Capillary Electrophoresis Laser-Induced Fluorescence Investigations of Individual Molecules of *Escherichia coli* β-Galactosidase

by

Ellert R Nichols

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Biochemistry and Medical Genetics

University of Manitoba

Winnipeg, Manitoba

Abstract

Fifteen years following the first demonstration of single molecule fluorescence detection, single molecule studies of enzymes have revealed that nominally identical individual enzyme molecules are functionally heterogeneous. Different individual molecules exhibit different catalytic rates under identical conditions, and individual enzyme molecules show fluctuating rates over broad timescales. The structural basis and the biological sources for such heterogeneity remains poorly understood. Herein, studies are presented of the β-galactosidase from *Escherichia coli*, using capillary electrophoresis with laser-induced fluorescence (CE-LIF), to investigate the sources of catalytic heterogeneity at the single molecule level. Two new single molecule CE-LIF assays for β-galactosidase were developed based upon the synthetic fluorogenic substrates 9H-(1,3-dichloro-9,9-dimethyacridin-2-one-7-yl)-β-D- galactopyranoside (DDAO-gal) and fluorescein-β-D-digalactopyranoside (FDG). Limited proteolysis as a possible source for single molecule heterogeneity, and for the changes in activity of a population of individual molecules over time, was investigated by inducing enzyme expression in two *E.coli* strains in the presence of a broad spectrum of protease inhibitors. The effect of protease inhibitors was found to be limited. β-Galactosidase was expressed from a *lacZ* linear template from two different *E. coli* strains using an *in vitro* protein expression system to determine if *in vitro* synthesized enzyme was identical to its *in vivo* counterpart. *In vitro* synthesized enzyme was found to be less active than *in vivo* sources. The differences were attributed to deficient N-terminal methionine removal and the higher rates of translation error associated with *in vitro* protein synthesis. Single molecule separations revealed that individual molecules of β-galactosidase were

electrophoretically distinct, and that the electrophoretic heterogeneity was independent of source of enzyme, method of measurement or of capillary coating. Electrophoretic modeling indicated that slight variations of shape and charge could account for the observed range. The magnitude of these differences, and published rates for translation error suggested that translation error could be a possible source of the variations. The extent of single molecule catalytic variation was reduced in a mutant with a hyperaccurate translation phenotype implying that translation error is a source of the heterogeneity. Streptomycin-induced translation error reduced average activity, but did not lead to an increase in catalytic heterogeneity. No relationship between translation error and electrophoretic heterogeneity was observed. A novel CE-LIF assay was developed for the continuous monitoring of the catalytic activity and electrophoretic mobility of individual β-galactosidase molecules. Thermally-induced catalytic fluctuations were observed suggesting that individual enzyme molecules were capable of conformational fluctuations that supported different catalytic rates.

Acknowledgements

I would like to thank my advisor Dr. Doug Craig for his support and guidance throughout my study period. His indulgence regarding my unanticipated experimental excursions made possible several very interesting paths of enquiry.

I would also like to thank the members of my advisory committee, Dr. Doug Goltz, Dr. Steven Pind, and Dr. John Wilkins for all their helpful suggestions and willingness to discuss in detail particular aspects of the work presented here. I also thank Dr. Edgar Arriaga for traveling to Winnipeg for my PhD oral defence.

In particular I appreciate the unrestricted access provided to me by the Biology Department at the University of Winnipeg to their laboratory preparatory areas and the unlimited use of their research equipment. I also want to thank Jens Franck and Etienne Leygue for their assistance in assembling DNA sequences.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

AEBSF	4-(2-aminoethyl)benzenesulfonylfluoride hydrochloride
ANOVA	analysis of variance
APD	avalanche photodiode
A_{pk}	area incubation product peak
A_{sd}	area of injected standard peak
Asn	asparagine
Asp	aspartate
ATCC	American Type Culture Collection
atm	atmosphere
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
C_a	concentration of analyte
CCD	charge-coupled device
CE-LIF	capillary electrophoresis laser-induced fluorescence
D	Aspartate
D	diffusion constant
Da	daltons
DDAO	7-hydroxy-9H-(1,3-dichloro-9,9-dimethyacridin-2-one)
DDAO-gal	9H-(1,3-dichloro-9,9-dimethyacridin-2-one-7-yl)-β-D- galactopyranoside
ddH_2O	double-distilled water
DEAE	diethylaminoethyl

DFP	diisopropylfluorophosphate
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
E	separation potential
e	electronic charge
E_i	injection potential
EDTA	ethylene dinitrilotetraacetate
EOF	electroosmotic flow
f	frictional coefficient
$f(\kappa r_p)$	Henry function
FAD	flavin adenine dinucleotide
FDG	fluorescein-β-D-digalactopyranoside
fL	femtolitre
fM	femtomolar
FMG	fluorescein-β-D-monogalactopyranoside
FMN	flavin mononucleotide
FRET	Förster resonance energy transfer
F_z	charge suppression factor
G	glycine
GFP	green fluorescent protein
Glu	glutamate
GS-6	Genescan polymer 6^{TM}
G6P	glucose 6-phosphate

G6PDH	Glucose 6-phosphate dehydrogenase
HEPES	*N*-2 Hydroxymethyl-piperazine-*N′*-2-ethanesulfonic acid
His	histidine
His_6-tag	hexahistidine tag
HPLC	high performance liquid chromatography
Hz	hertz
I	ionic strength
IEF	isoelectric focusing
I/O	input/ouput
IPTG	isopropyl-β-D-thiogalactoside
k_B	Boltzmann constant
K_m	Michaelis-Menten constant
KEGG	Kyoto encyclopedia of genes and genomes
L	length of capillary
LDH	lactate dehydrogenase
M	molecular weight
MALDI	matrix-assisted laser desorption/ionization
MDa	mega daltons
$MgCl_2$	magnesium chloride
MWCO	molecular weight cutoff
MΩ	megaohm
N	asparagine
N	theoretical plate number

N_A	Avogadro's number
N.A.	numerical aperture
NAD	nicotinamide adenine dinucleotide
Ni^+-NTA	nickel-nitrilotriacetic acid
oNPG	*ortho*-nitrophenyl-β-D-galactopyranoside
PCR	polymerase chain reaction
PE	pepstatin A
PEG	polyethylene glycol
pI	isoelectric point
PI	pentamidine isethionate
pL	picoliter
PMSF	phenylmethanesulfonylfluoride
PMT	photomultiplier tube
pNPG	*para*-nitrophenyl-β-D-galactopyranoside
PVP	polyvinylpyrrolidone
Q_a	quantity of analyte injected
rpm	revolutions per minute
r	radius of particle
r_b	average radius of buffer ions
r_c	radius of capillary
r_p	Stokes radius of protein
R_s	separation resolution between two analytes
RES-gal	resorufin-β-D-galactopyranoside

RNA	ribonucleic acid
S	axial ratio of protein
S_0	ground electronic state
S_1	first excited electronic state
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SiOH	silanol
SmFRET	single molecule FRET
S-30	30,000 x *g* supernatant fraction
T	temperature
t	analyte time to detector
t_{ass}	assay time
t_b	time required for baseline shift to reach detector
TBE	Tris-Borate-EDTA
t_{DDAO}	DDAO time to detector
t_i	injection time
t_{pk}	product peak time to detector for extended separations
t_{sd}	injected standard time to detector
TIM	triosephosphate isomerase
TIRF	total internal reflective fluorescence
TOF	time of flight
TLCK	N-α-p-tosyl-L-lysinechloromethylketone hydrochloride
TPCK	N-p-tosyl-L-phenylalaninechloromethylketone

t_{pk}	time required for product peak to reach detector
Trp	tryptophan
t_s	mobilization time for short separations
t_{sep}	total separation time for extended assay
Tyr	tyrosine
UV	ultraviolet light
V	volt
v	velocity of an ion in solution under an applied field
v_{act}	actual velocity of analyte
v/v	volume to volume ratio
V_c	capillary volume
v_{app}	apparent velocity of enzyme
v_{eof}	velocity of EOF
v_{rp}	velocity of enzyme relative to DDAO
w	peak width at half maximum
w_{pt}	plateau width for continuous assay
w/v	ratio of weight to volume
Z_a	actual valence
Z_c	calculated valence
Z_p	net charge on protein
ε_r	relative permittivity
ζ	zeta potential
η	viscosity

κ	inverse screening length
μ	electrophoretic mobility
μ_{eof}	electrophoretic mobility of EOF
μg	microgram
μL	microliter
μM	micromolar
μm	micrometer
μ_p	electrophoretic mobility of protein
ψ	asymmetry factor
Å	angstrom

List of Copyright Materials Used

Page 18: Figure 3. Cholesterol oxidase dynamic heterogeneity. From "Lu, H.P., Xun, L., Xie, X.S. (1998) Single-molecule enzymatic dynamics. Science, 282(5395): 1877-1882." Reprinted with permission from AAAS.

Page 19: Figure 4. Distribution of cholesterol oxidase waiting on-times. From "Lu, H.P., Xun, L., Xie, X.S. (1998) Single-molecule enzymatic dynamics. Science, 282(5395): 1877-1882." Reprinted with permission from AAAS.

Page 40: Figure 8. Ribbon representation of β-galactosidase. From "Juers, D.H., Jacobson, R.H., Wigley, D., Zhang, X.J., Huber, R.E., Tronrud, D.E., Matthews, B.W. (2000) High resolution refinement of β-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for α-complementation. *Protein Sci.*, ***9*:** 1685-1699." Used with permission.

Page 42: Figure 9. Schematic of reaction mechanism for hydrolysis of galactopyranoside by β-galactosidase. Reprinted with permission from "Juers, D.H., Heightman, T.D., Vasella, A., McCarter, J.D., Mackenzie, L., Withers, S.G. and Matthews, B.D. (2001). A structural view of the action of *Escherichia coli* (*lacZ*) β-galactosidase. *Biochemistry* **40**(49): 14781-14794". Copyright 2001 American Chemical Society.

1. Introduction

1.1. Background

Late in 1959 the acclaimed American physicist Richard Feynman gave a talk at the annual meeting of the American Physical Society entitled "There's plenty of room at the bottom" where he urged researchers to explore the possibilities of studying and manipulating physical systems at the smallest possible scales—atoms and molecules [1]. To spur others to take the first steps to realize his conviction that ultra small detection and manipulation was feasible, he provided two one thousand dollar prizes to the first practical demonstration of a rotating electrical motor with a size of 1/64 of a cubic inch, and the 25,000-fold reduction of the page of a book. The prizes were duly awarded in 1960 and 1986 respectively. His exhortations inspired a generation of researchers to devise the tools to study molecules at their fundamental level, and for this he is frequently regarded as the intellectual grandfather of single molecule studies.

1.2. Why study single molecules?

Prior to the advent of single molecule techniques, chemistry necessarily entailed making measurements on a large number, or ensembles, of molecules simultaneously, and thus the experimental results represented average values for all the components of the system. By measuring the properties of molecules one molecule at a time this ensemble averaging is eliminated, and any underlying complexity and heterogeneity becomes accessible to experimental interrogation. Single enzyme molecule experiments have revealed that this complexity comes in two varieties: complexity of participants, known as static heterogeneity, and complexity of processes, known as dynamic heterogeneity.

Static heterogeneity addresses the presumed homogeneity of structure of the individual molecules that comprise the different chemical species that participate in a chemical reaction. For purified small simple molecules this is not a significant concern, however, for large polymeric structures such as proteins, structural homogeneity cannot be assumed. Single molecule experiments have revealed the presence of large differences in the activity of different, nominally identical enzyme molecules, and that these differences are stable and long-lived. [2, 3] Such results imply that notwithstanding their nominal identity, they are indeed different in some unspecified way. Dynamic heterogeneity encompasses the sequence of events culminating in the output of interest, and allows for the detection of rare or short-lived intermediates which are obscured by ensemble measurements. The detection of such species in an ensemble measurement would require that all the participants be synchronized, whereas single molecule studies eliminate this necessity. Moreover, different molecules exhibit unique reaction trajectories that would quickly confound any synchronization initially established in an ensemble-based experiment. Examples of the complexity of processes for biological molecules include the magnitude and longevity of structural fluctuations of an enzyme molecule along a reaction pathway [4, 5]; the discrete steps of a motor protein [6], the forces required to unfold proteins [7]; and the dynamic fluctuations of a protein [8].

It is altogether appropriate that it be noted here that the insights available to single molecule studies should in no way be construed as a disparagement of conventional biochemistry and molecular biology. The merits of traditional approaches cannot be gainsaid, and are evident in the wealth of knowledge and achievements of these fields. The results from single molecule experiments must accord with the results obtained from

classical biochemistry and molecular biology. The role for the single molecule field is to augment current knowledge, and contribute to a fuller understanding of the processes that underlie chemical transformations. This is more likely to be attained by examining reacting systems at the smallest scale that retains all the properties of the system—the individual molecules. At least two thousand single molecule papers have been published during the past decade, and such is the growth of the field that it prompted a prominent single molecule researcher to jocularly assert that if current trends were projected forward, in 30 years all research would be about single molecules [9].

1.3. Overview of single molecule studies

The two main branches of contemporary single biomolecule studies are based upon fluorescence spectroscopy and a variety of techniques that measure the response of molecules to stretching and twisting forces. Single molecule fluorescence spectroscopy is the dominant branch, and for the study of biological macromolecules the fluorescent signal is provided either by fluorescent probes that are covalently linked to the macromolecule of interest that report on the position or status of the molecule, or by substrates that are transformed by the molecule, in the case of enzymes, into fluorescent products. Fluorescent spectroscopy underlies the research that is presented here.

Single molecule force-based studies utilize atomic force microscopy, optical tweezers and magnetic tweezers technologies. Some examples of the types of investigations that have been conducted include the measurement of forces generated by molecular motors [10], DNA mechanics [11], polysaccharide elasticity [12], and the generation of ATP by the application of mechanical force to F1-ATPase [13]. However, since the measurement of forces was not a part of the investigations here, and are only of

tangential relevance to questions this work was attempting to address, this branch of single molecule studies will not be considered further.

2. Literature review and experimental rationale

2.1. Fluorescence spectroscopy

Intensely fluorescent organic molecules are characterized by extensive conjugated π-electronic systems [14]. Accordingly, fluorescent molecules are generally polycyclic aromatic hydrocarbons, and are the basis for most fluorescent probes used in single molecule studies [15]. In solution, the fluorescence process begins with the absorption of a photon (Figure 1a) and drives a π bond electron from the ground electronic state (S_0) to the first excited electronic state (S_1) at a rate that is proportional to the intensity of the excitation [16]. A rapid ($< 10^{-14}$ s) non-radiative relaxation step (b) to the lowest excited singlet state is followed by a slower (several ns) (c) radiative (fluorescence) or non-radiative (heat) relaxation back to the ground electronic state.

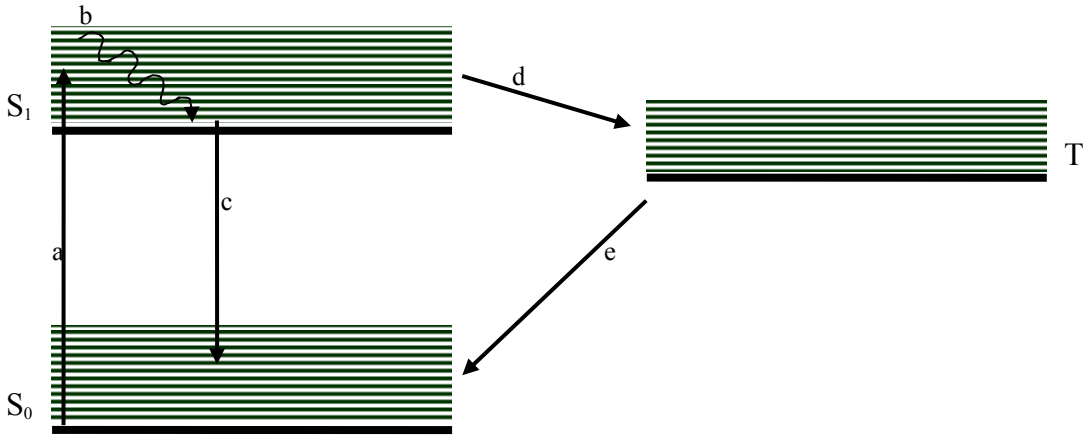


Figure 1. Schematic of energy levels for a fluorescent dye. S_0, S_1 and T are the ground, first excited and triplet states respectively; various horizontal lines represent vibrational and rotational levels. a) excitation b) internal relaxation c) decay d) inter-system crossing e) relaxation

At this point the molecule is ready for another cycle of absorption and decay. The fraction of decay events that are accompanied by photon emission is known as quantum yield, and the difference in the wavelength of the emitted and absorbed photons is the Stokes shift [14]. At saturation excitation intensities for molecules with a high quantum yield, greater than 10^6 photons can be emitted per second [15]. The amount of information that can be accessed from a single fluorophore is limited by quenching (non-radiative interactions with solvent molecules) and photobleaching (the photo chemical destruction of the dye that has a specific probability for a particular molecule for each absorption/emission cycle). Inter-system crossing (d) can occur where the excited S_1 state decays to the triplet state (T), and phosphorescence results if the decay back to the ground state is accompanied by photon emission.

2.2. Single molecule detection strategies

Single molecule detection requires surmounting formidable signal to noise challenges, and different methods for the measurement of single molecule fluorescence share common attributes. It is implicit that these strategies entail either maximizing the amount of signal that can be obtained from a specific fluorophore, and/or the minimization of signal from all other sources. The common features for most single molecule assay setups include the use of lasers for excitation; the minimization of background signal by physical and optical means; optimizing photon collection through high numerical aperture objectives; and maximizing their detection through the use of highly sensitive detectors.

2.2.1. Excitation intensity

The rate of photon emission, and hence the magnitude of the fluorescent signal, will depend on the rate of photon absorbance at the excitation wavelength—the more photons absorbed, the more photons emitted. The extremely high photon flux over a narrow wavelength range that characterizes laser output makes them an excellent excitation source, and their coherent output helps focus the beam to a tight spot. For these reasons their use for single molecule studies is ubiquitous. The intensity and the duration of the laser output can be tailored to minimize the effects of photobleaching if signal collection over an extended time frame is required for an individual fluorophore. Single biological macromolecule detection has been attained with alternate light sources, but is limited to setups where a large number of fluorescent product molecules produced by an enzyme molecule accumulate in sub-nL volume containers [17, 18].

2.2.2. Background signal minimization

The background signal arises from fluorescent impurities in the solvent, other fluorophores, and inelastic and elastic laser scatter from the solvent itself due to Raman and Rayleigh scattering. Impurities are reduced through the use of ultra pure water for all solutions, and by filtration with sub-micron pore size filters. The most common method to minimize background is to reduce the excitation volume to the smallest convenient size [19]. This is generally accomplished by the manipulation of optical geometry with prisms, mirrors and microscope objectives that focus the laser beam to a tight discrete spot. A second approach is to isolate the molecules of interest in ultrasmall containers using oil-dispersed droplets [20, 21], liposomes [22] and or by using fL microarrays [23] or modified fibre optic cable [24]. Spatial filters such as pinholes and slits placed

between the excitation chamber and the detector are used to reduce out of focus fluorescence from fluorophores outside the excitation volume [25]. Laser scatter is curtailed by the use of optical filters that block excitation light, but allow emission light to pass unimpeded.

One drawback of the small excitation volume is that the molecule of interest will diffuse out of the probe volume if it is not physically restricted, and a number of methods are used to confine the molecule within the probe volume for the duration of the experiment. The most common method entails tethering the molecule of interest to a microscope coverslip and then manoeuvring it and the attached molecule into the laser excitation volume, or by manoeuvring the excitation spot to the location of the affixed molecule. Typical tethering strategies exploit the high affinity of streptavidin for biotin [5] or use hexahistidine tags for Ni^{+}-NTA [26] immobilization. To prevent interactions between the molecule and the glass surface, the surface is usually passivated by BSA (bovine serum albumin) or with PEG (polyethylene glycol). Frequently a flexible linker such as PEG is used to increase the distance between the molecule and the surface and to promote unimpeded translational rotation [5]. Another commonly used confinement technique for single enzyme studies is to trap the molecule in a dilute agarose gel (1%) that allows substrate and product to freely diffuse [4]. The principal concern of confinement strategies is that they do not distort the properties that the experiment is attempting to measure, and most studies assert that this concern was assessed and determined to be of minimal effect.

2.2.3. Optical detectors

Since the signal produced by a single molecule is unavoidably minute, the use of highly sensitive detectors is paramount. High numerical aperture (N.A.) microscope objectives maximize photon collection efficiency and focus the signal towards the detector. There are two principle classes of optical detectors: point detectors such as photomultiplier tubes (PMT) and silicon avalanche photodiodes (APD), and wide-field detectors based upon charge-coupled devices (CCD) [27]. The point detectors are extremely sensitive and can have μs temporal resolution, but cannot differentiate between spatially distinct signals and therefore are limited to the examination of one molecule at a time. CCD-based devices are less sensitive with only ms temporal resolution, but can capture spatially distinct images simultaneously, or detect the spatial migration of a fluorescent analyte over time [27].

2.2.4. Confocal and total internal reflective fluorescence (TIRF) spectroscopy

To make the forthcoming discussion of various single molecule experiments easier to envision, a brief description will be provided at this point for two common experimental setups used to study single molecules. The technique used for the experiments that will be presented here, capillary electrophoresis with laser-induced fluorescence detection (CE-LIF) will be discussed in greater detail in a later section. The most commonly used optical geometries for single molecule studies of biological macromolecules are TIRF (total internal reflective fluorescence) and confocal microscopy. Figure 2 depicts the basic arrangement for the two different strategies where the molecule of interest is tethered to a coverslip in solution between the coverslip and a microscope slide.

For the TIRF arrangement (Figure 2a), laser light is directed by the prism at the microscope slide at an angle that exceeds the critical angle and is totally reflected. However an evanescent wave is propagated about 100 nm into the solution between the coverslip and the microscope slide that excites the fluorescent molecule(s) and emission is collected by a high numerical aperture microscope objective and passed through an optical filter to the detector. For confocal microscopy (Figure 2b), the excitation beam is reflected off a dichroic mirror and focused by an objective to the location of the molecule(s) of interest. Fluorescence is collected by the same high numerical aperture microscope objective. The dichroic mirror is opaque to wavelengths of the laser light, but transparent to the fluorescent wavelengths. This minimizes the laser light that reaches the detector, and the pinhole filter discriminates against florescence from molecules at the periphery of the excitation region.

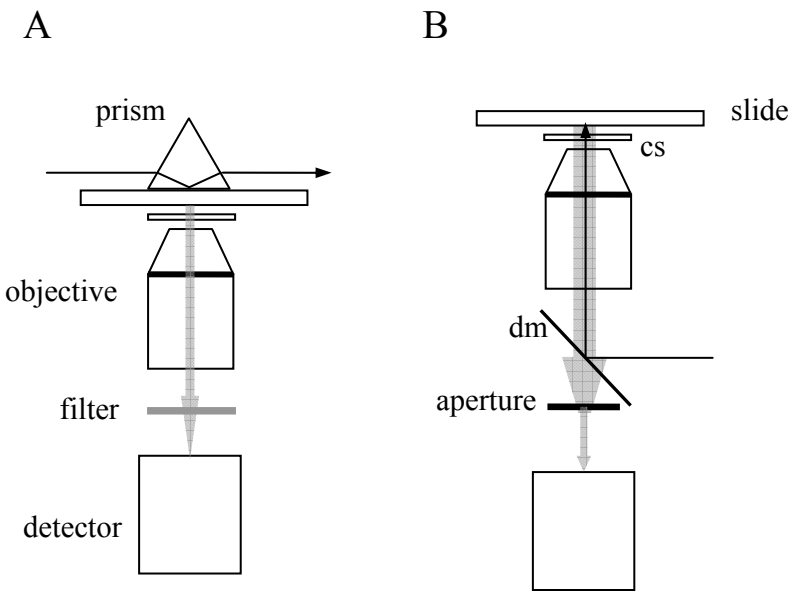


Figure 2. Schematic for TIRF and confocal microscopy. A) TIRF setup, B) Confocal setup. Laser beam depicted as narrow arrow-head line. Fluorescent emission depicted as semi-transparent grey arrow. Cover slip (cs), dichroic mirror (dm).

2.3. Single molecule studies

2.3.1. Historical background

In 1961 Rotman performed the first single molecule experiment based upon fluorescence microscopy and detected the activity of *E.coli* β-galactosidase by measuring the production of a fluorescent product that accumulated over a number of hours in fine droplets suspended in oil [17]. Fifteen years later Hirschfeld [28] detected the presence of individual γ-globulin molecules that had been labelled with 80 to 100 fluorescein molecules. Both of these experiments recorded the presence of individual molecules indirectly by detecting a large number of fluorescent molecules that were associated with them. The possibility for practical detection of individual fluorophore molecules at room temperature was predicted in 1984 [29], but was not achieved until the early 1990s when near-field scanning was used to image carbocyanine dye molecules embedded in a polymethylmethacrylate-coated coverslip [30]. The fluorescent lifetimes and dynamic properties of individual molecules of sulforhodamine 101 [31] and rhodamine 6G [32] were then measured by near-field scanning optical microscopy. This technique was extended to single molecule FRET [33] but has not gained wide acceptance [34]. The near-field scanning technique was considered suboptimal because the excitation source was an aluminium-coated tapered optical fibre that was positioned approximately 10 nm from the dye molecule which caused non-specific interactions between the probe and the dye [35]. Far-field techniques have since become dominant, and the first far-field detection of individual fluorophores at room temperature used confocal fluorescence microscopy to measure the influence of local environment on the fluorescent lifetime and emission maximum of individual rhodamine 6G molecules [16]. Almost immediately

upon the development of methods to detect single fluorophores it was recognized that single molecule fluorophores could be used to label and monitor biological macromolecules [36].

2.3.2. Studies of single biological macromolecules

Fluorescent labelling has been used to directly visualize motions of a variety of motor proteins, and an application of this technique, single molecule FRET (smFRET), has been used to examine conformational fluctuations associated with catalysis [37] and protein folding [38]. An analysis of the movement by myosin V labelled with rhodamine on immobilized actin using TIRF was used to demonstrate that the motor protein walked with a hand-over-hand rather than an inchworm mechanism along the filaments [6], and the same group found similar results for the reverse direction myosin VI [39]. TIRF was also used to measure the step size and rate of movement of kinesin and dynein by tracking the movement of GFP labelled peroxisomes in a live cell [40]. The rotational movement of the F_1-ATPase that is the motor component of the ubiquitous F_0F_1-ATPase responsible for coupled transmembrane proton flow to ATP synthesis/hydrolysis was directly visualized by an actin filament fluorescently labelled with tetramethyrhodamine that was linked to the central stalk γ-subunit through a streptavidin-biotin linkage [26]. The actin filament for every molecule was observed to rotate in a counter-clockwise direction (from the perspective of the membrane) in the presence of 2 mM ATP. The same group was subsequently able to demonstrate that a 120° rotation was associated with each ATP hydrolysis, and that this was comprised of two smaller steps corresponding to the binding and hydrolysis of ATP [41].

Single molecule FRET relies upon the non-radiative distance-based transfer between two fluorophores, and can be regarded as a continuously reporting nm scale ruler [34]. By judicious site-specific labelling, the distance between two molecules, or two regions of the same molecule, can be determined from the changing emission of the two fluorophores. SmFRET was used to find novel conformations along the reaction pathway for hairpin ribozyme [42], and was also used to obtained mechanistic information and conformation dynamics about T4 lysozyme [37]. A recent smFRET study of *E. coli* adenylate kinase revealed that the enzyme continuously fluctuated between an open and closed state where a lid-like structure moved over the active site even in the absence of substrate [43]. When substrate was present the equilibrium was shifted and favoured the closed conformation. The rate limiting step was identified as the opening of the lid to allow for product release.

The single molecule field where some of the earliest and most striking observations were made is that of single molecule enzymology. Although it overlaps with the forgoing discussion, it will be considered in more detail in a separate section because of its direct relevance to the work presented here.

2.3.3. Single molecule enzymology

2.3.3.1. First experiment:

The first single molecule enzymology report was that of the previously mentioned work of Rotman on the activity of *E.coli* β-galactosidase [17]. He sprayed a dilute solution of the enzyme and a fluorogenic substrate into a mist of fine droplets approximately 15 μm in diameter that were trapped in silicon oil on a microscope slide. After several hours the droplets were visualized by fluorescent microscopy and those that

contained enzyme molecules were filled with fluorescent product. By applying Poisson statistics to the ratio of droplets that contained enzyme to those that did not, he was able to calculate the concentration of the enzyme and make an estimation of its molecular weight that was nearer to the actual value than contemporary estimates. He also used this approach to address the question of whether the residual activity and high k_m values measured by spectrophotometric studies for a mutant enzyme were attributable to the presence of a small number of copies of wild type enzyme molecules contaminating a sample of otherwise inactive enzyme molecules. He found that the mutant enzyme did indeed have very low activity. Notwithstanding the innovative nature of this work, the single enzymology field was to remain dormant for 30 years.

2.3.3.2. Static Heterogeneity:

The next unambiguous report on the activity of a single enzyme molecule was presented in 1995 when Xue and Yeung recorded the activity of individual lactate dehydrogenase molecules (LDH) by incubating the enzyme in the presence of lactate and NAD^+ for one hour inside a capillary. The catalytic rate of the individual enzyme molecules was determined by measuring the production of the inherently fluorescent NADH using a laser-induced fluorescence detection system [2]. The product zone was electrophoretically driven past an argon ion laser based detection system to determine how much NADH had accumulated. They found that different molecules had different activity, and a four-fold range of activities was observed for the 79 molecules assayed. Double incubations of individual molecules revealed that the different activities were reproducible. This same group later showed that single molecule assays with LDH in fL microarray wells constructed by photolithography of fused silica produced a similar

range of activity as was seen for their earlier CE experiments [44]. They also demonstrated that if a single metal ion catalyst was used rather than a protein, no rate heterogeneity was observed. This affirmed their conclusions that variable reaction rates were catalyst dependent and not an artefact of the experimental setup. Such variation of catalytic activities between nominally identical single molecules is known as *static heterogeneity*.

The observation that different individual enzyme molecules were catalytically distinct was replicated shortly afterward by the Dovichi group, also using CE-LIF [3]. This group used a post-column sheath flow detection system to measure the synthetic and fluorescent hydrolysis product AttoFluor^{TM} produced by the activity of bovine alkaline phosphatase. The highly sensitive detection system enabled them to reduce assay times to as brief as one minute. A 10 fold range of activities was observed, and replicate incubations of the same molecule revealed that activity was independent of assay duration. The activation energy of catalysis was measured by incubating the same molecule at different temperatures and was shown to be heterogeneous between molecules. No correlation was observed between activation energy and catalytic rate; this prompted the speculation that ligand binding may have dominated catalytic rate. A decade later, Hsin and Yeung [22] fused liposomes that separately contained bovine alkaline phosphatase and the substrate fluoroscein-diphosphate and found 20 fold static heterogeneity between the individual enzyme molecules. The demonstration that different assays types yielded similar ranges of static catalytic heterogeneity for the same type of molecule provides further support that the phenomenon is not an experimental artefact.

The β-galactosidase from *E. coli* is the most extensively studied enzyme at the single molecule level because of its high catalytic rate and the commercial availability of high quality fluorogenic substrates. The Dovichi group made the first measurements using contemporary single molecule techniques by monitoring the hydrolysis of the synthetic fluorogenic substrate RES-gal with CE-LIF, and demonstrated that prokaryotic enzymes also exhibit static heterogeneity for single molecule catalytic activity [45]. The activity of individual β-galactosidase molecules obtained from enzyme crystals was measured using the same substrate, and the activity of individual molecules was reproducible to within 15% [46]. It was also shown that enzyme from the crystals had the same level of catalytic heterogeneity as enzyme that had never been crystallized. The enzyme has been found to be heterogeneous at the single molecule level with respect to a requirement for magnesium [47], and it has been demonstrated that the average activity of individual enzyme molecules is source dependent [48]. Rissin *et al.* recently presented simultaneous monitoring of multiple individual *E. coli* β-galactosidase using a high density fL array constructed from modified optical fibres [18]. Using RES-gal as substrate, they affirmed the static heterogeneity of the enzyme, and found a 15 fold range of individual enzyme activities at different substrate concentrations. The net accumulation of product was measured every 15 s and they observed that the different activities of the different individual molecules were stable for the duration of the 3 minute assay. Their analysis of the distribution of activities at different substrate concentrations led to a conclusion that differences in activity between individual molecules were attributable to differences in k_{cat}. The activity of single molecules of β-galactosidase has been reported for the substrate FDG, but the objective of the paper was

to demonstrate the compatibility of a microarray comprised of fL sized containers constructed from a silicone elastomer with single molecule enzyme detection [23]. There was no effort to use the device to investigate the properties of the enzyme *per se*. Recently, the direct detection of individual molecules of β-galactosidase by tryptophan fluorescence was achieved using ultraviolet confocal microscopy [49].

Other single molecule enzyme studies have revealed that static heterogeneity is a common feature of samples of individual molecules. The activities of alkaline phosphatase from *E. coli* [50] and *Thermus thermophilus* [51] have been detected at the single molecule level by CE-LIF, with a 12 fold range of activities measured for the latter. A dihydroorotate dehydrogenase from *E.coli*, with a tyrosine to leucine substitution necessary to facilitate the detection of its fluorescent FMN cofactor, was also shown to exhibit static heterogeneity [52]. A haloperoxidase from *Curvularia verruculosa* trapped in a dilute agarose gel was monitored by confocal microscopy and a 6 fold range of activities was observed for different molecules [53]. Additionally, although not a fluorescence-based study, λ-exonuclease has 4 fold catalytic heterogeneity for the digestion of phage DNA [54].

2.3.3.3. Dynamic heterogeneity:

A different single molecule enzyme assay strategy was introduced by the Xie group using the cholesterol oxidase from *Brevibacterium sp.* Rather than measuring the amount of fluorescent product produced by an enzyme molecule over an incubation period that lasted many minutes, they monitored individual catalytic turnovers from a single enzyme molecule [4]. Chemical transformations occur on the ps time scale and cannot be detected with current technology; instead they were able to exploit the

fluorescent properties of the enzyme's flavin adenine dinucleotide (FAD) co-factor for evidence that a transformation had occurred. The co-factor is fluorescent only when it is in the oxidized state, and by following the on/off status of the cofactor of a single enzyme trapped in an agarose gel they were able to indirectly observe individual turnovers of the enzyme (Figure 3). The waiting time for an enzyme spent in its "off" or "on" state was found to be highly variable or stochastic.

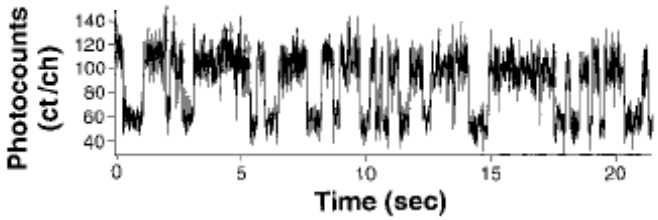


Figure 3. Cholesterol oxidase dynamic heterogeneity. Sample trace of fluorescent emission of cholesterol oxidase NAD cofactor showing fluctuations between the 'on/off' states. From "Lu, H.P., Xun, L., Xie, X.S. (1998) Single-molecule enzymatic dynamics. *Science,* **282**(5395): 1877-1882." Reprinted with permission from AAAS.

The fluorescent trace data was used to evaluate the relationship between waiting times for successive turnovers, and a memory effect was found. In other words, a short waiting time was more like to be followed by another short waiting time; and similarly, a long waiting time was more likely to be followed by a long waiting time. A histogram (Figure 4) of the frequency of waiting times and their duration was assembled from the fluorescent trace data and from this information, when considered in conjunction with the memory effect, led the authors to conclude that the enzyme was migrating through a series of different conformations with different catalytic rates. This stochastic fluctuation of turnovers is known as *dynamic heterogeneity*.

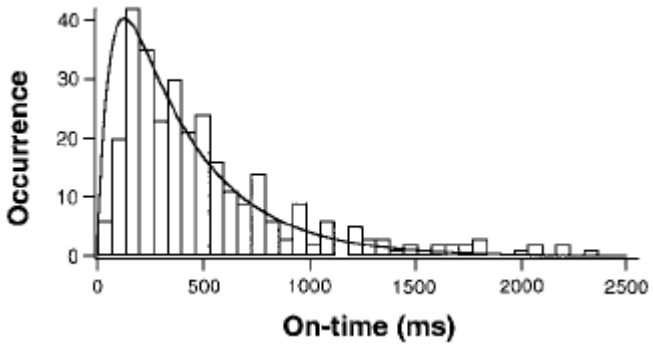


Figure 4. Distribution of cholesterol oxidase waiting on-times. From "Lu, H.P., Xun, L., Xie, X.S. (1998) Single-molecule enzymatic dynamics. *Science,* **282**(5395): 1877-1882." Reprinted with permission from AAAS.

In 2006 the same group examined the dynamic heterogeneity of *E. coli* β-galactosidase [5]. In this experiment the enzyme was tethered with a biotinylated flexible cysteine-reactive PEG linker molecule to a streptavidin coated bead that in turn was biotinylated to a microscope slide. A trajectory for the waiting time between successive turnovers was generated by detecting the production of each fluorescent resorufin product molecule produced by the enzyme from the fluorogenic RES-gal substrate using confocal microscopy. Applying the same type of analysis as was used for cholesterol oxidase, they demonstrated that the results were consistent with a model where the enzyme interconverts between different conformations with different catalytic rates, and it was estimated that more than ten interconverting conformers were required to explain the observations. The authors speculated that a dynamic network of hydrogen bonds could be responsible for the conformational fluctuations.

Since the molecule was tethered to a microscope slide, the experiment was performed at different substrate concentrations to examine the relevance of the Michaelis-Menten equation, which describes the hyperbolic relationship between substrate concentration and catalytic rate for ensemble enzyme kinetics, to single

molecule studies. At the single molecule level, the variables from the Michaelis-Menten relationship for the concentration of the enzyme and the concentration of the enzyme-substrate complex are meaningless because there is only one enzyme molecule. Instead, probability of the enzyme existing in either state is used. It was demonstrated that the inverse of the average waiting time (average because the enzyme fluctuates through conformations with different waiting times), and hence the overall rate for a particular molecule, followed the same hyperbolic relationship to substrate concentration as bulk measurements.

Dynamic heterogeneity has been observed by other research groups for molecules where individual turnover events can be detected. The conversion of individual fluorogenic dihydrorhodamine 6G to the fluorescent rhodamine 6G by a biotinylated horseradish peroxidase tethered to a coverslip coated with streptavidin was monitored with confocal microscopy [55]. The enzyme was observed to fluctuate in activity over a wide range of time scales, and the authors concluded that the enzyme had a large number of slightly different thermodynamically accessible conformations that were catalytically distinct. When the activity of this enzyme was reexamined using a TIRF setup and a different rhodamine-based substrate, dynamic heterogeneity was observed again, but the rate constants were an order of magnitude higher than the earlier report [56]. The authors attributed the differences to substrate hydrophobicity, charge density, and size. Other enzymes for which dynamic disorder has been discerned include hairpin ribozyme [42], an *E. coli* flavin reductase [57], and lipase B from *Candida antarctica* [58, 59]. It appears, therefore, that dynamic heterogeneity, as was seen for static heterogeneity, is common to many different types of enzymes.

2.3.3.4. Relationship between static and dynamic heterogeneity:

The review above outlined some of the experimental evidence for the presence of static and dynamic heterogeneity respecting the activity of individual enzyme molecules. Generally, studies have focused upon one form of catalytic heterogeneity to the exclusion of the other. In part, this occurs because the analytical technique used was not amenable to an examination of both types, and as a result the current understanding of the relationship between the two forms of heterogeneity is deficient. An exception to this bifurcation is found in the report by Lu *et al.*, where they reported that if a cholesterol analogue at saturating conditions was used as a substrate for their cholesterol oxidase experiments, five-fold static heterogeneity was present for the enzyme substrate complex to enzyme plus product step [4]. It was not possible to determine if this was a substrate analogue effect because the reaction proceeded too rapidly to discern individual turnovers if cholesterol was used. Edman *et al*. suggested that static and dynamic heterogeneity were different manifestations of the same phenomena, and that one implied the other [55]. The basis for this assertion rested on their result that the rate constant for different molecules converged for longer assays, and upon the ergodicity that is implicit in statistical mechanics. The former was based on only two separate runs so its significance for the justification of a trend may be overstated. Ergodicity means that the average value for a physical property should be the same regardless if it is measured on one molecule over time, or for a large number of molecules simultaneously [19]. Thus their apparent meaning is that static heterogeneity is observed because the limited time dimension of single molecule experiments samples only a portion of the different available conformational lifetimes and their associated catalytic rates. This interpretation

was recently reiterated, and the suggestion was made that the time required for convergence of the average activities of different molecules was 10^3 to 10^4 greater than the average turnover rate [60]. If this assertion is indeed correct, the significance of static heterogeneity requires reexamination. Regardless, at the time scales of relevance to a cell, static heterogeneity loses none of its significance [22].

Notwithstanding the caveat just made for the abiding significance of static heterogeneity for a cell, the contention that all the molecules are the same in the end requires closer scrutiny. Despite the dearth of experiments that explicitly test this prediction, sufficient information is available in the literature to evaluate its soundness. The only molecule for which extensive data pertaining to both types of heterogeneity is available is for the *E. coli* β-galactosidase. Data presented by Min *et al.* [61] from the same set of experiments published by English *et al.* [5], suggests that the average turnover rate for a long trace of β-galactosidase using RES-gal as a substrate was approximately 200 s^{-1}. Based upon the contention that convergence should appear after 10^4 turnovers, one would expect static heterogeneity to diminish as assay duration approached should 50 s. However, 15 minute CE-LIF assays of β-galactosidase using the same substrate indicate that static heterogeneity remains greater than 10-fold [45-48], and implies that static heterogeneity may have a more substantive basis than the merely snapshot effect suggested by some authors. This leads to the question of what is the basis for single enzyme molecule static heterogeneity.

2.3.3.5. Sources of static heterogeneity:

Chemical dogma requires differences in activity to reflect differences in structure. Despite a decade of single molecule enzymology investigations, the source(s) of static

catalytic heterogeneity remain an area largely dominated by speculation. Xue and Yeung suggested that the differences of activity for LDH were attributable to the enzyme adopting stable catalytically distinct conformations [2], whereas Craig *et al.* viewed post-translational modification, such as glycosylation, as a potential source for bovine alkaline phosphatase [3]. Polakowski *et al.* reported that purified *E. coli* alkaline phosphatase stored in the absence of protease inhibitors evinced greater catalytic heterogeneity than enzyme that was stored in their presence, and this led the authors to suggest that limited proteolysis may contribute to single molecule catalytic heterogeneity [50]. Whatever the structural differences are, they are likely to be subtle, as the extent of heterogeneity for redissolved β-galactosidase from crystals was indistinguishable from the original source used to make them [46]. Such a result is not surprising, as small changes of orientation of catalytically important residues or cofactors can have significant impact on activity [62].

2.4. Rationale and Objectives

The foregoing review indicates that individual enzyme molecules of many types are catalytically distinct, and there is evidence that static heterogeneity is also present for non-enzyme proteins [63], suggesting that it may a widespread phenomenon. A decade of research has provided many iterations of protein enzyme heterogeneity while leaving a number of significant issues unresolved. First, the structural basis for catalytic heterogeneity lacks an experimental foundation and is dominated by conjecture. Second, the gulf between static and dynamic heterogeneity persists, and how they relate to each other remains unclear. Third, the significance of heterogeneity at the single molecule level for cellular processes has not been established.

The primary objective of the work presented here was to test a number of hypotheses including limited proteolysis and cell-directed processes as possible sources of static catalytic heterogeneity using the β-galactosidase from *E. coli* as a model enzyme and CE-LIF as the principal analytical tool. A secondary objective was to develop new single molecule CE-LIF assays for β-galactosidase with the hope that new tools would yield novel experiments that might provide additional insights for this field. Due to their significance for this project, a more detailed review of CE-LIF and *E. coli* β-galactosidase will be provided in chapter three prior to a description of materials and methods in chapter four. The results chapters, (5-10), are presented roughly in the order that experiments were performed.

Chapter five examines the role of limited proteolysis as a source of single molecule β-galactosidase catalytic heterogeneity; chapter six compares the single molecule activity of β-galactosidase from *in vitro* and *in vivo* sources for the purpose of making inferences about the cellular processes that generate single molecule catalytic heterogeneity; chapter seven presents the development of two new single molecule β-galactosidase assays for CE-LIF; chapter eight demonstrates that it is possible to measure the electrophoretic mobility of individual enzyme molecules, and utilizes this information and a simple theoretical expression for protein electrophoretic mobility to make inferences about the structural differences that may underlie catalytic heterogeneity; chapter nine presents a novel single molecule continuous assay that is used to simultaneously follow static and dynamic heterogeneity and measure the influence of temperature for activity changes; and chapter ten investigates the role of translation error as a source of catalytic and electrophoretic heterogeneity, and reexamines the structural

differences required to account for single molecule electrophoretic heterogeneity. Chapter eleven presents a brief summarizing discussion of the results, and will include a few additional comments regarding static and dynamic heterogeneity; chapter twelve considers some additional CE-LIF experiments that could be performed to address questions still outstanding.

3. CE-LIF and *E. coli* β-galactosidase

3.1. CE-LIF

3.1.1. Background

Electrophoresis in narrow bore, open-ended glass tubes, known as capillary zone electrophoresis, was pioneered in the 1970s to redress free zone analyte band broadening associated with electrophoresis in the absence of solid supports, and to address the cumbersomeness and time requirements of gel-based electrophoretic separations [64]. At the most reductive level, the technique entails the immersion of both ends of a buffer-filled capillary into small vials of electrolyte solution with the interior of the capillary functioning as a single channel separation chamber. The system is connected to a high voltage power source, and routine separations are performed at potentials that exceed 100 Vcm^{-1}. The capillaries are constructed from fused silica and are externally coated with a thin polymer (often polyimide) to enhance flexibility and durability. Typical CE setups utilize capillaries with the diameter of the inner bore ranging from 10 to 75 μm, and lengths of 30 to 100 cm. Capillaries with an inner bore diameter as small as 2 μm are now available.

An early demonstration of CE used polytetrafluoroethylene rather than glass tubes as a separation chamber [65] and showed that it was possible to rapidly separate small anions. Jorgenson *et al.* demonstrated during the early 1980's that CE using drawn glass tubes was capable of extremely high separation efficiencies with short analysis times for dansyl derivatized amino acids [66], and they achieved highly efficient protein separations shortly afterwards. Nearly one million theoretical plates were achieved for a

mixture of peptides produced by tryptic digests, and it was shown that a complex mixture of fluorescently labeled components could be easily resolved [67].

Analyte detection is performed either on-column, where a small section of the capillary coat is removed at an appropriate distance from the injection end, and the passage of analytes is recorded, or post-column, as the analytes exit the capillary. On-column detectors are based upon UV absorption [67, 68], or fluorescence [66, 69]. Typically, post-column detection is used for electrochemical [70, 71], and mass spectrometry [72] based detection of analytes.

The introduction of the post-column sheath flow cuvette with laser-induced fluorescent excitation as a post-column detection cell greatly enhanced detection of fluorescent analytes [73], and its conjunction with CE [25] has lowered the limit of detection of analyte to the yoctomole level [74, 75]. As was discussed in the introduction to static heterogeneity, CE-LIF has been used to measure the quantity of product produced by single molecules of *E. coli* β-galactosidase and bovine alkaline phosphatase [45, 3]. CE-LIF is also an established analytic tool for protein studies, DNA sequencing [76], organelle heterogeneity [77], assaying of enzyme activity from individual cells [78], and the heterogeneity of protein expression in a cancer cell line [79].

When capillary electrophoresis is used as the platform for single molecule enzymatic assays, the reaction takes place in the capillary, and a zone of product molecules accumulate around the enzyme molecule. Quantification of the reaction product occurs subsequent to the incubation period as the product zone passes the detector. Accordingly, the usefulness of capillary electrophoresis for single molecule enzyme assays is highly dependent upon a minimization of distortion to the product zone.

The high separation efficiency capabilities of CE ensure that discrete regions of product molecules formed by an enzymatic reaction can be attributed to a particular molecule, and increases the concentration of molecules available for detection.

3.1.2. Mobility of a charged particle under an applied field

The application of a potential difference between two electrodes generates a potential gradient (E) which is the applied voltage (V) divided by the distance between the two electrodes. A charged particle in the presence of an applied field will accelerate under a force equal to the product of the potential E and Z_a, where Z_a is the net charge on the particle, until the frictional force (f) retarding the forward movement of the particle in the separating medium equals the driving force, and a steady state velocity (v) is achieved:

$$v = \frac{EZ_a}{f} \tag{3.1}$$

The frictional coefficient reflects the size and shape of the charged particle and is described by Stokes expression as:

$$f = 6r\pi\eta \tag{3.2}$$

where r is the radius of the charged particle, and η is the viscosity of the separating medium [80]. The migration of a charged particle is normally expressed as the ratio of its velocity to the electric field, and is termed electrophoretic mobility (μ) with units of $cm^2V^{-1}s^{-1}$. It follows therefore that electrophoretic mobility is:

$$\mu = \frac{Z_a}{6r\pi\eta} \tag{3.3}$$

The applicability of this expression for predicting the electrophoretic mobility of actual analytes is limited because it does not account for the interaction of the charged particle and solvent ions, and for irregularities of particle shape [80].

3.1.3. Separation efficiency

The chromatographic separation efficiency of analytes is described as a theoretical plate number which represents a sequence of notional equilibration zones. A higher number of plates signifies a superior separation capability. If diffusion is the only source of broadening as an analyte migrates, the theoretical number of plates (N) is governed by diffusion, and it can be shown to be proportional to the applied potential and is independent of the capillary length or separation time:

$$N = \frac{\mu V}{2D} \tag{3.4}$$

where D is the diffusion constant for the analyte and the other terms are as before [66]. The number of theoretical plates for an analyte under any particular separating conditions is calculated as

$$N = 5.54\left[\frac{t}{w}\right]^2 \tag{3.5}$$

where t is the time to the detector and w is the full peak width at half-maximum. In practice, however, electrophoretic separations will be strongly affected by convective currents that arise from the heat that is generated from passing a current through an electrolyte solution. This current-induced heat is known as Joule heating. The disruptive effects of convection can be minimized through the use of supports such as gels, or by effective dispersal of the heat when free zone electrophoresis is the mode of separation.

The capillaries used for CE typically have an inner diameter of 10 to 50 μm, and an outer diameter of 150 to 360 μm, and consequently their large surface area to volume ratios provide excellent heat dispersing properties [67]. This allows large potentials to be applied, and as equation 3.4 indicates, will lead to an increase in the number of theoretical plates. In a capillary there will be a small radial heat gradient which will affect electrophoretic mobility because of its influence upon viscosity, but because the channel is so narrow, and analytes are well-mixed radially, this effect is of minimal concern [67].

3.1.4. Electroosmotic flow

Electroosmotic flow (EOF) is the bulk flow of solvent within the capillary due to the application of the potential. The silanol groups (SiOH), which are present at the water-glass interface on the inner surface of the capillary, begin to ionize when the pH of the buffer inside the capillary is above two, and this causes the formation of a dense layer of negative charge at the interface of the capillary wall and the buffer [81] (Figure 5). As a result, a double layer of positively charged buffer counter ions accumulate along the interior of the capillary. This double layer is comprised of a static layer adjoining the wall, and a secondary diffuse layer varying in thickness depending upon the ionic strength of the solution. At an ionic strength of 10 mM, the double layer will be approximately 3 nm thick [82].

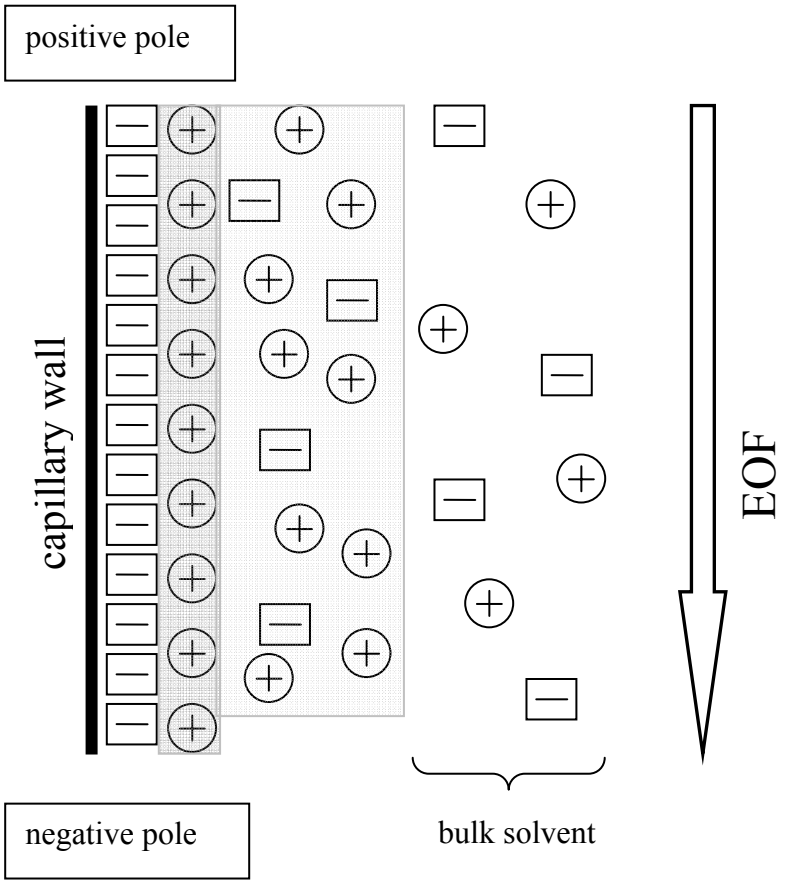


Figure 5. Capillary charge profile. Darker semi-transparent region is the static layer. Lighter semi-transparent region is the diffuse layer.

