyme conformation. The results of both competitive inhibition and a multiple substrate reaction also demonstrate the difference between substrate-inhibitor competition in which the active site is more selective for the substrate or the inhibitor and substrate-substrate competition in which the active site is not selective for either substrate.

2.4 Supporting information

2.4.1 Bulk Studies to characterize inhibitors and obtain rate constants

Bulk experiments to determine the inhibition constant (K_i) and on and off-rates (k_{on} and k_{off}) were performed on a Tecan microtiter plate reader. First, progress curves were obtained at different inhibitor concentrations by adding the enzyme to a mixture of substrate and inhibitor and recording the fluorescence intensity over time. These progress curves were fit to equation 2.1 to obtain K_{obs}[1] (Figure 2.6), the pseudo first order rate constant

$$[P] = \frac{v_0 - v_s}{k_{obs}}(1 - e^{-k_{obs}t}) + v_s t \qquad \text{(Eqn. 2.1)}$$

where v_0 and v_s are the initial and the steady state velocity respectively and [P] is the product concentration. The pseudo first order rate constant $K_{obs,}$ is defined as:

$$k_{obs} = [I]k_{on} + k_{off} \qquad \text{(Eqn. 2.2)}$$

where, k_{on} and k_{off} are the first order association rate constant and the dissociation rate constant, respectively. The k_{obs} was computed at different inhibitor concentrations using (1) and plotted against the inhibitor concentration. This plot was fit to equation 2.2 to obtain k_{on} and $\underline{k_{off}}$ (Figure 2.7). The k_{on} and k_{off} for inhibition of β-galactosidase by NpBHC were determined to be 7×10^5 $M^{-1}s^{-1}$ and 2.6×10^{-3} s^{-1}, respectively.

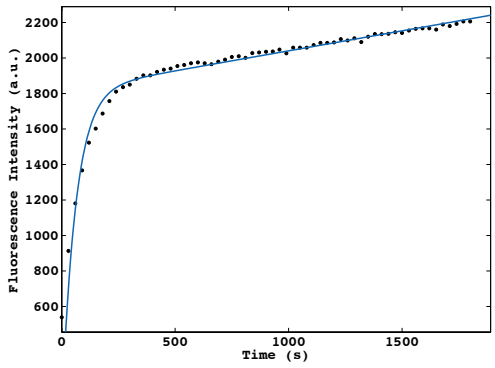


Figure 2.6—Progress curve showing the slow binding inhibition of β-galactosidase by NpBHC. 360 pM of final concentration of the enzyme is added to a substrate-inhibitor mixture containing 100 µM of substrate and 20 nM of the inhibitor and the fluorescence intensity is recorded every 30 s for 30 min. The progress curve is fit to equation 2.1 to obtain the pseudo first order rate constant, $k_{obs.}$ For 20 nM inhibitor, the k_{obs} is 0.0162 s^{-1}.

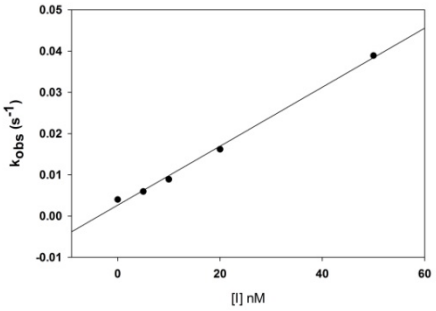


Figure 2.7—The plot shows pseudo first order rate constant k_{obs} obtained for different inhibitor concentrations. The plot can be fit to a straight line (eqn. 2.2) to determine k_{on} and k_{off}.

2. 5 Experimental

2.5.1 Materials

β-galactosidase from E. coli (grade VIII) was purchased from Sigma–Aldrich and dissolved and diluted in 1X PBS/1 mM MgCl2 buffer (pH=7.4). The enzyme was purified using HPLC and characterized on a native gel. Stock solutions of 100 mM D-galactal (1,5-anhydro-2-deoxy-d-lyxo-hex-1-enitol) (Sigma–Aldrich) in PBS/$MgCl_2$, 50 mM N-p-bromobenzylamino-hydroxymethyl cyclopentanetriol (NpBHC) in PBS/$MgCl_2$, 100 mM resorufin-β-D-galactopyranoside (RDG) (Invitrogen, Carlsbad, CA) in DMSO, 100 mM resorufin sodium salt (Invitrogen) in DMSO, 10 mM fluorescein di-β-D-galactopyranoside (Invitrogen) in DMSO and 100 mM fluorescein (Invitrogen) in DMSO were aliquotted and stored at −20°C. Further dilutions of all the substrates and inhibitors were made in reaction buffer containing 1X PBS, 1 mM $MgCl_2$, 0.005% w/v Tween-20 and 0.005% w/v BSA, pH =7.4. Optical fiber bundles (Ø=2mm) containing 50,000 optical fibers (Ø=4.5μm) were purchased from Schott (Southbridge, MA).

2.5.2 Microchamber array fabrication

Array fabrication was performed as described by Rissin et al[21]. Fiber bundles that are 4.5 cm long were sequentially polished on a Techprep fiber polisher (Allied Hightech) using lapping films with a grid size of 30, 12, 9, 3, 1, and 0.3 μm (Allied Hightech products Inc.) One end of the fiber bundle was then etched for 115 s with 0.025 M HCl with stirring and immediately washed with DI water. The cladding and the cores of the fiber bundle are made of differently doped silicas resulting in selective etching of the core over the cladding. Etching for 115 s results in 46 fL size wells in the optical fibers, thereby affording an array of 50,000 femtoliter sized reaction vessels.

2.5.3 Single molecule isolation

To obtain representative single molecule information, hundreds of individual *E. coli* β-galactosidase molecules were isolated and enclosed with the substrate resorufin-β-D-galactopyranoside (RDG) in a large array of 46-fL containers etched into the surface of an optical fiber bundle[29]. Individual β-galactosidase molecules turned over non-fluorescent RDG to highly fluorescent resorufin (Figure 2.6). As the reaction components were confined in each 46 fL-container, the resorufin production of individual enzyme molecules could be monitored by fluorescence microscopy.

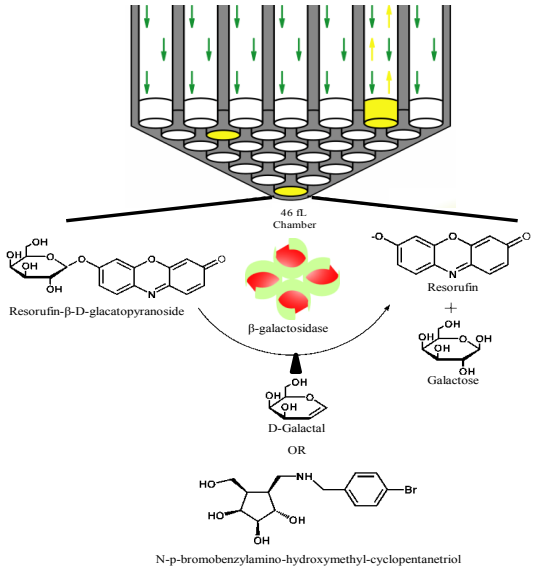


Figure 2.6 Single molecule setup to study the kinetics of β-galactosidase molecules in the presence of inhibitor under steady state conditions[21]. Femtoliter sized wells generated by acid etching of fiber bundles are filled with a solution containing 100 µM substrate, 20 µM of the inhibitor D-galactal and 3.6 pM β-galactosidase resulting in 1 molecule of β-galactosidase per 10 wells. The probability, *p*, of a well containing n molecules of the enzyme can be calculated using the Poisson equation $\boldsymbol{p(n) = e^{-x}x^{n}/n!}$;$p(n) = e^{-x}x^{n}/n!$

where x is the average number of enzyme molecules per well. For x = 0.1, we obtain $p(0) = 0.904; p(1) = 0.090; p(2) = 0.0045$, meaning that the probability of finding more than 1 enzyme in a well is negligible. Therefore, each well has either 1 or 0 molecules of β-galactosidase and an excess of substrate and inhibitor molecules. By sealing the fiber wells, the fluorescent product generated by the enzyme is contained within the wells. (b) Snapshot of the fiber bundle after 30 minutes of the steady state experiment and the zoomed in image of a portion of the fiber bundle. Each well contains either 1 or 0 molecules of β-galactosidase[5], 20 μM of the inhibitor D-galactal, and 100 μM of the substrate RDG.

2.5.4 Pre-steady state inhibition experiments

Pre-steady state experiments were used to investigate whether the inhibitor release from the tetrameric enzyme happens in a cooperative or sequential manner. 1 μL of a pre-incubated mixture of enzyme and inhibitor (3.6 nM of enzyme and 100 μM of D-galactal incubated for 15 minutes or 3.6 nM of enzyme and 10 nM of NpBHC incubated for 20 minutes) was diluted into 1 mL of 100 μM RDG solution to give a final enzyme concentration of 3.6 pM and inhibitor concentration at least an order of magnitude lower than K_i. After the final dilution, the reaction mixture was immediately sealed into optical fiber wells and time-lapse imaging was performed.

2.5.4 Steady state inhibition experiments

Steady state experiments were used to compare the activities of the individual enzyme molecules before and after inhibitor binding. For D-galactal experiments, 1 μL of 3.6 nM enzyme was diluted into the 1 mL solution containing 100μM of RDG and 20 μM of D-galactal. For NpBHC experiments, 1 μL of 3.6 nM enzyme was diluted into the 1 mL solution containing 100μM of RDG and 0.5 nM of NpBHC). These inhibitor concentrations were chosen because they are equal to the inhibition constant (K_i) of the inhibitor. After the final dilution, the reaction mixture was immediately sealed into the optical fiber wells and time-lapse imaging was recorded.

2.5.5 Multiple substrate reactions

Multiple substrate reactions were conducted by diluting 1 μL of 3.6 nM enzyme solution into 1 mL of a solution containing 10 μM each of both RDG and FDG. After the final dilution, the reaction mixture was immediately sealed into the optical fiber wells and time lapse imaging was recorded every 30 s for at least 20 min with an exposure time of 2 s for both fluorophores. The filter wheel was set to alternate between the resorufin and fluorescein cubes for each measurement.

2.5.6 Imaging

Time lapse images were recorded for at least 30 minutes with an upright Olympus (Tokyo, Japan) BX61 microscope equipped with a short arc mercury lamp (Ushio, Tokyo, Japan) and a CCD camera (Sensicam QE; Cooke Optics, Romulus, MI). A filter set with λ_{ex} = 571 nm and λ_{em} = 584 nm (Chroma Technology, Rockingham, VT) was used for the resorufin product, and a filter set with λ_{ex} = 475 nm and λ_{em} = 540 nm was used for the product fluorescein (Chroma Technology, Rockingham, VT). Images were recorded every 15 s or 30 s using an exposure time of 2 s under reduced excitation light (ND = 1). The total reaction time lasted for at least 15 minutes.

2.8 References

1. Q. Xue and E. S. Yeung, *Nature*, 1995, 373, 681-683.
2. D. B. Craig, E. A. Arriaga, J. C. Y. Wong, H. Lu and N. J. Dovichi, *Journal of the American Chemical Society*, 1996, 118, 5245-5253.
3. H. P. Lu, L. Xun and X. S. Xie, *Science*, 1998, 282, 1877-1882.
4. T.-M. Hsin and E. S. Yeung, *Angewandte Chemie International Edition*, 2007, 46, 8032-8035.

5. D. M. Rissin, H. H. Gorris and D. R. Walt, *Journal of the American Chemical Society*, 2008, 130, 5349-5353.

6. J. Li and E. S. Yeung, *Analytical Chemistry*, 2008, 80, 8509-8513.

7. H. H. Gorris and D. R. Walt, *Journal of the American Chemical Society*, 2009, 131, 6277-6282.

8. A. M. van Oijen, P. C. Blainey, D. J. Crampton, C. C. Richardson, T. Ellenberger and X. S. Xie, *Science*, 2003, 301, 1235-1238.

9. K. Velonia, O. Flomenbom, D. Loos, S. Masuo, M. Cotlet, Y. Engelborghs, J. Hofkens, A. E. Rowan, J. Klafter, R. J. M. Nolte and F. C. de Schryver, *Angewandte Chemie International Edition*, 2005, 44, 560-564.

10. B. P. English, W. Min, A. M. van Oijen, K. T. Lee, G. Luo, H. Sun, B. J. Cherayil, S. C. Kou and X. S. Xie, *Nat Chem Biol*, 2006, 2, 87-94.

11. G. Luo, M. Wang, W. H. Konigsberg and X. S. Xie, *Proceedings of the National Academy of Sciences*, 2007, 104, 12610-12615.

12. Y. Wang and H. P. Lu, *The Journal of Physical Chemistry B*, 2010, 114, 6669-6674.

13. L. C. Tabares, D. Kostrz, A. Elmalk, A. Andreoni, C. Dennison, T. J. Aartsma and G. W. Canters, *Chemistry – A European Journal*, 2011, 17, 12015-12019.

14. N. S. Hatzakis, L. Wei, S. K. Jorgensen, A. H. Kunding, P. Y. Bolinger, N. Ehrlich, I. Makarov, M. Skjot, A. Svendsen, P. Hedegard and D. Stamou, *Journal of the American Chemical Society*, 2012, 134, 9296-9302.

15. Y. Choi, I. S. Moody, P. C. Sims, S. R. Hunt, B. L. Corso, I. Perez, G. A. Weiss and P. G. Collins, *Science*, 2012, 335, 319-324.

16. R. J. Davenport, G. J. L. Wuite, R. Landick and C. Bustamante, *Science*, 2000, 287, 2497-2500.

17. T. R. Strick, V. Croquette and D. Bensimon, *Nature*, 2000, 404, 901-904.

18. Z. Zhang, P. T. R. Rajagopalan, T. Selzer, S. J. Benkovic and G. G. Hammes, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, 101, 2764-2769.

19. Y. Rondelez, G. Tresset, T. Nakashima, Y. Kato-Yamada, H. Fujita, S. Takeuchi and H. Noji, *Nature*, 2005, 433, 773-777.

20. S. F. Tolić-Nørrelykke, M. B. Rasmussen, F. S. Pavone, K. Berg-Sørensen and L. B. Oddershede, *Biophysical Journal*, 2006, 90, 3694-3703.

21. H. H. Gorris, D. M. Rissin and D. R. Walt, *Proceedings of the National Academy of Sciences*, 2007, 104, 17680-17685.

22. D. N. Fuller, D. M. Raymer, V. I. Kottadiel, V. B. Rao and D. E. Smith, *Proceedings of the National Academy of Sciences*, 2007, 104, 16868-16873.

23. Y. He, Y. Li, S. Mukherjee, Y. Wu, H. Yan and H. P. Lu, *Journal of the American Chemical Society*, 2011, 133, 14389-14395.

24. D. F. Wentworth and R. Wolfenden, *Biochemistry*, 1974, 13, 4715-4720.

25. O. M. Viratelle and J. M. Yon, *Biochemistry*, 1980, 19, 4143-4149.

26. D. B. Craig, T. T. Morris and C. M. Q. Ong-Justiniano, *Analytical Chemistry*, 2012, 84, 4598-4602.

27. J. N. Greul, M. Kleban, B. Schneider, S. Picasso and V. Jäger, *ChemBioChem*, 2001, 2, 368-370.

28. J. F. Morrison, *Trends in Biochemical Sciences*, 1982, 7, 102-105.

29. N. J. Carter and R. A. Cross, *Nature*, 2005, 435, 308-312.

30. D. B. Craig, M. J. Eggertson, M. Chikamatsu and C. A. Horwood, *Analytical Letters*, 2011, 44, 1835-1841.

CHAPTER 3

Enzyme Fragment Complementation of β-Galactosidase

3.1 Introduction

In genetics, complementation is a phenomenon when loss in biological function due to mutation in a gene is restored in the presence of a complementing species that has a mutation in a different gene. Protein complementation is an extension of this phenomenon where a protein that has lost its function due to a genetic mutation can recover its function in the presence of another protein molecule that has a mutation in a different part of the protein[1]. Protein complementation can be intercistronic or intracistronic. In intercistronic complementation the complementing proteins have mutations in different genes, for example in different subunits of the protein. In intracistronic complementation the complementing proteins have mutations in different parts of the same gene. For example, a protein having a terminal mutation can complement with the protein that contains the missing sequence[1].

The α-complementation of β-galactosidase is an intracistronic complementation in which an inactive N-terminal deletion mutant of the enzyme (enzyme acceptor, EA or α-acceptor) and the inactive deleted peptide (enzyme donor, ED or α-donor) associate to generate an active enzymatic species[2-4]. This construct is commonly utilized in the blue-white screening technique for bacteria[5], CEDIA immunoassays[6], the detection of *in vitro* and *in vivo* protein-protein interactions[7], and in high throughput screening assays[8,9]. In this chapter, we examine the α-complementation reaction of β-galactosidase (referred to as EFC – enzyme fragment complementation) at the single molecule level to characterize the active species that are generated in the complementation reaction and compare their activity to the native β-galactosidase enzyme.

Native ß-galactosidase from *E. coli* is a homo-tetramer, each made up of 1023 amino acids[10]. Various N-terminal deletion mutants lacking residues in the α-region (23-31 or 11-41) have been shown to be inactive and gain activity in the presence of peptides containing the deleted amino acid sequence (3-41 or 3-92) to generate β-galactosidase activity[2]. Activity can be restored via complementation to the acceptors lacking as many as the first 85 N-terminal amino acids[11]. It was shown for α-complementation,

residues 3-41 are the most important for the donor, and therefore can be presumed to contribute to the activity of the enzyme.

The molecular basis of α-complementation has been discussed from a structural viewpoint[10,12,13]. β-galactosidase is a relatively large (465kda) enzyme made up of four identical monomers, each having five distinct structural domains (depicted as different colors) as shown in Figure 3.1a. Two subunit-binding interfaces can be seen in the structure, the long interface and the activating interface. Residues 3-41 which are an essential part of the α-donor are shown in orange. A simplified schematic is shown in Figure 3.1b, where the long interface and the activating interface between the four monomers are shown. The sequences corresponding to the α-donor are shown as distinct thicker and darker lines. As obvious from Figure 3.1a, the α-donor does not directly contribute to the active site of the enzyme (domain 2, yellow), which is shown as pockets in schematic in Figure 3.1b. However, the α-donor contributes to the activating interface. The activating interface is critical for catalytic activity as the contacts along this interface enable loops from adjacent monomers to extend into the active sites of the neighboring monomers, thereby completing the active site. This mutual completion of active site is depicted in Figure 3.1b. Thus, in the absence of the α-donors, inter-subunit contacts are only established along the long interface resulting in inactive dimers due to incomplete active sites. In the presence of α-donors, contacts along the activating interface are established resulting in active tetramers. The α-donors may provide dimer-dimer binding along the activating interface in two ways: residues 13-20 of the donors contribute directly to the interface contacts (Figure 3.1b), and/or residues 29-33 can provide for intra-subunit contacts by occupying a tunnel between domain 1 and rest of the monomer. This occupation of a tunnel between domain 1 and the monomer may potentially result in conformational changes within the monomer that promote dimer-dimer binding.

