STRUCTURE AND FUNCTION OF THE NES RELAXASE AND C-TERMINAL DOMAINS REQUIRED FOR VANCOMYCIN RESISTANCE TRANSFER

Jonathan Seaver Edwards

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2012

Approved By:

Matthew Redinbo, Ph.D.

Gary Pielak, Ph.D.

Edward Collins, Ph.D.

Stephen Frye, Ph.D.

Steven Matson, Ph.D.

©2012
Jonathan Seaver Edwards
ALL RIGHTS RESERVED

Abstract

JONATHAN SEAVER EDWARDS: Structure and Function of the NES Relaxase and C-terminal Domains Required for Vancomycin Resistance Transfer
(Under the direction of Matthew R. Redinbo)

Antibiotic resistance has become a large burden in the healthcare setting due to increasing rates of mortality and increased healthcare costs due to extended treatment. *Staphylococcus aureus* is one of the most prevalent healthcare associated infections (HAI) and has recently become resistant to vancomycin (VRSA). This is a serious problem as vancomycin is one of the last lines of defense against resistant gram-positive bacteria. Given this new resistance combined with the fact that novel antibiotic approval has steadily been decreasing we sought to better understand the process of conjugative plasmid transfer (CPT) for the vancomycin resistance plasmid in *S. aureus*, pLW1043.

We focused on a key enzyme involved in CPT, the relaxase. The relaxase on pLW1043 has been identified through sequence analysis and is termed the nicking enzyme of *S. aureus* (NES). The relaxase initiates and terminates CPT. When CPT has been initiated the relaxase will bind to and *nic* the DNA at a specific location on the plasmid termed the *oriT*.

We hypothesized that inhibition of the relaxase would halt the spread of antibiotic resistance and could result in a novel class of antibiotics. To investigate this hypothesis, we set out to determine the structure of NES, understand its biological function, and investigate its interaction with the *oriT* DNA. Little work has been conducted in gram-positive bacteria and by understanding this critical enzymes role in CPT we hoped to provide new insights to the scientific community.

In the following chapters we discuss the structure of the N-terminal relaxase domain in complex with DNA. Furthermore, we will use this structural information combined with

biochemical assays to identify how NES recognizes and site specifically cleaves the DNA at the proper location. We will then investigate the activity of NES with different metals and shed light on the elements required for proper cleavage and religation. Using this data, we then propose a novel strategy for the inhibition of NES and the development of novel antibiotics. Finally, we will analyze the structure of the C-terminal domain, which has no known function, and propose a role for its function in the DNA processing reaction.

Dedication

To my grandfather, Charles (Kroko) Edwards, who provided many special childhood memories and instilled a respect for education from a young age. I can only hope to one day posses, the amazing passion and intellect you display with such humility. Thank you.

Acknowledgements

It would be impossible to adequately thank or list all of those who have supported me in my endeavors and helped to shape the person I am today. I consider myself extremely lucky to have so many loving and motivating people in my life. Without them, I would not have been able to complete my graduate work and have the motivation to continue on. I will do my best to thank those who have been instrumental in this journey of life and I hope to impart my sincerity and appreciation to you all.

I first have to thank my wonderful family. From a young age education, honesty, and respect were highlighted as key components of our beliefs. My parents, Peter and Joanne, were amazingly patient and humble always keeping me focused and supporting me with love. I can still remember wanting to be just like my Dad, hoping one day to be as smart and funny as he was. His intellect is unparalleled and he leads his life with a humility that is an example for all. My mother is my backbone, who has taught me more about life than any textbook could possibly hope to do. Still to this day, I find myself thinking, "Wow, back then I didn't understand what she was doing, but now it all makes sense." Without her, I am not quite sure where I would be today. My sister and brother, Sarah and Ethan, are both equally amazing people; a sign of what a wonderful job my parents have done raising us. Sarah is a smart and funny woman who always keeps me in line. Ethan could be one of the nicest people on the planet and his unrivaled enthusiasm for life is a joy to be a part of. I am excited to see where we all end up, and I know the journey will be the best part. Thank you all for your support and love during this arduous process.

I am fortunate to have a large and diverse family. It would require another ten pages if I was to list and thank them all. To all of you, I have learned something different from each of you

and I am forever grateful to the life lessons you have instilled upon me. From learning how to be an honest and genuine person to experiencing cultures throughout the world, what you have taught me will stay with me forever. I consider myself lucky to have known each one of you and I am honored to be part of our loving family.

There is one special person I have to acknowledge that means more to me than he will ever know. Peter D'Amato was instrumental in my development from a young boy into a man. The amount of people's lives he touched in a positive light is impossible to count. He was a loving father, devoted coach, and an excellent mentor. I still remember the life lessons he taught us on the way to soccer games or during sleepovers at his house. I love to cook because of him and I have a serious respect for Stevie Ray Vaughan, imparted through evenings of listening and watching him play the guitar. Mr. D'Amato's love of life was amazing and his sense of what was important never waivered. I know he is looking down on all of us that were in his life, and I hope I can be half as good of a man as he was. Thank you.

I would like to acknowledge Bob Lenzi, my high school biology teacher. I still talk to him today and he was instrumental in allowing me to develop my passion for science. Were it not for him, I would have never majored in biology and gone down the path I am on today. I am still puzzled as to how and why he put up with my antics, but I am forever grateful for it. He took me under his wing, taught me how to play racquetball after school, and is a great friend. I will never forget his patience and wisdom.

I must acknowledge the UNC community as well. This is a special place, and I will have some tarheel blood in me forever. The opportunities that were afforded to me during my time at this University were nothing short of spectacular. I will look back on this time with fond memories. I would like to thank Barry Lentz for giving me the opportunity to be a part of the biochemisty and biophysics program at UNC. Although his classes were a nightmare, his patience and thorough love of education kept me going. I have become a better scientist due to his training and enthusiasm.

I could write an entire thesis about why my P.I., Matt Redinbo, is the best advisor on this campus. Without him I would not have had nearly as many chances to learn new aspects of science and business or grow as a scientist like I did. None of this would have been possible without the environment he provided. He leads our lab as a friend and mentor. His scientific insights are amazing and I have already been able to see the effects of what he has imparted on me during my interviews. Through my time in his lab I have not only been able to grow as a scientist but as a person, through the independence he allows. We have had some great times together, even if he did bring that guy Bill to dinner in San Francisco. I consider him a great friend and look forward to seeing what new scientific discoveries come from his work. I would also like to thank my committee: Gary Pielak, Ed Collins, Steve Matson, and Stephen Frye. They were instrumental in helping me through graduate school, both scientifically and on a personal level. Their foresight and knowledge set me up for success.

Lastly, I would like to thank everyone that I worked with in the Redinbo lab. It was a wonderful five years and I learned something from each one of you. Specifically I would like to thank Mike Miley, Yuan Cheng, and Rebekah Nash. Mike introduced me to the lab and taught me the basics about science and life in grad school. His advice has been so helpful through the years. Yuan taught me about work ethic and scientific rigor. He may be one of the smartest and hardest working individuals I have ever met. In the words of Yuan, "Just do it." Rebekah was my predecessor and helped me learn many of the techniques I used in my research. She is one of the most thorough and rigorous researchers I have encountered. She was an excellent mentor and an even better person. In closing, without the support of my friends, family, and mentors I would not have been nearly as successful or enjoyed my time at UNC as I have. Thank you all for your support and love throughout the years.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

%	percent
°	degree
4-MUA	4-methylumbelliferone
6FAM	fluorescein
A	percent activity at time t
α	alpha-helix
A_0	initial activity at t=0
Å	angstrom
AChE	acetylcholinesterase
A_{max}	maximal percent recovery
Au	gold
AU	asymmetric unit
β	beta-strand
BChE	butyrylcholinesterase
BSA	bovine serum albumin
C	Celsius
CD	circular dichroism
CDC	center for disease control
cDNA	complementary deoxyribonucleic acid
cm	centimeter
CPD	cysteine protease domain
CPT	conjugative plasmid transfer
ddH_2O	double distilled water
DLS	dynamic light scattering
DMPNPP	dimethyl-*para*-nitrophenylphosphate

DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	double stranded deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
FA	fluorescence anisotropy
g	g-force
HAI	healthcare acquired infection
hCE1	human carboxylesterase 1
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGT	horizontal gene transfer
HL1	hairpin loop 1
HL2	hairpin loop 2
HPLC	high pressure liquid chromatography
hr(s)	hour(s)
HUH	histidine, hydrophobic amino acid, histidine
I	intensity
IC_{50}	the concentration at 50% inhibition
I_0	intensity at q=0
ICP-MS	inductively coupled plasma mass spectrometry
IPTG	isopropyl-β-D-thiogalactopyranoside
K	kelvin
k_2	rate of enzyme phosphonylation
k_3	rate of enzyme dephosphonylation
kB	kilo base
k_{cat}	rate of enzyme turnover

K_D	dissociation constant
kDa	kilodalton
KF	potassium fluoride
K_i	dissociation constant for the inhibitor
k_i	bimolecular rate of inhibition
K_m	the Michaelis-Menten constant
k_{obs}	observed rate of enzyme reactivation
LB	lysogeny broth
LIC	ligation independent cloning
M	molar
µ	micro
MBP	maltose binding protein
µg	microgram
MIC	minimum inhibitory concentration
µL	microliter
µM	micromole
µm	micron
mg	milligram
min	minute
mL	milliliter
mm	millimeter
mM	millimole
MOB	set of mobility enzymes
MRSA	methicillin resistant *Staphylococcus areus*
NaCl	sodium chloride
NAG	N-acetylglucosamine

NAM	N-acetylmuramic acid
NES	nicking enzyme of *Staphylococcus aureus*
nic	specific site where a relaxase cleaves the DNA substrate
nm	nanometer
N-terminus, N-term	amino terminus
OD	optical density
oligo	oligonucleotide
OP	organophosphate
oriT	origin of plasmid transfer
PBP2A	penicillin-binding protein transpeptidase
PCR	polymerase chain reaction
Pd	palladium
PEG	polyethylene glycol
PEI	polyethyleneimine
pH	negative log (base 10) of the molar concentration of hydronium ions
pKa	acid dissociation constant
pM	picomole
pNPB	para-nitrophenyl butyrate
PVDF	polyvinylidene fluoride
R5	NH_2-SSKKSGSYSGSKGSKRRIL-CO_2H
R_g	radius of gyration
rpm	revolutions per minute
s	second
SAXS	small-angle X-ray scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel
SEC	size exclusion chromatography

SEM	scanning electron microscopy
σ	error
t	time
$t_{1/2}$	half-time of reactivation
T4SS	type four secretion system
TAE	tris base, acetic acid, EDTA
TBE	tris, borate, EDTA
TEM	transmission electron microscopy
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus
T_m	melting temperature
TMOS	tetramethylorthosilicate
Tris	tris(hydroxymethyl)aminomethane
u	unit
v	remaining percent enzyme activity
v/v	volume to volume
VISA	vancomycin-intermediate *Staphylococcus areus*
V_{max}	maximum enzyme rate
VRSA	vancomycin resistant *Staphylococcus areus*
wt	wild type

Chapter 1

Staphylococcus aureus Antibiotic Resistance and Conjugative DNA Transfer

1.1 *Staphylococcus aureus* Effect on Healthcare and Antibiotic Resistance

Staphylococcus aureus (*S. aureus*) is a gram-positive, facultative anaerobic, coccal bacteria. It typically resides as part of the normal flora of the skin, especially around the nasal passages [1]. Nosocomial infections, or healthcare acquired infections (HAIs), are one of the top 10 causes of deaths per year in the United States and cost the healthcare system more than $20 billion annually [2]. *S. aureus* is one of the four most common HAIs. One of the major reasons these infections are so difficult to treat is due to the presence of antibiotic resistance. Shortly after the introduction of the first antibiotic, penicillin, it was observed that bacteria are able to gain resistance to these compounds [3]. Furthermore, statistics from the Center for Disease Control (CDC) demonstrate that more people die from methicillin resistant *S. aureus* than AIDS [4]. Because of these issues, the World Health Organization has declared antibiotic resistance a top priority for the health of the human population. Antibiotic resistance results in a large cost to our healthcare system by extending the average length of stay, making low cost drugs ineffective, and allowing secondary illnesses to progress faster [5]. Antibiotics select for bacteria that are able to evade the effects of these compounds, and this selection has placed significant pressure on pharmaceutical companies to stop development of antibiotics because it is difficult to make a return on the investment. This lack of development results in an innovation gap for novel compounds [6] (Figure 1.1).

One of the major concerns for healthcare professionals is the acquisition of vancomycin resistance by bacteria. Vancomycin is a glycopeptide antibiotic (Figure 1.2) that is efficient at killing recalcitrant infections in gram-positive bacteria and is typically used as a

last resort due to drug administration and toxicity issues [7]. It was discovered in 1953 in a soil sample from the jungles of Borneo. The primary mechanism of action is to inhibit cell wall synthesis in gram-positive bacteria. It does this by binding to terminal D-alanyl-D-alanine moieties found on the N-acetylmuramic acid (NAM) and N-acetylglucosamine (NAG) peptides. This results in a large hydrophilic molecule being present in the forming cell wall that prevents the efficient crosslinking and elongation of polymers, and results in cell death.

S. aureus has shown the ability to rapidly develop resistance to numerous antibiotics. As previously mentioned, after penicillin was introduced in the 1940's it was found that 60% of isolates contained a β-lactamase that contributed to penicillin resistance [8]. Methicillin was introduced in 1961 as semisynthetic version of a β-lactam that would be able to evade the effects of penicillin resistance. However, methicillin-resistant *S. aureus* (MRSA) was soon identified. This resistance was possible because of the acquisition of the mecA gene that encodes for an alternate penicillin-binding protein transpeptidase (PBP2A) [9,10]. Furthermore, other classes of antibiotics have been developed that had an effect on *S. aureus* but resistance soon emerged. These antibiotics include macrolides, lincosamides, aminoglycosides and quinolones, as well as newer antibiotics such as linezolid, daptomycin, and quinupristin-dalfopristin. As a result of this resistance, vancomycin, licensed in 1958, has become the drug of choice for difficult to treat *S. aureus* infections. Contrary to the other antibiotics it took nearly 40 years for any type of resistance to vancomycin to emerge. In 1997 the first reported case of a clinical isolate of *S. aureus* that had reduced susceptibility to vancomycin (MIC = 8 μg/mL) was identified and termed vancomycin-intermediate *S. aureus* (VISA) [11]. This strain of *S. aureus* is able to evade the effects of vancomycin through producing more D-ala-D-ala targets, which vancomycin recognizes, in its peptidoglycan layer. By creating a thicker cell wall the bacteria are able to evade the effects of vancomycin before it is able to completely disrupt the cell wall resulting in cell death [12-14]. In 2002 a *S. aureus* isolate with a MIC for vancomycin of 1024 μg/mL was found in a dialysis patient in Michigan indicating high-level resistance [15]. Upon sequencing the resistance plasmid

(pLW1043) in this isolate the *vanA* gene cluster was found [16]. This strain of vancomycin-resistant *S. aureus* is now referred to as VRSA.

The *vanA* gene cluster allows for the resistance pathway to become inducible in VRSA. This is a great benefit to the bacteria because VISA isolates and other resistant bacteria often spend considerable resources evading the effects of antibiotics. By only supporting vancomycin resistance pathways when the drug is present, these bacteria are at significant advantage to other resistant bacteria. Upon sensing vancomycin the bacteria are able to initiate this alternate pathway for cell wall synthesis. Effectively, the bacteria eliminate the old D-ala-D-ala targets and replace them with D-ala-D-lac [17]. This allows for the continual growth of the cell wall while in the presence of vancomycin (Figure 1.3).

1.2 Conjugative Plasmid Transfer in Gram Positive Bacteria

Bacteria store their genetic material on either a chromosome or an extra-chromosomal element. The extra-chromosomal elements are often mobilizable and are able to move between bacterial hosts as well as into chromosomes. Extra-chromosomal elements include gene cassettes, transposons, bacteriophages, and single- or double-stranded plasmids [18-20]. Bacteria are able to transfer their genetic material through two routes depending on the direction of transfer. Vertical gene transfer is the process through which bacteria will duplicate the genetic material and distribute a copy of both chromosomal and extra-chromosomal elements to each of the daughter cells. Horizontal gene transfer (HGT) is the process in which the extra-chromosomal elements are copied and distributed to a recipient bacterial cell. HGT is the primary pathway responsible for the dissemination of extra-chromosomal elements that contain antibiotic resistance genes [21].

HGT transfer of double-stranded DNA (dsDNA) plasmids between bacteria is referred to as conjugative plasmid transfer (CPT). dsDNA plasmids are classified into two major categories based on their ability for transfer: conjugative or mobilizable. Both types of plasmids contain the mobility (MOB) set of enzymes that allow for DNA replication. Conjugative plasmids contain the genes for the type 4 secretion system (T4SS) required to process and transfer the DNA from the

donor to recipient bacteria during CPT. Mobilizable plasmids can be transferred but do not posses all of the genes necessary for self transfer requiring another plasmid or chromosome with T4SS capabilities [19].

To date, more information is known about the process of HGT in gram-negative bacteria, and as such, that information is used as a model system for gram-positive bacteria. The major differences in the transfer mechanism are believed to be due to the difference in the formation of the cell walls and membranes between gram-negative and gram-positive bacteria. Gram-negative bacteria have a cytoplasmic membrane followed by a thin layer of peptidoglycan and an outer membrane. In contrast, gram-positive bacteria have a cytoplasmic membrane covered by a thick outer layer of peptidoglycan. The main differences in HGT between the species is believed to arise in the way these cells form cell-cell contact to initiate the transfer process [21].

When cells make tight cell-to-cell contact CPT can begin. During conjugation, two multi-protein complexes form to process the DNA. The first complex, T4SS in gram-negative species, assembles to mediate DNA transfer from donor to recipient. The second complex, the relaxosome, forms in unison with the T4SS and prepares the DNA for transfer [21]. Together these complexes allow the cell to transfer a plasmid via conjugation (Figure 1.4). The components of the relaxosome vary between species and plasmids but have similar functions. In order to have successful conjugation two factors are required of the plasmid. An origin of transfer (oriT) must be present to act as a cis regulator and a site where conjugation initiates and terminates termed *nic*. Also, there must be a gene encoding a site-specific relaxase that cleaves a phosphodiester bond of single stranded DNA at the *nic* site [22-24]. These enzymes are present in all relaxosomes and provide a novel target for antibiotics.

1.3 The Relaxase NES – Domain Organization and Mechanism

Relaxases are an integral part of the relaxosome and are responsible for carrying out the DNA cleavage and religation reactions. They are transferases that can cleave and religate ssDNA through transesterification in a site- and strand-specific manner. They contain an N-terminal

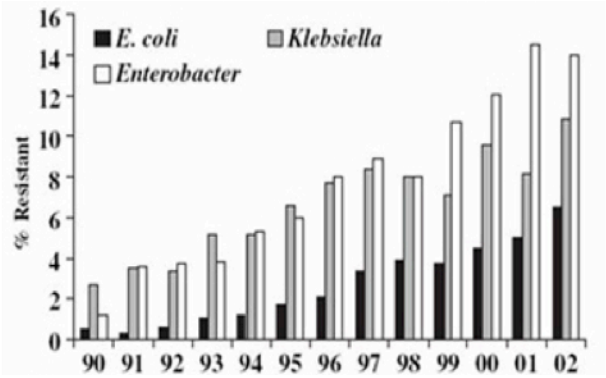
relaxase domain while the C-terminal domains can have helicase, primase, or other activities that depend on the host in which the relaxase resides [25]. The relaxase present in the VRSA plasmid pLW1043 is named nicking enzyme of *S. aureus* (NES). Interestingly, in NES the C-terminal domain shows no relationship to these previously described functions based on sequence homology or secondary structure prediction. Relaxases are separated into two large families, Mob or Rep, based on the order of the elements contained in their conserved motifs [26]. Motif I consists of one to four conserved catalytically active tyrosines. Motif II will have either a conserved aspartic acid or glutamic acid [27]. Motif III contains two conserved histidine residues that are proposed to coordinate a divalent metal ion in the transesterifcation reaction [28]. The typical histidine motif is HUHUUU where U represents a bulky hydrophobic amino acid. The one catalytically active tyrosine of NES is N-terminal to the HUH motif, thus classifying it as a Mob type relaxase.

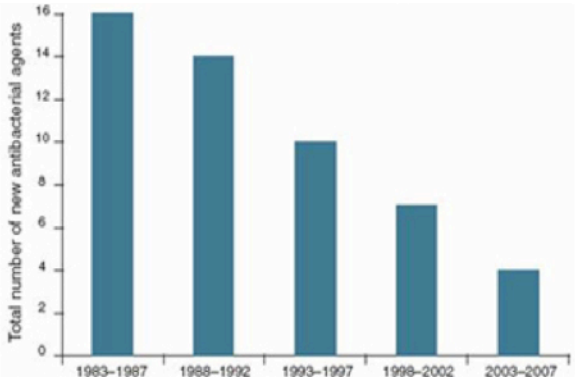
Relaxases initiate and terminate conjugation through their single stranded nicking and ligation activity. They recognize and bind DNA at a specific site on the plasmid called the oriT. By recognizing specific bases present in the oriT, the relaxase can bind to the proper location on the plasmid. Once the relaxase has bound the oriT it will catalyze the cleavage of the scissile phosphodiester bond [29]. The active site will position the DNA so that the tyrosine hydroxyl can act as a nucleophile to attack the scissile bond in the DNA backbone. After cleavage, the 5' end of the DNA is covalently bound to the relaxase via a phosphodiester bond. Figure 1.5 depicts the proposed mechanism for NES. The HUH motif is shown coordinating a Ni^{2+} cation as was determined from ICP-MS (Table 3.3). A specific base, such as a water molecule, is most likely abstracting the hydrogen from the hydroxyl on the catalytic tyrosine because the structure did not reveal any amino acid side chains within the appropriate distance. Once the DNA is positioned by the divalent cation the tyrosine can attack the scissile phosphate at the *nic* site between Gua28 and Cyt29. The 3' end of the cut DNA now has a hydroxyl and is released from the active site. The

DNA will then be transferred in a 5' to 3' direction into the recipient bacteria. The reaction is reversible with the 3' hydroxyl acting as a nucleophile [29].

NES was first classified by Climo *et. al* in 1996 [25]. They were able to demonstrate that NES is required for the successful conjugative transfer of the plasmid pGO1. They were also able to locate the oriT on the plasmid and show that the relaxase cleaves where it would be predicted to based upon other sequence similarities of well characterized conjugative plasmids. The N-terminal region of NES on plasmid plw1043 has a 26% sequence similarity with the MobA relaxase but no sequence similarity is present for the C-terminal domain. The C-terminal could prove to be a dimer interface as suspected for TraI of the plasmid RP4 [22] or act as a docking domain in the relaxosome. Once bound, the relaxase will catalyze the transesterifcation of the DNA at the *nic* site. This will allow the DNA to be unwound and transferred into the recipient cell. Many structures of gram-negative relaxases have been solved with inhibitors or in complex with DNA. However, there is only one structure of a Mob relaxase containing a single tyrosine in the active site. Furthermore, it was in apo form and from a gram-negative organism [30]. The goal of this research is to learn more about the biochemical function of NES as well as determine the crystal structure of the enzyme for use in structural analysis through mutations and to develop novel antibiotics that inhibit its function.