When the voltage potential is applied, the positive counter ions in the secondary diffuse layer and their associated solvation shells migrate towards the negatively charged pole. This causes a bulk flow of buffer (the EOF) away from the positive electrode and towards the negative pole. The flow profile of the bulk mobility is flat, or plug-like, if the radius of the capillary is greater than seven times the thickness of the double layer [82]. Figure 6 depicts a cartoon for the difference between laminar and plug-like flow. A flat flow profile minimizes band broadening and contributes to the high separation efficiencies of CE.

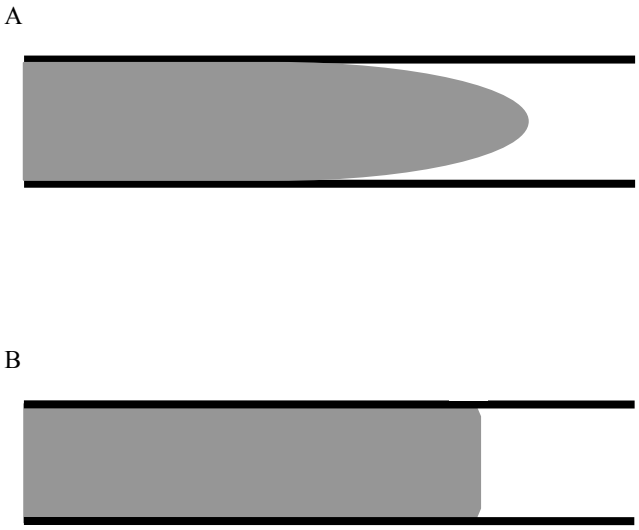


Figure 6. Laminar and plug-like flow profiles of solvent in a capillary. A) Laminar flow B) plug-like flow.

Laminar flow is characteristic of HPLC and other forms of liquid column chromatography and contributes to band broadening.

The linear velocity of the EOF (v_{eof}) is determined by:

$$v_{eof} = \frac{\varepsilon_r}{4\pi\eta} E\zeta \tag{3.6}$$

where ε_r is the relative permittivity of the buffer, ζ is the zeta potential, and the other terms are as before. Zeta potential is the potential difference between the slipping plane (the region that separates static layer from the mobile regions of the diffuse layer) and the bulk solvent. The net migration, or apparent velocity, of any particle will be the sum of the velocity of EOF and the actual velocity due to its electrophoretic properties:

$$v_{app} = v_{act} + v_{eof} \tag{3.7}$$

Under normal polarity conditions the detection end of the capillary is held at a negative potential and the EOF is directed towards the detector. At neutral pH, the magnitude of

the EOF is sufficient to overcome the electrophoretic migration of anions towards the injection end of the capillary, and all types of charged species can be migrated past the detector. The resolution between two different analytes (R_s) is given by:

$$R_s = 0.177(\mu_1 - \mu_2)\left[\frac{\mathrm{V}}{D(\overline{\mu} + \mu_{\mathrm{eof}})}\right]^{1/2} \tag{3.8}$$

where μ_1 and μ_2 are the electrophoretic mobilities of the two analytes [66]. This implies that maximal resolution between two analytes can be achieved when EOF is in the opposite direction and approaches their mean mobility.

The separation efficiency of CE is diminished by a number of well established processes and includes the width of the injection zone, conductivity differences between the injection zone and separating buffer, interactions between analyte and the capillary wall, and hydrodynamic flow [83]. The possible significance of these influences upon the experiments here will be discussed in chapter 8 that examines the electrophoretic mobility of individual β-galactosidase molecules.

The high charge density on the interior surface of the capillary is not always desirable and can be minimized when necessary by shielding the surface with a variety of polymers. Capillary coatings are commonly required for the separation of basic proteins because their positive charge leads to interaction with the negative surface of the capillary wall and results in band broadening or permanent retention on the column. Capillary coats are also used when EOF suppression is required to perform an assay. Capillaries can be coated permanently with polyacrylamide or polyoxyethylene, but these methods can be difficult, expensive and time consuming to prepare [84]. Also, these viscous polymers do not work well with the narrow bore capillaries used for single molecule analysis because the coating process results in the capillary becoming plugged.

Alternatively, capillaries can be dynamically coated by flushing the capillary with a solution containing the coating polymer. Phospholipid bilayers [85], polyvinylpyrrolidone (PVP) [86] and variety of proprietary polymers [87] have been used for dynamic coating purposes. The appeal of dynamic coating is the ease of their application and that they are relatively inexpensive; their principal drawback is variable coating effectiveness and degradation which causes inconsistent EOF. The development of capillary coating technologies continues to be an area of extensive research efforts.

3.1.5. CE-LIF setup

The high sensitivity of CE-LIF with a sheath flow cuvette is attained using the same strategies discussed earlier for single molecule detection. These include high fluorescent emission through the use of laser excitation of high quantum yield fluorophores; reduction of excitation volume; efficient photon collection by high numerical aperture microscope objectives; optical and spatial filtering to minimize background signal, and highly sensitive photon detectors.

A schematic of the CE-LIF set up used for the experiments described in later chapters is depicted in Figure 7.

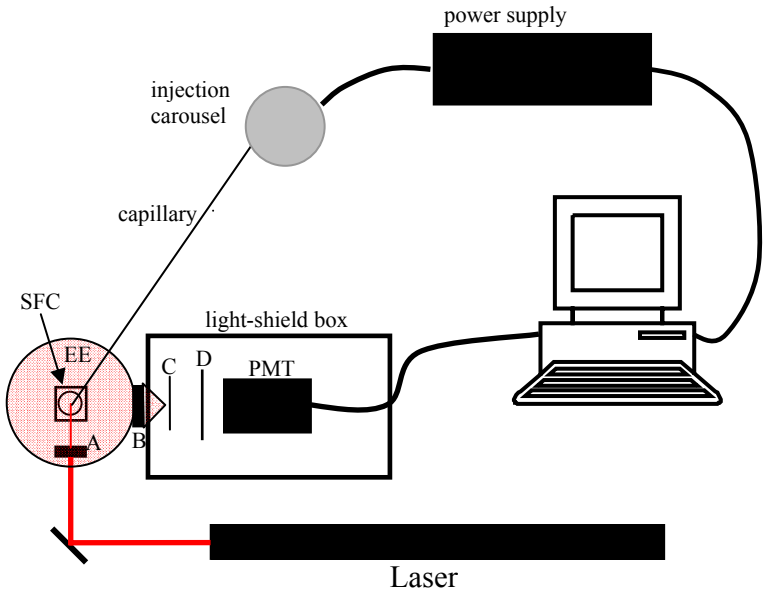


Figure 7. Schematic drawing for CE-LIF instrument. Lightly shaded circle (lower left of diagram) represents fluorescent emission. SFC sheath flow cuvette; EE excitation emission; A- 6.3X 0.2 N.A. objective; B- 60X 0.7 N.A. objective; C-optical filter, D-pinhole filter.

The injection end of the capillary is placed into a buffer containing vial in an injection carousel equipped with a platinum wire connected to a high voltage power supply. A safety lock is used to ensure that the circuit is broken when the platinum electrode is not immersed in the sample vial. The system is grounded through the sheath flow buffer and the cuvette holder at the detection end of the capillary.

Sample can be introduced into the capillary by electrokinetic or pressure injection. Electrokinetic injection is performed by inserting the injection end of the capillary into the sample and applying a potential that causes analytes to migrate into the capillary forming a small plug of sample at the injection end. Following the injection the capillary is returned to a buffer-containing vial. The quantity of analyte injected (Q_a) can be calculated by:

$$Q_a = \mu_a E_i t_i \pi r_c^2 C_a \tag{3.9}$$

where μ_a is the net electrophoretic mobility of the analyte, t_i and E_i are the electrokinetic injection time and potential respectively, r_c is the internal radius of the capillary and C_a is the concentration of the analyte [88]. Pressure injection is performed by inserting the injection end of the capillary into a syringe filled with sample where a column of compressed air between the sample and plunger forces sample into the capillary. This type of pressure injection is used to completely fill the capillary with sample, and is used because it is either faster than electrokinetic injection for filling of the capillary, or because the electrophoretic properties of a component of the sample are not amenable to electrokinetic injection. The latter scenario occurs when EOF is oriented towards the injection end of the capillary and a component of the sample has a positive charge or is neutral.

The detection end of the capillary is inserted approximately 1 cm into a sheath flow cuvette that has been adopted from the flow cytometry field and is constructed from quartz with good optical quality. The cuvette has an interior chamber with dimensions of 250 μm x 250 μm. The capillary and cuvette are fastened to a stainless steel cuvette holder that is bolted to a translation stage for precise *x y z* axis positioning. The cuvette and cuvette holder are encased in a retractable light shield to block ambient light. The light shield has a small opening for entry and escape of the excitation beam. Sheath buffer is the same as the separation buffer, and is introduced into the top of the cuvette and surrounds the detection end of the capillary in laminar flow. Laminar flow hydrodynamically focuses the emerging analyte and minimizes any post-capillary sample broadening [25]. The sheath flow buffer and capillary eluent are collected in a waste container. The laser beam is focused through a microscope objective to a spot

approximately 10 μm below the capillary in the emerging sample stream. Laser scatter (which increases background signal) is minimized by the optical quality of the cuvette; the laminar flow of sheath buffer which has the same refractive index as the sample stream; by collection of fluorescent emission at 90° to the excitation beam; and by passing the collected light through a bandpasss filter that is more transparent to emission but far less so for excitation light. A high numerical aperture microscope objective is used to maximize the collection of emission photons, and is passed through a ~ 100 μm pinhole aperture set to limit the field of view of the detector to the excitation volume [89]. Fluorescence is collected by a PMT at 10 Hz, and the PMT output is digitized by an I/O board of a desktop computer which provides a graphical representation of the change in fluorescence over time known as an electropherogram.

3.2. *E. coli* β-galactosidase

3.2.1. Background

The β-galactosidase from *E. coli* (E.C. 3.2.1.23) encoded by the *lacZ* gene is a 465 kDa homotetramer retaining glycosidase that functions to hydrolyze the disaccharide lactose to glucose and galactose [90].

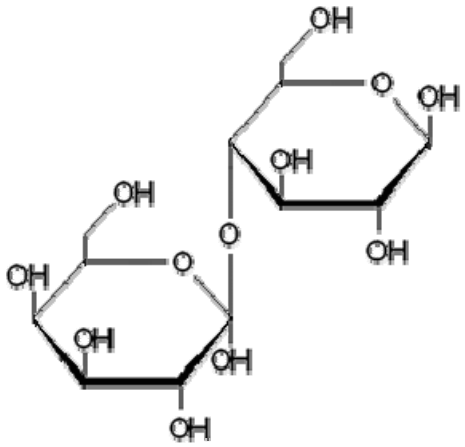


Figure 8. Structure of lactose

β-Galactosidase is specific for galactose, but not for the glucose, or aglycon, part of the substrate, and as a result a wide variety of synthetic aglycons have been substituted for the glucose moiety with chromo or fluorogenic properties [91]. The enzyme also catalyzes the transglycosylation of the (β1→4) linkage between galactose and glucose to a (β1→6) linkage to form allolactose. Allolactose is the natural inducer of the enzyme [92] and can be substituted for induction purposes with non-hydrolysable analogues such as isopropyl-β-D-thiogalactoside (IPTG). The study of the increase in β-galactosidase activity in response to the presence of lactose led to the formulation of the operon model for gene regulation by Jacob and Monod [93]. The *lac* operon is comprised of three genes: *lacZ*, *lacY*, which codes for a lactose transmembrane transporter, and *lacA* which codes for a transacetylase that is not required for lactose metabolism [94]. When lactose is present in the growth media it is transported into the bacterium by the permease and some of the lactose molecules are transglycosylated to allolactose by the small number of β-galactosidase molecules that are present. Allolactose, or in the case of a synthetic inducer such as IPTG, binds to a repressor protein that normally occupies the operator

region of the operon [95], and the conformation changes that result from allolactose or IPTG binding lower its affinity for DNA. As a result the operator is liberated, and providing that glucose levels are low, transcription of the *lac* operon commences leading to a rapid increase in β-galactosidase levels. Copy levels of β-galactosidase can increase 100 fold within an hour [94].

Ullman *et al.* observed that inactive mutants, for which the operator proximal portion of the protein was missing, could be compensated by another peptide containing the missing residues and catalytic activity would be recovered [96]. This phenomenon is known as α-complementation and can restore a substantial portion of the lost catalytic activity; the peptide that donates the missing region is called the α peptide [97]. Residues 11-41 have been identified as responsible for α-complementation [98]. β-Galactosidase fusion proteins have been used extensively as purification tags [99], and as a reporter genes [100,101] to normalize transfection efficiency.

3.2.2. Structure and catalysis

The amino acid [102] and nucleotide sequence [103] for *E. coli* β-Galactosidase have been determined, and the crystal structure has been solved [104, 105]. The enzyme is comprised of four identical 1023 amino acid monomers, and each monomer has five independent folding domains [104]. The enzyme is active as a tetramer [90], and higher order oligomers disaggregate easily [106]. Dimers and monomers are not active [91]. Circular dichroism analysis indicates that 14% of the protein has α-helix secondary structure, and 48% is β sheets [107]. There are no disulfide bonds [108], and Mg^{2+} is required for tetrameric stability [107] and catalytic activity [109, 110], but paradoxically magnesium levels greater than 10 mM impair tetramer formation [111]. The four

monomers are organized along three mutually perpendicular axes of symmetry (Figure 9) forming interfaces between pairs of monomers.

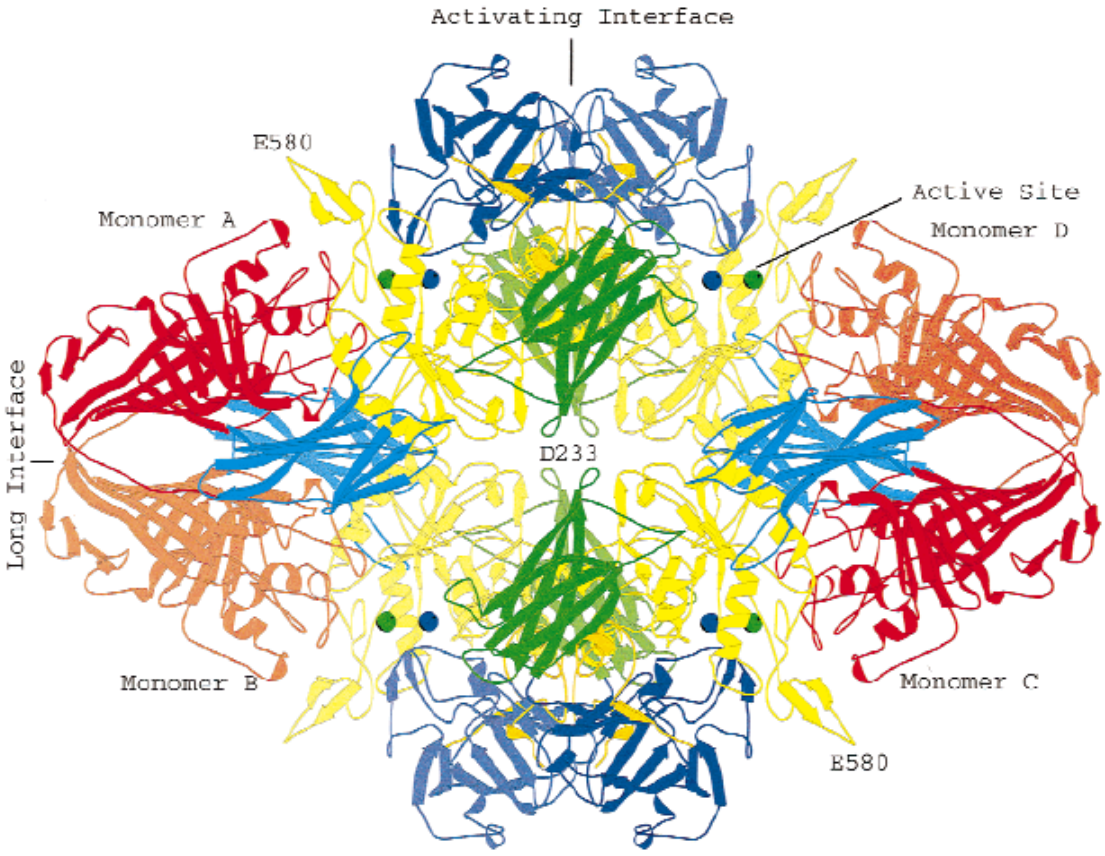


Figure 9. Ribbon representation of β-galactosidase. From "Juers, D.H., Jacobson, R.H., Wigley, D., Zhang, X.J., Huber, R.E., Tronrud, D.E., Matthews, B.W. (2000) High resolution refinement of β-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for α-complementation. Protein Sci., 9: 1685-1699." Used with permission.

The long interface relates monomer A with B and D with C, has an area of approximately 4000 $Å^2$, and is planar and relatively unstructured. The activating interface relates monomer A with D, and B with C. The area of this interface is 4600 $Å^2$, and is more contiguous than the long interface. This interface is also essential for catalytic activity because residues 272-to 288 from one monomer extend across the activating interface to complete the active site for the adjacent monomer [104]. The third interface relates

monomers A with C and B with D, and is approximately 230 $Å^2$. The importance of the contiguous interface for formation of the active site precludes active monomers. Although the long interface is characterized by weak interactions, dimers form along the long rather the activating interface. Presumably rearrangements induced by long interface interactions are necessary for the formation of the contiguous interface and explains the absence of active dimers.

The four independently acting active sites are comprised from residues of the first, third and fifth domains for a particular monomer, and the loop described above from domain two of the adjoining monomer across the activating interface. The active site(s) is a deep pit that sits in a TIM barrel of domain three and is associated with a Mg^{2+} and a Na^{+}. Residues Glu461 [112], Tyr 503 [113], Glu537 [114], His540 [115], and Trp999 [116] have all been identified as catalytically crucial. The residues of the loop from the adjoining monomer are not believed to be catalytically active *per se*, but are thought to stabilize the protein backbone in the region of Mg^{2+} binding residues [104].

The hydrolysis of substrates proceeds by a double displacement reaction through a covalent intermediate assisted by acid/base catalysis [117]. Figure 10 depicts a schematic of the reaction where OR represents glucose, or the aglycon in the case of synthetic substrates. The initial substrate interaction with the enzyme occurs by a shallow binding mode at the opening of the active site where the glucose moiety stacks on Trp999 [116]. Substrate then moves further into the active site in the deep binding mode positioning the substrate for nucleophilic attack by Glu537. There is concerted transfer of the galactosyl group to Glu537 with cleavage of the glycosidic bond to form an α-D-galactosyl enzyme intermediate and concomitant release of the aglycon. The

breaking of the glycosidic bond is assisted by acid catalysis, either by Glu461 or by a magnesium ion. The release of the galactosyl is mediated by removal of a proton from the acceptor molecule by Glu461 and the enzyme is ready for another catalytic cycle.

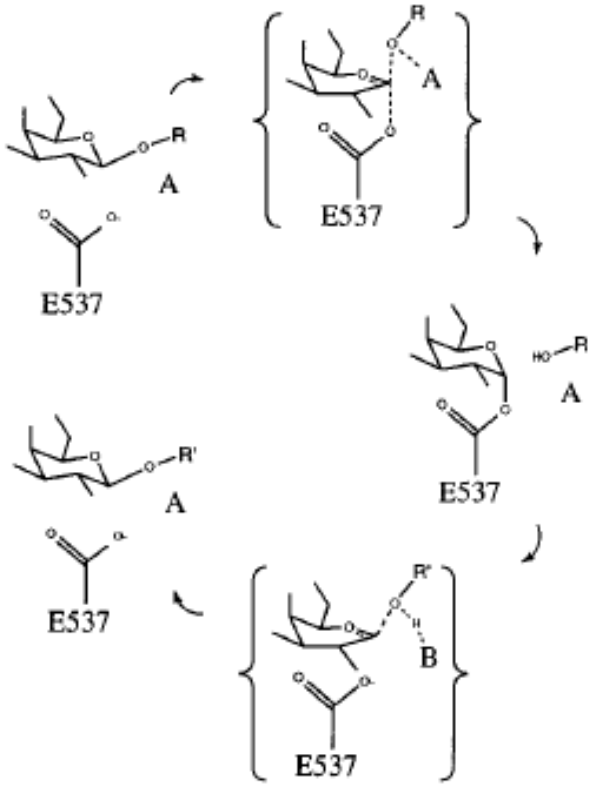


Figure 10. Schematic of reaction mechanism for hydrolysis of galactopyranoside by β-galactosidase. OR represents aglycon, HOR´ represents the acceptor. A and B signify acid and base respectively. Reprinted with permission from "Juers, D.H., Heightman, T.D., Vasella, A., McCarter, J.D., Mackenzie, L., Withers, S.G. and Matthews, B.D. (2001). A structural view of the action of *Escherichia coli* (*lacZ*) β-galactosidase. *Biochemistry* **40**(49): 14781-14794". Copyright 2001 American Chemical Society.

β-galactosidase is a good molecule for single molecule studies because of the availability of high quality fluorogenic substrates, its relatively high catalytic rate, and the inducibility that facilitates the expression of numerous copies of the enzyme under defined conditions.

4. Materials and methods

4.1. List of chemicals

Table 1. Chemicals used and supplier

Name	Supplier
2-mercaptoethanol	Sigma (St. Louis, MO)
acrylamide	Sigma (St. Louis, MO)
AEBSF	Sigma (St. Louis, MO)
agarose	Sigma (St. Louis, MO)
benzamidine hydrochloride	Sigma (St. Louis, MO)
bis-acylamide	Sigma (St. Louis, MO)
boric acid	Sigma (St. Louis, MO)
bromophenol blue	Sigma (St. Louis, MO)
chloroform	Sigma (St. Louis, MO)
coomassie brilliant blue	Sigma (St. Louis, MO)
DDAO	Invitrogen (Eugene, OR)
DDAO-gal	Invitrogen (Eugene, OR)
DEAE cellulose	Sigma (St. Louis, MO)
DMSO	Sigma (St. Louis, MO)
DNA MW markers	Roche (Indianapolis, IN)
dNTPs	Roche (Indianapolis, IN)
E-64	Sigma (St. Louis, MO)
ethidium bromide	Sigma (St. Louis, MO)
FDG	Invitrogen (Eugene, OR)
ferrous sulfate	Sigma (St. Louis, MO)
FMG	Invitrogen (Eugene, OR)
GS-6	Applied Biosystems (Foster City, CA)
glacial acetic acid	Sigma (St. Louis, MO)
glucose	Sigma (St. Louis, MO)
glycerol	Sigma (St. Louis, MO)
HEPES	Sigma (St. Louis, MO)
HF buffer	New England Biolabs (Can)
hydrolyzed casein	Difco (BD, Sparks, MD)
InstaGeneTM	Bio-Rad (Hercules, CA)
IPTG	Sigma (St. Louis, MO)
LB media	Difco (BD, Sparks, MD)
leupeptin	Sigma (St. Louis, MO)
M9 mineral salts	Sigma (St. Louis, MO)
magnesium chloride	Sigma (St. Louis, MO)
methanol	Sigma (St. Louis, MO)
nitric acid	EMD (Darmstadt, Germany)
oNPG	Sigma (St. Louis, MO)

Name	Supplier
p-aminobenzyl 1-thio-β-D-glactopyranoside	Sigma (St. Louis, MO)
PCR primers	Invitrogen (Eugene, OR)
pentamidine isethionate	Sigma (St. Louis, MO)
pepstatin	Sigma (St. Louis, MO)
phenanthroline	Sigma (St. Louis, MO)
phenylmethanesulfonylfluoride	Sigma (St. Louis, MO)
phosphoramidon	Sigma (St. Louis, MO)
Phusion DNA polymerase	New England Biolabs
protease inhibitor cocktail	Sigma (St. Louis, MO)
Protein MW marker	Sigma (St. Louis, MO)
PVP	Aldrich (Milwaukee, WI)
RES-gal	Invitrogen (Eugene, OR)
resorufin	Invitrogen (Eugene, OR)
sodium citrate	Sigma (St. Louis, MO)
sodium hydroxide	Sigma (St. Louis, MO)
sodium tetraborate	Sigma (St. Louis, MO)
streptomycin sulfate	Sigma (St. Louis, MO)
sucrose	Sigma (St. Louis, MO)
TLCK	Sigma (St. Louis, MO)
toluene	Sigma (St. Louis, MO)
TPCK	Sigma (St. Louis, MO)
Tris base	Sigma (St. Louis, MO)
zinc chloride	Fluka (Buchs, Germany)

4.2. Capillaries and instrumentation

4.2.1. Capillaries

Capillaries were made of fused silica and polyimide-coated (Polymicro Technologies, Phoenix, AZ). Capillaries with 10 μm inner diameter and 165 μm outer diameter were used for all single molecule activity and separation assays, except for the continuous assay experiments where a 2 μm inner diameter capillary was used. The length of the capillary varied from experiment to experiment, and will be indicated when specific experiments are presented. The detection and injection ends of the capillary were cut perpendicular to the longitudinal axis of the capillary using a sapphire capillary cutter (Supelco, Oakville, ON), and approximately 1 mm of the polyimide coating was

removed from the detection end by flame. The detection end of the capillary was examined under a microscope prior to installation on the CE instrument to ensure that no rough edges or other defects were present.

4.2.2. Instrumentation

All single molecule assays were performed using an in-laboratory built CE instrument equipped with a post-column laser-induced fluorescence detection system. The injection end of the capillary and a 0.5 mm diameter platinum wire connected to a high voltage power supply (Spellman model CZE 2000, Hauppauge, NY) were placed in a buffer-containing vessel in the injection carousel. The detection end of the capillary was inserted into a quartz sheath flow cuvette with a 250 X 250 μm inner bore (Hellma, Concord, ON) that was mounted onto a stainless steel custom-machined cuvette holder attached to translational stages for precise positioning in the *x*, *y* and *z* axis. The beam from a laser (see below) was focused with a 6.3 X, N.A. 0.2 microscope objective (Melles Griot, Nepean, ON) approximately 10 μm below the detection end of the capillary. Emission was collected at 90^{o} to excitation using a 60X, N.A. 0.7 microscope objective (Universe Kogaku, Oyster Bay, NY), passed through an optical filter (see below) and a pinhole and onto a photomultiplier tube (PMT) (Hamamatsu model 1477, Bridgewater, NJ). The analog PMT signal was collected and digitized using a Pentium 2 computer with an AT-MIO-16XE I/O board utilizing LabView™ software (National Instruments, Austin, TX) at 10 Hz. The same board was used to control the electrophoresis voltage and PMT bias. RES-gal assays used the 543.5 nm line from a 1 mW green HeNe laser for excitation and emission was passed through a 540df40 optical filter (Omega Optical, Brattleboro, VT); DDAO-gal assays used the 633 nm line from a

10 mW red HeNe laser and a 660BP10 optical filter (Omega Optical); and FDG assays used the 488 nm line from a 25 mW Ar^+ laser and a 535df45 filter (Omega Optical).

For the continuous flow assay a Pentium 4 computer containing a PCI-MIO-16XE-50 I/O board was used.

4.2.3. Enzyme assay buffers

All buffers used for CE-LIF and spectrophotometric assays were prepared with 18 $M\Omega cm^{-1}$ deionised water and passed through 0.45 μm filters (Nalgene, Lima, OH) to remove particulates. Separation and sheath buffers were 10 mM or 50 mM *N*-2, Hydroxymethyl-piperazine-*N′*-2-ethanesulfonic acid (HEPES), adjusted to pH 7.3 with NaOH, containing 1 mM $MgCl_2$ and 1 mM sodium citrate. For assays which used coated capillaries, separation and sheath buffers also contained 0.02% (w/v) polyvinylpyrrolidone (PVP), average molecular weight 1.3 MDa, or 0.0007% (w/v) Genescan polymer 6™ (GS-6). For the RES-gal assay, assay buffer was 62.5 mM HEPES (adjusted to pH 7.3 with NaOH) and 1.25 mM $MgCl_2$. Extraction buffer was 5 mM sodium citrate (adjusted to pH 4.8 with NaOH). Coating buffers were the same as separation buffer except the PVP was 0.5% (w/v) or 0.07% (w/v) GS-6. All buffers, vessels and pipette tips were autoclaved prior to use.

4.2.4. Capillary preparation

Capillaries were rinsed each day by pressure injection with 100mM NaOH for 10 minutes and then dynamically coated by pressure injection (approximately 5 atm) for 30 minutes with coating buffer if required. Excess coating buffer was removed by flushing the capillary with separation buffer for 30 minutes. If no capillary coat was necessary,

the capillary was flushed by pressure for 30 minutes with separation buffer following the NaOH rinse.

4.3. β-Galactosidase single molecule CE-LIF assays

4.3.1. Standards

The instrument response was calibrated each day by injection of resorufin, DDAO or FMG reference standards depending upon the substrate that would be used for enzyme assays. Concentrations of standards were 1 x 10^{-8} M for resorufin; 1 x 10^{-8} M for DDAO, and 2 x 10^{-8} M for FMG, and were prepared by serial dilution of stock preparations in DMSO into separation buffer. Standards were injected at 100 Vcm^{-1} for 5 s. Following injection, the injection end was briefly rinsed twice by immersion in two different vials containing separation buffer and then placed in a third separation buffer containing vial. The standard was mobilized towards the detector at 400 Vcm^{-1} until the peak emerged. The injection end was held at a negative potential, if capillaries were coated, or at a positive potential (normal polarity) if capillaries were uncoated. Standards were analyzed in triplicate daily.

4.3.2. RES-gal β-galactosidase assay

All RES-gal-based assays were performed on 40 cm PVP coated capillaries with the injection end held at a negative potential. Immediately prior to each run, substrate was prepared by mixing 2 μL of 100 mM RES-gal stock solution in DMSO with 198 μL of extraction buffer in a 0.6 mL microcentrifuge tube to which 200 μL chloroform was added. The mixture was vortexed at maximum speed for 1 minute and then centrifuged briefly on a microfuge (Hermle Z180M, Labnet). The RES-gal aqueous layer was removed without disturbing the chloroform and was transferred to a clean 0.6 mL

microcentrifuge tube. The substrate wash with chloroform was performed three times. After the third wash 100 μL of the substrate was added to 400 μL of assay buffer containing enzyme at a final enzyme concentration of approximately 1 x 10^{-15} M, and a final substrate concentration of 200 μM. Note that the addition of 100 μL of extraction buffer to 400 μL assay buffer produces a buffer of the same composition as the separation buffer. A total of 200 μL of sample was placed in a sterile syringe and then pressure injected (~ 5 atm) into the capillary for 5 minutes filling the capillary. After the injection the syringe was removed, the injection end was placed in a 0.6 mL microcentrifuge tube containing only separation buffer, and the contents of the capillary were statically incubated for 15 minutes. At the end of the incubation the contents of the capillary were mobilized past the detector by the application of a 400 Vcm^{-1} potential. Blanks were the same except that no enzyme was present.

4.3.3. FDG β-galactosidase assay

All runs were performed using a 40 cm PVP-coated capillary with the injection end held at a negative potential. Substrate was prepared fresh immediately prior to each run by adding by adding 5 μL of 10 mM FDG in DMSO to 495 μL of separation buffer and mixed by vortexing at high speed for 5 s. Two hundred μL of substrate was added to a clean sterile syringe and pressure injected into the capillary for 5 minutes at ~ 5 atm. The syringe was removed and the injection end of the capillary was placed in a 0.6 mL microcentrifuge tube in the injection carousel that contained approximately 1 x 10^{-15} M β-galactosidase in separation buffer. The enzyme solution was electrokinetically injected into the capillary by application of 600 Vcm^{-1} potential for 800 s. The injection end was then placed in a vessel containing only separation buffer, and the contents of the capillary

were statically incubated for 30 minutes followed by electrophoretic mobilization of the capillary contents past the detector by the application of a 400 Vcm^{-1} electric field. Blanks were the same except that no enzyme was present during the 800 s injection.

In the double incubation, the sample was prepared as previously. Following the first 30 minute incubation the sample was mobilized for 30 s at an electric field of 400 Vcm^{-1}. This was followed by a second 30 minute incubation. After the second incubation the sample was mobilized past the detector by an electric field of 400 Vcm^{-1}.

4.3.4. DDAO-gal β-galactosidase assay

A number of different assays based upon the DDAO-gal substrate were devised. The protocols for the different types are described separately. Blanks were the same as their respective assays except that no enzyme was added.

4.3.4.1. Activity-only assays:

Single molecule assays that measured only β-galactosidase activity were performed on a 40 cm PVP coated capillary with the injection end held at a negative potential. Substrate was prepared immediately prior to each run by adding 1 μL of 20 mM DDAO-gal in DMSO to 99 μL of separation buffer and 100 μL of toluene in a 0.6 mL microcentrifuge tube. The mixture was mixed by vortexing for 1 minute at high speed, and then centrifuged briefly in a bench top microfuge. Approximately 75 μL of the aqueous layer containing the DDAO-gal was removed and transferred to a new 0.6 mL microcentrifuge tube and spun briefly on a bench top centrifuge. Fifty μL of this solution was added to 50 μL of separation buffer with β-galactosidase for a final enzyme concentration of 1×10^{-15} M, and a final DDAO-gal concentration of 100 μM. The enzyme/substrate solution was pressure injected (~ 5 atm) into the capillary for 5 minutes

following which the syringe was removed and the injection end was placed in a vial containing separation buffer. The contents of the capillary were statically incubated for 5 minutes. Post-incubation the capillary contents were electrophoretically swept past the detector at an electric field of 400 Vcm^{-1}.

4.3.4.2. Short separation assays:

Assays were performed on 50 cm PVP coated capillaries with the injection end held at a negative potential. Substrate preparation, and substrate/enzyme introduction into the capillary was the same as the single molecule activity assay described above except the concentration of β-galactosidase was approximately 1 x 10^{-16} M. After the first 5 minute incubation the enzyme was mobilized by the application of a 400 Vcm^{-1} potential for 45 s and then reincubated for 5 minutes. Separations and incubations were repeated two, three or four times followed by a final 400 Vcm^{-1} mobilization past the detector.

4.3.4.3. Extended separation assays on coated capillaries:

Extended separations were performed on 50 cm PVP or GS-6 coated capillaries with the injection end held at a negative potential. Preparation of 100 μM DDAO-gal was the same as above except substrate contained no enzyme. Substrate was introduced into the capillary by pressure injection (~ 5 atm) for 5 minutes using a syringe. The injection end was then placed in a solution of 1 x 10^{-13} M β-galactosidase and 4x10^{-8} M DDAO in separation buffer and electrokinetically injected at 100 Vcm^{-1} for 7 s. The injection end was then rapidly rinsed twice by a brief immersion into two different vessels containing only separation buffer and then placed in an injection vessel containing separation buffer. The contents of the capillary were subjected to a 400 Vcm^{-1}

electric field for 10 minutes followed by a 5 minute static incubation. The capillary contents were then swept past the detector at 400 Vcm^{-1}.

For multiple extended separations and incubations substrate and enzyme were prepared and introduced into the capillary as described in the previous paragraph. Rather than a single 10 minute separation, three separations of 2, 3.5 and 5 minutes at 400 Vcm^{-1} with each separation followed by a 5 minute static incubation were performed. After the final incubation the capillary contents were swept past the detector at 400 Vcm^{-1}

4.3.4.4. Extended separations on an uncoated capillary:

Extended separations of this type were performed on a 50 cm capillary with the injection end held at a positive potential. Substrate was prepared immediately prior to each run by adding 4 μL of 20 mM DDAO-gal in DMSO to 196 μL of separation buffer and 200 μL of toluene in a 0.6 mL microcentrifuge tube. The mixture was mixed by vortexing for 1 minute at high speed and then centrifuged in a microfuge. Approximately 150 μL of the aqueous layer was removed and transferred to a clean 0.6 mL microcentrifuge tube and centrifuged briefly on a bench top microfuge; 125 μL of this was added to 375 μL of separation buffer and mixed by vortexing for 5 s at high speed for a final substrate concentration of 100 μM. A run was started by electrokinetic injection of $1x10^{-13}$ M β-galactosidase and $4x10^{-8}$ M DDAO for 7 s in separation buffer at 100 Vcm^{-1}. The injection end was rinsed twice by immersion into different vessels with separation buffer and then placed in the tube containing the substrate. Enzyme, DDAO and substrate were mobilized into the capillary for 6 minutes by the application of a 400 Vcm^{-1} potential that was followed by a five minute static incubation, and then a final 400 Vcm^{-1} mobilization. Separation and activity assays for the experiments in chapter 10

were the same except that enzyme and reference standard were injected for 5 s, incubations were for three minutes, and a 10 mM DDAO-gal in DMSO stock was used resulting in a final substrate concentration of 50 μM.

4.3.4.5. Continuous assay with DDAO-gal:

Continuous assays were performed using a 60 cm long, 2 μm inner diameter 150 μm outer diameter capillary with the injection end held at a positive potential. A 500 μL solution of 100 μM DDAO-gal was prepared as described above in section 4.3.4.4 immediately prior to the start of a run except that the 375 μL of separation buffer contained approximately 2 x 10^{-15} M β-galactosidase. A run was performed by placing the injection end of the capillary in the vial containing enzyme and substrate and continuously injecting the sample into the capillary at 300 Vcm^{-1}. Where required, regions of the capillary were heated by sandwiching it between an aluminum block and insulating foam. The aluminum block was heated to the desired temperature by passing water from a Julabo model F12 circulating water bath through an internal coil.

4.3.4.6. Peak analysis:

Run data for all single molecule assays was collected using Labview™ software and was loaded into IgorPro™ for data analysis and then copied into PeakFit™ for peak integration. Overlapping peaks were estimated by fitting to a Gaussian. If enzyme product peaks could not unambiguously be identified as originating from single molecules, they were not included in the analysis. For chapter 10 experiments, if any peak could not be resolved the entire run was discarded.

4.4. Ensemble Assays

Bulk ensemble assays for β-galactosidase with substrates oNPG, RES-gal, DDAO-gal and FDG were performed in 50 mM HEPES separation buffer at 22°C using a 8453 UV-Visible spectrophotometer (Agilent Technologies) and 4.5 mL polystyrene cuvettes with 10 mm path length. Catalytic rate was determined by measuring the change in absorbance over time at 406 nm for oNPG, 570 nm for RES-gal, 485 nm for FDG, and 620 nm for DDAO-gal for enzyme at various concentrations of substrate. All assays were performed in triplicate. K_m was calculated by Lineweaver-Burke double reciprocal plots.

4.5. Cell cultures and β-galactosidase induction

4.5.1. Culturing *E. coli*

E. coli strains ATCC 35321, ATCC 8677 and ATCC 33588 were obtained from the American Type Culture Collection (ATCC) (Manassas, VA). To provide uniform *E. coli* sources for long term storage, *E. coli* strains ATCC 35321 and ATCC 8677 were grown for 18 hours at 37°C with 50 rpm constant shaking in sterile growth media that contained 12.29 g/L M9 mineral salts, 1% (w/v) glycerol, 0.2% (w/v) hydrolyzed casein, 1 mM $MgCl_2$, 50 µM $FeSO_4$, and 50 µM $ZnCl_2$. A 20 mL aliquot from each culture was removed and added to 20 mL sterile glycerol and mixed until homogeneous. From the 50% glycerol/culture mixtures, 1 mL aliquots were transferred to sterile tubes and stored at –70°C. For experiments that used either of these two *E. coli* strains, a single 1 mL tube was used to inoculate 500 mL of growth media and cultures were grown for 18 hours under the conditions described above.

4.5.2. β-Galactosidase induction and harvesting

To induce β-galactosidase synthesis, actively growing *E. coli* culture was added to an equal volume of pre-warmed sterile growth media and incubated 30 minutes prior to the addition of IPTG to a final concentration of 250 μM. After 30 minutes or one hour of induction, 30 mL of *E. coli* culture was harvested by centrifugation at 10,000 x *g* at 4°C for 10 minutes. The fresh pellets were resuspended in 100 μL of cold 100 mM HEPES pH 7.3, 2 mM $MgCl_2$, 2 mM citrate, 2% (v/v) Sigma protease inhibitor cocktail (no metal chelators) and frozen in $N_{2(l)}$. To extract β-galactosidase, the pellets were ground, briefly thawed and then refrozen in $N_{2(l)}$. After 5 freeze/thaw cycles the cold slurry was diluted with 900 μL of cold 100 mM HEPES pH 7.3, 2mM $MgCl_2$, 2 mM citrate and 2% (v/v) protease inhibitor cocktail, and then separated by centrifugation for 1 minute at 10,000 x *g* to remove cellular debris. The supernatant was passed through a sterile 0.45 μm filter, and the filtrate was diluted with an equal volume of sterile glycerol and stored at –20°C until assayed.

4.5.3. Protease inhibition experiments

4.5.3.1. Single protease inhibitor treatments:

E. coli strains ATCC 8677 and 35321 cultures were grown for 18 hours in growth media at 37°C with 50 rpm constant shaking. Fifteen mL aliquots of each culture were removed and added to an equal volume of sterile fresh media to which a different protease inhibitor had been added. The final concentrations of the protease inhibitors after the addition of the growing cultures are outlined in Table 2. Controls contained no protease inhibitors. After the addition, the samples were incubated for 30 minutes and

then induced for β-galactosidase expression for 60 minutes with 250 μM IPTG and harvested and stored as described.

Table 2. Final concentration of protease inhibitors.

Protease inhibitor	μM	Protease class targeted
E-64	10	cysteine
PE (pepstatin A)	10	aspartic
PMSF (phenylmethanesulfonylfluoride)	1000	serine
LP (leupeptin)	100	cysteine/serine
TPCK (N-p-tosyl-L-phenylalaninechloromethylketone)	100	serine
P (phosphoramidon)	10	metalloproteinase
AEBSF (4-(2-aminoethyl)benzenesulfonylfluoride))	10	serine
TLCK (N-α-p-tosyl-L-lysinechloromethylketone hydrochloride	100	serine
PI (pentamidine isethionate)	500	serine
BH (benzamidine hydrochloride)	1000	serine
PA (phenanthroline)	1500	metalloproteinase
SPI (Sigma protease inhibitor cocktail)[a]	(1%v/v)	all

[a]Sigma protease inhibitor cocktail contains (AEBSF, PE, E-64, Bestatin, LP, Aprotinin)

4.5.3.2. Time since induction:

E. coli strain ATCC 8677 was grown for 18 hours in growth media at 37°C with 50 rpm constant shaking. The culture was divided into two 250 mL fractions to which an equal volume of sterile fresh growth media was added. Leupeptin, TLCK, E-64 and AEBSF were added to one culture flask to final concentration of 100 μM for each inhibitor except AEBSF, which had a final concentration of 1 mM. The cultures were incubated for an additional six hours followed by β-Galactosidase induction for one hour using 250 μM IPTG. Enzyme induction was halted by harvesting the cells by centrifugation at 10,000 x *g* for 10 minutes and resuspending the pellet in growth media that contained no IPTG. The wash and resuspension step was repeated three times with the final resuspension in growth media that contained 100 mM glucose instead of

glycerol. The cells that had been cultured and induced in the presence of protease inhibitors were washed and resuspended in the glucose containing growth media that also contained the protease inhibitors at the same concentration as used during the induction. Immediately prior to and at 0.5, 1.5, 3.0, 6.0 16.5, 27, 46 and 69 hours post-induction, 30 mL aliquots were removed, and enzyme was harvested and stored in the same manner as described above.

4.5.4. Role of translation error experiment

4.5.4.1. E. coli growth conditions:

E. coli strains ATCC 35321 and the ATCC 8677 were grown for 18 hours in 500 mL growth media at 37°C with 50 rpm constant shaking. Two hundred mL from each strain was re-suspended in an equal volume of pre-warmed fresh sterile growth media from which aliquots were immediately transferred into seven separate 50 mL sterile tubes for each strain. Prior to the addition of resuspended culture, sufficient 10 mg/mL streptomycin sulfate in growth media had been added to the tubes to attain final streptomycin concentrations of 0, 2, 10, 25, 50, 150 and 300 μg/mL. The cultures were incubated at 37°C with 50 rpm shaking in the presence of the various concentrations of streptomycin for one hour, and then induced with 250 μM IPTG for 30 minutes to promote β-galactosidase synthesis. After induction, β-galactosidase was harvested and stored as described previously.

The *E. coli rpsL* mutant ATCC 33588 was grown overnight in LB media that contained 100 μg/mL streptomycin at 37°C with constant shaking at 50 rpm and then transferred to an equal volume of fresh media that contained 100 μg/mL streptomycin for

one hour prior to induction with 250μM IPTG for 30 minutes. β-Galactosidase was harvested and stored as before.

4.5.4.2. E. coli Streptomycin sensitivity:

E. coli strains ATCC 8677 and 35321 were grown overnight in growth media at 37°C with 50 rpm shaking. One mL aliquots were removed and added to different tubes with 39 mL growth media that contained streptomycin at final concentrations of 0, 2, 10, 25, 50, 150 and 300 μg/mL. Cell density was estimated by measuring light scatter approximated by measuring absorbance at 550 nm with a path length of 1cm using an Agilent 8453 UV/visible spectrophotometer with a 1 nm slit. One mL aliquots were removed at 0, 3, 4, 5.5, 7.5 and 9 hrs following addition of streptomycin, mixed with 2 mL of sterile fresh growth in a cuvette, and absorbance readings taken.

4.5.4.3. β-Galactosidase thermolability:

Two hundred μL samples of enzyme from strains ATCC 8677, 35321 and 33588 were incubated at 56, 60 or 62°C (see results for details) using a Julabo model F12 circulating water bath. Eighty μL aliquots were removed after 5 and 15 minutes at the denaturing temperature, and immediately assayed for β-galactosidase activity at 22°C in 2.5 mL of 1.5 mM oNPG in 10 mM HEPES separation buffer by monitoring the change in absorbance at 406 nm. Assays were performed in triplicate.

4.6. *In vitro* β-galactosidase synthesis

4.6.1. *E.*coli genomic DNA extraction

The procedures described in this section were used to express β-galactosidase from *lacZ* from *E. coli* strains ATCC 8677 and 35321. Growth media was inoculated with 1 mL samples from long term storage and grown overnight at 37°C with constant

shaking at 50 rpm. One mL of actively growing culture was removed for each strain and a pellet was obtained by centrifugation for 5 minutes at 10 000 x *g* in a 1.5 mL microcentrifuge tube. The supernatant was discarded and the pellet was resuspended in 200 µL of InstaGene^{TM} matrix (Bio-Rad) and incubated for 30 minutes in a 56°C water bath. The mixture was mixed by vortexing at high speed for 10 s and reincubated in boiling water for 8 minutes. The preparation was mixed a second time at high speed for 10 s and separated by centrifugation for 3 minutes at 10 000 x *g*. The supernatant was then stored at – 20°C until further use as template DNA for *lacZ* amplification.

4.6.2. *LacZ* amplification

Three different *lacZ* amplifications were required for each strain depending on the presence or absence of a His_6-tag, and whether the location of the tag was at the N or C terminus of the protein. Table 3 provides the volume and order of addition of reagents for a 20 µL PCR for the amplification of *lacZ* for all three constructs.

Table 3. Volume and concentration of reagents of PCR mix for *lacZ* amplification.

Reagent	Volume (µL)	Final Conc.
ddH_2O	14.4	
5X amplification buffer	4	1X
dNTPs (10 mM)	0.4	200 µM
Sense upstream primer(10 µM)	0.3	150 nM
Anti-sense downstream primer (10 µM)	0.3	150 nM
Template DNA	0.4	
Phusion DNA polymerase (2 U/µL)	0.2	0.02 U/µL

Primer sequences and their respective annealing temperatures are provided in Table 4. The grey highlighted region of the primers indicates primer sequences for overlap extension specific for the type of His_6-tag.

Table 4. Primer sequence and annealing temperature for *lacZ* amplification for various β-galactosidase constructs. Grey highlighted region of primers indicates primer sequences for overlap extension.

His_6-tag location	**sense primer (5'-3')**	**antisense primer (5'-3')**	**Annealing Temp (°C)**
no tag	CTT TAA GAA GGA GAT ATA CCA TGA CCA TGA TTA CGG ATT CA	TGA TGA TGA GAA CCC CCC CCT TAT TTT TGA CAC CAG ACC AA	64.0
C-terminal	CTT TAA GAA GGA GAT ATA CCA TGA CCA TGA TTA CGG ATT CA	TGA TGA TGA GAA CCC CCC CCT TTT TGA CAC CAG ACC AA	58.5
N-terminal	CGC TTA ATT AAA CAT ATG ACCATG ATT ACG GAT TCA	TTA GTT AGT TAC CGG ATC CCT TAT TTT TGA CAC CAG ACC AA	62.0

The PCR mix was prepared in thin-walled, round top 0.2 mL PCR tubes, and all reagents and PCR mixtures were kept on ice during preparation. Immediately following the addition of the DNA polymerase, the tubes were mixed on a vortex at the lowest instrument setting for 2 s and then by centrifugation for 5 s on a microfuge and then immediately transferred to a PT-150 MiniCycler for amplification. Amplification conditions are listed in Table 5.

Table 5. *LacZ* amplification conditions

Step	Temp (°C)	Time	No. of Cycles
Initial denaturation	94	2 minutes	1
	98	30 s	1
Denaturation	98	10 s	
Annealing	Table 4	15 s	30
Elongation	72	90 s	
Final elongation	72	8 minutes	1
Cool/hold	4	unlimited time	

Amplification products were stored at –20°C until required as a template for the ligation and amplification of T7 regulatory elements.

4.6.3. Ligation of T7 regulatory elements by overlap extension PCR

Preparation of the DNA template for *in vitro* protein synthesis used the Linear Template Generation Set, His_6-tag from Roche. The DNA template was the *lacZ* product of the first PCR from section 4.6.2. T7 promoter, T7 terminator and C or N-terminal His_6-tags were obtained from Roche in the RTS *E. coli* Linear Template Generation Set, and prepared according to the manufacturer's instructions. For native enzyme and enzyme with a C-terminal His_6-tag, the C-terminal His_6-tag was used. The volume of reagents for a 20 μL reaction is provided in Table 6. Reagent handling was as described for the first PCR.

Table 6. Volume and concentration of reagents for ligation of His_6-tag and regulatory elements to *lacZ*.

Reagent	Volume (µL)	Final Conc.
ddH_2O	14.0	
5X amplification buffer	4.0	1X
dNTPs (10 mM)	0.4	200 µM
T7 promoter (6 µM)	0.4	120 nM
T7 terminator(6 µM)	0.4	120 nM
C or N-terminal His_6-tag	0.3	
template DNA	0.3	
Phusion DNA polymerase (2 U/µL)	0.2	0.02 U/µL

Amplification cycling times and temperatures were identical to the amplification of *lacZ* except that the annealing temperature was 68.5°C for all constructs.

4.6.4. β-Galactosidase expression

β-Galactosidase was synthesized using reagents prepared according to the manufacturers instructions from the RTS 100 *E. coli* HY kit from Roche. Expression of all three constructs was identical, and the volume of reagents for a 20 µL reaction is set out in Table 7. The DNA template is the product of the second PCR step, and all reagents and mixtures were kept on ice until incubation.

Table 7. Volume of reagents for *in vitro* synthesis of β-galactosidase.

Reagent	Volume (µL)
ddH_20	2.0
E. coli lysate	4.8
reaction mix	4.0
amino acids	4.8
methionine	0.4
reconstitution buffer	2
DNA template	2

Expression reactions were prepared in 200 µL PCR tubes and incubated for 4 hours at 30° with constant shaking at 200 rpm, and then immediately diluted 10 fold in 50% (v/v) glycerol, 50 mM HEPES (pH 7.3), 1mM $MgCl_2$, 1mM citrate and 1% (v/v) protease inhibitor cocktail (no metal chelators) and stored at -20°C until assayed for enzyme activity.

4.6.5. DNA gel electrophoresis

The purity of PCR amplification products was assessed by agarose gel electrophoresis. Gels were prepared by dissolving 0.4 g of agarose in 50 mL of boiling TBE buffer and pouring the mixture into a gel casting tray when the temperature of the mixture was 55°C. TBE contained EDTA (0.05 M) and 10.8 g/L Tris Base and 5.5 g/L boric acid. Five µL of sample was loaded into each well and separated by electrophoresis for 1 hour at 5 Vcm^{-1}. Gels were stained with 1 µg/mL ethidium bromide for 10 minutes, destained with distilled water for 5 minutes. Gels were imaged using a Biorad UV Gel Analyzer, Universal Hood II.

4.6.6. DNA sequencing

PCR products from the first set of amplification reactions were sent to the University of Toronto Sick Children's Hospital for DNA sequencing. Contiguous sequences were assembled using the CAP3 Sequence Assembly Program [118]. Final assembled sequences for *lacZ* from ATCC 8677 and 35321 were compared to the published sequence in the KEGG data base using the *lalign* program from ch.EMBnet.org [119].