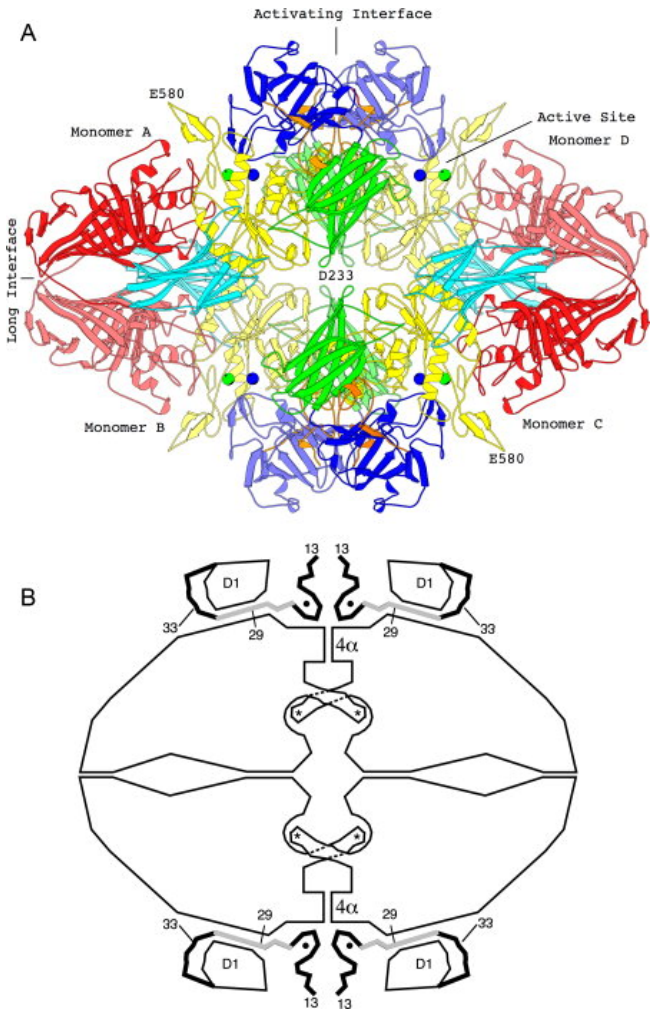


Figure 3.7 (a) Crystal structure of *E. coli* β-galactosidase reprinted from reference 12. Each monomer contains 5 distinct domains that are shown in different colors. Domains from an adjacent monomer are shown in a lighter or darker shade. Domains 1-5 are shown in blue, green, yellow, cyan and red respectively. Domain 3 is a TIM barrel containing the active site of the enzyme. The vertical axis of symmetry is called the activating interface and contains monomer-monomer contacts that make up the active site. Residues corresponding to the α-peptide are shown in orange and contribute to the activating interface. The horizontal axis of symmetry is called the long interface. The contacts along this interface remain unperturbed in the absence of the α-peptide. (b) Simplified sketch of β-galactosidase tetramer emphasizing the α-complementation features. Residues 13-50 which belong to the donor are shown as thicker lines. Residues 13-20 are shown in black and seem to contribute directly to the activating interface. Residues 29-33 are shown in grey and pass through a tunnel between domain 1 and the rest of the monomer. This sketch also shows

the role of the activating interface for β-galactosidase activity. The active sites are shown at the center of the sketch. A loop containing residues 272-288 from the adjoining monomer contributes to each active site. Therefore when the activation interface is disrupted, the loop from the adjoining monomer can no longer complete the active site resulting in all the four active sites being inactive.

Reports are contradictory as to whether and to what extent oligomerization is necessary for β-galactosidase to display activity. Some studies, as discussed above, suggest that β-galactosidase is active only as a tetramer, proposing that the α-acceptor exists primarily as dimers and is consequently inactive, and the deleted fragment therefore restores activity by forming β-galactosidase tetramers[14,15]. In other studies, active dimers of β-galactosidase have been identified and have been shown to be part of the complementation reaction suggesting that the α-donor can replenish the activity without tetramerization[16]. The reports on the stoichiometry of the complementation reaction are also contradictory. While studies based on chromatography have determined an α-donor:monomer ratio of 1:1[17], fluorescence fluctuation analysis of the complementation reaction using a fluorescently tagged α-donor has indicated the presence of only one donor per active species.[18]

In the current work we apply the high-density optical fiber bundle platform described previously[19,20] to isolate and monitor the activity of individual active species formed in the α-complementation reaction of β-galactosidase. The complementation reaction between EA and commercially available α-donors (which are modified versions of ED) from DiscoveRx was studied at the single molecule level. Specifically, two α-donors named Prolabel and Prolink were used in this study. Prolabel is a truncated version of ED and has been used as a protein fusion tag for over 200 proteins to study protein expression, translocation and secretion from cells[21]. Prolink is a low affinity version of Prolabel, which has been primarily used to study protein-protein interactions where Prolink and EA are each fused to the interacting proteins[21]. Due to the low affinity between Prolink and EA, the presence of complementation indicates the presence of a protein-protein interaction. Making use of the unique capability of our optical fiber system to observe multiple reactions with single molecule resolution, we

were able to directly count the number of active species generated in the complementation reaction and deduce the ratio of EA to ED in the complementation species. In addition, we were able to directly compare the activity of the complementation species to that of native tetrameric β-galactosidase.

3.2 Results and Discussion

To study the complementation reaction, the active species was first formed in bulk solution by incubating EA and ED (Prolabel or Prolink) for 1 hour. Single molecules of the active species were then isolated in the optical fiber wells along with excess substrate.

3.2.1 Single Molecule experiments

Single molecule studies were performed using the high-density array of femtoliter-sized wells generated on the optical fiber bundle platform[19,20,22]. The fabrication of the femtoliter-sized well array is described elsewhere[19]. Isolation of single molecules of the active species was achieved by sealing picomolar concentrations of the active enzyme species in the presence of 100 μM RDG in the 46 fL optical fiber wells (Figure 3.2). Using Poisson distribution statistics $P_\mu(x) = e^{-\mu} . \mu^x/x!$ where $P_\mu(x)$ gives the probability of finding *x* molecules per well and μ is the mean number of molecules per well, it can be determined that for a concentration of 3.6 pM, there is almost zero probability of finding more than one active species per well. Hence, every well contains either one or zero active molecules. The turnover of resorufin β-D-galactopyranoside to resorufin is used to detect the presence of the active species, as shown in Figure 3.2.

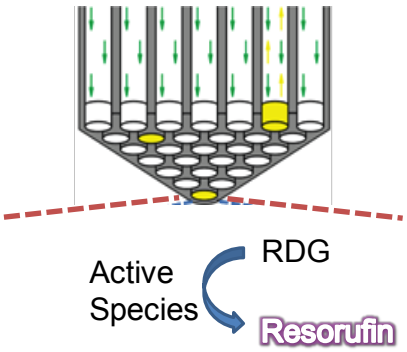


Figure 3.2 Isolation of single molecules of active species. The acceptor and donor molecules were pre-incubated for at least 30 minutes to allow for complementation. The incubated mixture was then diluted to 3.6 pM of the active species (assuming the reaction went to completion and EA and ED associate in a 1:1 ratio) and was sealed into 46 fL optical fiber wells.

In order to isolate single molecules of the product generated by the complementation reaction, we must ensure that the final concentration of the product is ≤ 3.6 pM. First, the active species of the complementation reaction were generated by incubating 10 µL each of 200 µM of EA and ED for 30 minutes. The maximum concentration of active species that can be formed is 100 µM (in this case EA:ED=1:1). Assuming that 100 µM of active species were formed, serial dilutions were performed in the reaction buffer such that the final concentration of the active species was 100 pM. 3.6 µL of this solution was then diluted in 96.4 µL of 100 µM substrate solution to give a final concentration of 3.6 pM of active species. If the association between EA and ED did not go to completion or the association is not 1:1, then the final concentration of each of the molecules would be less than 3.6 pM. This diluted solution of complementation species (concentration ≤ 3.6 pM) was used in the single molecule experiments described.

3.2.2 Isolation of single molecules of the active species

The results for single molecule complementation experiments are shown in Figure 3.3. A schematic of the single molecule experiment is shown in Figure 3.3a. As described above the active

species are first formed by incubating the acceptor and donor solutions followed by dilution to single molecule level. Figure 3.3b shows a snapshot of the signal generated in the optical fiber wells after 30 minutes. Figure 3.3c shows the fluorescence intensity profiles of the active wells during the time course of the reaction. Fluorescence intensity steadily increased due to the activity of single molecules in the fiber wells during the observation time without any abrupt pauses in activity. This suggests that the active species is stable and does not dissociate during the time course of observation. In the event that the complemented enzyme dissociates, the probability of re-association of a single donor with four acceptor molecules during the time of observation is very low (assuming apparent reaction association rate of ~200 $M^{-1}s^{-1}$).[18]The formation of the active species is therefore an irreversible reaction, at least during the time course of the experiment.

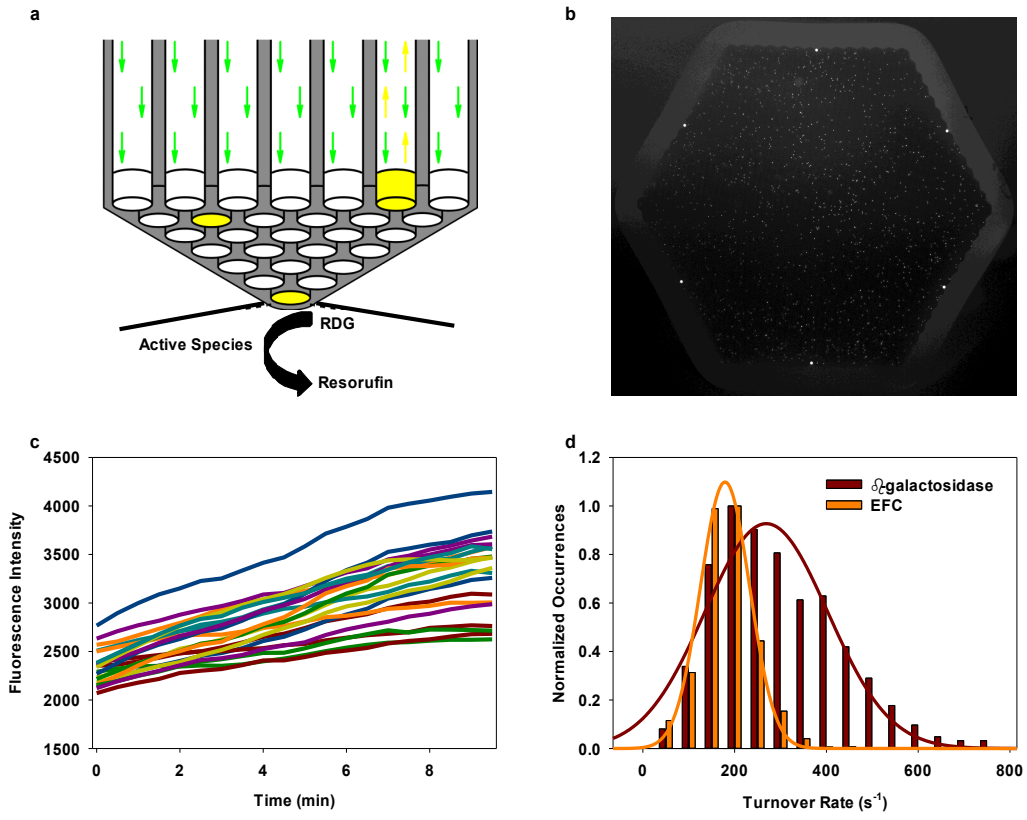


Figure 3.3 (a) Schematic for isolation of single molecules of the active complementation species in optical fiber wells. After sealing single molecules of active species in the presence of excess substrate within the optical fiber wells, time lapse imaging was performed for 30

minutes from the non-etched end of the fiber. (b) Snapshot of the optical fiber bundle consisting of 50,000 fibers at the end of the time lapse imaging. The wells that contain the active species fluoresce due to the turnover of the fluorogenic substrate into the product resorufin. (c) Fluorescence time traces of some of the active wells in (b). The imaging was performed with 30 seconds between each image capture. (d) Turnover rate histograms of the active species generated from the complementation reaction (Prolabel) and a typical turnover rate histogram of single molecules of β-galactosidase. The histograms are fit to Gaussian distributions $y = a * \exp\left(-\frac{(x-b)^2}{2c^2}\right)$ ($a = 0.9$, $b = 269$ and $c = 134$ for β-galactosidase; ($a = 1.1$, $b = 179$ and $c = 53.7$ for EFC). To obtain the turnover rates, the fluorescence intensities in (c) were first converted into numbers of resorufin molecules using a calibration factor of 120 molecules of resorufin/1 unit of fluorescence intensity. The slope at each point of the trajectory gives the instantaneous turnover rate, which was then corrected for photobleaching and averaged over the reaction time to obtain the average turnover rate.

The substrate turnover rates of the active species were calculated from the fluorescence intensity trajectories (see Data Analysis). Histograms of the turnover rates of the active species generated in the complementation reactions for both Prolink and Prolabel were plotted (Figure 3.3d) along with a typical turnover rate histogram of β-galactosidase. As can be seen in the Figure, the activity distribution is much narrower than the native tetrameric enzyme, and is skewed toward the lower activities. The activity of the native enzyme is clearly not recovered completely in the complementation reaction. The average activity of the β-galactosidase computed from the histogram in Figure 3.3d is 268 ± 9 s^{-1} which is about 1.5 times the average activity of EFC active species (172 ± 6 s^{-1}.for Prolabel and 144 ± 18 s^{-1} for Prolink). The time constant of the acceptor and donor association was determined to be 10.6 ± 2.8 min using the single molecule platform (SI 3.6.2) and matches up well with the time constant reported in the literature[18] (8.8 ± 1.7 min).

3.2.3 Determination of Stoichiometry

To obtain the kinetics of the irreversible formation of the active species, the single molecule experiments were performed with varying concentrations of the donor, keeping the acceptor concentration constant at 100 μM (Figure 3.4a) and vice versa (Figure 3.4b). For each experiment, first a 10,000-fold dilution of the complementation mixture was made in the reaction buffer followed by 3.6:96.4 dilution into substrate solution. The number of active species generated was counted by counting the number of active wells (data analysis), and plotting them as a percentage of the total wells in the fiber bundle. The activity distribution was plotted whenever single molecule activity was observed and was always identical to that in Figure 3.3d. The average activity of the single molecule species generated in all these experiments is 200 s^{-1}, showing very little variability.

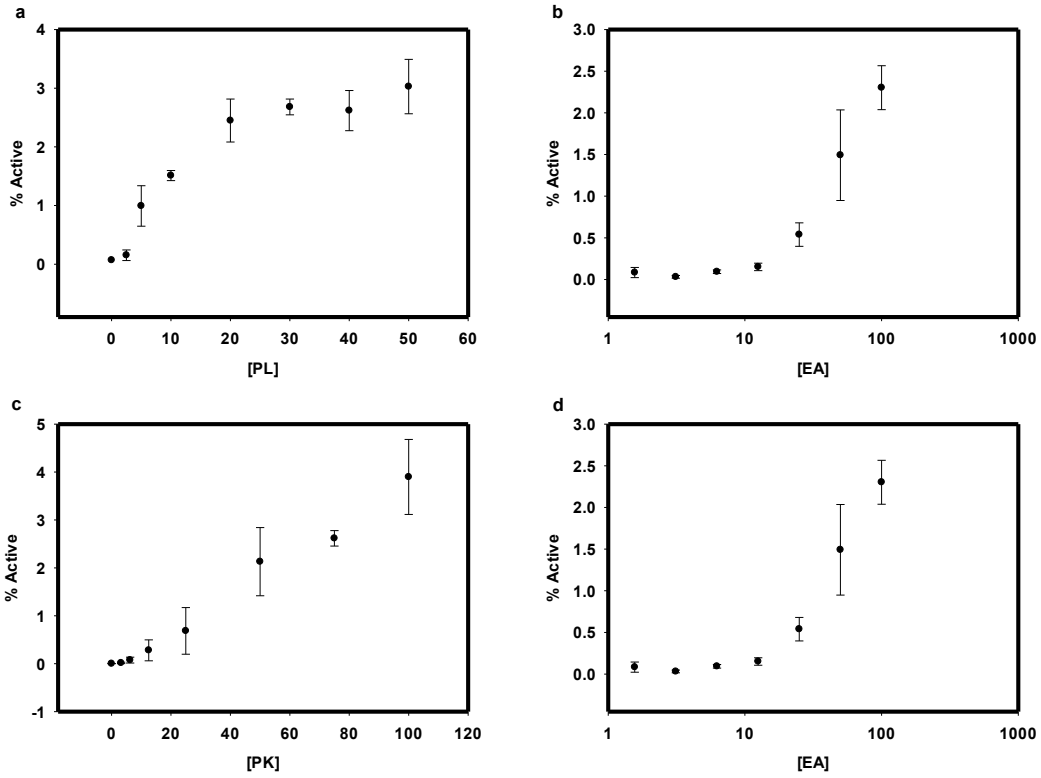


Figure 3.4 a) **Varying concentrations of Prolabel (PL) were incubated with 100 μM of EA for 30 minutes followed by a 10,000 fold dilution. The diluted solution was then mixed with substrate in the ratio of 3.6:96.4, sealed in the optical fiber bundle wells, and imaged. The**

numbers of active wells observed are plotted as a percentage of total wells (% active). (b) The experiment is repeated for varying concentrations of EA. The % active wells observed when the concentration of EA is varied while the concentration of PL is held constant at 100 µM. (c) Plot of % actives obtained for experiments with varying concentrations of Prolink (PK) with EA held constant at 100 µM. (d) Plot of % actives obtained for varying concentrations of EA with PK held constant at 100 µM.

The data plotted in Figure 3.4 are also summarized in Tables 3.1 and 3.2 for Prolabel and Tables 3.3 and 3.4 for Prolink (Experimental). Table 3.1 shows the percentage of active wells expected assuming different ratios of EA:ED in the complementation species. For example, at 100 µM concentration of EA and 50 µM concentration of ED, at 1:1 ratio of EA:ED, the expected concentration of active species is 50 µM. A 10,000 fold dilution results in an active species concentration of 50 pM, followed by 3.6:96.4 dilution into substrate resulting in a final concentration of 1.8 pM. At this concentration, 5% of the wells are expected to show activity. Similar calculations were made for the rest of the concentrations at different ratios of EA:ED. From Tables 3.1 and 3.2, it can be deduced that the ratio of EA:ED of 4:1 best matches the experimental data from both experiments shown in Figure 3.4a and b for Prolabel. To be more accurate, the number of ED molecules per complementation species was determined for each experiment by dividing the expected % active in the case of one ED molecule with the experimentally determined percent active. A weighted average of the number of ED molecules calculated in this manner (the square inverse of standard deviations were used as weights) results in 0.8 ± 0.2 molecules per complex. Similar analysis for Table 2 results in 4.4 ± 0.6 molecules of EA per complex. Therefore, the ratio of EA to ED is 4.4 ± 0.6:0.8 ± 0.2. The ratio of 4:1 best matches the experimental results as well as the structure of β-galactosidase known from the literature. Similarly, for Prolink the EA:ED ratio was calculated to be 4.6 ± 0.9: 2.8 ± 0.1. The ratio of 4:3 best matches the experimental data.