A. Increase in Drug Resistance Among Pathogens B. Decreased Number of New Drugs in the Market





Source: Clinical Micro Infect 2004; 10 (Suppl. 4): 1-9 Source: Nat Biotech 2006 24: 1521

Figure 1.1 Increases in Antibiotic Resistance Compared to New Drugs in the Market

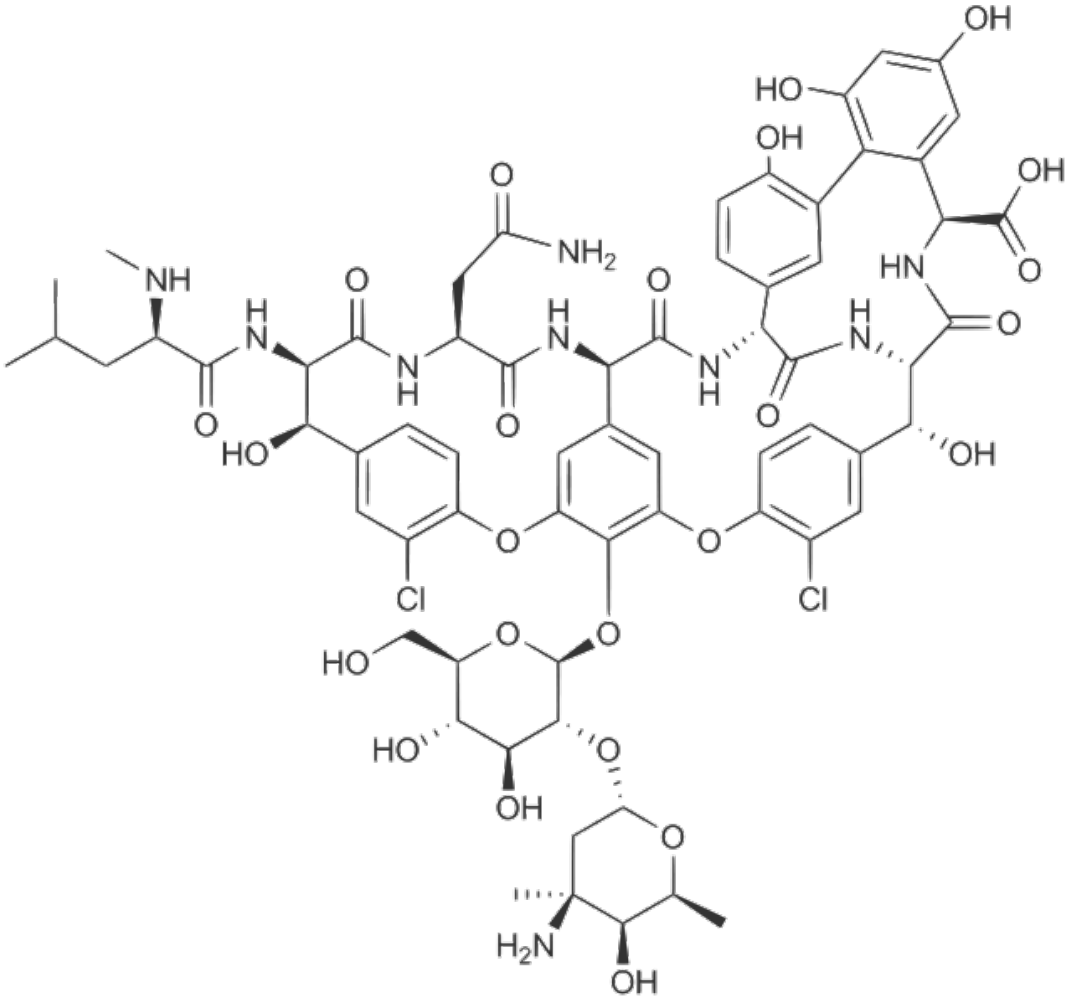
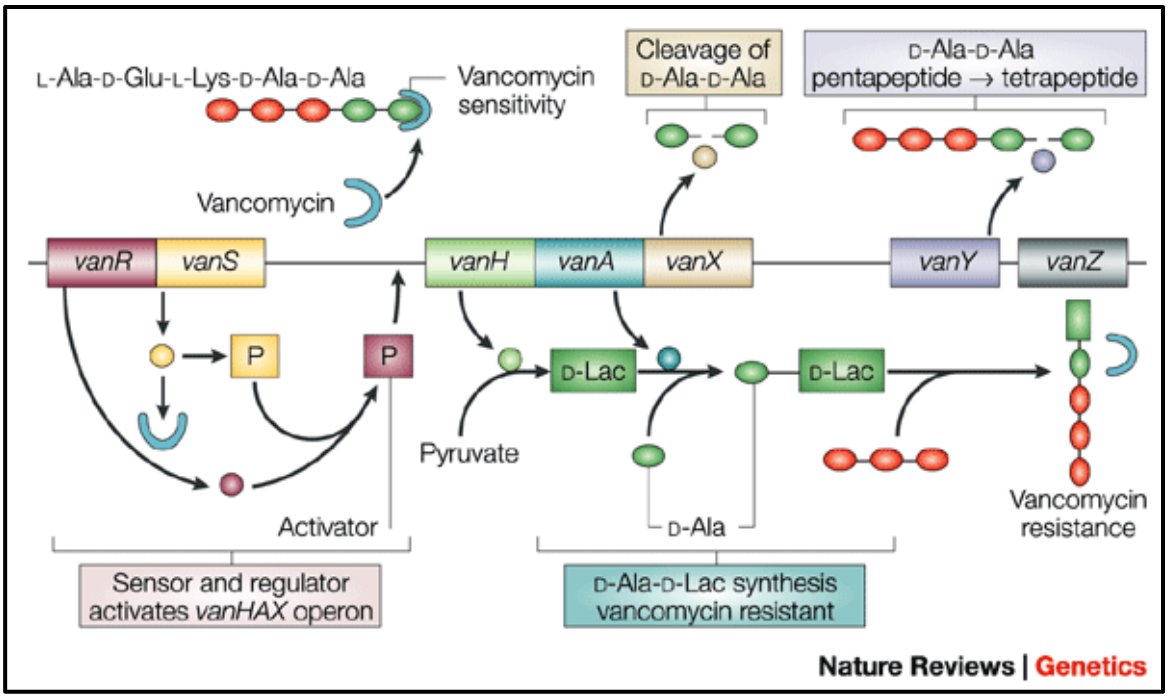


Figure 1.2 Structure of Vancomycin



Source: Hughes, D. (2003) Exploiting genomics, genetics and chemistry to combat antibiotic resistance, *Nature reviews 4*, 432-441.

Figure 1.3 VRSA and VRE Vancomycin Resistance Mechanism

A two-component regulatory system VanR–VanS regulates vancomycin resistance in vancomycin-resistant enterococci (VRE) and vancomycin-resistant *Staphylococcus aureus* (VRSA) strains. VanS is a membrane-associated sensor (of vancomycin) that controls the level of phosphorylation of VanR. VanR is a transcriptional activator of the operon encoding VanH, VanA and VanX. VanH is a dehydrogenase that reduces pyruvate to D-Lac, whereas VanA is a ligase that catalyses the formation of an ester bond between D-Ala and D-Lac. Vancomycin does not bind to D-Ala-D-Lac, which leads to vancomycin resistance. VanX is a dipeptidase that hydrolyses the normal peptidoglycan component D-Ala-D-Ala, which prevents it from causing vancomycin sensitivity. VanY is a D,D-carboxypeptidase that hydrolyses the terminal D-Ala residue of late peptidoglycan precursors that are produced if elimination of D-Ala-D-Ala by VanX is not complete. So, D-Ala-D-Lac replaces the normal dipeptide D-Ala-D-Ala in peptidoglycan synthesis resulting in vancomycin resistance150. VanZ confers resistance to teicoplanin by an unknown mechanism.

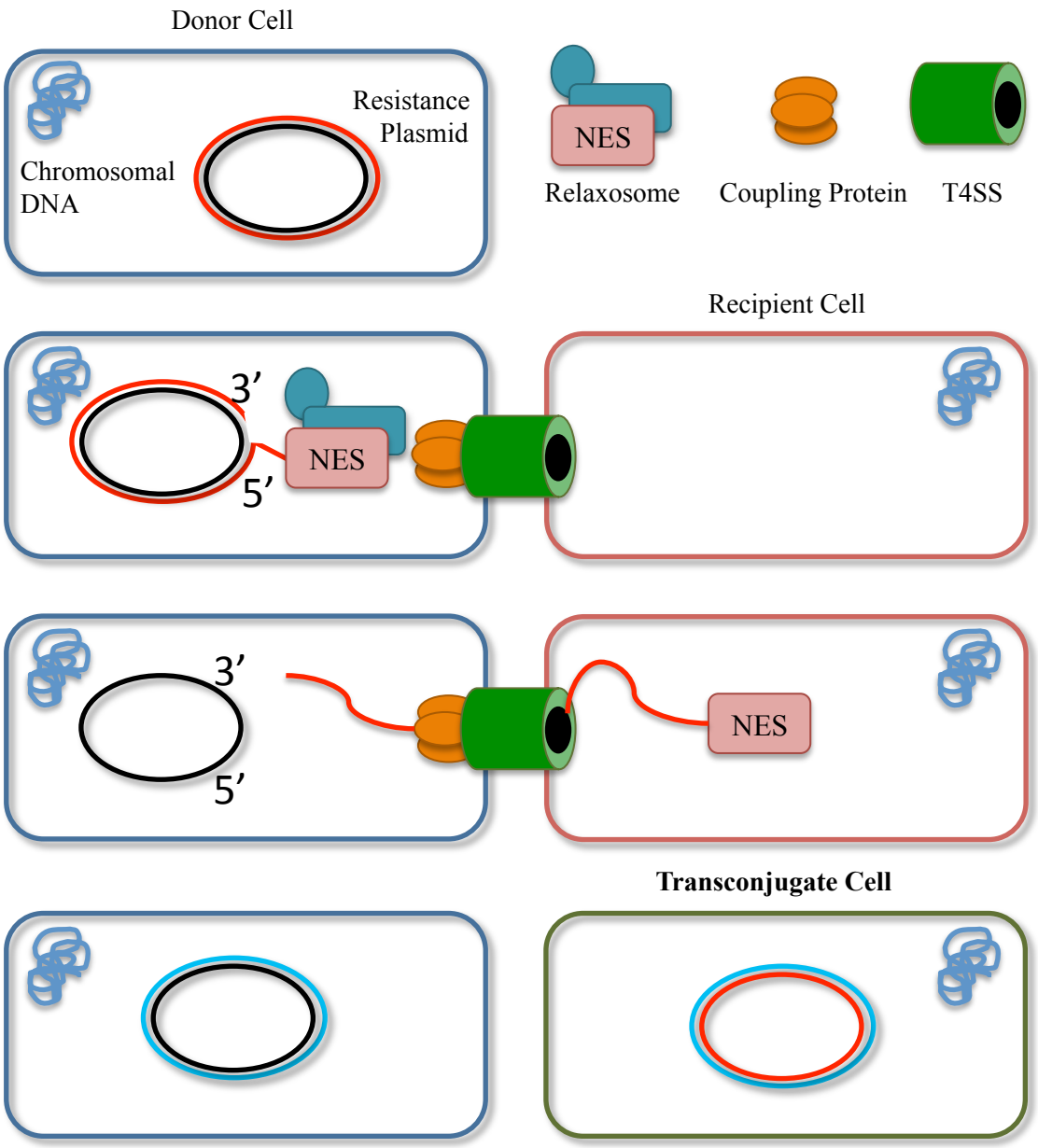


Figure 1.4 Conjugative Plasmid Transfer

Schematic representation of CPT in bacteria. A donor cell is shown with chromosomal DNA and a resistance plasmid that is capable of transfer. Upon the formation of the Type 4 Secretion System (T4SS) the relaxosome, including NES, will bind to the oriT and perform the transesterification reaction at the *nic* site to form the covalent complex. The components in the relaxosome will join with the coupling protein to allow NES to transfer the DNA into the recipient cell in a 5' to 3' direction. Once transfer is complete the 3' hydroxyl on the transferred strand can attack the tyrosine-phosphate bond to recircularize the DNA. The host replication machinery will then duplicate the DNA so each will have double stranded resistance plasmids.

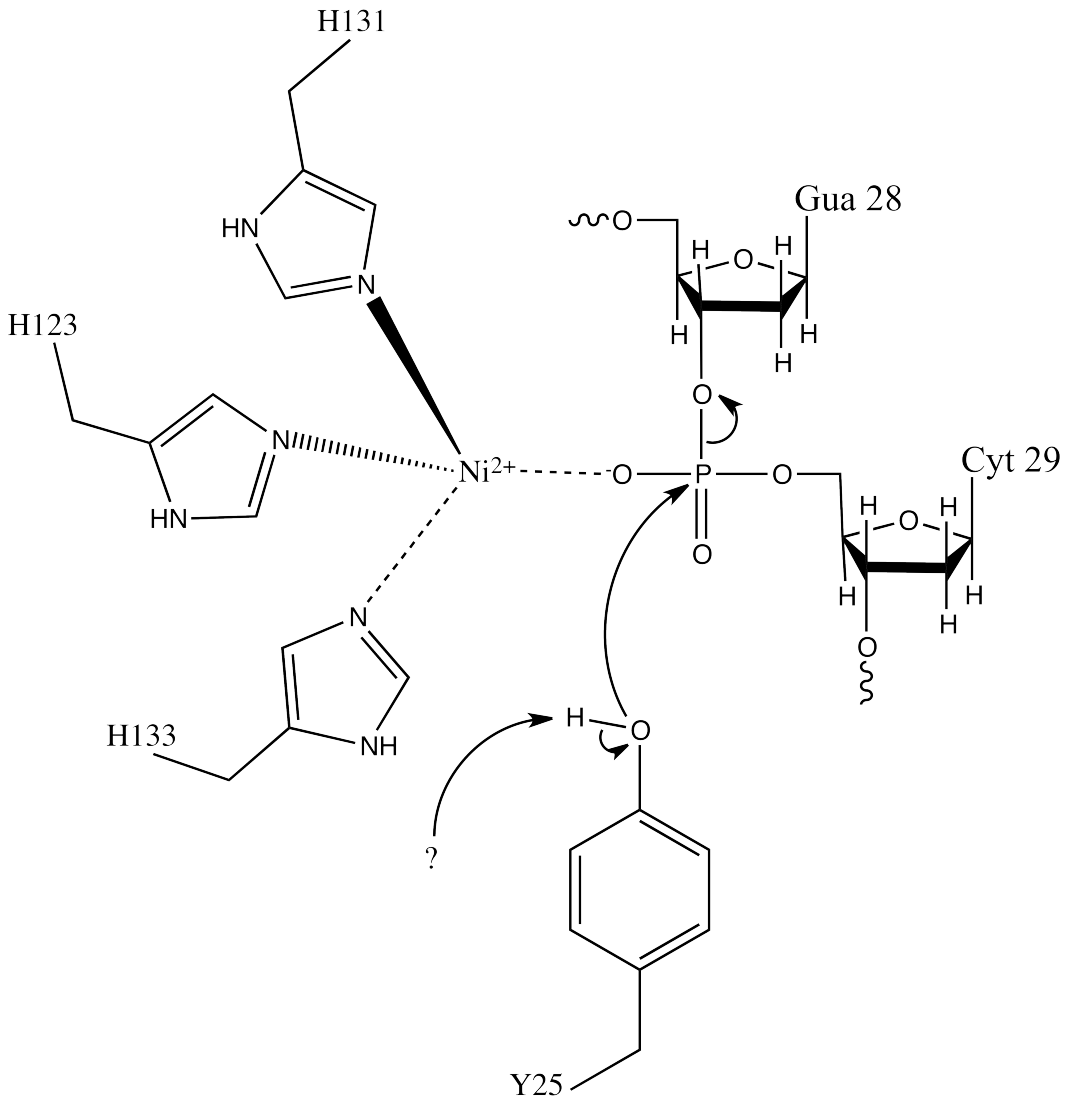


Figure 1.5 Schematic Diagram for the Proposed Catalytic Mechanism of NES

The tyrosine residue of NES initiates a S_N2 attack on the scissile phosphate between the 5' guanine and 3' cytosine of the *nic* site. The hydroxyl is believed to be activated by a specific base such as an ordered water molecule. The HUH motif is shown coordinating a Ni^{2+} used to position the DNA for the cleavage reaction.

1.5 References

1. Kluytmans, J., van Belkum, A. & Verbrugh, H. Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. *Clin Microbiol Rev* **10**, 505-520 (1997).

2. *US Department of Health and Human Services Agency for Healthcare Research and Quality. 2010. Healthcare-Associated Infections*, http://www.ahrq.gov/qual/hais.htm, (October 2011).

3. Rammelkamp, C. H., Maxon, T. Resistance of Staphylococcus aureus to the action of penicillin. *Proc. Royal Soc. Experim. Biol. Med.* **51**, 386-389 (1942).

4. *Centers for Disease Control and Prevention. ABCs Report: Methicillin-Resistant Staphylococcus aureus, 2005 (Updated)*, http://www.cdc.gov/abcs/reports-findings/survreports/mrsa05.html.

5. World Health Organization (2001). WHO Global Strategy for Containment of Antimicrobial Resistance.

6. Walsh, C. Where will new antibiotics come from? *Nat Rev Microbiol* **1**, 65-70, doi:10.1038/nrmicro727 (2003).

7. Knox, J. R. & Pratt, R. F. Different modes of vancomycin and D-alanyl-D-alanine peptidase binding to cell wall peptide and a possible role for the vancomycin resistance protein. *Antimicrob Agents Chemother* **34**, 1342-1347 (1990).

8. Barber, M. & Rozwadowska-Dowzenko, M. Infection by penicillin-resistant staphylococci. *Lancet* **2**, 641-644 (1948).

9. Matsuhashi, M. *et al.* Molecular cloning of the gene of a penicillin-binding protein supposed to cause high resistance to beta-lactam antibiotics in Staphylococcus aureus. *J Bacteriol* **167**, 975-980 (1986).

10. Katayama, Y., Ito, T. & Hiramatsu, K. A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in Staphylococcus aureus. *Antimicrob Agents Chemother* **44**, 1549-1555 (2000).

11. Hiramatsu, K. *et al.* Dissemination in Japanese hospitals of strains of Staphylococcus aureus heterogeneously resistant to vancomycin. *Lancet* **350**, 1670-1673, doi:10.1016/S0140-6736(97)07324-8 (1997).

12. Cui, L. *et al.* Cell wall thickening is a common feature of vancomycin resistance in Staphylococcus aureus. *J Clin Microbiol* **41**, 5-14 (2003).

13. Cui, L., Murakami, H., Kuwahara-Arai, K., Hanaki, H. & Hiramatsu, K. Contribution of a thickened cell wall and its gultamine nonamidated component to the vancomycin resistance expressed by Staphylococcus aureus Mu50. *J Antimicrob Chemother* **44**, 2276–2285 (2000).

14. Sieradzki, K., Pinho, M. G. & Tomasz, A. Inactivated pbp4 in highly glycopeptide-resistant laboratory mutants of Staphylococcus aureus. *J Biol Chem* **274**, 18942-18946 (1999).

15. Chang, S. *et al.* Infection with vancomycin-resistant Staphylococcus aureus containing the vanA resistance gene. *N Engl J Med* **348**, 1342-1347 (2003).

16. Weigel, L. M. *et al.* Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus. *Science* **302**, 1569-1571 (2003).

17. Hughes, D. Exploiting genomics, genetics and chemistry to combat antibiotic resistance. *Nat Rev Genet* **4**, 432-441 (2003).

18. de la Cruz, F. & Davies, J. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol* **8**, 128-133 (2000).

19. Lanka, E. & Wilkins, B. M. DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* **64**, 141-169, doi:10.1146/annurev.bi.64.070195.001041 (1995).

20. Smillie, C., Garcillan-Barcia, M. P., Francia, M. V., Rocha, E. P. & de la Cruz, F. Mobility of plasmids. *Microbiol Mol Biol Rev* **74**, 434-452, doi:10.1128/MMBR.00020-10 (2010).

21. Grohmann, E., Muth, G. & Espinosa, M. Conjugative plasmid transfer in gram-positive bacteria. *Microbiol Mol Biol Rev* **67**, 277-301, table of contents (2003).

22. Pansegrau, W. & Lanka, E. Mechanisms of initiation and termination reactions in conjugative DNA processing. Independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPalpha plasmid RP4. *J Biol Chem* **271**, 13068-13076 (1996).

23. Llosa, M. & de la Cruz, F. Bacterial conjugation: a potential tool for genomic engineering. *Res Microbiol* **156**, 1-6, doi:10.1016/j.resmic.2004.07.008 (2005).

24. Wilkins, B. M. a. L., E. *Bacterial Conjugation*. Clewell, D. B., Ed. edn, 105-136 (Plenum Press, 1993).

25. Climo, M. W., Sharma, V. K. & Archer, G. L. Identification and characterization of the origin of conjugative transfer (oriT) and a gene (nes) encoding a single-stranded endonuclease on the staphylococcal plasmid pGO1. *J Bacteriol* **178**, 4975-4983 (1996).

26. Koonin, E. V. & Ilyina, T. V. Computer-assisted dissection of rolling circle DNA replication. *Biosystems* **30**, 241-268 (1993).

27. Francia, M. V. *et al.* A classification scheme for mobilization regions of bacterial plasmids. *FEMS Microbiol Rev* **28**, 79-100 (2004).

28. Pansegrau, W., Schroder, W. & Lanka, E. Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J Biol Chem* **269**, 2782-2789 (1994).

29. Agundez, L., Gonzalez-Prieto, C., Machon, C. & Llosa, M. Site-specific integration of foreign DNA into minimal bacterial and human target sequences mediated by a conjugative relaxase. *PLoS One* **7**, e31047, doi:10.1371/journal.pone.0031047 (2012).

30. Monzingo, A. F., Ozburn, A., Xia, S., Meyer, R. J. & Robertus, J. D. The structure of the minimal relaxase domain of MobA at 2.1 A resolution. *J Mol Biol* **366**, 165-178 (2007).

Chapter 2

The Structure of the NES Relaxase in Complex with DNA

2.1 Introduction

The relaxase domain of NES was examined to elucidate its mechanism of DNA binding and cleavage during conjugative plasmid transfer (CPT). Indeed, because relatively little detailed biochemical and structural work has been conducted on plasmid systems from Gram-positive bacteria, a central goal was to understand if the relaxase on pLW1043 interacts with DNA in the same manner as the relaxases found on other plasmids. The plasmid pLW1043 was found in a clinical VRSA strain; however, it is 99% identical to other, less virulent plasmids, including pGO1 and pSK41 that do not encode vancomycin resistance. The NES genes from these three plasmids are identical in sequence. To date, the structures of the relaxases from the F plasmid (TraI), R388 (TrwC), pCU1 (TraI), and R1162 (MobA) have been determined [1-6]. NES was predicted to contain only one functional tyrosine residue in the active site, similar to that observed for MobA. Prior to this work, only relaxases from the plasmids F and R388 (TraI and TrwC, respectively) with multiple redundant tyrosines had been determined in complex with DNA [1-4]. By analyzing the differences in binding and specificity, the long-term goal of this project is to use potential subtle changes in structure to design specific inhibitors that can impede the function of the relaxase. An additional central objective is to learn more about the relatively under-studied Gram-positive conjugative process.

We determined the 2.9 Å resolution structure of the relaxase domain of NES (residues 1-220) in complex with DNA containing 28 nucleotides including a hairpin and single-stranded region leading to the *nic* site. This represents the first structure of a single active site tyrosine relaxase in complex with DNA substrate. While the DNA was observed in the same general

orientation as seen in TrwC-DNA complex, the NES-DNA complex structure reveals some overall key differences, as well as more detailed distinctions. Previous relaxases have been described as having an overall architecture of a "hand" with the β-sheet core being the "palm" and the surrounding α-helices representing the "thumb" and "fingers" holding the DNA in place. We observed our enzyme was able to bind to the DNA substrate without this thumb feature as well as numerous differences in the residues that provide specificity for the DNA. Before the more detailed distinctions unique to the NES relaxase are outlined, the methods and results related to the characterization of NES will be described.

2.2 Construct Design and Cloning of the NES Relaxase and Mutants

To determine the structure of the NES enzyme, multiple attempts at crystallization were made. Initially the full-length protein was used for crystallization trials, but significant degradation was found in the resulting crystals. (This will be covered in Chapter 4 when discussing the C-terminal domain structure.) The amino acid sequence of NES (GenBank: AE017171) was input in the Phyre server [7] and the relaxase minMobA was found to be the best sequence and structural match for N-terminal portion of NES. Analysis of the resulting alignment showed that the characteristic motifs of a relaxase were present and that NES should have a similar fold compared to MobA. The structure of minMobA incorporated the N-terminal 186 amino acid fragment of the enzyme. Primers were designed for the NES relaxase domain using the alignments to search for optimal places to end the fragment such as loops or disordered regions and 5 constructs were selected. Integrated DNA Technologies (IDT) commercially synthesized the primers and all cloning or mutagenesis was verified using Eton Biosciences Inc. The primers were diluted to 100 mM stock solutions in ddH_20 and subsequently diluted to 10 mM working stock solutions. Standard ligation independent cloning techniques, as described by Stols *et al.* [8], were used to generate the expression plasmids encoding the desired NES fragments. The PCR amplified DNA fragments were inserted into empty pMCSG7-Lic and pMCSG7-Lic-MBP

vectors. The N-terminal His_6 or His_6 maltose-binding protein (MBP) tag was cleaved off after purification using tobacco etch virus (TEV) protease.

NES primers were also designed for incorporation into the CPD expression system that was generated in our lab by Joseph Lomino and Daniel Yao. The initial CPD expression system described by Shen *et al.* [9] was modified through codon optimization, changing the hexa-his tag to a deca-his tag, and adding an longer glycine linker between the protein and cleavage site (CPD-Lasso). Again, the primers were ordered from IDT, amplified, and cloned into the CPD-Lasso expression vector using the *NdeI* (5') and *BamHI* (3') restriction endonuclease sites. The constructs result in a C-terminal CPD-His_{10} tag on the protein that was cleaved using 2 mM inositol hexakisphosphate ($InsP_6$) incubated for 3 hours at 4 °C.

All point mutations were done using site-directed mutagenesis in the corresponding expression plasmid. Again, sequencing for the mutations was confirmed using Eton Biosciences, Inc. for DNA sequencing.

2.3 Expression and Purification of the NES Relaxase

The cloned constructs were transformed into chemically competent *E. coli* BL21 (DE3) AI cells. Transformed cells were grown in 1.5 liters of lysogeny broth (LB), in the presence of ampicillin, at 37 °C, with shaking, to an optical density of 0.6-0.8. An L-Arabinose solution was added to the growth at a final concentration of 0.2% v/v, and the temperature was reduced to 18 °C for 30 minutes. Protein expression was induced with 100 μM isopropyl-β-D-thiogalactopyranoside (IPTG) and cells were allowed to grow for 16 hours. The resulting growths were spun at 4,500 x *g* for 30 minutes at 4 °C and split into six pellets that were stored at -80 °C. Individual pellets were resuspended in nickel A buffer (500 mM NaCl, 25 mM imidazole, 20 mM potassium phosphate buffer pH 7.4, and 0.02% v/v sodium azide) along with protease inhibitor tablets (Roche), DNase, and lysozyme. The slurry was lysed using a Fischer Scientific Sonic Dismembrator at 50% power for 3 cycles of 1 minute with 1 second intervals at 4 °C. The lysed cells were then spun at 18,500 x *g* for 1 hour. The supernatant was filtered and purified over a

HisTrap column (GE Healthcare). None of the N-terminal His_6 or His_6 MBP constructs expressed solubly after the removal of the tag by TEV. For the CPD constructs they were allowed to bind to the HisTrap column and incubated with $InsP_6$ for 3 hours at 4 °C. The resulting protein was then further purified over a Superdex 200 column (GE Healthcare) pre-equilibrated in sizing buffer (300 mM NaCl, 50 mM Tris buffer pH 7.4, and 0.02 % v/v sodium azide). Fractions containing NES 1_220, which was found to contain the functional and soluble relaxase domain of NES, were combined and used for site directed mutagenesis as well as the following experiments.

To obtain protein in complex with DNA an active site mutant Y25F was used. A piece of DNA with the sequence 5'-ACGCGAACGGAACGTTCGCATAAGTGCGCC-3' was synthesized by IDT and used as the substrate. After nickel purification the DNA was added in a 2.25 molar excess to the enzyme and size exclusion chromatography was performed for final purification. The presence of DNA was confirmed measuring the absorbance of A260/280 on a nanodrop (Thermo Scientific).

2.4 Circular Dichroism and Dynamic Light Scattering of NES 1_220

Circular dichroism (CD) and dynamic light scattering (DLS) were used in conjunction to confirm that the protein was folded, stable, and monodispersed in solution. CD measurements were made on a Chirascan CD Spectrometer (AppliedPhotophysics) with the protein diluted into CD buffer (200 mM KF, 10 mM potasim phosphate buffer pH 7.4) to a concentration of 5 μM. A wavelength scan from 190-260 nm revealed a protein that was folded containing both α-Helices and β-Sheets. This was determined by the presence of a positive peak at 190 nm and a double negative peak at 208 and 225 nm. However, the double negative peaks were reduced as to what would be expected for only α-Helices and this can be attributed to the presence of β-sheets that have a negative peak at 215 nm which will affect the curves (Figure 2.2.A). The resulting ensemble of peaks made us believe that the protein was in fact folded and a combination of both sheets and helices which is predicted when comparing NES 1_220 to similar structures. Thermal denaturation was performed while monitoring the 209 nm wavelength that corresponded to the

most distinguishable peak from the wavelength scan. The melting temperature (T_m) was determined to be approximately 58 °C by observing where half of the protein was in a folded versus unfolded state from the data (Figure 2.2.B).