4.7. Miscellaneous

4.7.1. SDS-PAGE and IEF

SDS-PAGE was performed according the method of Laemmli [120] Discontinuous gels were prepared with a stacking gel of 3.75% 37:1 acrylamide/bis-acrylamide, 0.1% SDS in 0.125 M Tris-HCl pH 6.8 and resolving gel of 7.5% 37:1 acrylamide/bis-acrylamide, 0.1% SDS in 0.375 M Tris-HCl pH 8.8 using a Fisherbrand Gel Casting Unit. Five or 10 μg of commercial β-galactosidase in 0.0625 M Tris-HCl pH 6.8, 10% (w/v) sucrose, 1% SDS, 3% (v/v) 2-mercaptoethanol, 0.0005% bromophenol blue was loaded on to the discontinuous gel. Gels were run at 200 V for one hour, and then stained with 0.25% (w/v) Coomassie brilliant blue in 45% (v/v) methanol and 9% (v/v) glacial acetic acid for 20 minutes. Destaining was overnight in 7% (v/v) methanol and 7% (v/v) glacial acetic acid. Gel imaging was performed using a Biorad UV Gel Analyzer, Universal Hood II.

Isoelectric focusing (IEF) of β-galactosidase was performed using a Bio-Rad Protean^{TM} IEF system. Five μg of sample was focused on a 7 cm pH 4-7 ReadyStrip^{TM} IPG for 10,000 Vhr^{-1} and then stained for 30 minutes with Coomassie brilliant blue, and destained in 10% acetic acid. The pI for the native enzyme was calculated using the Scripps Institute Protein Calculator V3.3 [121]. The effect of charge changes of the protein on pI were calculated using aspartate to valine substitutions to model single charge alterations.

4.7.2. β-Galactosidase purification

β-Galactosidase from *E. coli* lysate or *in vitro* synthesis sources was purified according the method of Steers and Cuatrecasas [122]. Two hundred μL of crude lysate

or *in vitro* reaction mixture was diluted 5-fold by mixing with 800 µL of 50 mM Tris chloride pH 7.5, 0.1 M NaCl, 10 mM $MgCl_2$ and 100mM 2-mercaptoethanol (loading buffer). The suspension was fractionated by adding 0.18 g of ammonium sulphate (30% saturated) and then separated by centrifugation for 10 minutes at 10,000 x *g*. One mL of the supernatant was removed and an additional 0.13 g of ammonium sulphate (total 50% saturated) was added and separated by centrifugation for 10 minutes at 10,000 x *g*. The pellet was resuspended in 10 mL of loading buffer and then added drop wise to a p-aminobenzyl 1-thio-*β*-D-glactopyranoside immobilized on agarose affinity column that had been previously equilibrated with 5 volumes of loading buffer. β-Galactosidase was eluted by the addition of 100 mM sodium tetraborate. A one mL fraction collected after the addition of the elution buffer was concentrated using a 10K MWCO filter (Palco). The protein captured on the filter was resuspended in 100 µL 50 mM HEPES pH 7.3, 1mM $MgCl_2$ and 1 mM citrate for future analysis by mass spectrometry.

4.7.3. Mass Spectrometry

Samples were subjected to tryptic digest and analyzed by tandem mass spectrometry at the Manitoba Centre for Proteomics and Systems Biology using a Sciex prototype MALDI quadropole/TOF mass spectrometer as described by Krokhin *et al*. [123].

5. Proteolysis as a source of single molecule catalytic heterogeneity

Parts of the data used for this chapter was originally published as:

Nichols, E.R., Gavina, J.M.A., Mcleod, R.G., Craig, D.B. (2007) Single molecule assays of β-galactosidase from two wild-type strains of *E. coli*: effects of protease inhibitors on microheterogeneity and different relative activities with differing substrates. *Protein J.*, **26:** 95-105. Necessary copyright permission has been obtained.

5.1. Introduction

The central paradigm of chemistry requires that structure define function, and it follows therefore that identical enzymes should have identical activity. To test this prediction, Polakowski *et al*. [50] purified the dimeric *E. coli* alkaline phosphatase by IEF to yield its three isoenzymatic forms. Single molecule assays were performed on the purified isoenzymes using CE-LIF and the individual molecules were found to be catalytically homogeneous. During the course of their investigations, the authors found that storage of the enzyme at 4°C in the absence of protease inhibitors led to single molecule catalytic heterogeneity, and up to 12 active bands when separated by IEF. This result suggested that some of the generally observed single molecule catalytic heterogeneity may be the result of limited proteolysis.

Bacterial cells are adept at adapting to changing environmental conditions, and directed proteolysis provides an additional level of control to transcription and translation for cellular regulation of protein phenotype. Proteolysis in bacterial cells serves primarily to minimize the accumulation of proteins that are normally only required at very low levels; to eliminate regulatory or repair proteins that are no longer necessary; to eliminate abnormally folded proteins as part of the heat shock response; and degrade defective protein arising from truncated protein synthesis [124]. Proteolysis is also

important for the recycling of amino acids in response to starvation [125]. Proteases also have a number of housekeeping functions including N-terminal methionine removal [126, 127] and signal peptide excision for protein exported outside of the cell [128].

E. coli contains at least 40 different proteases [125]. There are four families of ATP-dependent proteases that share features with the 26S proteases of eukaryotes that include the Lon, Clp and Alp type proteases. These families share similar features where the ATPase domain regulates access to the protease domain. They are processive proteases that typically produce peptide products 10-15 amino acids in length. The Lon and Clp proteases are responsible for the majority of protein degradation for both abnormal protein removal and for general protein regulation [124]. These proteases are unlikely to contribute to the protein structural heterogeneity that produces catalytic heterogeneity because the proteolytic product is comprised of small peptide fragments. Other *E. coli* proteases include the cytoplasmic endoproteases (II, In, Fa, So), periplasmic and membrane associated endoproteases (I, III, SohB, Prc, Mi, IV, V, VI, HflA, HflB, OmpT, OrfX, signal peptidases I and II), dipeptidases (D, E, Q), tripeptidase T, aminopeptidases (A, B, M, N, P, Iap), dipeptidylcarboxypeptidase, oligopeptidase A and metalloendopeptidase QG. The specific proteases thought responsible for single molecule catalytic heterogeneity have not been speculated upon.

A wide spectrum of protease inhibitors have been developed to minimize protein degradation during cell lysis which exposes proteins to proteases from which they are often segregated. Many of these proteases fall into the broad class of serine proteases and are susceptible to inhibition by compounds such as DFP (diisopropylfluorophosphate). DFP inhibits many of the proteases found in *E. coli*, including proteases Lon, Clp, II, Fa,

So, and OmpT. However it is particularly toxic, so AEBSF (4-(2-aminoethyl) benzenesulfonylfluoride hydrochloride) and PMSF (phenylmethanesulfonylfluoride), which are safer alternative serine protease inhibitors, are generally used. TPCK (N-p-tosyl-L-phenylalaninechloromethylketone) and TLCK (N-p-tosyl-L-lysinechloromethylketone hydrochloride) are inhibitors of chymotrypsin and trypsin respectively. These proteases are not present in *E. coli*; however, TPCK and TLCK have been shown to be effective against protease II. Benzamidine is an inhibitor of trypsin-like proteases and has been found to be effective against protease VI [125]. Leupeptin is an inhibitor of serine and cysteine proteases. Pentamidine isethionate is also a trypsin inhibitor [129]. E-64 is an inhibitor of cysteine proteases but does not affect cysteine residues of other proteins nor react with low molecular weight thiols. Several proteases are metallopeptidases and therefore can be inhibited by chelators [125]. Phenanthroline is a chelator of Fe^{2+}, Zn^{2+} and other divalent cations and an inhibitor of metalloproteases. Phosphoramidon is also an inhibitor of metalloproteases.

The effect of limited proteolysis as a source of single molecule catalytic heterogeneity of β-galactosidase was investigated by inducing enzyme expression in the presence of various protease inhibitors. If proteases do contribute to the observed single molecule variability, their inhibition should produce a reduction in the measured catalytic heterogeneity. Because the role of proteases had not been investigated previously in this context, a broad spectrum of inhibitors was chosen to increase the probability that evidence of their role would be detected. It has also been shown that the average activity of a population of enzyme molecules changes over time [48]. The possibility that proteases may contribute to this phenomenon was investigated by growing *E. coli* ATCC

8677 for an extended period in the presence of various protease inhibitors, and then periodically assaying the sample for single molecule β-galactosidase activity.

5.2. Methods

5.2.1. Protease treatments

5.2.1.1. Single Protease Treatment:

E. coli strains ATCC 8677 and 35321 cultures were grown for 18 hours as described, aliquots of each culture were removed and added to fresh media to which a different protease inhibitor had been added, incubated and then induced with IPTG and enzyme harvested. The final concentrations of the protease inhibitors after the addition of the growing cultures are outlined in table 8, (Note, this table is a reproduction of table 2 from chapter 4 section 5.3).

Table 8. Final concentration of protease inhibitors.

Protease inhibitor	μM	Protease class targeted
E-64	10	cysteine
PE (pepstatin A)	10	aspartic
PMSF (phenylmethanesulfonylfluoride)	1000	serine
LP (leupeptin)	100	cysteine/serine
TPCK (N-p-tosyl-L-phenylalaninechloromethylketone)	100	serine
P (phosphoramidon)	10	metalloproteinase
AEBSF (4-(2-aminoethyl)benzenesulfonylfluoride))	10	serine
TLCK (Nα-p-tosyl-L-lysinechloromethylketone hydrochloride	100	serine
PI (pentamidine isethionate)	500	serine
BH (benzamidine hydrochloride)	1000	serine
PA (phenanthroline)	1500	metalloproteinase
SPI (Sigma protease inhibitor cocktail)[a]	(1% v/v)	all

[a]Sigma protease inhibitor cocktail contains (AEBSF, PE, E-64, Bestatin, LP, Aprotinin)

5.2.1.2. Time since induction:

E. coli strain ATCC 8677 was grown as described, subdivided and leupeptin, TLCK, E-64 and AEBSF were added to one culture to final concentration of 100 μM for each inhibitor except, AEBSF which had a final concentration of 1 mM and then incubated for 6 hours prior to IPTG induction. Induction was halted after 1 hour by transfer to glucose rich media. Aliquots were removed immediately prior to and at various times post-induction, and enzyme was harvested and stored in the same manner as described.

5.2.2. Single molecule assays

All single molecule β-galactosidase assays were based upon the RES-gal assays. Substrate was prepared fresh immediately prior to each run by triple chloroform wash to a final concentration of 200 μM in separation buffer that contained approximately 1 fM enzyme. The enzyme substrate mixture was pressure injected for 5 minutes and statically incubated for 15 minutes, following which the capillary contents were mobilized past the detector by the application of 400 Vcm^{-1} potential (injection end negative). Triplicate resorufin reference standards and a RES-gal blank were run daily.

5.2.3. Statistical analysis

Comparison of means for the effect of a particular protease inhibitor was performed using Duncans pairwise comparison on square root transformed data. The analysis of the variation in activity of the individual molecules within a particular treatment was performed using Levene's test for homogeneity of variance. Differences at the 95% confidence level were considered to be significant.

5.3. Results

5.3.1. Description of resorufin-based single molecule assay

The RES-gal based single molecule β-galactosidase assay has been previously described [45]. The cartoon in Figure 11 depicts how the single molecule β-galactosidase assays were performed. First, the PVP-coated capillary was filled by pressure injection with a sample mixture containing 200 μM of substrate RES-gal and approximately 1 x 10^{-15} M enzyme. At this concentration, approximately 20 enzyme molecules would, on average, be randomly distributed along the 40 cm substrate-filled capillary, with an average distance between molecules of approximately 2 cm. After the capillary was filled, enzyme and substrate were statically incubated for 15 minutes. Motion within the capillary during the incubation was limited to diffusion. The enzyme was essentially static for the relatively brief duration of the assay because the diffusion constant for β-galactosidase is 3.3 x 10^{-11} m^2 s^{-1} [130] and the hydrolysis product with its diffusion constant of 5 x 10^{-10} m^2 s^{-1} [46], will only migrate approximately 1 mm from the enzyme molecule. As a result, the fluorescent product resorufin accumulated in discrete pools around the individual enzyme molecules.

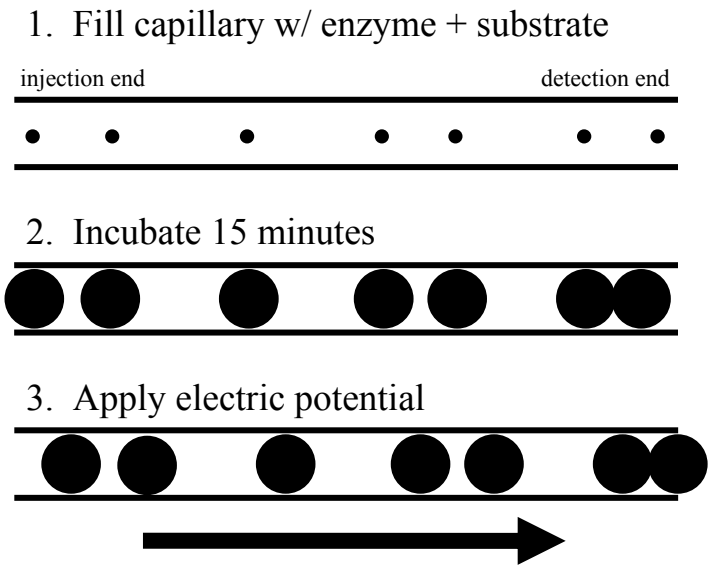


Figure 11. Schematic for RES-gal single molecule assay. 1. Capillary pressure-filled with 200 μM RES-gal and β-galactosidase (small black circles). 2. Discrete pools (large black circles) of resorufin accumulate around individual enzyme molecules. 3. Product pools are migrated past the detector by activating electric potential (injection end –ve)

It is improbable that enzyme activity would be affected by depletion of substrate because if 20 enzyme molecules were present in the capillary and each hydrolyzed 100 000 RES-gal molecules per minute during the 15 minute assay, less than 0.001% of the substrate would be consumed. Application of the 400 Vcm^{-1} potential (injection end negative) across the capillary at the end of the incubation period swept the negatively charged resorufin product molecules out of the capillary where they were excited by the laser.

Figure 12 shows a typical electropherogram for the single molecule β-galactosidase RES-gal assay. The lower trace for the blank assay shows a plateau that is present in all assays of this type, and indicates the amount of resorufin that was present in the capillary at the conclusion of the incubation. The resorufin plateau has three sources: first, the incomplete removal by the chloroform wash of the resorufin which was present

as an impurity in the resorufin-gal; second (except for the blank), resorufin that was produced by enzymatic hydrolysis of the substrate during the period between the addition of the substrate to the enzyme solution and the end of the pressure injection; and third, that which was formed by spontaneous hydrolysis of the substrate during the injection and incubation. The pools of product that formed around the individual enzyme molecules during the incubation appear as peaks sitting atop the plateau.

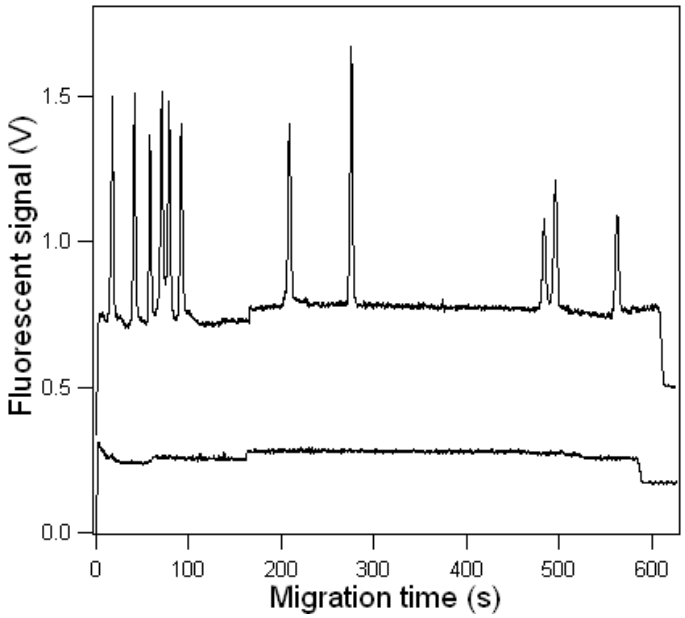


Figure 12. RES-gal single molecule β-galactosidase assay electropherogram. Traces are offset for clarity. *Bottom trace:* capillary pressure-filled with 200 μM RES-gal and incubated 5 minutes. Capillary contents mobilized past the detector by the application of 400 Vcm^{-1} potential. *Top trace:* same, except also includes approx. 1 fM β-galactosidase.

The area of each peak represents the activity of a single β-galactosidase molecule. The differences in observed peak areas reflects the range of activities the individual enzyme molecules have, rather than merely reflecting poor precision of the assay system [45]. For β-galactosidase, this range is greater than 10-fold [46].

The area of the peaks is proportional to the amount of product produced, and the catalytic activity of the individual enzyme molecules can be calculated by:

$$\text{activity} = \frac{[A_{\text{pk}}][t_{\text{i}}][E_{\text{i}}][V_{\text{c}}][C_{\text{sd}}][N_A]}{[A_{\text{sd}}][t_{\text{sd}}][E][t_{\text{ass}}]} \tag{5.1}$$

where A_{pk} is the area of the enzyme product peak; t_{i} and E_{i} are the injection time and injection potential for the daily resorufin standards respectively; V_{c} is the volume of the capillary, C_{sd} is the concentration of the standard; N_A is Avogadro's number; A_{sd} is the area for the injected standard peak on the electropherogram; t_{sd} is the migration time for the standard to the detector, E is the potential under which the standard was migrated, and t_{ass} was the incubation time for the enzyme assay. Essentially the equation calculates the volume of the injected standard which enables a straightforward conversion of the concentration to a number of resorufin molecules that the standard peak represents. The enzyme activity is the ratio of the peak area for the product produced during the incubation to the area of the standard divided by the incubation time. Injection bias as characterized by Huang *et al*. [88] is not relevant here because the same type molecule is used for measuring catalytic activity as was used for determining the quantity of the standard injected.

If an uncoated capillary were used for this assay the RES-gal would exit with the resorufin pools because its mobility is the same as the EOF which on an uncoated capillary is towards the detector. Although the substrate is nominally fluorogenic, it does fluoresce, and because its concentration is much greater than the product pools, it saturates the detector. To circumvent this problem, a coated capillary is used and the polarity of the capillary reversed so that EOF is towards the injection end of the capillary [45]. The PVP coating masks a significant fraction of the ionized silanol groups on the interior surface of the capillary, and consequently the injection end oriented EOF is greatly reduced. The substrate is neutral and migrates towards the injection end when the

potential is applied at the end of the incubation. The electrophoretic mobility of the negatively-charged resorufin product molecules is greater than the residual mobility of the EOF; as a result the product pools had net migration towards the detector.

5.3.2. Effect of protease inhibition upon catalytic heterogeneity

When *E. coli* is grown in a lactose-free media the β-galactosidase load is several molecules per cell which can increase a hundred fold following induction [48]. Since protease inhibitors were added to the media prior to induction of β-galactosidase with IPTG, the large majority of the β-galactosidase present when harvested were synthesized in the presence of the given protease inhibitor. None of the protease inhibitors were found to prevent the induction of the enzyme.

Figure 13 shows the average (mean ± σ) single molecule activities for the enzyme obtained from *E. coli* ATCC 8677 and 35321. The number of enzyme molecules assayed for each sample ranged from approximately 100 to 150 and are provided in the figure. The two control cultures for each bacterial strain had no protease inhibitors present at any time during the induction, and both times the samples yielded nearly identical average catalytic activities. The grouped average catalytic rate for β-galactosidase from strain 35321 was 105,600 ± 52,100 molecules/minute/enzyme, and for enzyme from strain 8677 it was 58,300 ± 34,800 molecules/minute/enzyme. A difference in the average activity of enzyme from different strains has been previously reported [48]. The large standard deviations reflect the underlying catalytic heterogeneity between individual enzyme molecules.

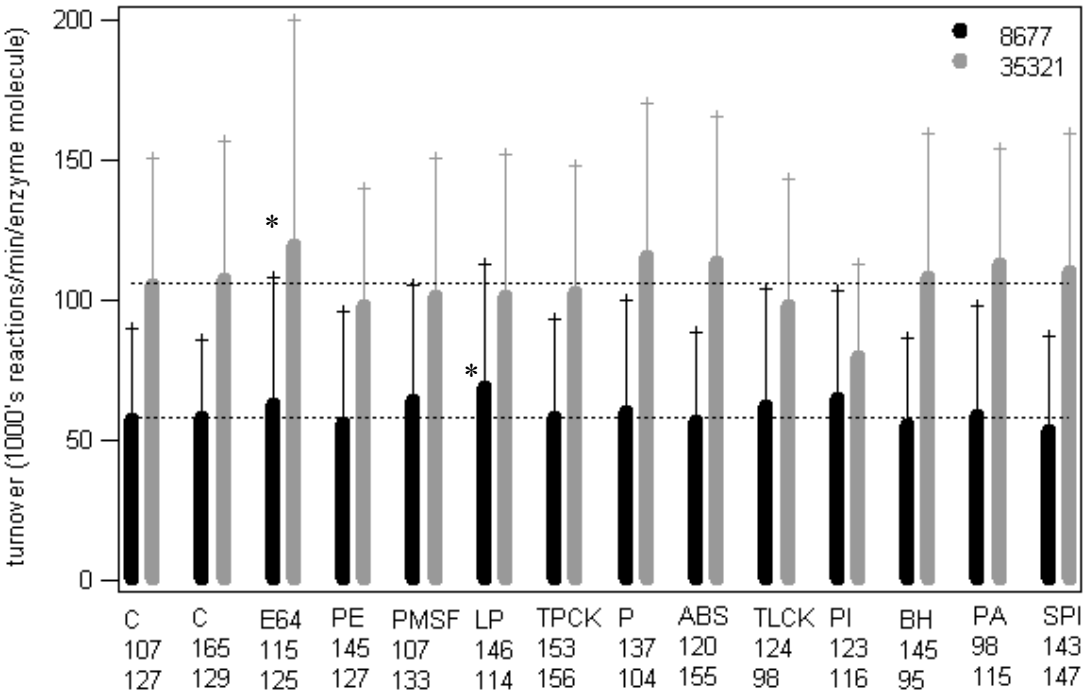


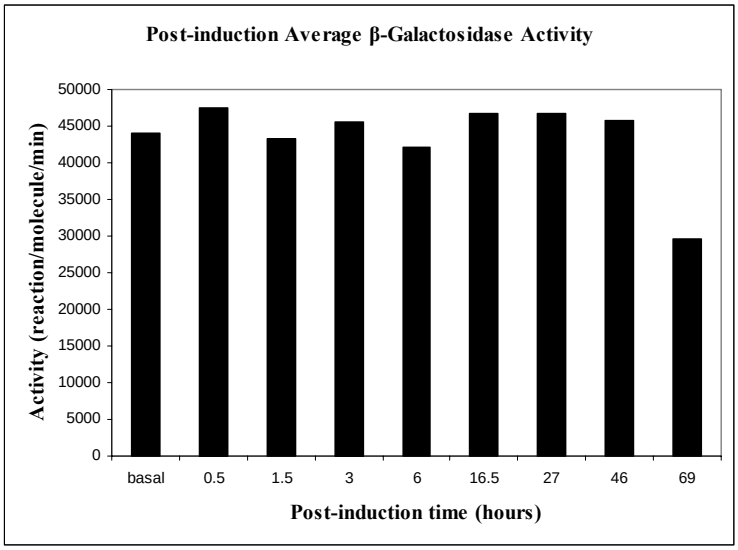
Figure 13. Effect of various proteases on the average activity of β-galactosidase. The average activities of β-galactosidase molecules induced from *E. coli* strains ATCC 8677 and 35321 in the presence of various proteases are given along with their standard deviations. The abbreviations at the base of each bar represents the inhibitor present for each induction are as follows: E64 = E-64, PE = pepstatin, PMSF = phenylmethanesulfonylfluoride, LP = leupeptin, TPCK = N-p-tosyl-L-phenylalaninechloromethylketone, P = phosphoramidon, ABS = 4-(2-aminoethyl) benzenesulfonylfluoride, TLCK = N-α-p-tosyl-L-lysinechloromethylketone , PI = pentamidine isethionate, BH = benzamidine hydrochloride, PA = phenanthroline and SPI = Sigma protease inhibitor cocktail. The top number below is the number of molecules assayed from the 8677 strain and the bottom number that from the 35321 strain. Significant difference at 95% confidence interval indicated by *.

Comparison of the average activities of β-galactosidase showed that induction in the presence of protease inhibitors had little effect, with only leupeptin in the 8677 strain, and pentamidine in the 35321strain producing a significant difference. Comparison of variance to determine if the catalytic heterogeneity of a population of molecules was altered by protease inhibition revealed that only the enzyme from strain 35321 induced in the presence of E-64, which inhibits cysteine proteases, was found to be significantly different.

5.3.3. Time from induction

The experiment to assess if the change in average activity over time of a population of enzyme molecules was affected by proteases was performed subsequent to the experiment presented above. To address the possibility that the absence of any protease inhibitor effect was attributable to a failure of the inhibitors to penetrate to the cytoplasm of the cells, or for the possibility of compensatory protease responses that evaded the single inhibitor approach, a combination of four different protease inhibitors and a longer incubation period in their presence prior to induction was performed. Figure 14 shows the average activity for the individual β-galactosidase molecules sampled at the various post induction time points for the enzyme induced in the presence and absence of the protease inhibitors.

A



B

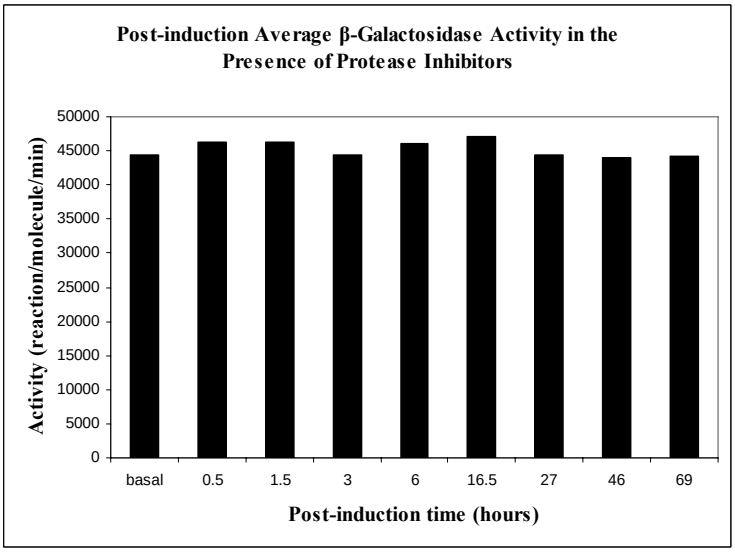


Figure 14. Effect of protease exposure over time on β-galactosidase activity. Average activity for samples of *E.coli* ATCC 8677 β-galactosidase at various times post-induction in the presence (B) and absence (A) of four different protease inhibitors (LP, TLCK, E-64, AEBSF).

The results indicate that there was little variation in the average activity of enzyme samples taken at the various times post-incubation. The reduction of average catalytic activity observed for the 69 hour sample from the culture grown in the absence of protease inhibitors but not for the culture exposed to protease inhibitors does suggest that some starvation induced proteolytic degradation of enzyme may have taken place.

The previous section indicated that the effect of leupeptin alone upon the average β-galactosidase activity was significantly different for 8677. However, there was no effect observed when it was in a mixture of additional protease inhibitors. The distribution of activities over time was examined to determine if the enzyme population became more heterogeneous over time, but no differences were present.

5.4. Discussion

The results presented here indicate that endogenous proteases do not contribute to observed single molecule catalytic heterogeneity and provides no support for the hypothesis suggested by Polakowski *et al* [50] about a putative role for proteases. There are several possible explanations for this discrepancy. One possibility is that the protease inhibitors that were present in the growth media were unable to enter the cell. No evaluation of *E. coli* permeability to the various protease inhibitors was made, nor were the inhibitory effects of the various inhibitors against *E.coli* proteases assessed. However, many of the protease inhibitors used here have been used by others [129] in the growth media and found to illicit an effect on the overall rate of protein degradation in the cell. Second, the earlier study examined alkaline phosphatase and not β-galactosidase, and it has been shown that β-galactosidase that are lacking the final ten C-terminal residues are proteolytically degraded through an 837 amino acid intermediate [131, 132]. Perhaps limited proteolytic attacks on β-galactosidase lead to its complete degradation rather than leaving a partially active enzyme. If β-galactosidase is specifically exempt, then this hypothesis needs to be tested using other enzymes. If new single molecules studies with different enzyme molecules were to support the limited proteolysis hypothesis, such a result would imply that there must still be other sources of catalytic heterogeneity that

generate the variability observed for β-galactosidase. Third, the authors [50] observed their effect of the absence of protease inhibitors for commercially sourced enzyme in extended storage at 4°C for 60 days. The emergence of catalytic heterogeneity in storage that was not present in the original sample suggests that the authors observed a storage-related issue that may not be relevant to what is happening *in vivo*. Moreover, it is not clear if there was a systematic assessment of the relationship between the duration of storage and presence and absence of protease inhibitors on the observed heterogeneity because no evidence was presented that the extent of heterogeneity increased as the storage duration also increased.

5.5. Conclusion

These results indicate that proteolysis is not likely a significant contributor to *E. coli* β-galactosidase single molecule catalytic heterogeneity. Although the results here yielded a negative result, the copious use of protease inhibitors during cell growth, β-galactosidase induction, harvesting and storage of enzyme indicates that proteolytic attack during cell lysis is unlikely to be a source of artificial single molecule heterogeneity.

6. Analysis of single molecule β-galactosidase synthesized *in vitro*

The data used for this chapter was originally published as:

E.R. Nichols and Craig, D.B. (2008) Single molecule assays reveal differences between *in vitro* and *in vivo* synthesized β-galactosidase. *Protein J.*, **26**: 376-383. Necessary copyright permission has been obtained.

6.1. Introduction

6.1.1. Background

Cells invest substantial metabolic resources to ensure proteins are produced correctly, yet single molecule enzyme studies indicate that there are considerable catalytic differences between the individual molecules [2-4]. There are several possible origins for these differences. They may be the result of processes directed by a cell to attain a particular metabolic objective; they are present due to the inherent conformational flexibility of protein molecules; or they arise from defects of their synthesis. If the single molecule heterogeneity is the consequence of the essential nature of the individual proteins, and not the manner of their synthesis or modification, enzyme synthesized *in vivo* should be indistinguishable from protein expressed by an *in vitro* translation system. However, if the mode of synthesis generates functionally distinct enzyme, differences between the two protein expression methods may shed some light on the actual source(s) of enzyme heterogeneity in cells.

In vitro or cell-free protein expression techniques of varying degrees of sophistication have been available since the early 1960s. They are generally based upon optimized lysates from *E. coli* [133], wheat germ [134, 135], or rabbit reticulocyte [136] into which exogenous DNA coding for the protein of interest under the control of a specific promoter on a plasmid, or amplified by PCR, is introduced [137]. The lysates

provide the basic molecular machinery for transcription and translation. Energy for translation is commonly provided by a phosphoenolpyruvate/pyruvate kinase system, and the ribonucleotide and amino acid building blocks for mRNA and protein synthesis are exogenously added [133].

The first cell-free expression systems were based upon the 30,000 x *g* supernatant fraction (S-30) of *E. coli* and were initially devised to elucidate the flow of genetic information from DNA to proteins [138] and were crucial for the deciphering of the genetic code [139]. The first natural protein to be synthesized by a cell-free system used RNA from f2 coliophage as a template, and protein synthesis was monitored by the incorporation of C^{14} leucine [140]. Fragments of β-galactosidase from *E. coli* were the first peptides to be synthesized directly from DNA templates [141]. The relative simplicity and flexibility of cell-free expression position these systems as excellent tools to investigate a wide variety of research interests. They have been used to unravel complicated metabolic interactions [142]; to investigate role of specific ribosomal mutations for translation fidelity [143]; explore tRNA discrimination kinetics [144]; determine the kinetic and conformational effects of streptomycin upon translation fidelity [145]; study the functional domains of proteins that are potentially toxic or insoluble [146]; for the incorporation of $^{13}C/^{15}N$-labelled amino acids for NMR spectroscopy [147, 148]; and to investigate the role of infrequently used codons for regulation of protein translation [149]. Their amenability to automation suggests that *in vitro* expression will likely be an integral tool for future integrated proteome-scale investigations for structural biology [150].

E. coli-based systems generally produce the largest quantities of protein, and the broad appeal of cell free expression has driven their ongoing refinement [151]. The gene of interest is placed under the control of a T7 promoter and transcribed by a T7 RNA polymerase [152]. The S-30 fraction of an *E. coli* lysate supplies the ribosomes, tRNAs, amino-acyl synthetases, elongation factors, and other elements required for translation [153]. The *E. coli* strain A19 is commonly used as the lysate source because it lacks RNAase I activity, and therefore minimizes digestion of mRNA transcripts which would lead to diminished protein yields [154]. Polyethylene glycol is added to mimic the viscosity and macromolecular crowding of the cytoplasm and to stabilize mRNA [155]. Poor expression yield has been identified as the principal drawback of cell-free expression, and this limitation is due to the accumulation of waste products, especially phosphate, from the energy generating system. To mitigate this shortcoming, a continuous exchange cell-free system was introduced by Spirin *et al*. in 1988, where a semipermeable membrane confines the molecular machinery for transcription and translation, but allows for the continuous influx of small molecule reagents and efflux of waste products [156]. When combined with a creatine phosphate/creatine kinase energy source, these continuous exchange systems can produce as much as 6 mg of protein/mL [148]. However, because continuous exchange systems waste considerable resources, additional optimization of batch mode cell-free systems has been pursued by using pyruvate as an energy source, and by attempting to recreate an environment that mimics the *E. coli* cytoplasm [155]. A cell-free system based exclusively upon highly purified recombinant components was recently introduced to provide enhanced control over the expression environment [157].

6.1.2. Rationale

The comparison at the single molecule level of β-galactosidase synthesized *in vitro* to enzyme synthesized *in vivo* was undertaken with the aim that differences between the enzyme from the two modes of expression might provide insights about the source(s) of single molecule heterogeneity *in vivo*. Also, because purification tags are easily introduced into a cell-free expressed protein, β-galactosidase was expressed with a C or an N-terminal His_6 tag as any possible catalytic effects from this modification provided an opportunity to investigate the hypothesis that small changes to the primary sequence affect the activity of single enzyme molecules [46].

An incidental objective of these experiments was to test the unarticulated major premise for *in vitro* protein synthesis that the reaction products are functionally identical to their counterparts synthesized *in vivo*. The products of *in vitro* protein synthesis reactions are commonly compared to their *in vivo* counterparts by SDS-PAGE, or if they are enzymes, by examining specific activity. Single molecule enzymology makes possible the direct comparison of the products of *in vitro* protein synthesis to their *in vivo* counterparts. Because a sample of the protein product is examined molecule-by-molecule, one can determine accurately the activity of active molecules. By contrast, bulk enzyme activity assays coupled to protein content determinations are less reliable because the latter cannot differentiate between active and inactive enzyme molecules and frequently makes use of standards that are different from the molecule being investigated. Also, ensemble measurements provide only average enzyme activities, whereas single enzyme molecule assays allow one to assemble histograms that describe a distribution of

activities of a population of enzyme molecules that may reveal additional differences between populations of enzyme molecules.

6.2. Methods

6.2.1. Cell-free β-galactosidase synthesis

Genomic DNA was extracted from *E. coli* strains 8677 and 35321, and a linear DNA template construct coding for *lacZ* and the regulatory elements necessary for coupled transcription/translation was synthesized by two successive PCR reactions (Figure 15). The first PCR amplified the *lacZ* gene using primers that had 5' extensions (designated in gray) that were specific for the location (if any) of the His_6-tag.

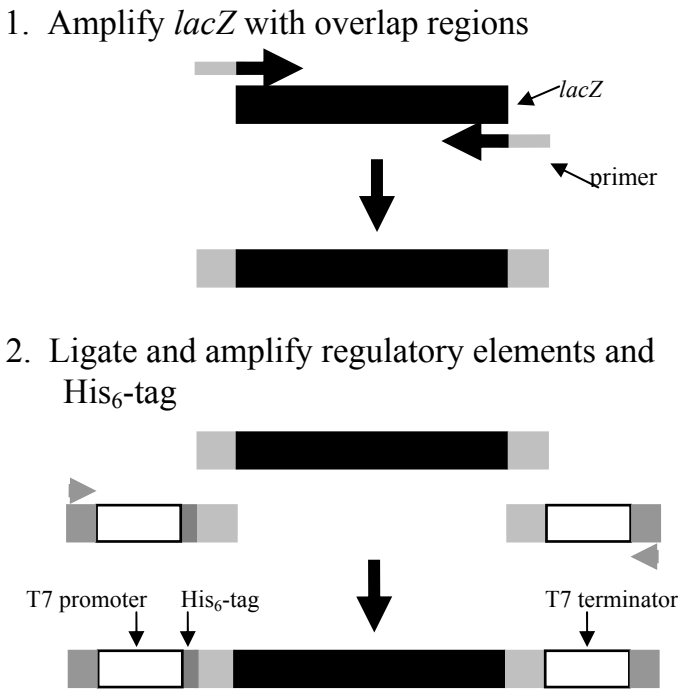


Figure 15. Schematic for two-step overlap extension PCR.

The product from the first PCR was mixed with oligonucleotides that contained the T7 promoter, T7 terminator and a ribosomal binding site required for coupled

transcription/translation that were complementary to the first set of primer extensions, and the 3' ends were extended. The entire *lacZ* gene and regulatory element construct was amplified using flanking primers supplied by the kit manufacturer in a second PCR step. The product of the second PCR was used as the template for the coupled transcription and translation of β-galactosidase with the *E. coli* 100 HY kit.

6.2.2. Single Molecule Assays

Single molecule β-galactosidase measurements were based on the RES-gal single molecule assay. Instrumentation, assay protocol, standards and electropherogram analysis were identical to the description in chapter 5. Samples were assayed in an arbitrary order.

6.3. Results and discussion

6.3.1. *LacZ* amplification

To minimize the introduction of sequence heterogeneity to the *in vitro* synthesized enzyme due to heterogeneous amplification products, an empirical approach was used to optimize PCR conditions. Initial efforts to produce pure first PCR products by utilizing various combinations of annealing temperature, Mg^{2+} concentration and amounts of genomic DNA template were not successful. However when the DNA polymerase that was originally used was replaced with the high fidelity Phusion® DNA polymerase, PCR product purity improved dramatically. Figure 16 shows typical reaction products for the first and second PCR amplification steps. The *lacZ* gene is 3075 bp, and the purity of the first reaction product was similar for all constructs. The product from the second PCR amplification is shifted up from *lacZ* band because of the ligated oligonucleotides containing the regulatory elements. This product consistently exhibited

secondary banding. Sequential varying of template and primer concentrations in conjunction with different annealing temperatures did not yield superior results to those displayed in the figure. The origin of the secondary bands is unclear; they may be a common by-product as extensive secondary banding was also present in depictions of the second PCR product in the manufacturer's instructional materials.

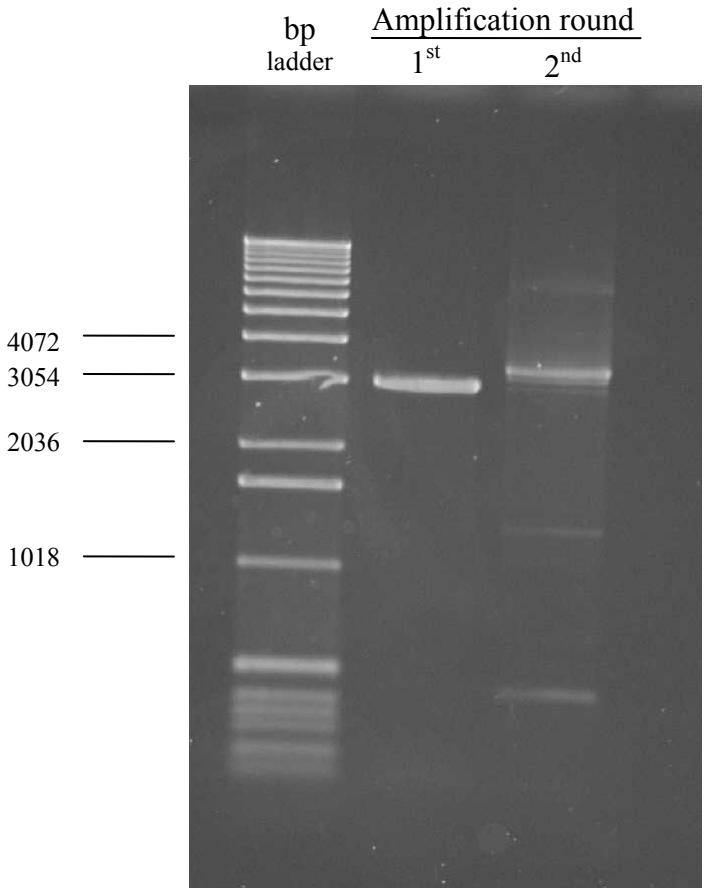


Figure 16. First and second *lacZ* PCR amplification products. Five µL of sample was separated on a 0.8% agarose gel, stained with 1 µg/mL ethidium bromide and visualized using a Biorad UV Gel Analyzer, Universal Hood II.

6.3.2. DNA sequencing

Confirmation that *lacZ* had been amplified was provided by DNA sequencing of the first PCR product for both strains. The experiments presented in the chapter on the

role of proteases indicated that there were differences in β-galactosidase activity between the *E. coli* strains ATCC 8677 and 35321, so to determine if differences in the primary sequences of β-galactosidase enzymes studied were present, the complete *lacZ* gene from each strain was sequenced and compared to the sequence available from the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. The only difference between *lacZ* from ATCC 35321 and the published sequence was a silent mutation for a glycine at residue 20. Four overlapping sequences for 8677 showed a point mutation resulting in the substitution of aspartate at position 280 with asparagine. This residue has not been previously implicated as being catalytically active *per se*, but it does contribute to the stability of the active site [104]. β-Galactosidase is a tetramer, and each monomer contains an active site that involves a loop of residues from an adjacent monomer. Asp280 is part of this loop and forms an H-bond to the backbone amide on residue 423, which is relatively close to the Mg^{2+} in the active site [105]. The mechanism for the effect that this substitution has on catalytic activity is examined in chapter7.

6.3.3. *In vitro* β-galactosidase expression

It was belatedly discovered that the *E. coli* lysate that is part of the protein expression kit contains an endogenous β-galactosidase. By assaying the lysate using CE-LIF, the actual concentration of this endogenous β-galactosidase could be calculated based upon the dilution factor, the volume of the capillary, and the average number of molecules per run. Only those reactions that yielded newly synthesized β-galactosidase at a concentration greater than 100 times background enzyme were used for the results presented here to minimize the influence of the endogenous enzyme. This concentration threshold was surpassed for two individual reactions for both strains for native enzyme,

and for enzyme that had a C-terminal His_6-tag. However, despite extensive manipulation of reaction conditions, functional enzyme with an N-terminal His_6-tag could not be synthesized for either strain. There was no evident difference in the quality of the second PCR products as determined by agarose gel electrophoresis. The N-terminal region is crucial for the inactive dimers to form active tetramers [91], so it is possible that the presence of the tag interfered with the assembly of the quaternary structure. It should be noted that this explanation is difficult to reconcile with α-complementation and fusion proteins which attach a novel amino acid sequence to the N-terminal side of β-galactosidase [97].

6.3.4. Comparison of β-galactosidase from two strains of E. coli cells

Figure 17 shows an electropherograms from an assay of β-galactosidase obtained from bacterial cultures induced by IPTG for strains ATCC 8677, ATCC 35321 and a blank.

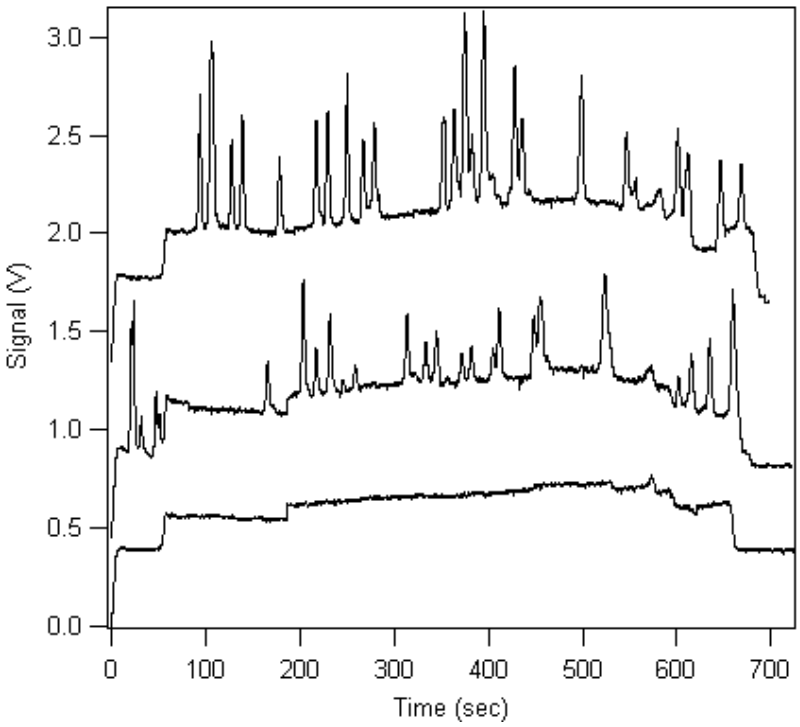


Figure 17. Electropherograms for β-galactosidase assays for enzyme from 8677 and 35321. Fifteen minute incubation of β-galactosidase from two sources with 200 μM RES-gal. *Lower trace:* blank; *middle trace:* 8677; *top trace*: 35321.

The electropherogram exhibits the same features as were observed for the RES-gal assays described previously for the protease inhibitor experiments. Peaks representing the activity of individual enzyme molecules sit atop the resorufin plateau. It is evident that the different molecules have different catalytic activity, and that enzyme molecules from ATCC 35321 (top trace) are more active than those from 8677 (middle trace).

A total of 139 β-galactosidase molecules from ATCC 35321 and 181 ATCC 8677 were assayed by CE-LIF (Table 9), and their average catalytic rates for the hydrolysis of 200 μM RES-gal to resorufin were 53,400 ± 18,400 and 34,300 ± 17,800 reactions/minute/enzyme respectively. The difference in average activity for β-galactosidase from different strains of wild type *E. coli* has been previously reported [48].

The difference in activity with respect to these two particular strains is similar to the results presented in the protease inhibitor chapter.

Table 9. Average catalytic activity for single β-galactosidase molecules from *in vitro* and *in vivo* sources.

Enzyme Source	N	Activity
ATCC 35321	139	53,400 ± 18,400
in vitro 35321 no tag	302	35,800 ± 20,900
in vitro 35321 His_6 tag	315	31,700 ± 17,700
ATCC 8677	181	34,300 ± 17,800
in vitro 8677 no tag	288	29,000 ± 17,900
in vitro 8677 His_6 tag	240	25,200 ± 12,600

6.3.5. Effect of *in vitro* synthesis for β-galactosidase activity

Table 9 also summarizes the average activities for individual β-galactosidase molecules synthesized *in vitro* for both strains of *E. coli*, with and without a C-terminal His_6 tag. The average activity of 302 single molecules of β-galactosidase synthesized *in vitro* using the *lac Z* template from *E. coli* 35321 was 35,800 ± 20,900 reactions/minute/molecule. The average activity of 315 individual molecules of β-galactosidase with a C-terminal His_6-tag synthesized *in vitro* using the *lacZ* template from *E. coli* 35321 was 31,700 ± 17,700 reactions/minute/molecule. The average activity of both forms of 35321 β-galactosidase synthesized *in vitro* are significantly less active than the 53,400 ± 18,400 for enzyme produced by the bacteria (for both strains the p value at 95% confidence is <0.0005). Also, the average activity of β-galactosidase synthesized *in vitro* with a C-terminal His_6-tag was significantly less active than the enzyme synthesized *in vitro* without a tag ($p = 0.0098$). Similar results were obtained for *E. coli* strain 8677. The average activity of 288 *in vitro* synthesized β-galactosidase with

no tag was 29,000 ± 17,900 reactions/minute/molecule and 25,200 ± 12,600 reactions/minute/molecule for β-galactosidase synthesized *in vitro* with a C-terminal His_6-tag. Both are significantly less active than the 34,300 ± 17,800 hydrolysis rate for β-galactosidase obtained from the bacteria (p values 0.002 and <0.0005 respectively). The C-terminal His_6-tag 8677 enzyme was also significantly less active than the *in vitro* β-galactosidase for 8677 without a tag (p = 0.0043).

The *K*m values of the *in vivo*, *in vitro* and His_6-tag containing enzyme from both strains for RES-gal were determined by bulk assay under the same conditions as the single molecule assay. For both strains the *in vitro* synthesized enzyme showed slightly higher *K*m values. The *K*m of the 35321 strain enzyme produced *in vivo* was 330 μM which is similar to the previously reported *K*m range of 340-380 [158]. The *K*m values for the 35321 enzyme synthesized *in vitro* with no tag and with the His_6-tag were 360 and 370 μM respectively. For the 8677 strain, the *in vivo, in vitro* and *in vitro* His_6-tag had *K*m's of 220, 240 and 240 μM respectively. The differences in *K*m between the modes of synthesis cannot account for the differences of the average catalytic rates, and this implies that there are also differences in k_{cat} amongst the samples. The specific structural basis for these differences between enzymes synthesized *in vitro* and *in vivo* is unclear.

Histograms for the distributions of single molecule activity for *in vivo* and *in vitro* synthesized enzyme from both strains are presented in Figure 18. The distribution of activity for enzyme expressed *in vivo* from 35321 is more symmetrical than *in vivo* enzyme from 8677. The shape of the distributions of single molecule activities for enzyme synthesized *in vitro* is similar for both strains. There were no individual molecules for *in vivo* 35321 enzyme that had a catalytic activity less than 10,000

reactions/minute/molecule, whereas there were molecules with such low activity for both variants of the *in vitro* enzyme. The enzyme expressed *in vitro* from 35321 *lacZ* did produce molecules that were as active as the most active *in vivo* molecules, but they were less common. The most abundant class for both *in vitro* 35321 enzyme is for molecules with activity between 21,000 to 30,000 reactions/minute/molecule which accounts for approximately a third of all the molecules assayed for both sources, while this class accounted for only 6% of the *in vivo* enzyme. A similar pattern for the differences between the *in vivo* and *in vitro* is also seen for 8677 β-galactosidase. However, for this source, the shift to lower activity molecules is less pronounced. It may be noteworthy that the distribution of 8677 *in vivo* enzyme that has the D280N substitution discussed earlier closely resembles the distribution for the 35321 *in vitro* no tag enzyme.

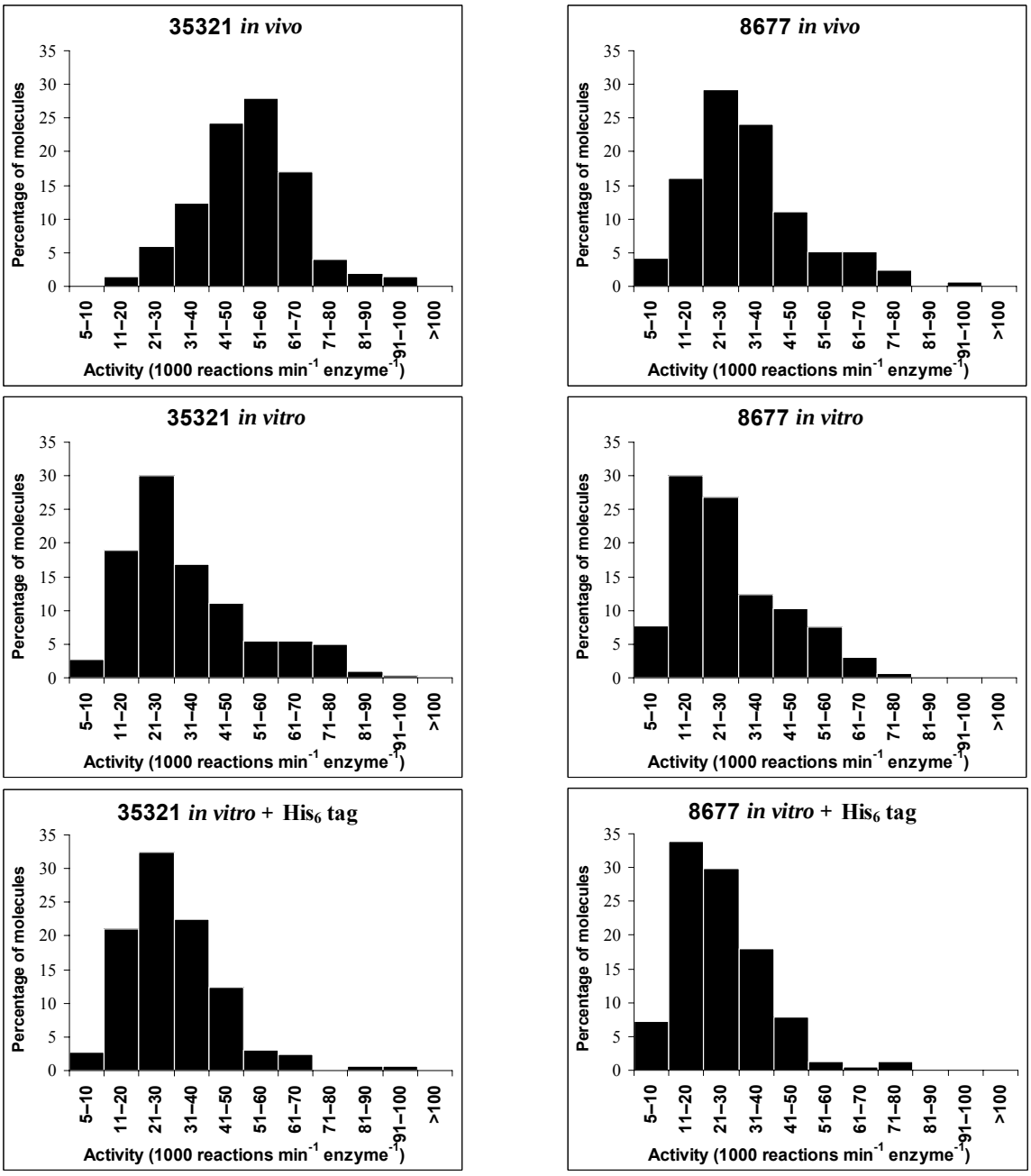


Figure 18. Distribution of activities for *in vitro* and *in vivo* synthesized β-galactosidase. The number of molecules analyzed is indicated in table 9.