[ED] (μM)→	50	40	30	20	10	5	2.5	0
Experimental	**3.02 ± 0.46**	**2.62 ± 0.34**	**2.68 ± 0.13**	**2.45 ± 0.37**	**1.51 ± 0.09**	**0.99 ± 0.35**	**0.15 ± 0.09**	**0.07**
EA:ED								
for 1:1	5	4	3	2	1	0.5	0.25	0
for 1:2	2.5	2	1.5	1	0.5	0.25	0.125	0
for 2:1	5	4	3	2	1	0.5	0.25	0
for 2:2	2.5	2	1.5	1	0.5	0.25	0.125	0
for 4:2	2.5	2	1.5	1	0.5	0.25	0.125	0
for 4:1	**2.5**	**2.5**	**2.5**	**2**	**1**	**0.5**	**0.25**	0

Table 3.2 The expected % active wells for the experiment in Figure 3.4a for different ratios of EA:ED. The ratio of 4:1 best matches the experimentally determined % actives.

[EA] (μM)→	100	50	25	12.5	6.25	3.13	1.56	0
Experimental	**2.30 ±0.26**	**1.49 ±0.54**	**0.54 ±0.14**	**0.15 ±0.04**	**0.09 ±0.02**	**0.03 ±0.02**	**0.08 ±0.06**	**0.03**
EA:ED								
for 1:1	10	5	2.5	1.25	0.6	0.31	0.15	0
for 1:2	5	5	2.5	1.25	0.6	0.31	0.15	0
for 2:1	5	2.5	1.25	0.625	0.3	0.15	0.07	0
for 2:2	5	2.5	1.25	0.625	0.3	0.15	0.07	0
for 4:4	**2.5**	**1.25**	**0.625**	**0.3**	**0.15**	**0.078**	**0.037**	**0**
for 4:2	**2.5**	**1.25**	**0.625**	**0.3**	**0.15**	**0.078**	**0.037**	**0**
for 4:1	**2.5**	**1.25**	**0.625**	**0.3**	**0.15**	**0.078**	**0.037**	**0**

Table 3.2. The expected % active wells for the experiment in Figure 3.4b for different ratios of EA:ED. The ratios of 4:4, 4:2 and 4:1 best matches the experimentally determined % actives. EA:ED ratio of 4:1 therefore matches the experimental data from both Figure 3.4a and 3.4b.

[ED] (μM)→	100	75	50	25	12.5	6.25	3.2	0
Experimental	**3.9 ±0.8**	**2.6 ±0.2**	**2.1 ±0.7**	**0.7 ±0.5**	**0.3 ±0.2**	**0.08 ±0.06**	**0.02 ±0.01**	**0.002**
EA:ED								
for 1:2	5	3.75	2.5	1.25	0.625	0.3	0.15	0
for 2:1	5	5	5	2.5	1.25	0.625	0.32	0
for 2:2	5	3.75	2.5	1.25	0.625	0.3	0.15	0
for 4:2	2.5	2.5	2.5	1.25	0.62	0.3	0.125	0
for 4:1	2.5	2.5	2.5	2	1	0.5	0.25	0
for 4:3	**2.5**	**2.5**	**1.6**	**0.8**	**0.4**	**0.2**	**0.1**	**0**

Table 3.3. The expected % active wells for the experiment with Prolink in Figure 3.4c for different ratios of EA:ED. The ratio of 4:3 best matches the experimentally determined % actives.

[EA] (μM)→	100	50	25	12.5	6.25	3.13	1.56	0
Experimental	**2.3 ±0.3**	**1.5 ±0.5**	**0.5 ±0.1**	**0.15 ±0.04**	**0.09 ±0.02**	**0.03 ±0.02**	**0.08 ±0.06**	**0.03 ±0.04**
EA:ED								
for 1:2	5	5	2.5	1.25	0.6	0.31	0.15	0
for 2:1	5	2.5	1.25	0.625	0.3	0.15	0.07	0
for 2:2	5	2.5	1.25	0.625	0.3	0.15	0.07	0
for 4:4	**2.5**	**1.25**	**0.625**	**0.3**	**0.15**	**0.078**	**0.037**	**0**
for 4:2	**2.5**	**1.25**	**0.625**	**0.3**	**0.15**	**0.078**	**0.037**	**0**
for 4:3	**2.5**	**1.25**	**0.625**	**0.3**	**0.15**	**0.078**	**0.037**	**0**

Table 3.4 The expected % active wells for the experiment with Prolink in Figure 4d for different ratios of EA:ED. The ratios of 4:4, 4:2 and 4:3 best matches the experimentally determined % actives. EA:ED ratio of 4:3 therefore matches the experimental data from both Figure 3.4a and 3.4b.

Therefore, we conclude the complementation reaction results in a pentamer consisting of four acceptors and one donor molecule in case of Prolabel. In case of the This result confirms the EA:ED ratio determined using fluorescence fluctuation analysis[18]. In the case of Prolink, which is a low affinity version of Prolabel, the complementation reaction results in a hexamer consisting of four acceptors and two donor molecules.

3.3 Conclusions

The α-complementation reaction of β-galactosidase results in stable, irreversible active species that can be observed for long periods of time at the single molecule level. The complementation species is a pentamer consisting of four acceptors and one donor in the case of Prolabel. The acceptor dimers are themselves inactive because a loop from a different dimer is required to complete the active site. The α-donor mediates the dimer-dimer binding which enables the dimers to complete the active site of the adjoining dimer. The crystal structure of *E. coli* β-galactosidase predicts four donors per tetramer such that each donor anchors to one monomer[12,15]. The donors promote dimer-dimer binding by stabilizing a four helix bundle that forms a part of the dimer-dimer binding interface, also called the activation interface[15]. In addition, contact between two donor molecules directly contributes to the activation interface[12]. From our results we can deduce that in case of Prolabel only one donor is sufficient to recover the activation interface. Therefore, direct contribution of the donors to the contacts that make up the activation interface is not critical[15]. The donor potentially stabilizes the conformation of the dimer that is required for dimer-dimer binding. Consequently, this new conformation readily associates with other dimers to form a stable tetramer. In the case of the low-affinity donor, Prolink, three donor molecules are required to form the active complementation species. Therefore only dimers that are bound to the donor Prolink associate with each other to form an active enzyme. Activity histograms show that only 66% of the activity of native β-galactosidase is restored in the complementation species. The lack of complete recovery of activity could be an effect of reconstitution of the tetramer with only one or three donors. Direct comparison of the activities of wild type β-galactosidase and the complementation species has only

been studied at the bulk level and suggests that the activities are similar[12,16]. This result contrasts with the results from single molecule experiments reported in this paper, thus demonstrating the significance of single molecule studies.

3.4 Supporting Information

3.4.1 Bulk experiments

Enzyme activity measurement experiments were performed on a Magellan microtiter plate reader to test the presence of activity in the complementation reaction before performing a single molecule reaction. 10 µL of both EA and ED were incubated for 1 hour and added to 100 µL of the substrate solution (100 µM of RDG). The fluorescent product generated was monitored by measuring the fluorescence at λ_{ex} = 558 nm and λ_{em} = 590 nm. Measurments were taken after 30 minutes of incubating with the substrate.

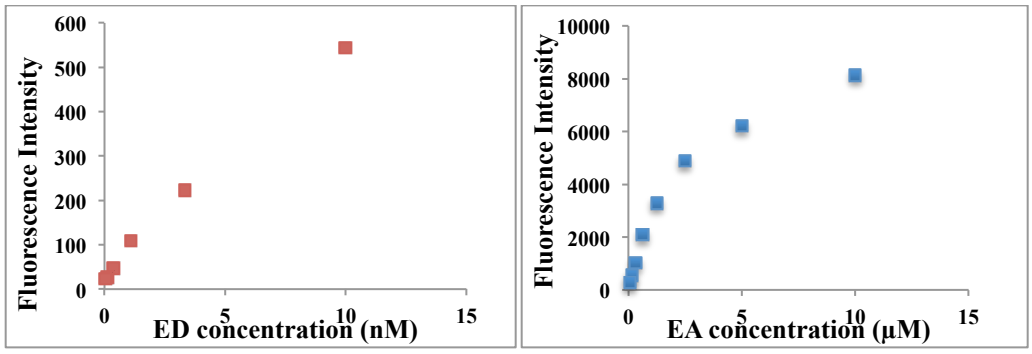


Figure 3.6 (a) 0.15 µM of EA was incubated with varying concentration of ED for 1 hour, followed by a 30 minute incubation with the substrate. The final fluorescence intensities measured are plotted against the ED concentration. The intensity increases with ED concentration as expected. (b) Varying concentrations of EA were incubated with 10 nM of ED for 1 hour followed by a 30 min incubation with substrate. The intensity increases with increasing concentration of EA. . At lower concentrations of EA (<2.5 µM) the fluorescence intensity increases linearly with EA concentration. At higher concentrations saturation is observed in the fluorescence intensity due to consumption of the substrate.

3.4.2 Single molecule experiments to determine the time constant of association

100 μM of both EA and ED were incubated for different time intervals (Figure 3.7) and immediately diluted to 3.6 pM into a solution containing 100 μM of substrate. The number of active species generated were then counted using the single molecule platform. Figure 3.7 was fit to an exponential function of time, $y = a*(1-e^{-t/\tau})$ to obtain the time constant of association (τ).

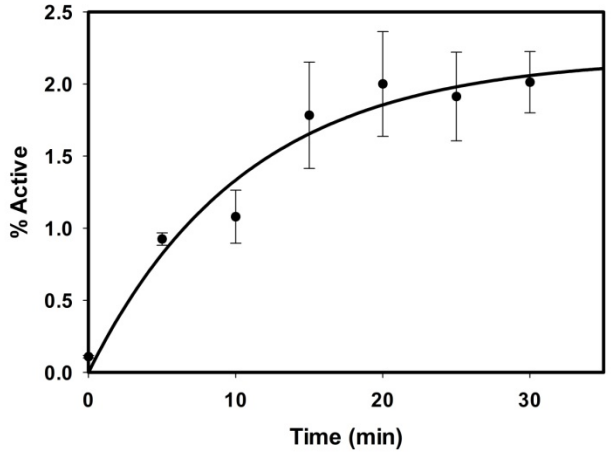


Figure 3.7 Number of active molecules generated in the EFC reaction as a function of time. The number of active molecules is determined by counting the number of wells that light up, and are plotted as a percentage of the total wells in the optical fiber bundle. The plot was fit to $y = a*(1-e^{-t/\tau})$; a = 2.2±0.2 and τ = 10.6±2.8 min.

3.5 Experimental

3.5.1 Materials

The Enzyme Fragment Complementation (EFC) detection assay kit containing the acceptor (EA) and donors Prolabel (ED) and Prolink (PK) were obtained from DiscoveRx Corporation. The sequence of Prolabel is MSSNSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEEARTDRPSQQLRSNLGE[21] and the sequence of Prolink is DSLAVVLQRRDWENPGVTQLNRLAARPPFASWRNSEEARTDR.[23] β-galactosidase from *E. coli* (grade VIII) was purchased from Sigma–Aldrich and purified on Zorbax-450 HPLC column. All protein dilutions were made in the reaction buffer 1x phosphate buffered saline (PBS; 0.14 M NaCl, 2.7 mM KCl, 8 mM NaH_2PO_4 and 2 mM K_2HPO_4, autoclaved and filtered) + 1 mM $MgCl_2$

(pH=7.4). 10x PBS was purchased from Ambion. Resorufin-β-D-galactopyranoside (RDG) and resorufin standard sodium salt were purchased from Invitrogen and stored at a stock concentration of 100mM in DMSO at -20^{o}C.

3.5.2 Bulk experiments

The complementation reaction was tested in bulk format before performing single molecule experiments to verify activity. The bulk experiments are described in the supporting information (SI 3.6.1).

3.5.3 Imaging

For imaging the single molecule experiments on the optical fiber, the optical fiber was fixed on the stage of an inverted microscope such that the non-etched end of the bundle was facing the objective. A 30 μL droplet of the substrate solution containing the diluted complementation species (concentration ≤ 3.6 pM) was placed on a well-cleaned PDMS gasket on a glass slide base which was positioned on an automated z-direction stage directly beneath the fiber bundle (Figure 3.5). The z-direction stage was then used to mechanically seal the reaction solution within the optical fiber wells. Time lapse imaging of the non-etched end of the optical fiber bundle was subsequently performed with a custom built imager (Quanterix, Lexington, MA) equipped with a 200 Watt metal arc lamp (Prior Scientific, Rockland, MA) and a CCD camera (Infinity4-11, Lumenera, Ontario, Canada). A filter set with λ_{ex} = 558 nm and λ_{em} = 577 nm (Semrock, Rochester, NY) was used for imaging the resorufin product.

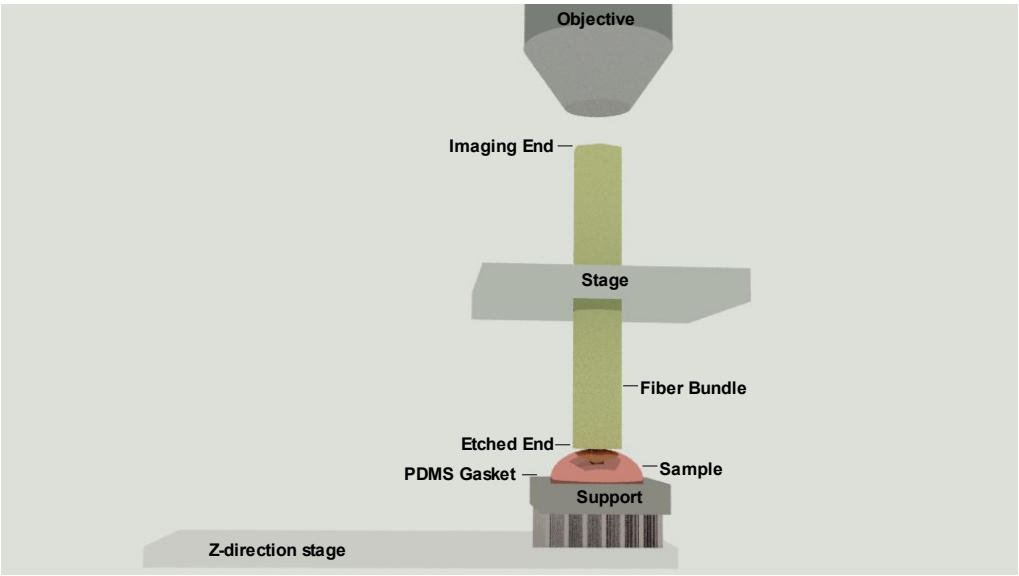


Figure 3.5 Schematic showing the imaging set-up for the single molecule experiments.

3.5.4 Data Analysis

Images were analyzed using MATLAB image processing toolbox. An algorithm was developed to identify the wells and extract the mean intensity of each well. The percentage increase in intensity between the first and last frames (30 minutes) was calculated for each well and a threshold of 20% was used to determine active wells. To convert the fluorescence intensities obtained from imaging to the number of resorufin molecules, a calibration curve was first obtained by measuring the fluorescence intensity generated by sealing different concentrations of resorufin standard solution in the fiber wells. The fluorescence intensities (F(t)) were then converted into number of resorufin molecules (P(t)) by multiplying by the calibration factor. The slope at each point of the trajectories was computed to obtain the instantaneous turnover rate (P'(t)). To correct for photobleaching, 10 µM of resorufin was sealed in optical fiber wells and time lapse imaging was performed. The fluorescence trajectories from the photobleaching experiment were fit to a mono-exponential decay curve to obtain the photobleaching rate $k_{ph.}$ The photobleaching corrected instantaneous turnover rates were computed using the equation: $P_{corr}(t)$

= P'(t) + K_{ph}xP(t). The instantaneous turnover rates of a given trajectory were then averaged to obtain the average turnover rate corresponding to that trajectory.

3.7 References:

(1) Zabin, I.; Villarejo, M. R. *Annual Review of Biochemistry* **1975**, *44*, 295.

(2) Ullmann, A.; Jacob, F.; Monod, J. *J Mol Biol* **1967**, *24*, 339.

(3) DeVries, J. K.; Zubay, G. *Journal of Bacteriology* **1969**, *97*, 1419.

(4) Langley, K. E.; Fowler, A. V.; Zabin, I. *Journal of Biological Chemistry* **1975**, *250*, 2587.

(5) Vieira, J.; Messing, J. *Gene* **1982**, *19*, 259.

(6) Henderson, D. R.; Friedman, S. B.; Harris, J. D.; Manning, W. B.; Zoccoli, M. A. *Clinical Chemistry* **1986**, *32*, 1637.

(7) Rossi, F.; Charlton, C. A.; Blau, H. M. *Proceedings of the National Academy of Sciences* **1997**, *94*, 8405.

(8) Naqvi, T.; Lim, A.; Rouhani, R.; Singh, R.; Eglen, R. M. *Journal of Biomolecular Screening* **2004**, *9*, 398.

(9) Eglen, R. M. *ASSAY and Drug Development Technologies* **2002**, *1*, 97.

(10) Juers, D. H.; Jacobson, R. H.; Wigley, D.; Zhang, X. J.; Huber, R. E.; Tronrud, D. E.; Matthews, B. W. *Protein science : a publication of the Protein Society* **2000**, *9*, 1685.

(11) Moosmann, P.; Rusconi, S. *Nucleic Acids Research* **1996**, *24*, 1171.

(12) Matthews, B. W. *Comptes Rendus Biologies* **2005**, *328*, 549.

(13) Juers, D. H.; Matthews, B. W.; Huber, R. E. *Protein Science* **2012**, *21*, 1792.

(14) Juers, D. H.; Heightman, T. D.; Vasella, A.; McCarter, J. D.; Mackenzie, L.; Withers, S. G.; Matthews, B. W. *Biochemistry* **2001**, *40*, 14781.

(15) Juers, D. H.; Jacobson, R. H.; Wigley, D.; Zhang, X.-J.; Huber, R. E.; Tronrud, D. E.; Matthews, B. W. *Protein Science* **2000**, *9*, 1685.

(16) Kaneshiro, C. M.; Enns, C. A.; Hahn, M. G.; Peterson, J. S.; Reithel, F. J. *Biochem. J.* **1975**, *151*, 433.

(17) Zabin, I. *Mol Cell Biochem* **1982**, *49*, 87.

(18) Meyer-Almes, F. J.; Wyzgol, K.; Powell, M. J. *Biophysical Chemistry* **1998**, *75*, 151.

(19) Rissin, D. M.; Gorris, H. H.; Walt, D. R. *Journal of the American Chemical Society* **2008**, *130*, 5349.

(20) Gorris, H. H.; Rissin, D. M.; Walt, D. R. *Proceedings of the National Academy of Sciences* **2007**, *104*, 17680.

(21) Olson, K. R.; Eglen, R. M. *ASSAY and Drug Development Technologies* **2007**, *5*, 137.

(22) Gorris, H. H.; Walt, D. R. *Journal of the American Chemical Society* **2009**, *131*, 6277.