Dynamic light scattering (DLS) was performed on a Dynapro Dynamic Light Scattering Plate Reader (Wyatt Technology Corp.). Prior to measurement the samples were spun at 13,000 x *g* to ensure no dust or sediments would affect the measurement. The sample of NES 1_220 was measured at 10 mg/mL and showed a single monodisperse peak with a polydispersity of 10.7% and a radius of hydration of 2.6 nm (Figure 2.3). The molecular weight calculated from this reading was 33 kDa, larger than the actual molecular weight of 26 kDa for NES 1_220 but perhaps indicative of an elongated protein shape in the absence of DNA. Taken together, though, these data are representative of a well folded, stable, and pure protein.

2.5 Crystallization, Data Collection, and Data Processing

Selenomethionine-substituted protein was used in crystal trials to obtain phase information upon data collection. NES 1_220 + DNA (10 mg/mL) was crystallized in a solution containing 16% PEG 8,000, 120 mM calcium acetate, 80 mM sodium cacolydate (pH 6.5), and 20% glycerol v/v. The trays were set with drop ratios of 1:1, 1:2, and 2:1 of protein to mother liquor respectively. Vapor diffusion hanging-drop trays were set at 20 °C to obtain well formed crystals. The crystallization solution was found to be an adequate cryoprotectant in subsequent diffraction experiments performed at the UNC Biomolecular X-ray Crystallography Facility. The crystals were flash frozen in liquid nitrogen for data collection at 100K.

Data sets were obtained on the 23_ID_B beamline maintained by The National Institute of General Medical Sciences and National Cancer Institute Collaborative Access Team (GM/CA-CAT) at the Advanced Photon Source (APS) part of the Biosciences Division (BIO) at Argonne National Laboratory (ANL). A fluorescence scan was performed at the selenium absorption edge to confirm that the seleno-methione substitution was successful. The collection strategy used included 1-second exposures with a 1° rotation between frames and inverse beam every 10° to

collect 100° of data. The data sets were indexed and scaled using HKL-2000 processing software [10]. Indexing resulted in resulted in a space group of $P6_422$ with one NES relaxase-DNA complex per asymmetric unit. The best data was obtained when the resolution was cut to 2.9 Å (Table 2.1). Initial phases were determined using the Phenix software suite with the AutoSol function [11]. Two of the three heavy atom sites were successfully located. This was a result of the first amino acid, methionine, not being observable in the final structure, most likely as a result of it being processed away during expression as was seen in the MobA structure [5]. The model was built using Coot [12] by placing the seleno-methiones and iteratively building the protein from these starting points. Refinements were performed using the Phenix refine application [11].

2.6 Analysis of the NES Relaxase in Complex with DNA Structure

2.6.1 The Overall Fold of the Relaxase Domain NES

One NES relaxase-DNA complex was observed per crystallographic asymmetric unit, indicating that the enzyme interacts and functions with the DNA as a monomer, a conclusion supported by size exclusion chromatography studies (data not shown). The residues 2_196 of NES 1_220 were built into the electron density. The terminal methionine was most likely processed away during expression and the 24 C-terminal residues are at the end of a loop that is likely disordered in this construct. Additionally, the 5' DNA adenine and the two 3'-terminal cytosine bases were not visible in the electron density. The overall fold of NES 1_220 was similar to the other relaxases studied in regard to the active site and position of the catalytic tyrosine. As seen in Figure 2.1, the core of the enzyme is composed of 5 central anti-parallel β-strands. The remaining two β-strands and four α-helices are structural elements and position the catalytic tyrosine for nucleophilic attack. The two DNA binding loops clamp on either side of the hairpin for proper substrate recognition and positioning. The central β-sheet is ordered β1-β5-β7-β6-β4 when looking down into the active site of the molecule. The HUH motif (H131 and H133) is found at the beginning of β7 and the third H (H123) is located at the 3' end of strand β6. This motif is conserved in relaxases and coordinates the metal ion at the active site. The catalytic

tyrosine is located on α1 in an orientation that would allow it to perform its nucleophilic attack on the scissile phosphate at the *nic* site. Ni^{2+} was built into the density adjacent to the chelating histadine residues, as it was clear there was a metal bound from the difference density maps; additional ICP-MS data indicated this metal was Ni^{2+} and was most highly incorporated after purification over the Ni-resin (data not shown). However, H131A mutants were made to disrupt the metal chelation site in the HUH motif and the Ni^{2+} concentration was significantly reduced (data not shown). From Figure 2.1, it is clear the single stranded stretch of the DNA curves around the enzyme and is held in a U turn conformation where it would be possible for the nucleophilic attack to occur from the oxygen on the catalytic tyrosine. In the structure one can see that the 3' guanine appears to be flipped away from the active site. This was due to a crystal-packing defect where the 3' cytosine (C27) and guanine (G28) bases stacked with the corresponding 3' terminal bases in another molecule in the crystal (Figure 2.4).

Secondary structure alignments comparing NES to the other relaxase structures reveal some key differences between the enzymes. Figure 2.5.A compares the structure-annotated sequences of NES_220 to MobA (PDB: 2NS6), F TraI (PDB: 2A0I), TrwC (PDB: 1QX0), and pCU1 TraI (PDB: 3L57). The structures were aligned using DaliLite [13-15] for pairwise comparison of the protein structures. When the relaxase domain of NES was aligned to MobA, the RMSD was 2.2 Å over 162 equivalent Cα positions, and 28% sequence identity was observed. The RMSDs for F TraI, TrwC, and pCU1 TraI were all 3.1 Å over 158, 155, and 153 equivalent Cα positions, respectively, with 10%, 8%, and 5% sequence identities, respectively. Despite the relatively low sequence identity between structures, especially for the latter three, there were significant similarities in the overall fold. As mentioned previously, the secondary structure elements are arranged in an analogous fashion for all of the enzymes aligned. For example, the N-terminus is located at the beginning of β1, the first strand in the "palm" of the hand. α1 is conserved in all structures; however, the catalytic tyrosine is present on the helix in NES, MobA,

and F TraI, while in TrwC and pCU1 TraI the catalytic tyrosines are present on the loop between α1 and β4 in Figure 2.5.A. This may be due to the fact that the enzymes with multiple tyrosines have functional redundancy, and by placing such tyrosines in loops, mobility can be enhanced. In NES exclusively, there is a short loop containing β2,3 before reaching the conserved β4 which forms the outer edge of the other side of the "palm". α3 is highly conserved in all of the aligned structures, is located at the top of the beta-sheet core, and is responsible for positioning hairpin loop 1 that interacts with the DNA substrate in NES. The loop between α3 and β5 is variable between the structures and will be discussed in the next section. Again, β5 is highly conserved, as it is part of the central core of the enzyme. F, TrwC, and pCU1 have an extra helix at the base of the β-sheet core where NES and MobA only contain the conserved α4. Another difference between NES and MobA, though, when compared to other structures is the presence of a loop containing two short β-strands between α4 and β6. This loop is responsible for interacting with the minor groove of the DNA substrate as seen in the structure of TrwC and NES in complex with DNA and will be discussed in more detail in the next section. β6 and β7 are highly conserved, as they are the final two strands present in the β-sheet core. Another major difference can be seen between β7 and α5 in Figure 2.5.A. In both NES and MobA there is a long loop that interacts with the DNA hairpin. In the other three, this loop is shorter and contains two small β-strands. Finally, α5 is highly conserved in all of the structures and packs against α4 to form the back of the enzyme. The subtle differences that were found when comparing the hairpin Loop1 and Loop2, as well as the interactions with the single stranded stretch of the DNA substrate, explain how these enzymes are able to achieve specificity for their specific substrates, as outlined below.

2.6.2 Structural Elements for the Recognition of the DNA Substrate

The DNA substrate has extensive interactions with the relaxase that are centered on the two major DNA regions – the hairpin and the stretch of single-strand. Interestingly, the majority of the interactions are found to be at the base of the hairpin and the section of single-strand DNA closest to the hairpin (Figure 2.5.B). The hairpin present at the 5' side of the oriT has extensive

interactions with two loops in NES 220 termed Hairpin Loop1 and Hairpin Loop2 (Figure 2.5.A,B). The single-stranded portion of the DNA interacts primarily with key residues found on α5 and β1. In F TraI and TrwC, there are extra α-helices present at the C-terminus that interact with the single stranded portion of the DNA, acting as a 'thumb" clamping it in place [16]. These helices are not present in the smaller, single-tyrosine relaxases NES and MobA. After the *nic* site, we speculate that the single-strand of DNA would continue over the active site and be held in place through interactions with α1, β2, β3, and the loop between the β-strands in NES 220.

Each enzyme will have direct interactions along the single-stranded portion of the DNA. These interactions will be specific to the DNA sequence, how the bases are presented to the enzyme will confer specificity. Figure 2.6 illustrates how two bases on the single stranded portion (Ade20 and Thy21) are exclusively recognized by the DNA bases present, while no interactions occur with the phosphate backbone. Q82 and Y156 form the hydrogen bond network that recognizes Ade20. N8 and R85 are hydrogen donors for the carbonyl oxygen on Thy21. Not all of the bases on the single stranded portion make interactions with the DNA, for example Ade24 and Gua25. Specific bases that do interact with the enzyme have been studied through mutational analysis and will be discussed in the next chapter.

More surprising was the finding that the loops responsible for binding the hairpin portion of the DNA (Figure 2.7) are vastly different between the two sets of enzymes, NES 220, MobA vs. F TraI, TrwC, pCU1 TraI. When comparing the loops that interact with the DNA hairpin portion of the substrate it is necessary to look at the enzymes in the two different groups mentioned above. This is because the DNA hairpin structures are significantly different between the two groups. Currently only two structures of relaxases in complex with the DNA hairpin portion of the substrate have been determined, TrwC and NES 220. When comparing NES 220 + DNA to the MobA structure, two major differences can be seen. Figure 2.8 depicts the alignment for NES and MobA (RMSD 2.2 Å). In Figure 2.8.A the major difference arises in the β2, β3 region of NES. This is the area of the protein where we believe the 3' side of the *nic* site would be

positioned by the enzyme. When NES interacts with the DNA this portion of the enzyme might order to form the β-strands from the small α-helix that is seen in the MobA structure (Figure 2.8.A). This would allow the enzyme to maintain tight contact with the covalently bound portion of the enzyme during substrate transfer into the recipient cell. Another significant change in the enzyme architecture can be observed when looking at the interactions made with the DNA that can be observed from the NES 220 structure. Figure 2.7 depicts the two key loops that clamp the DNA hairpin in order to bind the substrate. Loop1 interacts with the minor groove that is formed in the hairpin while the longer Loop2 interacts with the major groove (Figures 2.5.B, 2.7). These loops have extensive hydrogen bond interactions with both the phosphate backbone as well as a few key bases. Loop1 is in an almost identical position in both the MobA and NES structures. However, Loop2 has a significantly different conformation when bound to DNA. In the NES structure Loop2 is a long loop that packs into the major groove of the DNA acting as one side of the clamp. In the MobA structure this loop forms two short α-helices and is positioned in the open conformation (Figure 2.8.B). One can imagine that when the DNA comes into position on the enzyme the α-helices rearrange into loops, and the new conformation will be stabilized through interactions with the DNA forming the double sided clamp that holds the substrate in place for the attack at the *nic* site.

When comparing NES to TrwC, the larger relaxase construct, more differences can be noted. In the comparison of the overall structural alignment (Figure 2.9) the "thumb" region of TrwC is the first noticeable difference. This region is not present in the NES construct and could either be due to the fact that the construct crystallized for NES was not long enough to observe this region, the C-term may perform the thumb function in NES, or the thumb region regulates the action of the multiple catalytic tyrosines since it is only present on those enzymes. The biochemical data discussed in the next chapter would seem to rule out the second point as this portion of the enzyme is able to specifically bind this substrate with low nanomolar affinity and is functional. However, this region is present in both TrwC and F TraI and may play a role in adding

specificity or binding efficiency for those enzymes. Another difference, noted in MobA as well, is the β2, β3 region of the enzyme. As mentioned previously, this structural arrangement will most likely be dependent on the substrate sequence and has not been thoroughly investigated. Figure 2.5 highlights the significant sequence and structural divergence that occurs in the hairpin Loop1 and Loop2 regions in TrwC and NES. The differences can be seen in the structures presented in Figure 2.10, however they have been separated for visual simplicity. Figure 2.10.A shows the Loop1,2 interactions that were previously discussed for NES. When comparing that to TrwC, Figure 2.10.B, significant differences can be observed both in the enzyme secondary structure as well as the DNA hairpin conformation. In TrwC, Loop1 is significantly longer and across the face of the minor groove interacting with both the phosphate backbone as well as the DNA bases. Loop1 in NES also interacts with the minor groove in the same fashion but is much shorter in length. In TrwC the hairpin appears to be squished leading to a shorter and less well defined minor groove. Loop2 in both constructs interacts with the DNA in a similar fashion despite the structural differences. In NES Loop2 interacts with the DNA backbone along the left side of the loop. The loop winds its way into the major groove where base specific interactions occur. Despite the significant structural and sequence divergence TrwC interacts with the DNA in the same manner. The β-strand on the left side of Loop2, Figure 2.10.B, interacts with the DNA backbone while the loop that makes the turn to the next β-strand makes some specific base interactions. Furthermore, both loops interact with the 5' terminal portion of the substrate.

2.7 Conclusions

By determining the structure of NES 220 in complex with DNA we were able to make some interesting structural observations. Firstly, we are the first group to report the structure of a single tyrosine relaxase in complex with DNA. This result will be critical for our analysis of the DNA binding and cleavage specificity in the following chapter as well as for our efforts to develop novel antibiotics. The overall fold of this structure was similar to all other relaxase structures determined to this date. The central β-sheet core is consistent with the known

architecture and the main differences arise due to key loops required for DNA binding. In the case of NES and MobA the DNA results in subtle conformational changes between Hairpin Loop1 and Loop2 to allow for the binding of the substrate to the enzyme. When looking at the structures found in the TrwC group of enzymes more drastic differences were observed. Again, despite significant lack of sequence homology there was maintenance of the enzyme core. However, the loops required for hairpin binding achieve the same end goal while being in completely different conformations. These differences lead to proper substrate recognition and provide novel sites to design specific compounds that could inhibit the enzyme-substrate interaction that could result in the development of novel antibiotics.

Data Collection	
X-ray source	APS ANL GM/CA-CAT
Space Group	P 6_4 2 2
Unit cell: a,b,c (Å); α,β,γ (°)	63.2, 63.2, 313.8; 90, 90, 120
Wavelength (Å)	0.95471
Resolution (Å) (highest shell)	48.5-2.90 (2.95-2.90)
I/σ	47.9 (5.96)
Completeness (%)	97.7 (79.7)
Redundancy	20.3 (13.2)
Refinement	
Resolution (Å)	48.5-2.90
No. reflections (unique)	8782
R_{work}	22.6
R_{free}	28.5
Molecules per asymmetric unit (AU)	1
No. of amino acids per AU (DNA nucleotides)	195 (27)
No. of waters per AU	43
Average *B*-factors ($Å^2$)	60.8
R.M.S. deviations (Å)	2.08
Bond lengths (Å)	0.008
Bond angles (°)	1.407
Ramachandran (%)	
Preferred	96.3
Allowed	3.7
Outliers	0.0

Table 2.1 Crystallographic Statistics

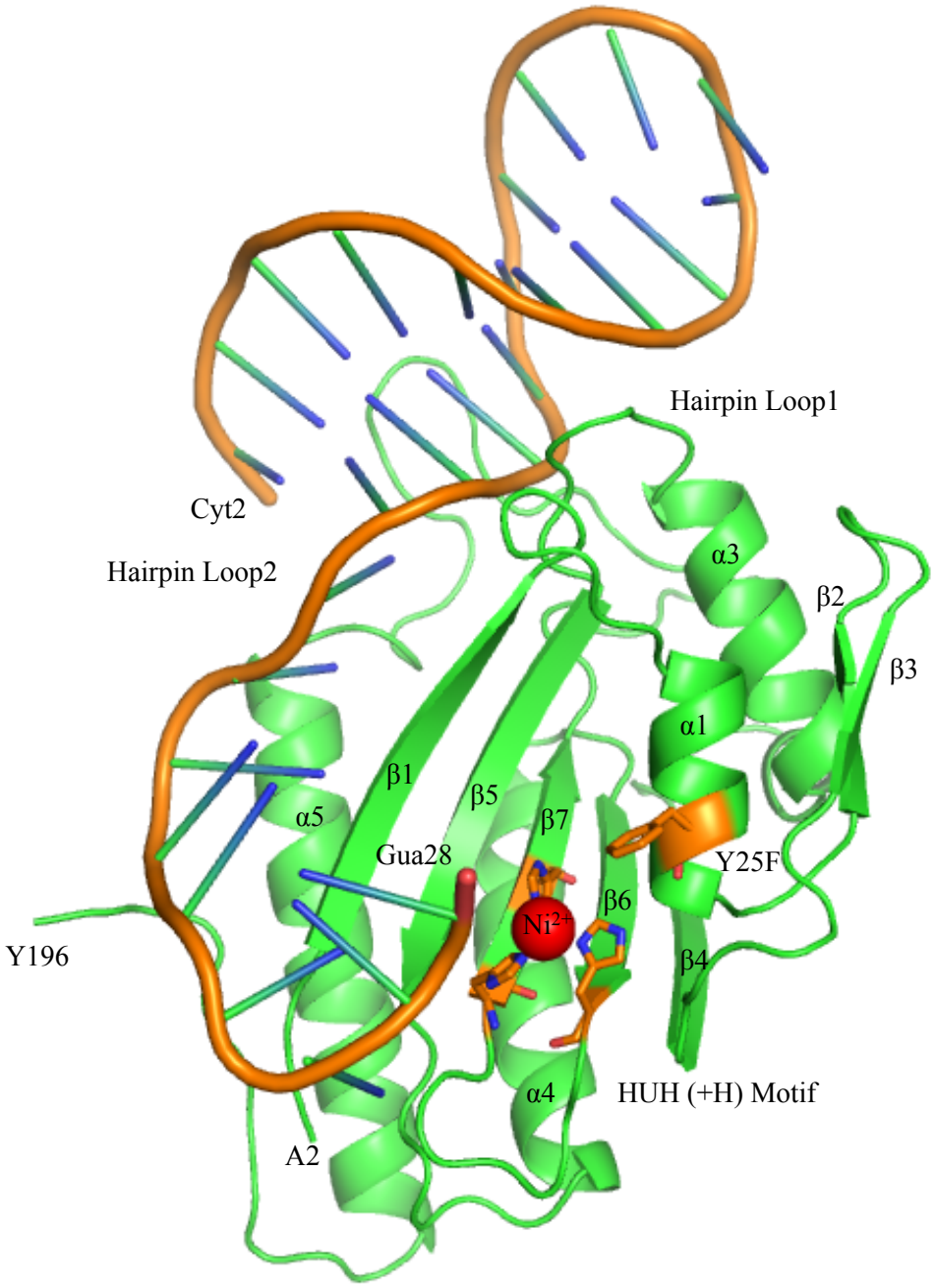


Figure 2.1 Structure of NES 1_220 in complex with DNA

The 2.9 Å crystal structure of NES 220 in complex with DNA. The N-term and C-term amino acids are labeled as well as the 5' and 3' nucleotides. The α-helices, β-strands, and hairpin loops are labeled by the order in which they appear in the sequence. α2 is hidden behind α3. The side chains for the HUH(+H) motif and the catalytic residue are shown.

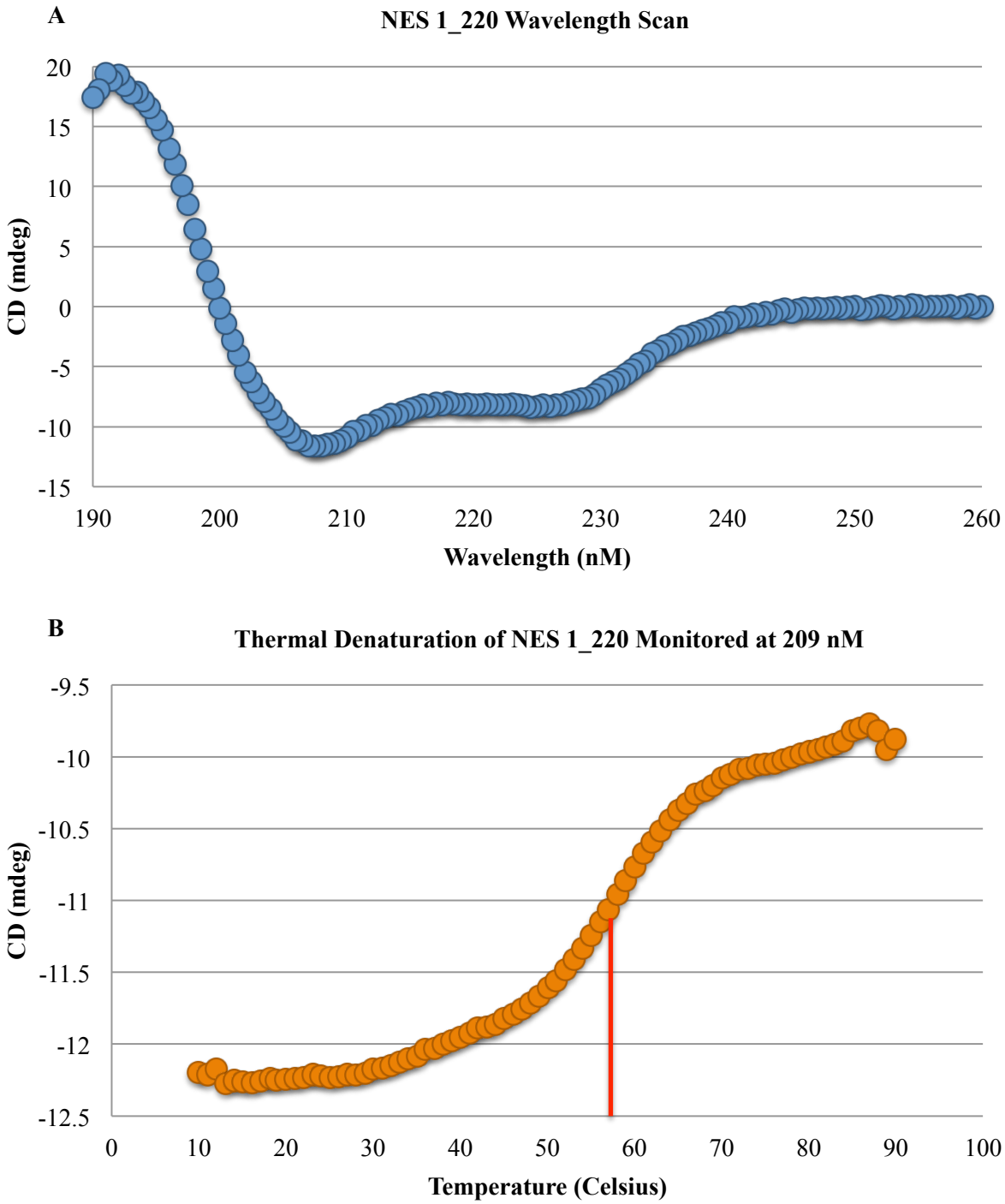
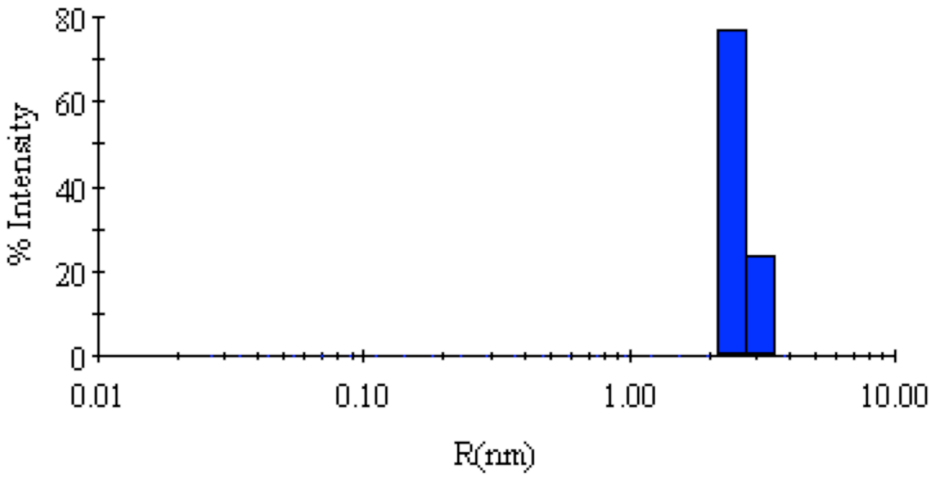


Figure 2.2 Circular Dichroism Analysis of NES 220

A. CD wavelength scan of NES 220 at 5 μM. B. Thermal denaturation of NES 220 monitoring the change in CD signal (mdeg) at 209 nM.



Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
Peak 1	2.6	10.7	33	100.0	100.0

Figure 2.3 Dynamic Light Scattering of NES 1_220

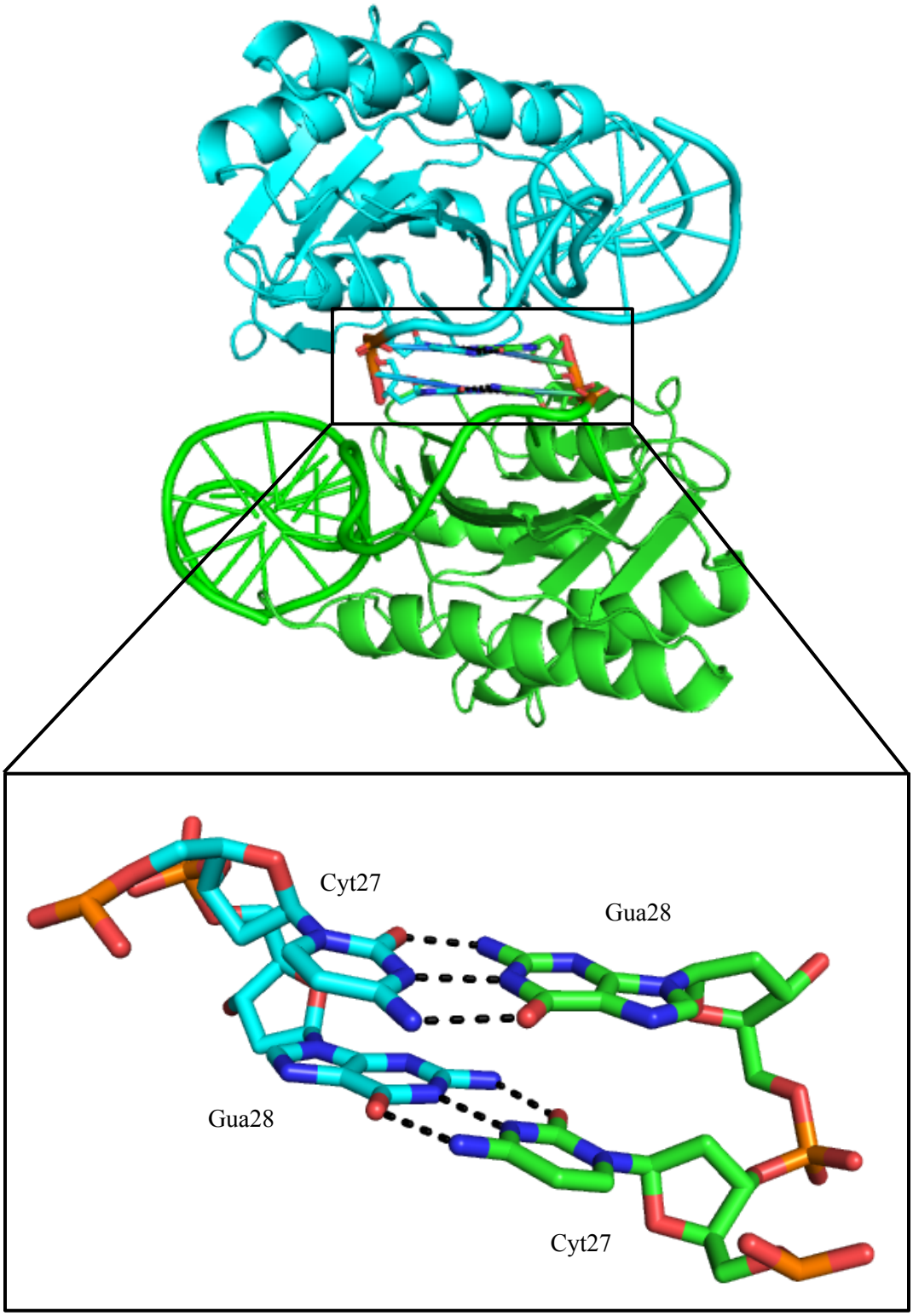


Figure 2.4 Crystal Packing Effect on the 3' End of the DNA Substrate

Two symmetry mates from the crystal packing are depicted in green and cyan. Highlighted above it is clearly visible that the DNA is being pulled out of the active site. The inset depicts the Watson-Crick base paring interactions for the each molecule's terminal DNA bases.

```
                 β1        α1    β2    β3      β4     α2
NES_220          myhfqnkfv-skangqsataksaynsasrikdfkenefkdysnkqcdyseillpnnaddk
MobA             iyhltaktgs--ggqsarakadyiqregk----------yardmdevlhaesghmpefve

F TraI           ----msiaqvr-----sagsagnyytdkdnyyvlgsm----------gerwagrgaeqlg
TrwC DNA         ----lshmvltrqdigraasyyedgaddyyakdgda------------sewqgkgaeelg
pCU1 TraI        ----ldittitrqnvtsvvgyysdakddyyskdssf------------tswqgtgaealg

                 α2   α3      HAIRPIN LOOP1   β5

NES_220          fkd--reylwnkvhdve-------------nrknsqvareiiiglpnefd----------
MobA             ----rpadywdaadlye-----------rangr---lfkevefalpvel-----------

F TraI           lqgsvdkdvftrllegrlpdgadlsrmqdgsnrhrp-gydltfsapksvsmmamlggdkr
TrwC DNA         lsgevdskrfrellagnigeghrimrsatrqdskerigldltfsapksvslqalvagdae
pCU1 TraI        lsgdvesarfkellvgeidtfthmqrhvg-dakkerlgydltfsapkgvsmqalihgdkt

                 α4                          β6      β7

NES_220          -pnsnielakefaeslsnegmi------------------vdlnihkineenp---hahl
MobA             tldqqkalasefaqhltgaerlp-----------------ytlaihaggge---nphchl

F TraI           lidahnqavdfavrqv-ealastrvmtdgqsetvltgnlvmalfnhdtsrdqepqlhtha
TrwC DNA         iikahdravartleqae-araqarqkiqgktriettgnlvigkfrhetsrerdpqlhtha
pCU1 TraI        iieahekavaaavreae-klaqarttrqgksvtqntnnlvvatfrhetsraldpdlhtha

                 HAIRPIN LOOP2                α5

NES_220          lctlrgldknnef------------------epkrkgndyirdwntkekhnewrkrwenv
MobA             miserindgierpaaqwfkryngktpekggaqkt------ealkpkawleqtreawadha

F TraI           vvanvtqhngewk------------------tlssdkvgkt-gfienvyanqiafgrlyr
TrwC DNA         vilnmtkrsdgqw------------------ralk------ndeivkatrylgavynael
pCU1 TraI        fvmnmtqredgqw------------------ralk------ndelmrnkmhlgdvykqel

                 α5

NES_220          qnkhlekngfsvrvsadsy                      Beta Strands
MobA             nraleragh----------                      Alpha Helix

F TraI           eklkeqvealgyetevvgk
TrwC DNA         ahelqklgyqlrygkdgnf
pCU1 TraI        aleltkagyelrynsknnt
```

Figure 2.5.A Secondary Structure Alignment

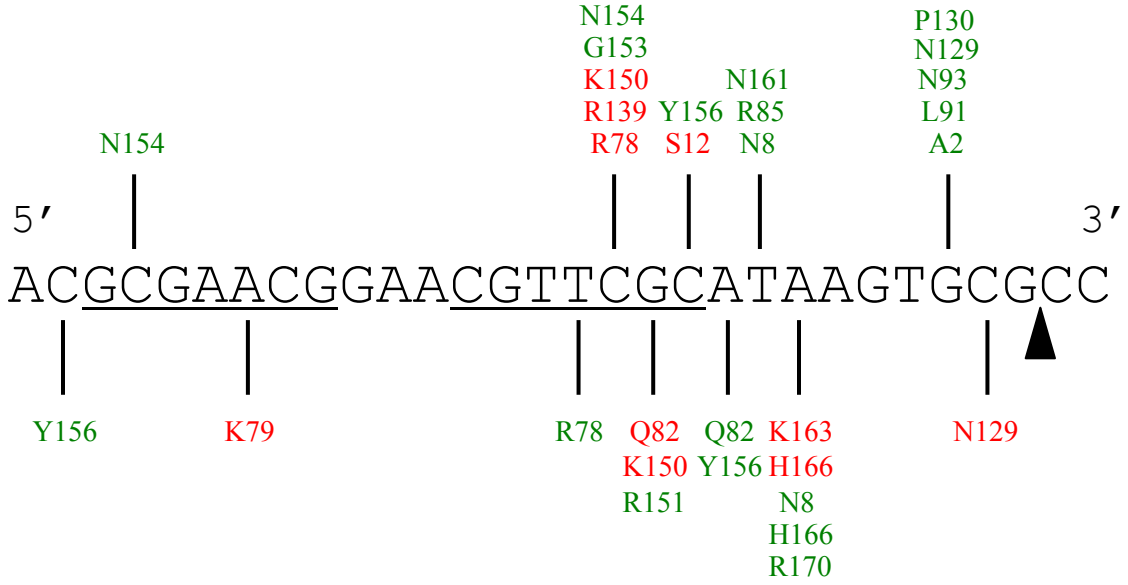


Figure 2.5.B Protein Residue Interactions with the DNA Bases

The DNA substrate is shown in the 5' to 3' direction with the hairpin section underlined and the *nic* site indicated by the triangle. Protein residues colored red indicate interactions via the DNA phosphate backbone. Protein residues colored green indicate interactions via the DNA bases.

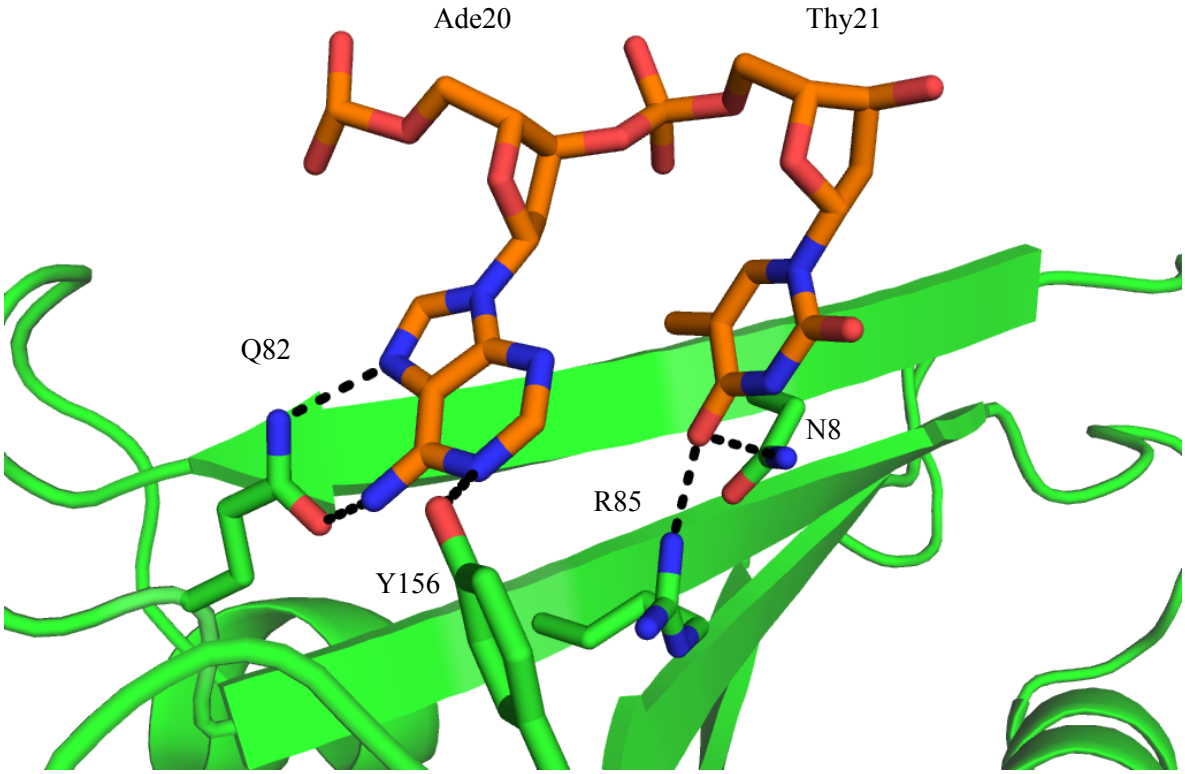


Figure 2.6 Single Strand DNA Base Specific Interactions

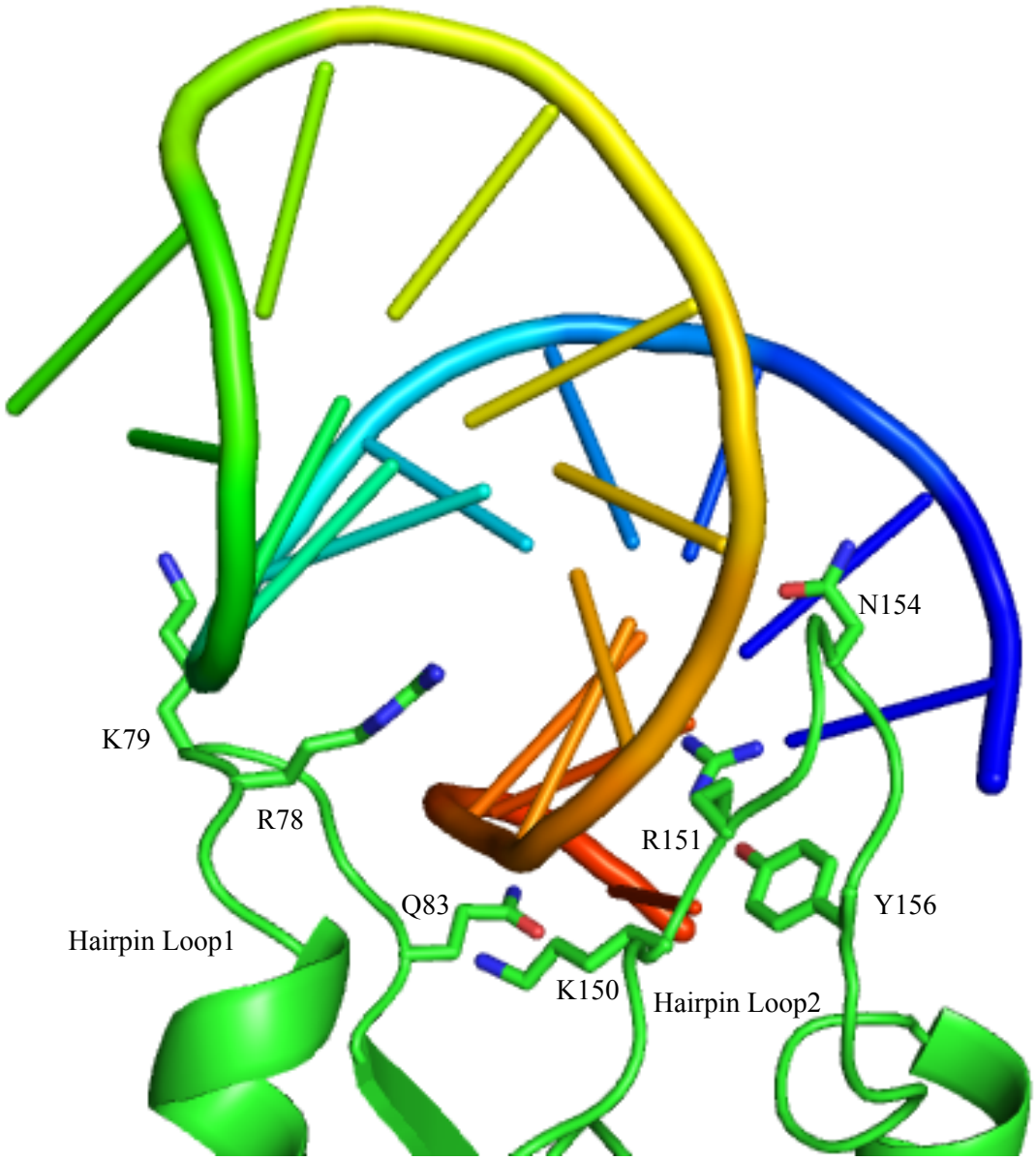


Figure 2.7 NES Hairpin Loop Interaction with the DNA Hairpin

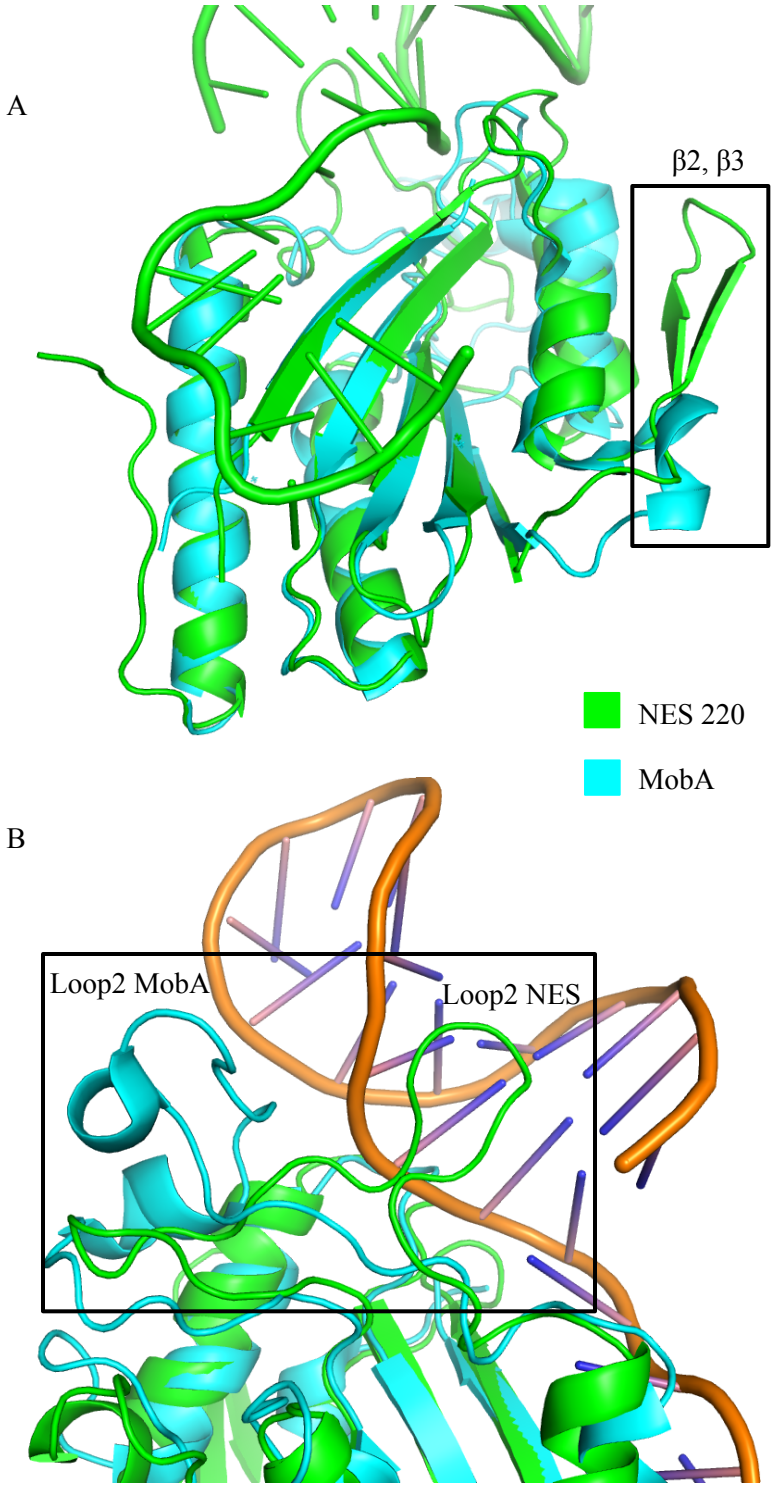


Figure 2.8 Comparison of NES 220 to MobA

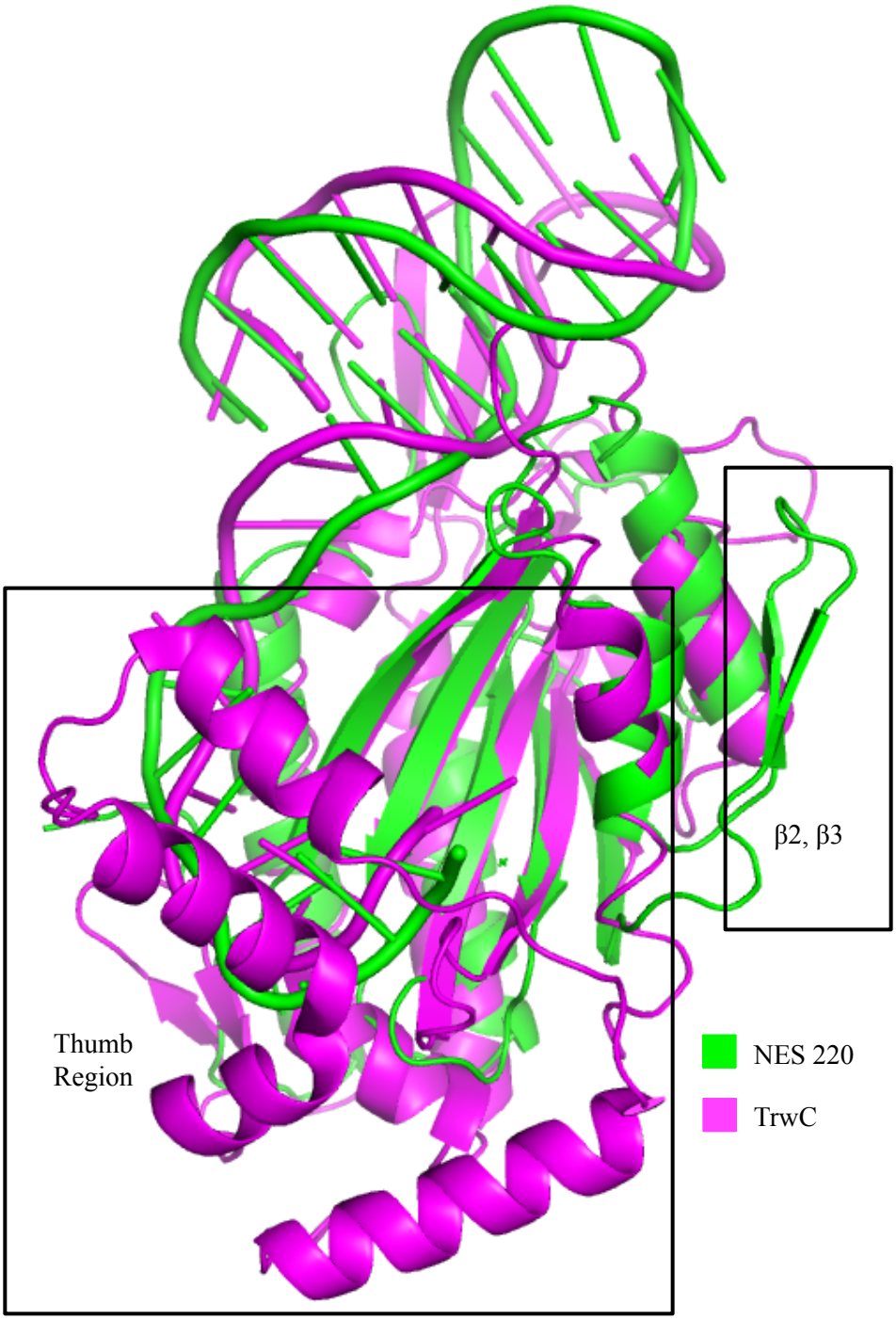


Figure 2.9 Comparison of NES 220 to TrwC

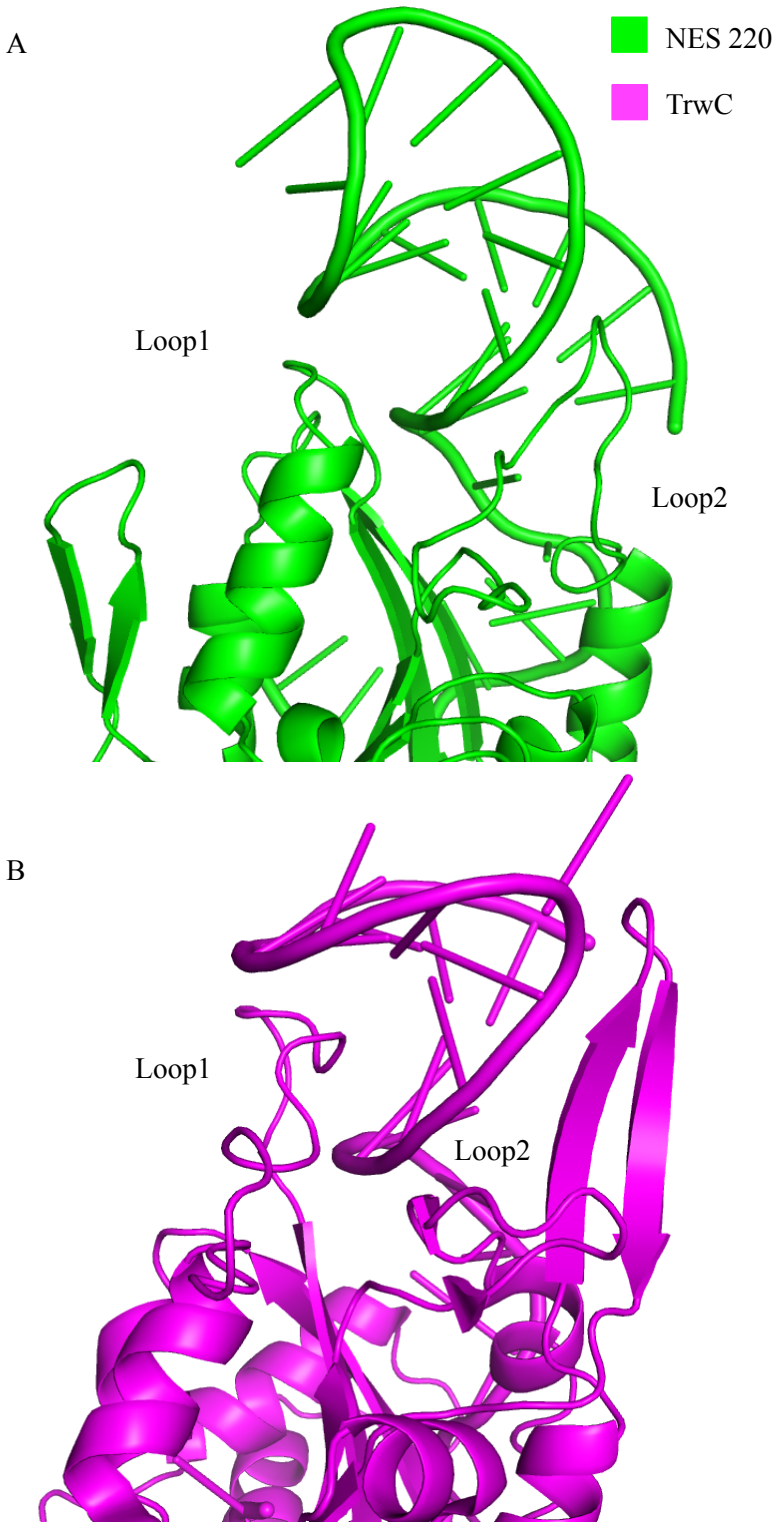


Figure 2.10 Comparison of NES 220 to TrwC – DNA Hairpin Regions

2.9 References

1. Boer, R. *et al.* Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J Mol Biol* **358**, 857-869, doi:10.1016/j.jmb.2006.02.018 (2006).

2. Datta, S., Larkin, C. & Schildbach, J. F. Structural insights into single-stranded DNA binding and cleavage by F factor TraI. *Structure* **11**, 1369-1379 (2003).