6.3.6. Source of differences between *in vitro* and in *vivo* synthesized enzyme

These single molecule results reveal that although the cell-free translation systems produce functional β-galactosidase, the molecules are not identical to the enzyme expressed *in vivo*. This prompts the question of why are the molecules different. One

possible explanation for the difference in the average activity and the distribution of activities of enzyme synthesized *in vitro* is deficient N-terminal formylmethionine excision of enzyme from the *in vitro* sources. A significant fraction of *E. coli* cytosolic proteins, including β-galactosidase, have the N-terminal methionine used for translation initiation removed by methionyl-aminopeptidase [159, 127]. *E.coli in vitro* expression systems generally lack the capacity for effective N-terminal methionine removal [160], and various attempts have been made to remedy this shortcoming [161,162]. Mass spectrometry of the *in vitro* synthesized protein was undertaken to determine if the observed differences in activity were attributable to deficient N-terminal processing. The presence of the β-galactosidase N-terminal methionine containing sequence MTMITDSLAVVLQR was detected in a 1:5 ratio with the properly excised formylmethionine sequence TMITDSLAVVLQR, whereas only the TMITDSLAVVLQR sequence was detected from the *in vivo* source. Torizawa *et al.* [160] also found that an *E.coli* cell-free system failed to remove approximately 75% of the N-terminal methionine from calmodulin as determined by mass spectrometry. Assuming that there is no preferential assortment of the monomers that retained the methionine, a binary distribution of the monomers would suggest that approximately 50% of the *in vitro* enzyme would contain at least one subunit with an N-terminal methionine. Whether this can account for the observed differences between the *in vivo* and *in vitro* enzyme is uncertain. Various studies indicate that a lack of the removal of the N-terminal methionine affects the behavior of different proteins differently. Examples of the effects when the N-terminal methionine is not removed include unusual interactions between the N and C terminus of interleukin-2 [163], and reduced lysozyme solubility [164]. NMR

studies of calmodulin synthesized *in vitro* indicated that the location of numerous residues at varying distances from the N-terminus differed depending on the presence or absence of an N-terminal methionine [160].

A second possible source of structural differences between the *in vivo* and *in vitro* proteins is due to translation error. A comparison of error rates measured *in vitro* [145] and *in vivo* [165] suggests that *in vitro* synthesis has a 10-fold higher error rate than *in vivo*. HIV-1 reverse transcription *in vivo* is less error prone than predicted from the fidelity of purified reverse transcriptase in cell-free studies [166]. This difference was attributed to the *in vivo* presence of accessory proteins or cellular mismatch repair systems, and suggests that accessory protein deficiency may contribute to higher error rates in cell-free translation systems. The higher translation error rates might account for the differences in average activity and distributions of activity of the cell-free enzyme. Errors are likely to produce less active enzyme, and the higher error rates for the cell-free system should produce larger amounts of enzyme with reduced activity. However, some error free copies would continue to be synthesized. This could account for the reduction in the average activity of the *in vitro* protein and the positive skew of the distribution of activities. The presence of a positive skew in the distribution of activities of *in vivo* 8677 that has the D280N substitution provides some support for this explanation.

Variable protein truncation may also contribute to the differences between the two modes of synthesis. C-terminal truncated β-galactosidase are degraded *in vivo* [167], but, Manley [168] has shown that *in vitro* synthesis of β-galactosidase produces greater amounts of prematurely terminated β-galactosidase fragments. Presumably the responsible degradative pathway is less efficient in cell-free extracts. Because it is the N-

terminal region of the monomers that mediate the interactions required for active tetramers, and the majority of the residues that constitute the active site reside in the first half of the protein [117], C-terminal truncation could still yield partially active tetramers. Whether a particular truncated monomer contributes to the measured catalytic rate would depend upon how close the truncation was to the stop codon. Tryptophan 999 is important for substrate binding [116], so truncation before this residue would presumably eliminate activity for the particular monomer. However, because the four active sites are independent, complete loss of activity for a particular monomer might not extinguish tetramer activity provided the truncation did not unduly disturb the other active sites.

The initial amplification of the genomic DNA is an unlikely source of significant protein sequence heterogeneity. The error rate for the *Taq* DNA polymerase used here is reported to be 4.4×10^{-7} (New England Biolabs package insert). This error rate would result in a replication error in 1 out every 700 copies of *lacZ*—without considering that some of the mutations would be silent. A replication error that occurred early in the chain reaction would have a greater impact on the overall reaction fidelity than one that occurred later, but presumably the contribution of this source of error is minimal compared to translation error.

6.3.7. Effect of C-terminal His_6-tag on β-galactosidase Activity

The introduction of fusion tags serves a variety of purposes. They are used to assist isolation and purification, and they facilitate the continuous monitoring of a molecule of interest *in vivo* when fused with a fluorescent reporter molecule. The addition of a His_6-tag to a protein of interest enables the researcher to isolate the protein by affinity chromatography on a Ni^{2+}-NTA column or to identify the protein in a complex

mixture with anti-His specific antibodies [99]. Some investigations suggest that the presence of a His_6-tag has no effect on activity [146], whereas other studies have demonstrated that failure to remove the His_6-tag affects different proteins differently. It has been shown that a C-terminal His tag on PhoP, the response regulator of the PhoP/PhoQ two-component response regulator from *Salmonella enterica*, affects its dimerization and DNA-binding properties [169]. A recombinant kinase expressed in *E. coli* exhibited unusual autocatalytic properties on serine residues adjacent to the tag [170], and the relative position of the tag affected the binding properties of a recombinant antibody [171]. *In vitro* synthesized β-galactosidase with a C-terminal His_6-tag based upon *lacZ* templates from strains 35321 and 8677 are significantly less active than their native *in vitro* counterparts, and functional protein with an N-terminal His_6-tag could not be synthesized. These results reiterate the observation that the effect of a tag on function depends upon the protein and the location of the tag. The significant reduction of activity for the C-terminal tagged protein provides some support for the hypothesis that changes to the amino acid sequence of even a very large protein like β-galactosidase can have measurable effects on activity, and affirms at the single molecule level that the effect of tags on protein properties are not benign.

6.4. Conclusion

The single molecule assays on *in vivo* and *in vitro* synthesized β-galactosidase revealed that enzyme from all sources is heterogeneous. *In vitro* synthesized enzyme was harvested from the reaction mixtures without the need for any chemical or mechanical disruption of cells. While mindful of the possible sources of heterogeneity that may be present in the *in vitro* enzyme as discussed above, this result suggests that the observed

heterogeneity of enzyme liberated from functional *E. coli* cells is not an artifact introduced when obtaining the enzyme. The results here also unambiguously demonstrate that a population of protein synthesized *in vitro* is not identical to its *in vivo* counterpart. Because *in vitro* synthesized enzyme was found to be significantly less active, and that the presence of a fusion tag further reduced catalytic activity, it suggests that caution is required if one assumes that synthesis of proteins *in vitro* for a specific purpose will behave exactly the same as their *in vivo* counterparts. Finally, the known deficiencies of cell free translation systems also point towards translation error as a possible source for the catalytic heterogeneity that is present in enzyme obtained from living cells.

Appendix I Sense strand sequences added by overlap extension PCR

Bold regions represent overlap regions for the addition of T7 RNA polymerase regulatory elements.

Native non-tagged protein

upstream-Addition:
5´cggtcacgcttgggactgccataggctggcccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatc
ccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaa**ctttaagaaggagatata**
cc + ATG

downstream-Addition:
5´TAA +
gggggggg**ttctcatcatca**tcatcatcattaataaaagggcgaattccagcacactggcggccgttactagtggatccggctg
ctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacgg
gtcttgaggggttttttgctgaaaggaggaactatatccggagcgactcccacggcacgttggcaagctcg

C-terminal His-tag DNA

upstream-Addition:
5´cggtcacgcttgggactgccataggctggcccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatc
ccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaa**ctttaagaaggagatata**
cc + ATG

downstream-Addition:
5´**gggggggg****ttctcatcatca**tcatcatcattaataaaagggcgaattccagcacactggcggccgttactagtggatccggc
tgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataaccccttggggcctctaaacg
ggtcttgaggggttttttgctgaaaggaggaactatatccggagcgactcccacggcacgttggcaagctcg

N-terminal His-tag DNA

upstream-Addition:
5´cggtcacgcttgggactgccataggctggcccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatc
ccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatatacc
atgtctggttctcatcatcatcatcatcatagcagcggcatcgaaggccgcggc**cgcttaattaaacatatgacc**

downstream-Addition:
5´**gggatccggtaactaactaa**gatccggtaagatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccac
cgctgagcaataactagcataaccccttggggcctctaaacgggtcttgaggggttttttgctgaaaggaggaactatatccgga
gcgactcccacggcacgttggcaagctcg

7. Two novel β-galactosidase single molecule CE-LIF assays

Some of the data used for this chapter was originally published as:

Nichols, E.R., Gavina, J.M.A, Mcleod, R.G., Craig, D.B. (2007) Single molecule assays of β-galactosidase from two wild-type strains of *E. coli*: effects of protease inhibitors on microheterogeneity and different relative activities with differing substrates. *Protein J.*, **26:** 95-105. Copyright permission to reuse the data here has been obtained.

7.1. Introduction

A basic requirement of single molecule assays is that the concentration of the molecules of interest is low enough that only one or several individual molecules are present in the assay volume. The dilution required to attain the necessary concentration is usually determined empirically. To expedite this process in our lab when new β-galactosidase samples for single molecule CE-LIF assays are prepared, spectrophotometric assays are performed using the chromogenic substrate oNPG to obtain a rough estimate of the amount of active enzyme in the sample. This colorless β-galactosidase substrate produces a yellow product when hydrolyzed by the enzyme, and the rate of hydrolysis can be monitored by measuring the change of absorbance at 406 nm. This preliminary protocol was observed for the all the samples harvested for the protease inhibitor experiments presented in chapter 5. When the experiments were completed it was evident that the preliminary spectrophotometer data and the results of the single molecule assays provided a ready opportunity to compare the average bulk activity of the enzyme with the activity that was measured for the single molecule experiments—although with different substrates. Determination of single molecule activity based upon bulk assays is a dubious proposition because, regardless of purity, a precise determination of the number of active molecules cannot be made. However, the single molecule assays for the protease inhibitor experiment provided explicit

information about the concentration of active β-galactosidase molecules that were present in the samples, and because the same samples were used for the bulk assay, the average rate of oNPG hydrolysis per individual enzyme molecule could be determined. The calculation for the concentration of the enzyme was based upon the volume of the capillary, the dilution factor and the average number of molecules present in a run. The average activity of individual molecules for the hydrolysis of oNPG could be determined from the rate of product formation and the number of molecules that were present in the bulk assay. On this basis the average catalytic rate for β-galactosidase from ATCC 35321 for the hydrolysis of oNPG was found to agree with the rate for the hydrolysis of RES-gal determined by single molecule assay to within 5%. However, for enzyme from the ATCC 8677 strain, the rate of hydrolysis for oNPG was twice the rate as that observed for RES-gal. Also, the rate for oNPG hydrolysis for 8677 enzyme was higher than that of 35321—a scenario that was never seen with the RES-gal substrate. The aglycon for oNPG is a substituted phenyl ring, whereas the RES-gal aglycon is based upon three fused rings (Figure 19).

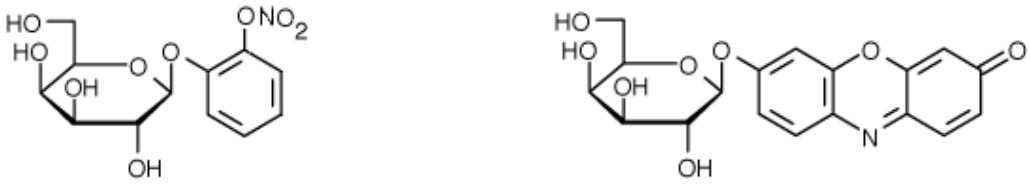


Figure 19. Structures of oNPG and RES-gal.

This difference led to the formulation of the hypothesis that the catalytic rate for enzyme from the 8677 strain was more sensitive to the properties of the aglycon than enzyme from the 35321 strain. To test this hypothesis two new single enzyme molecule assays

were devised using the commercially available β-galactosidase fluorogenic substrates DDAO-β-D-galactoside (DDAO-gal) [172], and fluorescein-β-D-digalactoside (FDG). A secondary motivation for the development of the DDAO-gal assay was to provide a substrate that could be used to compete with RES-gal for planned two optical channel single molecule kinetic assays. This chapter describes the development of the new assays, and presents the results for the effect that aglycon substitution has for the relative hydrolysis rate for β-galactosidase from *E. coli* strains ATCC 35321 and 8677.

7.2. Methods

7.2.1. Cell culture and β-galactosidase harvesting

E. coli strains were cultured and β-galactosidase was harvested and stored as described previously. Commercial enzyme was used without further purification.

7.2.2. Capillaries, instrumentation, and reference standards

Capillary length was either 40 or 50 cm; all buffers, daily capillary preparation and coating with PVP were as previously described. Instrumentation was the same, except for the laser source and optical filters, with the differences identified below. Product reference standards were 1.0×10^{-8} M for DDAO and 2.0×10^{-8} M FMG, and were injected for 5 s at 100 Vcm^{-1} (injection end negative) and mobilized at 400 Vcm^{-1} until the peak passed the detector. Triplicate reference standards were run daily. Substrate purity was assessed by CE-LIF on uncoated capillaries by a 5 s injection at 100 Vcm^{-1} and 400 Vcm^{-1} separation until the product peak reached the detector.

7.2.3. Single molecule assays

The protocols and the development of the DDAO-gal and FDG assays are described in detail in the results section. Single molecule assays based upon resorufin were as described in chapter 5.

7.3. Results

7.3.1. DDAO-gal single molecule assay

Figure 20 depicts the structure of DDAO-gal. The enzyme hydrolyzes the ether bond between C-1 on galactose and the DDAO aglycon thereby liberating the negatively charged DDAO (at experimental pH), and enzyme activity is monitored by measuring the production of DDAO. The 633 nm line of a 10 mW HeNe laser is used for excitation, and emission is collected through a 660BP10 filter to minimize detection of scattered laser light.

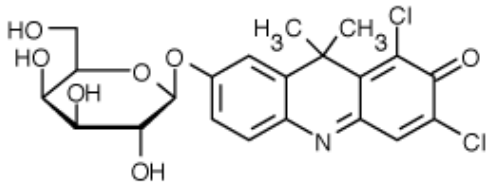


Figure 20. Structure of DDAO-gal.

The first step for evaluating a new substrate as a candidate for a single molecule assay is to evaluate its purity by injection and separation on an uncoated capillary with normal polarity. For this setup, the injection end of the capillary is held at a positive potential with EOF towards the detector, and because the capillary is uncoated, the magnitude of the EOF is sufficient to overcome the electrophoretic mobility of anions and they will also migrate towards the detector. The lower trace in Figure 21 is for an

injection of 1 x 10^{-8} M DDAO reference standard. The top trace shows the result for the injection and separation of 100 μM DDAO-gal in separation buffer. The small peak at ~ 300 s is the fluorescent signal produced by DDAO-gal, and the large peak at ~ 525 s with the same migration time as the DDAO suggests that the large peak in the substrate electropherogram is DDAO that is present as an impurity. Comparison of the area of the standard peak to the peak of the DDAO impurity in the substrate reveals that the DDAO-gal is 99.98% pure, but because DDAO is highly fluorescent, this slight impurity will significantly increase the background noise, and requires removal.

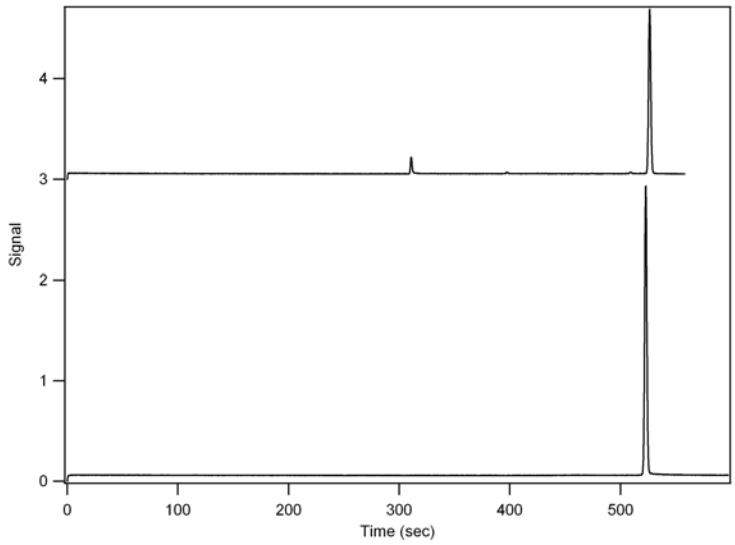


Figure 21. Electropherogram for DDAO and DDAO-gal. *Lower trace:* 1 x 10^{-8} M DDAO; *top trace*: 100 μM DDAO-gal. Both samples injected 5 s at 100 Vs^{-1}, separated at 400 Vs^{-1} on a 50 cm capillary.

Resorufin is a significant contaminant of RES-gal, but the resorufin impurity can be extracted from RES-gal dissolved in separation buffer by three rounds of washing with chloroform. Resorufin is soluble in chloroform but RES-gal, with its galactose moiety, is not [45]. The structural similarity of RES-gal and DDAO-gal (Figures 19 & 20) prompted an initial attempt to remove the DDAO impurity with chloroform. DDAO was

soluble in chloroform, but, to a lesser extent DDAO-gal was also soluble; this disqualified chloroform as a solvent candidate. An empirical approach was then used to identify an organic solvent that could separate DDAO-gal from DDAO. DDAO-gal was not soluble in highly non-polar solvents such as hexane and carbon tetrachloride, but neither was DDAO. Eventually toluene was identified as a suitable solvent. Washing 100 μM DDAO-gal in separation buffer with an equal volume of toluene using a vortex for one minute removed approximately 90% of the impurity, and two one minute washes with an equal volume of toluene removed 98% of the DDAO (Figure 22).

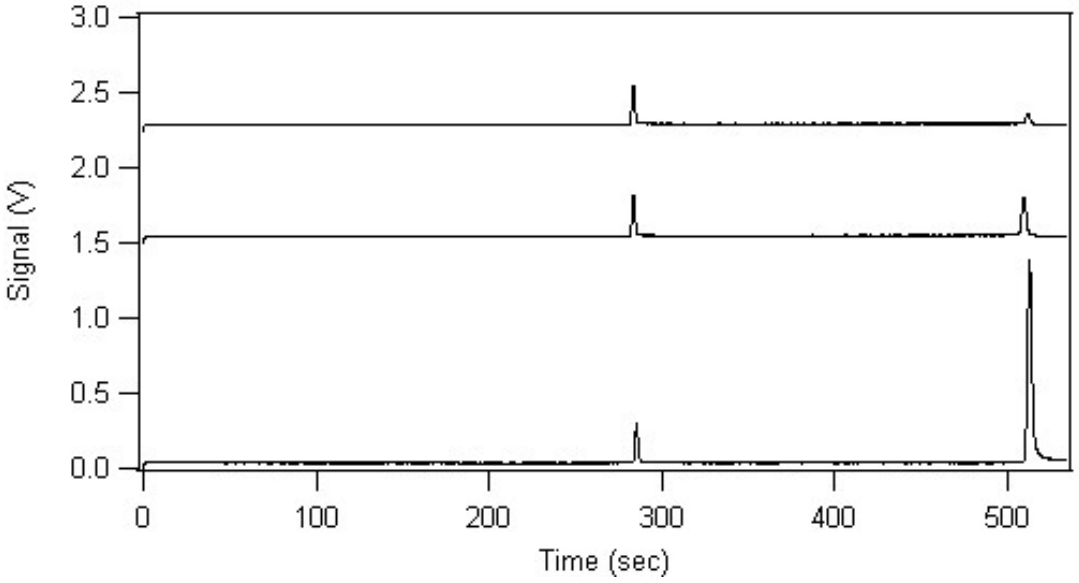


Figure 22. Electropherograms of DDAO-gal washed with toluene. *Lower trace:* 100 μM DDAO-gal in separation buffer injected 5s at 100 Vcm^{-1}, separated at 400 Vcm^{-1} on a 50 cm uncoated capillary; *middle trace*: 100 μM DDAO-gal extracted once with an equal volume of toluene; *upper trace:* 100 μM DDAO-gal extracted 2X with an equal volume of toluene.

The DDAO-gal was not soluble in toluene because its concentration was unchanged after two washes as determined by integrating the area of the substrate peak. Since the background hydrolysis rate of DDAO-gal is considerable, the amount of DDAO that is spontaneously formed during a typical enzyme assay negated a substantial portion of the

DDAO removed by the second toluene wash. For this reason, a single toluene wash was deemed sufficient.

The single molecule β-galactosidase assay using DDAO-gal as a substrate is similar to the RES-gal assay. A coated capillary was pressure-filled with washed substrate and dilute enzyme and then allowed to statically incubate for 5 minutes. During the incubation discrete pools of enzymatically formed DDAO accumulated around each enzyme molecule. Following incubation, the potential was applied and the pools of negatively-charged DDAO product were swept out of the capillary where they were excited by the laser beam and presented as peaks on the electropherogram. The new assay provides better signal to noise which allowed a reduction of the assay time to 5 minutes rather than the 15 minutes typically used for RES-gal. A coated capillary was used with reversed polarity (injection end negative) to eliminate the signal from substrate fluorescence. However, since this substrate produces much less signal than RES-gal, the step of coating the capillary and reversing polarity is less important.

Figure 23 shows an electropherogram for a single molecule assay of a commercially sourced (Sigma) *E. coli* β-galactosidase using the substrate DDAO-gal. In the blank a plateau is formed due to the presence of residual product present in the substrate as an impurity and that which formed non-enzymatically during the assay setup and incubation. In the presence of the enzyme, peaks are observed sitting atop this plateau. Each peak represents product formed from a single enzyme molecule.

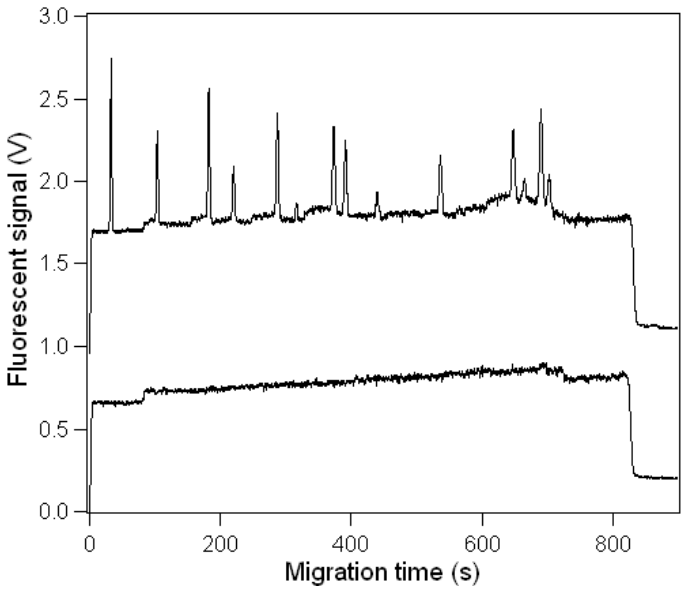


Figure 23. Representative electropherogram for the single β-galactosidase molecule DDAO-gal assay. *Bottom trace:* Five minute incubation of 100 μM DDAO-gal when no enzyme is present; *top trace:* five minute incubation of ~1 fM *E. coli* β-galactosidase from Sigma with 100 μM DDAO-gal.

To confirm that the product peaks are proportional to enzyme activity, double incubations were performed. Figure 24 shows the electropherogram for the double incubation of two molecules of the enzyme. In this assay the enzyme is first incubated with substrate, and then the enzyme is electrophoretically moved from the product it formed into fresh substrate and incubated again. A pair of peaks is formed for each enzyme molecule. The second peak is shorter and wider than the first. This is due to it having formed in the first incubation and therefore having more time for diffusion to broaden it. For all the individual molecules assayed, the areas of the two peaks were within 5% of each other, demonstrating the reproducibility of the assay.

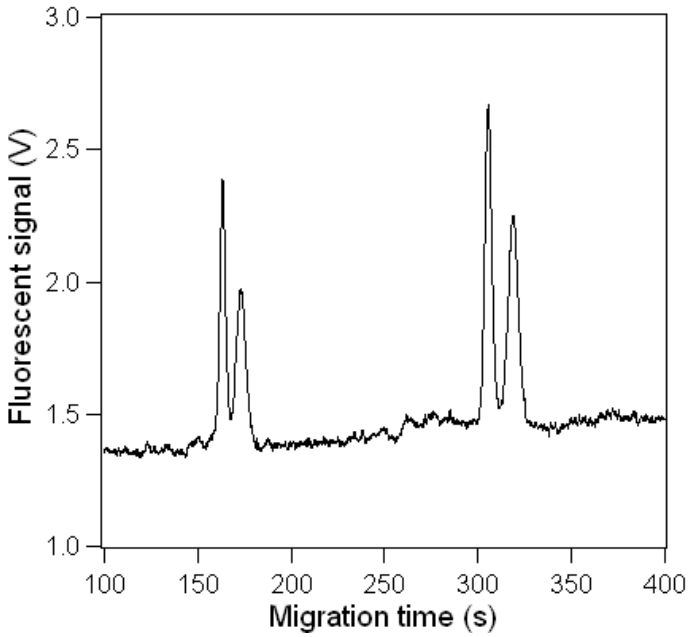


Figure 24. Double β-galactosidase incubation in 100 μM DDAO-gal. A section of the resultant electropherogram for two 5 minute incubations of two different *E. coli* β-galactosidase molecules. Enzyme mobilized into fresh substrate following first incubation by application of 400 Vcm^{-1} potential for 60 s.

Assays were performed on 195 molecules of the commercially sourced enzyme. The average turnover rate was 39,000 ± 18,000 reactions/minute/molecule of enzyme (mean ± σ). The large variation in the activity is not due to poor reproducibility of the assay, in fact the multiple incubations show the reproducibility is good, but rather it reflects the heterogeneous nature of the individual enzyme molecules.

7.3.2. FDG single molecule assay

The β-galactosidase fluorogenic substrate FDG has been used previously for bulk solution β-galactosidase CE-LIF assays [173]. The structure is shown in Figure 25, and unlike the three previous β-galactosidase substrates, FDG has two galactose moieties. The enzyme hydrolyzes the substrate to the fluorescent fluorescein-β-D-monogalactopyranose (FMG), and then a second hydrolysis reaction converts FMG to the highly fluorescent fluoroscein. Fluorscein is not a significant product until the

concentration of FMG is high enough to effectively compete with FDG for the enzyme's active sites. In the early 1990s there was a suggestion that FDG was hydrolyzed in a two step reaction directly to fluorescein on the same enzyme molecule by a channelling mechanism [174], but this hypothesis was repudiated by a careful examination of kinetic data [175], and by the monitoring the time course accumulation of the two fluorescent products by CE-LIF [173].

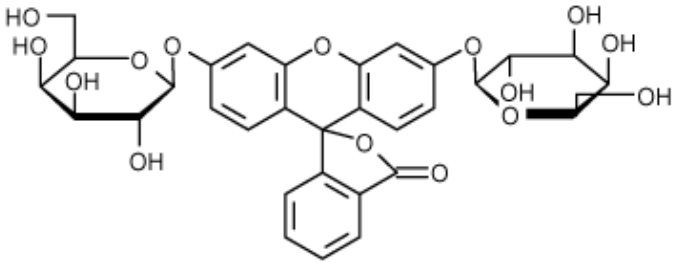


Figure 25. Structure of FDG.

The middle trace from figure 26 depicts an electropherogram for the separation of FDG on an uncoated capillary. The substrate FDG elutes at ~ 250 s and as was the case for the other substrates, the principal FDG impurity was its hydrolysis product FMG which elutes at ~ 350 s. The magnitude of the FMG peak is large and would contribute excessive background signal to a single molecule assay and therefore must be minimized. The FMG impurity in the substrate is not amenable to organic solvent extraction because of the presence of its galactose moiety, and therefore required a different strategy than was used for RES-gal and DDAO-gal. The pK_a for FMG is 4.9, and hence it carries a negative charge at the experimental pH (7.3); the substrate however is uncharged. This presented a basis for purification. FDG was dissolved in separation buffer and mixed with DEAE cellulose anion exchange beads to bind the negatively charged product

impurity. The mixture was centrifuged at 10 000 x *g* for 5 minutes, and the supernatant was removed and injected and separated on the capillary. The top trace from figure 26 shows that the anion exchange beads were highly effective at removing the FMG from the FDG.

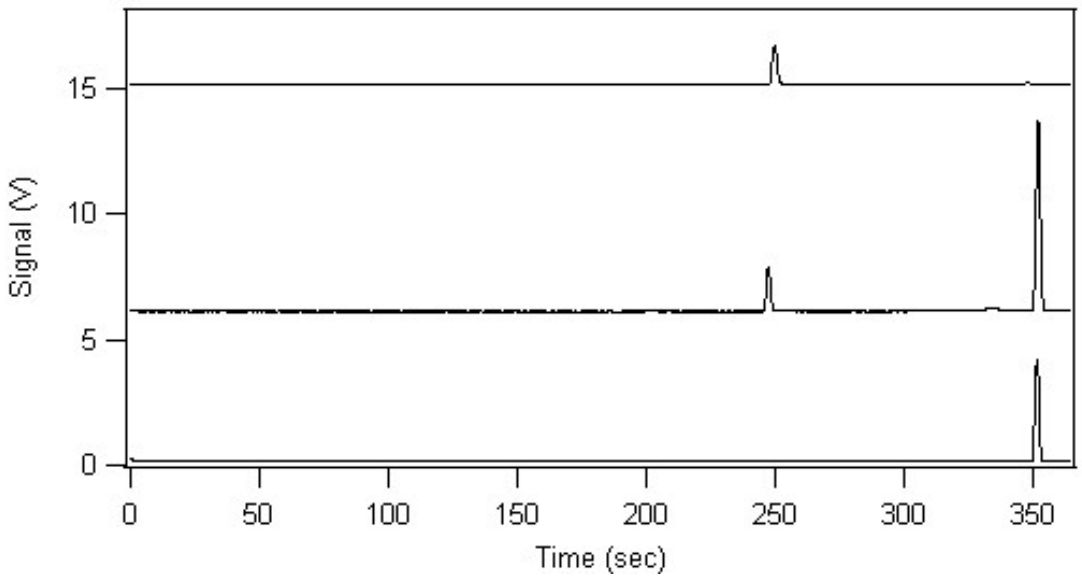


Figure 26. FMG extraction with DEAE cellulose. *Lower trace:* 2 x 10^{-8} M FMG in separation buffer; *middle trace:* 100 µm FDG in separation buffer; *top trace:* 100 µm FDG in separation buffer following DEAE cellulose extraction. All samples electrokinetically injected for 5 s at 100 Vcm^{-1} and mobilized at 400 Vcm^{-1} on a 40 cm uncoated capillary.

When β-galactosidase was mixed with DEAE extracted FDG and pressure injected into the capillary at a concentration previously determined to yield single molecule peaks, no individual enzyme product peaks were evident, and an increase in the enzyme concentration and assay duration still failed to produce any product peaks. Bulk assays indicated that substrate extraction with the beads did not inhibit the enzyme reaction. The CE-LIF assays did however exhibit an unusual characteristic--at the end of the FMG plateau (which is analogous to the resorufin and DDAO plateaus) there was a large peak that usually saturated the detector. Because β-galactosidase is negatively charged at the

experimental pH, and because a capillary coating does not fully mask the negatively charged capillary wall, this persistent product spike was attributed to individual enzyme molecules binding on DEAE fragments that adhered to the injection end of the capillary. During the incubation this collection of enzyme molecules would convert FDG into a concentrated pool of FMG that would saturate the detector at the end of the FMG plateau. Extensive efforts to remove the fragments of the DEAE from the substrate using 0.2 μM filters and centrifugation were unsuccessful. Under no circumstances was it possible to inject enzyme into the capillary with substrate that had been exposed to DEAE cellulose. To perform single molecule assays with FDG it was necessary to exploit the different electrophoretic mobilities of the substrate, product and enzyme.

The cartoon in Figure 27 depicts how the different mobilities were used to achieve single molecule detection. FDG, like the other substrates, is uncharged, while the hydrolysis product FMG and β-galactosidase are both negatively charged at pH 7.3.

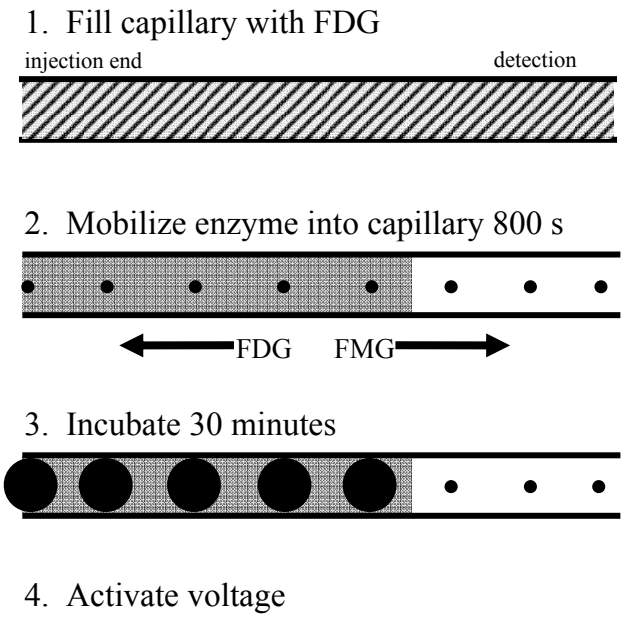


Figure 27. FDG single molecule assay. Capillary pressure filled with FDG (solid grey) contains FMG (diagonal lines) as an impurity. Mobilization for 800 s at 600 Vcm^{-1} drives FMG out of capillary and enzyme in; FMG product restricted to first 2/3 of capillary.

First, substrate was pressure injected into the PVP coated capillary for 5 minutes filling it with FDG and the FMG product impurity. The injection end of the capillary was then placed in a vial containing separation buffer and approximately 1 fM β-galactosidase and mobilized for 800 s at a field strength of 600 Vcm^{-1}(injection end negative). At the end of the mobilization the FMG has migrated out of the capillary; the enzyme has migrated along the full length of the capillary; and the neutral substrate has migrated back towards the injection end of the capillary at the same velocity as the EOF. However, because the EOF has been suppressed nearly 10 fold by the PVP coating, approximately 2/3 of the capillary from the injection end still contains substrate. The sample was then statically incubated for 30 minutes. During this period, the product accumulates around the individual enzyme molecules. Product is only formed in that region of the capillary

which contains both enzyme and substrate, which corresponds to approximately 2/3 of the capillary starting at the injection end. It is for this reason that no enzyme product peaks are ever observed in the first approximately 400 s of the 1200 s electropherogram.

Figure 28 shows the resultant electropherogram from a single incubation of individual molecules of β-galactosidase using FDG as substrate. In the blank a plateau is observed, as was with the assays using the other substrates, and in the presence of enzyme, peaks are seen sitting on top of this plateau.

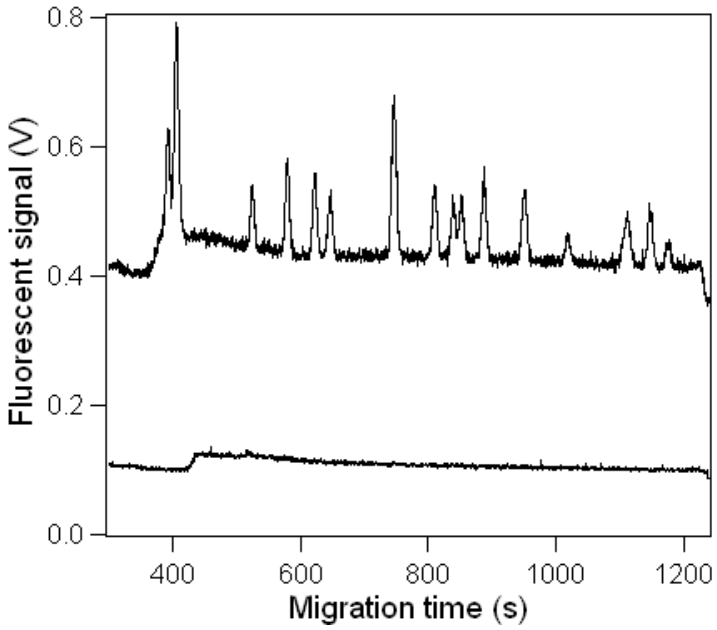


Figure 28. FDG single molecule β-galactosidase assay. *Lower trace:* 800 s injection at 600 Vcm^{-1} of 100 μM FDG followed by a 30 minute incubation and final mobilization at 400 Vcm^{-1} on a 40 cm PVP coated capillary. *Upper trace:* same as lower trace except includes approximately 1 fM *E. coli* β-galactosidase.

Figure 29 shows the resultant electropherogram from a double incubation of a single molecule of β-galactosidase, performed in a similar manner to that using the DDAO-based substrate. Pairs of peaks are seen for each enzyme molecule. Integration of peak areas showed the peaks of each set to be within 5% of each other.

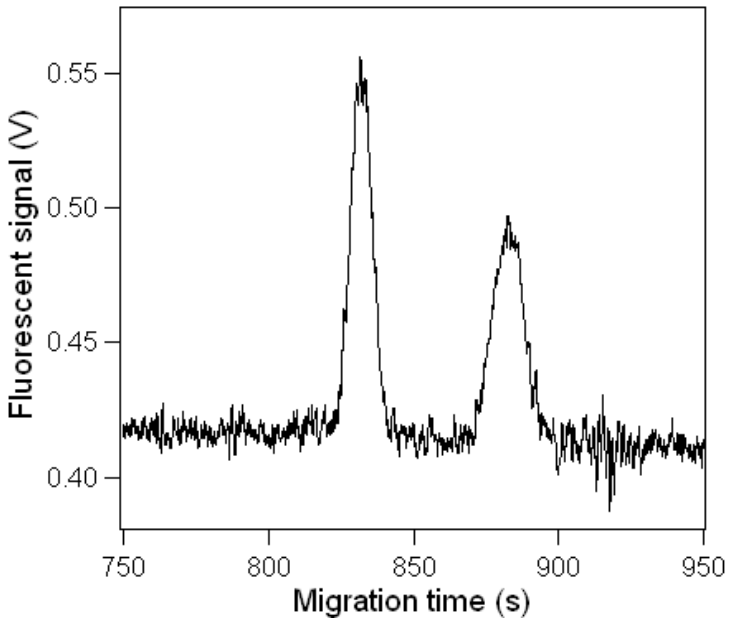


Figure 29. FDG double incubation electropherogram. A section of the resultant electropherogram from two 30 minute incubations of approximately 0.5 fM *E. coli* β-galactosidase separated by a 60 s electrophoresis period is shown.

At total of 196 molecules of commercially obtained purified β-galactosidase were assayed yielding an average activity of 4,900 ± 3,000 reactions/minute/molecule of enzyme (mean ± σ). Again, the relatively large range of rates obtained reflects the heterogeneity of individual enzyme molecules and not poor reproducibility of the assay.

7.3.3. Effect of aglycon on relative catalytic rate for β-galactosidase from 35321 and 8677

β-Galactosidase was expressed in *E. coli* strains ATCC 8677 and 35321 and single molecules assays were performed using substrates RES-gal, DDAO-gal and FDG. Approximately 200 molecules were assayed for each strain for every substrate. The average activities are shown in Table 10. Bulk solution assays were also performed using oNPG and the average activities were calculated using an absorptivity coefficient of 3500 $M^{-1}cm^{-1}$. As stated in the introduction, the concentration of active enzyme in the 8677

and 35321 samples was determined based on the average number of product peaks in the RES-gal and DDAO-gal single molecule assays, the capillary volume, and the dilutions used. The determined rates using oNPG were 107,000 ± 3,000 and 93,000 ± 2,000 reactions/minute/molecule of enzyme (mean ± σ) for the 8677 and 35321 strains respectively. The rates obtained with oNPG are somewhat different in their meaning than those obtained in the single molecule assays. In the oNPG assays, a very large number of enzyme molecules were assayed in bulk, and the average activities calculated based on the calculated enzyme concentration. The variation shown reflects the reproducibility of the measurement itself. No differences between individual molecules can be measured in such an assay. In the single molecule assays, the variation shown represents the heterogeneity of the enzymes themselves as well as the variation of the measurements, which according to the multiple incubation data shown here, and elsewhere with RES-gal [46], is relatively small.

Table 10. The average single molecule catalytic rates[a] of β-galactosidase from 8677 and 35321 for various substrates.

	Substrate (aglycon p*K*a)			
Source	oNPG (7.2)[b]	RES-gal (6.0)[c]	DDAO-gal (5.5)[c]	FDG (4.9)[d]
ATCC 35321	93,000 ± 2,000	98,000 ± 40,000 N =175	104,000 ± 47,000 N = 200	8,100 ± 3,500 N = 239
ATCC 8677	107,000 ± 3,000	53,000 ± 31,000 N = 190	44,000 ± 23,000 N = 173	3 300 ± 2,000 N = 212
8677/35321[e]	1.15	.54	.42	.41

[a] reactions per molecule^{-1}minute^{-1}
[b] oNPG aglycon p*K*a from [176], oNPG rates based upon triplicate assays
[c] aglycon p*k*a obtained from Molecular Probes
[d] aglycon p*K*a from [177]
[e] ratio of average activity of β-galactosidase from 8677 to 35321 for each substrate

Comparison of the relative rates of hydrolysis by β-galactosidase from the two strains of *E. coli* for the four substrates was performed by weighted least squares on a two-factor fixed effects ANOVA model using logarithmic transformed data. The level of confidence was 95%. The ratio of the average rate of the 8677 enzyme to the average rate of the 35321 enzyme using oNPG, RES-gal and FDG were found to be significantly different from one another. The ratio for DDAO-gal fell between that for RES-gal and FDG so that its confidence interval overlapped with theirs, but it was significantly different from that for oNPG. The enzyme from the 8677 strain showed 115, 54, 42 and 41% of the catalytic rate of that from the 35321 strain when using oNPG, resorufin-gal, DDAO-gal and FDG as substrates. The data obtained for the assay using RES-gal from this experiment was compared with the control data from the experiment with the two strains and the proteases (Chapter 5), and in both cases, the enzyme from the 8677 strain showed 54% of the average activities of the 35321 strain's enzyme. The average activities with DDAO-gal and FDG differ with the enzyme obtained from these *E. coli* strains is different to that obtained using the commercially sourced enzyme for the development of the assay. Differences in average activities between strains and commercial preparations have been reported elsewhere [48].

7.4. Discussion

7.4.1. DDAO-gal and FDG assays

The principal challenge for the development of new single molecule enzyme assays for CE-LIF, or any high sensitivity measurements, is to sufficiently minimize the background noise to allow detection of the signal above it. The optical and mechanical refinements that confer the high sensitivity detection to CE-LIF have been considered in

the literature review, and the discussion here addresses the properties of good single molecule assay substrates.

A useful fluorogenic substrate must meet several requirements for consideration as a candidate for on-column single molecule assays. First, because the substrate will be at a much greater concentration than the concentration of product produced by a single molecule, the substrate must produce minimal signal upon laser excitation. If this is not the case, strategies must be devised to prevent the substrate from occupying the excitation volume at the same time as the product. Second, the enzymatic transformation of the substrate to product must occur at a rate high enough to produce sufficient signal for detection within a reasonable time frame. The amount of signal also depends upon the fluorescent properties of the product molecule including its absorbance coefficient at the wavelength of the laser, and quantum yield. Third, if the product molecule is a weak acid and the protonated and deprotonated states have different spectral properties, the pK_a of the product molecule must be such that enzyme activity occurs at a pH that the product molecule is present in sufficient amounts of the detectable form. Fourth, highly purified substrates are preferable to minimize background fluorescence and to minimize the possibility of secondary reactions. Fifth, because substrates invariably contain product contaminants which increase the background signal, it is desirable that substrates and products have physical property differences that facilitate their rapid and efficient separation. And finally, if the enzyme is hydrolytic, as is the case with β-galactosidase, the substrate is subject to non-enzymatic hydrolysis; the rate of this hydrolysis must be relatively low under the conditions where the enzyme is active.

The substrates RES-gal, DDAO-gal and FDG are high quality substrates from Molecular Probes/Invitrogen and satisfy criteria number four. The deprotenated form of these fluorescent dyes is the fluorescent form, and in all cases the pK_a of the dyes is more than one pH unit below assay pH, thereby satisfying the third criteria. The quality of the substrates with respect to the first, second, fifth and last criteria were more ambiguous and are the basis for differentiating between their relative usefulness.

The single molecule assay based upon DDAO-gal is superior to the earlier RES-gal assay in every respect. First the fluorescence of the DDAO-gal substrate is minimal when compared to that of RES-gal (Figure 30). This eliminates the need to use coated capillaries, which is advantageous because it eliminates the daily coating requirement and reduces intra and inter day run variability due to variegated EOF that was the result of the coating process. Also, the viscous coating polymer invariably led to clogging of the sheath flow cuvette that can contribute to peak tailing—deterioration that eventually required cuvette disassembly and cleaning. More importantly, the compatibility of this substrate for single molecule assays on uncoated capillaries made possible novel assay strategies that will be presented in subsequent chapters.

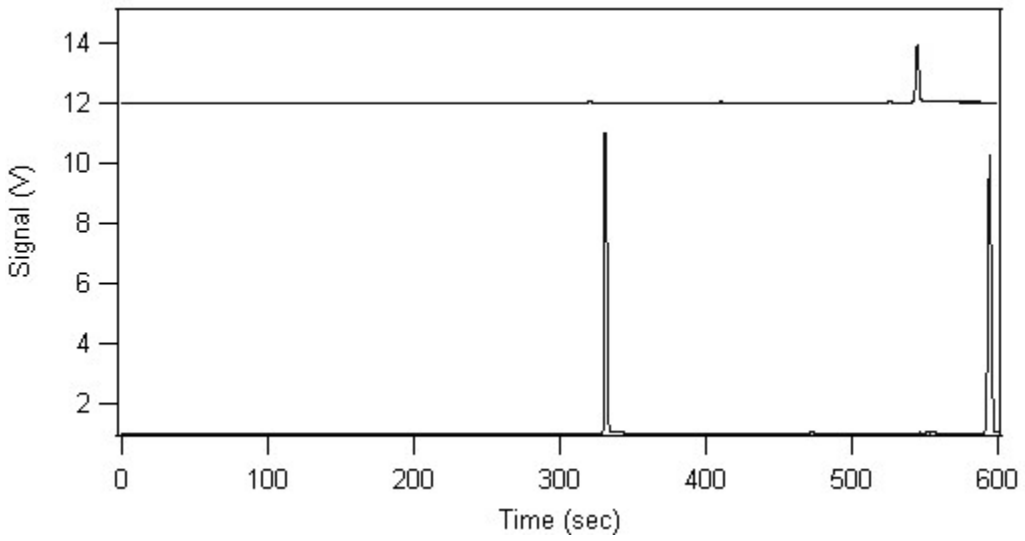


Figure 30. Comparison of DDAO-gal and RES-gal fluorescence. 100 μM sample of both substrates electrokinetically injected 5 s at 100 Vcm^{-1} and separated at 400 Vcm^{-1} on different uncoated 50 cm capillaries. *Lower trace:* RES-gal, peak at 330 s is RES-gal and peak at 600 s is the resorufin impurity; *upper trace:* DDAO-gal. The DDAO-gal peak at ~ 325 s is barely visible because of the scale of the electropherogram.

DDAO-gal is also superior to RES-gal on the basis of the second requirement because the catalytic rate is similar to RES-gal, but the laser for this assay has a higher output and produces easily quantifiable peaks with a 5 minute rather than 15 minute incubation. Moreover, subsequent experimentation revealed that excellent single molecule peaks can be attained with incubations as short as three minutes in substrate concentrations as low as 25 μM. If capillaries with 5 μM diameter were to be used, assay times could easily be reduced to less than one minute. Considering that this would reduce peak broadening due to diffusion, a greater number of separate incubations could be performed with the same molecule before it exited the capillary. This also allows a greater number of individual molecules to be assayed during a single run. With respect to the fifth consideration, effective removal of the DDAO impurity in DDAO-gal could be achieved with a single toluene wash. In contrast, extraction of resorufin from RES-gal required three one

minute washes with chloroform. The sole shortcoming of this substrate is the high background hydrolysis rate. This precludes daily preparation of substrate in bulk quantities and may limit its usefulness as a substrate for low activity enzymes. It was also found that it was best to aliquot the stock solution in DMSO into volumes less than 50 μL because larger volumes lasted longer and tended to accumulate sufficient water from the air which increased the amount of DDAO contaminant. However, the high background hydrolysis was turned to our advantage by using it as a proxy for the substrate front when measuring EOF on a coated capillary (Chapter 8).

FDG is a poor substrate with respect to the first, second and fifth criteria. Fluorogenic FDG is fluorescent (Figure 26) and therefore coated capillaries are required. The rate of its hydrolysis by β-galactosidase is about 10 fold lower than DDAO-gal and RES-gal, and consequently required considerably longer incubation times despite the higher laser power. Removal of the product impurity was not amenable to solvent extraction because of the presence of the second galactose and therefore required a cumbersome separation strategy. Also, the apparent mobility of FMG is low on a coated capillary that further lengthened the time required to perform a single run. Finally the assay utilizes an Ar^+ laser for excitation which is substantially more expensive, in terms of both initial cost and energy requirements, than HeNe laser types. As a result, the FDG assay generally is not particularly useful, and except for the purpose here for evaluating the role of the substrate aglycon, has found no further use in our laboratory.

7.4.2. Effect of aglycon substitution

The active site for β-galactosidase sits at the bottom of a deep pocket [104], and substrate binding proceeds through shallow and deep modes. Binding for the shallow

mode is located at the mouth of the active site and a several Å movement is required to bring the galactose moiety into the deeper mode that is necessary for cleavage of the glycosidic bond and the formation of the galactosyl enzyme intermediate [117]. There are specific interactions between the hydroxyls of the galactose moiety in both modes, but the interactions with the glucose aglycon are non-specific. It is for this reason that the enzyme can utilize a large number of substrates where the glucose unit is replaced by a variety of chromogenic or fluorogenic groups [91].

The proposed mechanism for the hydrolytic reaction catalyzed by β-galactosidase involves two steps after the substrate moves into the deeper binding mode. In the first step, Glu461 donates a proton to the glycosidic oxygen in concert with the transfer of the galactosyl group to Glu537. The pK_a of Glu461 may be modulated by interaction with the Mg^{2+} in the active site. The glucose unit, or in the case of non-physiological substrates the aglycon group, is released. In the second step the galactosyl group on Glu537 is transferred to a water molecule, releasing it [117]. One would predict, based on the proposed mechanism involving the acidic cleavage of the ether bond that the rate for the first step would increase with increasing acidity of the aglycon group. The second step is identical regardless of which substrate is used. Sinnott and Souchard [178] showed that k_{cat} for the reaction is broadly correlated with a decrease in the pK_a of the aglycon. Literature values for the pK_a's of oNP [176], resorufin and DDAO are 7.2, 6.0 and 5.5 (Molecular Probes, product information) respectively. The pK_a of FMG was determined to be 4.9, as measured by monitoring the change in its spectrum with varying pH. If aglycon acidity were the sole determinant of catalytic rate, hydrolysis of FDG

would occur at the highest rate and oNPG the slowest. The results presented here do not accord well with this prediction. Presumably other variables are more significant.