(23) Wehrman, T. S.; Raab, W.; Loh, C. Y.; Google Patents: 2012.

CHAPTER 4

Thermoswitchability of β-Glucuronidase Variants Revealed Through Single Molecule Studies

4.1 Introduction

The effect of temperature on single enzyme molecules was first reported by Rotman, who studied the thermal denaturation of single molecules of β-galactosidase[1]. Single enzymes were found to either fully retain their activity or lose their activity when heated to the denaturation temperature. Thus, while at the ensemble level there is an average decrease in activity, at the single molecule level thermal denaturation seemed to be digital. The same result was observed later for single molecules of alkaline phosphatase, where a fraction of the single enzyme molecules became inactive at the denaturing temperature where as the surviving molecules retained full activity[2]. Thus single enzymes are irreversibly converted into an inactive form upon denaturation. When single enzymes were subjected to a heat pulse at a temperature lower than the denaturation temperature, their activity was shown to be different after the pulse[3]. The activity of single enzymes either increased or decreased, with no correlation to their activity before the pulse. The random change in activity was associated with a conformational change in the enzyme, where each conformation was hypothesized to have a distinct time-averaged activity. Thermal studies of single enzymes so far focused mainly on the denaturation of the enzymes. In this chapter we report the effect of temperature on β-glucuronidase when the temperature is increased from room temperature (RT) to 37°C, which is the optimal temperature of the enzyme. We show that the single molecules of the wild type (WT) are shown to increase in activity from 1-4 fold and that the activity at 37°C has no significant correlation with the activity at RT. Single molecules of β-glucuronidase variants on the other hand show a thermal switching behavior where enzymes that are completely inactive at RT gain activity at 37°C.

β-glucuronidase - an enzyme reported in a wide range of organisms - belongs to the glycosyl hydrolase family 2 and is known to hydrolyze glucuronides. Commercially known as

GUS reporter, β-glucuronidase has several applications including its use as a marker for bacterial water contamination and as a reporter for gene expression studies in plants. *E. coli* β-glucuronidase is a tetramer, where each monomer contains a sugar binding domain, an immunoglobulin like domain, and an α/β barrel that contains the active site[4]. The enzyme has four active sites each at the interfaces between monomers. Two glutamate residues within the active site act as the nucleophilic acid/base catalyst pair[5]. The β-glucuronidase variants used in this study were designed and expressed by the Ellington lab at the University of Texas, Austin. The variants were primarily redesigned on the surface of the protein so as to have only one solvent exposed cysteine in order to enable site-specific conjugation of oligonucleotides or labels to the enzyme.

Each monomer of β-glucuronidase has six potentially reactive surface cysteine residues. In order to be able to site specifically conjugate β-glucuronidase with DNA, the wild type enzyme was redesigned using Rosetta to generate single cysteine variants. The Rosetta protein modeling and structure prediction software was used to assess the impact of altering the cysteines with different amino acids. The overall structural stability of the proteins calculated for each design was compared to the wild type and the designs that exhibited the most stability were expressed. The four selected variants - D1, D2, D3 and D4 – have all the surface cysteines altered, except for either C133 or C262 (Figure 4.1a). Two additional serine variants – 133 and 262 – with all the cysteines altered to serine except for C133 and C262 were also synthesized as controls.

In this chapter, activities of single molecules of β-glucuronidase are compared at different temperatures. Single molecules studies show that β-glucuronidase variants display varying levels of thermal switching ability of β-glucuronidase variants. When the variants were tested for single

molecule activity at RT and 37°C, several variants demonstrated a significant increase in activity at 37°C at single molecule level. Single molecule experiments revealed reversible or irreversible thermal switching properties of a fraction of the population for these variants.

4.2 Results and discussion

4.2.1 Bulk experiments

Bulk experiments were performed for WT, Rosetta derived mutants D1-4 and serine variants (133, 262). While the Rosetta mutants show less activity than the WT, they outperform the serine variants (Figure 4.1b). The turnover rates and Michaelis constants of the variants are discussed in supplementary information (section 4.4.1). From Figure 4.1b, we can see that all the mutants display an initial lag in activity. Selected velocity-substrate curves are shown in Figure 4.8. While the WT displays the classic hyperbolic velocity-substrate curve (Figure 4.8a), all the Rosetta mutants display sigmoidal velocity-substrate curves (Figure 4.8b). Sigmoidal velocity-substrate traditionally indicates the presence of co-operativity in the catalytic mechanism. Accordingly, the velocity-substrate curves for the Rosetta derived mutants were fit to the Hill equation as described in Section 4.4.1. The serine variants also display a significant lag period in their progress curves. The velocity-substrate curves for the serine variants could not be fit to the Hill equation. The origin of lag phase for the Rosetta and the serine variants is not clear. We can hypothesize that the initial lag is a result of the enzyme undergoing a slow conformational change in the presence of the substrate. At a higher temperature (37°C), where the transition between conformations is expected to be faster, no lag was observed for the Rosetta mutants. In the case of the serine controls, the lag period is reduced at 37°C, though not completely overcome. This observation supports the slow conformational change hypothesis.

Variant	Mutations
D1	28K, 133D, 197D, 253V, 527M
D2	28L, 133D, 197D, 253V, 527M
D3	28L, 133D, 197D, 253I, 527M
D4	28K, 197D, 253I, 262Y, 527M
133	28S, 197S, 253S, 262S, 527S
262	28S, 133S, 197S, 253S, 527S

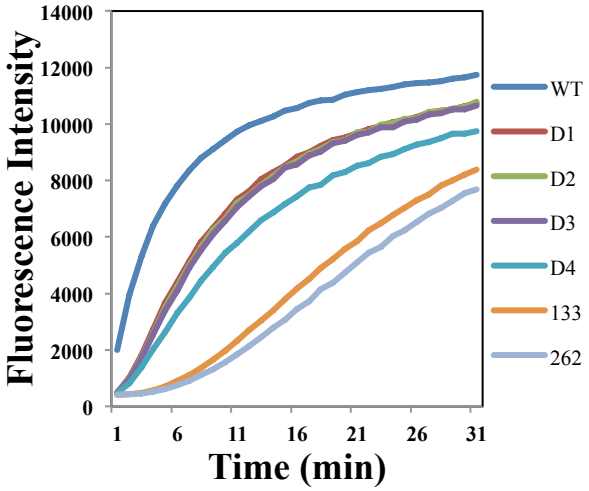


Figure 4.1 Rosetta design variants (D1-4) and serine controls (133 and 262) used in this study. The mutations at the solvent exposed cysteines are shown in the table. C133 is unaltered in D1, D2 and D3, while C262 is unaltered in D4. The graph shows the time course of activity of 1 nM of variants in the presence of 250 µM substrate.

4.2.2 Single molecule experiments

Single molecule experiments were performed for all variants at RT and at 37°C. The enzyme was diluted into substrate solution to a final concentration of 1 pM and sealed within the optical fiber wells. A substrate concentration of 250 µM was used for all the experiments, which is at least 2 times higher than Michaelis constant for most β-glucuronidase variants (Only K_m that could be calculated with standard Michaelis-Menten equation was considered). The only exception is D3 which displays a K_m of 387 µM at RT. However to avoid high background intensities substrate concentrations was limited to 250 µM in the single molecule experiments.

Figure 4.2 shows a comparison of average single molecule velocities to the velocities at 250 µM substrate concentration determined from the bulk experiments (Tabulated in Table S2). Triplicate experiments were performed for each variant. For the wild type and Rosetta mutants the reaction velocities are in agreement with the bulk velocities. Rosetta variant D4 is the only

exception. At 37 °C single molecule velocities of D4 are higher than the values calculated in bulk experiments.

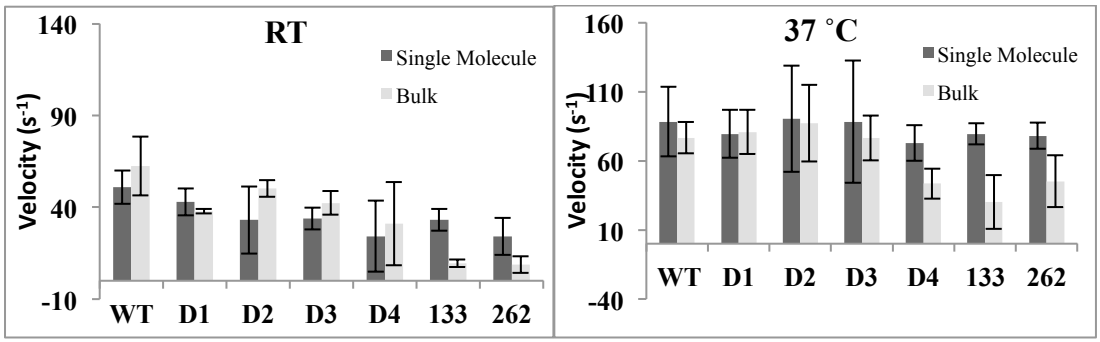


Figure 4.2 Comparison of single molecule and bulk activities at RT and 37°C. The activities are in agreement with each other except D4 and the serine variants 133 and 262.

The single molecule velocities calculated for the serine variants are significantly higher than the velocities determined from bulk experiments. For the serine variants, higher single molecule activities at RT is possibly due to only highly active enzymes observed at single molecule level. The RT bulk activities of these variants are less than 10 s^{-1}, which is close to the least observable activity in the optical fibers for this system (7.6 s^{-1} ; This number is calculated assuming the least observable resorufin concentration, 0.5 μM, generated over a period of 30 minutes). As a result, it is likely that only molecules higher than this activity are detected at single molecule level, leading to seemingly higher average single molecule velocities. The higher activity of D4 and serine variants at 37°C cannot be explained with the same reasoning as the bulk activities at 37°C are well above the single molecule activity detection limit. One possibility is that as the temperature is increased the tendency of a protein to aggregate is higher, which could potentially result in a loss of activity. However, in the single molecule experiments, the enzymes are isolated in different chambers eliminating the possibility of aggregation. As a result one can expect the single molecule activities to be higher than the bulk experiments.

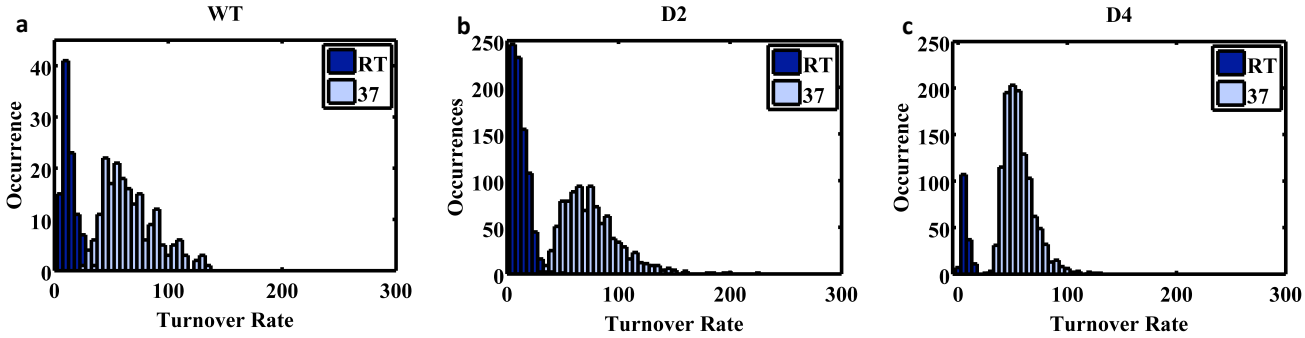


Figure 4.3 Single molecule activity histograms at RT and 37°C for WT, D2 and D4 (a, b and c respectively)

The reaction velocities of single molecules at RT and 37°C showed a distribution for all the variants. Histogram of single molecule activities at RT and 37°C are shown for WT, D2 and D4 in Figure 4.3. At 37°C all variants demonstrate a 2-3 fold increase in the average activity compared to room temperature (Table 4.1). To investigate the effect of temperature on single enzyme molecules, the fold increase in activity for individual enzymes was determined as described below. The increase in activity was found to be random and in some cases we found that a fraction of the population that was completely inactive at RT gained activity at 37°C.

	WT	D1	D2	D3	D4	133	262
Ratio	1.7	1.9	2.7	2.6	3.0	2.4	3.2

Table 4.5 Fold increase in average activity of enzymes at 37°C, when compared to the average activity at RT.

4.2.3 Temperature shift and pulse experiments

In order to compare single molecule activities at RT and 37°C, temperature shift experiments were performed. Temperature shift experiments were designed such that single molecule activities could be observed before and after a temperature change from RT to 37°C. A peltier plate was incorporated into the sealing platform[3] (see methods) to control the temperature.

Single molecules of GUS variants were trapped in optical fiber wells along with 250 μM of substrate. The activity of the enzyme was first monitored at RT for 5 minutes followed by an increase in temperature to 37°C as shown in Figure 4.4a.

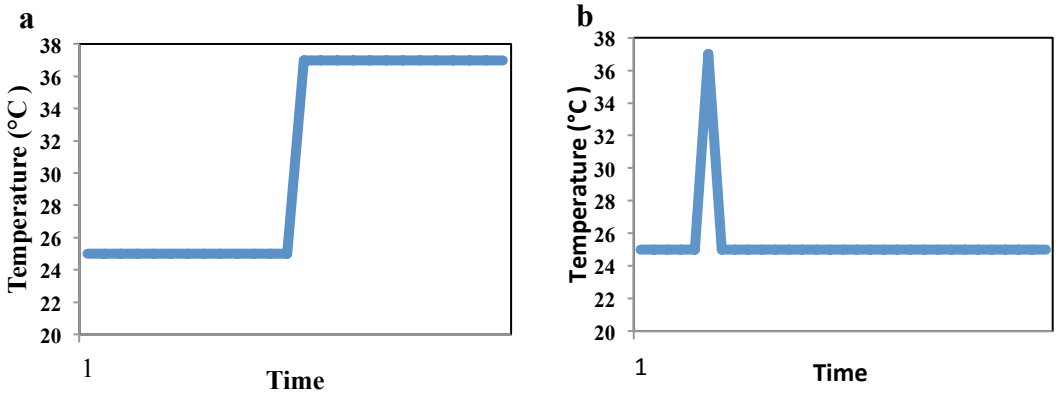


Figure 4.4 Temperature profiles for shift (a) and pulse (b) experiments.

When the activities of single molecules of WT are compared before and after the temperature shift we observe that the activity of nearly all the enzymes increased at 37°C. Figure 4.5a shows a histogram of fold activity increase for the single molecules. A scatter plot comparing the activities of individual molecule at RT and 37°C (Figure 5b) did not show any significant correlation between these activities (Spearman coefficient = 0.3, $p<0.005$). Therefore the increase in activity is random for individual molecules.

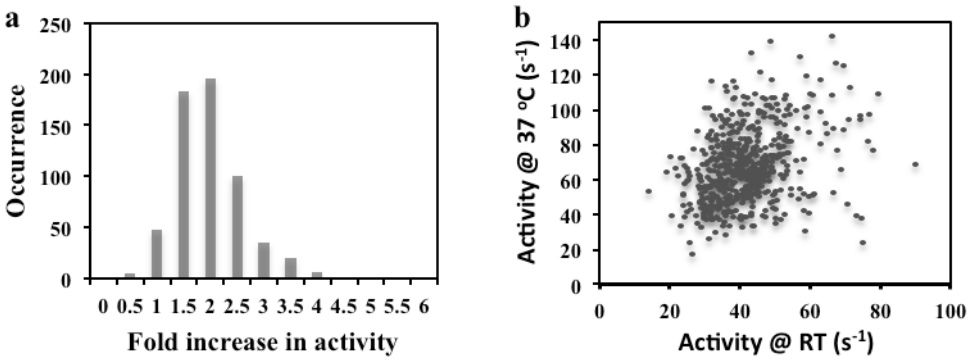


Figure 4.5 (a) Histogram of fold increase in activity for single enzyme molecules of the WT. (b) Scatter plot comparing the single molecule activities of the WT at RT and 37°C.

Shift experiments for D2 and D4 surprisingly showed a significant increase in single molecule activities at 37°C with average fold increase of 13.2 and 10.9 respectively. For both D2 and D4 we can identify several enzymes that are inactive or have undetectable activity at RT which gain activity at 37°C. For example, in the scatter plots in Figure 4.6, several enzymes that have no activity at RT (<0 on the x-axis) have high activities at 37**°C**. 18 % of the D2 enzymes gain activity at 37**°C** while 20 % of D4 enzymes gain activity at 37**°C**. Serine mutants showed a similar behavior where nearly 90% of the enzymes gained activity at 37**°C**. Therefore a fraction of the cysteine mutants exhibit thermal switching behavior where activity can be turned on at a higher temperature. It can be hypothesized that the single enzymes functioning as a thermal switch undergo a conformation change from an inactive to active conformation at 37**°C**. This observation of thermal switching behavior supports the slow conformation change hypothesis that results in an initial lag phase in bulk experiments as discussed in section 4.2.1.

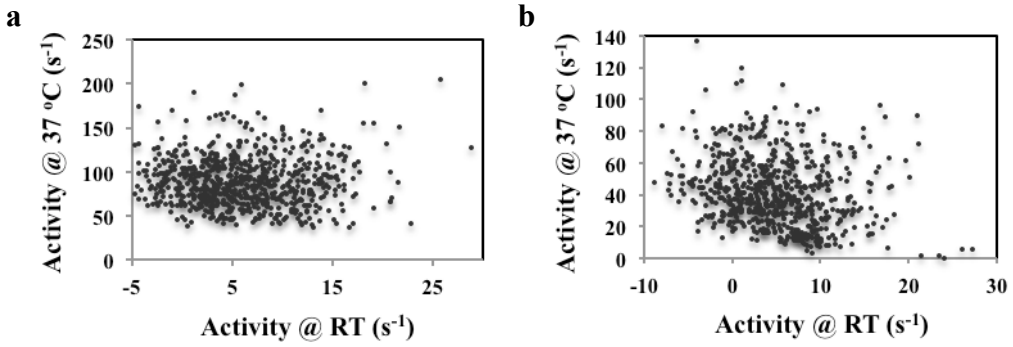


Figure 4.6 Scatter plot of activity of individual enzyme molecules at RT and 37°C for D2 and D4. Several enzymes that were inactive at RT gain activity at 37°C. No significant correlation could be observed between single molecule activities at RT and 37°C.

Thermal pulse experiments were performed to study the reversibility of thermal switching properties of the single enzyme molecules. Single enzyme molecules were first monitored at RT followed by increasing the temperature to 37°C. The temperature was held at

37°C for 2 s and reduced to RT. The activity of the single enzyme molecules before and after the thermal pulse was compared. For the Rosetta variants, several enzymes lost activity after the heat pulse resulting in reduced average activity. The serine mutants on the other hand gained activity due to the heat pulse and remained active after the heat pulse (Figure 4.7). Therefore, for the Rosetta mutants one can hypothesize that several single enzymes are converted into the inactive conformation when cooled back to RT. On the other hand the serine mutants retained the active conformation when cooled back to RT. Therefore the thermal switch behavior is irreversible in the serine mutants.