3. Larkin, C., Datta, S., Nezami, A., Dohm, J. A. & Schildbach, J. F. Crystallization and preliminary X-ray characterization of the relaxase domain of F factor TraI. *Acta Crystallogr D Biol Crystallogr* **59**, 1514-1516 (2003).

4. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W. & Redinbo, M. R. Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proc Natl Acad Sci U S A* **104**, 12282-12287 (2007).

5. Monzingo, A. F., Ozburn, A., Xia, S., Meyer, R. J. & Robertus, J. D. The structure of the minimal relaxase domain of MobA at 2.1 A resolution. *J Mol Biol* **366**, 165-178 (2007).

6. Nash, R. P., Habibi, S., Cheng, Y., Lujan, S. A. & Redinbo, M. R. The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res* **38**, 5929-5943, doi:10.1093/nar/gkq303 (2010).

7. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J. & Kelley, L. A. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**, 611-625 (2008).

8. Stols, L. *et al.* A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* **25**, 8-15, doi:10.1006/prep.2001.1603 (2002).

9. Shen, A. *et al.* Simplified, enhanced protein purification using an inducible, autoprocessing enzyme tag. *PLoS One* **4**, e8119, doi:10.1371/journal.pone.0008119 (2009).

10. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**, 307-326 (1997).

11. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

12. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

13. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780-2781, doi:10.1093/bioinformatics/btn507 (2008).

14. Holm, L., Kaariainen, S., Wilton, C. & Plewczynski, D. Using Dali for structural comparison of proteins. *Curr Protoc Bioinformatics* **Chapter 5**, Unit 5 5, doi:10.1002/0471250953.bi0505s14 (2006).

15. Holm, L. & Park, J. DaliLite workbench for protein structure comparison. *Bioinformatics* **16**, 566-567 (2000).

16. Guasch, A. *et al.* Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Biol* **10**, 1002-1010, doi:10.1038/nsb1017 (2003).

Chapter 3

DNA Binding and Nicking Analysis of the Relaxase Domain of NES

3.1 Introduction to Relaxase Function

As mentioned previously, relaxases are an integral part of the relaxosome allowing for the initiation and termination of CPT [1]. The relaxosome is the set of proteins that are responsible for processing the DNA, joining with the coupling protein, and allowing the transfer of the DNA into the recipient cell [2]. In order for consistent CPT the relaxase must be able to recognize the DNA substrate at the *oriT* and cleave exclusively at the *nic* site. This requires a high affinity and sequence specific interaction to ensure the cleavage reaction occurs at the correct site, every time, during CPT. All relaxases studied to date have high affinity, sequence specific interactions with their DNA *oriT* regions except for TraI of pCU1 [3]. Through investigating the DNA binding characteristics and determining the specific base interactions with the protein we hoped to gain two main pieces of information. The first would be to determine the optimal substrate for use in X-ray crystallization trials. This would improve our chances of obtaining the structure of NES 220 in complex with DNA. The rational being that if we can find the highest affinity, shortest ligand we will have the best chance of ensuring the enzyme is in complex with DNA in the crystal. The second is to learn more about how NES, a gram-positive plasmid relaxase, recognizes and binds to the DNA. This information will be useful by aiding the field in understanding the subtle differences found between relaxase' recognition of DNA and to provide a map of key DNA bases that can be explored as novel drug targets. Through understanding which DNA bases interact with NES, we can focus our efforts on developing strategies or compounds that can impede the ability of NES to bind DNA, and hopefully disrupt CPT.

Another long-standing question in the field is to determine which divalent cation is actually used *in vivo* for the transesterifcation reaction. Although we do not answer that question here, we have analyzed the divalent cation preference in vitro, determined the actual *nic* site for NES through cleavage assays, and looked at the effects of mutations and metals on the cleavage reaction. Mutant forms of NES designed from the structure, as well as mutant DNA oligos, were analyzed to shed light on the specific role certain segments of the enzyme play in the reaction. Taken together, these data will allow us to understand the effects of mutations and provide a system to test novel inhibitor efficacy.

3.2 Construct and Oligonucleotide Design, Mutagenesis, Expression and Purification

The relaxase domain of NES, NES 220, was cloned and purified as previously mentioned in Chapter 2. Briefly, all constructs in the following work were cloned and expressed into the CPD-Lasso expression vector using the *NdeI* (5') and *BamHI* (3') restriction endonuclease sites. All mutations were made by site directed mutagenesis and sequences were confirmed by Eton Biosciences, Inc. The resultant cloned proteins or mutant variants were purified in the same fashion as used in crystallization trials. After HisTrap and size exclusion chromatography the pure fractions were combined, flash frozen in liquid nitrogen, and frozen in stocks stored at -80 °C for use in experiments.

Integrated DNA Technologies (IDT) manufactured oligonucleotides for binding and *nic* assays. The oligos were designed based on the sequence for pLW1043 (GenBank: AE017171) around the *oriT* that was identified based on sequence comparisons to pSK41 and pGO1 [4,5]. Oligos used in direct binding experiments were labeled with a 5' 6-FAM fluorescein and HPLC purified for fluorescence anisotropy experiments. All oligos used in binding or *nic* assays were diluted to 100 mM in oligo annealing buffer (50 mM NaCl, 0.05 mM EDTA, 10 mM Tris pH 7.5). The oligos were then annealed by bringing them to 95 °C for 10 min. and lowering the temperature by 1 °C each min. until 4 °C was reached using a thermocylcer (Eppendorf). The annealed oligos were then diluted to 10 mM working stocks in the oligo annealing buffer.

Annealing was done to ensure the proper formation of the hairpin duplex present on some of the oligos.

3.3 DNA Binding Data Introduction

We wanted to learn more about the specific DNA substrate that is recognized by NES during CPT in order to determine which bases are responsible for recognition, what DNA structural elements are required, and what the minimal length of the substrate is necessary. For other relaxases studied, it has previously been shown that these enzymes bind a portion of the transferred strand (T-strand) at a specific location called the *oriT*. They will then cleave the substrate at the specific *nic* site to begin transfer [3,6-8]. Obtaining this information for NES aided us in identifying the optimal crystallization substrate as well as learning more about the recognition characteristics of the enzyme. In order to truly observe the binding characteristics of the enzyme a catalytically dead mutant (Y25F) was used in the experiments. This would ensure that no secondary effects due to cleavage and religation of the substrate were interfering with the measurements. Once the binding affinity for the initial oligo was determined, competition assays were employed to probe the effects of shifting the sequence base by base, shortening the sequence, or removing the hairpin. Upon obtaining the crystal structure, mutations were made to the enzyme based on the structural analysis to investigate critical residue-DNA base interactions.

3.3.1 Methodology

DNA Binding Studies – Direct Binding. In order to determine if in fact our relaxase bound in a specific, high affinity fashion we employed fluorescence anisotropy to measure the binding interaction. NES 220 and NES 220 Y25F were used to determine the dissociation constant, K_D, for the initial probe used. The annealed DNA oligo selected for the initial experiments had the sequence 5' 6-FAM-ACGCGAACGGAACGTTCGCATAAGTGCGCC 3' which incorporated the full inverted repeat and the 5' side of hypothesized *nic* site (terminal C base). For each assay, purified protein was diluted to 2X maximal assay concentration into a buffer of 150 mM NaCl, 0.1 mg/mL bovine serum albumin (BSA), 5 mM Mg^{2+} Acetate, 25 mM Tris Acetate pH 7.5. The

protein stock was then serial diluted into 40 µL of the final reaction buffer 100 mM NaCl, 0.1 mg/mL BSA, 5 mM Mg^{2+} Acetate, 25 mM Tris Acetate pH 7.5. This was done in a 384-well black assay plate (Costar) to generate 16 concentrations of protein, 8 dilutions from 500 nM protein and 8 from 400 nM protein. The DNA probe was diluted to 5X final assay concentration into the reaction buffer. 10 µL of the DNA stock solution was added to the 40 µL resulting in a final concentration of DNA probe at 50 nM, 1X protein, and a total volume of 50 µL in each well. The no-protein control well was generated by adding 10 µL of DNA to 40 µL of reaction buffer.

Fluorescence anisotropy (FA) was employed to measure the dissociation constant for the DNA oligo. The FA of the fluorescently-labeled substrate at 520 nM was measured following excitation at 485 nM using a PHERAstar plate reader (BMG Labtech) in a T format. Measurements were made in triplicate and the normalized data was plotted as average FA vs. protein concentration. Equation 3.1 was used to fit the data using Graphpad PRISM v5.03 (Graphpad, 2010) and calculate the K_D for the substrate.

Equation 3.1:

$$f = min + (max - min)\frac{\left\{(T + x + K) - [(-T - x - K)^2 - 4Tx]^{\frac{1}{2}}\right\}}{2T}$$

Where *f*, average FA signal detected; *T*, total DNA concentration (set to 50 nM); *x*, total protein concentration; *K*, K_D; *min*, average FA signal of no protein control; and *max*, average FA signal of sample at saturating concentration of protein. A single binding site was assumed and all experiments were run in triplicate with standard error reported for each measurement.

DNA Binding Studies – Competition. To determine the effect of altering the DNA substrate sequence DNA competition assays were employed. Again, fluorescence anisotropy was measured, this time to determine the effect of unlabeled competitor oligos on the dissociation constant of the substrate used in the direct binding studies. The assay was setup in a similar fashion to the direct binding studies except the protein and labeled DNA oligo were held constant while the unlabeled competitive oligo was titrated. The same 384 well assay plate was used in the

experiment. The final reaction buffer was the same (100 mM NaCl, 0.1 mg/mL BSA, 5 mM Mg^{2+} Acetate, 25 mM Tris Acetate pH 7.5) with the addition of the labeled DNA oligo at 62.5 nM. The labeled DNA oligo was also added to the unlabeled stock of competitor oligo at 125 nM so that the concentrations were not affected during titration. The maximal 2X concentration for the competitor oligo was 20 µM. As in the direct binding experiments, 40 µL of the final reaction buffer was placed in each of the 16 wells. For the competition experiments, 40 µL of the competitor olio was sequentially titrated down to the final concentration. The final well contained the no-protein control that only contained reaction buffer, labeled, and un-labeled competitor oligo. Lastly, 10 µL of protein at 5X reaction concentration was added to the wells. This resulted in an assay with 100 nM protein, 50 nM labeled DNA oligo, and a titration of competitor oligo concentrations from 10 µM to 0. Reductions in anisotropy were measured using the PHERAstar plate reader (BMG Labtech) in a T format. Measurements were made in triplicate and the normalized data was plotted as average FA vs. log protein concentration. Equation 3.2 was used to fit the data using Graphpad PRISM v5.03 (Graphpad, 2010) and calculate the IC_{50} for the competitor.

Equation 3.2

$$f = min + (max - min)/\left[1 + 10^{(x - LogIC_{50})}\right]$$

Where *f*, average FA signal detected; *x*, competitive oligo concentration; $LogIC_{50}$, the concentration at which the maximal signal has decreased by half; *min*, average FA signal of no protein control; and *max*, average FA signal of sample at saturating concentration of protein. Once the IC_{50} has been determined it can be used in the following equation to determine the dissociation constant for the inhibitor (K_i) using equation 3.3.

Equation 3.3

$$K_i = [I]_{50}/\left(\frac{[L]_{50}}{K_d} + \frac{[P]_0}{(K_d + 1)}\right)$$

Where K_i, dissociation constant for the inhibitor, $[I]_{50}$, concentration of free inhibitor at 50% inhibition, $[L]_{50}$, concentration of free labeled ligand at 50% inhibition, $[P]_0$, concentration of free protein at 0% inhibition, and K_d, dissociation constant of the protein-ligand complex [9,10]. All experiments were run in triplicate with standard error reported.

3.3.2 DNA Binding Sequence and Mutation Results

The first step in investigating the sequence requirements for NES to bind the DNA substrate was to confirm it specifically bound the *oriT* as would be expected for a relaxase. Fluorescence anisotropy was employed to determine the dissociation constant for the selected substrate. The initial DNA substrate was chosen based on sequence alignment to the pGO1 and pSK41 plasmids that had known *oriT* sites and were identical to plw1043 in this region [11]. The substrate contained a 5' 6-FAM fluophore attached to the inverted repeat (IR) hairpin, the single stranded portion of the DNA, and the 5' cytosine of the hypothesized *nic* site (Table 3.1, Oligo 1). The *nic* site for NES was currently unknown and hypothesized to cut at that specific site based on sequence alignments to plasmids with the same sequence at the *oriT*. A catalytically inactive mutant, Y25F, of NES was used to ensure only binding interactions were measured. Five individual experiments run in triplicate were used to determine the initial affinity for the substrate. The data presented in Figure 3.1.A shows there was a high affinity interaction with a K_D of 18.24 nM ± 2.71. To confirm that this interaction was specific to the substrate used, a control experiment was run using a randomly generated sequence of the same length without the hairpin as well as a random sequence that contained an IR hairpin with a different sequence (Figure 3.1.A.B). The results demonstrated that the recognition of the substrate by NES was due to the specific sequence and the length or hairpin structure are not enough for the enzyme to bind the substrate.

Next we wanted to determine the specific part of the plasmid sequence that is required for efficient binding and recognition and if there were any sites upstream or downstream of the sequence initially used. To do this we employed a competition assay using our reference substrate

as the fluorescent ligand and ordered a set of oligos that maintained the same length but varied in sequence. The sequence was started upstream of the hairpin and sequentially shifted 3' by a single base until past the *nic* site (Table 3.1, Oligos 2-9). By determining the K_i for the competitors we would then be able to tell if there were other sites on the sequence that were recognized better by NES. Figure 3.2 shows representative binding curves for a competitor that has a weak affinity for NES (A) compared to a substrate that is an efficient competitor with the fluorescent probe (B). As seen in previous work with other relaxases, we found that the presence of the hairpin, the single stranded section, and the *nic* site were required for binding. Oligo 2 in Table 3.1 has a K_i ~17 fold weaker than our reference substrate. Any oligos shifted further in the 5' direction had even weaker affinities until no binding was detected. Oligos 3-6 have an increase in affinity as the sequence is shifted in the 3' direction. This demonstrates that as we maintain the hairpin structure and extend the length of the single stranded DNA 3' of the hairpin, the enzyme is interacting with those bases preferentially. Oligo 6 has the highest affinity for the NES relaxase and it incorporates the full-length hairpin, the single stranded DNA up to the *nic* site, the hypothesized *nic* site, and 3' thymine after the *nic* site. As the sequence continues to shift 3' in oligos 7-9 it can be seen that the hairpin is now affected. By shortening and finally disrupting the hairpin the affinity decreases. Any oligos shifted further than 9 showed no significant DNA binding.

Next we were curious to know if a shorter DNA oligo would be suitable as a substrate for NES. We decided to truncate the competitor oligos from both the 3' and 5' ends of the DNA. These results are presented in Table 3.1 (Oligos 10-16). From these results it is clear that truncating the construct from the 5' side, and thus disrupting the hairpin, greatly reduces the enzyme's ability to bind the substrate as seen in oligos 10-12. Interestingly, truncation from the 3' side of the substrate did not initially have much of an effect. NES still has a specific and high affinity interaction with oligo 13. This result will be discussed further in the following results section. The subsequent truncations from the 3' end result in a decreased binding affinity. These

data confirm that a structurally intact hairpin and a single stretch of DNA that is at least 9 bases long must be present for NES to efficiently recognize the substrate. From these data, we chose to select oligo 4 for use in crystallization trials because it has a high affinity (8 nM), had some protection bases for the 6-FAM 5' to the hairpin, and contained half of the hypothesized *nic* site.

Using the same competition assay we designed a set of 93 oligonucleotide competitors that contained DNA base mutations to further explore the specific bases required for the interaction between NES and the DNA. Oligo 4 from Table 3.1 was used as the template and each bases was mutated to one of the other 3 DNA bases. When the hairpin was involved the bases on each side of the duplex were mutated to ensure the hairpin could still form. A mutant hairpin was also designed that maintained the 3' sequence of the hairpin but disrupted the 5' sequence so that hairpin formation was not possible. This was done to determine if the base interactions were specific to one side of the hairpin or the structure of the hairpin was required for binding. Another set of mutations was made on a longer template that extended 6 bases 3' of the *nic* site. Mutations were made in this area to see if the DNA following the *nic* site played a significant role in binding.

For simplicity only a select set of the data have been presented in Table 3.1 because not all of the mutations had interesting effects on the binding characteristics. As discussed previously in Chapter 2 (Figure 2.5.B) there are specific DNA bases that make interactions with NES. Oligo 17 is a hairpin mutant that disrupts the 5' side of the sequence with random bases that will not allow duplex formation. The K_i for this construct is ~33 fold weaker than for the *oriT* reference probe confirming again that the sequence and secondary structure of the *oriT* are required for efficient binding.

The biochemical data was collected prior to obtaining the structure of NES in complex with DNA so no inferences could be made as to which bases would be key in the interactions. By looking at the resultant K_i values that were determined for each mutant competitor there was a clear pattern that emerged. Firstly, there was no significant effect on the enzymes ability to bind the DNA when mutations were made to the terminal 8 bases on the 3' end of the construct. This

agreed well with Figure 2.5.B that was determined from the structure. As seen in that figure, there are almost no interactions with the DNA bases on this portion of the DNA except for Gua26. We were surprised when we looked at the structure and saw the confirmation of the base at position Gua26 as it is protruding directly into an enzyme pocket and has numerous base specific contacts with NES (Figure 3.3). However, since guanine is the bulkiest base being a purine it is possible to imagine that substitutions into this pocket may not maintain the hydrogen-bonding network but will not clash with the enzyme residues. Furthermore, given that the base will still be able to fit into this cleft it is foreseeable that all of the other correct contacts will be maintained along the DNA substrate leading to efficient binding of these mutant oligos. An abasic site was designed into the oligo at this position and the effects will be discussed in the cleavage assay results section below. However, the biochemical data seem to agree with the structural analysis implying that the eight 3' bases only contribute to binding through being present and not being dependent on the specific base sequence. Another region on the DNA substrate that did not affect the enzymes ability to bind the substrate was the stem loop region present on the top of the DNA. When bases were inserted into this region (Table 3.1, Oligo 30) or mutation to the three unpaired bases were made (data not shown) the enzyme was able to bind with high affinity.

Mutations made 5' of the Ade23 began to have significant effects on the enzymes ability to bind the DNA. When the bases from Cyt17 to Ade22 were mutated almost every mutation decreased the enzymes ability to bind by at least 4-fold with some being much as ~ 235-fold worse. The lone exception in this region was when mutations were made to Gua3 and Cyt19 at the base of the hairpin. As seen in Figure 2.5.B there are two interactions made with the protein but one of them is a main-chain hydrogen bond. The bases around Cyt19 have extensive interactions with the enzyme, including multiple base specific interactions, and may compensate for the loss of this one interaction from the mutation.

Since the rest of the mutations significantly affect DNA binding by NES, oligos 18-23 were chosen for Table 3.1 because of the interesting differences in binding affinity for the

mutations within the DNA substrate. The mutations in the sequence at WT Thy21 to Ade21, Cyt21, and Gua21 and WT Ade20 to Thy20, Cyt20, and Gua20 were highlighted because their binding affinities did follow our initial thoughts that purine to purine substitutions would be better tolerated than purine to pyrimidine mutations. For example, one might expect that a mutation from Ade to Gua might be tolerated easier than a mutation from Cyt to Ade based on the structure of the purine and the pyrimidine bases. However, the K_i for the mutant substrates did not reflect this in all cases. Oligos 18-23 contain mutations in Ade20 and Thy21. These bases interact with four polar residues of NES and no contacts are made to the phosphate backbone (Figure 3.4). For oligo 18 we see that the binding affinity decreases ~4 fold due to the mutation. However, for oligo 19 there is a ~7 fold decrease in binding affinity. Figure 3.5.B,C help to explain this finding based on the structure. Even though the Thy21Ade mutation is a pyrimidine to a purine one can see from the structural analysis that in both cases the hydrogen bonds are disrupted. Figure 3.5.B shows the position of the Ade based on a model created from the difference maps. Ade21 is out of position to maintain both hydrogen bonds with R85 and N8 as seen in Figure 3.5.A. For Cyt21, Figure 3.5.C, it is clear that the structure of the base will allow for the base to be positioned in the same way as Thy21 but the primary amine on the cytosine is not a hydrogen acceptor resulting in a disrupted network. This would cause the base to not fit as well as Ade21 which is simply out of place. Oligo 20 had the most severe effect on the binding of the substrate of NES. Similarly to Ade21, Gua21 is out of position but also contains a primary amine that will have a steric clash with the loop that is above the base recognition site. From the binding data this clash has the most significant effect on the enzymes ability to bind to the substrate.

Oligos 21-23 represent mutations made to Ade20 that also had interesting binding affinities. Figure 3.6 compares the WT base position from the structure to the modeled mutations. Thy20 and Cyt20 have dissociation constants that are ~5 fold weaker than for the native substrate. It is clear from 3.6.B,C that the structure of the pyrimidine base is not sufficient for them to

maintain the hydrogen bond network between Y156 and Q82. For the mutant Gua20, which again had the most significant effect on binding, the structure would allow for the base to placed in the same position but the chemical composition does not allow for the maintenance of the WT hydrogen bond network as seen in Figure 3.6.A. Taken together, these data demonstrate that the bases recognized by NES on the single stretch of DNA, 3' of the hairpin, have specific interactions and mutations from purine to pyrimidine as well as pyrimidine to pyrimidine have significant effects on the binding affinity.

The next set of oligos involves mutations made to the hairpin structure that change the specific base but maintain the secondary structure (Table 3.1, Oligos 24-29). The regions of NES primarily responsible for binding the hairpin portion of the DNA substrate have been termed Hairpin Loop1 (HL1) and Hairpin Loop2 (HL2). The residues present on these loops primarily have interactions with the deoxyribose-phosphate backbone of the DNA, but there are a few base specific hydrogen bonds as well. Figure 3.7 highlights the interactions made between these two loops and the DNA bases that were mutated in oligos 24-29. The interactions are a combination of side chain and main chain hydrogen bonds that form a clamp from either side of the DNA. The interesting observations made when analyzing the data were that mutations in Cyt17 did not significantly affect substrate binding where mutations in the next base, Gua18, severely decreased the enzymes ability to bind the DNA.

For oligos 24-26 the WT Gua18 and Cyt4 were mutated and the dissociation constants were calculated. Mutations to this base resulted in affinities that were ~10 – 215 fold weaker than the WT interaction. Figure 3.8.A focuses on the interaction with Gua18 and highlights the hydrogen bonds made through the side chain of R151 with the guanine base as well as the hydrogen bonds with the phosphate made from the side chain of K150 and the main chain interactions with S81 and Q82. As previously, the mutant bases were modeled into the density maps for comparison. This time the two largest effects come from mutating a purine to a pyrimidine. Both the Cyt18 and Thy18 mutations (oligos 24,25) position the base too close to

R151 resulting in a steric clash. This would either push the DNA base out of the binding pocket or require R151 and HL2 to move out of the space altering the hydrogen bond network for the other bases in the substrate. The Ade18 mutation does not result in a steric clash but it does disrupt the hydrogen bond with R151 by replacing the carbonyl group from guanine with a primary amine resulting in a ~10 decrease in binding affinity.

Structural analysis of the Cyt17 mutation provided an explanation for the lack of effect that mutations had on this position. Figure 3.9.A depicts the interactions made with HL1 and HL2 with the WT Cyt17 base. Cyt17 hydrogen bond network primarily consists of interactions made with the phosphate backbone and the deoxyribose moiety. G153 does make a base specific interaction with Cyt17, but the majority of the recognition comes from the 6 backbone interactions. The mutation in oligo 27 has the largest effect due to the fact that the carbonyl group on guanine will require some movement by G153 to fit the base in the enzyme clamp (Figure 3.9.B). The mutation in oligo 28 only results in the loss of the hydrogen bond with G153, thus the binding remains at WT levels. Mutating to Ade18 in oligo 29 again results in the loss of the bond with G153 and may have subtle effects on the hydrogen bond network due to the size difference (Figure 3.9.D) that leads to a decrease in binding affinity.

3.4 Divalent Cation Analysis and DNA Cleavage Data Introduction

Once the relaxase has bound the *oriT* it must then catalyze the transesterification reaction. This reaction occurs at a specific *nic* site on the DNA that is aligned in the active site so that scissile phosphate can be attacked by the catalytic tyrosine [6,12,13]. For the reaction to proceed the negatively charged phosphate must act as an electrophile. The HUH(+H) motif coordinates a divalent cation that is positioned next to the scissile phosphate. This positive charge on the cation will serve to neutralize the negatively charged phosphate, allowing to act as an electrophile [6,12,14-16]. The third requirement for successful cleavage is that the tyrosine must act as a nucleophile (Figure 1.5). At physiological pH the tyrosine hydroxyl is protonated because the pK_a of tyrosine is 10. Deprotonation of the hydroxyl results from a neighboring residue or solvent molecule

acting as a base. Another possibility is that the local pH of the active site is higher than physiological pH (7-7.5).

The structure of NES 220 in complex with DNA reveals that in the active site is arranged in the typical fashion. In our structure, the HUH(+H) motif was seen to be coordinating a Ni^{2+} metal in the active site. The Y25F residue was positioned on α1 in a manner that it could be directed to attack the scissile phosphate (Figure 2.1). Unfortunately, our structure had a crystallographic defect that pulled the DNA out of the active site (Figure 2.4). However, it can easily be imagined that the DNA would lie on top of the active site and proceed along α1 and finally between α3 and the β2,3 loop (Figure 2.1). Similarly to the relaxase MobA of plasmid R1162, it is not possible to find an amino acid that is acting like a base to abstract the hydrogen from the carboxyl on Y25 [17]. We hypothesize that an ordered water molecule is acting as a specific base to accomplish this goal (Figure 1.5).

In the following sections we will outline the assays and the approach used to identify the true *nic* site for NES. We also examine the metal preference for the enzyme *in vivo* and determine the enzyme activity towards the substrate when in complex with each divalent cation. Following those experiments the effects of mutations on NES 220, determined from structural analysis, are also presented and discussed. Finally, we will propose a new strategy for inhibiting relaxases through the use of minor groove binding compounds such as the Hoechst 33528 stain.

3.4.1 Methodology

Electrophoretic Mobility Shift Assays (EMSA). To determine what the true *nic* site in the *oriT* sequence was we used EMSA assays to visualize the equilibrium products of the cleavage reaction. Commercially synthesized oriT DNA substrates (IDT) were ordered with 5' 6-FAM fluorescein molecules attached. Each oligo was resuspended and annealed as detailed in section 3.2 and the sequences used are listed in Table 3.2.