For the substrates studied here, aglycon bulkiness dominates the effect of aglycon acidity. oNPG is the smallest aglycon; RES-gal and DDAO-gal are both built from three fused rings, but DDAO-gal is bulkier with its methyl and di-chloro substituents; and FMG is the largest. The rate of substrate hydrolysis for β-galactosidase from strain 35321 does weakly follow the predicted relationship for decreasing pK_a of the aglycon with rates of 93,000 minute^{-1} for oNPG, 98,000 minute^{-1} for RES-gal and 104,000 minute^{-1} for DDAO-gal. However with FDG, which has the most acidic aglycon, there is a greater than 10-fold reduction of the catalytic rate to 8,100 minute^{-1}. It appears therefore that aglycon bulkiness attenuates the influence of aglycon acidity for β-galactosidase from strain 35321 as shown by the slight increase of activity as the aglycon becomes more acidic. However, for FDG the bulkiness of FMG supersedes the influence of aglycon acidity and the rate collapses. For the enzyme from strain 8677 the predicted effect of aglycon acidity fails completely with catalytic rate consistently decreasing with a decrease in aglycon pK_a. The rate of hydrolysis for oNPG, with the highest pK_a, is 107, 000 minute^{-1}, and for FDG which is the most acidic aglycon, the catalytic rate is 3,300 minute^{-1}. For the two substrates with the intermediate-sized aglycons—RES-gal and DDAO-gal, the effect of the increased bulkiness of DDAO compared to resorufin causes a decrease in the rate from 53,000 minute^{-1} for RES-gal to 44,000 minute^{-1} for DDAO-gal despite the increased acidity of DDAO. Aglycon bulkiness is possibly asserting its affect upon catalytic rate as steric clashes are generated as the substrate moves into the deeper binding mode necessary to affect catalysis. It has been previously proposed that the 10-

fold increase in k_{cat} when oNPG is substituted for pNPG is attributed to the added bulkiness that the nitro group in the *ortho* position compared to the *para* confers to the aglycon [117]. With respect to the difference in the effect that aglycon size has upon catalytic rate for enzyme from 35321 and 8677, the conversion of Asp280, which is located on a loop close to the aglycon binding site [104], to asparagine in strain 8677 likely produces a change in the geometry of the active site that results in a decreasing ability of the enzyme to accommodate substrates with larger aglycon groups.

7.5. Conclusion

Two new single molecule assays of β-galactosidase for CE-LIF using FDG and DDAO-gal as substrates were devised. The assay based upon DDAO-gal was shown to be superior to assays based upon FDG and RES-gal with respect to ease of preparation, sensitivity and assay duration. The compatibility of the DDAO-gal assay with uncoated capillaries provides flexibility for assay methodology. The influence of aglycon size upon the relative activity of β-galactosidase from strains ATCC 35321 and ATCC 8677 was investigated, and it was shown that substrates with bulkier aglycons were hydrolyzed more slowly, and that this effect is more pronounced in strain ATCC 8677 where a D280N substitution is present.

8. Single molecule separations

The data presented in this chapter was originally published in:

Nichols, E.R. and Craig, D.B. (2008) Measurement of the differences in electrophoretic mobilities of individual molecules of *E. coli* β-galactosidase provides insight into structural differences which underlie enzyme microheterogeneity. *Electrophoresis*, **29:** 4257-4269. Copyright permission to reuse the data here has been obtained.

8.1. Introduction

8.1.1. Overview

The current absence of techniques for the determination of molecular structure one molecule at a time confounds our understanding of the chemical basis for individual enzyme catalytic heterogeneity. Consequently, extensive speculation substitutes for experimental evidence regarding the actual structural differences between the molecules that produce different catalytic behavior. Examples of the hypotheses include the 'landscape' model for protein conformation dynamics where Xue and Yeung [2] apply the work of Freunfelder [179] to their results and attribute the observed differences to the occupation by otherwise identical enzyme molecules of different regions of the 'conformational landscape'. This explanation was challenged by the investigations of the Dovichi group where their efforts at driving individual alkaline phosphatase molecules into new conformational states with different catalytic behavior by thermal forcing were unsuccessful [3]. These authors suggested that structural differences such as glycosylation or spontaneous deamidation explained catalytic heterogeneity. There was a suggestion that limited proteolysis may contribute to enzyme heterogeneity [50], but the investigation of this possibility as a source for *E. coli* β-galactosidase in chapter 5 did not lend much support for it. Other authors suggest that there are no differences, and that the

observed differences are an artifact of sampling individual molecules for limited periods as they migrate through their conformational space [60]. If the individual molecules are structurally different, the differences must be subtle, because redissolved enzyme crystals exhibit the same range of catalytic heterogeneity as the enzyme source that was used to generate the crystal [46].

8.1.2. Single molecule electrophoresis:

Investigators have shown that it is possible to measure the apparent velocity of a single molecule of an analyte under an applied electric field, but there have been no systematic investigations of its structural significance, or its implications for single molecule heterogeneity. CE-LIF was used to measure the approximate sizes of individual double-stranded DNA molecules over a range of 2-49 kb that were intercalated with a fluorescent probe [180]. The migration time of individual molecules of the inherently fluorescent phycoerythrin was measured by CE-LIF and was used to investigate the role for molecular shot noise upon peak reproducibility [181]. The electrophoretic mobility of individual molecules of this protein was determined by measuring the time required for molecules to traverse a known distance between two lasers, but there was no reported attempt made to determine if the mobilities of individual molecules were reproducible [182]. An examination of the data collected by this study suggests that different molecules had different electrophoretic mobility but the authors did not comment upon it. The electrophoretic mobility of individual fluorescently labeled F-actin was measured by fluorescent microscopy and it was shown that mobility was dominated by the orientation of the rod-shaped molecules [183]. In this chapter, the combination of the high resolution capability of free zone capillary electrophoresis and

the single molecule detection of CE-LIF is used to measure the electrophoretic mobilities of individual molecules of *E. coli* β-galactosidase to investigate the possible structural basis for single molecule catalytic heterogeneity.

8.2. Methods

8.2.1. Single molecule assays, instrumentation and capillaries

All assays were based upon the β-galactosidase substrate DDAO-gal. Excitation was provided by the 633 nm line of a 10 mW HeNe laser and fluorescence was collected through a 660BP10 optical filter. All fused silica capillaries were 50 cm long with 10 μm internal diameter, and daily capillary preparation was as described in section 4.2.3. Separation and sheath buffers were 50 mM HEPES (pH 7.3), 1 mM $MgCl_2$ and 1mM citrate. Capillaries were used either uncoated, or, if EOF suppression was required, they were dynamically coated by pressure injection with either 0.5% (w/v) PVP or 0.07% (w/v) GS-6 in separation buffer. Substrate was prepared as described previously. Although the DDAO-gal assay was discussed in the previous chapter, significant changes were introduced that enabled the measurement of the electrophoretic mobility of individual molecules. How these assays were performed, and how the electropherograms were analyzed will be presented in detail in the results section.

8.2.2. β-Galactosidase sources

Wild type *E. coli* strain 35321 was cultured and β-galactosidase was induced, harvested and stored as described previously. Commercial enzyme was used as received from Sigma. The commercial enzyme migrated as a single band as determined by SDS-PAGE (Appendix II). Stock enzyme solutions from both sources were stored at – 20°C and diluted immediately prior to each assay.

8.3. Results

8.3.1. Reproducibility of single molecule mobility

The utility of electrophoretic mobility as a tool to investigate single molecule structural heterogeneity requires that the measurement of the mobility of any particular molecule be reproducible. This was assessed using two different methods. The first method involved multiple mobilizations of individual enzyme molecules randomly located within a capillary for 45 s. The second method mobilized enzyme molecules originating together at the injection end of the capillary for an extended period. These two methods will be identified as the short and extended separations respectively.

Figure 31 illustrates how the short separations with multiple incubations were performed.

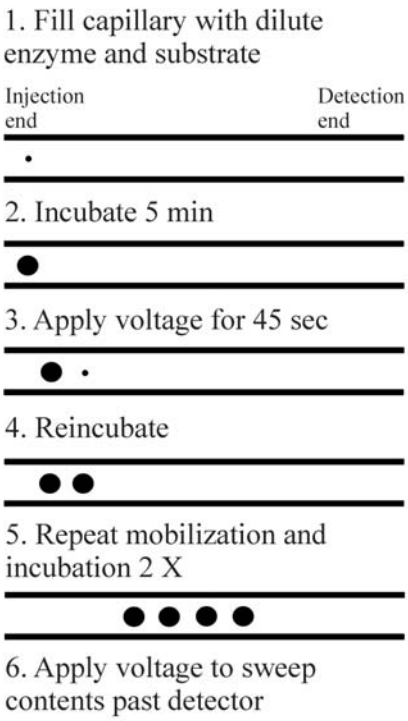


Figure 31. Schematic for short separations. Four incubations and three separations for an individual enzyme molecule from a random position in a capillary.

Approximately $1x10^{-16}$ M enzyme and 100 μM DDAO-gal in separation buffer were pressure injected for 5 minutes to fill the capillary followed by a 5 minute static incubation. At the end of the incubation the enzyme was subjected to an electric field of 400 Vcm^{-1} (injection end negative) for 45 s to move the enzyme molecule out of its product pool and into fresh substrate. This was followed by a second 5 minute incubation. The mobilization/incubation cycle was repeated two additional times yielding four distinct product zones for each enzyme molecule. Figure 32 depicts the resultant electropherogram from the repeated mobilization and incubation of an individual β-galactosidase molecule from a random location in a PVP coated capillary. The negatively charged enzyme and DDAO product molecules migrate towards the detector end of the capillary since EOF has been suppressed. β-Galactosidase has a higher mobility than the product molecules because it has a pI of approximately 5 and carries substantial negative charge at the buffer pH of 7.3. As a result, the peaks closer to the detector elute first, and are narrower and taller because they were formed later, and there was less band broadening due to diffusion during subsequent incubations.

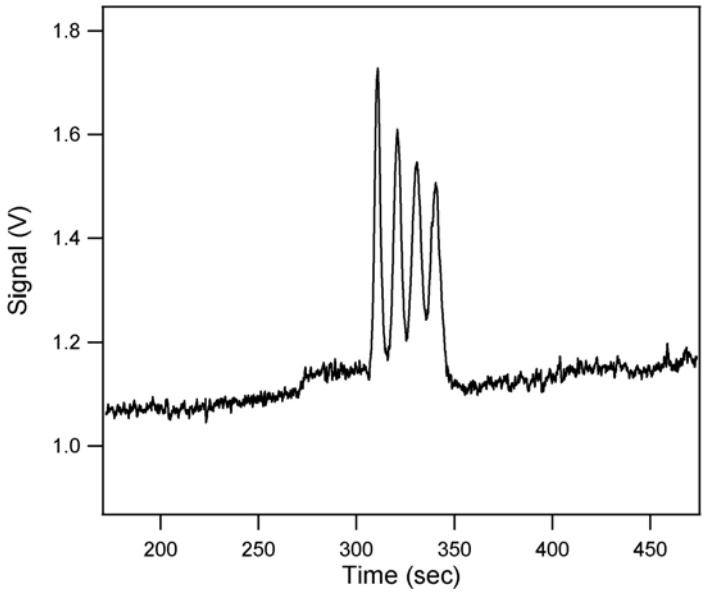


Figure 32. Electropherogram for three short separations. The capillary was filled with 100 μM DDAO-gal and ~ 1 x 10^{-16} M β-galactosidase. Enzyme was incubated in the substrate and then moved by an applied field of 400 Vcm^{-1} (injection end negative) for 45s to a new location in the capillary and incubated for 5 minutes. The separation/incubation cycle was repeated two more times. The capillary contents were swept past the detector by the application of a 400 Vcm^{-1} electric field.

The areas for all the peaks should be identical if enzyme activity is reproducible, and in all cases the variation was found to be less than 5%. The similar time intervals between the peaks indicate that the enzyme molecule migrated with the same average velocity for each mobilization. When comparing two different sets of peaks, the difference in the spacing indicates that different β-galactosidase molecules migrated at different velocities. The relative velocity of any individual enzyme molecule compared to the DDAO it produced was calculated as follows:

$$v_{rp} = \frac{\text{time between peaks} + t_s}{t_s} \tag{8.1}$$

where v_{rp} is the velocity of the enzyme molecule relative to the product molecules, time between peaks is the time between any two adjacent peaks produced by the same enzyme

molecule, and t_S is the mobilization time. Sixty-five individual ATCC 35321 β-galactosidase molecules were assayed in this manner. The average variation of apparent velocity for an individual enzyme mobilized multiple times on a capillary was 0.82%, and indicates that a mobility measurement for any particular molecule was reproducible. Relative to the DDAO, the molecules with the highest velocity migrated at 1.17 times the mobility of the molecules with the lowest velocity. The average variability of mobility for a single enzyme molecule was less than 5% of the range observed for the different molecules.

Figure 33 illustrates how the extended separations were performed using a PVP coated capillary.

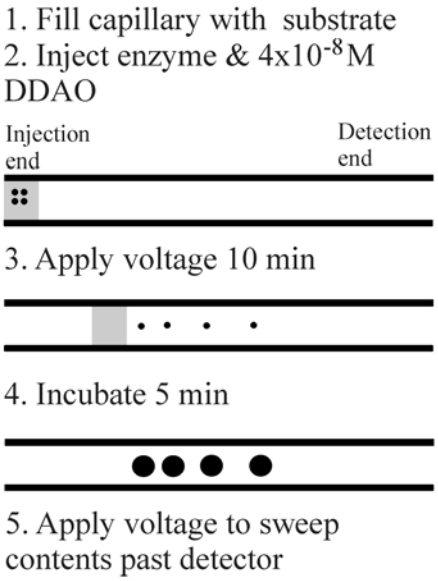


Figure 33. Schematic for extended separation. Ten minute separation and 5 minute incubation of 4 molecules in a substrate filled capillary. Grey rectangle represents injected DDAO standard. Small circles are enzyme, and large circles are pools of product formed by enzymatic hydrolysis.

The capillary was filled by pressure injection with 100 μM DDAO-gal for 5 minutes. A solution of approximately $1x10^{-13}$ M β-galactosidase and $4x10^{-8}$ M DDAO reference standard was introduced into the capillary by electrokinetic injection at 100 Vcm^{-1} (injection end negative) for 7 s. The injection end was then placed into separation buffer followed by electrophoresis for 10 minutes at 400 Vcm^{-1} (injection end negative) to drive the enzyme molecules into the substrate-filled capillary. At the end of the 10 minute mobilization, the enzyme molecules were statically incubated for 5 minutes followed by the reapplication of the 400 Vcm^{-1} electric field until the product peaks were swept past the detector. Figure 34 depicts two different runs for separations of this type and a blank.

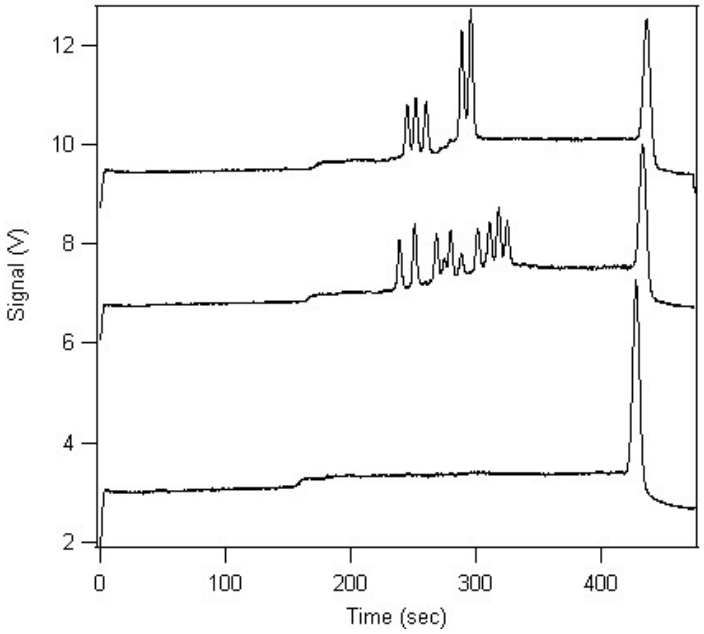


Figure 34. Electropherograms for three extended separations. The traces have been offset for clarity. *Lower trace:* Blank. The capillary was pressure filled with 100 μM DDAO-gal in separation buffer. A $4x10^{-8}$ M DDAO reference standard in separating buffer was injected for 7 s at 100 $Vcm^{-1}$and mobilized for 10 minutes at 400 Vcm^{-1}, followed by a 5 minute static incubation and a 400 Vcm^{-1} mobilization. The peak at ~450 s is the DDAO reference standard. *Middle trace and upper trace:* same as lower trace except a ~$1x10^{-13}$ M solution of β-galactosidase was injected with the DDAO reference standard. The peaks distributed from ~240 to 330 s are pools of DDAO product formed by individual β-galactosidase molecules.

The electropherograms shown do not include the 10 minute initial mobilization. Therefore the actual time of mobilization required to reach the detector is the migration time shown in the electropherogram plus 600 s. As previously, the enzyme migrated more quickly towards the detector end of the capillary during the 10 minute mobilization compared to the reference standard. The large peak at approximately 450 s in each run is the injected DDAO reference standard. The smaller peaks from approximately 230 s to 330 s are pools of DDAO formed by individual β-galactosidase molecules during the 5 minute incubation. These pools required different amounts of time to reach the detector after the voltage was reactivated because different individual enzyme molecules had migrated different distances during the 10 minute separation.

Figure 35 depicts a series of electropherograms for extended separations of a thousand-fold concentration range of enzyme to demonstrate that single molecules were detected. The principal indicator of single molecule events occurs when the number of peaks, but not the size of peaks, is reduced as the concentration of analyte is reduced [3]. The third trace through the seventh trace show fewer, but not smaller peaks as the dilution of the enzyme stock is increased. The peaks from the third and fourth runs appear larger because the concentration of enzyme is high enough that a number of enzyme molecules occupy the same region of the capillary and their respective pools of product merge.

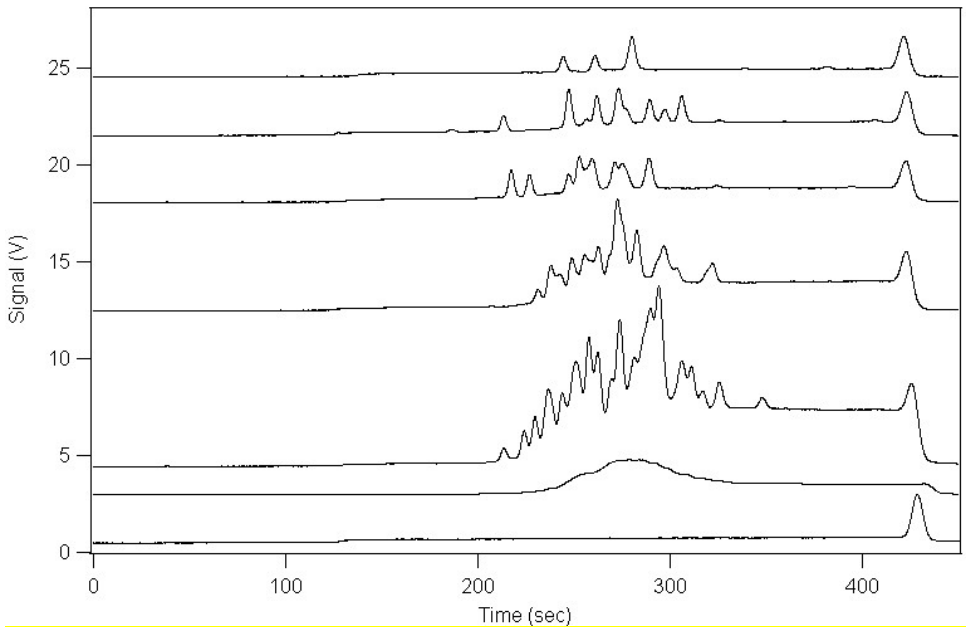


Figure 35. Extended separations of β-galactosidase at varying dilutions. Enzyme and $4x10^{-8}$ M DDAO were injected into a capillary pre-filled with 100 μM DDAO-gal for 7 s at 100 Vcm^{-1} mobilized at 400 Vcm^{-1} for 10 minutes followed by a 5 minute static incubation and second mobilization *Lowest trace:* Blank. *Second trace:* 5 x 10^5 dilution (note PMT bias lowered to avoid saturating the detector). *Third trace:* 3.33 x 10^7 dilution. *Fourth trace:* 6.67 x 10^7 dilution. *Fifth trace:* 1.25 x 10^8 dilution. *Sixth trace:* 1.67 x 10^8 dilution. *Seventh trace:* 5 x 10^8 dilution. Traces have been offset for clarity.

This figure also illustrates that the electrophoretic mobility range of the individual molecules matches that of a bulk sample (second trace).

The second basis for the justification that these are single molecule events rests upon the calculation of the required dilution to yield single molecules agrees with the dilution that was necessary to produce single peaks for the short separations. The time required for a DDAO molecule to reach the detector when mobilized at 400 $Vcm^{-1}$was approximately 1050 s. In other words a total of 4.2 x 10^5 Vs^{-1} were required to fill the capillary with DDAO. A 7 s, 100 Vcm^{-1} injection would mean that the injection volume of DDAO was only 1/600 of the capillary volume. Ignoring the difference of mobility for the enzyme and DDAO [88], the dilution factor required to put the same number of

enzyme molecules in the capillary for the extended separation would be 600 fold less than was required for the short separations. To place six molecules on average in a capillary for the short separation required a 2 x 10^{11} fold dilution of enzyme. Therefore a 3.33 x 10^{8} fold dilution would be required to assay six molecules for the extended type. The electropherograms in Figure 35 correspond reasonably well to this prediction.

Calculation of the mobility of individual enzyme molecules was based upon the simultaneous injection of the enzyme and the DDAO reference standard. The apparent velocity of any particular enzyme molecule could be calculated based upon its relative velocity to the reference standard, and the apparent velocity of the standard:

$$v_{app} = \frac{t_{DDAO} - t_{pk} + t_{sep}}{t_{sep}} \text{ x } \frac{L}{t_{DDAO} + t_{sep}} \tag{8.2}$$

where v_{app} is the apparent velocity (unadjusted for EOF) of an enzyme molecule in units of cm s^{-1}; t_{DDAO} is the electropherogram time required for the injected reference standard to reach the detector in s; t_{pk} is the electropherogram time required for any particular peak to reach the detector in s; t_{sep} is the total separation time (600 s); and L is the length of the capillary in cm. The first quotient from equation 8.2 determines the relative mobility of the enzyme molecule to the DDAO standard, and the second quotient from the equation is the apparent velocity of the DDAO molecules

To demonstrate mobility reproducibility for extended separations, a variation of the separations was performed where the enzyme molecules were mobilized three times with a 5 minute incubation following each mobilization. (Figure 36) The duration, in order, of the three mobilizations was 2, 3.5 and 5 minutes, and after the third incubation the product pools were swept past the detector. This extended separation with multiple incubations was otherwise identical to that for the single extended separation with the

exception that the enzyme concentration in the injected plug was approximately $5x10^{-14}$ M.

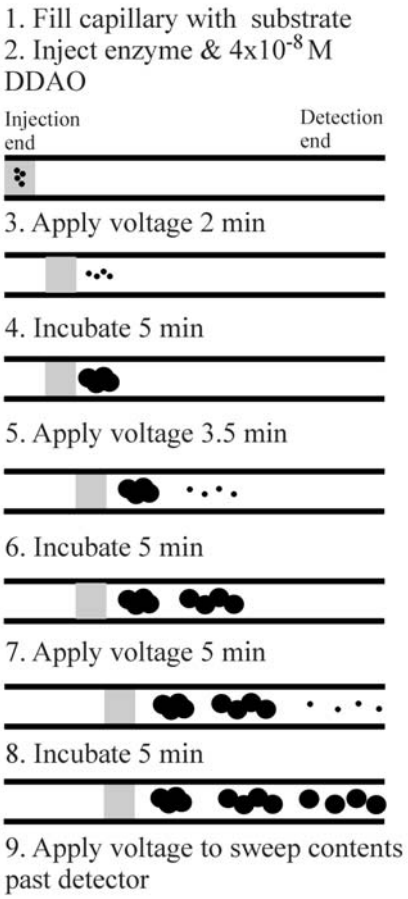


Figure 36. Schematic for multiple extended separations and incubations. Two minute, 3.5 minute and 5 minute separation of 4 molecules. Each separation followed by a 5 minute incubation. Grey rectangle is injected DDAO standard, small circles are individual enzyme molecules, and large circles are pools of product formed by enzymatic hydrolysis.

Enzyme molecules and the DDAO reference standard were mixed together at the injection end of the capillary prior to the separation. When the voltage was activated the enzyme molecules migrated faster than the reference standard plug and slightly separated from each other during the first two minutes of electrophoresis. During the incubation, their respective product pools intermingled because of the close proximity of the individual enzyme molecules. The second application of voltage caused the enzyme

molecules to migrate ahead of their respective pools of product and further separate from each other. Repetition of this separation/incubation process should produce three clusters of peaks. The peaks that appear first on any electropherogram were from the final incubation.

If different individual molecules have different but reproducible catalytic activity and mobility, the total area under a cluster of peaks should be equivalent for each of the incubations, and mobilities of the individual molecules should be constant, although different from each other. The electropherogram depicted in Figure 37 shows the results of three different runs performed in this manner.

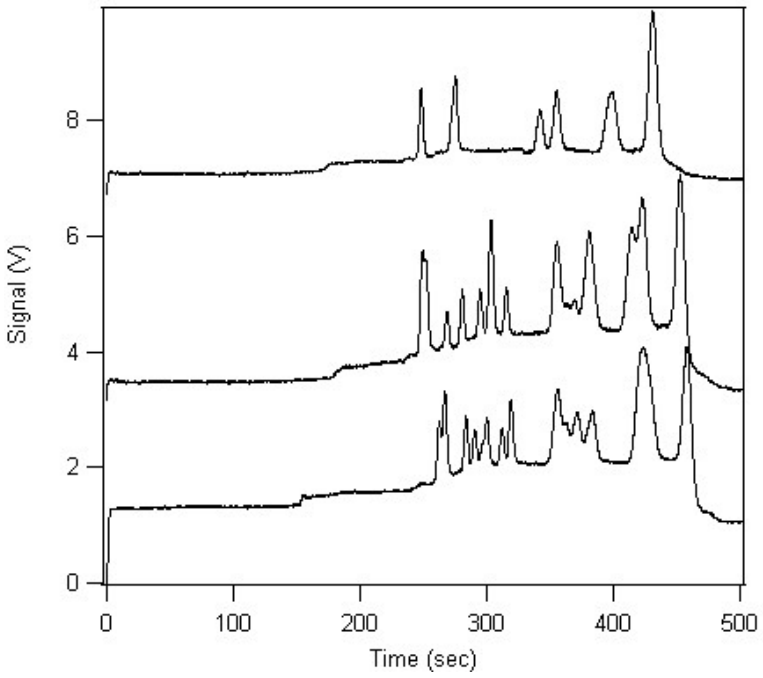


Figure 37. Electropherograms for multiple extended separations and incubations. *Lower trace:* the capillary was pressure filled with 100 μM DDAO-gal in separation buffer. A $4x10^{-8}$ M DDAO and $5x10^{-14}$ M β-galactosidase solution was injected for 7 s at 100 Vcm^{-1} immediately followed by mobilizations of 2, 3.5 and 5 minutes at 400 Vcm^{-1}. Mobilizations followed by 5 minute static incubation and then a final 400 Vcm^{-1} mobilization to sweep contents past detector. The peak at ~450 s is the DDAO product standard. The 3 clusters of peaks for each trace at ~300, 375 and 420 s are pools of product formed by the same enzyme molecules after a total of 10.5, 5.5 and 2.0 minutes of total elapsed mobilization. *Middle trace:* same as lower trace. *Upper trace:* same as lower trace except the concentration of enzyme solution is ~ $3x10^{-14}$ M.

In all cases the sum of the peak areas for each of the three incubations for any particular run were found to vary by less than 5%. It is evident that the unresolved features of the initial 2 minute separation are revealed in the pattern of the peaks in the subsequent incubations as the individual enzyme molecules become separated along the capillary. Peak analysis indicated that the apparent velocity for any particular individual enzyme molecule that could be reliably resolved varied by less than 1%. This result, considered in conjunction with the reproducibility of the short 45 s separations, indicates that individual β-galactosidase molecules have distinct, different, and reproducible velocities in a capillary. Moreover, it indicates that a single velocity measurement provides an accurate estimation of the velocity of any particular molecule. It appears therefore that different individual β-galactosidase molecules have different electrophoretic mobility.

8.3.2. Influence of well-known sources of peak broadening in CE

The previous section established the reproducibility of the single molecule β-galactosidase separations, but before proceeding to a further analysis of the significance of these results, it necessary to discount any effect upon the separations by the well-established extracolumn processes that contribute to the broadening of analyte peaks in free zone CE. These processes include variances from injection width, diffusion, Joule heating, conductivity, hydrodynamic flow and adsorption [83]. Intuitively it is difficult to envisage these generating the measured differences in velocity because, apart from analyte adsorption, their effects would also be evident in the reference standard peak which is clearly narrower than what is observed for the range of individual enzyme molecules.

The width of the injection plug can be calculated by the method of Huang *et al.* [88], and for the separations here it is approximately 0.8 mm. For comparison, the enzyme peaks for a typical 10 minute run span approximately 5 cm. The width of the injection plug is 0.16% of the length of the capillary, and broadening effects are minimal if the injection zone is less than 1% of the capillary length [184].

The influence of diffusion for band broadening is proportional to the diffusion constant of the analyte and the duration of the separation. The diffusion constants for β-galactosidase and DDAO are 3.3×10^{-11} m^2s^{-1} [130] and 5×10^{-10} m^2s^{-1} [46] respectively. Inspection of the DDAO reference standard peak width (which would be more susceptible to the influence of diffusion) shows that even for the extended separations this influence was small.

The application of high voltage across the capillary results in current-generated heat called Joule heating. Dissipation of this heat occurs from the surface of the capillary and results in the generation of a radial temperature gradient. This affects the mobility of analytes because the viscosity of the separating buffer will be lower in the warmest regions of the capillary, and therefore apparent enzyme velocity will be the highest in the center of the capillary. The effect for this on the enzyme mobility variations is presumably minimal because small radius capillaries dissipate heat quickly, analytes are well-mixed radially [67], and because the short multiple separations of an individual molecule show mobility reproducibility. At the end of a five minute incubation for a short separation, it is extremely improbable that the enzyme molecule would have diffused back to a similar position with respect to the center of the capillary that it occupied during a previous mobilization.

Conductivity differences between an analyte zone and the separating buffer can produce distortions of peak shape commonly resulting in fronting or tailing. The DDAO peaks that each enzyme molecule produces during the incubation show little evidence of distortion, and it follows therefore that no enzyme 'peak' fronting or tailing is present. Even if moderate DDAO peak distortion was present, one of the approaches to curtail this extracolumn effect is to minimize the concentration of the analyte [83]. Considering that single molecule analysis requires sub-pM analyte concentrations, this is not a concern.

Hydrodynamic flow arises because of a difference in the height of the buffer level at the injection end of the capillary relative to the sheath flow level at the detection end. Hydrodynamic flow contributes to band broadening because this flow is parabolic. This effect was minimized by adjusting the height of the sheath flow level and monitoring the effect upon peak size for a 5 s siphon-only injection of DDAO. The likelihood that this process contributed to enzyme separability is very low because of the reproducibility of the mobility for any particular molecule.

Analyte adsorption imparts chromatographic qualities to CE separations and intuitively this is the most likely source for the differences in the mobility of the individual molecules. The possible effects of this were extensively investigated and will be considered separately in the section below that addresses the role of capillary coating.

8.3.3. Measuring EOF in a coated capillary

To determine if the differences in mobility of individual enzyme molecules were attributable to exogenous factors such as capillary lumen irregularities, coating polymer-enzyme interactions, or unspecified intermolecular interactions between the enzyme molecules and components of the cell lysate, it was necessary to assess the effects of

source, coat and separation method upon the observed separation heterogeneity. To meaningfully compare results required the determination of the actual electrophoretic mobility for every enzyme, which in turn necessitated a method to accurately measure EOF. The apparent velocity of an analyte during a CE separation is the sum of its electrophoretic velocity and the EOF.

$$v_{app} = v_{act} + v_{eof} \tag{8.3}$$

The classical method for the measurement of EOF in CE is to use a molecule that is neutral at the separation pH and can be readily detected by the experimental apparatus [68]. The methods described below enable EOF to be measured using the intrinsic components of the single molecule assay for each run. This method was efficient, and any variation in EOF from day to day, from capillary coat variability, or from coating degradation during any particular day, was corrected for.

The velocity of the EOF for any particular extended separation run on a coated capillary was determined by exploiting the non-enzymatic background hydrolysis of the uncharged substrate, DDAO-gal to DDAO (Figure 38). Prior to the separation, the capillary was pressure filled with the substrate, but during the 10 minute separation period the neutral substrate migrated slowly away from the detector at the same speed as the EOF which had been suppressed due to the coat. After the separation there was a zone at the detection end of the capillary that contained only separation buffer. During the 5 minute incubation a small fraction of the substrate that was present behind this zone (towards the injection end), spontaneously hydrolyzed to the fluorescent DDAO. Therefore, at the end of the incubation a short zone contained only separating buffer, and behind this zone the remainder of the capillary was filled with substrate and a low

concentration of DDAO molecules. When the voltage was reapplied the negatively charged DDAO product migrated towards the detector and a baseline shift was observed that corresponded to the boundary between the two zones.

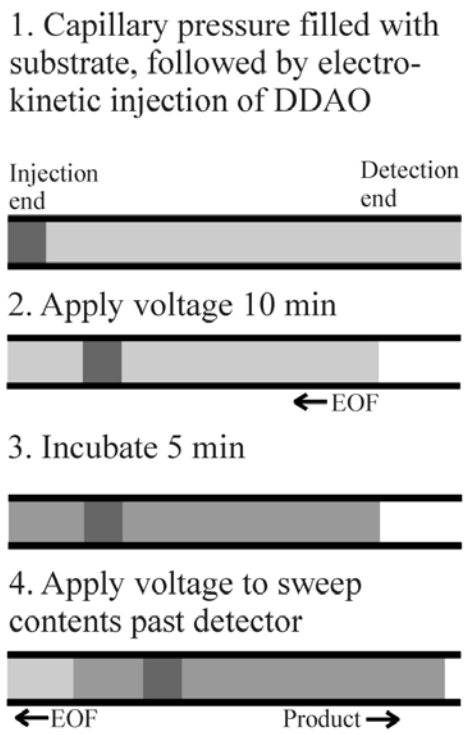


Figure 38. Measuring EOF in a coated capillary. During the 10 minute mobilization the neutral substrate (light shaded region) migrates towards the injection end of the capillary at the same velocity as the EOF leaving a short region at the detector end of the capillary free of substrate. The dark shaded rectangle represents injected DDAO reference standard. During the 5 minute incubation low level spontaneous hydrolysis of substrate to product (medium shaded region) occurs. Reapplication of electric field drives product standard and spontaneously formed product towards the detector. Time required for the product front to reach detector indicates the distance the substrate moved during the original separation.

The velocity of the EOF can be calculated as:

$$v_{eof} = \frac{t_b \text{ x } L}{(t_{DDAO} + t_{sep}) \text{ x } t_{sep}} \tag{8.4}$$

where t_b is the electropherogram time required for separation buffer/spontaneous substrate hydrolysis boundary to reach the detector, and the rest of the notation is as previously described. The accuracy of this approach to determine EOF was confirmed by two different methods. First, the magnitude of the baseline shift was proportional to incubation time; and second, it was possible to measure the EOF directly because of the slight fluorescence of the substrate. The time required for the signal to return to baseline on a coated capillary filled with substrate set at normal polarity (injection end positive) was equal to EOF, and EOF measured this way yielded a result similar to the baseline shift method. Adjusting each run for EOF in this manner reduced the calculated inter-run variability of the actual velocity of the DDAO to less than 0.5 percent. The actual electrophoretic mobility of individual β-galactosidase molecules was then calculated as:

$$\mu = \frac{v_{app} - v_{eof}}{E} \tag{8.5}$$

where v_{app} is the apparent velocity of a single β-galactosidase molecule, v_{eof} is the velocity of EOF, and E is the strength of the applied electrical field in Vcm^{-1}. Note that v_{app} is assigned a negative value because it oriented opposite to the EOF.

It was not possible to measure the EOF for the short separations directly, and EOF was based on the information obtained from the extended assays. For the short separation method the peaks sit upon a plateau of DDAO that is present because of incomplete extraction and background substrate hydrolysis. The time for the plateau to pass the detector indicates how much time DDAO molecules that were present at the injection end of the capillary required to migrate the length of the capillary under the applied voltage. The apparent velocity of DDAO molecules was calculated based upon the length of the capillary. The data for EOF and the apparent electrophoretic velocity of

the DDAO product standard on a PVP coated capillary obtained from the extended separation experiments was used to establish a relationship between the apparent velocity of DDAO and EOF. The EOF for the short separations could be calculated on the basis of this relationship.

8.3.4. Effect of the separation method

To test if the differences of single enzyme molecule migration were inherent to the molecules, or were artifacts attributable to the particular location of an enzyme molecule in the capillary during a separation, the electrophoretic mobility of individual molecules were compared for the short and extended mobilization methods. Table 11 shows the results for the separation of 163 and 119 purified individual β-galactosidase molecules from the short and extended mobilizations respectively. The highest electrophoretic mobility for a single β-galactosidase molecule measured by the short separation technique was -1.86×10^{-4} $cm^2V^{-1}s^{-1}$, and for the extended separation technique the highest value was -1.85×10^{-4} $cm^2V^{-1}s^{-1}$. The lowest measured electrophoretic mobilities for the short and extended techniques were -1.61×10^{-4} $cm^2V^{-1}s^{-1}$ and -1.63×10^{-4} $cm^2V^{-1}s^{-1}$ respectively. The mean electrophoretic mobility for both the short and extended mobilizations was -1.75×10^{-4} $cm^2V^{-1}s^{-1}$. The distribution of electrophoretic mobilities of 119 individual β-galactosidase molecules determined by the extended separation method is shown in Figure 39. Similar distributions were obtained for all methods used to measure mobility in this study. A two-tailed T-test indicated that there was no statistically significant difference ($p = 0.84$) for the electrophoretic mobility of a large sample of individual β-galactosidase molecules between the two techniques. The relative range of measured electrophoretic mobilities, which was calculated as the

ratio of the electrophoretic mobility of the fastest to the slowest molecule, was 1.15 and 1.14 for the short and extended mobilizations respectively.

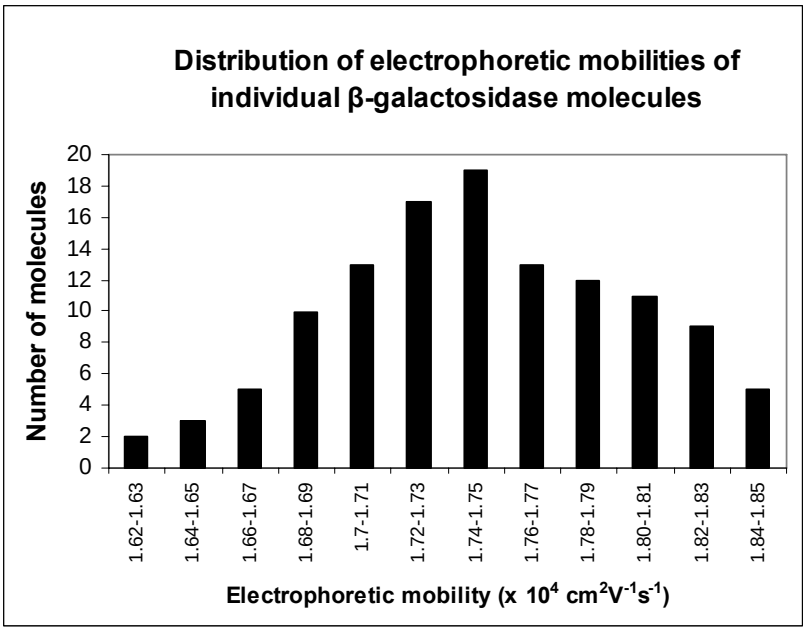


Figure 39. Distribution of β-galactosidase electrophoretic mobilities. Mobilities were determined by the extended separation method.

The samples of individual β-galactosidase molecules exhibit reproducible electrophoretic heterogeneity independent of how the molecules were introduced into the capillary, where within the capillary the mobilization occurred, or how long they were mobilized. However, because the capillaries used were coated, it remained possible that the heterogeneity could be an artifact from unspecified interactions between the enzyme molecules and the coating polymer.

Table 11. Summary of electrophoretic mobilities.

				Electrophoretic mobility ($\times 10^4$ $cm^2V^{-1}s^{-1}$)			Relative range
Enzyme Source	Separation method	Coat	N	Max	Min	Mean	(max/min)
Sigma	short	PVP	163	– 1.86	– 1.61	– 1.75 ± 0.05	1.15
Sigma	extended	PVP	119	– 1.85	– 1.63	– 1.75± 0.05	1.14
Sigma	extended	GS-6	137	– 1.87	– 1.63	– 1.77± 0.05	1.14
Sigma	extended	uncoated	117	– 1.94	– 1.63	– 1.81±0.06	1.19
ATCC 35321	short	PVP	141	– 1.85	– 1.60	– 1.74± 0.06	1.16

8.3.5. Effect of capillary coat

8.3.5.1. Effect of coat overview:

There are several mechanisms by which the observed individual enzyme molecule electrophoretic velocity heterogeneity could be attributed to the capillary coat that was applied to suppress EOF. An enzyme molecule could have become entangled in the coating polymer with the result that its migration down the capillary was impaired. If different enzyme molecules became entangled to different degrees, they would have exhibited different apparent mobility. It is difficult to consider this a probable explanation because of the multiple mobilization experiments. Because individual enzyme molecules exhibited reproducible mobility, this would have required different enzyme molecules to become entangled to different degrees, which is conceivable, but that this variable entanglement was consistent for a particular molecule for every mobilization, which is much less likely. A more plausible possibility is that coating polymer from either the separation buffer or from coat degradation, associated with different individual molecules to different degrees for the duration of a mobilization, and altered the enzymes' charge to frictional coefficient ratios. Previous experiments have

shown that PVP does not interact with anionic analytes, but this work is of uncertain applicability to large proteins such as β-galactosidase [86]. A different coating polymer would presumably interact differently with the enzyme and therefore one would expect to see different absolute electrophoretic mobilities, and a different distribution of mobilities. Experiments performed using uncoated capillaries would eliminate this possibility completely.

8.3.5.2. Genescan coated capillaries:

Genescan polymer 6^{TM} (GS-6) has been previously used to coat capillaries for single molecule β-galactosidase assays [45]. The method for single molecule mobilizations and calculations for electrophoretic mobility on a GS-6 coated capillary were performed identical to those described for PVP coated capillaries. Figure 40 shows a representative electropherogram for a 10 minute mobilization of β-galactosidase molecules on a GS-6 coated capillary, and includes a similar run performed on a PVP coated capillary for comparison.

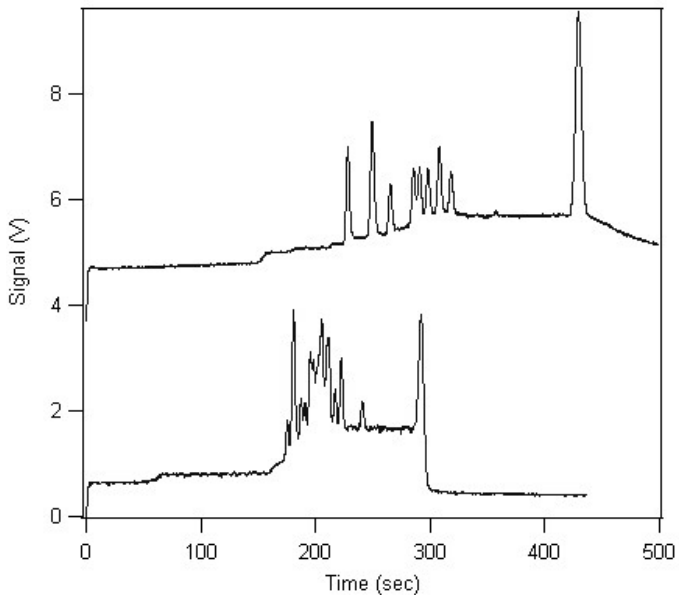


Figure 40. Extended separation on GS-6 and PVP coated capillaries. Both mobilizations were performed in an identical manner: $4x10^{-8}$ M DDAO reference standard and enzyme mobilized 10 minutes at 400 Vcm^{-1} in 100 μM DDAO-gal in separation buffer. Five minute static incubation followed by second mobilization at same potential. The traces have been offset for clarity. *Lower trace:* Capillary coated with 0.07% (w/v) GS-6. *Upper trace:* Capillary coated with 0.5 % (w/v) PVP.

The DDAO reference standard and individual enzyme peaks reached the detector more quickly compared to the PVP run because GS-6 suppressed EOF more effectively. The greater suppression of EOF was also evidenced by closer proximity to the detection end of the capillary by the background hydrolysis baseline shift. The individual enzyme molecules appeared to exhibit less electrophoretic heterogeneity on the GS-6 coated capillary because separations for CE are more effective when EOF and electrophoretic mobility are similar in magnitude but in opposite directions [67].

The reproducibility for both types of separations performed on GS-6 coated capillaries was similar to PVP coated capillaries, but because the difference in apparent mobility of the enzyme and product was diminished on the GS-6coated capillary, it was easier to perform separations using the extended separation method. The electrophoretic

mobilities of 137 individual β-galactosidase molecules from the commercially purified source were measured on a GS-6 coated capillary (Table 11). The highest, lowest and mean calculated electrophoretic mobility were -1.87×10^{-4} $cm^2V^{-1}s^{-1}$, -1.63×10^{-4} $cm^2V^{-1}s^{-1}$, and -1.77×10^{-4} $cm^2V^{-1}s^{-1}$ respectively. These values are very similar to those measured for enzyme molecules from the same source on a PVP coated capillary. The relative range of electrophoretic mobilities was 1.14.

8.3.5.3. Uncoated capillaries:

Capillary coat effects are most easily evaluated by comparison to an uncoated capillary. The inner surface of an uncoated capillary at pH 7.3 has a high negative charge density, and the pI for β-galactosidase is approximately 5, and therefore at pH 7.3 the enzyme also has considerable negative charge. Studies of the behavior of proteins on charged surfaces at the single molecule [185] and bulk level [186] clearly indicate that there is no interaction between the protein and the charged surface when the pH is more than one pH unit above the pI of the protein. Extended separation assays on uncoated capillaries required a modified approach because the injection end of the capillary was held at a positive potential (normal polarity), so the direction of EOF was towards the detector.

Figure 41 is a schematic that depicts the extended separation protocol on an uncoated capillary. A run is started by introducing approximately 1×10^{-13} M β-galactosidase and 4×10^{-8} M DDAO reference standard in separation buffer into the capillary by electrokinetic injection for 7 s at 100 Vcm^{-1}. The injection end was then placed in a vessel containing 100 μM DDAO-gal in separation buffer. The enzyme, the

reference standard, and the substrate were mobilized for 6 minutes by the application of a 400 Vcm^{-1} potential.

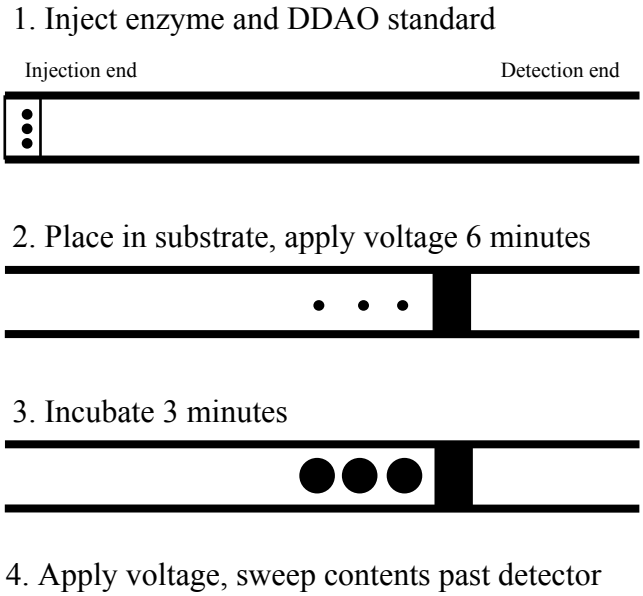


Figure 41. Schematic for extended separations on an uncoated capillary. Shaded rectangle is injected DDAO reference standard. Small circles represent individual enzyme molecules, and the large circles represent pools of product formed by enzymatic hydrolysis.

The DDAO plug and the enzyme are negatively-charged at the experimental pH and consequently have lower net mobility towards the detector than the neutral DDAO-gal substrate. As a result the capillary was filled with substrate before the enzyme or DDAO reached the detection end of the capillary. The apparent mobility of β-galactosidase is less than that of the DDAO standard because of the enzyme's higher electrophoretic mobility, and consequently the enzyme migrated a shorter distance along the capillary during the 6 minute mobilization. As before, because different individual enzyme molecules have different electrophoretic mobilities, they also separated from each other. The contents of the capillary were then statically incubated for 5 minutes, and pools of DDAO product formed around the individual enzyme molecules. At the end of the

incubation the potential was reapplied and the pools of the fluorescent product and the DDAO standard were driven pass the detector by mobilization at 400 Vcm^{-1}.

The electrophoretic mobility of the individual enzyme molecules was calculated by determining their relative migration to the DDAO standard during the initial 6 minute mobilization. Because the substrate DDAO-gal is uncharged, the velocity of its migration in the capillary will be identical to the velocity of the EOF. By measuring the time required for the weakly fluorescent DDAO-gal to reach the detector, the velocity of the EOF can be calculated as:

$$v_{eof} = \frac{L}{t_{DDAO\text{ - }gal}} \tag{8.6}$$

where L is the length of the capillary and $t_{DDAO\text{-}gal}$ is the time required for the substrate to reach the detector. This again allowed for the calculation of the velocity of the EOF for every run independently. The apparent velocity of an individual enzyme molecule was calculated by the same method as for coated capillaries:

$$v_{app} = \frac{t_{DDAO} - t_{pk} + t_{sep}}{t_{sep}} \text{x} \frac{L}{t_{DDAO} + t_{sep}} \tag{8.7}$$

where t_{DDAO} and t_{pk} were the electropherogram times required for the DDAO standard and product pools respectively to reach the detector following the incubation; t_{sep} is 360 s. The actual electrophoretic mobility of individual β-galactosidase molecules is therefore the apparent velocity of the enzyme molecules less the velocity of the EOF with the difference divided by the field strength of 400 Vcm^{-1}:

$$\mu = \frac{v_{app} - v_{eof}}{E} \tag{8.8}$$

Short mobilizations were attempted on an uncoated capillary, but because the large EOF dominated the apparent mobility of the enzyme and its associated product

molecules, the difference in their respective net mobilities was small. Short mobilizations provided insufficient time to effectively separate an enzyme molecule from its prior product pool, and if longer times were used, enzymes and their product pools were swept off the capillary. Figure 42 shows a series of extended mobilizations on an uncoated capillary.

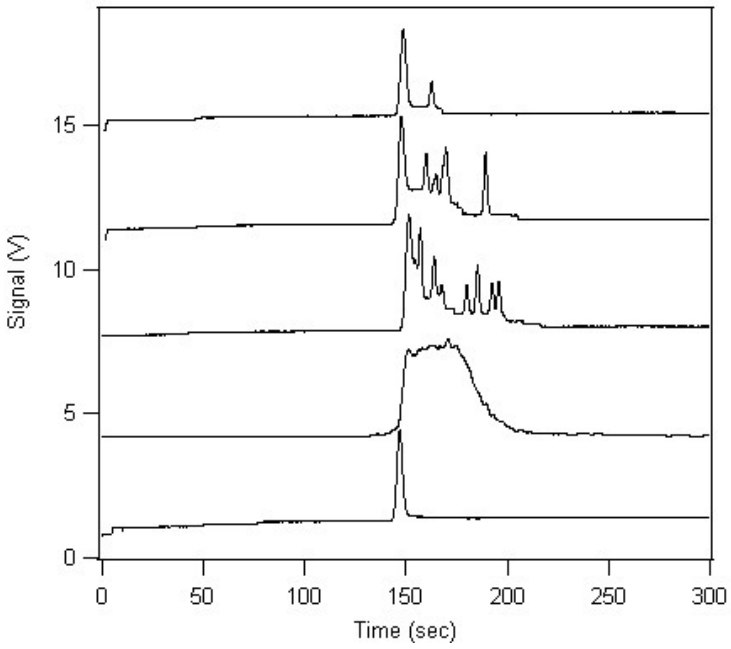


Figure 42. Extended separations on an uncoated capillary. Various concentrations of enzyme and $4x10^{-8}$ M DDAO were injected into the capillary for 7 s at 100 Vcm^{-1} (injection end positive). Injection end was then placed in 100 μM DDAO-gal and capillary contents mobilized at 400 Vcm^{-1} (injection end positive) for 6 minutes followed by 5 minute incubation. Following incubation capillary contents were swept past the detector by the reapplication of the 400 Vcm^{-1} electric field. *Lowest trace:* Blank. *Second trace:* 1 x 10^6 dilution of enzyme stock. (note PMT bias lowered to avoid saturating the detector). *Third trace:* 1 x 10^8 fold dilution. *Fourth trace:* 2 x 10^8 fold dilution. *Fifth trace:* 4 x 10^8 fold dilution.

Again, the decreasing number of peaks as the concentration of injected enzyme was reduced is a hallmark of single molecule assays, and confirms that the DDAO peaks were produced by single enzyme molecules. This series of electropherograms also illustrates

that the range of mobilities on an uncoated capillary of enzyme molecules measured at the single molecule level is indistinguishable from a bulk separation.