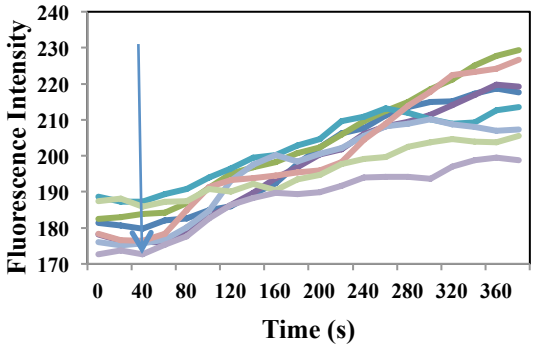


Figure 4.7 Time traces of single molecules of serine variant 133 in a thermal pulse experiment. The enzymes were monitored at RT for 50s followed by a thermal pulse (Figure 4.4). The enzymes were monitored at RT for the rest of the experiment. After the heat pulse, the enzymes retain increased activity at RT as well.

4. 3 Conclusions

In this chapter the effect of temperature on activity of single molecules of WT and single cysteine variants of β-glucuronidase was studied. Using the optical fiber bundle platform, we are able to compare the activity of single enzyme molecules at RT and 37°C. For several variants, a fraction of the population of single enzymes were shown to have thermal switching properties. That is, these enzymes had no activity at RT and gained activity at 37°C. For the serine variants

the thermal switching property was found to be irreversible. These enzymes retained the increased activity even after cooling back to RT. We hypothesize that the thermal switching property of the enzymes is a result of the single enzymes existing in an inactive or active conformation. When the temperature is increased to 37°C, enzymes that were in the inactive conformation switch to the active conformation resulting in a gain in activity.

4.4 Supporting Information

4.4.1 Bulk kinetics of GUS variants

The activity of the variants were measured at 0, 25, 50, 100, 150, 200 and 250 μM of the substrate and the initial velocities were used to obtain the velocity Vs. substrate concentration curve. For the WT, the velocity Vs. substrate curves could be fit to the Michaelis-Menten equation (Equation 4.1). The curves, which could not be fit to the Michaelis-Menten equation (Equation 4.1), were fit to the Hill equation (Equation 4.2). Figure 1a and 1b show examples of the velocity vs. substrate curves for WT at RT (h = 1) and variant D1 at RT (h = 1.85). For variants 133 and 262 the velocity-substrate could not be fit to Hill Equation 4.2 (Figure 1c and d).

$$\boldsymbol{v} = \frac{\boldsymbol{k_{cat}[E][S]}}{\boldsymbol{[S]+K_m}} \qquad \text{Equation 4.1}$$

$$\boldsymbol{v} = \frac{\boldsymbol{k_{cat}[E][S]^h}}{\boldsymbol{[S]^h+K'}} \qquad \text{Equation 4.2}$$

In the equations above, v is the velocity of the reaction, k_{cat} is the turnover rate, [E] and [S] are the enzyme and substrate concentrations respectively. K_m in Equation 4.1 is the Michaelis constant, K' in Equation 4.2 is a constant that relates to the dissociation constant and finally, h is the Hill coefficient.

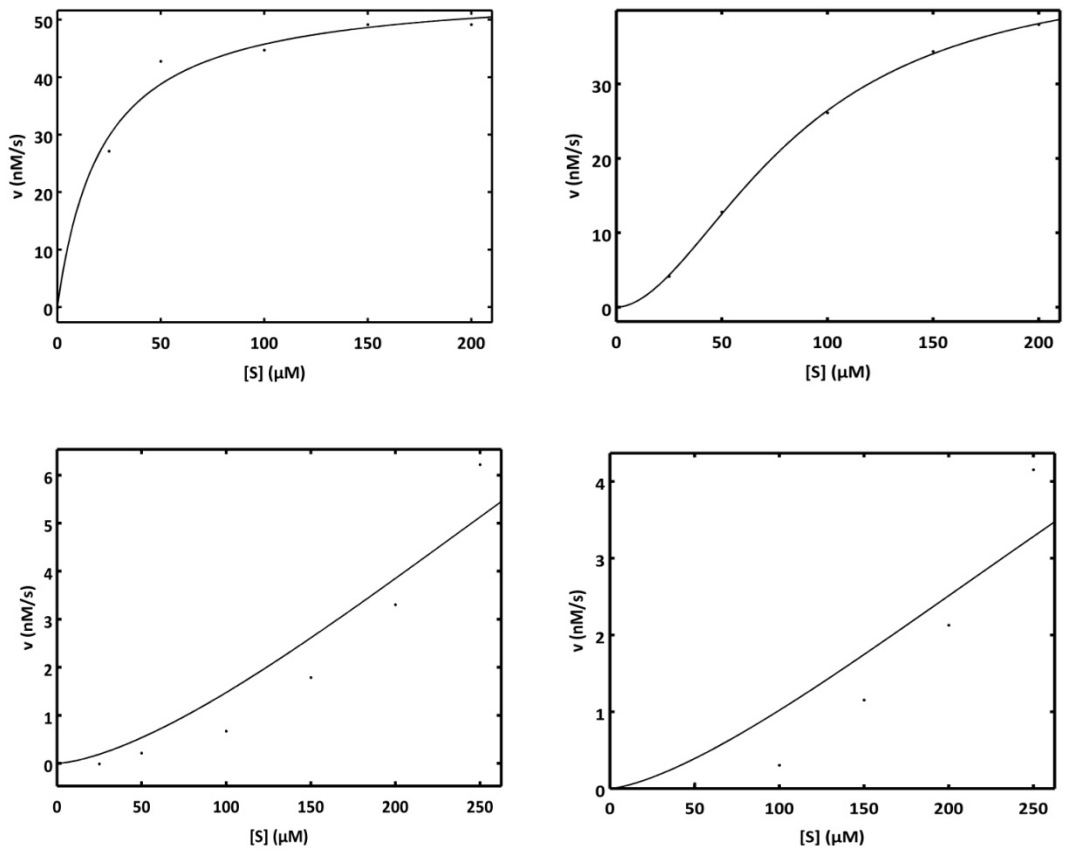


Figure 4.8 Velocity Vs. substrate concentration curves for (a) WT (b) D1 (c) 133 and (d) 262. All the experiments were performed at RT. (a) was fit to Equation 4.1, while (b) was fit to Equation 4.2. (c) and (d) could not be fit to Equation 4.2.

4.4.2 Comparison of single molecule and bulk activity

	Velocity @ RT (s^{-1})		Velocity @ 37°C (s^{-1})	
	Single Molecule	Bulk	Single Molecule	Bulk
WT	50.9 ± 9.0	62.5 ± 16.1	88.3 ± 25.1	74.8 ± 11.3
D1	42.8 ± 7.3	37.8 ± 1.2	79.6 ± 17.4	80.9 ± 16.1
D2	33.1 ± 18.3	50.2 ± 4.5	90.3 ± 38.3	87.3 ± 27.8
D3	33.8 ± 5.9	42.4 ± 6.4	88.2 ± 44.4	76.4 ± 16.1
D4	24.3 ± 19.4	31.3 ± 22.8	72.8 ± 12.8	43.5 ± 10.9

133	33.16 ± 6.0	9.6 ± 2.1	79.5 ± 7.9	30.3 ± 19.6
262	24.1 ± 10.0	8.8 ± 4.5	78.1 ± 7.9	45.2 ± 18.7

Table 4.6 Single molecule and bulk velocities calculated for different variants at RT and 37°C. Results from triplicate experiments were used to calculate the errors. For the bulk velocities, Equations 4.1 or 4.2 were used to calculate the velocities at 250 µM substrate concentration.

4.5 Methods and materials

4.5.1 Materials

GUS proteins received from UT Austin were aliquotted and stored in -80^{o}C. 5 mM stock solution of resorufin glucuronide sodium salt (Sigma Aldrich) was prepared in DMSO, and aliquots were stored in -20^{o}C.

4.5.2 Bulk experiments

Bulk experiments were performed on a microtiter plate reader to test for enzymatic activity. 100 µM of substrate solution was prepared in the reaction buffer (50 mM Na_2PO_4, 0.1% Triton X-100, 10 mM EDTA), and then 10 µL of 2 nM enzyme was mixed with 90 µL of 100 µM substrate solution. The fluorescence intensity (λ_{ex}=558 nm, λ_{em}=590 nm) was measured in the plate reader at 30 second intervals for 0.5 hours. All of the experiments were performed both at room temperature and at 37°C.

4.5.3 Imaging and data analysis

Time lapse imaging of the fluorescent product generated as a result of enzyme catalysis, was performed with an upright Olympus BX61 microscope equipped with a short arc mercury lamp (Ushio, Tokyo, Japan) and a CMOS camera (ORCA-Flash 4.0, Hamamatsu, Japan), and a

custom built automated sealing stage. A filter set with λ_{ex} = 558 nm and λ_{em} = 577 nm (Semrock, Rochester, NY) was used for imaging the resorufin product. Images of the distal end of the fiber were captured every 30 seconds for a period of 30 minutes or less. The images were analyzed using the MATLAB software Image Processing Toolbox.

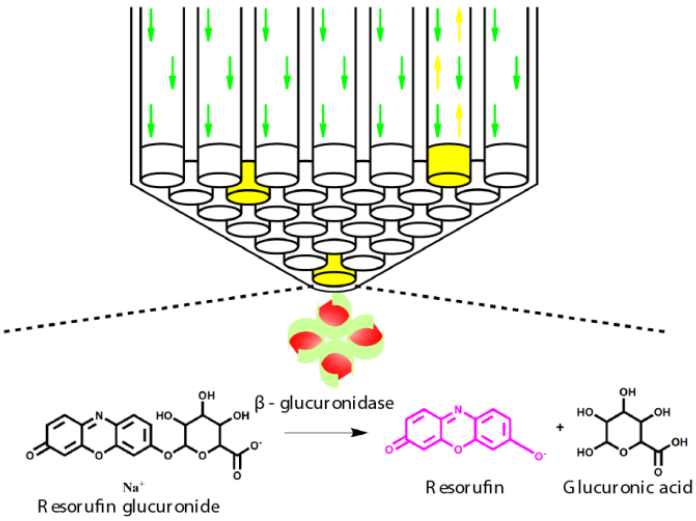


Figure 4.9 46 *f*L optical fiber microwells are generated at the distal end of the fiber bundle by etching the fibers for 115 seconds with 0.025 M HCl. The enzyme is diluted into a solution containing excess of the substrate (REG) such that the final concentration of the enzyme is 2 pM. This solution is immediately sealed into the array of 46fL wells. Using the Poisson distribution statistics $P_\mu(x) = e^{-\mu} X\mu^x/x!$ where $P_\mu(x)$ gives the probability of finding x molecules per well, μ being the mean number of molecules per well, it can be determined that for a concentration of 2 pM, there is a very low probability of finding more than one enzyme per well. Hence every well contains either one or zero enzyme molecules. The turnover of REG to resorufin is used to detect the presence of the active species, as shown in the figure.

4.6 References

(1) Rotman, B. *Proceedings of the National Academy of Sciences of the United States of America* **1961**, *47*, 1981.

(2) Craig, D. B.; Arriaga, E. A.; Wong, J. C. Y.; Lu, H.; Dovichi, N. J. *Journal of the American Chemical Society* **1996**, *118*, 5245.

(3) Rojek, M. J.; Walt, D. R. *PLoS ONE* **2014**, *9*, e86224.

(4) Wallace, B. D.; Wang, H.; Lane, K. T.; Scott, J. E.; Orans, J.; Koo, J. S.; Venkatesh, M.; Jobin, C.; Yeh, L.-A.; Mani, S.; Redinbo, M. R. *Science* **2010**, *330*, 831.

(5) Islam, M. R.; Tomatsu, S.; Shah, G. N.; Grubb, J. H.; Jain, S.; Sly, W. S. *Journal of Biological Chemistry* **1999**, *274*, 23451.

CHAPTER 5

Screening Oligomer Libraries for Rare Biomimetic Catalysts

5.1 Introduction

The goal of this project is to develop novel biomimetic catalysts for reactions that are extremely difficult to achieve in a test tube in the absence of a catalyst. Developing catalysts that mimic the high efficiency of enzymes with desired substrate selectivity, while overcoming some limitations of enzymes (thermal instability, incompatibility with organic solvents, narrow substrate selectivity) remains a long-term aim in the biomimetic catalysis field[1]. Various scaffolds such as supramolecular assemblies[2], polymers[1], nanomaterials[3] and small molecule compounds[2,4] have been employed to design biomimetic catalysts. Moderate success has been achieved in the field of biomimetic catalysis, including some cases where biomimetic catalysts had reaction rates exceeding those of enzymes[2].

Enzymes and other naturally-occurring catalytic biomolecules such as poly-D-ribonucleotides are all folded polymers with specific sequences and structures. The specific structures of these biomolecules enable them to have substrate binding sites, conformational flexibility, and a hydrophobic microenvironment, all of which are important features in enzyme catalysis[5]. Therefore, it can be hypothesized that polymeric backbones that are able to adopt a specific conformation are ideal molecular scaffolds for the design of effective biomimetic catalysts. This project explores polymeric backbones other than poly-L-α-peptides and poly-D-ribonucleotides to identify specific sequences that are expected to have a discrete folding pattern and can act as catalysts. Examples of catalysis by oligomers with unnatural backbones can be found in the literature[6-9]. Polymers that were shown to adopt a specific conformation were used as scaffolds to incorporate catalytic groups and tested for catalytic activity. For example m-phenyleneethynylene (mPE) polymers were shown to adopt a helical conformation in certain solvents[10,11]. 4-dimethylaminopyridine (DMAP) modified mPE was demonstrated to have high

reactivities with methyl sulfonates[6]. Peptoid helices were used as oligomer backbones to attach TEMPO (2,2,6,6-tetramethylpiperidine-1-oxyl), a catalyst known to promote oxidation reactions[7]. The designed catalysts were employed for oxidative kinetic resolution of racemic mixtures of secondary alcohols. β-peptide oligomers with sequences that were shown to adopt specific conformations were used as backbones for the synthesis of aldolase mimics[8]. In these and other studies, rational design of the catalysts was carried out based on knowledge of the backbone structure.

Due to our incomplete understanding of the relationship between sequence and structure for unnatural oligomer backbones, it is challenging to design sequences that are expected to fold into specific secondary structures. As an alternative strategy, large oligomer libraries can be screened to identify rare active catalysts. Combinatorial synthesis strategies have been described in the literature to generate large libraries of polymers that have functional moieties attached to a polymer backbone in both random orders and random ratios. For example, different ratios of carboxylic acids were covalently attached to a polyallylamine backbone to generate combinatorial libraries of polymers that were screened for phosphodiesterase activity[12] and a poly(acrylic anhydride) backbone was used for hydratase activity[13]. In these studies, the polymer pools were tested for bulk catalytic activity instead of isolating the individual catalytic polymers within the pool. The Miller group was able to generate a large library of α-amino acid oligomers on beads using solid phase peptide synthesis. They used the split and pool method[14] which enabled them to have only one member of the library to be attached per bead. In this manner, by immobilizing a fluorogenic indicator for catalytic activity on the beads, it was demonstrated that the active members of the library can be selected by looking for the brightest beads. By sequencing the α-amino acids on the beads, the sequences of the active members of the large

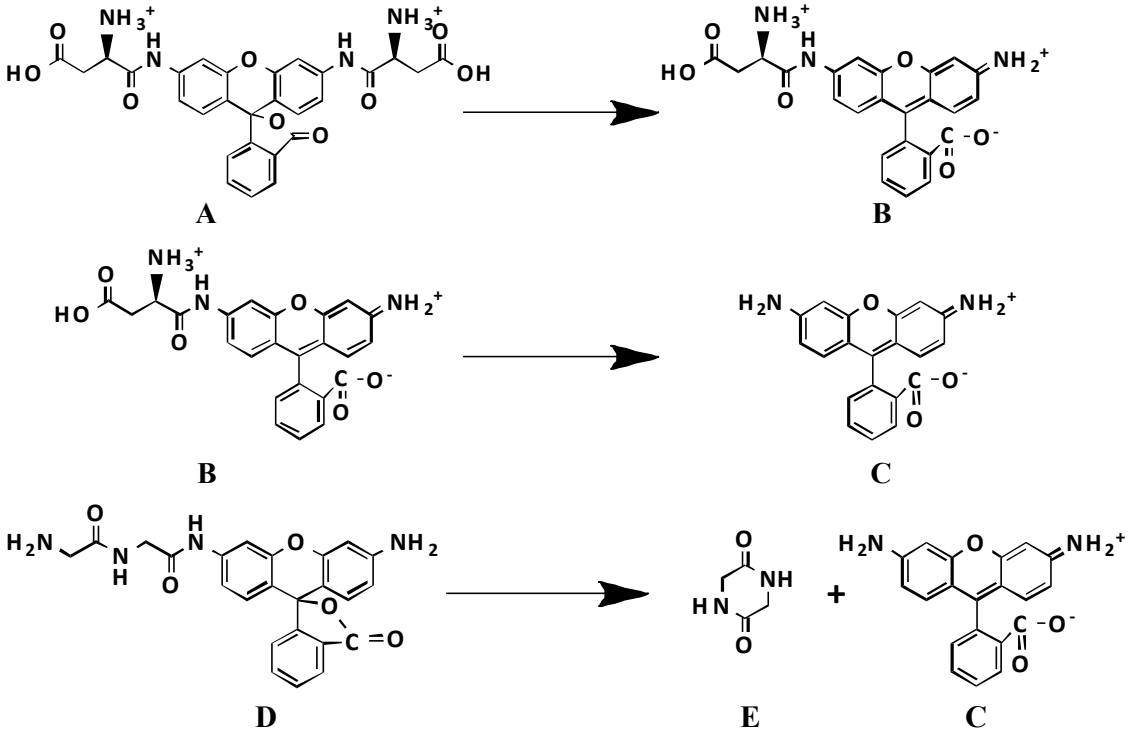
library could be pinpointed[15-17]. Inspired by this work, significant research has been done towards identification of α-amino acid catalysts[18]. However, the preliminary aim in these examples was to find synthetic catalysts that demonstrate high regio- or stereo-selectivity and not necessarily to discover catalysts with high reactivity.

The goal of this project is to screen large oligomer libraries containing unnatural building blocks for catalysts that (a) have a negligible background reaction rate in water and in the absence of a catalyst (b) are difficult to achieve in a test tube and (c) are known to proceed at a recognizable rate in the presence of an evolved biocatalyst[19]. Examples of such reactions include amide hydrolysis, phosphate ester hydrolysis and glycoside hydrolysis. While these reactions are difficult to achieve in a test tube (the uncatalyzed reactions proceed at rates of ~$10^{-11}s^{-1}$, $10^{-20}s^{-1}$ and $10^{-15}s^{-1}$ respectively[19]), they can be carried out with ease by evolved enzymes such as proteases, phosphatases and glycosidases respectively. A screening strategy for the identification of possible rare catalysts from large pools of hetero-oligomers is described and the preliminary research towards the development of novel biomimetic catalysts is presented in this chapter.