Each reaction contained 8 μL of EMSA buffer (50 mM NaCl, 20 mM Tris pH 7.4, 0.02% v/v sodium azide), 1 μL of 50 μM NES 220 in sizing buffer, and 1 μL of 10 μM 5' 6-FAM

labeled DNA substrate. The reaction was incubated at 37° C for 1 hr and quenched with the addition of 10 µL 2X quenching solution (0.01% xylene cyanol, 85% formamide, 20 mM EDTA, 2X TAE, 0.2% SDS). The 20 µL were run through a denaturing 16% polyacrylamide gel (40 mL 16% acrylamide gel stock (8M urea, 16% polyacrylamide/bisacrylamide, 1X TBE), 400 µL 10% ammonium persulfate (APS), 40 µL tetramethylethylenediamine (TEMED)) in 1X TBE running buffer to separate the substrate and products. The fluorescent oligos were visualized using a VersaDoc Imaging System, 4400 MP (BioRad) and the accompanying Quantity One software (BioRad). Band intensities were quantified using ImageJ 1.42 (Rasband, W.S., NIH 2008). Prior to quantification, standard background subtraction was performed on all gels. NES 220 was determined to be active the smaller product band was present in the gel. Activity was reported as a percent of the product band intensity divided by the product plus substrate intensities.

For the metal studies it was necessary to remove the natively bound metal through incubation of NES 220 in 500 µM EDTA overnight at 4° C. The metals of interested were then added in 10, 100, 1000 µM excess and activity was measured. The reaction conditions for the metal studies contained 7 µL of EMSA buffer (50 mM NaCl, 20 mM Tris pH 7.4, 0.02% v/v sodium azide), 1 µL of 50 µM NES 220 in sizing buffer incubated in 500 µM EDTA, 1 µL of 10 µM 5' 6-FAM labeled DNA substrate, and 1 µL of the metal in EMSA buffer at 10X the desired fold excess. The reaction was incubated at 37° C for 1 hr and quenched with the addition of 10 µL 2X quenching solution (0.01% xylene cyanol, 85% formamide, 20 mM EDTA, 2X TAE, 0.2% SDS). Gels were run in the same fashion as for the *nic* site studies.

Inductively Coupled Plasma Mass Spectrometry (ICP-MS). NES 220 and NES 220 H131A were analyzed by ICP-MS to determine the identity and concentration of metals present in the samples. Data were collected by the UNC-CH Chemistry Department Mass Spectrometry Core Facility on a Varian 820-MS (Palo Alto, CA). Protein was first digested for 4 hrs. in 70% nitric acid (HNO_3). Each sample was then diluted in 5.5 ml of 1.4% HNO_3. Data were collected as part per billion (ppb) and converted to µM. The original concentration of each metal in each

sample was then determined in Excel 2007 (Microsoft, 2011). Data were then presented as percent bound of the total concentration of protein.

3.4.2 DNA Cleavage, Metal Preference, and Mutant Analysis

The other relaxases studied to date (TraI F, TrwC R388, MobA R1162, and TraI pCU1) have been active against the DNA substrate in a metal dependent manner but the optimal metal is still under debate [12,13,16,18,19]. We set about to determine if NES 220 was able to cleave the product, the exact position of the *nic* site, and which metals allowed activity *in vitro*. To determine if NES was active against the *oriT* substrate oligo 1 in Table 3.2 was used. This oligo contained the hairpin, single stretch of DNA, and 8 DNA bases 3' of the *nic* site. The intial experiment was run in EMSA buffer with the DNA probe and NES. A ladder was made that consisted of the original substrate and a shorter probe that was designed to be the same length as the cleaved product (Table 3.2 – oligos 1,2; Figure 3.10). Controls were also run that contained NES incubated in 500 μM EDTA overnight, incubated NES plus Ni^{2+}, and a mutant form of NES that disrupted the metal binding motif NES H131A. The initial results in Figure 3.11.A were surprising because the product of the reaction did not have the correct predicted length based on previous alignment studies of NES from pGO1 [20]. However, we were able to determine that incubation with EDTA will inhibit activity, addition of a divalent cation will regain function, NES H131A disrupts the HUH motif and inhibits function (Figure 3.11.A), and NES Y25F is not capable of cleavage. In order to determine the exact *nic* site truncation oligos were ordered (Table 3.1, oligos 3,4) and used in the assay. Oligo 4 corresponded to the correct length for the products observed (Figure 3.11.B). This allowed us to determine that the true *nic* site for NES 220 of pLW1043 was between Gua28 and Cyt29. This finding also corroborated our DNA binding data in Table 3.1 because the affinity was improved by the presence of the true *nic* site in oligos 2,3, and 13.

Now that we had established to the true site at which the DNA was cleaved and that our assay was functional we wanted to investigate the enzyme's preference for different metals. To accomplish this goal NES 220 was incubated with EDTA to remove all bound metals. We chose

to investigate the effects of $NiCl_2$, $MgCl_2$, $MnCl_2$, $CaCl_2$, $ZnCl_2$, $FeCl_2$, $CoCl_2$, and $CuCl_2$ by adding them in 10 μM, 100 μM, or 1000 μM excess to the concentration of EDTA in the NES stocks. When NES 220 was incubated with the substrate and a natively bound metal it cleaved ~16% of the substrate with some variability (Figure 3.12). The metals that were able to support the cleavage reaction were Ni^{2+}, Mn^{2+}, Co^{2+}, and Cu^{2+} as seen in Figure 3.12. Mn^{2+} was the only metal that showed an increase in activity in a concentration dependent manner. However, Ni^{2+} had the most consistent activity at the concentrations tested. The ICP-MS data presented in Table 3.3 leads us to believe that NES is coordinating a nickel ion due to the purification method. When the NES H131A sample was analyzed no metals were found to be present in significant concentrations. Since Ni^{2+} concentrations in the cell cytoplasm are very low [21,22], we believe the presence of nickel does not indicate its use *in vivo*. Furthermore, Mn^{2+} was able to support cleavage (Figure 3.12), which correlates with what was seen in MobA [19], the most similar enzyme to NES. Co^{2+} and Cu^{2+} were able to support cleavage in NES but showed a decrease in activity at higher concentrations. Co^{2+} and Cu^{2+} are not used prevalently in biological systems and we conclude that these metals are most likely not used *in vivo*. From these data we believe NES most likely uses Mn^{2+} *in vivo* to help position the DNA into the active site and achieve cleavage of the DNA substrate.

Next we were curious to understand the effects of mutating some of the key bases identified in the previous section by observing their effect on the cleavage equilibrium. To do this oligos 5-10 from Table 3.2 were used in the EMSA assay. Oligo 5 contains mutations that will disrupt the hairpin. Oligos 6 and 7 contain mutations in Ade20 and Thy21. Oligos 8 and 9 contain mutations in the two hairpin bases discussed previously. Lastly, oligo 10 contains an abasic site at position Gua26. In Figure 3.13.A the percent of the fluorescent product formed in all cases was higher than WT substrate. For oligos 5-8 this was most likely due to the disruption in the ability of the enzyme to bind the DNA. The amounts of DNA seen in the EMSA assay represent the equilibrium that is in place between the cleavage and ligation activities of NES. For WT in

3.13.A this means that 19% of the shorter product is either free in solution or bound to NES. Oligos 5-8 have a lower affinity for NES and this will actually drive the equilibrium to produce more of the shorter product. This is due to the fact that once the enzyme cleaves the DNA substrate the remaining piece must be held in the enzyme in the correct position to allow for religation. Since the enzyme is not able to bind these mutants as well (Table 3.1) it will have a hard time maintaining the proper positioning, thus driving the equilibrium to contain more truncated product. Oligo 9 was interesting because it did not show a decrease in binding affinity but still had increased product formation. This could be due to the fact that even though the enzyme can bind the DNA, the optimal contacts are not present and still results in a shift of the equilibrium. An abasic site was inserted into oligo 10 to better understand the role of Gua26. In the structure this base appears to act as a post helping to aid the DNA in forming the U-turn into the active site (Figure 3.3). When the abasic site oligo was used in the EMSA assay the equilibrium was shifted even further towards product formation to ~60%. One can imagine that a loss of the base at position Gua26 would result in Cyt27 and Gua28 moving away from the active site. This is because there is only one interaction, phosphate backbone of Cyt27 with N129, for the six bases on the single strand of DNA. This would most likely result in large movements of these bases making it difficult to successfully perform the religation reaction because the 3' base of the *nic* site will not be in the proper position to attack the tyrosine phosphate bond.

Mutant forms of NES were purified as previously described. We wanted to investigate the effects of mutations near the active site of the enzyme to shed more light on the catalytic mechanism. In figure 3.14 K22, E86, and Y42 were mutated to alanine. This was to determine side chains of the residues were playing a role in the active site chemistry. A21 was mutated to a glutamic acid to see if a bulky residue would affect cleavage by disrupting the binding with the 3' side of the *nic* site. The results of the EMSA assays for these mutant enzymes are shown in Figure 3.13.B. The K22A mutation resulted in a ~1.5 fold increase in activity where the E86A mutation results in a ~3 fold decrease in activity. This result indicates that K22 is most likely not

playing a role in the catalytic mechanism and removal of this side chain may allow another water molecule to be ordered in the hydrogen bond network required to activate the carbonyl group on Y25. Furthermore, this provides evidence that E86 is involved in the reaction and most likely orders a water molecule to act as a specific base to allow the hydrogen abstraction (Figure 3.14). The Y42A mutation resulted in a loss of activity. This residue is found on a loop between β3 - β4 and may have enough mobility to position itself to act as a secondary catalytic tyrosine as seen in pCU1 and TrwC [23,24]. However, this is most likely not the case as the Y25F mutation in NES results in no cleavage product (Figure 3.11.A) leading to the assumption that no other residue can act like a nucleophile in this system. This would indicate that Y42 is most likely playing a role to bind and position the DNA 3' of the *nic* site. The A21E mutation was selected to determine if placing a larger residue in the groove where we believe the DNA 3' of the *nic* must be positioned by the enzyme. Similarly to Y42A, this mutation resulted in a decrease of activity. This results seems to implicate both Y42 and A21 as being in areas that are responsible for binding to the single portion of the DNA 3' of the *nic* site. The final mutant enzyme tested was the HL2 loop deletion. This loop was removed through site directed mutagenesis and purified in the same as previously described. The HL2 deletion enzyme was able to cleave the substrate DNA and resulted in the most drastic increase in activity. Again, this is due to favoring the product side of the equilibrium for the reaction. Given that these residues are so critical in recognizing and binding the hairpin the increased cleavage must be do to the fact that the enzyme will have random interactions with the DNA. On the chance that the substrate enters the active site properly and is cleaved, there is no hope of the enzyme maintaining the proper contacts with the hairpin portion of the substrate to allow for religation.

Using the information gained from the biochemical analysis we sought out to investigate novel methods to inhibit NES. The data clearly show that a few DNA bases are extremely important and extensive interactions are made with the DNA through HL1 and HL2. We sought to go a different route than inhibiting the active site as for F TraI [18] because this route proved to

be non-specific. The hairpin is long enough to form a small minor groove and it is located directly between HL1 and HL2. We sought out molecules that could bind in the minor groove and wanted to test the effect of these molecules. The first two compounds tried, berberine and netropsin, did not show an effect for DNA cleavage (data not shown). We next tried Hoechst 33258 (Figure 3.15.A) and found some interesting results. To test the effects of this molecule we incubated 20X DNA stocks with ~ 47, 23, and 12 fold excess of Hoechst. Figure 3.15.B shows that 47 fold excess of Hoechst was able to inhibit the enzymes function *in vitro*. The large excess is required due to the fact that the minor groove is short and does not contain many Ade or Thy which Hoechst prefers to bind to. Further work will need to be done to confirm the exact reason for the inhibition and to develop novel more specific inhibitors. However, this strategy provides hope to that it may be possible to create novel inhibitors towards the relaxase through disrupting an area of functional importance on the protein other than active site.

3.5 Conclusions

From the structural analysis combined with the biochemical we have been able to understand how the relaxase domain of NES is able to bind the DNA and cleave at the appropriate position. The direct binding studies revealed that NES 220 has an 18.24 nM K_D for the *oriT* DNA substrate. Furthermore we determined that for successful binding the hairpin secondary structure in the DNA must be present. The competition assays corroborated the structural evidence that key DNA bases are responsible for the highly sequence specific interaction. The hairpin has multiple bases and backbone interactions that allow the complex to maintain the correct orientation while there are few interactions along the single strand stretch of DNA leading up to the *nic* site. We were able to determine that one critical base along this stretch, Gua26, has multiple interactions with the enzyme and is required for the proper alignment of the *nic* site in the DNA. Deletion of this base drives the reaction equilibrium to product formation highlighting its critical role in position the substrate. Through the EMSA cleavage assays we were able to determine the true *nic* site between Gua28 and Cyt29 on the *oriT* sequence.

Metal assays revealed that Ni^{2+}, Mn^{2+}, Co^{2+}, and Cu^{2+} were able to facilitate the cleavage reaction. We determined that Mn^{2+} was most like the metal used *in vivo* based on these findings and the fact that the other biologically relevant metals, Mg^{2+}, Ca^{2+}, and Zn^{2+} showed no activity in our cleavage assay. Through using mutant DNA oligos to investigate cleavage we determined that efficient binding is necessary for proper religation of the substrate. This was due to the fact that mutant oligos that decreased the binding affinity resulted in greater product formation. When the active site of the enzyme was probed through structural mutations we were able to learn more about the catalytic mechanism employed by NES. Specifically, we determined that E86 is involved in the hydrogen bond network that allows for the deprotonating of the tyrosine carboxyl group. Mutations, Y42A and A21E, in the part of the enzyme that interact with the DNA substrate 3' of the *nic* site reduced activity. From this result we conclude that this region of the enzyme is responsible for maintaining the correct position of the DNA during the cleavage reaction. Finally, we used this information to aid us in a novel approach to inhibit relaxase enzymes. Through the use of a minor-groove binding molecule, Hoechst 33258, we were successful in halting the cleavage reactions. These data provide novel insights into how NES functions to allow successful CPT and a unique strategy to inhibit this vital process in bacteria. Through these experiments we hope to develop novel and specific compounds that can be used as antibiotics to help fight the spread of antibiotic resistance.

	5' Oligonucleotides 3'		
	Direct Binding	**K_D(nM)**	**R^2**
1	6FAM-CACGCGAACGGAACGTTCGCATAAGTGCGC*C*	18.24 ± 2.71	0.97
	Competition Binding	**K_i**	**R^2**
2	GCACGCGAACGGAACGTTCGCATAAGTGC**G**	313.0	0.99
3	CACGCGAACGGAACGTTCGCATAAGTGC**GC**	14.8	0.99
4	ACGCGAACGGAACGTTCGCATAAGTGC**GC***C*	8.1	0.99
5	CGCGAACGGAACGTTCGCATAAGTGC**GC***CC*	9.2	0.99
6	GCGAACGGAACGTTCGCATAAGTGC**GC***CC*T	4.0	0.99
7	CGAACGGAACGTTCGCATAAGTGC**GC***CC*TT	14.2	0.99
8	GAACGGAACGTTCGCATAAGTGC**GC***CC*TTA	16.8	0.99
9	AACGGAACGTTCGCATAAGTGC**GC***CC*TTAC	1,990.1	0.99
	Truncation Competition	**K_i(nM)**	**R^2**
10	GAACGGAACGTTCGCATAAGTGC**GC***CC*	1,085.8	0.99
11	CGGAACGTTCGCATAAGTGC**GC***CC*	940,533.4	0.99
12	GAACGTTCGCATAAGTGC**GC***CC*	N.B.	
13	CGCGAACGGAACGTTCGCATAAGTGC**G**	6.0	0.99
14	CGCGAACGGAACGTTCGCATAAGT	97.5	0.99
15	CGCGAACGGAACGTTCGCATA	838.0	0.99
16	CGCGAACGGAACGTTCGC	N.B.	
	Mutation Competition	**K_i**	
17	AC**CTATCGT**GAACGTTCGCATAAGTGCGC	564.6 nM	0.99
18	ACGCGAACGGAACGTTCGCA**A**AAGTGCGCC	86.6 nM	0.99
19	ACGCGAACGGAACGTTCGCA**C**AAGTGCGCC	139.2 nM	0.99
20	ACGCGAACGGAACGTTCGCA**G**AAGTGCGCC	531.4 nM	0.99
21	ACGCGAACGGAACGTTCGC**T**TAAGTGCGCC	80.3 nM	0.99
22	ACGCGAACGGAACGTTCGC**C**TAAGTGCGCC	97.9 nM	0.99
23	ACGCGAACGGAACGTTCGC**G**TAAGTGCGCC	448.5 nM	0.99
24	ACG**G**AACGGAACGTTC**C**CATAAGTGCGCC	4.28 µM	0.99
25	ACG**A**AACGGAACGTTC**T**CATAAGTGCGCC	1.93 µM	0.99
26	ACG**T**AACGGAACGTTC**A**CATAAGTGCGCC	0.21 µM	0.99
27	ACGC**C**AACGGAACGTT**G**GCATAAGTGCGCC	0.29 µM	0.99
28	ACGC**A**AACGGAACGTT**T**GCATAAGTGCGCC	0.02 µM	0.99
29	ACGC**T**AACGGAACGTT**A**GCATAAGTGCGCC	0.09 µM	0.99
30	ACGCGAACG**C**GAA**T**CCGTTCGCATAAGTGCGCC	0.04 µM	0.99

Table 3.1 DNA Competition Oligonucleotides and Binding Affinities

This is a representative list of the DNA oligonucleotides used in the binding experiments. A zigzag underline corresponds to the bases involved in forming the hairpin. For oligos 1-13 bold

bases correspond to the true *nic* site where italicized bases represent the hypothesized *nic* site. Bold bases are mutations made to the WT sequence for oligos 17-30. N.B. stands for no binding observed.

	5' EMSA DNA Oligo Substrate Sequences 3'
1	6FAM-ACGCGAACGGAACGTTCGCATAAGTGC**GC***CC*CTTACGG
2	6FAM-ACGCGAACGGAACGTTCGCATAAGTGC**GC***C*
3	6FAM-ACGCGAACGGAACGTTCGCATAAGTGC**GC**
4	6FAM-ACGCGAACGGAACGTTCGCATAAGTGC**G**
5	6FAM-AC***CTATCGT***GAACGTTCGCATAAGTGC**GC***CC*CTTACGG
6	6FAM-ACGCGAACGGAACGTTCGCA***G***AAGTGC**GC***CC*CTTACGG
7	6FAM-ACGCGAACGGAACGTTCGC***G***TAAGTGC**GC***CC*CTTACGG
8	6FAM-ACG***G***GAACGGAACGTTC***C***CATAAGTGC**GC***CC*CTTACGG
9	6FAM-ACGC***A***AACGGAACGTT***T***GCATAAGTGC**GC***CC*CTTACGG
10	6FAM-ACGCGAACGGAACGTTCGCATAAGT C**GC***CC*CTTACGG

Table 3.2 DNA EMSA Oligonucleotides

Labeled fluorescent DNA oligonucleotids used in the EMSA cleavage assays. A zigzag underline corresponds to the DNA hairpin. Italicized bases represent hypothesized *nic* site. Bold bases represent the true *nic* site. Bold and italicized bases represent mutations made to the WT sequence.

	NES 220 Y25F	NES 220 Y25F, H131A
Ni	20.68 %	0.16%
Mg	0.75 %	1.40 %
Mn	0.00 %	0.01 %
Ca	0.24%	0.44 %
Zn	2.85 %	0.11 %
Fe	0.32 %	0.39 %

Table 3.3 ICP-MS Results

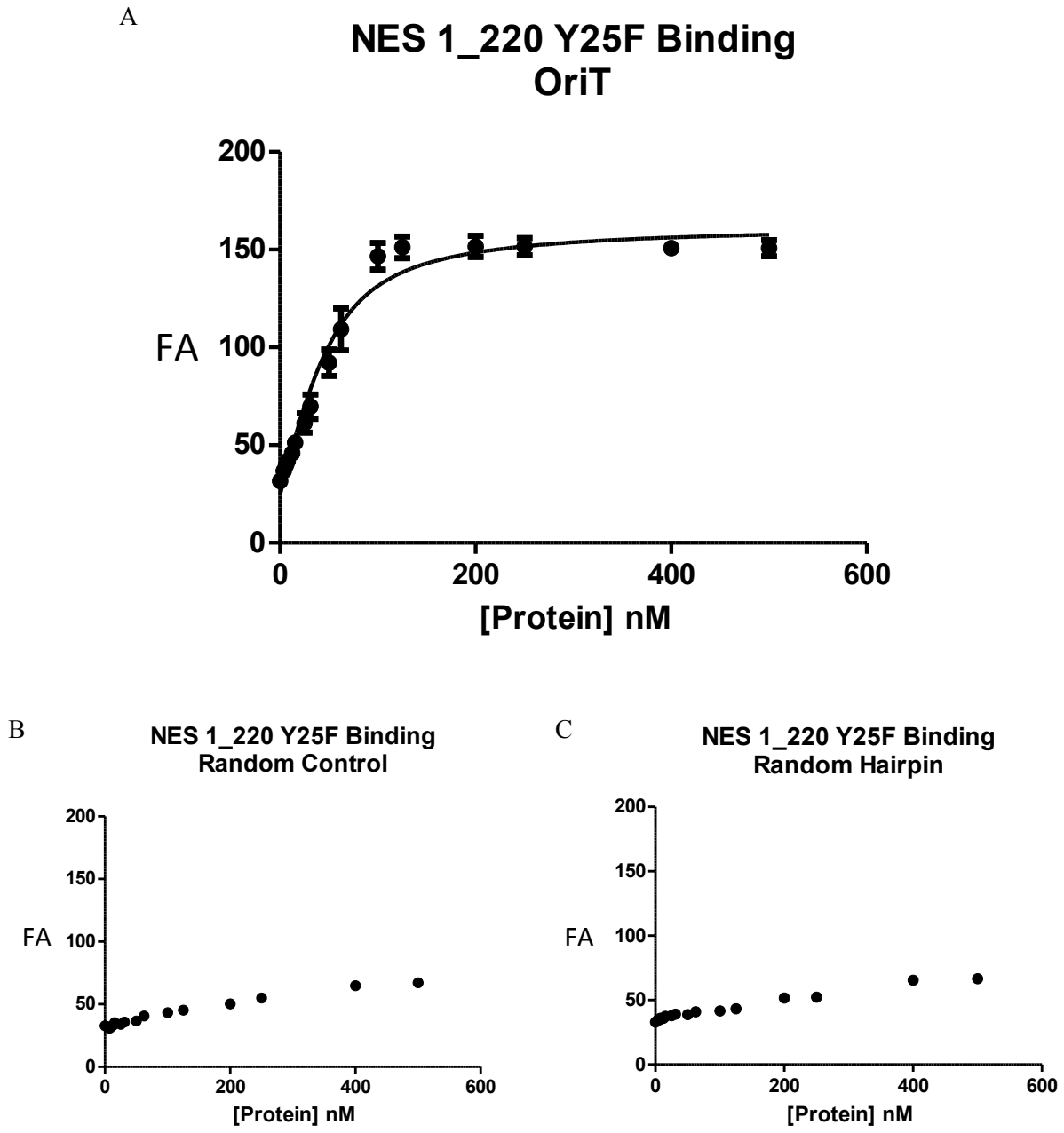
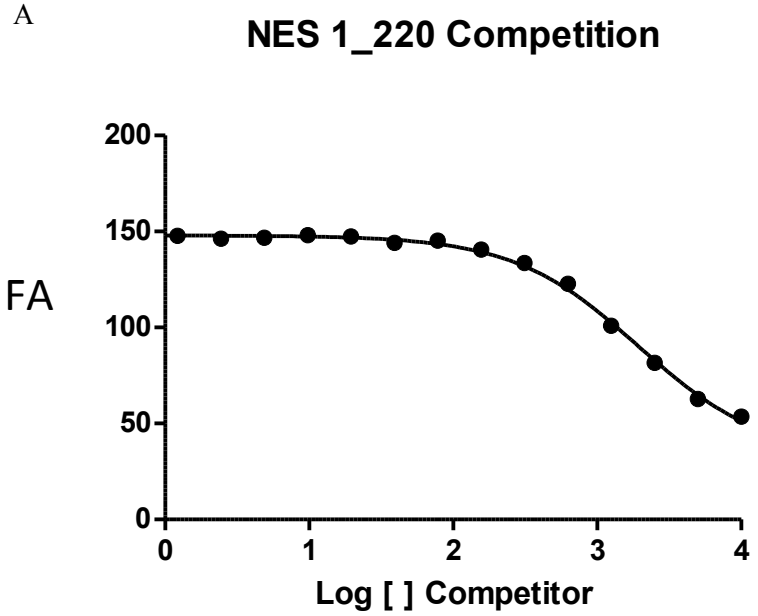


Figure 3.1 NES Relaxase Direct Binding to DNA Substrate

Fluorescence anisotropy experiments examining the binding affinity of NES for the *oriT* and control oligos. A. Binding data from 5 experiments run in triplicate. NES 220 has a K_d of 18.24 nM for the *oriT*. B. Control experiment with a random sequence DNA oligo of the same length. No specific binding was observed. C. Control experiment with a random sequence of DNA that maintains the hairpin secondary structure. No specific binding was observed.

A



B

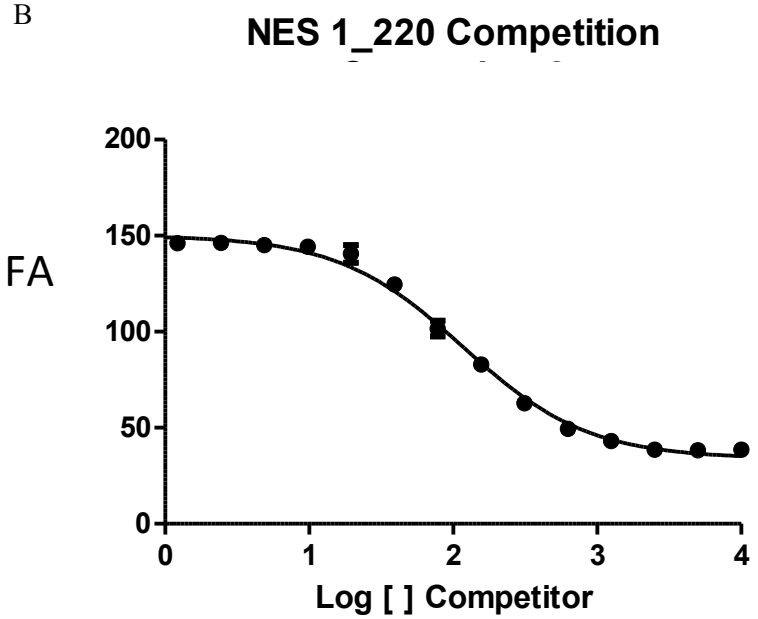


Figure 3.2 Representative DNA Competition Experiment Data

A. An example of a competitor oligo that cannot compete with the reference substrate when encountered by NES. B. An example of a competitor oligo that can compete with the reference substrate when encountered by NES. All experiments were run in triplicate and standard error is graphed for each concentration.