A total of 117 β-galactosidase molecules were assayed (Table 11), and the highest and lowest individual observed single molecule electrophoretic mobilities were -1.94×10^{-4} $cm^2V^{-1}s^{-1}$ and -1.63×10^{-4} $cm^2V^{-1}s^{-1}$ respectively. The mean value was -1.81×10^{-4} $cm^2V^{-1}s^{-1}$ and the relative range of mobilities was 1.19. The higher absolute electrophoretic mobility values are likely attributable to the lower viscosity of the separating medium due to the absence of any coating polymer. The greater range of mobilities may be due to the very low value for the slowest moving molecules. Only two molecules had a value this low, and if they are excluded, the relative range falls to 1.16, which is similar to all the other separations. The slight increase in the overall range of mobility of enzyme molecules on the uncoated capillary suggests that there may be a very modest coat interaction component to the observed single molecule electrophoretic heterogeneity, and that this interaction actually reduces the extent of electrophoretic heterogeneity.

8.3.6. Effect of enzyme source:

To determine if the observed electrophoretic heterogeneity varies with source, the electrophoretic mobility of 141 individual molecules of β-galactosidase obtained in the laboratory from *E. coli* strain ATCC 35321 were measured on a PVP coated capillary by the short multiple separation method. No attempt was made to purify the enzyme from bacteria—crude lysate was simply diluted to the extent necessary to measure individual molecule activity. The electrophoretic mobility of every enzyme was measured twice and the average for the two mobilizations was used. The results for these mobilizations are also presented in Table 11. The highest measured electrophoretic mobility for any

single enzyme molecule from this source was -1.85×10^{-4} $cm^2V^{-1}s^{-1}$, and the lowest was -1.60×10^{-4} $cm^2V^{-1}s^{-1}$, with mean value of -1.74×10^{-4} $cm^2V^{-1}s^{-1}$. The relative range of mobilities was 1.16, which is very similar to the relative range values of 1.15 and 1.14 that were obtained for the commercially prepared enzyme. A two tailed T-test showed no significant difference ($p = 0.12$) between the commercially prepared enzyme and the enzyme obtained from a crude lysate. Purification therefore does not create electrophoretic heterogeneity. Moreover, this discounts the possibility that the heterogeneity arises due to unspecific interactions between the enzyme molecules and molecules liberated during enzyme harvesting. If unspecified interactions were contributing to electrophoretic heterogeneity, these interactions would be different between the two sources, and should result in a different distribution of electrophoretic mobilities between the sources. Electrophoretic heterogeneity at the single molecule level for β-galactosidase is observed because the molecules are different.

8.4. Discussion[1]

8.4.1. Significance of electrophoretic mobility heterogeneity

The results presented here reveal that individual molecules of β-galactosidase are electrophoretically heterogeneous, and electrophoretic heterogeneity implies structural heterogeneity. The velocity of an ion under the influence of an electric field is given by the relationship:

$$v = \frac{EZ_a}{f} \tag{8.9}$$

[1] The structural implications of electrophoretic mobility heterogeneity were reexamined for the results obtained from investigations regarding the role of streptomycin and translation error upon catalytic heterogeneity presented in chapter 10. The separation is preserved here because they were performed under different experimental conditions and to reflect the evolution of my understanding of their significance. The extent of the compatibility of the two analyses is assessed in the chapter 10 analysis.

where v is the velocity of the ion, E is the potential applied, Za is the charge on the ion, and f is the frictional coefficient of the charged particle as it migrates through the separating medium. Molecules that are structurally identical should be electrophoretically identical.

The measured heterogeneity itself does not directly differentiate between changes of electrophoretic mobility that are due to alterations of charge, mass or combinations of both. Differences in electrophoretic mobility of nearly identical proteins are predominantly because of charge heterogeneity [80], and it is probable that the variability measured here is attributable to charge differences. The application of theoretical descriptions of protein electrophoresis make it possible to estimate the magnitude of the changes that are required to give rise to the observed single molecule heterogeneity, and to suggest its possible source. A simplified proportionality for describing protein electrophoretic behavior is given by the relationship:

$$\mu \propto \frac{Z_c}{M^s} \tag{8.10}$$

where Zc is the calculated valence of the protein based upon the Henderson-Hasselbach equation, and M is the protein's molecular mass, and s varies between 1/2 and 2/3 [80]. Zc often overstates charge because the calculated valence is based upon the pK_a's of free amino acids and fails to account for the influence of the local microenvironment on the ionization of residue side chains, and the surface electrostatic charge suppression of proteins [187]. To rectify this shortcoming, it has been shown that for any particular protein there is a pH-independent proportionality Fz between the calculated charge Zc and the actual charge Za, called the charge suppression factor [188].

$$F_z = \frac{Z_c}{Z_a} \tag{7}$$

If the mean electrophoretic mobility of the enzyme is assumed to have the native charge and mass, the magnitude of a charge or mass change that would be required, when the other is held unchanged, to produce the observed range of electrophoretic mobility can be calculated. From a *Z*c for β-galactosidase at a pH of 7.3 of -145.6 (scripps.edu/~cdputnam/protcalc.html), and a previously determined F_Z for β-galactosidase of approximately 6[188], the *Z*a of the enzyme is -24.3. When the F_Z for a protein is known, the best value for the *s* coefficient has been shown to be between 0.6 and 0.67 [80]. If the mean mobility of 119 purified β-galactosidase molecules (Table 11) is deemed to carry the native charge (-24.3), mass (465 kDa), and a value of 0.67 is used for *s*, when the mass is assumed to be constant, approximately 1.6 suppressed charge units had to be gained or lost by an individual protein molecule to reach the limits of the measured electrophoretic variability. Such charge heterogeneity may be possible. Conversely, if the charge were held constant, the protein would need approximately a hundred amino acid residues to be gained or lost to exhibit the observed mobility range. Such a result is less plausible. This latter number presumes the protein density is constant. It is possible that different individual molecules have no change to their primary structure, but instead have different regions with variable amounts of disorder, and that such disorder alters the effective radius of different enzyme molecules differently. It is also difficult to envision how such a disordered molecule would retain catalytic activity, and it seems doubtful that such quaternary heterogeneity is easily reconciled with conventional explanations for protein folding.

IEF was performed on the commercially purified β-galactosidase sample to determine if charge heterogeneity was present. The analysis yielded a series of closely

spaced bands corresponding to a pI range of 5.18 to 5.43, and this suggests that differences in charge are indeed an important component of the observed differences in mobility during CE. The calculated pI for the native sequence of β-galactosidase is 5.51. The observed result may be different because the calculated pI is based upon the ionization constants for individual amino acids and does not account for alteration of residue pK_a's. If one uses a charge suppression factor of 6, then 10 single charge alterations in either direction are required to approximate the suppressed charge difference of ±1.6. This yields a calculated pI change of ±0.09, which is similar to the observed range. However, protein charge ladder studies have shown that modifications of proteins that are expected to alter the charge by 1 unit, as predicted from pH and the pK_a's of the functional groups, alter the observed net suppressed charge of the protein by values ranging from 0.8 to close to unity [187]. If such is the case, then a charge alteration of two residues would be sufficient to yield a suppressed charge difference of 1.6. However, the calculated effect of such a change results in a pI difference of ±0.01, which is substantially less of a pI range than was observed. Protein charge ladder analyses are typically performed using proteins smaller than 50 kDa [189], so possibly the change in net suppressed charge for single amino acid substitutions is less substantial for extremely large proteins such as β-galactosidase.

8.4.2. Relationship between catalytic and electrophoretic heterogeneity

The range of electrophoretic mobilities for a sample of individual β-galactosidase molecules was approximately 15%. This is considerably less than the 10-fold differences in catalytic rates that have been measured for individual enzyme molecules. However, there was no strong relationship between individual enzyme mobility and catalytic

activity, apart from a small number of low mobility molecules with low activity. This result is not surprising because substitutions with minimal impact to the electrophoretic properties of the protein at sites near the active site could significantly affect substrate binding or transition state stabilization, but a more substantial (from the perspective of mobility) alteration at a location distant from the active site might have trivial effect upon activity.

The observed electrophoretic mobilities fall within a limited range for all the experiments. The detection limit for the assay allows for the detection of enzyme molecule activities is much less than the least active molecules observed. The absence of detection of any catalytically active individual enzyme molecules outside this apparent range suggests that such molecules do not exist, and therefore that such excess structural heterogeneity is catastrophic for β-galactosidase function.

8.4.3. Possible sources of charge heterogeneity

If one considers that *E. coli* error rates for transcription and translation are approximately 10^{-4} [190] and 10^{-3} to 10^{-4} respectively [191], and that a β-galactosidase tetramer has 4092 residues, it is likely that every molecule will contain at least one error in primary sequence. If the substitutions entailed the change of two like-charged amino acid residues to that of neutral ones, or a change of a single charged residue to one with opposite charge, this could make up a significant portion of the observed total range of electrophoretic mobilities on a suppressed charge basis. Those errors that do not entail the substitution of a charged residue for an uncharged one, or vice versa, could mediate their effects by subtly altering the ionization constants of neighboring residues that are charged. Because changes to the primary structure could occur anywhere, and the effect

of the change would depend upon the location and the nature of the change, one would observe a distribution of mobilities. The role that error may have for the electrophoretic and catalytic heterogeneity of β-galactosidase will be considered in chapter 10.

8.5. Conclusion

To elucidate possible structural bases for single molecule β-galactosidase catalytic heterogeneity, the electrophoretic mobility of single molecules was measured, and it was demonstrated that individual molecules have distinct, different, and reproducible mobilities. The differences could not be attributed to any exogenous capillary or purification effects, and implies that individual enzyme molecules are structurally distinct. Application of theoretical models for protein electrophoretic mobility to the results suggests that charge heterogeneity is likely to be a significant source of the observed variation. Although there was no evident relationship between electrophoretic and catalytic heterogeneity, this result supports the contention that conformational differences are insufficient to account for single molecule catalytic heterogeneity.

CE is commonly used in the separation and analysis of protein mixtures. Considerable work has been reported on the development of capillary coatings in order to reduce band broadening resultant from interactions between the proteins of interest and the capillary wall [84, 85]. Detection limits have been improved using fluorophore labeled proteins and laser-induced fluorescence detection in CE separations [192]. One concern has been band-broadening due to differential labeling of individual protein molecules [193] and reagents have been developed to minimize this effect [194]. If the finding that β-galactosidase is inherently heterogeneous with respect to electrophoretic mobility can be extended to other proteins, such heterogeneity may represent a

fundamental limit to the resolution of protein separations by capillary zone electrophoresis.

8.6. Appendix II

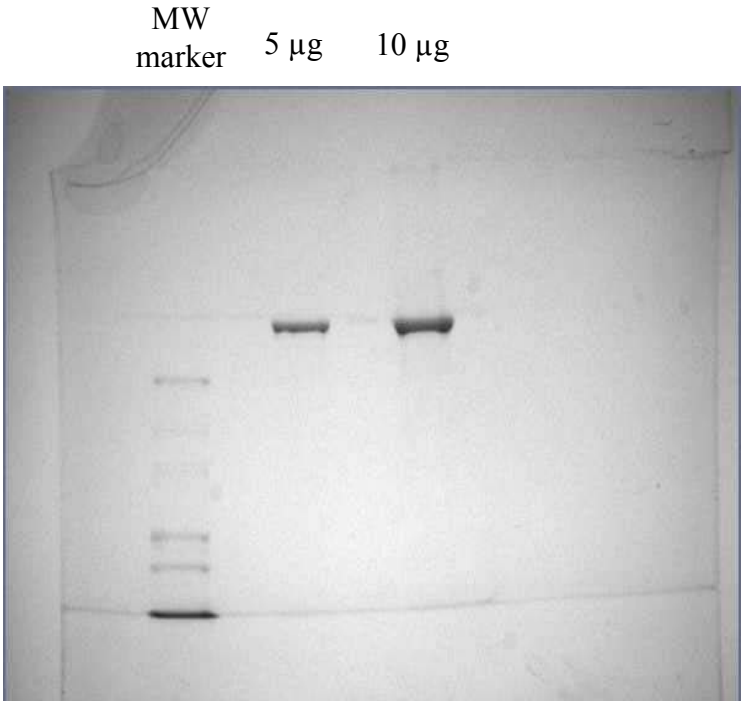


Figure 43. SDS-PAGE for Sigma β-galactosidase. Stacking gel was 3.75% 37:1 acrylamide/bis-acrylamide, 0.1% SDS in 0.125 M Tris-HCl pH 6.8 and resolving gel was 7.5% 37:1 acrylamide/bis-acrylamide, 0.1% SDS in 0.375 M Tris-HCl pH 8.8. Five or 10 µg of commercial β-galactosidase in 0.0625 M Tris-HCl pH 6.8, 10% (w/v) sucrose, 1% SDS, 3% (v/v) 2-mercaptoethanol, 0.0005% bromophenol blue was loaded. Gels were run at 200 V for one hour, stained with 0.25% (w/v) Coomassie brilliant blue in 45% (v/v) methanol and 9% (v/v) glacial acetic acid for 20 minutes. Destaining was overnight in 7% (v/v) methanol and 7% (v/v) glacial acetic acid. Gel imaging was performed using a Biorad UV Gel Analyzer, Universal Hood II.

9. Continuous flow assay and thermally-induced conformational fluctuations support different catalytic rates

The data presented in this chapter was originally published as:

Craig, D.B. and Nichols, E.R. (2008) Continuous flow assay for the simultaneous measurement of the electrophoretic mobility, catalytic activity and its variation over time of individual molecules of *Escherichia coli* β-galactosidase. *Electrophoresis*, **29**: 4298-4303. Copyright permission to use the data has been obtained.

9.1. Introduction

Single molecule enzyme studies can be divided into two distinct classes. One class consists of trajectory-based studies that monitor the catalytic activity of a confined molecule by detecting the release of fluorescent product molecules [5, 56, 195]; by recording conformational fluctuations from intrinsic cofactor fluorescence [4]; by detecting blinking from covalently attached fluorophores [196, 58] or by smFRET [42, 37] that are assumed to reflect catalytic turnovers. These stochastic events are assembled into histograms for the time interval between catalytic turnovers, and from this information the mathematical deduction is made that individual enzyme molecules undergo extensive interconversions of different states with different catalytic rates.

The second class of single enzyme molecule experiments has been performed using CE, and non-CE based techniques, and measure time-averaged activity. These studies have consistently demonstrated that the activities of individual molecules can vary by greater than 10-fold, and that multiple incubations of these molecules reveals that the large differences of activity are reproducible [2, 3, 18, 46]; see also chapters 5 and 6 this document. This static heterogeneity suggests that any particular enzyme molecule is capable of only a single rate under any particular set of conditions. If enzyme molecules are capable of conformational changes over a broad range of timescales as suggested by

the experiments that measure dynamic heterogeneity, it leaves unanswered the question of why assay types that measure static heterogeneity have not shown an interconversion between conformations that support different catalytic rates. A legitimate critique of these latter assays is that they do not continuously monitor activity, and therefore cannot determine if there are fluctuations of activity underlying the time-averaged measurements [197, 198]. It has been suggested that different molecules progressing through different conformational sequences will converge on a single time-averaged rate for an assay that is 10^3 to 10^4 longer than the average turnover time [60]. Yet assays of long duration using β-galactosidase with FDG, or RES-gal as substrates show no tendency for a convergence of single molecule rates. This would suggest that time scales greater than 10^4 the average turnover time are required for convergence. However, if enzyme molecules are continuously changing conformation, a single molecule should exhibit detectable changes of time-averaged activity when the assay is short. The reproducible time-averaged activities for the 5 minute quadruple assays with DDAO-gal presented in chapter 8 indicate that any transient variations in activity are already well-averaged. This incongruity is difficult to understand without impugning published results. Modification of the techniques used for static heterogeneity measurements to collect continuous data could help resolve these discrepancies, and place the contention that enzyme molecules adopt multiple conformations on a sturdier basis than mathematical inferences.

Avila and Whitesides [199] demonstrated an ensemble CE-based enzyme assay in which a plug of glucose 6-phosphate dehydrogenase (G6PDH) was introduced into a capillary that was continuously electrokinetically flushed with the substrates NAD and G6P. As the enzyme migrated along the capillary the NAD was reduced to NADH. The

enzyme and reaction product have different electrophoretic mobilities, and as a result, a broad plateau of NADH was formed which could be detected by absorbance at 340 nm. The single molecule DDAO-gal assays described in chapter 8 also exhibit a local plateau feature that reflects the catalytic activity of individual enzyme molecules as they migrated through the substrate before and after the static incubation. An example of such a plateau is evident in the electropherogram in Figure 44 on both sides of the peak at ~ 185 s.

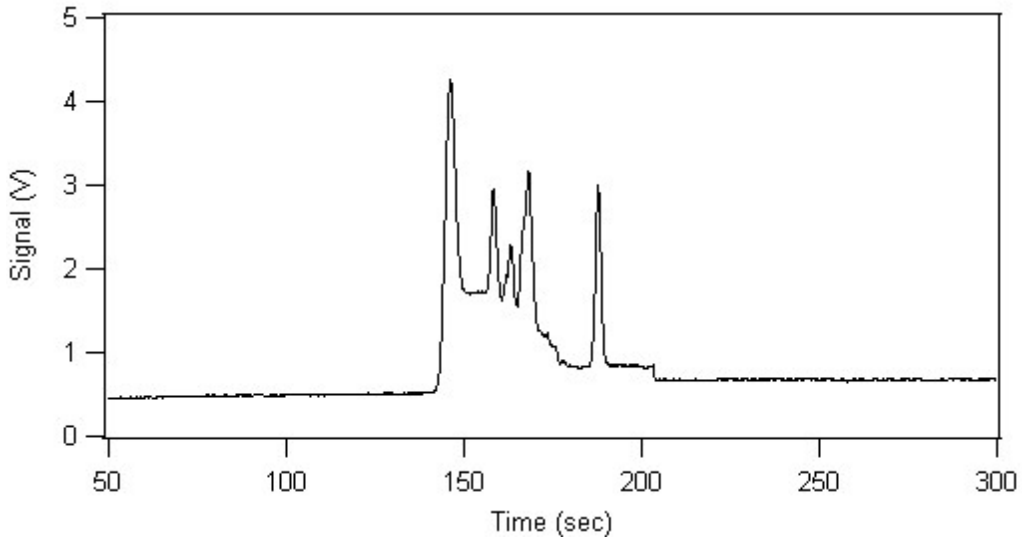


Figure 44. Electropherogram depicting single molecule peak and plateau.

The plateau that is seen after the peak was formed by the enzyme following the static incubation. It was recognized that these plateaus represented the same phenomena described by Avila and Whitesides, and this suggested a technique that would make it possible to continuously monitor single enzyme activity. This chapter presents a novel CE-LIF based assay that allows for the simultaneous measurement of the electrophoretic mobility, catalytic activity and the variation in activity over time for individual molecules of *E. coli* β-galactosidase which may help bridge the experimental gap described above.

9.2. Methods

9.2.1. Instrumentation, buffers and β-galactosidase

The assay is based upon the DDAO-gal β-galactosidase substrate using the 633 line of the 10 mW HeNe laser for excitation. Instrumentation for this assay was as previously described, except the analog PMT signal was collected and digitized using a Pentium 4 computer containing a PCI-MIO-16XE-50 I/O board utilizing LabView™ software. Capillaries were 60 cm long and had an inner diameter of 2 μm. Sheath and separation buffers were 50 mM HEPES (pH 7.3), 1mM $MgCl_2$ and 1 mM citrate and contained no coating polymers.

9.2.2. Continuous flow assay

Stock 20 mM DDAO-gal in DMSO was prepared to 100 μM in a final volume of 500 μL as described previously. The sample was continuously injected into the capillary at 300 Vcm^{-1} (injection end positive). Where necessary, regions of the capillary were heated by confining it between an aluminum block and insulating foam. The aluminum block was heated to the desired temperature by passing water from a Julabo model F12 circulating water bath through an internal coil.

9.3. Results

9.3.1. Continuous flow single molecule assay

Figure 45 depicts how the single molecule continuous assay was performed. At the start of an assay a buffer-filled capillary was placed in an injection vessel that contained approximately 2 x 10^{-15} M β-galactosidase and 100 μM DDAO-gal in separation buffer and a 300 Vcm^{-1} potential was applied. The substrate has no charge and it immediately entered the capillary because of EOF. An enzyme molecule would

eventually enter the capillary and begin its migration towards the detector. As the enzyme molecule moved through the capillary it continuously converted substrate into product. As discussed in chapter 8, the enzyme and the DDAO product molecules both carry a negative charge at the experimental pH, and both would tend to electrophoretically migrate towards the injection end of the capillary but for the EOF, which was of greater magnitude than both because the capillary was uncoated. Since the electrophoretic mobility of the enzyme is greater than DDAO, its apparent mobility towards the detector was lower than the product molecules that it produced, and therefore the product molecules continuously moved away from the enzyme towards the detection end of the capillary. This resulted in the formation of a zone of product by each individual enzyme molecule.

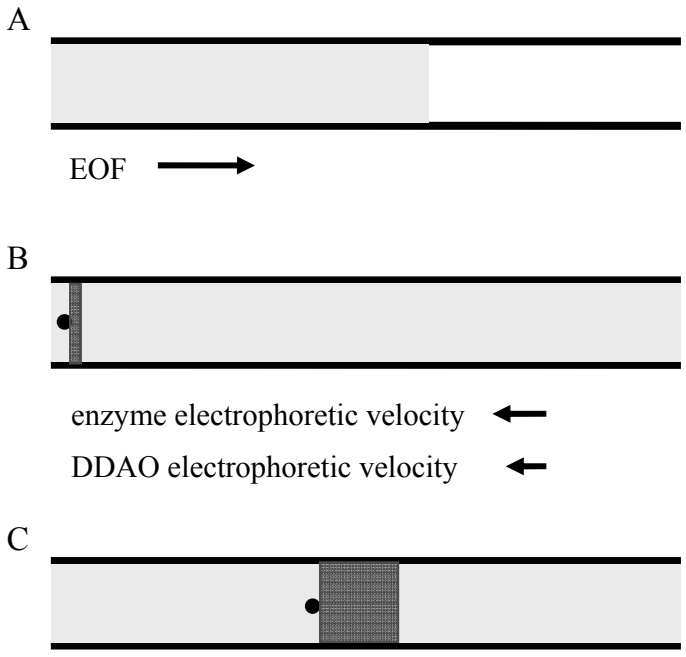


Figure 45. Continuous assay protocol. *A*. DDAO-gal enters capillary by EOF. *B*. Hydrolysis of substrate by migrating enzyme produces faster migrating product zone. *C*. Product zone approaches detector. Light shaded region is DDAO-gal; black circle is individual enzyme molecule; dark shaded region is DDAO product formed by enzymatic hydrolysis.

Figure 46 shows a portion of a continuous assay electropherogram for two individual β-galactosidase molecules performed at 24°C. The plateaus represent the product zones that were formed as the enzyme molecules traversed the capillary. The plateaus have different widths because different enzyme molecules have different electrophoretic mobility, and as a result their respective product zones migrated different distances ahead of the enzyme molecule. Enzyme molecules with higher electrophoretic mobility have broader plateaus. The area of a plateau is proportional to the catalytic activity of the individual enzyme molecule and its mobility, and variation of areas is due to differences in the activity and mobility of the individual enzyme molecules. The first product molecules to reach the detector provide information about the enzyme when it initially entered the capillary. The final product molecules to reach the detector provide information about the enzyme just prior to its exit. The zone in between provides a continuous record of the behavior of the enzyme as it migrated along the capillary.

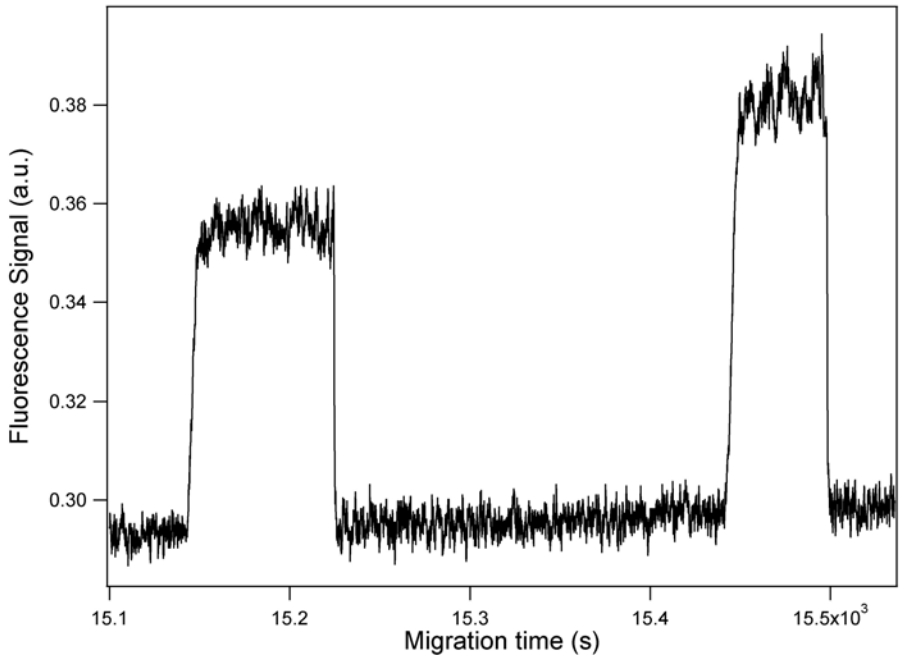


Figure 46. Continuous assay electropherogram. DDAO product plateaus produced by two β-galactosidase molecules mobilized at 300 Vcm^{-1} in 100 μM DDAO-gal in separation buffer.

The plateaus sit upon a background formed of DDAO that was always present because of the residual DDAO that was not removed by toluene extraction and that which formed by non-enzymatic hydrolysis, and by the enzymes in the injection vessel. DDAO-gal also contributes to the background because the EOF is oriented towards the detector. If the DDAO-gal produced a signal of similar magnitude as seen for FDG and RES-gal, this assay would not be possible because the plateaus would be invisible in the background signal produced by the substrate. The signal to noise ratios for the continuous assay is lower in comparison to the other CE-LIF assays because the enzyme molecule is migrating throughout the incubation and therefore its product pool is distributed over a broad zone rather than restricted to a small pool. This lowers the amount of DDAO produced by the enzyme that is passing through the excitation volume per unit of time, whereas the amount of background DDAO that is present is unchanged. The amount of non-enzymatic DDAO passing through the excitation volume can be reduced, and therefore the signal to noise ratio improved, if a narrower bore capillary is used. For this reason, the continuous assay is performed with 2 μm rather than 10 μm inner diameter capillaries.

9.3.2. Determination of electrophoretic mobility and catalytic rate

The neutral DDAO-gal substrate can be used as an EOF marker because the weakly fluorescent substrate traveled faster than the negatively charged DDAO in an uncoated capillary, and was detectable as a baseline shift when it reached the detector. Although the concentration of the DDAO impurity in the substrate was greatly reduced by toluene extraction prior to the assay, it was not completely removed. As such, there was a second baseline shift as the DDAO started to exit the capillary that can be used to

determine its mobility. The average DDAO-gal and DDAO retention times were 551 s and 992 s respectively. The velocity of the EOF, v_{eof}, can be determined from the mobility of the DDAO-gal using the equation:

$$v_{eof} = \frac{L}{t_{DDAO\text{ - }gal}} \tag{9.1}$$

Where L is the capillary length and $t_{DDAO\text{-}gal}$ is the migration time of the DDAO-gal to the detector. The electrophoretic mobility of the enzyme molecules can be calculated from the width of the box-shaped peaks using the equation:

$$\mu = \frac{\frac{L}{w_{pt} + t_{DDAO}} - v_{eof}}{E} \tag{9.2}$$

where w_{pt} is the width of the plateau, t_{DDAO} is migration time for the DDAO to the detector mentioned above, and E is the strength of the electric field (300 Vcm^{-1}). In this study, any peaks that appeared to result from the overlap of two plateaus were omitted from the data analysis. The electrophoretic mobility of the 69 enzyme molecules assayed was $-1.74 \times 10^{-4} \pm 4 \times 10^{-6}$ $cm^2V^{-1}s^{-1}$ (mean ± σ) with a range of –1.65 to -1.83×10^{-4} $cm^2V^{-1}s^{-1}$, and is similar to the measurements made using different methods in chapter 8. The weighted mean mobility for enzyme from the same source, in buffer of the same ionic strength as measured by the short and extended method on coated and uncoated capillaries as described in the separations chapter was -1.77×10^{-4} $cm^2V^{-1}s^{-1}$. The range of mobilities here, 1.11, is slightly lower than the 1.15 average for all the separation methods performed in chapter 8. This difference may be due to a smaller sample size for these separations.

The catalytic rate can be determined by comparison of the area of the plateaus to that of standards of DDAO divided by the incubation time. The incubation time, t_{ass}, is simply the time that each enzyme molecule spent in the capillary. This is calculated as:

$$t_{ass} = t_{DDAO} + w_{pt} \tag{9.3}$$

The average rate for the individual molecules was 29,000 ± 11,000 minute^{-1} with a range of 9,200 to 53,000 minute^{-1}. The catalytic rate of *E. coli* β-galactosidase has been found to vary with source [48] and these values fall within the normal range of the values observed.

Variation of the height of the plateau would indicate variation in the activity of the enzyme molecule during the incubation. Variation of the height could also reflect changes to the electrophoretic mobility of the enzyme—if the migration slowed, but the activity was unchanged, the plateau would be locally higher. This is viewed as improbable because multiple measurements of the mobility of individual molecules presented in chapter 8 indicate that their mobilities do not vary significantly as they traverse the capillary. The temporal resolution for activity variation is limited by the background noise and diffusion. The background noise for a plateau can be crudely judged by comparison to the noise of the non-plateau regions with the caveat that one must not unduly over interpret fluctuations in the height of the plateau. Figure 46 shows that the plateaus were relatively flat with moderately more noise than the background. This was typical of all peaks observed and indicates that at 24^{o}C the catalytic rates of the enzyme molecules, although different from each other, are relatively constant for the duration of the assay.

9.3.3. Assays at different temperatures

Figure 47 shows a similar assay performed at 40°C. The middle 46 cm of the 60 cm capillary was incubated at 40°C by heating with an aluminum block connected to a recirculating water bath. A 7 cm length of capillary on either end was at ambient room temperature, which in this case was 27°C. Two peaks are shown and as previously observed, the areas and widths differ. However, at 40°C the plateau shows discrete regions where the height changes abruptly although the background signal, as judged by the inter-plateau regions, was stable. This was typical of all the peaks observed (N=65) for assays of this type, and implies that the catalytic rate at the higher temperature is not constant, but rather, varies over time.

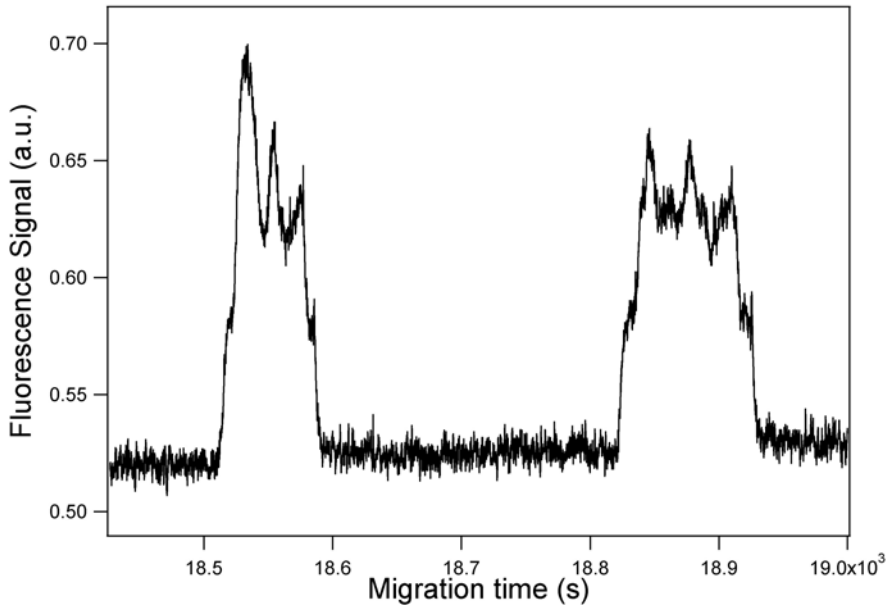


Figure 47. Continuous assay at 40°C. DDAO product plateaus produced by two β-galactosidase molecules mobilized at 300 Vcm^{-1} in 100 μM DDAO-gal in separation buffer.

The small shelves at both edges of the plateau are due to the decreased activity while the enzyme traveled the 7 cm of capillary at room temperature. The sharp change in activity

upon entering and exiting the heating block indicates that the temperature in the 2 μm internal diameter capillary equilibrates relatively quickly [3].

Figure 48 shows a double plateau resulting from the continuous flow assay of an enzyme molecule in a capillary where the first 7 cm and the final 25 cm were held at 28°C, and the internal 28 cm was at 37°C.

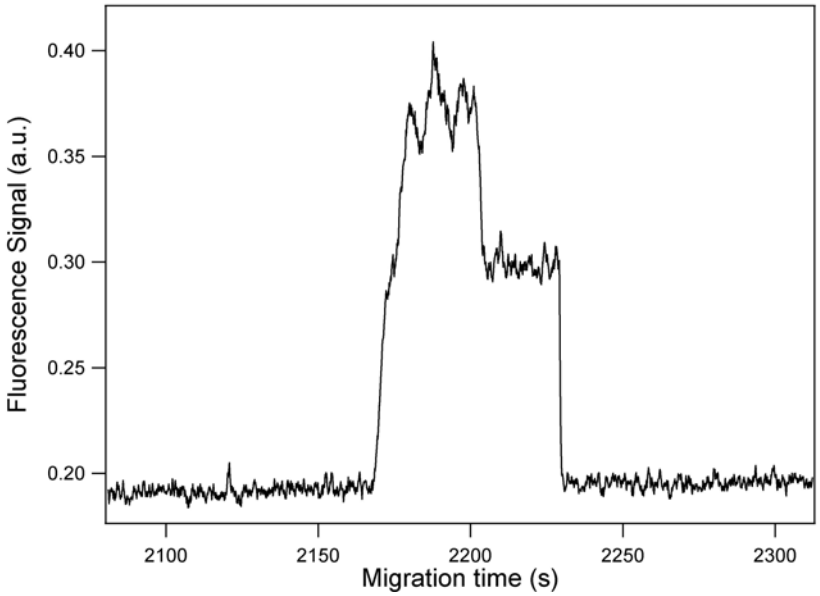


Figure 48. Continuous assay at 37°C and 28°C. A single β-galactosidase was continuously migrated at 300 Vcm^{-1} in 100 μM DDAO-gal in separation buffer. Initial 7 cm and final 25 cm of capillary at 28°C, and internal 28 cm at 37°C.

An initial shelf is observed for the 7 cm at 28°C, followed by a plateau corresponding to the 28 cm of capillary at 37°C that is in turn followed by a lower plateau for the 25 cm at 28°C. Fluctuation in the observed signal over the time axis at the higher temperature is substantially larger than that at the lower temperature, and this was typical of all molecules observed (N=25). In all cases, the small shelf corresponding to the initial 7 cm of the capillary was at approximately the same height as the plateau corresponding to the final 25 cm of the capillary. The difference in height of the two plateaus corresponds to the difference in activities supported at the two temperatures and was found to differ

between individual molecules. On average the height of the plateau at 37°C was 2.1 ± 0.5 (mean ± σ) times that at 28°C, with a range of 1.7 to 3.7-fold. Heterogeneity with respect to the change in activity with temperature has been reported previously with alkaline phosphatase [3].

In Figure 49 an enzyme molecule was passed through a capillary where the central 20 cm of the capillary was incubated at 45°C, and the 20 cm at both ends was incubated at 27°C. Once again, the height of the signal is seen to fluctuate at the higher temperature and remain relatively constant at the lower temperature.

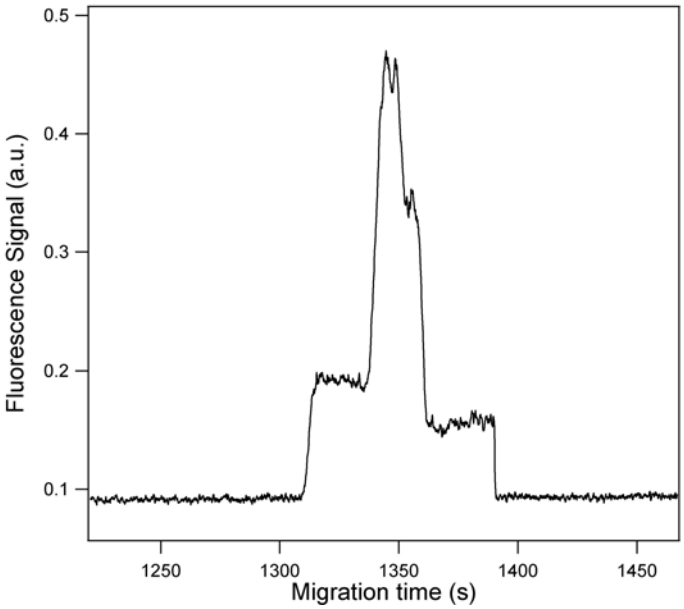


Figure 49. Continuous assay at 27°C, 45°C and 27°C. A single β-galactosidase molecule was continuously migrated at 300 Vcm^{-1} in 100 μM DDAO-gal in separation buffer. First and third portions of plateau incubated at 27°C; the middle region incubated at 45°C.

This pattern was observed for the most of the 82 molecules assayed; however for 13 molecules there was a complete loss of activity during the incubation at 45°C. For the majority of the molecules, the heights of the two plateaus corresponding to the 27°C were within 10% of each other. The signal for the molecule depicted in Figure 49 is that of

one where a lower activity was obtained subsequent to incubation at the higher temperature. Of those molecules where the difference in activity was greater than 10%, more were found to have a lower temperature following the period at the higher temperature than were found to have a higher activity. On average, the second plateau at 27^{o}C was 91 ± 22%, (mean ± σ) that of the first with an observed range of 42 to 130% (Table 12). This may imply that it is possible for an enzyme molecule in a given active conformation to be thermally converted to a different conformation with either a higher or lower activity, although conversion to that with a lower activity may occur more often.

The temperature of the middle region of the capillary was increased to 50^{o}C to determine if the loss of activity for the 13 molecules mentioned above was due to thermal denaturation. Figure 50 depicts a typical assay for this type.

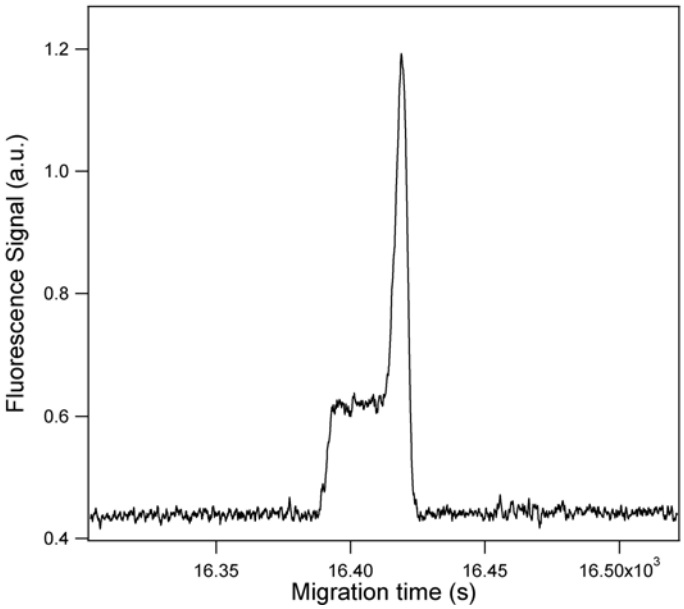


Figure 50. Continuous assay at 27°C, 50°C and 27°C. A single β-galactosidase molecule was continuously migrated at 300 Vcm^{-1} in 100 μM DDAO-gal in separation buffer. First and third portions of plateau incubated at 27°C; the middle region incubated at 50°C.

When the enzyme molecule reaches the heated central region of the capillary the signal increases dramatically with the higher temperature, but then is followed by a rapid drop back to the inter-plateau baseline. This implies a sudden loss of activity and suggests that thermal denaturation may be rapid and catastrophic. Of the 53 molecules assayed, 45 showed a complete loss of signal at the elevated temperature. For the 8 molecules that did not denature, every one showed a decrease in activity during the second incubation at 27^{o}C, with an average decrease during the second incubation to 56 ± 10% of the first and a range of 43 to 69% (Table 12).

Table 12. Summary of the fraction of enzyme molecules retaining activity and the residual activity after incubation at two different elevated temperatures.

	Incubation temperatures	
	27/45/27°C	**27/50/27°C**
# of molecules retaining activity	69	8
# of molecules denatured	13	45
Residual activity (%)	91 ± 22	56 ± 10
Range (%)	42–130	43–69

9.4. Discussion

9.4.1. Advantages of the continuous assay

The new assay is more efficient and more versatile than the previous assays. The original single molecule CE-LIF assays were limited to measuring only time-averaged activity. The range of the single molecule assay was expanded by the work in chapter 8 to measuring the electrophoretic mobility of individual molecules. The new single molecule continuous flow assay accesses the same information that the previous incarnations of the assay were able to obtain, and now extends to the continuous measurement of activity over time. As shown, this permits increased flexibility of assay

conditions under which a single enzyme molecule can be studied, and allows for new insights into the properties of individual molecules. The non-continuous flow single β-galactosidase molecule assays are labor intensive and require constant operator input. For every run the operator must prepare fresh enzyme and substrate that is injected electrokinetically by placing the capillary into a specific injection vessel, or manually by pressure using a syringe. Some automation is possible as runs can be programmed, but because the duration of runs is usually brief, the operator is generally in continual attendance. In contrast, the new continuous flow assay, once set up, runs independently for several hours. The principal limitation to the duration of the assay is the steady rise in the background signal because of the non-enzymatic hydrolysis of the substrate.

9.4.2. Thermally-driven activity fluctuations

The energy required for enzymatic conformational fluctuations originates from its surroundings, and the continuous assays performed at 24°C suggest that at that temperature the energy required to generate them is limited. If individual enzyme molecules are continuously interconverting between conformers with discrete rates as suggested by proponents for dynamic heterogeneity, it is below the temporal resolution of the continuous assay. A similar result was obtained by Rissin *et al.* [18] who monitored the activity of individual molecules of β-galactosidase confined within 40 fL wells hydrolyzing RES-gal, and found that the activity of individual molecules was a stable property. Their assays continuously measured the aggregate resorufin produced and therefore were not sensitive to short-term fluctuations of activity as the technique here, but if substantial changes of activity had occurred, they would have been detected as alterations to the slope of the product production curve.

Fluctuations of activity occurred when enzyme molecules were mobilized through a capillary heated to 40°C (Figure 47). This suggests that the absence of activity fluctuations at lower temperatures was not because the assay cannot detect them, but rather there was a lack of energy in the system to break weak interactions to allow for the formation of different interactions that produce variable catalytic activity. The constantly changing activity suggests that the enzyme molecules were frequently breaking and forming new interactions. Subjecting the same molecule to two different incubation temperatures where fluctuations of activity were observed for the higher incubation temperature, and not the lower, (Figure 48) affirms that any particular molecule is capable of fluctuating activities if the system has enough thermal energy. The results for the triple incubation where activity is measured at ambient temperature, heated, and then reassayed at the first temperature (Figures 49 and 50) indicates that the same molecule can support different catalytic rates at near room temperatures if additional energy is supplied to the system to drive conformational transformations. Similarly, the switching rates for GFP mut2 between different states was a stable property for a particular molecule at room temperature, but could be perturbed to adopt new switching rates following exposure to denaturing conditions [63]. It is unclear however if the thermally-driven changes observed here represent the same phenomenon as the conformational fluctuations associated with dynamic heterogeneity which are inferred from the trajectory studies. The fluctuations for the latter are based upon experiments that are conducted at room temperature, and the time scale of interconversion is estimated to range from ms to tens of seconds [5, 58]. Activity fluctuations are seen here only at elevated temperatures, and the time scale of the fluctuations is longer—though extremely rapid interconversions

would not be detectable by this system. Also, the interconversions detected here were biased towards lower activities and suggests that structural degradation was contributing to the differences. The results for the experiments conducted at 50°C further suggest that this is the case. A final basis upon which to query whether the same phenomena are being observed is the transformations here required the input of considerable thermal energy, and once the transformation had occurred, the enzyme exhibits stable activity at the new rate. This result is more consistent with the stable conformation differences proposed by Xue and Yeung [2]. The implications of this latter possibility are discussed further in the next section.

9.4.3. Relevance for time-averaged assays

The continuous assays performed at 24°C provide two important insights for time-averaged single molecule β-galactosidase assays performed at near room temperature. First, because the catalytic activity of individual molecules varied little with time, it suggests that under similar conditions time-averaged measurements are indeed representative of the activity throughout the assay. Second, the six-fold range of activities for the molecules at near room temperature affirms that the static heterogeneity findings of time-averaged experiments are unlikely to be statistical artifacts of assaying individual molecules at the limits of their intrinsic catalytic ranges. The range of activities is lower than that of previous experiments, but the number of molecules assayed was also lower, so the rare highly active and low activity molecules were probably not assayed.

Although the triple incubation at two different temperature continuous assays provides ambiguous support for the conformational fluctuation hypothesis, they do

provide useful information about ongoing uncertainties for the single molecule field. If the ten-fold variation of activity characteristic of static heterogeneity were solely due to conformational differences as suggested by Xue and Yeung [2], it should be possible given sufficient time to force an individual molecule to migrate through its conformational space and exhibit such an activity range. Or alternatively, single thermal perturbations of a large enough number of individual molecules should produce transformation events that yield some molecules with several-fold differences of activity. The results in Table 12 indicate that at a transforming temperature of 45°C there was a limit to the extent to which a particular individual molecule could be perturbed into a conformational state with different activity, and thereby limits the extent to which long-lived conformational states can be considered as a full explanation for static heterogeneity. The maximum observed increase in activity was only 30%, and it is suggestive that for those molecules that survived the period of elevated temperature at 45 and 50°C that the floor for the change in activity after higher temperature incubation was similar (42 and 43%). The assay was sensitive enough to detect enzyme molecules with lower activity, but they do not seem to be generated. This may indicate that there is a limit to the extent of thermally-induced conformational reorganization that is possible without a catastrophic loss of catalytic capability, and it implies that any particular molecule already possesses its intrinsic near-optimal activity. A single molecule assay of α-chymotrypsin [195] found a step-wise deactivation of molecules punctuated by periods of tens of seconds where no activity was detected. The least active state of an active molecule for α-chymotrypsin was approximately 30% of that when fully active. This α-chymotrypsin experiment was performed at room temperature and suggests that different

molecules deactivate differently, or that spontaneous deactivation proceeds in a different manner than thermal denaturation.

9.5. Conclusion

A new single molecule assay was devised that allows for the continuous measurement of catalytic activity of β-galactosidase while simultaneously measuring electrophoretic mobility. Assays performed at 24°C suggest that time-averaged single molecules studies are representative of the activity of the molecule throughout an assay, and that purported conformational fluctuations do not affect the average activity within the temporal resolution of assays this type. Experiments performed at elevated temperatures indicate that fluctuations in activity occur, and these variations are possibly attributable to thermally interconverting conformers with different catalytic activity. The variation of activity for different thermal conformers is much lower than that observed for a population of single molecules from the same source, and implies that conformational fluctuations cannot fully account for static heterogeneity.

10. Contribution of translation error to catalytic and electrophoretic heterogeneity

The data presented in this chapter has been accepted for publication as:

Nichols, E.R., Shadabi, E., Craig, D.B. (2009) Effect of alteration of translation error rate on enzyme microheterogeneity as assessed by variation in single molecule electrophoretic mobility and catalytic activity. *Biochemistry and Cell Biology*, in press.

10.1. Introduction

In Chapter 8 it was demonstrated that it is possible to measure the electrophoretic mobility of individual molecules of β-galactosidase using a CE-LIF separation and detection system. Different individual molecules were shown to have distinct, different and reproducible electrophoretic mobilities. This mobility heterogeneity was shown to be independent of the method of measurement, source of the enzyme and the type of capillary coating. The range of the measured mobilities, and a simple expression for describing the electrophoretic behavior of proteins formed the basis for a hypothesis that charge heterogeneity arising from translation error could account for the differences. Because primary sequence changes of the proteins that constitute the β-galactosidase tetramer could potentially alter substrate binding, transition state stabilization, protein stability and monomer interactions, and that such changes would likely alter the activity of a particular molecule, it was inferred that error was also a likely source of catalytic heterogeneity.

Ribosomal function and protein translation error have been a research area of abiding interest and importance [191,200]. Expression of protein from the DNA template is a multi-step process and errors can occur at any stage. The error rate for the transcription of *E. coli* genes to mRNA by RNA polymerase has been estimated at 1.4 x 10^{-4} per codon [190]. Misacylation, the charging of a tRNA with an incorrect amino acid,

has been estimated at 2.5 x 10^{-5}[201]. The most significant source of error is misreading whereby the ribosome incorporates an incorrect amino acid into the polypeptide during translation. The *E. coli* misreading rate has been estimated at 3.4 x 10^{-4} per codon [165], and this implies that a considerable fraction of larger proteins may contain translation errors. Although some β-galactosidase with N-terminal substitutions have been shown to be susceptible to proteolysis [159], protein synthesized under starvation conditions for histidine and asparagine which result in high levels of glutamine and lysine substitutions do not result in higher levels of protein degradation [202]. This implies that proteins with misreading errors are not extensively degraded in *E.coli*. Considering that β-galactosidase is a 465 kDa homotetramer comprised of 4092 amino acids [104], the translation error rate suggests that every enzyme molecule has at least one translation error, and possibly other errors arising from the different sources discussed above. Because β-galactosidase mutations that result in single amino acid substitutions have been shown to both increase and decrease catalytic rate, [203] a population of enzyme molecules with myriad different amino acid substitutions should show a constellation of activities when assayed at the single molecule level. Single molecule experiments confirm that different molecules exhibit a wide range of activities [45, 18]. If translation error is a significant source of enzyme heterogeneity, it should be possible to test this hypothesis by manipulating error levels and determining the effects upon catalytic and electrophoretic heterogeneity.

The *rpsL* streptomycin resistant mutants exhibit a hyperaccurate translation phenotype due to a mutation of ribosomal protein S12 [204]. This mutation destabilizes the *ram* (ribosomal ambiguity) state with the result that the ribosome favors the

restrictive conformation which is associated with non-cognate tRNA discrimination and ribosomal proofreading [205]. This preference results in fewer translation errors. These mutants confer streptomycin resistance by offsetting its effects rather than preventing streptomycin binding [200]. The above hypothesis predicts that β-galactosidase expressed by an *rpsL* mutant would exhibit a reduction of catalytic and electrophoretic heterogeneity.

The antibiotic streptomycin can be used to increase the translation error rate [206]. This antibiotic is an aminoglycoside that binds the *E. coli* ribosome and impairs bacterial cell growth and division by disrupting protein synthesis [191]. At sub-lethal concentrations streptomycin increases translation error by stabilizing the ribosome in the *ram* state which promotes the binding of non-cognate tRNA, and it impairs proofreading by reducing the difference in the kinetic step that contributes to the discrimination of cognate and non-cognate tRNAs [145]. The hypothesis predicts that an increase in the translation error rates should result in an increase in single molecule catalytic and electrophoretic heterogeneity.

In this chapter the findings for the effect that an increase or decrease of translation fidelity has upon the single molecule electrophoretic and catalytic heterogeneity of β-galactosidase from three *E. coli* sources are presented. This hypothesis is tested by inducing *E. coli* β-galactosidase expression at various concentrations of streptomycin in a wild type *E. coli* (ATCC 35321), and in a strain with a *lacZ* mutation (ATCC 8677), and assaying the molecules with CE-LIF. Because any translation error for the strain with the *lacZ* mutation will be superimposed upon the DNA sequence error, it will serve as a model for enzyme with greater polypeptide primary sequence error. The effect that a

decrease in translation error has for catalytic and electrophoretic heterogeneity of β-galactosidase is examined by measuring these properties of enzyme from an *rpsL* mutant.

10.2. Methods

10.2.1. *E.coli* growth conditions and β-galactosidase thermolability

E. coli strains ATCC 35321 and ATCC 8677 were grown for 18 hours under the conditions described in the materials and methods, aliquots were removed and transferred to equal volumes of growth media that contained 0, 2, 10, 25, 50, 150 and 300 µg/mL streptomycin and cultured for an additional hour, following which β-galactosidase was induced by IPTG for 30 minutes. The *E. coli rpsL* mutant ATCC 33588 growth conditions were the same except cells were cultured and β-galactosidase was induced in LB media that contained 100 µg/mL streptomycin. β-Galactosidase was harvested and stored as described in chapter 4.5.2.

The sensitivity of the strains ATCC 8677 and 35321 to streptomycin was determined by measuring bacterial growth in cultures that contained the same concentrations of streptomycin used for induction. Growth was determined by following changes to cell density at various times intervals by measuring scatter at 550 nm. Susceptibility of the streptomycin resistant strain ATCC 33588 to the antibiotic was evaluated by its growth in the presence of 100, 200 and 300 µg/mL streptomycin.

The thermolability of the β-galactosidase from the three sources was determined by measuring catalytic activity using oNPG as a substrate after 5 and 15 minute exposure to denaturing temperatures.