5.2 Experimental strategy

The experimental strategy involves the synthesis of randomized hetero-oligomer libraries followed by screening these libraries for the presence of rare (as few as one) catalysts in the library using high density femto-well optical fiber arrays. This project is performed in collaboration with the Gellman lab at the University of Wisconsin-Madison, where the oligomer libraries were designed and synthesized followed by screening for activity in the Walt lab. For initial feasibility studies, oligomers of ~20 chain length were synthesized. Two synthesis

methods have been utilized so far – anionic ring-opening polymerization[20,21] (AROP) to generate β-lactam peptide pools and solid phase peptide synthesis to generate α/β-amino acid peptides[9].



Scheme 5.1 Some substrates employed for detection of catalytic activity (A, B and D). Hydrolysis of amide bonds in rhodamine 110, bis-(L-aspartic acid amide) A, results in B, the mono aspartic acid amide of rhodamine 110, which is moderately fluorescent. Further hydrolysis of B results in highly fluorescent rhodamine 110 C. D is an intramolecular transamidation substrate, which results in a fluorogenic product C and a cyclization product E. The cyclization of GlyGlyNHMe should make this reaction more facile than the first two reactions.

To screen the oligomer libraries for activity, the oligomer pool containing up to 10^9 individual oligomers is partitioned into the 60000 microwells of the femto-well array described in earlier chapters, and then sealed along with an excess of a fluorogenic substrate. The fluorogenic substrate is selected such that it requires one of the difficult transformations discussed above to occur in order for the fluorophore to be released (Scheme 5.1). As a result, if

the oligomer pool contains any catalyst that can release the fluorophore from the substrate, the femto-wells containing active catalysts will fluoresce. Using different substrates, the oligomer libraries can be screened for different catalytic activities. The initial goals of the project are to screen oligomer libraries for the presence of catalysts that demonstrate activity beyond a certain threshold, evaluate the number of catalysts present within a certain pool and estimate the range of activities within an active pool

5.3 Materials and Methods

5.3.1 Materials

All the substrates and fluorophores were synthesized at the University of Wisconsin-Madison. Rhodamine110 was purchased from Sigma-Aldrich. 10X PBS (phosphate buffered saline) was purchased from Ambion.

5.3.2 Method of sealing:

Two methods of sealing were employed as discussed in the Results.

Oil Sealing: Oil sealing protocols were adopted from Reference 22[22] (Figure 5.1). Briefly, the fiber bundle was first polished and then etched for 1 minute 55 seconds using 0.025 N HCl to generate an array of 46 fL sized wells (4.5 μm diameter). The fiber bundle was then functionlized with a hydrophobic silane (n-octadecyltrichloro silane – OTS) using a contact printing method so that only the cladding of the wells was coated with hydrophobic silane. A fresh solution of 2 mM OTS in anhydrous toluene was spincoated onto a thick layer (~3 mm) of PDMS followed by drying the PDMS using an air-gun for 1 minute 15 seconds. The etched fiber bundle was then brought into contact with the PDMS surface for 60 seconds. The contact angle

was maintained at 90 degrees using a z-axis microscope stage to mount the fiber bundle. After stamping, the fiber bundle was sonicated for 1 minute in methanol and then toluene to remove excess OTS that is not chemically bound to the fibers. The fiber bundle is then by dried using an airgun. The fibers were placed in the oven at 80 ^{o}C for 20 minutes. The functionalized fibers can be used for oil sealing for at least 1 week after preparation. To seal the samples within the femto-wells, the stamped end of the fiber bundle was sonicated in the reaction buffer for 1 minute followed by incubation in the polymer-substrate mixture for 1 hour. To ensure homogeneity of the sample, a stir rod was used for continuous mixing. The fiber was then attached to a drill press and spun in light mineral oil for approximately 30 seconds (Figure 5.2). The fiber bundle was then imaged from the imaging end.

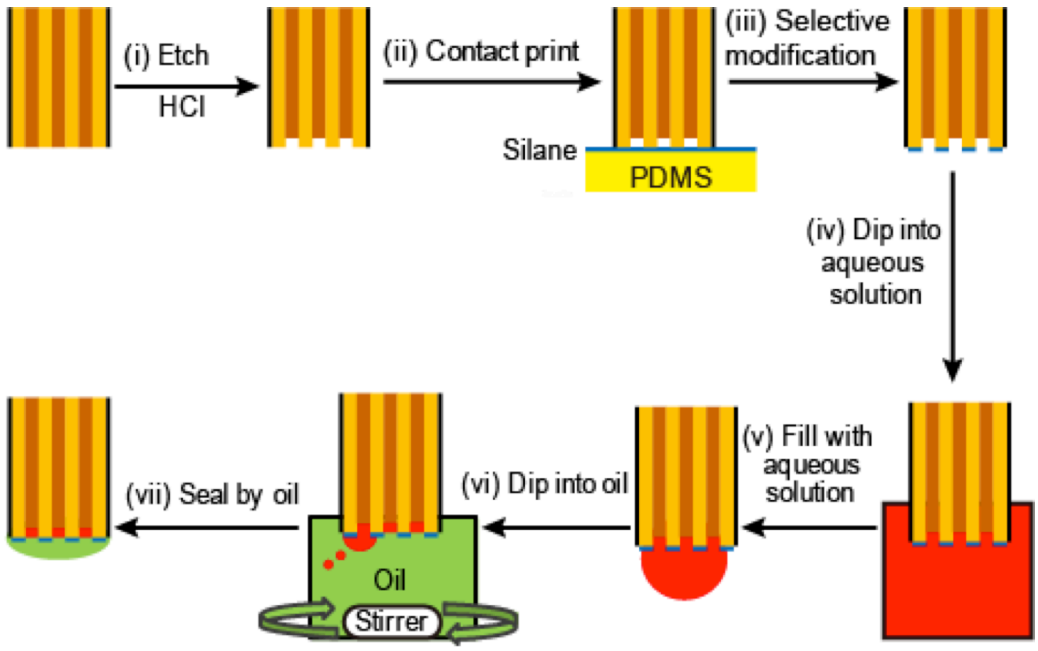


Figure 5.2 Schematic showing the procedure for oil sealing of the fiber bundles. Modification of the fiber bundle with OTS: First the fiber bundle is etched in HCl (i). Since the core (shown in red) etches faster than the cladding (shown in yellow), etching results in an array of femto-wells. A PDMS layer on which 2 mM solution of OTS has been freshly spin coated is brought into contact with the etched end of the fiber bundle for 1 minute (ii).

This procedure results in selective modification of the fiber bundle cladding into a highly hydrophobic surface (iii). The modified bundle is then dipped into the polymer substrate mixture for 1 hour (iv-v), followed by dipping into oil (vi). A shear force is required for the oil to seal the microwells. The shear force can be provided either by stirring the oil solution (vii) or fixing the fiber bundle to a drill press and spinning in an oil solution. The drill press method generates much higher shear force and was used in this study. Figure reprinted from Reference 22.

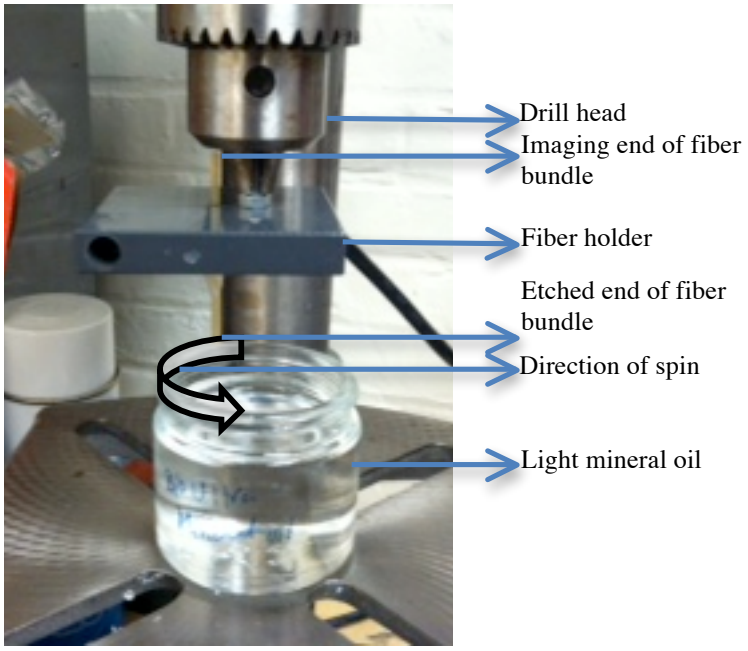


Figure 5.3 Drill press method employed to generate high shear forces between the fiber bundle and oil. The fiber bundle is fixed to a fiber holder which is in turn fixed to the head of the drill press (the etched and functionalized end of the fiber bundle is at the bottom). To seal the wells, the fiber bundle is immersed into the oil and spun for 30 seconds.

The second method of sealing used is the mechanical sealing discussed in Chapters 2 and 3. A custom built microscope with a sealing stage purchased from Quanterix Inc. was used for mechanical sealing in this project (Figure 5.3).

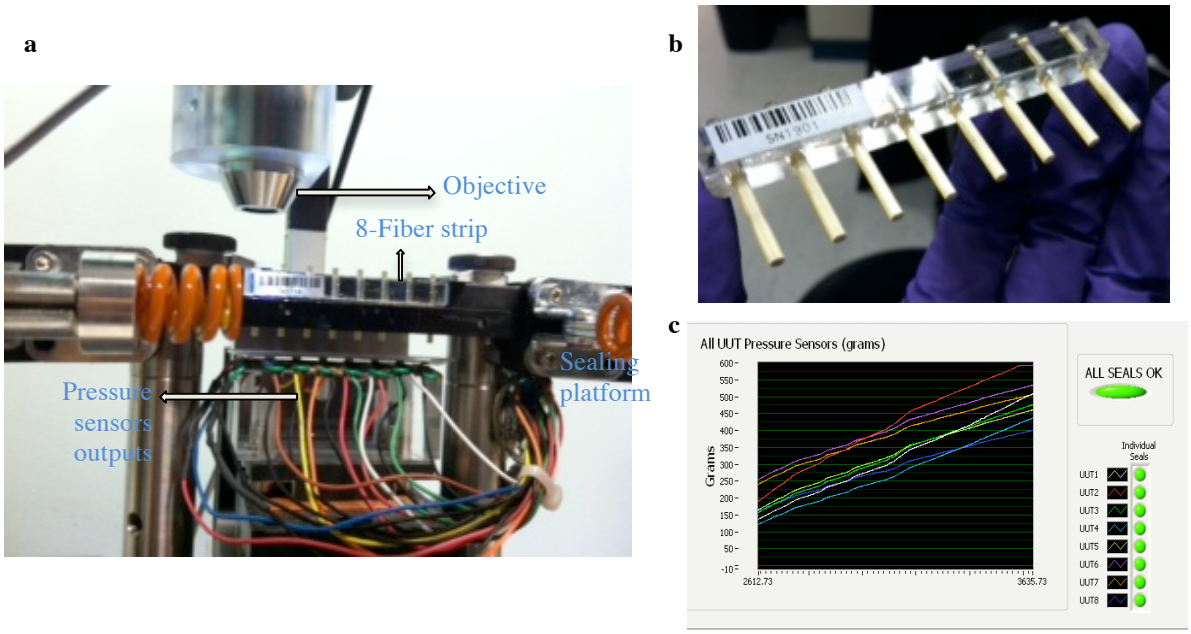


Figure 5.4 (a) Sealing component of the custom built microscope from Quanterix used for mechanical sealing. A strip containing 8 fiber bundles (shown in b) can be fixed to the microscope stage, which is programmed to move in both the x and y directions so that each fiber bundle can be imaged. The sealing platform is directly below the fiber strip and contains 8 diving board shaped platforms, each containing a pressure sensor underneath. A PDMS layer with 8 droplets of solution can be placed on the sealing platform. During the process of sealing, the sealing platform moves in the z-direction towards the fiber strip. Panel (c) shows the output from all the 8 pressure sensors once the sealing platform comes into contact with the fiber strip. When the output from each pressure sensor exceeds 400 grams, the sealing platform stops moving and all the fiber bundles are assumed to be sealed.

5.3.3 Imaging

For mechanical sealing described in Figure 5.3, the microscope is programmed to acquire the images of all the fiber bundles immediately after sealing. The custom build microscope is equipped with a 200 Watt metal arc lamp (Prior Scientific, Rockland, MA) and a CCD camera (Infinity4-11, Lumenera, Ontario, Canada). Fiber bundles sealed with oil were mounted on the microscope stage of an upright Olympus (Tokyo, Japan) BX61 microscope equipped with a short

arc mercury lamp (Ushio, Tokyo, Japan) and a CMOS digital camera (Orca-Flash 4.0; Hamamatsu).

In both cases, time lapse images were acquired with a 5 min interval and an exposure time of 1s. A filter set with λ_{ex} = 475 nm, λ_{em} = 540 nm and dichroic = 506 nm (Semrock) was used for the rhodamine as well as the pyranine product.

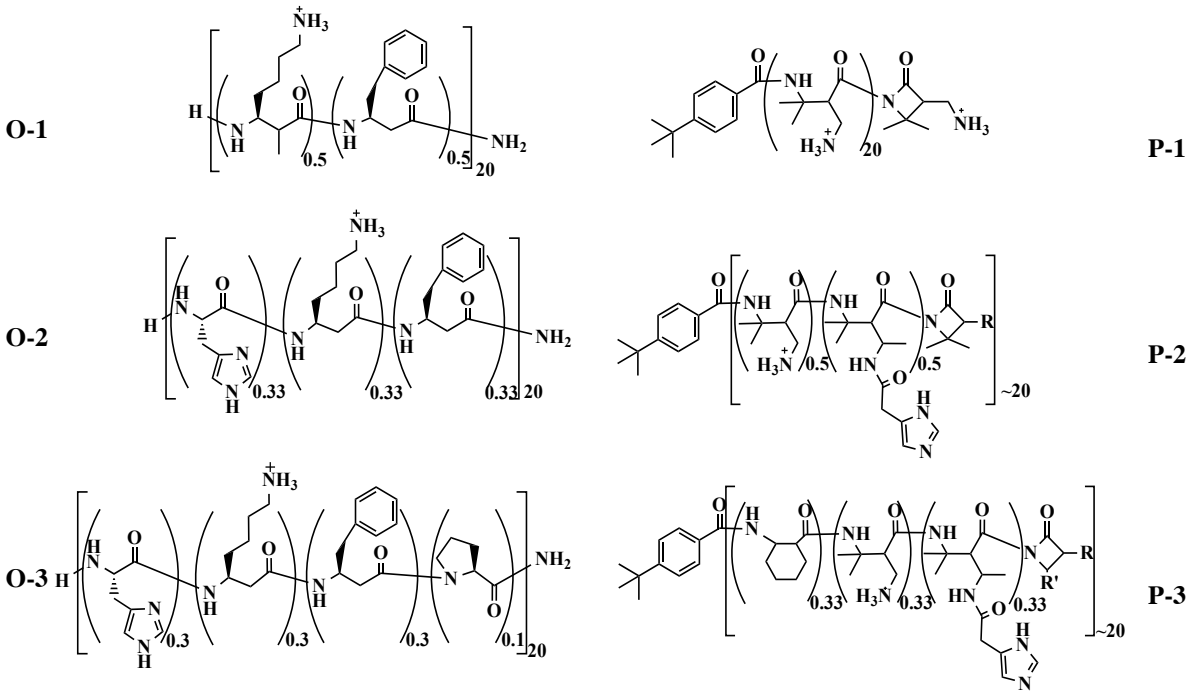
5.4 Results

5.4.1 Polymer pools

Preliminary results: Initial screens were performed on six different oligomer pools synthesized by researchers at University of Wisconsin-Madison, as shown in Scheme 2. Subunits such as imidazole and monomers with carboxyl or amine side groups were used in the synthesis of the pools as these sidechains were expected to promote catalysis. Oligomer pools O-1, O-2 and O-3 were prepared using solid phase peptide synthesis methods. Each oligomer is 20 amino acids long and the fraction of the amino acids used to synthesize each of the pools is indicated. For example, for pool O-1, at each coupling step, a 1:1 mixture of protected amino acids (S)-Fmoc-β^3-homophenylalanine (β^3-hPhe) and (S)-Fmoc-β^3-homolysine (β^3-hLys), was used. This pool is expected to have 2^{20} (1.05×10^6) oligomers. Similarly pool O-2 ($3^{20} = 3.5 \times 10^9$ oligomers) was prepared using a 1:1:1 mixture of protected β^3-hPhe, β^3-hLys and His. The expected pool size of O-3, which also contains proline in the amino acid mixture, is $4^{20}(1.1 \times 10^{12})$

Polymer pools P-1, P-2 and P-3 were prepared using AROP of the appropriate β-lactam. In pool P-1 only one monomer is used. Therefore the diversity arising in this pool is solely due to the stereochemistry. This pool is expected to have 2^{20} (1.05×10^6) diastereomers. P-2 and P-3 were prepared by copolymerization of multiple β-lactams and are expected to have pool sizes of

4^{20} (~10^{12}) and 6^{20} (~10^{15}) respectively. Unlike the solid phase synthesis methods used for generation of oligomer libraries, the polymerization methods for the polymer pools will generate a distribution of chain lengths. Therefore, the number of diastereomers generated could be larger.



Scheme 5.2 Structures of polymer pools obtained from the University of Wisconsin-Madison. The structures on the left correspond to β-peptide polymers synthesized with solid phase peptide synthesis and the structures on the right correspond to the AROP polymers.

For the initial experiments described in this chapter, 1-10 μM of the polymer was mixed with 10 μM of substrate and sealed within the fiber wells. For the smallest pool size, for example O-1 with 10^6 members, sealing 1 μM of polymer within the fiber array results in sealing 1.4 x 10^9 molecules, which samples the entire library 1000 times. Therefore, even when the pool contains as little as 1 oligomer, up to 1000 wells can be expected to light up if all the active catalysts are distributed in distinct wells. This statistic makes the femto-well array a novel, ideal platform for the screening proposed in this project. In the case of larger libraries, for example O-

2 (10^{12}) and O-3 (10^{15}), at a 1 μM polymer concentration the full library cannot be covered. In such cases, the polymer concentration must be increased to cover the whole library.

5.4.2 Experiments based on oil sealing technique

Initially, the activity screening of the polymers was performed using oil sealing. Since the reactions chosen are quite challenging, in the initial screens the catalytic rates of any possible catalysts are expected to be very low ($<10^{-3}$ reactions/s). With a low expected reaction rate, longer periods of observation are required to generate a detectable signal within any active femto-wells. Mechanical sealing imposes a limit on the experimentation time due to evaporation of the reaction mixture solution through the PDMS[23,24]. In addition, the use of PDMS for the gasket in mechanical sealing experiments limited each imaging step to less than 30 minutes due to absorption of organic compounds into the PDMS[25]. On the contrary, no limit to the experiment time has been observed for oil sealing[22], as evaporation of solution or adsorption of organic compounds is not an issue. Loss of substrate/product into the oil layer can be avoided by ensuring that none of the reaction components are miscible in oil. Therefore, oil sealing was the first sealing method used for the screening experiments.