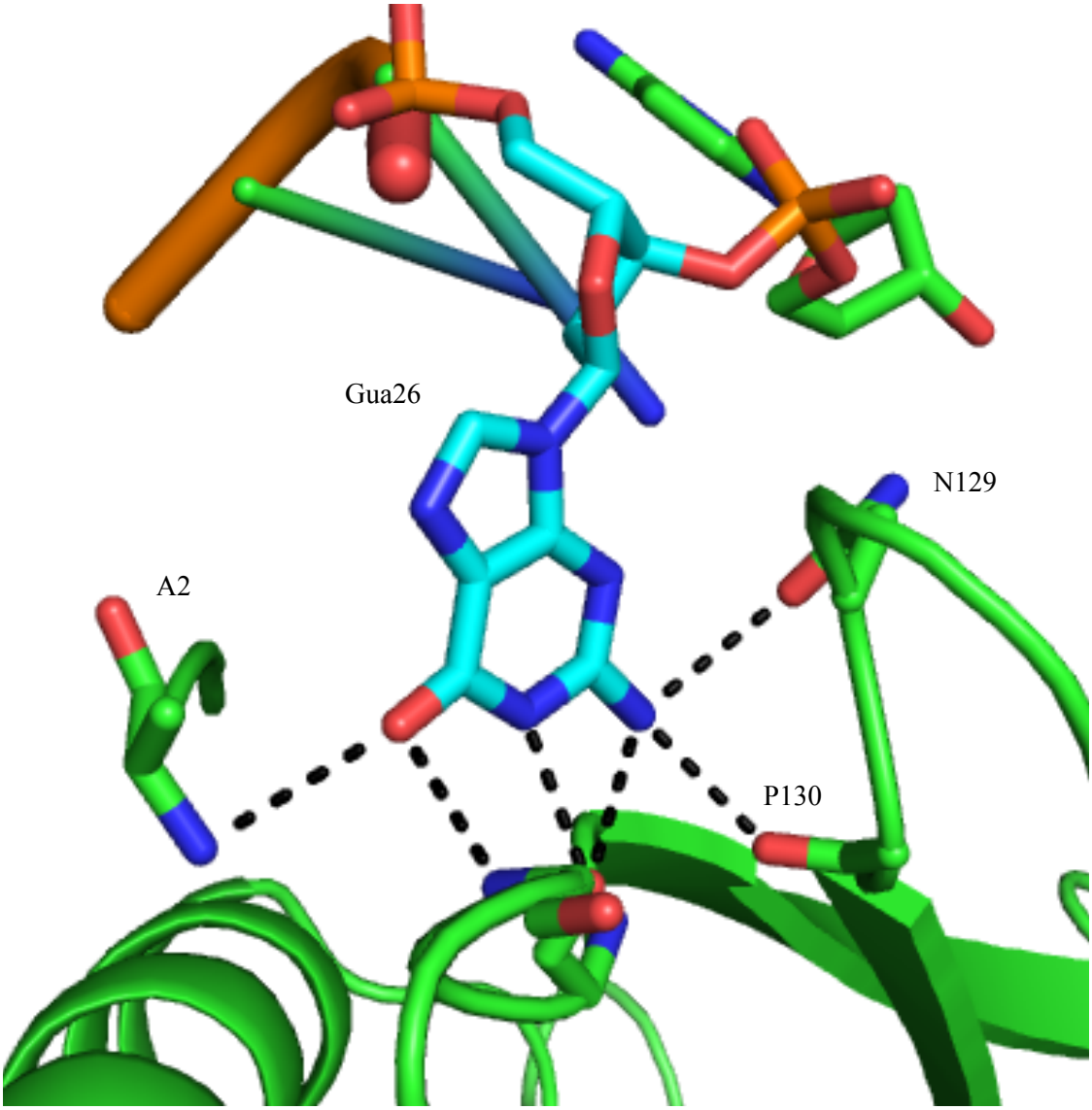


Figure 3.3 Guanine 26 Interaction with NES

The hydrogen bond network between residues in NES and Gua26. Gua26 acts like a post holding the DNA in the correct position cleavage and religation reactions. The N-terminal A2 and N129 make side chain interactions with the DNA base. P130 and L91,N93 (not labeled) make main chain interactions with the Gua26 base.

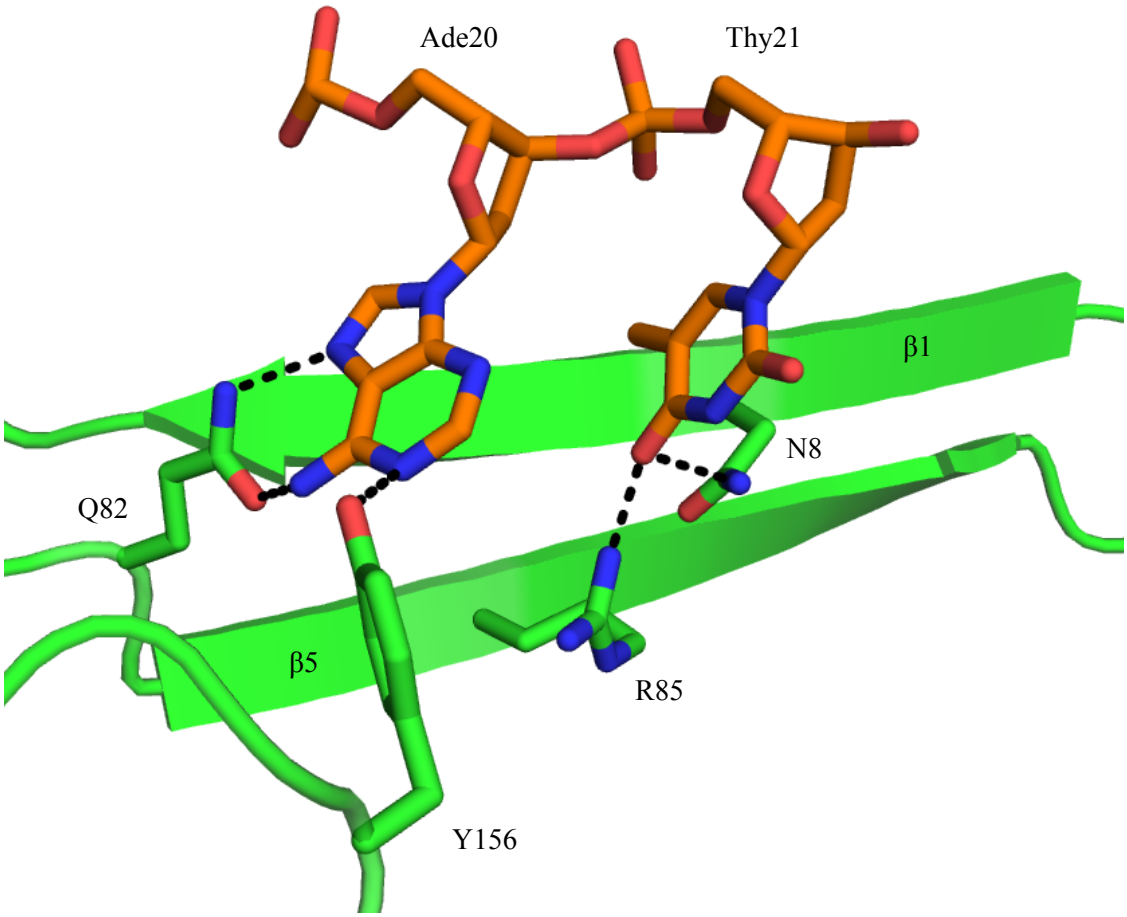


Figure 3.4 NES Residue Interactions with Ade20 and Thy21 of the Single Strand

Q82 and Y156 form hydrogen bonds with base of Ade20. N8 and R85 hydrogen bond with the carbonyl group on Thy21. The base specific interactions along the single stranded portion of the substrate allow for NES to bind in a sequence specific manner.

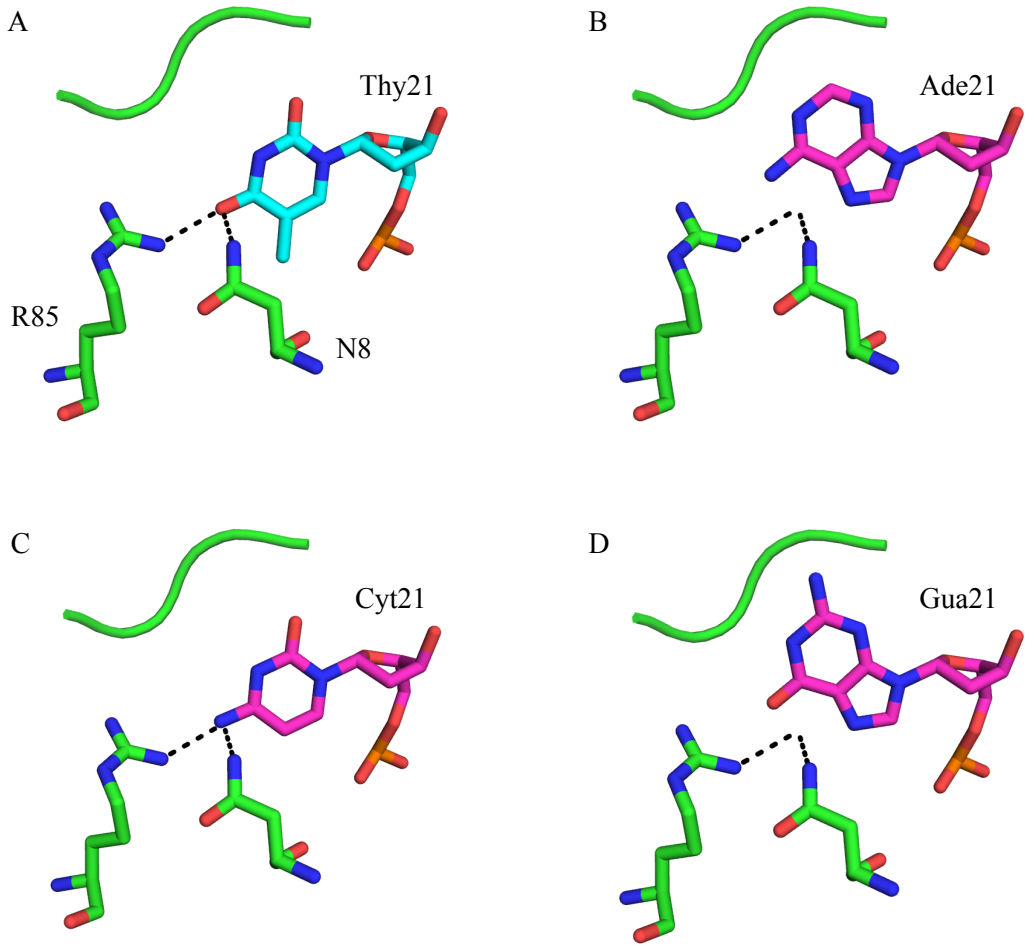


Figure 3.5 Structural Evidence for the Binding Effects of Thy21 Mutants

A. Diagram from the crystal structure showing R85 and N8 hydrogen bonding with the carbonyl group of Thy21. B. Diagram of Ade21 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 18. C. Diagram of Cyt21 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 19. D. Diagram of Gua21 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 20.

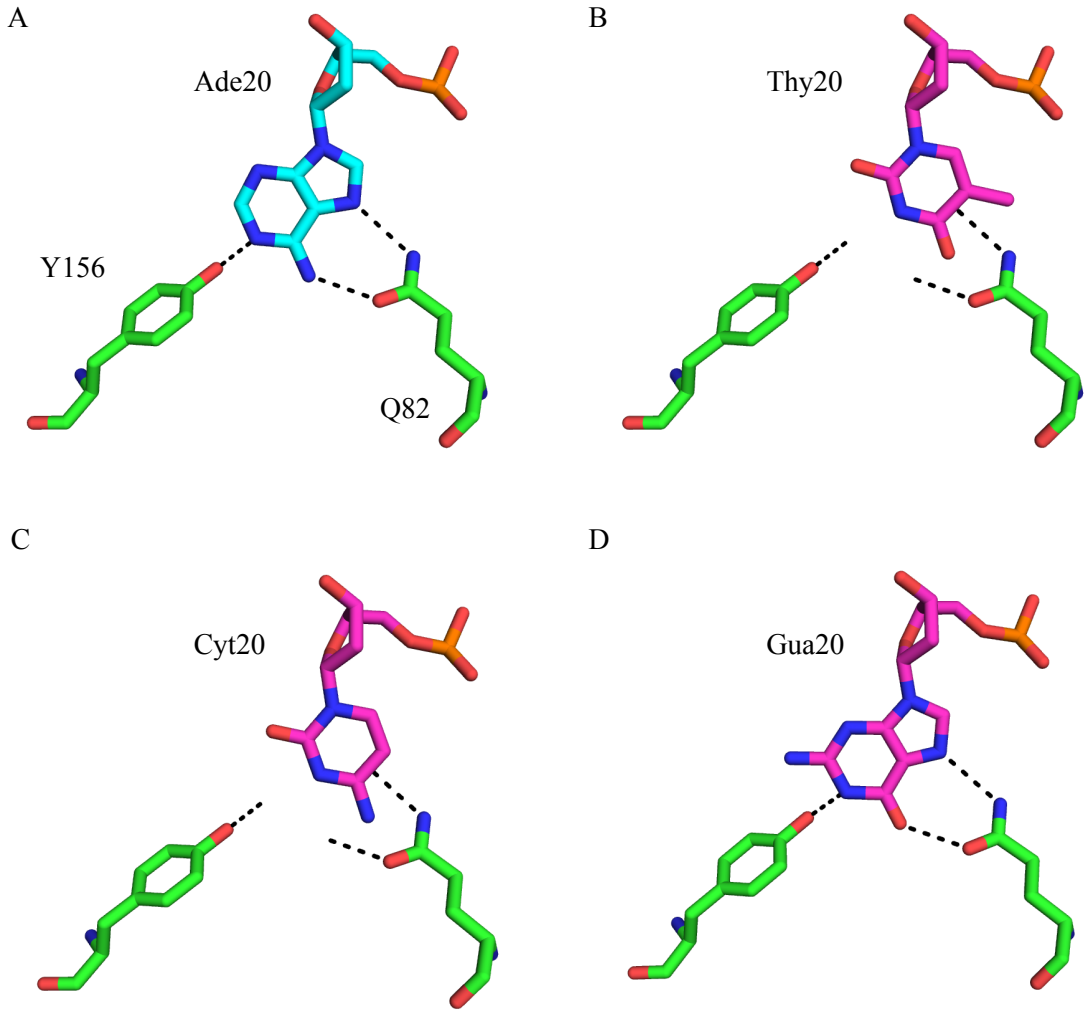


Figure 3.6 Structural Evidence for the Binding Effects of Ade20 Mutants

A. Diagram from the crystal structure showing Q82 and Y156 hydrogen bonding with the amines of Ade20. B. Diagram of Thy20 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 21. C. Diagram of Cyt20 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 22. D. Diagram of Gua20 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 23.

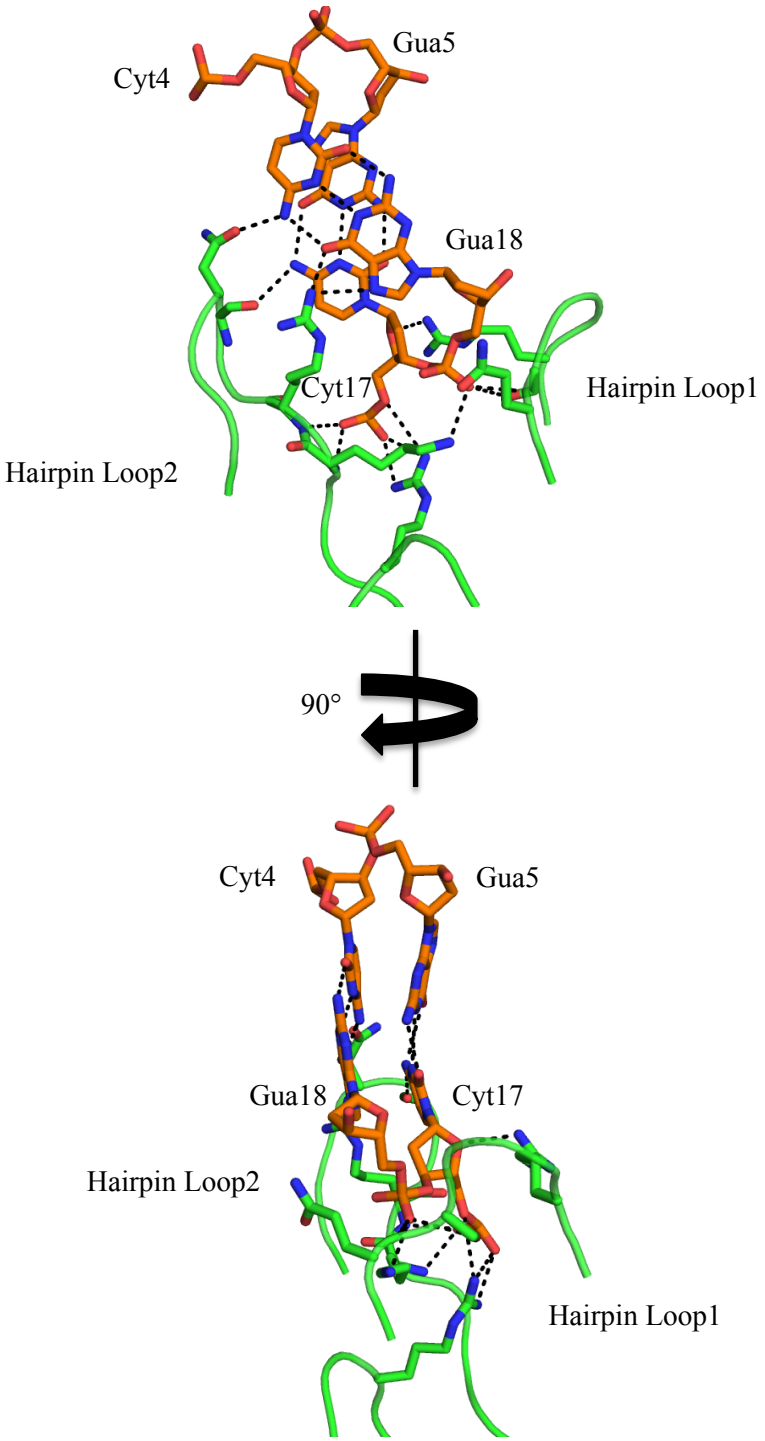


Figure 3.7 NES Binding Site for the Hairpin on the DNA Substrate

The hairpin is recognized by two loops in NES that for a clamp to hold the DNA substrate in place during CPT. Two base pairs are shown and the hydrogen bonds between Hairpin Loop1, Hairpin Loop2, and the DNA are shown in black.

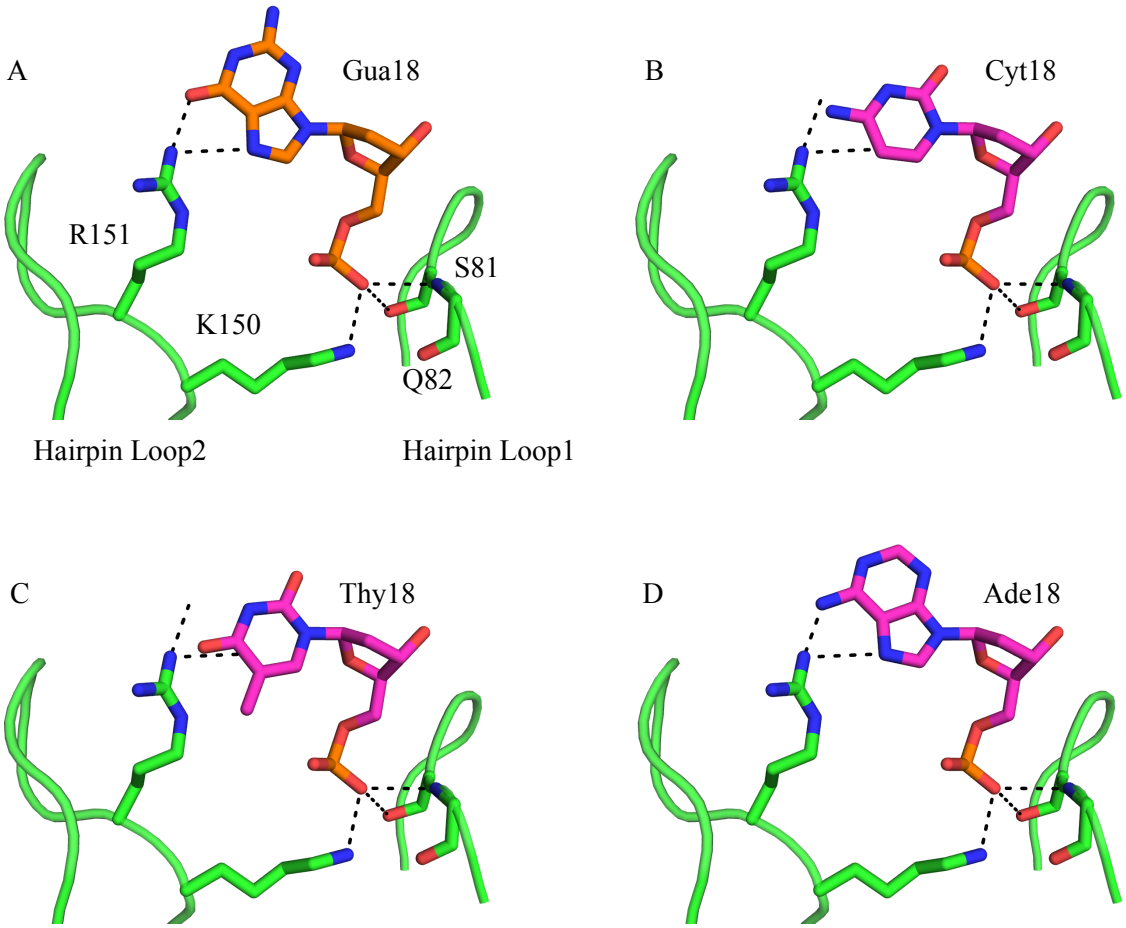


Figure 3.8 Structural Evidence for the Binding Effects of Gua18 Mutants

A. Diagram from the crystal structure showing R151 hydrogen bonding with the Gua18 base. S81, Q82, and K150 have interactions with phosphate backbone. B. Diagram of Cyt18 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 24. C. Diagram of Thy18 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 25. D. Diagram of Ade18 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 26.

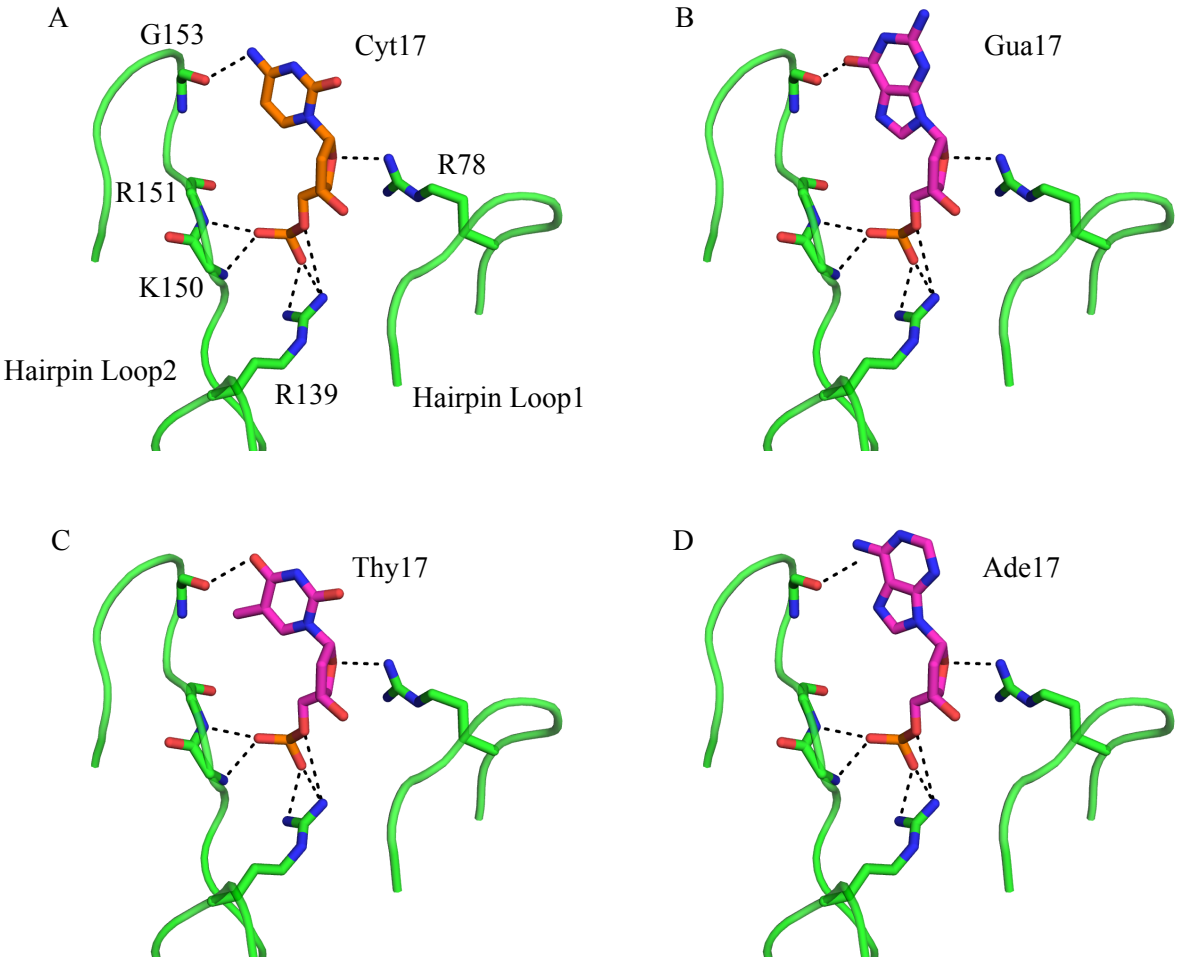


Figure 3.9 Structural Evidence for the Binding Effects of Cyt17 Mutants

A. Diagram from the crystal structure showing G153 hydrogen bonding with the Cyt17 base. R78, R139, and K150 have interactions with phosphate backbone. B. Diagram of Gua17 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 27. C. Diagram of Thy17 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 28. D. Diagram of Ade17 modeled into the difference density map. The binding affinity for this mutant can be found in Table 3.1, oligo 29.