10.2.2. CE-LIF Instrumentation and single molecule assays

The single molecule β-galactosidase mobility and activity assays were performed in 50 cm long and 10 μm inner diameter uncoated capillaries, and used the uncoated extended separation protocol based upon the DDAO-gal substrate described in detail in chapter 8 with the following differences: the HEPES concentration was 10 mM; the final DDAO-gal concentration was 50 μM; and the assay time was 3 minutes. Calculations for electrophoretic mobility were as described in chapter 8, and single molecule catalytic activity was determined as before.

10.3. Results

10.3.1. Effect of streptomycin on *E.coli* growth and β-galactosidase induction

Streptomycin had a pronounced, dose-dependent effect upon the growth of the streptomycin sensitive *E. coli* strains ATCC 8677 and ATCC 35321. Figure 51 indicates that a streptomycin concentration of 2 μg/mL was sufficient to significantly reduce the growth of strain ATCC 8677, and 10 μg/mL was sufficient to nearly completely inhibit growth. Strain ATCC 35321 was less sensitive to the effects of streptomycin as 25 μg/mL of the antibiotic was required to measurably reduce the growth rate. The growth of ATCC 35321 was almost completely inhibited at 50 μg/mL. As expected, the growth of the streptomycin resistant *rpsL* mutant, ATCC 33588, was not affected by streptomycin concentrations of up to 300 μg/mL, (data not shown).

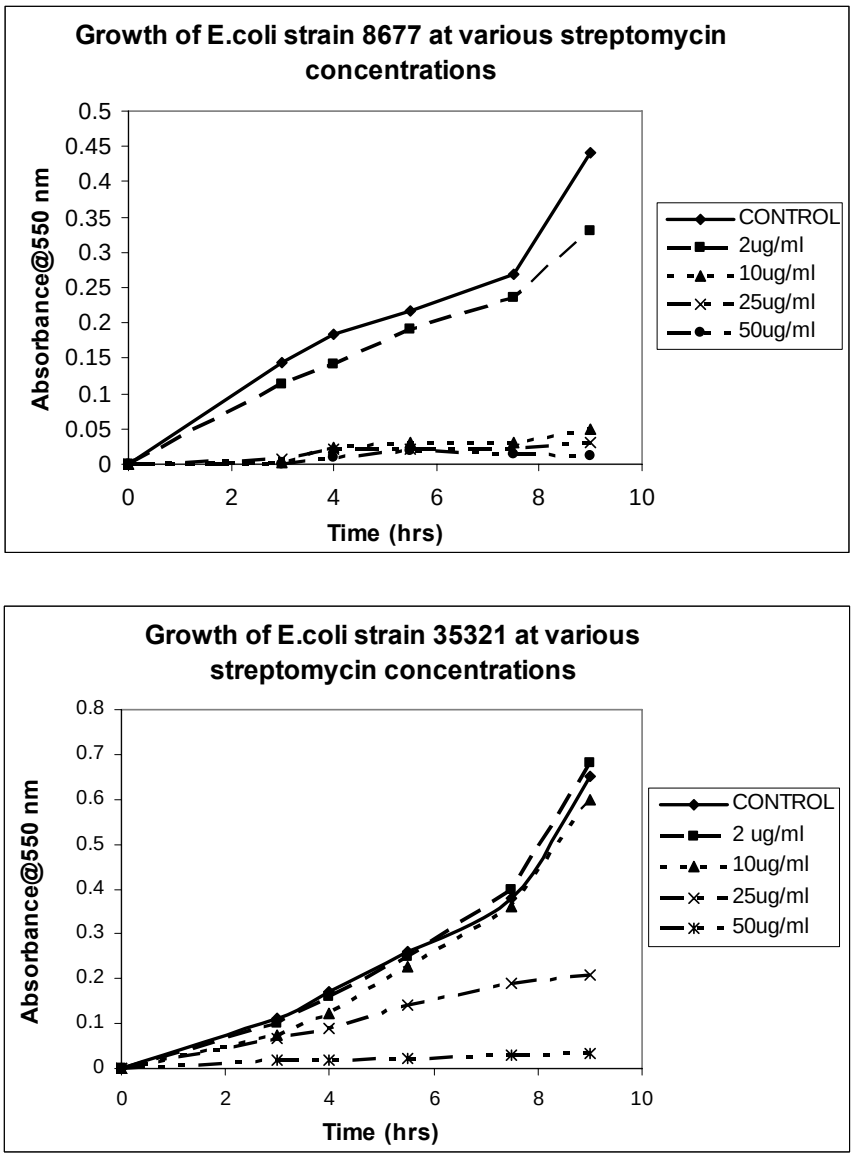


Figure 51. Effect of streptomycin on *E.coli* growth. The change of culture cell density measured by absorbance at 550 nm for *E.coli* strains ATCC 8677 and ATCC 35321 in the presence of varying concentrations of streptomycin.

The specific basis for the difference of streptomycin susceptibility between ATCC 8677 and ATCC 35321 is unclear, but phenotypic differences between ribosomes of naturally occurring *E. coli* have been reported [207].

The β-galactosidase yield from IPTG induction for the two streptomycin sensitive strains under varying concentrations of streptomycin exhibited a pattern that paralleled the growth curves (Figure 52).

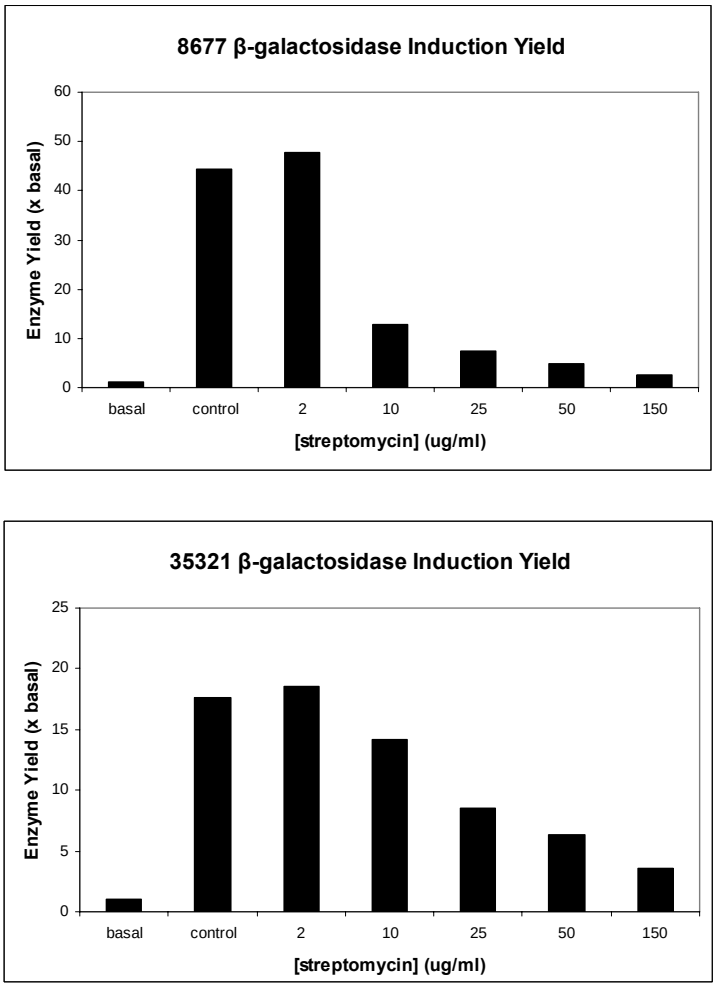


Figure 52. Effect of streptomycin on β- Galactosidase yield. *E. coli* cultures induced with 250 μM IPTG for 30 minutes. Enzyme yield is stated as a multiple of basal enzyme levels, and was determined on the basis of the average number of molecules per run, the volume of a capillary and the required dilution factor. Enzyme yield is unadjusted for differences in cell density.

The enzyme yield for strain ATCC 8677 dropped approximately four-fold from the control to 10 μg/mL streptomycin—the concentration of inhibitor at which growth also declined. The yield for strain ATCC 35321 fell two-fold at the antibiotic concentration of 25 μg/mL, and this was also the concentration of inhibitor that began to substantially

reduce growth for this strain. Both strains exhibited continued β-galactosidase synthesis (although at reduced levels) at concentrations of streptomycin higher than was required to completely impair measurable culture growth. This result was expected because the integrated metabolic circuits required for growth and reproduction are dependent upon protein synthesis. Small impairments to protein synthesis will still yield functional products, but their deficiencies will cascade with catastrophic consequences up through the higher order processes that depend upon them [208]. It is unlikely that the β-galactosidase molecules synthesized at the higher streptomycin concentrations are attributable to a brief period of high level induction that occurred prior to the effects of the inhibitor becoming manifest. Since the bacterial cells were exposed to the inhibitor for an hour prior to induction, and because previous studies have shown that streptomycin affects protein synthesis almost immediately upon introduction to the growth media [209, 210], it is probable that the protein expressed during the half hour induction was synthesized under the full effect of the streptomycin treatment. At streptomycin concentrations greater than 50 μg/mL, the enzyme yield for both strains was less than five fold that of basal levels. Because it was not possible to physically separate enzyme synthesized before and after the addition of streptomycin, the results presented hereafter will be restricted to streptomycin treatments of 50 μg/mL or less.

The induced enzyme yield for the hyperaccurate *rpsL* mutant exceeded basal levels by only a factor of two. For comparison, the yield for strain ATCC 8677 in the absence of streptomycin was approximately 50-fold higher than basal levels. This depressed yield is likely attributable to the reduced translation elongation rates associated with the *rpsL* restrictive mutants [211].

10.3.2. Enzyme Thermolability

The thermolability of the induced β-galactosidase was used to indirectly determine if the streptomycin was producing an increase in the frequency of translation errors. Enzymes require sufficient conformational flexibility under physiological conditions to bind substrate(s) and release product(s), and therefore a structural compromise is required between stability and flexibility. Error in the primary polypeptide sequence diminishes tertiary and quaternary structural stability, and consequently a sample of protein molecules that contains a high degree of error should be more susceptible to thermal denaturation than a population with relatively less error [209]. For example, a G794D substitution in β-galactosidase leads to increased thermolability of the enzyme [203].

Thermolability assays for β-galactosidase have been previously conducted at 53-55°C, [203, 209] but it was found that higher temperatures were required to denature a significant proportion of the enzyme from the *E.coli* strains used here. The different strains exhibited different susceptibility to thermal denaturation (Table 13). A temperature of 56°C caused a 50% loss of activity for the D280N mutant enzyme from ATCC 8677 after 5 minutes. If the temperature was raised to 60°C, 80% of the activity was lost after two minutes, and a total loss of activity occurred within six minutes, (data not shown).

Table 13. Thermolability of β-galactosidase from different sources and treatments at various denaturing temperatures.

			Residual Catalytic Activity[a]	
strain	[streptomycin] (μg/mL)	Temp °C	5 minutes	15 minutes
33588	none	62	0.79	0.78
33521	none	62	0.52	0.32
"	2	"	0.53	0.30
"	10	"	0.37	0.15
"	25	"	0.48	0.35
"	50	"	0.48	0.40
8677	none	56	0.51	0.44
"	2	"	0.52	0.41
"	10	"	0.52	0.31
"	25	"	0.45	0.27
"	50	"	0.41	0.25

Note [a] Residual activity represents the ratio of activity that remains after sample incubation at 5 or 15 minutes at specified temperatures to the activity that was present prior to heating the sample. Activity is the rate of change of absorbance at 406 nm of 1.5 mM ONPG in separation buffer. Assays were performed in triplicate.

To attain a 50% reduction of the activity of wild type β-galactosidase from ATCC 35321 after five minutes required a temperature of 62°C; however this temperature caused only a 20% loss of enzyme activity for the hyperaccurate ATCC 33588. These results affirm the general proposition that translation error and thermostability are correlated.

The difference of thermolability for the two streptomycin sensitive strains may be attributable to the D280N mutation that is present in the ATCC 8677 strain. This residue lies on a loop than extends from one monomer to the adjacent monomer and contributes to the stability of the active site [104]. These results suggest that it may also contribute to quaternary stability. Thermolability for ATCC 8677 β-galactosidase increased for enzyme synthesized at higher streptomycin concentrations and indicates that more translation error was present in these sources. The results were somewhat ambiguous for

ATCC 35321 as an inexplicable reduction of thermolability emerged at streptomycin concentrations of 25 μg/mL and higher.

10.3.3. Single molecule assays

Figure 53 depicts a typical electropherogram for a 6 minute separation and 3 minute incubation of β-galactosidase from *E. coli* strain ATCC 8677 in 50 μM DDAO-gal. The initial large peak is the 1 x 10^{-8} M DDAO standard and the succession of smaller peaks that follow represent the pools of product that formed around each enzyme molecule during the static incubation.

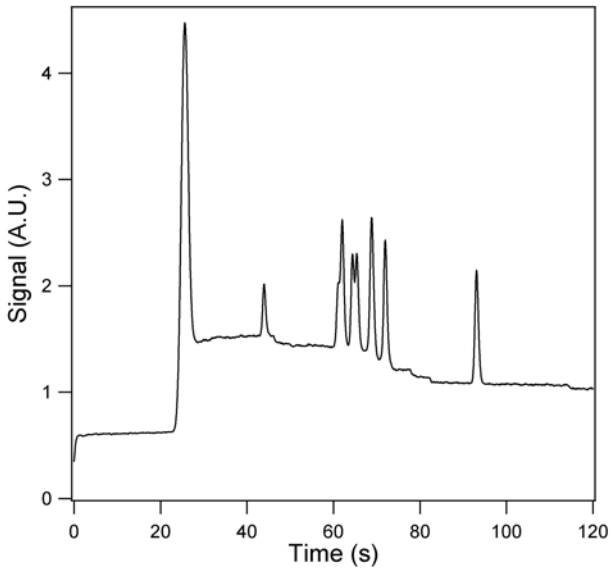


Figure 53. Separation of 8677 β-galactosidase in 50 μM DDAO-gal. Five s, 100 Vcm^{-1} injection of ~1 x 10^{-13} M enzyme and 1 x 10^{-8} M DDAO in separation buffer followed by a 6 minute separation at 400 Vcm^{-1} and 3 minute incubation in 50 μM DDAO-gal in 10 mM HEPES, 1 mM $MgCl_2$ and 1 mM citrate.

The peaks sit atop a plateau of non-enzmatically formed DDAO that was present due to the background hydrolysis of the substrate into product, and from the incomplete extraction by the toluene of the residual DDAO impurity that was originally in the substrate. Comparison of the area of the peaks produced by the individual enzyme

molecules to the area of the injected DDAO reference standard is the basis for calculating the catalytic activity of the individual enzyme molecules. The distance between the individual enzyme peaks and the reference standard is used to calculate the electrophoretic mobility of the molecules. Repeated separations and incubations have shown that the different mobilities and activities of individual molecules are stable and reproducible properties (chapter 8).

The catalytic and electrophoretic results for the individual β-galactosidase molecules from a single source can be assembled into scatterplots that relate the activity of each individual enzyme molecule to its mobility (Figure 54).

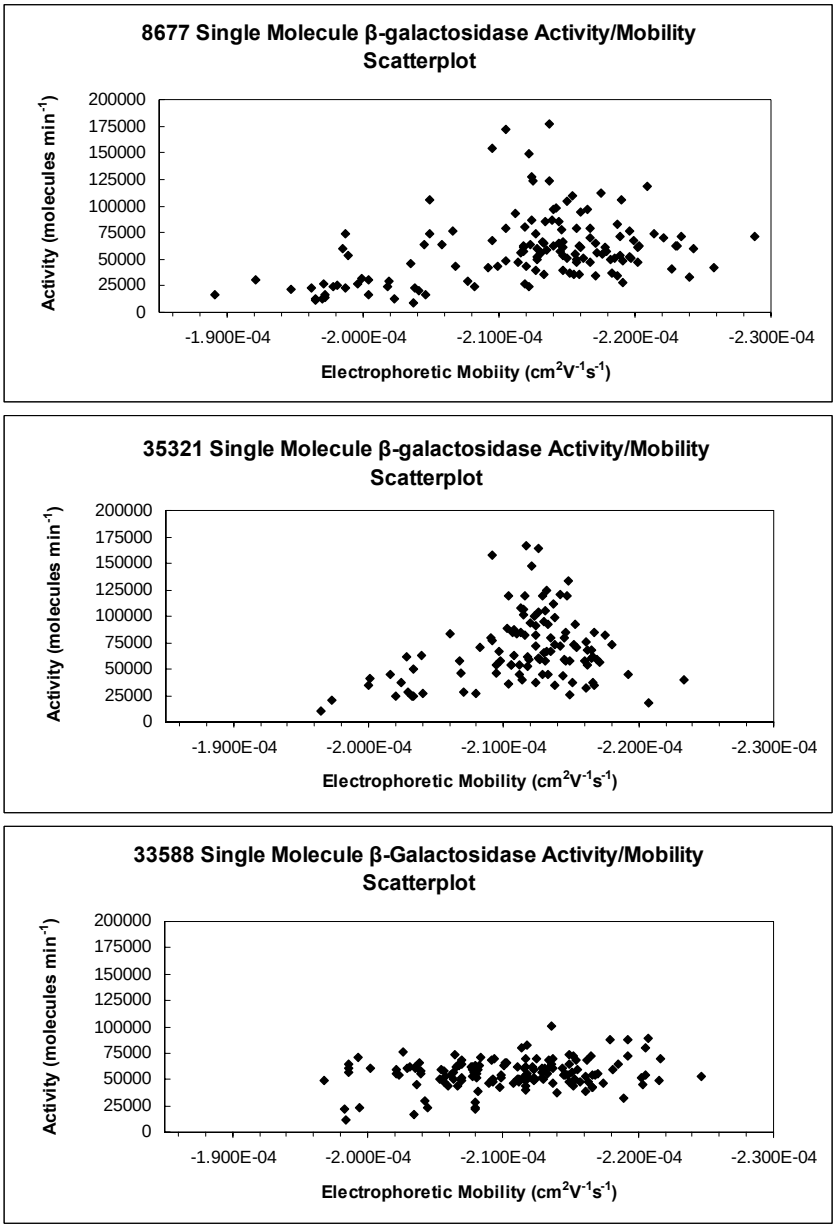


Figure 54. β-galactosidase activity and mobility scatterplots. Activity assays and separations were performed by CE-LIF in 10 mM HEPES, 1 mM $MgCl_2$ and 1 mM citrate. Activity is the rate of hydrolysis of 50 μM DDAO-gal per individual enzyme molecule. Number of molecules assayed indicated in table 14.

There are notable differences between the strains for the extent of apparent scatter in the data, with a ranking 8677 > 35321 > 33588. For enzyme from strains ATCC 35321 and 8677, there is little correspondence between activity and mobility for enzyme molecules that have mobilities from -2.09×10^{-4} $cm^2V^{-1}s^{-1}$ to -2.15×10^{-4} $cm^2V^{-1}s^{-1}$ (r^2 values for

35321 and 8677 are 0.1202 and 0.0019 respectively), but outside this range activity drops precipitously as the difference in mobility from this range increases. The scatterplots show that there are relatively few molecules from 35321 and 33588 with mobilities higher than – 2.20 x $10^{-4}cm^2V^{-1}s^{-1}$. Enzyme from 8677 does however show an abundance of molecules with higher mobility. For all three strains lower activity, lower mobility molecules are common. The low activity, low mobility molecules are likely those molecules that have accumulated sufficient structural differences of either charge or conformation that both mobility and activity are reduced. The assay is sufficiently sensitive to detect individual molecules that are substantially less active than the least active ones detected here. Their absence indicates that a catastrophic loss of activity occurs at some structural threshold.

The range of electrophoretic mobility (highest mobility/lowest mobility) is a proxy for electrophoretic heterogeneity, and for ATCC 35321 the value is 1.14. This is consistent with the range of mobilities that have been previously measured using CE-LIF (chapter 8&9). The range of electrophoretic mobilities for ATCC 8677 is 1.21.

The magnitude of the values for electrophoretic mobility here (approx. –2.1 x 10^{-4} $cm^2V^{-1}s^{-1}$) are higher than those reported in the chapters 8 and 9 (approx. –1.75 x 10^{-4} $cm^2V^{-1}s^{-1}$) because these experiments were performed at 10 mM HEPES rather than 50 mM. The ionic strength was reduced for these experiments to enhance the differences in the apparent electrophoretic mobilities between different enzyme molecules, and between the enzyme molecules and the DDAO reference standard. A reduction in ionic strength increases the protein's zeta potential with the result that the electrophoretic mobility of the protein is increased. The lower ionic strength also increases the electrophoretic

mobility of the DDAO, but the effect is less pronounced. If the average electrophoretic mobilities measured here are used to determine the average net charge on the protein [212], and this valence is used to recalculate what the mobility would be if the ionic strength for the current and previous experiments were the same, the mobilities agree to within 3%.

10.3.4. Effect of translation error rate

Figure 55 is a typical electropherogram showing the separation and incubation of five β-galactosidase molecules from the *rpsL* mutant. The similarity of the area under the peaks indicates that the five molecules were catalytically similar, whereas their separability indicates that they were electrophoretically distinct. For contrast, the electropherogram for enzyme from ATCC 8677 depicted in Figure 53 reveals significant catalytic and electrophoretic heterogeneity within a single run. The scatterplot for 33588 (Figure 54) shows a marked reduction in catalytic heterogeneity compared to the other strains, but little evident difference of electrophoretic heterogeneity. The range of mobilities for the *rpsL* mutant is 1.14, which is the same as the enzyme sample from ATCC 35321.

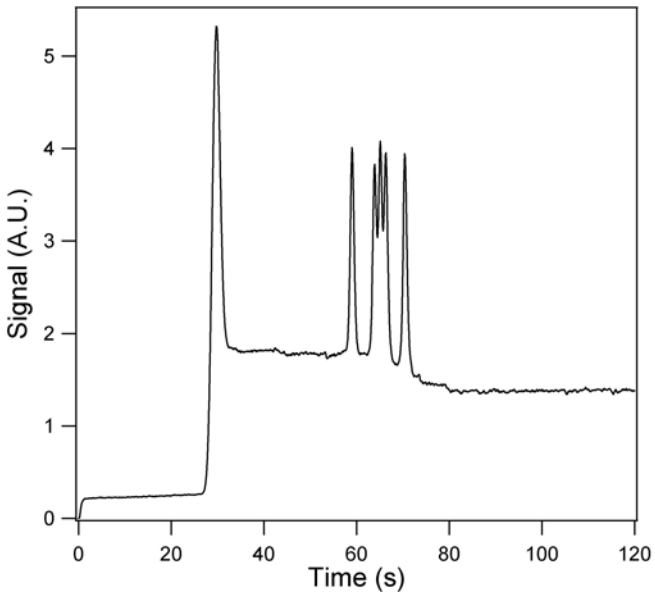


Figure 55. Separation of 33588 β-galactosidase in 50 μM DDAO-gal. Five s, 100 Vcm-1 injection of ~1 x 10^{-13} M enzyme and 1 x 10^{-8} M DDAO in separation buffer followed by a 6 minute separation 3 minute incubation at 400 Vcm-1 in 50 μM DDAO-gal in 10 mM HEPES, 1 mM MgCl2 and 1 mM citrate.

Table 14 presents the average single molecule catalytic and electrophoretic data for the three strains. The average rate for the hydrolysis of DDAO-gal per individual β-galactosidase molecule was (mean ± σ) 68,900 ±32300 minute^{-1} for ATCC 35321, 58,900 ± 31,900 for ATCC 8677, and 56,700 ± 13,800 for the *rpsL* mutant ATCC 33588. The average catalytic activity of enzyme from ATCC 35321 is significantly different from ATCC 8677 enzyme ($p = 0.016$) and the *rpsL* strain ($p < 0.0005$).

Table 14. Average single β-galactosidase molecule activity and electrophoretic mobility

Enzyme Source	N	Activity (molecules minute^{-1})	Electrophoretic mobility (x 10^4 cm^{2}V^{-1}s^{-1})
ATCC 35321	109	68,900 ± 32,300	– 2.12 ± .05
ATCC 8677	136	58,900 ± 31,900	– 2.12 ± .08
ATCC 33588 (*rpsL*)	147	56,700 ± 13,800	– 2.11 ± .06

There is no significant difference of activity between ATCC 8677 and 33588 ($p = 0.27$). The *F*-statistic for the comparison of standard deviations was used to evaluate if there were significant differences in heterogeneity of activity or mobility between the samples [213]. A significant difference of catalytic variation between the mutant *rpsL* strain and ATCC 8677 ($F = 5.36$) and ATCC 35321 ($F = 5.49$) was found. This suggests that a reduction of the translation error rate does reduce catalytic heterogeneity.

The average electrophoretic mobility for individual β-galactosidase molecules from strains ATCC 35321 and ATCC 8677 were $-2.12 \pm 0.05 \times 10^{-4}$ $cm^2V^{-1}s^{-1}$ and $-2.12 \pm 0.08 \times 10^{-4}$ $cm^2V^{-1}s^{-1}$ respectively, and for the *rpsL* mutant the mobility was $-2.11 \pm 0.06 \times 10^{-4}$ $cm^2V^{-1}s^{-1}$. For comparison, preliminary single molecule electrophoretic mobility results on the instrument used here under identical conditions for the dimeric β-galactosidase from *Lactobacillus bulgaricus* gives a mobility of approximately -1.81×10^{-4} $cm^2V^{-1}s^{-1}$ (results not shown). The average electrophoretic mobility for individual molecules of β-galactosidase from ATCC 35321 was not significantly different from the *rpsL* mutant ($p = 0.11$) or ATCC 8677 ($p = 0.76$). There was also no significant difference between ATCC 8677 and ATCC 33588 ($p = 0.11$). The *F*-statistic indicated that there were no significant differences between the standard deviations for electrophoretic mobility for any of the three strains. These results indicate that a reduction of translation error does not measurably contribute to a reduction of single molecule electrophoretic heterogeneity.

10.3.5. Effect of streptomycin and translation fidelity

Table 15 presents the average catalytic activity and average electrophoretic mobility results for the single β-galactosidase molecule assays for ATCC 8677 and

ATCC 35321 at varying concentrations of streptomycin. For ATCC 35321 the average catalytic activity was lower for all concentrations of antibiotic, and significantly lower for streptomycin concentrations greater than 2 μg/mL: 10 μg/mL, p = 0.036; 25 μg/mL, p = 0.0018 and 50 μg/mL p = 0.0083.

Table 15. Effect of streptomycin on average β-galactosidase catalytic and electrophoretic mobility. Activity is the rate of hydrolysis of 50 μM DDAO-gal per individual enzyme molecule. Mobility data is based upon separations performed in 10 mM HEPES (pH 7.3), 1 mM $MgCl_2$, and 1 mM citrate.

Enzyme source	N	[Streptomycin] (μg/mL)	Activity (molecules $minute^{-1}$)	Electrophoretic Mobility (x 10^4 $cm^2V^{-1}s^{-1}$)
ATCC 35321	109	none	68,900 ± 32,300	– 2.12 ± 0.05
"	118	2	61,400 ± 36,900	– 2.11 ± 0.06
"	112	10	60,400 ± 28,100	– 2.11 ± 0.06
"	111	25	55,400 ± 31,300	– 2.11 ± 0.06
"	135	50	58,100 ± 31,200	– 2.11 ± 0.05
ATCC 8677	136	none	58,900 ± 31,900	– 2.12 ± 0.08
"	143	2	49,000 ± 27,200	– 2.09 ± 0.10
"	142	10	44,100 ± 26,000	– 2.09 ± 0.10
"	121	25	45,000 ± 26,700	– 2.10 ± 0.10
"	140	50	42,200 ± 27,000	– 2.07 ± 0.12

There were no significant differences in average electrophoretic mobility for any of the ATCC 35321 samples treated with streptomycin compared to the control. For strain ATCC 8677 the average catalytic activity was significantly lower at all concentrations of streptomycin tested (all p values were < 0.01). Enzyme from strain 8677 did show significant differences of average electrophoretic mobility compared to the control at streptomycin concentrations of 2, 10 and 50 μg/mL (p values all lower than 0.025; p value for the 25 μg/mL treatment was 0.075). Comparisons of standard deviations using

the *F*-statistic indicated that there were no significant differences for catalytic or electrophoretic heterogeneity at any level of streptomycin exposure.

10.4. Discussion

10.4.1. Effect of translation accuracy for catalytic heterogeneity

Single molecule β-galactosidase catalytic heterogeneity was significantly reduced by the hyperaccurate *rpsL* mutant. Because enzyme from this source exhibited the lowest thermolability, and because every source of β-galactosidase that has been studied at the single molecule level exhibits multi-fold differences of the catalytic activity of single molecules, it is not likely that the reduction here is attributable to strain specific effects separate from the hyperaccurate phenotype. This result implies that translation error is a significant contributor to the observed catalytic heterogeneity of *E. coli* β-galactosidase. Error arising from transcription may explain why hyperaccurate translation does not extinguish catalytic heterogeneity to a greater extent.

An increase in error rate did lead to an increase in thermolability and a reduction in the average β-galactosidase activity for both streptomycin sensitive strains, but did not result in a significant increase in catalytic heterogeneity. The reduction of average catalytic activity is not surprising because on balance, errors are more likely to lower catalytic activity than raise it, though there is a report of a β-galactosidase mutation increasing activity [203]. Substitutions that replace residues critical for catalysis would likely extinguish measurable activity completely [114]. Less significant substitutions could affect subtle changes to the orientation of residues important for substrate binding, transition state stabilization and product release that result in a lowering of catalytic activity [214]. Efforts to generate greater catalytic heterogeneity may have been

unsuccessful because an enzyme the size of β-galactosidase may be reaching the limits of functional heterogeneity. The range of catalytic activity for the 8677 and 35321 controls was 15-fold, which is similar to previously reported ranges [46. 18]. The introduction of additional error may yield more inactive or extremely low activity enzyme that cannot be detected at the single molecule level. If there was an abundance of extremely low active molecules, their influence might be measurable in the oNPG bulk assays. The results were inspected for this possibility, but there was no evidence of their presence.

10.4.2. Effect of translation accuracy for electrophoretic heterogeneity

The hypothesis predicted that electrophoretic mobility heterogeneity would increase or decrease in tandem with error levels. The results presented here indicate otherwise as electrophoretic mobility was a robust property, and exhibited little variation between sources and treatments. Streptomycin had no effect on the average mobility or heterogeneity of mobility for 35321, although it did significantly reduce the average mobility for 8677. Mobility heterogeneity of enzyme from the hyperaccurate *rpsL* mutant was statistically indistinguishable from that of the other strains. These results suggest that mobility heterogeneity may be an inherent property of large multimeric structures, or that other protein degradative processes such as deamidation and oxidation. This leaves unanswered the question of what causes this heterogeneity, and what magnitude of structural changes is necessary to generate it.

10.4.3. Modelling free zone protein electrophoresis

Estimation of the structural variation that is necessary to account for the observed single molecule electrophoretic heterogeneity can be made by utilizing established

models for protein electrophoretic mobility. The electrophoretic mobility of a protein, u_p, in free zone capillary electrophoresis can be expressed as:

$$u_p = \frac{Z_p e f(\kappa r_p)(1+\kappa r_b)\psi}{6\pi\eta r_p\left[1+\kappa(r_p+r_b)\right]} \tag{10.1}$$

where Z_p is net protein charge or valence, r_p is the Stokes radius of the protein, r_b is the radius of buffer ions, e is the electronic charge, η is buffer viscosity, κ is the inverse screening length, ψ is a correction factor for protein asymmetry and $f(\kappa r_p)$ is the Henry function [212]. The inverse screening length is calculated as:

$$\kappa = \sqrt{\frac{8\pi N_A (2.998\text{x}10^9 e)^2 I}{1000 D(10^7 k_B)T}} \tag{10.2}$$

where N_A is Avogadro's number, I is the ionic strength of the separating medium, D is the dielectric constant, k_B is the Boltzmann constant, and T is temperature. The Henry function can be approximated to within 1% using:

$$f(kr_p) = 1 + \frac{0.5}{1+\exp\left[a\{1+\log(\kappa r_p)\}\right]} \tag{10.3}$$

where $a = 2.8$ for $\kappa r_a < 10$ and 2.5 for $\kappa r_a > 10$ [215]. The shape correction factor ψ is given by:

$$\psi = \frac{6\left[B+\ln\{(b+r_b)/b\}\right](1+\kappa r_p+\kappa r_b)}{\left[2/(3S)\right]^{1/3}(1+S)F'(\kappa b)f(\kappa r_p)(1+\kappa r_b)} \tag{10.4}$$

Where B and $F'(\kappa b)$ are graphically determined ratios [212], S is the axial ratio of the protein, and b is:

$$b = \left(\frac{2}{3S}\right)^{1/3} r_p \tag{10.5}$$

The net charge on β-galactosidase can be determined by rearranging equation 10.1 and incorporating the mobility data and the solutions to equations 10.2-10.5 to solve for Z_p..

The average electrophoretic mobility of β-galactosidase from *E. coli* strain ATCC 35321 was – 2.12 x 10^{-4} $cm^2V^{-1}s^{-1}$. Based upon buffer of ionic strength 0.0126 M and viscosity of 0.00089 cP, a gel-filtration determined Stokes radius of 6.9 nm [216], an average buffer radius of 0.25 nm [212], and an axial ratio of 1.31 determined from the enzyme crystal structure [104], the average net charge at pH 7.3 is – 62.9. Calculating net charge in this manner on previously collected mobility data (chapter 8) for β-galactosidase in 50 mM HEPES buffer (10 mM was used here) yields a net charge difference to that reported here of less than 3%. Although software is available to determine the net charge of a protein, its reliance upon the ionization constants of free amino acids yields dubious results [217]. The high negative charge explains the observation that the enzyme has higher electrophoretic mobility than the DDAO fluorescent dye.

10.4.4. Effect of changes of radius and charge for electrophoretic mobility

Figure 54 indicates that the electrophoretic mobility for β-galactosidase from 35321 resides within a range of -2 x 10^{-4} $cm^2V^{-1}s^{-1}$ to -2.25 x 10^{-4} $cm^2V^{-1}s^{-1}$. If the Stokes radius of the protein is fixed at 6.9 nm, the net charge must either fall to – 59.5 or rise to – 66.9 to attain this range. Conversely, if the protein's charge is fixed at –62.9, a 0.1 nm change in the radius alters the mobility to – 2.08 x 10^{-4} $cm^2V^{-1}s^{-1}$ (if the radius increases) or – 2.15 x 10^{-4} $cm^2V^{-1}s^{-1}$ (if the radius decreases). A change of 0.3 nm generates a range of – 1.97 x 10^{-4} $cm^2V^{-1}s^{-1}$ to – 2.30 x 10^{-4} $cm^2V^{-1}s^{-1}$ and would include almost every molecule assayed here. If a hybrid of the two scenarios is considered, the gain or loss of a single charge accompanied by a 0.1 nm alteration to the radius yields an electrophoretic range of – 2.05 x 10^{-4} $cm^2V^{-1}s^{-1}$ to – 2.19 x 10^{-4} $cm^2V^{-1}s^{-1}$; this range would account for 84% of the mobility heterogeneity for 35321. Changes to the molecular radii of 0.1 nm

have been shown to produce similar changes to electrophoretic mobility for the enzyme Staphylococcal nuclease [218]

Net charge has been traditionally viewed as the major determinant for protein mobility [80] and has been exploited to separate near identical molecules [188]. However, seemingly minor amino acid substitutions can have measurable mobility effects also. Site-directed mutagenesis studies of Staphylococcal nuclease showed that the loss of a hydrogen bond by replacing an asparagine with a glycine altered the CE-measured electrophoretic mobility of the mutant by $3 \times 10 \times 10^{-6}$ $cm^2V^{-1}s^{-1}$. A substitution of an alanine with a glycine resulted in a decrease of mobility of 2×10^{-6} $cm^2V^{-1}s^{-1}$ because the conformation of the peptide bond switched from *cis* to *trans* [218]. These substitutions resulted in 0.039 and 0.026 nm radius differences. It is interesting and surprising to note that the loss or gain of a single charge by β-galactosidase changes mobility by approximately 3.5×10^{-6} $cm^2V^{-1}s^{-1}$ which closely approximates the asparagine/glycine substitution discussed above and suggests that significant changes are possible with any amino acid substitution—not just those that entail a change of charge. Because proteins have an integrated and evolutionary optimized structure, most substitutions will be disruptive and result in an increase in the effective radius of the molecule with a corresponding decrease in electrophoretic mobility. This may explain why the electrophoretic mobility distribution skews towards lower mobilities.

10.4.5. β-Galactosidase Structure

The manipulation of the parameters that determine the electrophoretic mobility of β-galactosidase was undertaken to provide a semi-quantitative basis for the structural changes that are required to account for the enzyme's electrophoretic heterogeneity. It is

not known what the particular amino acid substitutions are, or what their effects may be, but the structure of the enzyme has been solved and it does yield clues for possible sources of structural heterogeneity.

β-Galactosidase is comprised of four identical 1023 amino acid monomers that each have 5 independent folding domains. The first 50 residues of each monomer are in an extended conformation making weak contacts with the 1st, 2nd, and 3rd domains from the same monomer, and the first 12 residues of each monomer are highly disordered [105]. There are numerous exposed solvent loops, and the specific exposed loops that consist of residues 578-583, 654-690 and 727-733 are all highly mobile with multiple conformational possibilities. The enzyme is unusual in that its exposed accessible surface area is much larger than would be expected for a protein of its size and quaternary structure. There are also deep solvent channels across the surface of the molecule and interconnected tunnels within and through the enzyme that are accessible to bulk solvent [105]. There are two major interfaces that mediate monomer-monomer, and dimer-dimer interactions required to form active tetramers. The long interface has relatively few points of contact and mediates monomer-monomer interactions that form inactive dimers. The activating interface is more contiguous and mediates the dimer interactions that complete the active site to produce active tetramers. The long interface, because it does not contribute to the active site, and because of its weak interactions, could mediate quaternary level heterogeneity without disrupting the integrity of the active site. It is evident that there are many regions of the molecule, at different organizational levels where structural heterogeneity is known to exist, that can directly affect the frictional radius of the enzyme. If protein sequence heterogeneity due to basal translation error

rates is considered in this context, it is not difficult to envision that small changes, depending on the nature and location of the change, could have a spectrum of structural effects that alter electrophoretic mobility.

The inability of our experiments to detect a substantial translation accuracy role for electrophoretic heterogeneity may reflect the possibility discussed above that β-galactosidase possesses an inherently dynamic structural range that may dominate and hide the error effect. The question remains as to why an improvement of translation accuracy can lead to a reduction in catalytic heterogeneity but not electrophoretic heterogeneity. If structural disorder from conformationally flexible exposed solvent loops and the unstructured region of the first 50 residues is a major contributor to electrophoretic heterogeneity but has little effect upon the active site, catalytic activity can remain relatively homogeneous without a reduction in electrophoretic heterogeneity. Random octameric oligonucleotide insertions into *lacZ* [214] do suggest that peripheral [105] structural changes to the enzyme have modest effects on catalytic activity. Alternatively, the four active sites are constituted mainly by residues from the monomer in which the active site resides, and from a loop from an adjacent monomer across the highly contiguous activating interface. Hyperaccurate translation would provide consistency for these interactions from molecule to molecule that would be evident as catalytic homogeneity. Weak interactions at the long interface could produce quaternary structure conformational flexibility that would affect the frictional radius of the enzyme without unduly disturbing the activating interface and hence the active sites. This combination would yield electrophoretic but not catalytic heterogeneity—the pattern seen for the *rpsL* mutant.

In chapter 8 where the single molecule separations data was initially presented, the observed range in mobilities was attributed to differences of charge between individual molecules. The more sophisticated analysis here suggests that charge heterogeneity is unlikely the major source of electrophoretic heterogeneity. The modeling of the earlier data to determine net charge used a pH independent charge suppression factor specific to β-galactosidase that was designed to adjust calculated protein valences that were based upon the ionization pK_as of free amino acids. This yielded a net charge for the protein that was 38% of the value obtained here. This had the effect of distorting the relative importance of charge variation. Additionally, the previous model did not take into account possible differences in protein conformation and treated the protein as a uniformly dense sphere which cannot adequately accommodate the effects of disordered regions near the surface of the protein described above.

10.5. Conclusion

The electrophoretic mobility and catalytic activity of single molecules of β-galactosidase was determined for three strains of *E. coli* to assess the role of translation error for catalytic and electrophoretic heterogeneity. The rate of translation error was reduced by expressing enzyme using an *E. coli* mutant with a hyperaccurate phenotype, and error was increased by inducing enzyme expression in the presence of the antibiotic streptomycin. A decrease in translation error resulted in a reduction of catalytic heterogeneity and a reduction in thermolability, but no change to electrophoretic heterogeneity. An increase in translation error led to increased thermolability, lower activity, but generally no significant change to catalytic or electrophoretic heterogeneity. Modeling of the electrophoretic behavior of β-galactosidase suggests that variation of the

hydrodynamic radius was the most significant contributor to electrophoretic heterogeneity.

11. General Discussion

A growing body of experimental work has established that enzymes, when examined at the single molecule level, are functionally heterogeneous. Examples of these heterogeneities include reaction rates [2, 3], requirement for cofactors [47], and energy of activation [3]. Iterations of single molecule heterogeneities are routinely published, while the source and significance of these insights languishes. To address that shortcoming, the principal objective of the experiments undertaken here was to establish a foundation for the structural basis and biological sources of single molecule catalytic heterogeneity. Techniques for the determination of molecular structure one molecule at a time are unavailable, and thus an indirect approach of measuring the effect of expressing *E. coli* β-galactosidase under various treatments designed to interfere with specific biological processes was adopted. Any changes to the distribution of the enzyme properties would indicate what possible structural differences and biological sources may underlie catalytic heterogeneity.

In chapter 5, based upon published inferences, the role of limited proteolysis was investigated as a source of structural heterogeneity that is detected as catalytic heterogeneity. No persuasive evidence was found of any influence for proteolysis. In chapter 6 a different strategy was used to identify possible sources of heterogeneity. Rather than designing an experiment to test a specific hypothesis, β-galactosidase was synthesized *in vitro* and *in vivo*, and then assayed at the single molecule level to determine if unidentified cellular processes contributed to heterogeneity. It was found that enzyme synthesized *in vitro* was less active, and had a different distribution of

activities than enzyme synthesized *in vivo*. A consideration of known differences between *in vitro* and *in vivo* synthesis suggested that deficient N-terminal methionine removal, and low translation fidelity associated with *in vitro* translation systems were possible sources of the discrepancies between the sets of proteins. Differential N-terminal methionine removal was detected by mass spectrometry, although it was not possible to identify which particular molecules were defective. N-terminal methionine removal is likely not contributing to catalytic heterogeneity *in vivo*, but its effect *in vitro* suggested that single amino acid differences could alter catalytic properties by changing the precise geometry of the active site. Thus the work in chapter 6 provided a rational basis to examine error as a possible source for single molecule heterogeneity.

Serendipitous observations made during the development of new single molecule β-galactosidase assays for CE-LIF suggested that individual enzyme molecules had different electrophoretic mobilities. This prompted preliminary investigations seeking correlations between mobility and activity. Although no correlations were found, the high reproducibility of the measurements implied that physical information about the individual molecules was being obtained. Pursuant to a suggestion that electrophoretic modelling of the mobility data might yield insights into their structural differences, an extensive effort was made to demonstrate that the individual mobility differences were not attributable to unidentified artefacts associated with the measurements. Upon devising a method to accurately determine EOF on coated and uncoated capillaries, it was shown that individual β-galactosidase molecules had distinct, different and reproducible mobilities; differences that were independent of the source of enzyme, the mode of separation, and capillary conditions. These findings were closely corroborated by the

continuous assay separation method (chapter 9). The demonstration that individual molecules were electrophoretically heterogeneous was interesting, but alone, merely added to the list of the heterogeneous properties of single molecules without deepening an understanding of it. Application of models for protein electrophoretic mobility suggested that the range of observed mobilities were consistent with a small number of amino acid substitutions. Consideration of known *E. coli* translation error rates, and the number of residues in the enzyme reinforced the inference from the cell-free experiments that translation error might contribute to single molecule catalytic heterogeneity.

The translation error hypothesis was tested directly in chapter 10 examining the catalytic and electrophoretic heterogeneity of β-galactosidase from samples with translation error rates either greater or lesser than wild-type *E. coli*. Catalytic heterogeneity was significantly reduced for enzyme that was synthesized by a high fidelity mutant and is the first demonstration of a specific biological basis for single molecule heterogeneity. An increase in translation error lowered average activity, but did not yield greater heterogeneity of activity. This result was somewhat surprising, but suggests that the high level of heterogeneity present in wild-type enzyme is at the functional limits for it. The absence of a continuum of declining activities to the limit of detection provides some support for this hypothesis. No influence of translation error was evident upon electrophoretic mobility heterogeneity. This too was a surprising result, but a consideration of the structure from crystallography studies, and a more sophisticated electrophoretic model indicated that enough conformational diversity may inherently be present to produce variations in the hydrodynamic radius sufficient to produce the observed range of mobilities. Identification of translation error as a source of

catalytic heterogeneity does not itself identify the specific structures producing it, but the range of activities from different site-directed mutagenesis studies, mutant enzymes such as the D280N present for β-galactosidase in strain ATCC 8677, and the cell free experiments does indicate that small changes to the primary sequence have measurable effects. Considering that the D280N substitution mirrors the spontaneous deamidation of asparagine residues, it suggests that common non-specific protein degradative processes such as deamidation and oxidation may also contribute to single molecule heterogeneity.

A secondary objective was to develop a better understanding of the relationship between static and dynamic catalytic heterogeneities, and the design of the continuous flow assay, (Chapter 9), was motivated in part to address this question. With respect to the *E. coli* β-galactosidase, the dynamic heterogeneity hypothesis is unpersuasive on a number of grounds. First, a testable prediction has been made [60], namely, that fluctuations in activity should converge towards an average value over time. This was shown unambiguously not to be the case. Second, the results from the continuous assay experiments provided tenuous support for dynamic heterogeneity. Fluctuations of activity were observed, but only at elevated temperatures where non-physiological strains are placed upon structural stability. The range of activities for single molecules at elevated temperatures where activity fluctuations were present was substantially less than what has been reported for the dynamic heterogeneity experiments performed at room temperature [5]. Furthermore, the room temperature continuous assays showed invariant activity for individual molecules, although fluctuations would have been detectable if they lasted more than a few seconds. With respect to the paper by English *et al*. [5], the experimental design and data interpretation does not properly account for the nature of

the enzyme itself. The experiment was designed such that each product molecule (resorufin) produced by a tethered enzyme molecule resulted in a fluorescent burst in the probe volume that was detected by confocal microscopy. The detector cannot differentiate between turnover events from the four independent active sites, and does not take into consideration that the monomers may have different kinetic properties. Also, the bin size, or sampling window where photons were counted was 500 μs; this means that the maximum detectable rate is 120 000 minute^{-1} which is similar to the average wild-type rate with the RES-gal substrate, (Chapters 5, 6 and 7), at similar substrate concentrations. In view of the fact that there is no reason for the turnovers in the different active sites to be synchronized, multiple product molecules could be present in the probe volume within a single bin time. The authors reported that higher intensity fluorescent bursts were detected, but these were attributed to different paths by the product out of the probe volume. Considering that it is the distribution of time intervals between successive fluorescent bursts that is the basis for the assertion that the enzyme fluctuates between conformations with different catalytic properties, these concerns may qualify the analysis. A final reason to query dynamic heterogeneity was recently published examining this experiment and suggested the results could be accounted for by viewing substrate arrival and catalysis as two discrete stochastic processes [219]. This had the effect of eliminating the requirement for multiple interconverting conformers. Unaddressed is the fact that the catalytic process can be further subdivided into stochastic binding of galactose to Glu537 in concert with dye release, and then a second stochastic step of enzyme degalactosylation to return the enzyme to its original form.

In retrospect it is not surprising that such extensive catalytic heterogeneity has been found at the single molecule level. The integrated dynamic short and long-range interaction networks that govern protein stability [220], affect binding affinities [116] and conformation changes [221] necessary for catalysis, also make enzyme molecules sensitive amplifiers of even slight structural variations [62]. The magnitude of structural perturbations does not always easily correlate with their impact on activity or conformations, as single amino acid substitutions can extinguish activity or induce extensive quaternary rearrangements. Conversely, fusion constructs can leave enzyme activity largely unaltered. The recently described phenylalanine to leucine substitution in human porphobilinogen synthase morpheeins that causes a quaternary rearrangement to hexamers from octomers, shows that even single amino acid substitutions can result in extensive conformational changes [222]. There is no evidence that β-galactosidase is undergoing such extensive rearrangements. Native gels show diffuse bands but no discrete bands, and Rissin *et al.* purified β-galactosidase by gel filtration and found a similar range of single molecule catalytic activities as seen in the various experiments presented here.

There is no reason in principle that the heterogeneities exhibited by single enzyme molecules are not also pertinent to proteins generally. Evidence of extensive heterogeneities has been characterized in enzymes first because they are such effective reporters of structural variation. It has been suggested that single molecule heterogeneity constitutes a physiological 'tax', and it has long been speculated that polypeptide size is partially constrained by translation error. This raises the question of why translation error rates have not been lowered by natural selection. The hyperaccurate mutants demonstrate

that higher translation fidelity is possible--albeit at the expense of rates of translation. High fidelity rapid translation presumably could have been attained unless not attaining it conferred some subtle survival advantage. It has been suggested that modest levels of translation error provides a mechanism by which organisms can explore future evolutionary space without a genomic commitment [223], or adapt to stressful and changing environments through mistranslation of DNA polymerase to accelerate sub-population mutation rates [208]. Thus the catalytic heterogeneity observed for β-galactosidase may simply be an incidental effect where the differences generally have minimal fitness effects [224].

It remains unclear what significance, in any, single molecule heterogeneity has for biological systems. Dynamic heterogeneity has been proposed as a mechanism by which allosteric responses can be attained by individual molecules that do not exhibit it at the ensemble level [225]. This hypothesis suggests that substrate concentration influences the lifetimes of the various catalytically distinct conformations for which the enzyme molecule is competent. At low substrate concentrations the less active, and at higher concentrations the more active conformations predominate—an arrangement that would yield the characteristic S-shape relationship between rate and substrate concentration for allosteric enzymes. This scenario would require that the copy number per cell is low, and that static heterogeneity was present amongst different molecules. If this were not the case, the reaction profiles would be indistinguishable from bulk kinetics.

Single molecule events may contribute to stochastic processes in cells that produce single cell heterogeneity amongst genetically identical cells. In microbiology, the phenomenon of bistability refers to a situation where identical cells in homogeneous

media bifurcate into stable subpopulations when a critical threshold is surpassed for a gene product [226]. It is not difficult to envision how heterogeneity of enzymes present in low or single copy numbers, or variable binding constants for regulator proteins, could contribute to such transformations. This would confer an adaptable heterogeneity to a rapidly dividing system such as bacteria or cancer where physiological bets can be placed, or hedges set against the future. Such processes also have evolutionary and embryological implications.

12. Future Directions

The utility of the development of the single molecule β-galactosidase CE-LIF assay based upon the DDAO-gal substrate was demonstrated by its application for single molecule separations and the continuous assay. The flexibility of this substrate for various assays protocols position it as the preferred substrate for CE-LIF β-galactosidase single molecule assays that will allow some of the questions left unanswered by this work to be addressed in the future.

One rationale for the *in vitro* experiments, (Chapter 6), was the capacity of single molecule assays to determine the actual activity of enzyme by eliminating the uncertainties of reliance upon specific activity to compare different samples. The additional sensitivity to differences in activity this provides can be used to further test the hypothesis that random amino acid substitutions produce heterogeneities of catalytic activity at the single molecule level. Commercial site-directed mutagenesis kits could be used to produce a series of *lacZ* mutants with amino acid substitutions at a variety of residues not already established as affecting β-galactosidase activity. Analysis of the various mutants by CE-LIF would be expected to show a distribution of averages of activity. This approach could also be used to investigate the structural differences necessary to generate the observed electrophoretic heterogeneity, and mimic deamidation by substituting asparagine residues with aspartate.

The results from chapter 10 suggested that the quaternary structure of *E.coli* β-galactosidase may contribute to the heterogeneity of electrophoretic mobility. There is no evident way to test the hypothesis by modifying the enzyme as only tetramers are active. However, the single molecule assays are compatible with β-galactosidases from

other bacteria, and assaying enzyme from species for which the tertiary and quaternary structure is known provides an opportunity to test the hypothesis. β-Galactosidases from other bacterial species are generally less active than enzyme from *E.coli*, but the high sensitivity of the DDAO-gal assay should allow detection and quantification with reasonable analysis times. Preliminary single molecule assays measuring catalytic activity and electrophoretic mobility with the 228 kDa homo dimeric β-galactosidase from *L. bulgaricus* [227] have been performed using a 15 minute assay. The electrophoretic mobility of the enzyme is different from the *E.coli* enzyme, but sufficient data had not yet been collected to evaluate the hypothesis. Other β-galactosidases for which the quaternary organization is known include a 107 kDa heterodimer from *L reuteri* [228]; a 280 kDa homotetramer from *L. helveticus* [229]; and a 145 kDa monomer from *Bacillus. circulans* [230]. Examining these enzymes at the single molecule level will provide an opportunity to evaluate the role of quaternary structure for both catalytic and electrophoretic heterogeneity.