Initial screens were performed for amide hydrolysis using commercially available rhodamine 110, bis(L-aspartic acid) amide substrate (D_2R110). The standard curve for rhodamine 110 is shown in Figure 6.9. One unit of intensity corresponds to 32 molecules of rhodamine. Based on the calibration curve, 1 μM of rhodamine 110 is considered as the limit of detection for the product, because a minimum of ~1000 units of intensity is required to be generated to be considered as signal. Therefore, if the solutions were observed for over 24 hr, the polymers with a minimum activity of 0.4 s^{-1} can be detected. For initial experiments, 1 μM of

polymers were sealed with 10 μM of D_2R110 in the femto-wells. Images were taken 2 hours and 18 hours after sealing the solutions in the fiber wells. Signal was observed after 2 hours of sealing the mixtures in the fiber wells. Figure 6.4 shows fiber snapshots for oligomer pool O-1 at 0 and 2 hours after sealing the polymer/substrate mixture in the fiber wells.

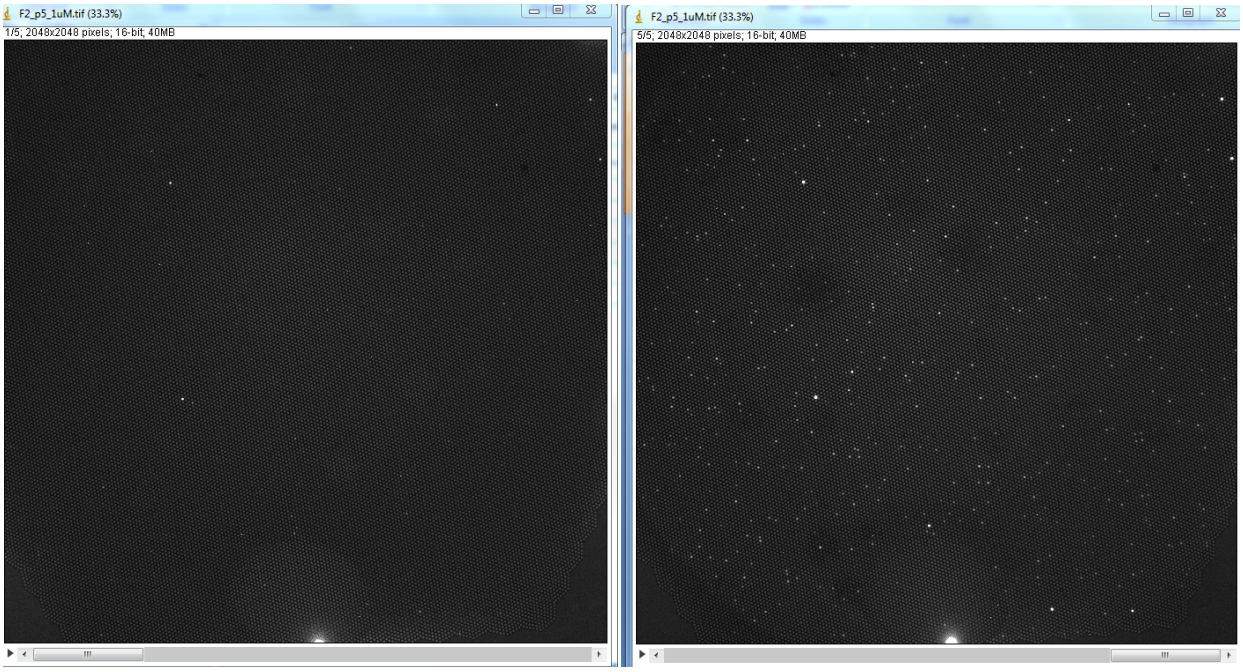


Figure 5.4 Snapshots of the fiber bundle taken 0 hours (left) and 2 hours (right) after sealing 1 μM of pool O-1 with 10 μM of the rhodamine bis aspartic acid substrate. Several wells exhibit a dramatic increase in fluorescence.

The activity was quantified by counting the number of wells that were observed to be fluorescent. Figure 5.5 summarizes the results obtained from oil sealing experiments performed for pools shown in Scheme 6.1. Pool P-1, which is mainly composed of different diastereomers displayed the least amount of activity (fewer than 20 active wells), while pool O-1 displayed most activity. However, the results could not be reproduced. No signal was observed in other fibers prepared by the same technique for the same experiments. It is not clear if this result caused by (a) unsuccessful stamping of OTS resulting in improper sealing of solutions within the

wells, or (b) lack of activity within the oligomer pools. Due to the inconclusive nature of the results and the lack of reproducibility within several trials, the minimal signal observed was determined to be an artifact. Due to the lack of reproducibility in the oil sealing technique itself (stamping of the fiber onto the OTS layer and the actual sealing between wells), in addition to the ambiguity of whether the observed signals were artifacts or due to polymer activity, mechanical sealing was determined to be a superior approach for this project. Mechanical sealing does not require any fiber functionalization, has better throughput (eight experiments can be performed simultaneously using an eight fiber strip) and is reproducible for well-to-well sealing, and should provide more conclusive results. The disadvantage of mechanical sealing is the limitation on the experimental time (less than 30 minutes). As a result only highly active polymers within the oligomer pool can be detected using mechanical sealing.

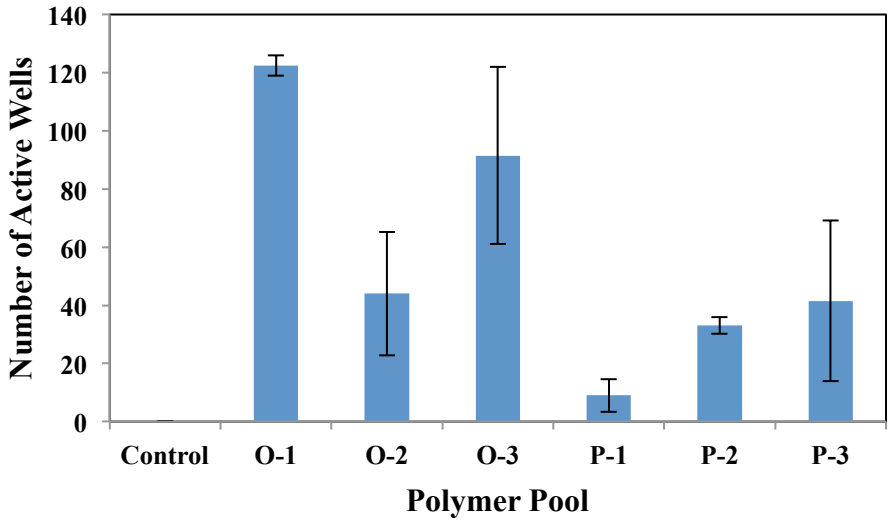


Figure 5.5 The number of active wells observed for different oligomer pools after sealing 1 µM of oligomer with 10 µM of the rhodamine bis aspartic acid substrate.

5.4.3 Mechanical sealing experiments

The polymer pools in Scheme 5.1 were tested with different substrates for amide hydrolysis (summarized in Table 5.1). The polymer concentrations were varied for each substrate. With mechanical sealing no activity was detected for any the polymer/substrate mixtures in Table 5.1.

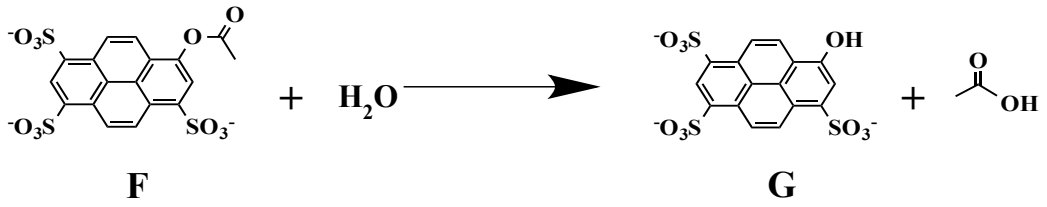
Substrate	**Substrate Concentration**	**Polymer Concentration**
Rhodamine bis-(L-aspartic acid amide)	10 μM	1 μM
		10 μM
		100 μM
Rhodamine mono-(L-aspartic acid amide)	10 μM	100 nM
		1 μM
		10 μM
		100 μM
GlyGlyRhodamine	10 μM	1 μM
		10 μM
		100 μM
Nile Blue aspartic acid amide	100 μM	1 μM
Rhodamine bis-urea substrate	10 μM	1 μM

Table 5.7 Panel of experiments conducted using mechanical sealing.

5.4.4 Positive control experiment

Before moving on to new polymer pools and substrates, the activities of the initial six polymer pools were tested with an ester substrate (Scheme 5.3). Since hydrolysis of the ester in scheme 5.3 should be easier than the amide, all the oligomers are expected to hydrolyze this

substrate. As a result, this substrate can be used as a control to test the performance of the mechanical sealing platform. The control substrate ensures that the results obtained for the more challenging substrates are due to a lack of active oligomer and not due to improper sealing.



F G

Scheme 5.3 The pyranine ester substrate F was used with polymers to test their ester hydrolysis activity. Successful hydrolysis releases pyranine G which is fluorescent.

Figure 5.6 shows the results obtained with the ester substrate containing 10 μM of polymer and 10 μM of ester substrate. Snapshots of the fiber bundle at 0 and 30 minutes after sealing are shown. Several wells begin to light up right after sealing and increase in intensity with time. When compared to the standard curve of pyranine (Figure 5.10) under the same imaging conditions, the intensities generated in this experiment were much higher than the intensities corresponding to 10 μM of pyranine. This result suggests that over 100% reaction yield has been achieved, which is not possible. Performing a standard curve of the pyranine in the presence of 10 μM of oligomer pool O-1 revealed a few discrepancies (Figure 5.7). First, in the presence of the polymers, the fluorescence of all the wells was not uniform, even though all the wells were filled with the same concentration of pyranine product (2.5, 5, 7.5 or 10 μM product). As shown in Figure 5.7, a significant fraction of the wells are much brighter than the rest of the wells. Some wells are saturated with respect to the sensitivity of the detector. Second, the fluorescence of pyranine enclosed within the wells increases with time in the presence of the polymer. The intensity of the wells increases with increasing fluorophore concentration and with

increasing polymer concentration as well. In the absence of any polymer, the fluorescence in the wells is uniform for all concentrations of pyranine. The intensity observed in the wells could be attributed to the precipitation of the polymers. The precipitation might increase with time resulting in an increase in fluorescence over time. On the other hand the pH within the wells may not be uniform among the wells, although that is unlikely due to the use of a strongly buffered pH (1X PBS, pH = 7.4). The difference in pH in the wells can result in different fluorescence signals; a higher pH would result in a higher fluorescence signal.

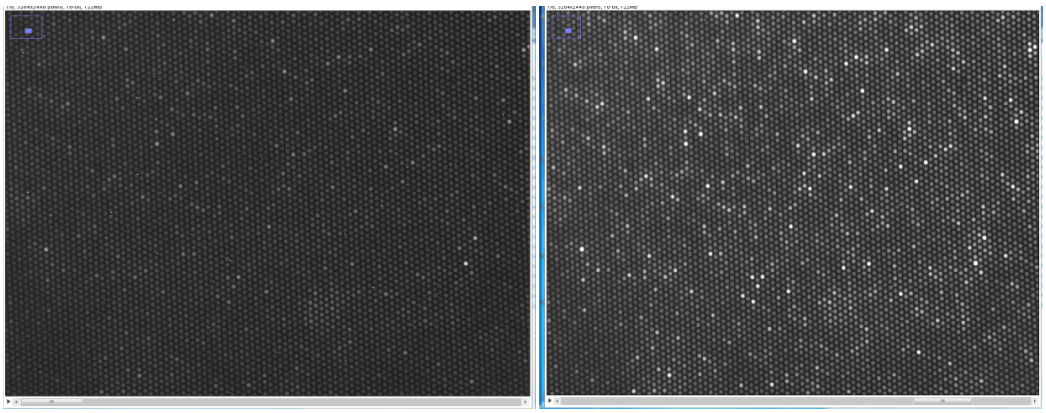


Figure 5.6 Snapshots of fiber bundle taken at 0 mins and 30 minutes after sealing 10 µM of the pool O-1 with 10 µM of the ester substrate.

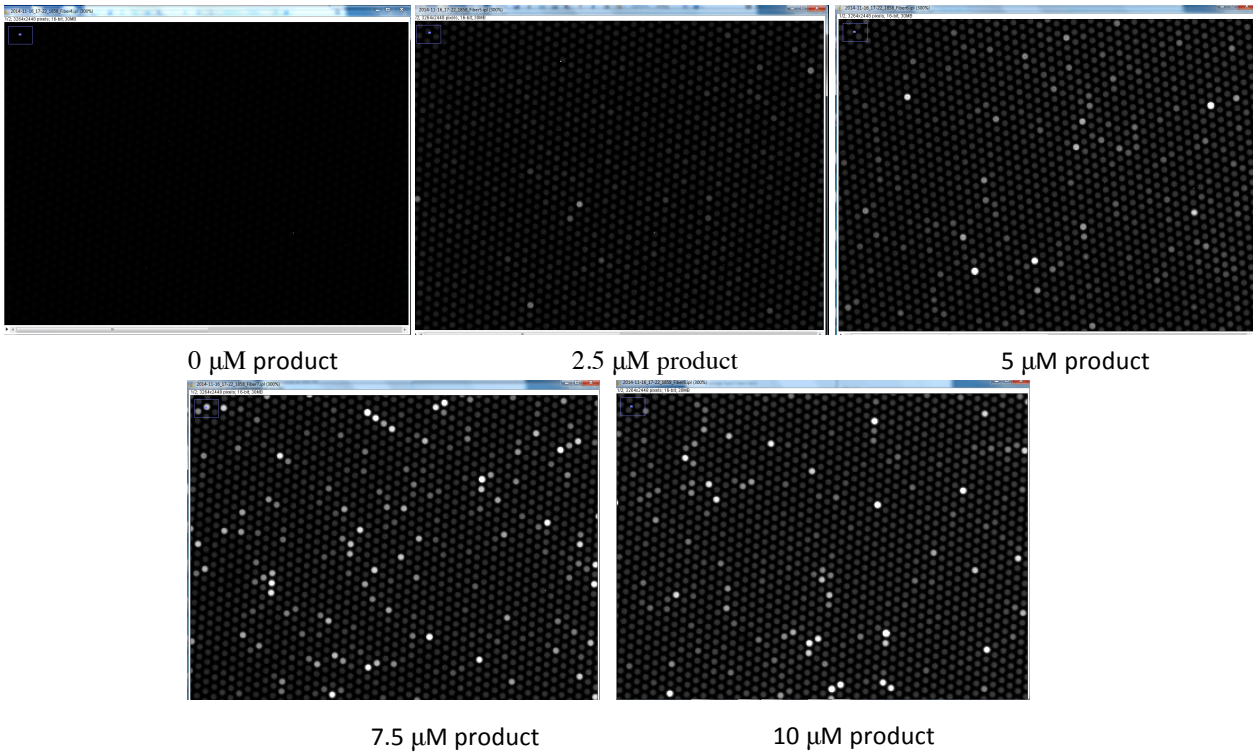


Figure 5.7 Snapshots of the fiber bundle after sealing the fiber wells with a mixture of 10 µM of the polymer pool O-1 and different concentrations of pyranine (product of the ester hydrolysis reaction). As the concentration of the pyranine increases, several wells are observed to have high intensities compared to the background.

To avoid these discrepancies, the polymer concentration was decreased to 1 µM. When the polymer concentration was reduced, activity can be observed with the pyranine ester substrate. In this case, the fluorescence of all the wells increases with time. Figure 5.8 shows a histogram of pyranine produced within the wells in the presence of 1 µM polymer and 10 µM of the substrate after 30 minutes (Figure 5.8a & b). A histogram of the concentration of pyranine in the wells for a control experiment (in the absence of polymer) is shown in Figure 5.8c. The concentration of the pyranine produced is higher in the experiments with polymers. Therefore activity could be observed for this polymer.

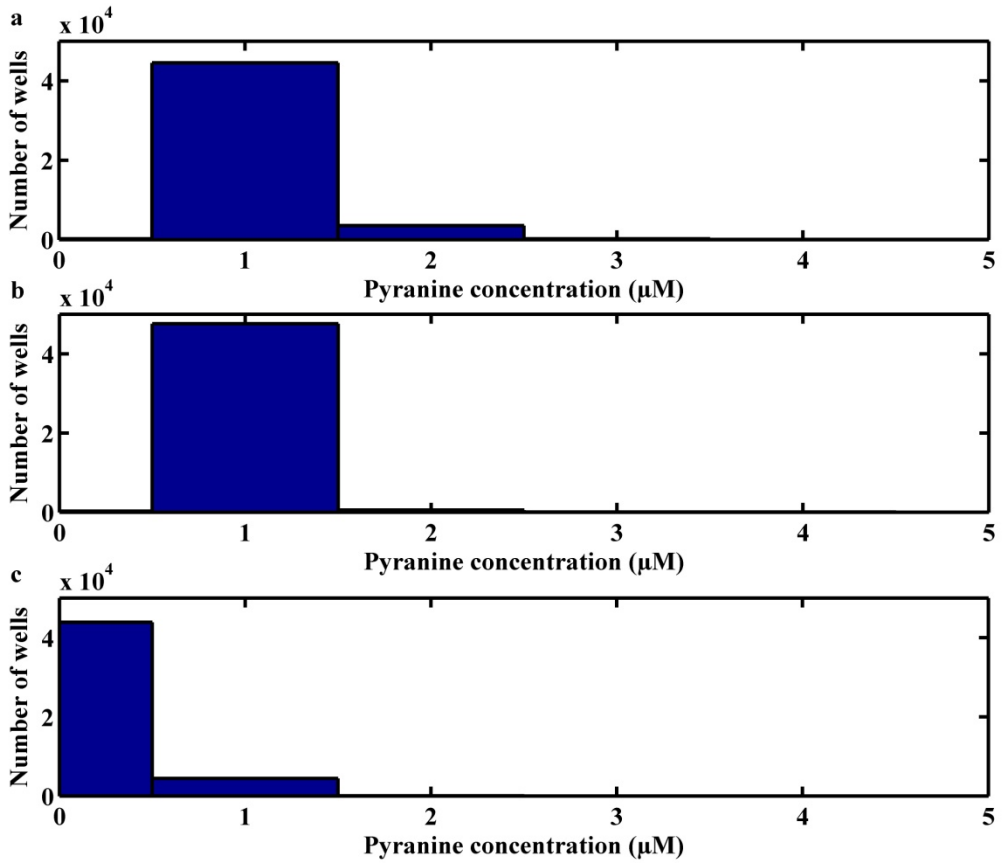


Figure 5.8 Histograms of the concentration of product generated in the wells after sealing 10 µM of the substrate with (a) 1 µM of polymer O-2 (b) 1 µM of polymer O-3 and (c) no polymer.

5.5 Conclusions and future work

In this work, mechanical sealing was established as a reliable preliminary platform compared to oil sealing for initial activity screening of polymers. Activity could be detected for two of the polymers tested in this study, and a positive control experiment was identified. The positive control experiment helps to determine the performance of the sealing strategies. After further experimental setup improvements, this method should enable the measurement of catalytic activity of polymer pools for longer times. In addition, as demonstrated in the results in

this chapter, the positive control experiment also helps us identify and understand possible sources of artifacts resulting from polymer aggregation and precipitation.

The progress described in this chapter is only a small step towards the ultimate goal of identifying biomimetic catalysts. Although the oil-sealing technique allows the reaction to be observed for a longer period of time, the sealing consistency of the platform needs to be improved in order to obtain reproducible results. This improvement is important because the biomimetic catalysts are expected to have slow reaction rates. The current mechanical sealing platform alone might not suffice as a screening platform.

5.6 Supporting Information

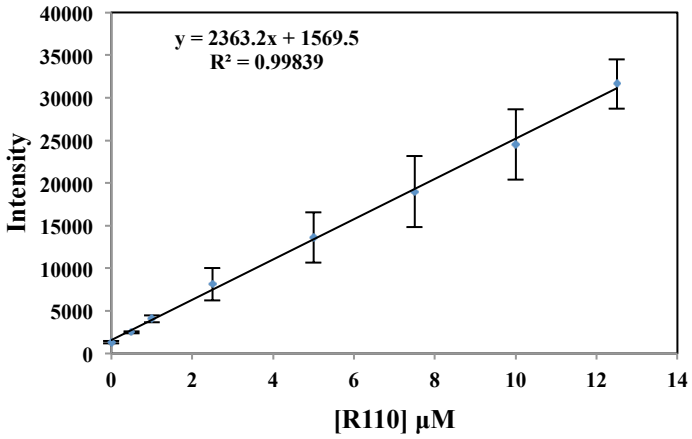


Figure 5.9 Calibration curve of rhodamine 110

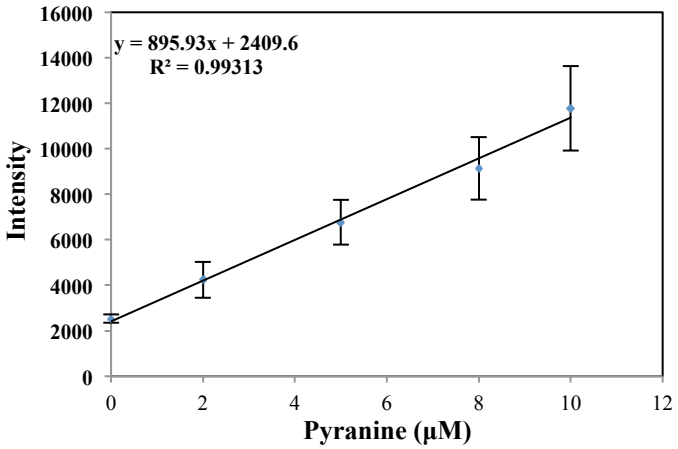


Figure 5.10 Calibration curve of pyranine

5.7 References

(1) Yin, Y.; Dong, Z.; Luo, Q.; Liu, J. *Progress in Polymer Science* **2012**, *37*, 1476.

(2) Marchetti, L.; Levine, M. *ACS Catalysis* **2011**, *1*, 1090.

(3) Lin, Y.; Ren, J.; Qu, X. *Accounts of Chemical Research* **2014**, *47*, 1097.

(4) Breslow, R.; Dong, S. D. *Chemical Reviews* **1998**, *98*, 1997.

(5) Kirby, A. J. *Angewandte Chemie International Edition in English* **1996**, *35*, 706.

(6) Smaldone, R. A.; Moore, J. S. *Journal of the American Chemical Society* **2007**, *129*, 5444.

(7) Maayan, G.; Ward, M. D.; Kirshenbaum, K. *Proceedings of the National Academy of Sciences* **2009**, *106*, 13679.

(8) Muller, M. M.; Windsor, M. A.; Pomerantz, W. C.; Gellman, S. H.; Hilvert, D. *Angewandte Chemie (International ed. in English)* **2009**, *48*, 922.

(9) Hayouka, Z.; Chakraborty, S.; Liu, R.; Boersma, M. D.; Weisblum, B.; Gellman, S. H. *Journal of the American Chemical Society* **2013**, *135*, 11748.

(10) Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G. *Science* **1997**, *277*, 1793.

(11) Prince, R. B.; Saven, J. G.; Wolynes, P. G.; Moore, J. S. *Journal of the American Chemical Society* **1999**, *121*, 3114.

(12) Menger, F. M.; Eliseev, A. V.; Migulin, V. A. *The Journal of Organic Chemistry* **1995**, *60*, 6666.

(13) Menger, F.; Ding, J.; Barragan, V. *The Journal of Organic Chemistry* **1998**, *63*, 7578.

(14) Furka, A.; Sebestyén, F.; Asgedom, M.; Dibó, G. *International journal of peptide and protein research* **1991**, *37*, 487.

(15) Harris, R.; Nation, A.; Copeland, G.; Miller, S. *Journal of the American Chemical Society* **2000**, *122*, 11270.

(16) Copeland, G. T.; Miller, S. J. *Journal of the American Chemical Society* **2001**, *123*, 6496.

(17) Lewis, C. A.; Miller, S. J. *Angewandte Chemie International Edition* **2006**, *45*, 5616.

(18) Davie, E. A. C.; Mennen, S. M.; Xu, Y.; Miller, S. J. *Chemical Reviews* **2007**, *107*, 5759.

(19) Wolfenden, R. *Annual Review of Biochemistry* **2011**, *80*, 645.

(20) Mowery, B. P.; Lee, S. E.; Kissounko, D. A.; Epand, R. F.; Epand, R. M.; Weisblum, B.; Stahl, S. S.; Gellman, S. H. *Journal of the American Chemical Society* **2007**, *129*, 15474.

(21) Lee, M.-R.; Stahl, S. S.; Gellman, S. H.; Masters, K. S. *Journal of the American Chemical Society* **2009**, *131*, 16779.

(22) Zhang, H.; Nie, S.; Etson, C. M.; Wang, R. M.; Walt, D. R. *Lab on a Chip* **2012**, *12*, 2229.

(23) Heo, Y. S.; Cabrera, L. M.; Song, J. W.; Futai, N.; Tung, Y.-C.; Smith, G. D.; Takayama, S. *Analytical Chemistry* **2006**, *79*, 1126.

(24) Yu, Y.-S.; Wang, Z.-Q.; Zhao, Y.-P. *Acta Mech Sin* **2013**, *29*, 799.

(25) Toepke, M. W.; Beebe, D. J. *Lab Chip* **2006**, *6*, 1484.

APPENDIX

6.1 Determination of "activity levels"

In a pre-steady state experiment, an inhibitor release event is accompanied with a step increase in the turnover rate of the enzyme. The turnover rate between two inhibitor release events is stable and is defined as an "activity level". The average turnover rate of each activity level is determined by taking the average turnover rate during the total time the enzyme dwells in that activity level. First, the time points at which an inhibitor release event occurs is determined by using a method based on student's t-test[2]. Then the turnover rates of all the time points between two subsequent inhibitor release events identified by the student's t-test method are averaged to obtain the turnover rate of the activity level.

To employ the student's t-test based method, the raw intensity data are first background subtracted and smoothed (Savitzky-Golay filter with a quadratic convolute and window length of 15) followed by conversion into turnover rates. To identify the inhibitor release events, the time points at which there is a sudden change in turnover rate have to be identified. At each time point the student's t-test is used to compare seven values before and after the point. When the turnover rate is constant within the background standard deviation, the t-value is close to 0. When there is a sudden change in the turnover rate, the absolute t-value calculated at that point is high. When a profile of the t-values is plotted, a spike in the t values indicates the inhibitor release events. By setting a threshold, the time points at which t-values are above the threshold can be assigned to inhibitor release events. Figure 6.1 demonstrates the use of the student's t-test method to identify inhibitor release events.

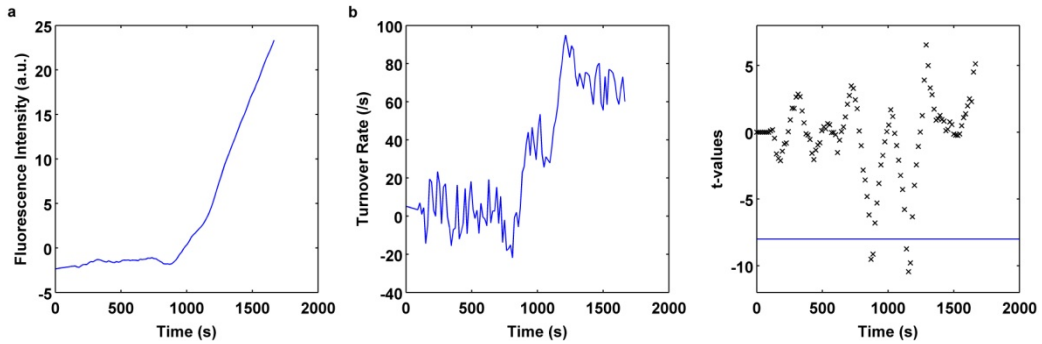


Figure 6.1—Demonstration of the application of t-test method to identify inhibitor release events. (a) shows the smoothed background subtracted intensity profile of a pre-steady state experiment. The intensity profile is then converted into a turnover rate (b) by multiplying the derivative of (a) by a calibration constant. The t-test is then applied to the turnover rate profile in (b) and the t-values are plotted in (c). By setting a threshold (blue line in (c), we can identify the inhibitor release events. Multiple points can be seen below the threshold for each inhibitor release event. The local minimum is assigned to an inhibitor release.

The t-test was also used to identify an inhibitor binding or release event in steady state experiments. In this case the t-value profile shows downward spikes for inhibitor release and upward spikes for inhibitor binding. Positive and negative thresholds are used to determine the inhibitor binding and release events respectively (Figure 6.2).

The t-test method was validated by application to a simulated data set consisting of background noise extracted from the turnover rate profiles of experimentally collected data with added steps (jumps in turnover rate) of known location and size. The limit of step size (i.e. the jump in turnover rate) that can be estimated using the t-test is 17 s^{-1} and the minimum dwell time for a step (i.e. an activity level) to be detected is 225 s.

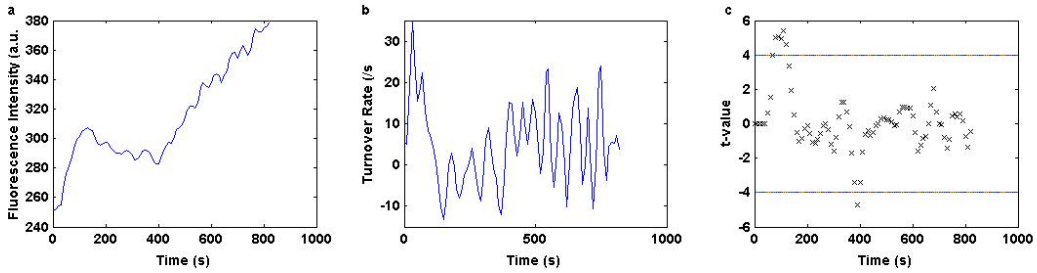


Figure 6.2—Demonstration of the t-test method to identify inhibitor binding and release events in a steady state experiment. (a) The smoothed fluorescence intensity trajectory from a steady state experiment. The fluorescence intensities are converted into turnover rates as shown in (b). The t-test is applied at each point in (b) where seven points before and after the point are compared using student's t-test. The t-values are plotted in (c). An upward spike indicates an inhibitor binding event, while a downward spike indicates an inhibitor release event. To accurately determine the events, the t-value profile is thresholded and the local maxima or minima beyond the thresholds are assigned to the binding or release events.

6.2 Determination of on and off times

To determine the on and off times in the pre-steady and steady state experiments, a velocity thresholding method was used. The turnover rate profile obtained in the t-test method was further smoothened using a boxcar filter (window length 5). A predetermined threshold was used to determine if the enzyme was in an on state or off state. In the turnover rate profile, all the points that are above the threshold are considered on and all the points below the threshold are considered off.

For the steady state experiments, the dwell time in each state as determined by the thresholding method provides the on and off times. To determine the off times for the pre-steady state experiments, the thresholding method was used to identify the first time point after which the turnover rate is above the threshold. This time point gives the off-time for the profile. Using the velocity thresholding method the on-times and off-times in the steady state experiments of NpBHC and D-galactal were computed. Figure 6.3 shows the on-time and off-time histograms.

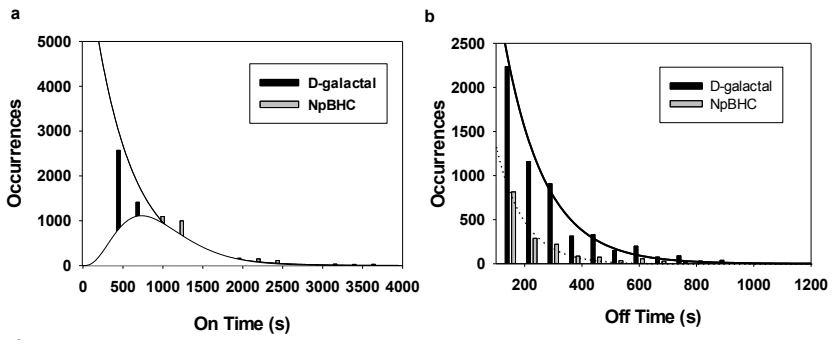


Figure 6.3—(a) and (b) On-time and off-time histograms for the steady state experiments in the presence of D-galactal and NpBHC. The histograms are each fit to a mono exponential decay as shown by the curves. The on-times were determined by measuring the duration the enzymes were active at the beginning of the steady state experiment before converting into an inactive state due to inhibitor binding. Since the association rates for D-galactal and NpBHC under steady state conditions are on the order of 10^{-3} s^{-1}, the majority of the

enzyme molecules were active at the beginning of the steady state experiments. The off-times are determined by measuring the duration the enzyme is inactive before it regains activity.

The histograms for D-galactal can be fitted to an exponential function to obtain an association rate constant of 10^2 $M^{-1}s^{-1}$ (Figure 6.3a). Histograms of the off-times (duration during which the enzyme is inactive) are plotted in Figure S4.1b and can be fit to an exponential decay as expected for a one step reaction. The dissociation rates calculated from the exponential fits are 7 x $10^{-3}s^{-1}$ for D-galactal and 1 x 10^{-2} s^{-1} for NpBHC. The on-time histogram for NpBHC shows a rise and decay profile as inhibitor binding is a multi-step process (Figure 6.3a). Assuming identical rates for each step of the multistep process, the on-time histogram is fit to the equation[3]

$$y = \frac{k^4 x^3}{4!} e^{-kx} \qquad \text{Eqn. 6.1}$$

Where k is the on-rate and is calculated to be 4.1 x 10^{-3} s^{-1}. The association rate constant is calculated to be 8.2 x 10^6 $M^{-1}s^{-1}$.

6.3 Autocorrelation Analysis

We determined the inhibitor exchange rates through the normalized autocorrelation function, $C_{(m)}T$, of the substrate turnover dwell times[4, 5] (Eqn. A4.1):

$$\mathrm{C_{(m)}T} = \frac{\langle \Delta \mathrm{T(0)} \Delta \mathrm{T(m)} \rangle}{\langle \Delta \mathrm{T^2} \rangle} = \frac{\sum_{\mathrm{i}} \Delta \mathrm{T(i)} \Delta \mathrm{T(i+m)}}{\sum_{\mathrm{i}} \Delta \mathrm{T^2(i)}} \qquad \text{(Eqn. 6.2)}$$

where m is the time between two measurements, T is the substrate turnover rate, $\Delta T(m) = T(m) - \langle T \rangle$, and the brackets denote averaging along a time trace. $C_{(m)}T$ has the following meaning for steady-state inhibition experiments: In the absence of inhibitor, $C_{(0)}T = 1$ and $C_{(m)}T = 0$ with $m > 0$, as the substrate turnover rates are long-lived and do not fluctuate. In the presence of inhibitor, $C_{(m)}T$ follows a single-exponential decay, if we apply classical chemical kinetics. The initial amplitude ($m = 1$) reflects the variance of the substrate turnover rates, $\langle \Delta T^2 \rangle$, and the decay rate indicates how rapidly an individual enzyme molecule binds ($\mathrm{k_{on}}$) and unbinds ($\mathrm{k_{off}}$) inhibitor molecules. Thus, the data points of the substrate turnover autocorrelation function of each individual β-galactosidase molecule can be fitted to Eqn. 6.3 to determine the inhibitor exchange rate k_c:[1]

$$\langle \Delta T(0) \Delta T(m) \rangle = \langle \Delta T^2 \rangle e^{-k_c m} \qquad \text{(Eqn. 6.3)}$$

$$k_c = k_{on}[I] + k_{off} \qquad \text{(Eqn. 6.4)}$$

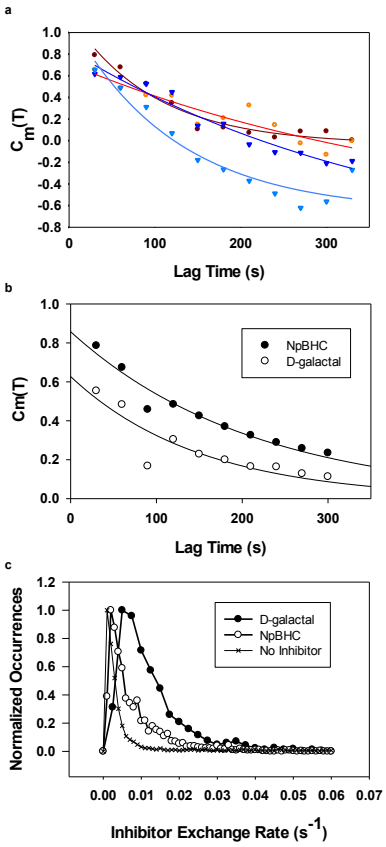


Figure 6.4—(a) Normalized autocorrelation of the substrate turnover rates shown in Figure 3b, computed as described in the data analysis. (b) Averaged autocorrelation function of the steady state turnover profiles of thousands of enzymes (4999 for D-galactal and 2993 for NpBHC). (c) Histogram of Inhibitor exchange rates for steady state reaction of β-galactosidase with D-galactal and NpBHC. The same data analysis was performed on a single molecule experiment of β-galactosidase in the absence of inhibitor and the extracted 'inhibitor exchange rates' are shown (No Inhibitor).

Figure 2.3b in the manuscript for example shows the noisy profile of substrate turnover rates of four individual enzyme molecules, where distinct states of activity can be observed. A plot of the corresponding autocorrelation function of the turnover rates to extract the inhibitor exchange rates is shown in Figure 6.4a. The autocorrelation function can be fit to a mono-exponential decay as shown. Histograms of the inhibitor exchange rates calculated from individual

trajectories using Eqn. 6.3 are shown in Figure 6.4c. The exchange rate computed in the absence of inhibitor is also shown.

6.4 References

1. D. F. Wentworth and R. Wolfenden, *Biochemistry*, 1974, 13, 4715-4720.
2. N. J. Carter and R. A. Cross, *Nature*, 2005, 435, 308-312.
3. D. L. Floyd, S. C. Harrison and A. M. van Oijen, *Biophysical Journal*, 2010, 99, 360-366.
4. S. Yang and J. Cao, *The Journal of Chemical Physics*, 2002, 117, 10996-11009.
5. H. H. Gorris, D. M. Rissin and D. R. Walt, *Proceedings of the National Academy of Sciences*, 2007, 104, 17680-17685.