English *et al*. [5] demonstrated that individual β-galactosidase molecules followed Michaelis-Menten kinetics, and Rissin *et al*. [18] suggested that the different activities of individual β-galactosidase molecules were attributable to differences of their respective k_{cat} values. This attribution was based on the observation that the distribution of single molecule activities was substrate concentration independent—a scenario that obtains when the K_m is identical for every β-galactosidase molecule. However it was shown in chapter 6 for bulk spectrophotometric assays with RES-gal that the K_m value for enzyme from strain 8677 with a D280N substitution was lowered to approximately 250 μM from the 330 μM for wild-type enzyme. This indicates that slightly different molecules have

different kinetic properties. This result considered in conjunction with the demonstration that translation error contributes to catalytic heterogeneity implies that different individual molecules are distinct with respect to K_m and k_{cat}. This discrepancy can be addressed by measuring these parameters directly using the extended separation method for DDAO-gal on uncoated capillaries. This can be accomplished by electrokinetically injecting a molecule into the capillary followed by an injection of a zone of substrate for 1 minute. The higher velocity of the substrate due to EOF will cause it to reach the vicinity of the capillary occupied by the enzyme molecule where static incubation can occur. The injection end can be place in a vial of a different substrate concentration and the new substrate mobilized until it reaches the enzyme molecule at which point a second incubation can be performed. If brief incubations (to minimize peak diffusion) and 5 μm diameter capillaries are used, it should be possible to obtain 4 different incubations of the same enzyme molecule from which k_{cat} and K_m can be determined for each molecule. This type of assay could also be used to determine if individual molecules are heterogeneous with respect to inhibitors. The experiment would be set up as just described, but for the successive incubations IPTG (which acts as a competitive inhibitor) could be added to an injection vial with substrate. IPTG is neutral, and would migrate in tandem with the substrate at the same velocity as the EOF catching up with the slower moving enzyme molecule. The expectation is that different enzyme molecules would be heterogeneous for inhibition. Such a result would have implications for highly heterogeneous drug targets.

The results from the continuous assay, and the 15 minute static assays using RES-gal as a substrate (chapters 5, 6 and 7), provide little support for the existence of β-

galactosidase dynamic catalytic heterogeneity. To reinforce the results from the continuous assay that short-lived room temperature rate fluctuations do not occur, it should be possible to perform multiple very brief static assays with DDAO-gal using the extended separation method. The results from chapter 10 indicated that 3 minute assays yield excellent signal to noise ratios, and 90 s incubations would have still produce quantifiable peaks. Since background signal is proportional to the cross-sectional area of the capillary, performing this experiment on 5 μm diameter capillaries could allow assay times of 20 s. Conceivably, assays of less than 10 s may be possible if 2 μm diameter capillaries are used, and substrate is extracted twice with toluene prior to injection.

A 2006 paper by Xie and coworkers [231]took single enzyme detection to the level of the individual bacterium when they monitored the stochastic expression of β-galactosidase in semi-permeabolized *E.coli* by measuring the rate of change of fluorescein efflux from the cells by the hydrolysis of FDG by newly synthesized enzyme molecules. The foundation for their analysis was that each individual β-galactosidase had the same catalytic rate. The extensive documentation of single molecule heterogeneity makes this assumption questionable, and a close examination of their data is not fully consistent with the assumption, but the experiment does raise an important point. Except for the just described study, all individual molecule enzyme studies conducted are performed upon enzyme molecules obtained from commercial preparations, or from lysed bacterial cultures grown in the lab. Considering that in any particular single molecule experiment only a couple hundred molecules may be assayed out of several billion harvested, it is extremely improbable that any two molecules assayed are from the same bacterium. Because the DNA replication error rate is estimated to be as low as 10^{-9},

it is generally assumed that all assayed enzyme molecules were synthesized from identical DNA templates from different bacteria, and that the heterogeneity of catalytic activity arises from events subsequent to replication, although one study [48] indicated that β-galactosidase from cloned *E.coli* were different from the parent culture. To test the possibility that single enzyme molecule heterogeneity may be partially attributable to random DNA polymorphisms of bases, single molecule assays on uncoated capillaries could be performed on β-galactosidase from individual *E. coli* cells by adapting the method recently described by the Dovichi group [232] for the on-capillary lysis and measurement of green fluorescent protein in a single bacterium using laser-induced fluorescence. In that report single GFP detection was not achieved as the limit of detection was approximately 100 yoctomole, but an enzyme assay does not require single fluorophore sensitivity as a large number of product molecules from a single molecule are measured and not the enzyme itself. The bacterium is injected into the capillary and sandwiched between two plugs of a proprietary lysis buffer which liberates the cell contents within 5 minutes. The short lysis time and non-denaturing properties of the lysis buffer should leave functional β-galactosidase molecules within close proximity to each other at the injection end of the capillary. The β-galactosidase copy number for non-induced cells is typically less than five molecules per cell, which should minimize the possible complication of excess enzyme and overlapping product peaks. Mobilization of the cell contents in the presence of DDAO-gal followed by a 5 minute incubation and second mobilization to sweep product pools past the detector should allow the measurement of single molecule activity of individual enzyme from the same bacterium.

It is expected that individual molecules of β-galactosidase from a single bacterium will exhibit electrophoretic and catalytic heterogeneity. However, the range of activities should be lower when contrasted to the range observed for individual molecules obtained from a large number of cells. The range is expected to be lower because these cells will be assayed for basal levels of β-galactosidase only, and are likely transcribed off a single mRNA. However, if the molecules from an individual cell are electrophoretically identical, it will not be possible to separate them, and therefore know how many are present, or assay their individual activity.

The proposed experiments provide straightforward means to further investigate some of the issues raised during the course of this work of specific relevance for β-galactosidase. Other experiments using techniques not yet devised may provide deeper insights for the basis of single molecule catalytic heterogeneity that in turn may lay a foundation for experiments examining protein heterogeneity generally and its implications for cellular responses and disease development. A greater understanding of the significance of individual biological macromolecule heterogeneity has for cellular processes presents formidable technological challenges that will need to be met at some stage if a complete description of cellular process is to be attained.

13. References

1. Feynmann, R.P. (1961) There's plenty of room at the bottom. In: Gilbert, H.D. (ed.) *Miniaturization*. Reinhold Publishing Corporation, NY.

2. Xue, Q. and Yeung, E.S. (1995) Differences in the chemical reactivity of individual molecules of an enzyme. *Nature*, **373:** 681-683.

3. Craig, D.B., Arriaga, E.A., Wong, C.Y., Dovichi, N.J. (1996) Studies on single alkaline phosphatase molecules: reaction rate and activation energy of a reaction catalyzed by a single molecule and the effect of thermal denaturation—the death of an enzyme. *J. Am. Chem. Soc.,* **118:** 5245-5253.

4. Lu, H.P., Xun, L., Xie, X.S. (1998) Single-molecule enzymatic dynamics. *Science,* **282**(5395): 1877-1882.

5. English, B.P., Min, W., van Oijen, A.M., Lee, K.T., Luo, G., Sun, H., Cherayil, B.J., Kou, S.C., Xie, X.S. (2006) Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited. *Nature Chem. Biol.*, **2:** 87-94.

6. Yildiz, A., Forkey, J.N., McKinney, S.A., Ha, T., Goldman, Y.E., Selvin, P.R. (2003) Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, **300**: 2061-2065.

7. Kellermayer, M.S., Smith, S.B., Granzier, H.L., Bustamante, C. (1997) Folding-unfolding transitions in single titin molecules characterized with laser tweezers. *Science*, **276**(5329): 1112-1116.

8. Min, W., Luo, G., Cherayil, B.J., Kou, S.C., Xie, X.S. (2005) Observation of a power-law memory kernel for fluctuations within a single protein molecule. *Phys. Rev. Lett.*, **94**(19): 198302(1)-198302(4).

9. Cornish, P.V., Ha,T. (2007) A survey of single-molecule techniques in chemical biology. *ACS Chem. Biol.*, **2**: 53-61.

10. Veigel, C., Coluccio, L.M., Jontes, J.D., Sparrow, J.C., Milligan, R.A., Molloy, J.E. (1999) The motor protein myosin I produces its working stroke in two steps. *Nature*, **398**: 530-533.

11. Smith, S., Cui, Y., Bustamante, C. (1996) Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, **271**: 795-799.

12. Marszalek, P.E., Oberhauser, A.F., Pang, Y.P., Fernandez, J.M. (1998) Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. *Nature*, **396**: 661-664.

13. Itoh, H., Takahashi, A., Adachi, K., Noji, H., Yasuda, R., Yoshida, M., Kinoshita K. (2004) Mechanically driven ATP synthesis by F1-ATPase. *Nature*, **427**(6973): 465-468.

14. Wehry, E.L. (1997) Molecular fluorescence and phosphorescence spectrometry. In: Settle, F.A. (ed.) *Handbook of instrumental techniques for analytical chemistry*. Prentice Hall, Inc. New Jersey, pp. 507-539.

15. Barnes, M.D., Whitten, W.B., Ramsay, J.M. (1995) Detecting Single Molecules in Liquids. *Anal. Chem.*, **67**(13): 418A-423A.

16. Nie, S., Chiu, D.T., Zare, R.N. (1994) Probing individual molecules with confocal fluorescence microscopy. *Science*, **266**:1018-1021.

17. Rotman, B. (1961) Measurement of activity of single molecules of β-D-galactosidase. *Proc. Natl. Acad. Sci*, **47**: 1981-1991.

18. Rissin, D.M., Gorris, H.H., Walt, D.R. (2008) Distinct and long-lived activity states of single enzyme molecules. *J. Am. Chem. Soc.*, **130:** 5349-5353.

19 Blom, H., Kastrup, L., Eggeling, C. (2006) Fluorescence fluctuation spectroscopy in reduced volumes. *Curr. Pharm. Biotech.*, **7**: 51-66.

20. Gratzl, M., Lu, H., Matsumoto, T., Yi, C., Bright, G.R. (1999) Fine chemical manipulations of microscopic liquid samples. 1. Droplet loading with chemicals. *Anal. Chem.*, **71**(14): 2751-2756.

21 Nakano, M., Komatsu, J., Takashima, K., Katsura, S., Mizuno, A. (2003) Single-molecule PCR using water-in-oil emulsion. *J. Biotechnol.*, **102**(2): 117-124.

22. Hsin, T. and Yeung, E.S. (2007) Single-molecule reactions in liposomes. *Angew. Chem. Int. Ed.*, **46**: 8032-8035.

23. Rondelez, Y., Tresset, G., Tabata, K.V., Arata, H., Fujita, H., Takeuchi, S., Noji, H. (2005) Microfabricated arrays of femtoliter chambers allow single molecule enzymology. *Nat. Biotechnol.*, **23**(3): 361-365.

24. Rissin, D.M. and Walt, D.R. (2006) Digital concentration readout of single enzyme molecules using femtoliter arrays and poisson statistics. *Nano Letters*, **6**(3): 520-523.

25. Cheng, Y.F., and Dovichi, N.J. (1988) Subattomole amino acid analysis by capillary zone electrophoresis and laser-induced fluorescence. *Science*, **242**: 562-564.

26. Noji, H., Yasuda, R., Yoshida, M., Kinoshita, K. (1999) Direct observation of the rotation of the F_1-ATPase. *Nature*, **386**: 299-302.

27. Walter, N.G., Huang, C., Manzo, A.J., Sobhy, M.A. (2008) Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nature Methods*, **5**(6): 475-489.

28. Hirschfeld, T. (1976) Optical microscopic observation of single small molecules. *Applied Optics*, **15**(12): 2965-2966.

29. Dovichi, N.J., Martin, J.C., Jett, J.H., Trkula, M., Keller, R.A. (1984) Laser-induced fluorescence of flowing samples as an approach to single-molecule detection in liquids. *Anal. Chem.*, **56**(3): 348-354.

30. Betzig, E. and Chichester, R.J. (1993) Single molecules observed by near-field scanning optical microscopy. *Science*, **262**(5138): 1422-1425.

31. Xie, X.S. and Dunn, R.C. (1994) Probing single molecule dynamics. *Science*, **265**(5170): 361-364.

32. Ambrose, W.P., Goodwin, P.M., Martin, J.C., Keller, R.A. (1994) Alterations of single molecule fluorescence lifetimes in near-field optical microscopy. *Science*, **265**(5170): 361-367.

33 Vickery, S.A. and Dunn, R.C. (1999) Scanning near-field fluorescence resonance energy transfer microscopy. *Biophys. J.*, **76**(4): 1812-1818.

34. Ha, T. (2001) Single-molecule fluorescence resonance energy transfer. *Methods*, **25**: 78-86.

35. Macklin, J.J., Trautman, J.K., Harris, T.D., Brus, L.E. (1996) Imaging and time-resolved spectroscopy of single molecules at an interface. *Science*, **272**: 255-257.

36. Funatsu, T., Harada, Y., Tokunaga, M., Salto, K., Yanagida, T. (1995) Imaging of single fluorescent molecule and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, **374**: 555-559.

37. Chen, Y., Hu, D., Vorpagel, E.R., Lu, H.P. (2003) Probing single-molecule T4 lysozyme conformational dynamics by intramolecular fluorescence energy transfer. *J. Phys. Chem. B*, **107**(31): 7947-7956.

38. Schuler, B. and Eaton, W.A. (2008) Protein folding studied by single molecule FRET. *Curr. Opin. Struc. Biol.*, **18**: 16-26.

39. Yildiz, A., Park, H., Safer, D., Yang, Z., Chen, L., Selvin, P.R, Sweeney, H.L. (2004) Myosin VI steps via a hand-over-hand mechanism with its lever arm undergoing fluctuations when attached to actin. *J. Biol. Chem.* **279**(36): 37223-37226.

40. Kura, C., Kim, H., Syed, S., Goshima, G., Gelfand, V.I., Selvin, P.R. (2005) Kinesin and dynein move a peroxisome *in vivo*: a tug-of-war or coordinated movement? *Science*, **308**: 1469-1472.

41. Yasuda, R., Yoshida, M., Kinoshita Jr., K., Itoh, H. (2001) Resolution of distinct rotational substeps by submillisecond kinetic analysis of F1-ATPase. *Nature*, **410**(6831): 898-904.

42. Zhuang, X., Kim, H., Pereira, M.J., Babock, H.P., Walter, N.G., Chu, S. (2002) Correlating structural dynamics and function in single ribozyme molecules. *Science*, **296**: 1473-1476.

43. Hanson, J.A., Duderstadt, K., Watkins, L.P., Bhattacharyya, S., Brokaw, J., Chu, J., Yang, H. (2007) Illuminating the mechanistic roles of enzyme conformational dynamics. *Proc. Natl. Acad. Sci. USA*, **104**(46): 18055-18060.

44. Tan, W. and Yeung, E.S. (1997) Monitoring the reactions of single enzyme molecules and single metal ions. *Anal. Chem.*, **69**(20): 4242-4248.

45. Craig, D.B. and Dovichi, N.J. (1998) *Escherichia coli* β-galactosidase is heterogeneous with respect to the activity of individual molecules. *Can. J. Chem*., **76:** 623-626.

46. Shoemaker, G.K., Juers, D.H., Coombs, J.M.L., Matthews, B.W., Craig, D.B. (2003) Crystallization of β-galactosidase does not reduce the range of activity of individual molecules. *Biochemistry*, **42**(6): 1707-1710.

47. Craig, D.B., Hall, T., Goltz, D.M. (2000) *Escherichia coli* β-galactosidase is heterogeneous with respect to a requirement for magnesium. *Biometals*, **13**: 223-229.

48. Craig, D.B., Nachtigall, J.T., Ash, H.L., Shoemaker, G.K., Dyck, A.C., Wawrykow, T.M.J., Gudbjartson, H.L. (2003) Differences in the average single molecule activities of *E. coli* β-galactosidase: Effect of source, enzyme molecule age and temperature of induction. *J. Prot. Chem.*, **22:** 555-561.

49. Li, Q. and Seeger, S. (2006) Label-free detection of single protein molecules using deep UV fluorescence lifetime microscopy. *Anal. Chem.*, **78**(8): 2732-2737.

50. Polakowski, R., Craig, D.B., Skelley, A., Dovichi, N.J. (2000) Single molecules of highly purified bacterial alkaline phosphatase have identical activity. *J. Am. Chem. Soc.,* **122:** 4853-4855.

51. Dyck, A.C. and Craig, D.B. (2002) Individual molecules of thermostable alkaline phosphatase support different catalytic rates at room temperature. *Luminescence*, **17**: 15-18.

52. Shi, J., Palfey, B.A., Dertouzos, J., Jensen, K.F., Gafni, A., Steel, D. (2004) Multiple states of the tyr328leu mutant of dihydroorotate dehydrogenase revealed by single molecule kinetics. *J. Am. Chem. Soc.*, **126**(22): 6914-6922.

53. Martinez, V., De Cremer, G., Roeffaers, M.B.J., Sliwa, M., Baruah, M., De Vos, D.E., Hofkens, J., Sels, B.F. (2008) Exploration of single molecule events in a haloperoxidase and its biomimic: localization of halogenation activity. *J. Am. Chem. Soc.*, **130**(40): 13192-13193.

54. van Oijen, A.M., Blainey, P.C., Crampton, D.J., Richardson, C.C., Ellenberger, T., Xie, X.S. (2003) Exonuclease reveals base dependence and dynamic disorder. *Science*, **301**(5637): 1235-1238.

55. Edman, L., Foldes-Papp, Z., Wennmalm, S., Rigler, R. (1999) The fluctuating enzyme: a single molecule approach. *Chem. Phys.*, **247**: 11-22.

56. Hassler, K., Rigler, P., Blom, H., Rigler, R., Widengren, J., Lasser, T. (2007) Dynamic disorder in horseradish peroxidase observed with total internal reflection fluorescence correlation spectroscopy. *Optics Express*, **15**(9): 5366-5375.

57. Yang, H., Luo, G., Karnchanaphanurach, P., Louie, T., Rech, I., Cova, S., Xun, L., Xie, X.S. (2003) Protein conformation dynamics probed by single-molecule electron transfer. *Science*, **302**: 262-266.

58. Velonia, K., Flomenbom, D.L., Masuo, S., Cotlet M., Engelborghs, Y., Hofkens, J., Rowan, A.E., Klafter, J., Nolte, R.J.M., deSchryver, F.C. (2005) Single-enzyme kinetics of CALB-catalyzed hydrolysis. *Angew. Chem. Int. Ed.,* **44**: 560-564.

59. Flomenbom, O., Velonia, K., Loos, D., Masuo, S., Cotlet, M., Engelborghs, Y., Hofkens, J., Rowen, A.E., Nolte, R.J.M., Van der Auweraer, M., de Schryver, F.C., Klafter, J. (2005) Stretched exponential decay and correlations in the

catalytic activity of fluctuating single lipase molecules. *Proc. Natl. Acad. Sci. USA*, **102**(7): 2368-2377.

60. Dan, N. (2007) Understanding dynamic disorder fluctuations in single-molecule enzymatic reactions. *Curr. Opin. Coll. Inter. Sci.*, **12**: 314-321.

61. Min, W., English, B.P., Luo, G., Cherayil, B.J., Kou, S.C., Xie, X.S. (2005) Fluctuating enzymes: lessons from single-molecule studies. *Acc. Chem. Res.*, **38**(12): 923-931.

62. Mesecar, A.D., Stoddard, B.L., Koshland, D.E. (1997) Orbital steering in the catalytic power of enzymes: small structural changes with large catalytic consequences. *Science*, **277**(5323): 202-206.

63. Baldini, G., Cannone, F., Chirico, G., Collini, M., Campanini, B., Bettati, S., Mozzarelli, A. (2007) Evidence of discrete substates and unfolding pathways in green fluorescent protein. *Biophys. J.*, **92**: 1724-1731.

64. Gordon, M.J., Huang, X., Pentoney, S.L., Zare, R.N. (1988) Capillary electrophoresis. *Science*, **242**(4876): 224-228.

65. Mikkers, F.E.P., Everaarts, F.M., Verheggen, T. (1979) High-performance zone electrophoresis. *J. Chromatogr.*, **169**: 11-20.

66 Jorgenson, J.W. and Lukacs, K.D. (1981) Zone electrophoresis in open-tubular glass capillaries. *Anal. Chem.*, **53**(8): 1298-1302.

67. Jorgenson, J.W. and Lukacs, K.D. (1983) Capillary zone electrophoresis. *Science*, **222**(4621): 266-272.

68. Lauer, H.H. and McManigill, D. (1986) Capillary zone electrophoresis of proteins in untreated fused silica tubing. *Anal. Chem.*, **58**(1): 166-170.

69. Gassmann, E., Kuo, J.E., Zare, R.N. (1985) Electrokinetic separation of chiral compounds. *Science*, **230**(4727): 813-814.

70. Ewing, A.G., Mesaros, J.M., Gavin, P.F. (1994) Electrochemical detection in microcolumn separations. *Anal. Chem.*, **66**(9): 527A-537A.

71. Sun, X., Jin, W., Li, D., Bai, Z. (2004) Measurement of alkaline phosphatase isoenzymes in individual mouse bone marrow fibroblast cells based on capillary electrophoresis with on-capillary enzyme-catalyzed reaction and electrochemical detection. *Electrophoresis*, **25**: 1860-1866.

72. Huck, C.W., Bakry, R., Huber, L.A., Bonn, G.K. (2006) Progress in capillary electrophoresis coupled to matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *Electrophoresis*, **27**: 2063-2074.

73. Dovichi, N.J., Martin, J.C., Jett, J.H., Keller, R.A. (1983) Attogram detection limit for aqueous dye samples by laser-induced fluorescence. *Science*, **219**(4586): 845-847.

74. Chen, D.Y. and Dovichi, N.J. (1994) Yoctomole detection limit by laser-induced fluorescence in capillary electrophoresis. *J. Chromatogr. B*, **657**: 265-269.

75. Chen, D.Y., Adelhelm, K., Cheng, X.L., Dovichi, N.J. (1994) A simple laser-induced fluorescence detector for sulforhodium 101 in a capillary electrophoresis system: Detection limits of 10 yoctomoles or six molecules. *Analyst*, **119:**349-352.

76. Dovichi, N.J. (1997) DNA sequencing by capillary electrophoresis. *Electrophoresis*, **18**: 2393-9.

77. Whiting, C.E. and Arriaga, E.A. (2006) CE-LIF analysis of mitochondria using uncoated and dynamically coated capillaries. *Electrophoresis*, **27**: 4523-4531.

78. Shoemaker, G.K., Lorieau, J., Lau, L.H., Gillmore, C.S., Palcic, M.M. (2005) Multiple sampling in single-cell enzyme assays using CE laser-induced fluorescence to monitor reaction progress. *Anal. Chem.*, **77**(10): 3132-3137.

79. Hu, K., Ahmadzadeh, H., Krylov, S.N. (2004) Asymmetry between sister cells in a cancer cell line revealed by chemical cytometry. *Anal. Chem.*, **76**(13): 3864-3866.

80. Adamson, N.J. and Reynolds, E.C. (1997) Rules relating electrophoretic mobility, charge and molecular size of peptides and proteins. *J. Chromatogr. B*, **699:** 133-147.

81. Bolt, G.H. (1957) Determination of the charge density of silica sols. *J. Phys. Chem.*, **61**(9): 1166-1169.

82. Stevens, T.S. and Cortes, H.J. (1983) Electroosmotic propulsion of eluent through silica-based chromatographic media. *Anal. Chem.*, **55**(8): 1365-1370.

83. Gas, B., Stedry, M., Kenndler, E. (1997) Peak broadening in capillary zone electrophoresis. *Electrophoresis*, **18**: 2123-2133.

84. Cunliffe, J. M., Baryla, N. E., Lucy, C. A. (2002) Phospholipid bilayer coatings for the separation of proteins in capillary electrophoresis. *Anal. Chem.*, **74**(4): 776-783.

85. MacDonald, A.M. and Lucy, C.A. (2006) Highly efficient protein separations in capillary electrophoresis using a supported bilayer/diblock copolymer coating. *J. Chromatog. A*, **1130**: 265-271.

86. Kaneta, T., Ueda, T., Hata, K., Imasaka, T. (2006) Suppression of electroosmotic flow and its application to determination of electrophoretic mobilities in a poly(vinylpyrrolidone)-coated capillary. *J. Chromatog. A.*, 1106: 52-55.

87. Chang, W.W.P., Hobson, C., Bomberger, D.C., Schneider, L.V. (2005) Rapid separation of protein isoforms by capillary zone electrophoresis with new dynamic coatings. *Electrophoresis*, **26**: 2179-2186.

88. Huang, X., Gordon, M.J., Zare, R.N. (1988) Bias in quantitative capillary zone electrophoresis caused by electrokinetic sample injection. *Anal. Chem.*, **60**(4): 375-377.

89. Zhao, J.Y., Dovichi, N.J., Hindsgaul, O., Gosselin, S., Palcic, M.M. (1994) Detection of 100 molecules of product formed in a fucosyltransferase reaction. *Glycobiology*, **4**(2): 239-242.

90. Alpers, D.H., Steers, Jr., E.S., Shifrin, S., Tomkins, G. (1965) Isozymes of the lactose operon of *Escherichia coli*. *J. Mol. Biol.*, **11**: 12-22.

91. Matthews, B.W. (2005) The structure of *E. coli* β-galactosidase. *C.R. Biologies*, **328:** 549-556.

92. Jobe, A. and Bourgeois, S. (1972) The natural inducer of the *lac* operon. *J. Mol. Biol.*, **69**: 397-408.

93. Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.*, **3**: 318-356.

94. Lewis, M. (2005) The *lac* repressor. *C.R. Biologies*, **328**: 521-548.

95. Gilbert, W. and Muller-Hill, B. (1966) Isolation of the *lac* repressor. *Proc. Natl. Acad. Sci. USA*, **56**: 1891-1898.

96. Ullman, A., Jacob, F., Monod, J. (1967) Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of *Escherichia coli*. *J. Mol. Biol.*, **24**: 339-343.

97. Fowler, A.V. and Zabin, I. (1983) Purification, structure and properties of hybrid β-galactosidase proteins. *J. Biol. Chem.* **258**(23): 14354-14358.

98. Welply, J.K., Fowler, A.V., Zabin, I. (1981) β-galactosidase α-complementation. *J. Biol. Chem.*, **256**(13): 6804-6810.

99. Sassenfeld, H.M. (1990) Engineering proteins for purification. *Trends Biotechnol.,* **8**: 88-93.

100. Alam, J. and Cook, J.L. (1990) Reporter genes: Application to the study of mammalian gene transcription. *Anal. Biochem.*, **188**: 245-254.

101. Schenborn, E., Groskreutz, D. (1999) Reporter gene vectors and assays. *Mol. Biotech.*, **13**: 29-44.

102. Fowler, A.V., Brake, A.J., Zabin, I. (1978) Amino acid sequence of β-galactosidase. VI. Limited tryptic digestion of the citraconylated protein and sequences of tryptic peptides. *J. Biol. Chem.*, **253**(15): 5490-5498.

103. Kalnins, A., Otto, K., Ruther, U., Muller-Hill, B. (1983) Sequence of the *lacZ* gene of *Escherichia coli*. *EMBO*, **2**(4): 593-597.

104. Jacobson, R.H., Zhang, X.J., DuBose, R.F., Matthew, B.W. (1994) Three dimensional structure of β-galactosidase from *E. coli*. *Nature*, ***369*:** 761-766.

105. Juers, D.H., Jacobson, R.H., Wigley, D., Zhang, X.J., Huber, R.E., Tronrud, D.E., Matthews, B.W. (2000) High resolution refinement of β-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for α-complementation. *Protein Sci.*, ***9*:** 1685-1699.

106. Marchesi, S.L., Steers Jr., E., Shifrin, S. (1969) Purification and characterization of the multiple forms of β-galactosidase of *Escherichia coli*. *Biochim. Biophy. Acta,* **181**: 20-34.

107. Edwards, R.A., Jacobsen, A.L., Huber, R.E. (1990) Thermal denaturation of β-galactosidase and of two site-specific mutants. *Biochemistry*, **29**(49): 11001-11008.

108. Craven, G.R., Steers Jr., E., Anfinsen, C.B. (1965) Purification, composition, and molecular weight of the β-galactosidase of *Escherichia coli* K12. *J. Biol. Chem.*, **240**(6): 2468-2477.

109. Tenu, J., Viratelle, O.M., Yon, J. (1972) Kinetic study of the activation process of β-galactosidase from *Escherichia coli* by Mg^{2+}. *Eur. J. Biochem.*, **26**: 112-118.

110. Huber, R.E., Parfett, C., Woulfe-Flanagan, H., Thompson, D.J. (1979) Interaction of divalent cations with β-galactosidase (*Escherichia coli*). *Biochemistry*, **18**(19): 4090-4096.

111. Ullman, A. and Monod, J. (1969) On the effect of divalent cations and protein concentration upon renaturation of β-galactosidase from *E. coli*. *Biochem. Biophys. Res. Comm.*, **35**(1): 35-42.

112. Bader, D.E., Ring, M., Huber, R.E. (1988) Site-directed mutagenic replacement of Glu-461 with Gln in β-galactosidase (*E. coli*): evidence that Glu-461 is important for activity. *Biochem. Biophys. Res. Commun.*, **153**(1): 301-306.

113. Ring, M., Bader, D.E., Huber, R.E. (1988) Site-directed mutagenesis of β-galactosidase (*E. coli*) reveals that Tyr-503 is essential for activity. *Biochem. Biophys. Res. Commun.*, **152**(3) 1050-1055.

114. Yuan, J., Martinez-Bilbao, M., Huber, R.E.(1994) Substitutions for Glu-537 of β-galactosidase from *Escherichia coli* cause large decreases in catalytic activity. *Biochem J.*, **299**(2): 527-531.

115. Roth, N.J. and Huber, R.E. (1996) The β-galactosidase (*Escherichia coli*) reaction is partly facilitated by interactions of His-540 with C6 hydroxyl of galactose. *J. Biol. Chem.*, **271**(24): 14296-14301.

116. Huber, R.E., Hakda, S., Cheng, C., Cupples, C.G., Edwards, R.A. (2003) Trp-999 of β-galactosidase (*Escherichia coli*) is a key residue for binding, catalysis and synthesis of allolactose, the natural *lac* operon inducer. *Biochemistry*, **42**(6): 1796-1803.

117. Juers, D.H., Heightman, T.D., Vasella, A., McCarter, J.D., Mackenzie, L., Withers, S.G. and Matthews, B.D. (2001). A structural view of the action of *Escherichia coli* (*lacZ*) β-galactosidase. *Biochemistry* **40**(49): 14781-14794.

118. Huang, X. and Madan, A. (1999) CAP3: A DNA sequence assembly program. *Genome Res.,* **9**: 868-877.

119. Huang, X. and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math*, **12**(3): 337-357.

120. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**(5259): 680-685.

121. http://www.scripps.edu/~cdputnam/protcalc.html

122. Steers Jr., E. and Cuatrecasas, P. (1974) β-galactosidase. *Meth. Enzymol.*, **34**: 350-358.

123. Krokhin, O.V., Ens, W., Standing, K.G. (2005) MALDI QqTOF MS combined with off-line HPLC for characterization of protein primary structure and post-translational modifications. *J. Biomol. Tech.,* **16**: 429-440.

124. Gottesman, S. (2003) Proteolysis in bacterial regulatory circuits. *Annu. Rev. Cell. Dev. Biol.*, **19**: 565-587.

125. Miller, C.G. (1996) Protein degradation and proteolytic modification. In: Neidhardt, F.C. (ed), *Escherichia coli and Salmonella: Cellular and Molecular Biology*. American Society for Microbiology, Washington, DC, pp. 938-954.

126. Ben-Bassat, A., Bauer, K., Chang, S.Y., Myambo, K., Boosman, A., Chang, S. (1987) Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. *J. Bacteriol.*, **169**(2): 751-757.

127. Hirel, P.H., Schmitter, J.M., Dessen P., Fayat, G., Blanquet, S. (1989) Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid. *Proc. Natl. Acad. Sci. USA*, **86**: 8247-8251.

128. Zwizinski, C. and Wickner, W. (1980) Purification and characterization of leader (signal) from *Escherichia coli*. *J. Biol. Chem.*, **255**(16): 7973-7977.

129. Prouty, W.F. and Goldberg, A.L. (1972) Effects of protease inhibitors on protein breakdown in *Escherichia coli*. *J. Biol. Chem.* **247**(10): 3341-3352.

130. Kues, T., Peters, R., Kubitscheck, U. (2001) Visualization and tracking of single proteins in the cell nucleus. *Biophys. J.*, **80**(6): 2954-2967.

131. McKnight, J.L. and Fried, V.A. (1983) A novel proteolytic activity apparently initiating degradation of β-galactosidase nonsense fragments in *in vitro* extracts of *Escherichia coli*. *J. Biol. Chem.*, **258**(12): 7550-7555.

132. Wang, S.S., and Fried, V.A. (1987) Early steps initiating a degradation pathway in *Escherichia coli*. *J. Biol. Chem.*, **262**(13): 6357-6364.

133. Zubay, G. (1973) *In vitro* synthesis of protein in microbial systems. *Annu. Rev. Genet.*, **7**: 267-287.

134. Roberts, B.E. and Paterson, B.M. (1973) Efficient translation of tobacco mosaic virus RNA and rabbit globin 9S RNA in a cell-free system from commercial wheat germ. *Proc. Natl. Acad. Sci. USA*, **70**(8): 2330-2334.

135. Vinarov, D.A., Newman, C.L.L., Markley, J.L. (2006) Wheat germ cell-free platform for eukaryotic protein production. *FEBS J.*, **273**: 4160-4169.

136. Roads, R.E., McKnight, G.S., Schimke, R.T. (1971) Synthesis of ovalbumin in a rabbit reticulocyte cell-free system programmed with hen oviduct ribonucleic acid. *J. Biol. Chem.*, **246**(23): 7407-7410.

137. Betton, J.M. (2003) Rapid translation system (RTS): a promising alternative for recombinant protein production. *Curr. Prot. Pep. Sci.*, **4**: 73-80.

138. Matthaei, J.H. and Nirenberg, M.W. (1961) Characteristics and stabilization of DNAase-sensitive protein synthesis in *E. coli* extracts. *Proc. Natl. Acad. Sci. USA*, **47**: 1580-1588.

139. Nirenberg, M.W. and Matthaei, J.H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. USA*, **47**: 1588-1602.

140. Nathans, D., Notani, G., Schwartz, J.H., Zinder, N.D. (1962) Biosynthesis of the coat protein of coliphage *f*2 by *E.coli* extracts. *Proc. Natl. Acad. Sci.*, **48**: 1424-1431.

141. DeVries, J.K., Zubay, G. (1967) DNA-directed peptide synthesis, II. The synthesis of the α-fragment of the enzyme β-galactosidase. *Biochemistry*, **57**: 1010-1012.

142. Zubay, G., Schwartz, D., Beckwith, J. (1970) Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc. Natl. Acad. Sci. USA*, **66**(1): 104-110.

143. Chumpolkulwong, N., Hori-Takemoto, C., Hosaka, T., Inaoka,T., Kigawa, T., Shirouzu, M., Ochi, K., Yokoyama, S. (2004) Effects of *Escherichia coli* ribosomal protein S12 mutations on cell-free protein synthesis. *Eur. J. Biochem.*, **271**: 1127-1134.

144. Gromadski, K.B. and Rodnina, M.V. (2004) Kinetic determinants of high fidelity tRNA discrimination on the ribosome. *Mol. Cell*, **13**: 191-200.

145. Gromadski, K.B. and Rodnina, M.V. (2004) Streptomycin interferes with conformational coupling between codon recognition and GTPase activation on the ribosome. *Nat. Struc. Mol. Biol.*, **11**(4): 316-322.

146. Jasiecki, J., Wegrzyn, G. (2006) Phosphorylation of Escherichia coli poly(A) polymerase I and effects of this modification on the enzyme activity. *FEMS Microbiol. Lett.,* **261**: 118-122.

147. Kigawa, T., Muto, Y., Yokoyama, S. (1995) Cell-free synthesis and amino acid-selective stable isotope labelling of proteins for NMR analysis. *J. Biomolec. NMR*, **6**: 129-134.

148. Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T., Yokoyama, S. (1999) Cell-free production and stable-isotope labelling of milligram quantities of proteins. *FEBS Lett.*, **442**: 15-19.

149. Chumpolkulwong, N., Sakamoto, K., Hayashi, A., Iraha, F., Shinya, N., Matsuda, N., Kiga, D., Urushibata, A., Shirouzu, M., Oki, K., Kigawa, T., Yokoyama, S. (2006) Translation of 'rare' codons in a cell-free protein synthesis system from *Escherichia coli*. *J. Struct. Func. Genom.*, **7**: 31-36.

150. Yokoyama, S. (2003) Protein expression systems for structural genomics and proteomics. *Curr. Opin. Chem. Biol.*, **7**: 39-43.

151. Kim, D., Kigawa, T., Choi, C., Yokoyama, S. (1996) A highly efficient cell-free synthesis system from *Escherichia coli*. *Eur. J. Biochem.*, **239**: 881-886.

152. Nevin, D.E. and Pratt, J.M. (1991) A coupled *in vitro* transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. *FEBS*, **291**(2): 259-263.

153. Schwarz, D., Junge, F., Durst, F., Frölich, N., Schneider, B., Reckel, S., Sobhanifar, S., Dötsch, V., Bernhard, F. (2007) Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems. *Nature Protocol*, **2**(11): 2945-2957.

154. Shirokov, V.A., Kommer, A., Kolb, V.A., Spirin, A.S. (2007) Continuous-exchange protein synthesizing systems. In: Grandi, G. (ed.) *In vitro transcription and translation protocols, 2nd edition*. Humana Press Inc., Totowa, NJ, pp 19-55.

155. Jewett, M.C. and Swartz, J. (2004) Mimicking the *Escherichia coli* cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol. Bioeng.,* **86**(1): 19-26.

156. Spirin, A.S., Baranov, V.I., Ryabova, L.A., Ovodov, S.Y., Alakhov, Y.B. (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. *Science*, **242**: 1162-1164.

157. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., Ueda, T. (2001) Cell-free translation reconstituted with purified components. *Nat. Biotech.*, **19**: 751-755.

158. Seong, G.H., Heo, J., Crooks, R.M. (2003) Measurement of enzyme kinetics using a continuous-flow microfluidic system. *Anal. Chem.*, **75**(13): 3161-3167.

159. Tobias, J.W., Shrader, T.E., Rocap, G., Varshavsky, A. (1991) The N-end rule in bacteria. *Science*, **254**(5036): 1374-1377.

160. Torizawa, T., Shimizu, M., Taoka, M., Miyano, H., Masatsune, K. (2004) Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. *J. Biolmolec. NMR*, **30**: 311-325.

161. Shibuya, N., Nishiyama, T., Nakashima, N. (2004) Cell-free synthesis of polypeptides lacking an amino-terminal methionine by using a dicistroviral intergenic internal ribosome entry site. *J. Biochem.* **136**(5): 601-606.

162. Liao, Y., Jeng, J., Wang, C., Wang, S., Chang, S. (2006) Removal of N-terminal methionine from recombinant proteins by engineered *E. coli* methionine aminopeptidase. *Protein Science*, **13**:1802-1810.

163. Endo, S., Yamamoto, Y., Sugawara, T., Nishimura, O., Fujino, M. (2001) The additional methionine residue at the N-terminus of bacterially expressed human interleukin-2 affects the interaction between the N- and C-termini. *Biochemistry*, **40**: 914-919.

164. Mine, S., Ueda, T., Hashimoto, Y., Imoto, T. (1997) Improvement of the refolding yield and solubility of hen egg-white lysozyme by altering the Met residue attached to its N-terminus to Ser. *Protein Eng.*, **10**: 1333-1338.

165. Kramer, E.B., Farabaugh, P.J. (2008) The frequency of translation misreading errors in *E.coli* is largely determined by tRNA competition. *RNA*, **13**: 87-96.

166. Mansky, L.M. and Temin, H.M. (1995) Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.*, **69**(8): 5087-5094.

167. McKnight, J.L. and Fried V.A. (1981) Limited proteolysis: early steps in the processing of large premature termination fragments of β-galactosidase in *Escherichia coli*. *J. Biol. Chem.*, **256**(18): 9652-9661.

168. Manley, J.L. (1978) Synthesis and degradation of termination and premature-termination fragments of β-galactosidase *in vitro* and *in vivo*. *J. Mol. Biol.*, **125**: 407-432.

169. Perron-Savard, P., De Crescenzo, G., Le Moural, H. (2005) Dimerization and DNA binding of the *Salmonella enterica* PhoP response regulator are phosphorylation independent. *Microbiology*, **151**: 3979-3987.

170. Du, P., Lolakis, P., Luo, C. (2005) Phoshorylation of serine residues in histidine-tag sequences attached to recombinant protein kinases: A cause of heterogeneity in mass and complications in function. *Protein Expr. Purif.*, **44**: 121-129.

171. Goel, A., Colcher, D., Koo, J., Booth, B.J.M., Pavlinkova, G., Batra, S.K. (2000) Relative position of the hexahistidine tag effects binding properties of a tumor-associated single-chain Fv construct. *Biochim. Biophys. Acta,* **1523**: 13-20.

172. Corey, P.F., Trimmer, R.W., Biddlecom, W.G. (1991) A new chromogenic β-galactosidase substrate: 7-β-D-Galactopyranosyloxy-9,9-dimethy-9*H*-acridin-2-one. *Angew. Chem. Int. Ed. Engl.*, **30**: 1646-1648.

173. Craig, D., Arriaga, E.A., Banks, P., Zhang, Y., Renborg, A., Palcic, M.M., Dovichi, N.J. (1995) Fluorescence-based enzymatic assay by capillary electrophoresis laser-induced fluorescence detection for the determination of a few β-galactosidase molecules. *Anal. Biochem.*, **226**: 147-153.

174. Huang, Z. (1991) Kinetic fluorescence measurement of fluorescein di-β-D-galactoside hydrolysis by β-galactosidase: intermediate channelling in stepwise catalysis by a free single enzyme. *Biochemistry*, **30**(35): 8535-8540.

175. Fiedler, F. and Hinz, H. (1994) No intermediate channelling in stepwise hydrolysis of fluorescein di-β-D-galactoside by β-galactosidase. *Eur. J. Biochem.*, **222**: 75-81.

176. Wingrove, A.S. and Caret, R.L. (1981) In: *Organic Chemistry*, Harper & Row Publishers, NY, p. 669.

177. Nichols, E.R., Gavina, J.M.A, Mcleod, R.G., Craig, D.B. (2007) Single molecule assays of β-galactosidase from two wild-type strains of *E. coli*: effects of protease inhibitors on microheterogeneity and different relative activities with differing substrates. *Protein J.*, **26:** 95-105.

178. Sinnot, M.L., and Souchard, I.J.L. (1973) The mechanism of action of β-galactosidase. Effect of aglycone nature and α-deuterium substitution on the hydrolysis of aryl galactosides. *Biochem. J.* **133**(1): 89-98.

179. Frauenfelder, H., Sligar, S.G., Wolynes, P.G. (1991) The energy landscapes and motions of proteins. *Science*, **254**(5038): 1598-1603.

180. Shortreed, M. R., Li, H., Huang, W., Yeung, E. S. (2000) High-throughput single-molecule DNA screening based on electrophoresis. *Anal. Chem.*, **72**(13): 2879-2885.

181. Chen, D. and Dovichi, N.J. (1996) Single-molecule detection in capillary electrophoresis: molecular shot noise as a fundamental limit to chemical analysis. *Anal. Chem.*, **68**(4): 690-696.

182. Castro, A. and Shera, E. B., (1995) Single-molecule electrophoresis. *Anal. Chem.*, **67**(18): 3181-3186.

183. Guanglai L., Wen, Q., Tang, J. X. (2005) Single filament electrophoresis of F-actin and filamentous virus fd. *J. Chem. Phys.*, **122**: 1044708-1044716.

184. Huang, X., Coleman, W.F., Zare, R.N. (1989) Analysis of factors causing peak broadening in capillary zone electrophoresis. *J. Chromatog. A*, **480**: 95-110.

185. Kang, S. H. and Yeung E. S. (2002) Dynamics of single-protein molecules at a liquid/solid interface: implications in capillary electrophoresis and chromatography. *Anal Chem.*, 74(4): 6334-6339.

186. Graf, M., Garcia, R.G., Watzig, H. (2005) Protein adsorption in fused-silica and polyacrylamide-coated capillaries. *Electrophoresis*, **26**(12): 2409-2417.

187. Gitlin, I., Carbeck, J. D., Whitesides, G. M. (2006) Electrostatic interactions in proteins. *Angew. Chem. Int. Ed.*, **45**: 3022-3060.

188. Compton, B.J. and O'Grady, E.A., (1991) Role of charge suppression and ionic strength in free zone electrophoresis of proteins. *Anal. Chem.*, **63:** 2597-2602.

189. Gao, J. and Whitesides, G.M. (1997) Using protein charge ladders to estimate the effective charges and molecular weights of proteins in solution. *Anal. Chem.*, **69**: 575-580.

190. Rosenberger, R.F. and Foskett, G. (1981) An estimate of the frequency of *in vivo* transcription errors at a nonsense codon in *Escherichia coli*. *Mol. Gen. Genet.*, **183**: 561-563.

191. Parker, J. (1989) Errors and alternatives in reading the universal genetic code. *Microbiol. Rev*., **53**: 273-298.

192. Pinto, D. M., Arriaga, E. A., Craig, D., Angelova, J., Sharma, N., Ahmadzedah, H., Dovichi, N. J., Boulet, C. A. (1997) Picomolar assay of native proteins by capillary electrophoresis precolumn labelling, submicellar separation, and laser-induced fluorescence detection. *Anal. Chem.*, **69**(15): 3015-3021.

193. Craig, D. B. and Dovichi, N. J. (1998) Multiple labelling of proteins. *Anal. Chem.*, **70**(13): 2493-2494.

194. Wetzl, B. K., Yarmoluk, S. M., Craig, D. B., Wolfbeis, O. S. (2004) Chameleon labels for staining and quantifying proteins. *Angew. Chem. Int. Ed.*, **43**(40): 5400-5402.

195. De Cremer, G., Roeffaers, M.B.J., Baruah, M., Sliwa, M., Sels, B.F., Hofkens, J., DeVos, D.E. (2007) Dynamic disorder and stepwise deactivation in a chymotrypsin catalyzed hydrolysis reaction. *J. Am. Chem. Soc.,* **129**(50): 15458-15459.

196. Rajagopalan, P.T.R., Zhang, Z., McCourt, L., Dwyer, M., Benkovic, S.J., Hammes, G.G. (2002) Interaction of dihydrofolate reductase with methotrexate: ensemble and single-molecule kinetics. *Proc. Natl. Acad. Sci.*, **99**: 13481-13486.

197. Xie, X.S. and Lu, H.P. (1999) Single-molecule enzymology. *J. Biol. Chem.*, **274**(23): 15967-15970.

198. Xie, X.S. (2001) Single-molecule approach to enzymology. *Single Mol.*, **2**(4): 229-236.

199. Avila, L.Z. and Whitesides, G.M. (1993) Catalytic activity of native enzymes during capillary electrophoresis: an enzymatic microreactor. *J. Org. Chem.*, **58**(20): 5508-5512.

200. Ogle, J.M. and Ramakrshnam, V. (2005) Structural insights into translational fidelity. *Annu. Rev. Biochem.*, **74**: 129-77.

201. Freist, W., Pardowitz, I., Cramer, F. (1985) Isoleucyl-tRNA synthetase from bakers' yeast: multistep proofreading in discrimination between isoleucine and valine with modulated accuracy, a scheme for molecular recognition by energy dissipation. *Biochemistry*, **24**: 7014-23.

202. Parker, J. (1981) Mistranslated Protein in *Escherichia coli*. *J. Biol. Chem.*, **256**: 9770-9773.

203. Martinez-Bilbao, M., Holdsworth, R.E., Edwards, L.A., Huber, R.E. (1991) A highly reactive β-galactosidase (*Escherichia coli*) resulting from a substitution of an aspartic acid for gly-794. *J. Biol. Chem.*, **266**: 4979-4986.

204. Funatsu, G. and Wittman, H.G. (1972) Location of amino-acid replacements in protein S12 isolated from *Escherichia coli* mutants resistant to streptomycin. *J. Mol. Biol.*, **68**: 547-550.

205. Carter, A.B., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T., Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interaction with antibiotics. *Nature*, **407**: 340-348.

206. Ruusala, T. and Kurland, C.G. (1984) Streptomycin preferentially perturbs ribosomal proofreading. *Mol. Gen. Genet.*, **198**: 100-104.

207. Mikkola, R. and Kurland, C.G. (1991) Is there a unique ribosome phenotype for naturally occurring *Escherichia coli*? *Biochimie*, **73**: 1061-1066.

208. Boe, L. (1992) Translational errors as the cause of mutations in *Escherichia coli*. *Mol. Gen. Genet.*, **231**: 469-471.

209. Menninger, J.R., Coleman, R.A., Tsai, L. (1994) Erythromycin, lincosamides, peptidyl-tRNA dissociation, and ribosome editing. *Mol. Gen. Genet.*, **243**: 225-233.

210. Chattopadhyay, S., Pal, S., Pal, D., Sarkar, D., Chandra, S., Das Gupta, C. (1999) Protein folding in *Escherichia coli*: role of 23S ribosomal RNA. *Biochim. Biophys. Acta*, **1429**: 293-298.

211. Ruusala, T., Anderson, D., Ehrenberg, M., Kurland, C.G. (1984) Hyper-accurate ribosomes inhibit growth. *EMBO*, **3**: 2575-2580.

212. Winzor, D.J., Jones, S., Harding, S.E. (2004) Determination of protein charge by capillary zone electrophoresis. *Anal. Biochem.*, **333**: 225-229.

213. Freund, J.E. (2001) Modern Elementary Statistics. Prentice Hall New Jersey 10th ed. pp. 345-367.

214. Breul, A., Kuchinke, W., von Wilcken-Bergmann, B., Muller-Hill, B. (1991) Linker mutagenesis in the *lacZ* gene of *Escherichia coli* yields variants of active β-galactosidase. *Eur. J.Biochem.*, **195**: 191-194.

215. Rodbard, D. and Chrambach A. (1971) Estimation of molecular radius, free mobility, and valence using polyacrylamide gel electrophoresis. *Anal. Biochem.*, **40**: 95-134.

216. Johansson, E., Majka, J., Burgers, M.J. (2001) Structure of DNA polymerase δ from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **276**: 43824-43828.

217. Winzor, D.J. (2004) Determination of the net charge (valence) of a protein: a fundamental but elusive parameter. *Anal. Biochem.*, **325**: 1-20.

218. Kalman, F., Ma, S., Fox, R.O., Horvath, C. (1995) Capillary electrophoresis of S. nuclease mutants. *J. Chromatogr. A*, **705**: 135-154.

219. Basu, M. and Mohanty, P.K. (2009) Stochastic modeling of single molecule Michaelis-Menten kinetics. epublished, arXiv:0901.2844.

220. Lockless, S.W. and Ranganathan, R. (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science*, **286**(5438): 295-299.

221. Juers, D.H., Hakda, S., Matthews, B.W., Huber, R.E. (2003) Structural basis for the altered activity of Gly794 variants of *Escherichia coli* β-galactosidase. *Biochemistry*, **42**(46): 13505-13511.

222. Tang, L., Breinig, S., Stith, L., Mischel, A., Tannir, J., Kokona, B., Fairman, R., Jaffe, E.K. (2006) Single amino acid mutations alter the distribution of human Porphobilinogen synthase quaternary structure isoforms (Morpheeins). *J. Biol. Chem.*, **281**(10): 6682-6690.

223. Whitehead, D.J., Wilke, C.O., Vernazobres, D., Bornberg-Bauer, E. (2008) The look-ahead effect of phenotypic mutations. *Biology Direct*, **3:** 3-18.

224. Dean, A.M., Dykhuizen, D.E., Hartl, D.L. (1988) Fitness effects of amino acid replacements in the β-galactosidase of *Escherichia coli*. *Mol. Biol. Evol.*, **5**(5): 469-485.

225. van Oijen, A.M. (2008) Cutting the forest to see a single tree? *Nat. Chem. Biol.*, **4**: 440-443.

226. Dubnau, D. and Losick, R. (2006) Bistability in bacteria. *Mol. Microbiol.*, **61**(3): 564-572.

227. Schmidt, B.F., Adams, R.M., Requadt C., Power S., Mainzer S.E. (1989) Expression and nucleotide sequence of the *Lactobacillus bulgaricus* β-galactosidase gene cloned in *Escherichia coli*. *J Bacteriol.*, **171**(2): 625-635.

228. Nguyen, T.H., Splechtna, B., Steinbock, M., Kneifel, W., Lettner, H.P., Kulbe, K.D., Haltrich, D. (2006) Purification and characterization of two novel β-galactosidases from *Lactobacillus reuteri*. *J. Agric. Food Chem.*, **54**: 4989-4998.

229. de Macias, M.E., Nadra, M.N., Saad, S., Holgado, P., Oliver, G. (1983) Isolation and properties of β-galactosidase of a strain of *Lactobacillus helveticus* isolated from natural whey starter. *J Appl. Biochem.*, **5**: 275-281.

230. Vetere, A. and Paoletti, S. (1998) Separation and characterization of three β-galactosidases from *Bacillus circulans*. *Biochim et biophys acta*, **1380**: 223-231.

231. Cai, L., Friedman, N., Xie, X.S. (2006) Stochastic protein expression in individual cells at the single molecule level. *Nature*, **440**(7082): 358-362.

232. Turner, E.H., Lauterbach, K., Pugsley, H.R., Palmer, V.R., Dovichi, N.J. (2007) Detection of green fluorescent protein in a single bacterium by capillary electrophoresis with laser-induced fluorescence. *Anal. Chem.*, **79**(2): 778-781.