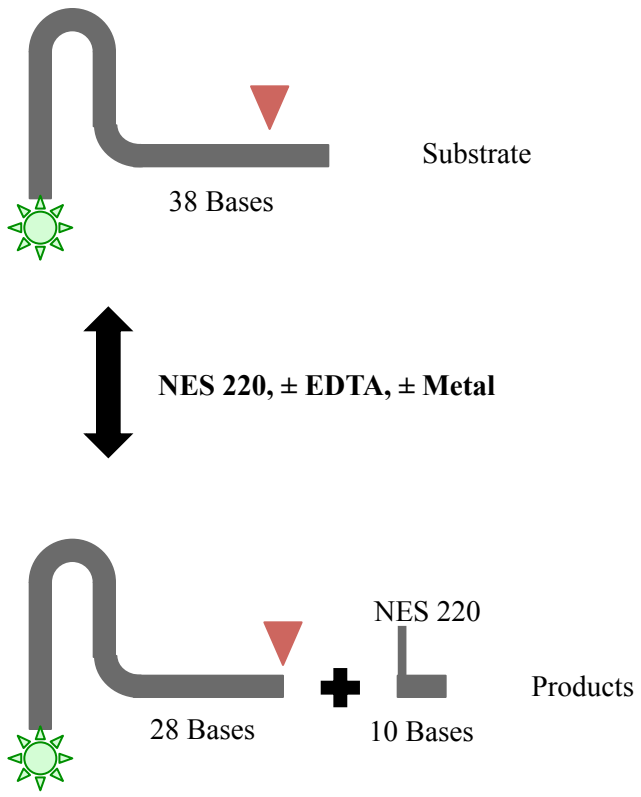


Figure 3.10 Schematic of Cleavage Activity EMSA Assays

The original DNA substrate with a 5' 6-FAM fluorophore is shown in grey. The *nic* site is located at the red arrow. Once the substrate is combined with NES 220 and incubated, NES will cleave the substrate resulting in a shorter fluorescent oligo and ## base DNA is covalently attached to NES. To investigate the metal preference EDTA and specific divalent metals can be added to the assay to measure their activity. The ladder for the experiments contained the original 38 base probe, the shorter 28 base probe, and no enzyme.

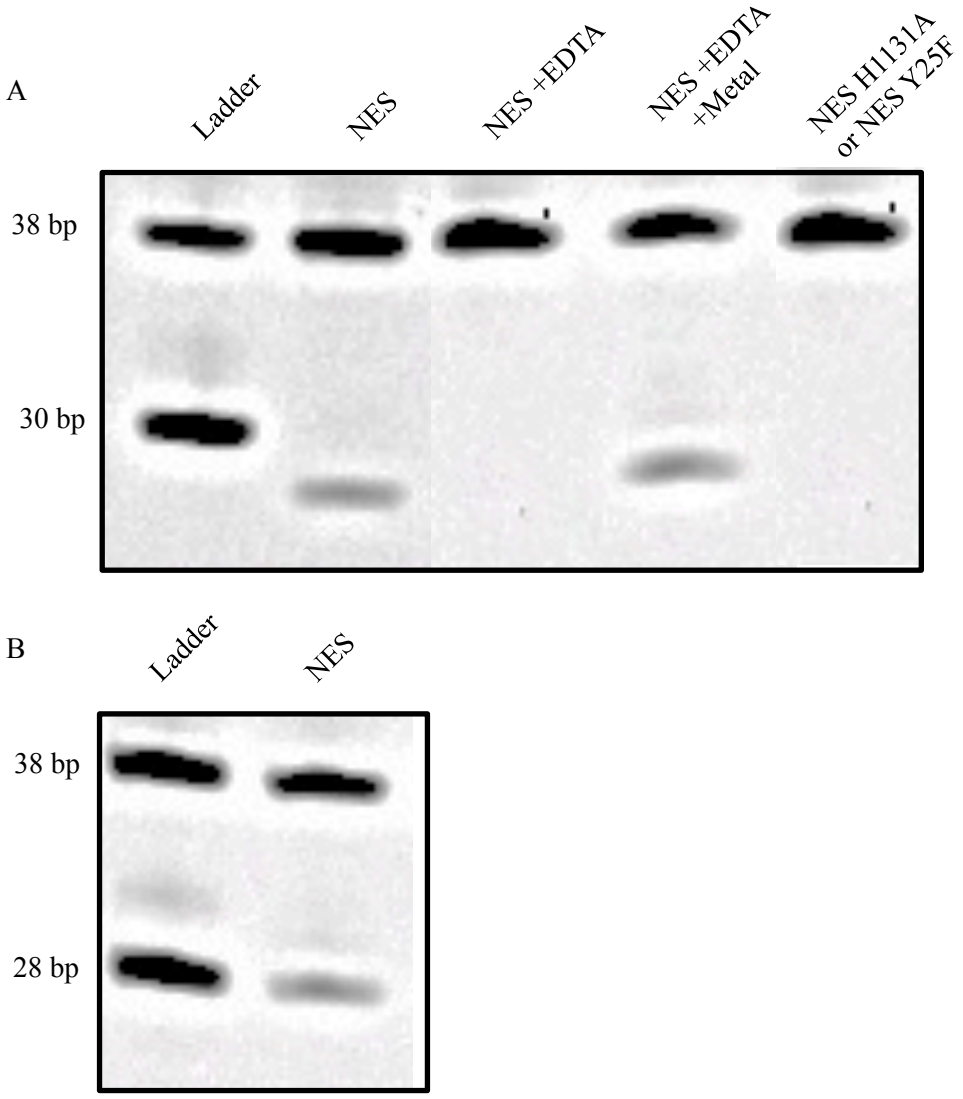


Figure 3.11 EMSA Assay and Controls

A. Initial EMSA using the *oriT* probe and the hypothesize nic site to create the ladder. Lane 1 shows NES is active but cutting at a different site. Lane 2 shows that EDTA can inhibit the activity of NES through chelating the metals. Lane 3 shows that adding in a divalent cation will reconstitute activity. Lane 4 shows that the Y25F mutation in NES is catalytically dead. B. New *nic* site is two bases shorter than previously reported for NES of pGO1.

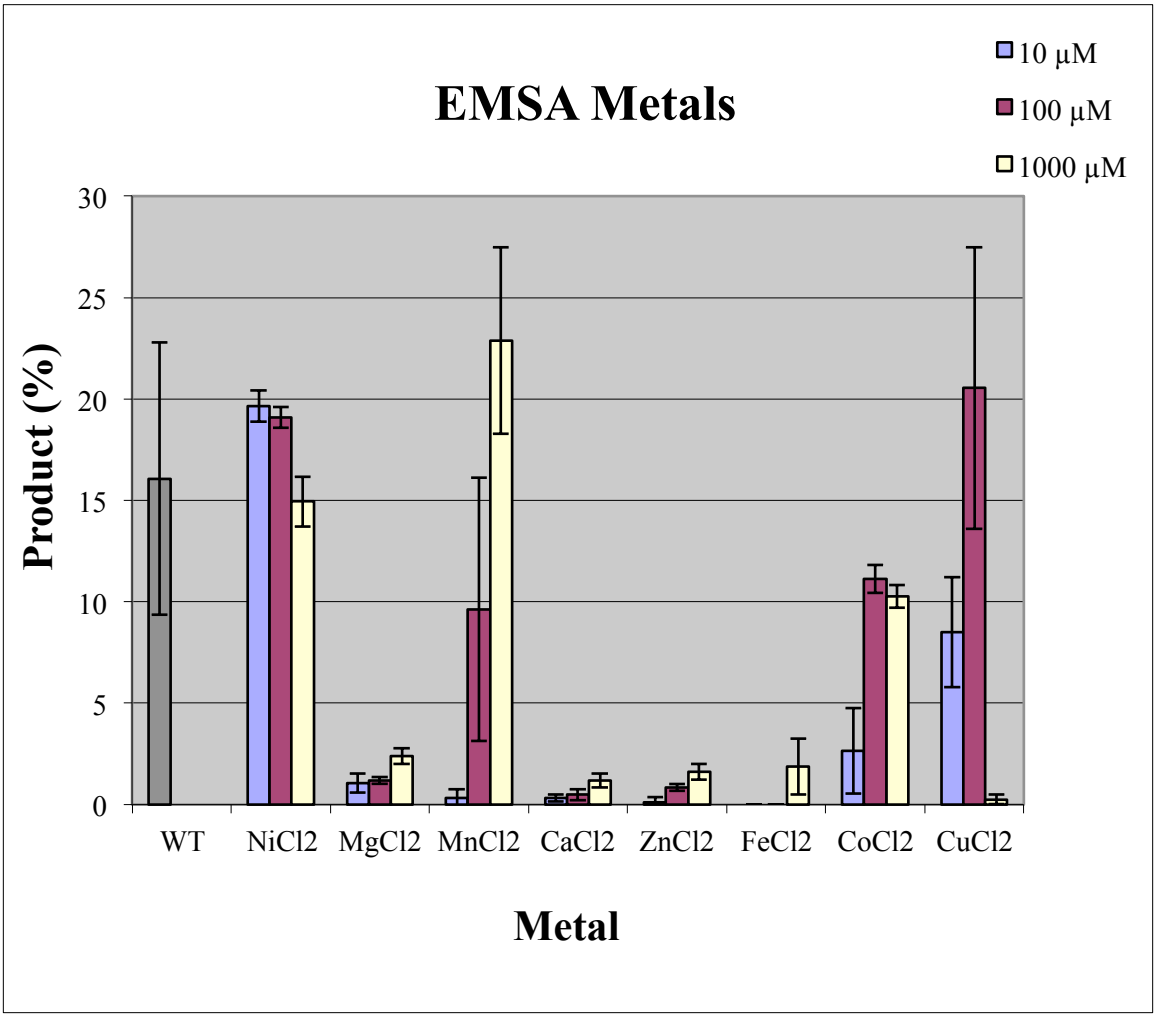
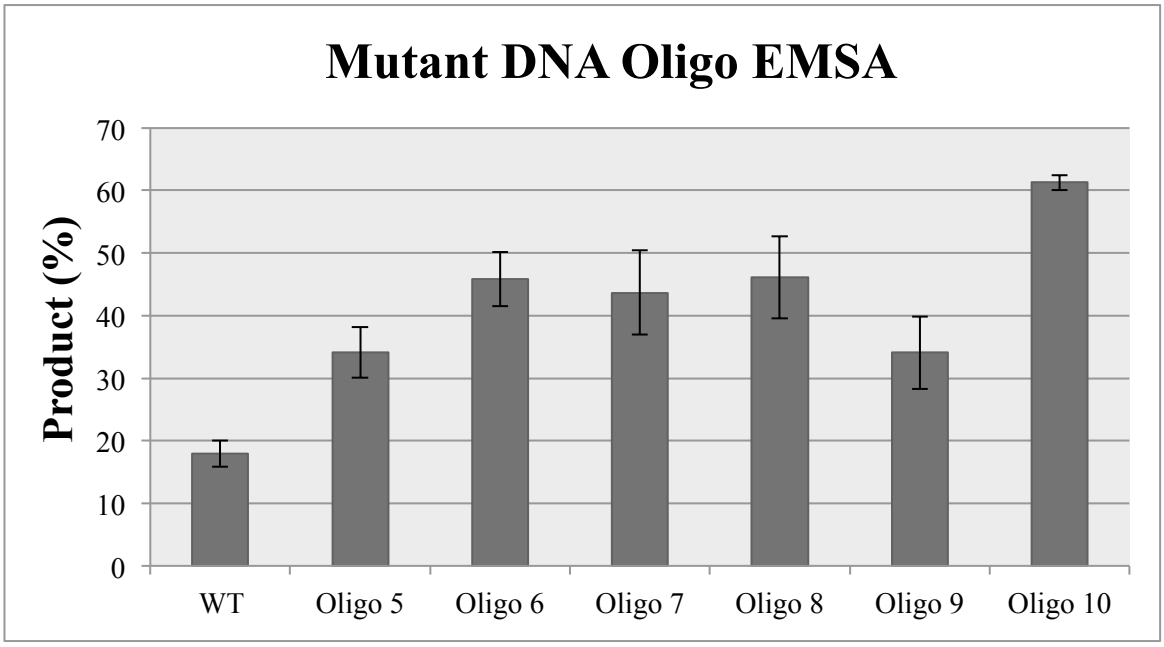


Figure 3.12 Divalent Cation Analysis

This figures depicts the activity of NES when giving various divalent cations in the EMSA experiments. Ni^{2+}, Mn^{2+}, Co^{2+}, and Cu^{2+} showed activity that was comparable to WT. From these data we hypothesize NES uses Mn^{2+} *in vivo* due to its biological relevance and what has been seen in other systems.

A



B

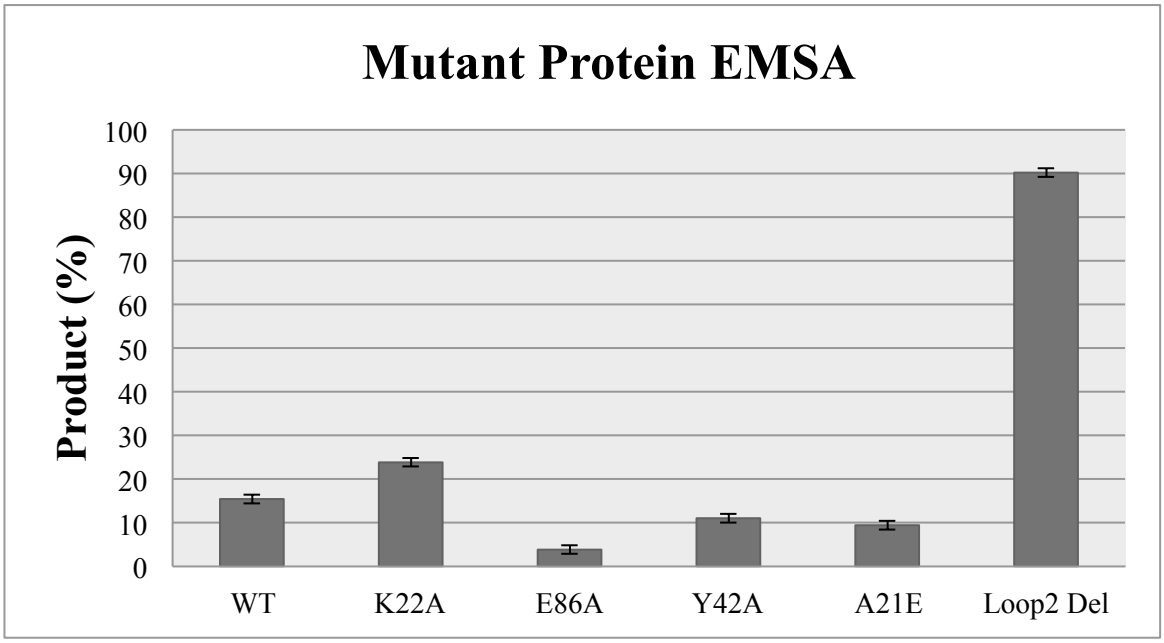


Figure 3.13 Mutant DNA Oligo and Protein EMSA Assays

A. Activity of NES when given mutant oligos as substrates. Each mutation drives the equilibrium towards product formation indicating that the enzyme is unable to efficiently bind to the truncated product. The oligos are from table 3.2. B. The activity of mutant forms of NES when given the WT substrate. All measurements were done in triplicate and S.E. is shown.

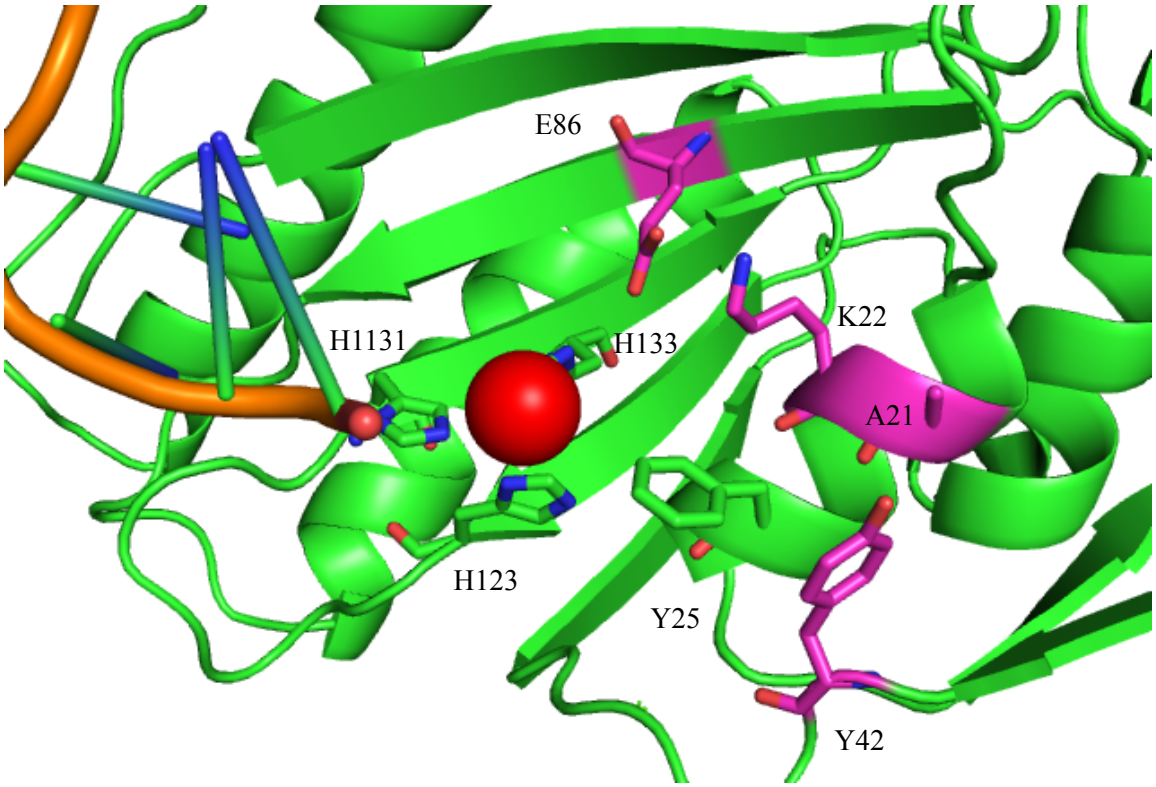
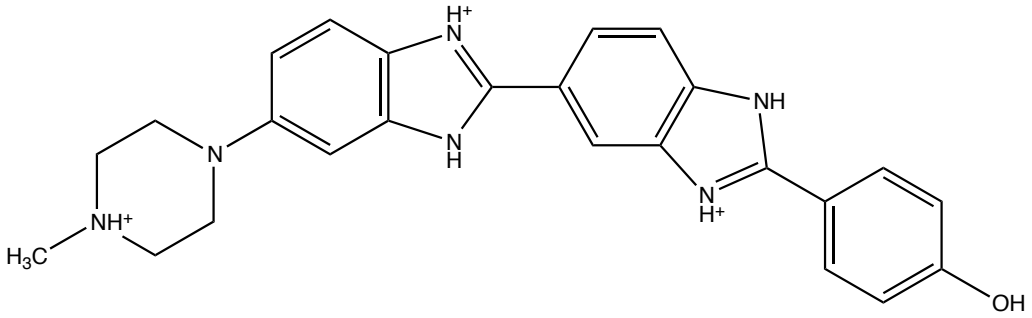


Figure 3.14 Mutations Made to the Active Site of NES

A diagram of the active site of NES. Mutated residues are shown in magenta.

A



B

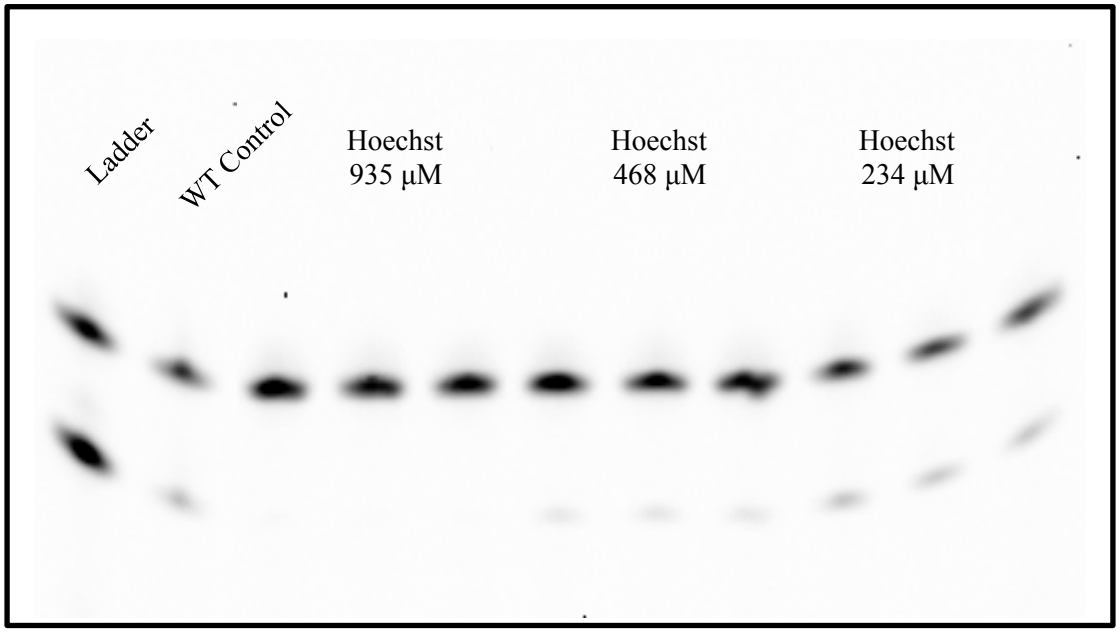


Figure 3.15 Hoechst 33258 Structure and Inhibition EMSA

A. Structure of Hoechst 33258. B. EMSA experiment after incubating the DNA substrate with Hoechst 33258.

3.7 References

1. Pansegrau, W. & Lanka, E. Mechanisms of initiation and termination reactions in conjugative DNA processing. Independence of tight substrate binding and catalytic activity of relaxase (TraI) of IncPalpha plasmid RP4. *J Biol Chem* **271**, 13068-13076 (1996).

2. de la Cruz, F., Frost, L. S., Meyer, R. J. & Zechner, E. L. Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**, 18-40.

3. Nash, R. P., Habibi, S., Cheng, Y., Lujan, S. A. & Redinbo, M. R. The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res* **38**, 5929-5943, doi:10.1093/nar/gkq303 (2010).

4. Morton, T. M., Eaton, D. M., Johnston, J. L. & Archer, G. L. DNA sequence and units of transcription of the conjugative transfer gene complex (trs) of Staphylococcus aureus plasmid pGO1. *J Bacteriol* **175**, 4436-4447 (1993).

5. Berg, T. *et al.* Complete nucleotide sequence of pSK41: evolution of staphylococcal conjugative multiresistance plasmids. *J Bacteriol* **180**, 4350-4359 (1998).

6. Guasch, A. *et al.* Recognition and processing of the origin of transfer DNA by conjugative relaxase TrwC. *Nat Struct Biol* **10**, 1002-1010, doi:10.1038/nsb1017 (2003).

7. Stern, J. C. & Schildbach, J. F. DNA recognition by F factor TraI36: highly sequence-specific binding of single-stranded DNA. *Biochemistry* **40**, 11586-11595 (2001).

8. Williams, S. L. & Schildbach, J. F. Examination of an inverted repeat within the F factor origin of transfer: context dependence of F TraI relaxase DNA specificity. *Nucleic Acids Res* **34**, 426-435, doi:10.1093/nar/gkj444 (2006).

9. Nikolovska-Coleska, Z. *et al.* Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization. *Anal Biochem* **332**, 261-273, doi:10.1016/j.ab.2004.05.055 (2004).

10. Nikolovska-Coleska, Z. *et al.* Design and characterization of bivalent Smac-based peptides as antagonists of XIAP and development and validation of a fluorescence polarization assay for XIAP containing both BIR2 and BIR3 domains. *Anal Biochem* **374**, 87-98, doi:10.1016/j.ab.2007.10.032 (2008).

11. Weigel, L. M. *et al.* Genetic analysis of a high-level vancomycin-resistant isolate of Staphylococcus aureus. *Science* **302**, 1569-1571 (2003).

12. Boer, R. *et al.* Unveiling the molecular mechanism of a conjugative relaxase: The structure of TrwC complexed with a 27-mer DNA comprising the recognition hairpin and the cleavage site. *J Mol Biol* **358**, 857-869, doi:10.1016/j.jmb.2006.02.018 (2006).

13. Larkin, C. *et al.* Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure* **13**, 1533-1544, doi:10.1016/j.str.2005.06.013 (2005).

14. Ilyina, T. V. & Koonin, E. V. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaebacteria. *Nucleic Acids Res* **20**, 3279-3285 (1992).

15. Koonin, E. V. & Ilyina, T. V. Computer-assisted dissection of rolling circle DNA replication. *Biosystems* **30**, 241-268 (1993).

16. Larkin, C., Haft, R. J., Harley, M. J., Traxler, B. & Schildbach, J. F. Roles of active site residues and the HUH motif of the F plasmid TraI relaxase. *J Biol Chem* **282**, 33707-33713, doi:10.1074/jbc.M703210200 (2007).

17. Monzingo, A. F., Ozburn, A., Xia, S., Meyer, R. J. & Robertus, J. D. The structure of the minimal relaxase domain of MobA at 2.1 A resolution. *J Mol Biol* **366**, 165-178 (2007).

18. Lujan, S. A., Guogas, L. M., Ragonese, H., Matson, S. W. & Redinbo, M. R. Disrupting antibiotic resistance propagation by inhibiting the conjugative DNA relaxase. *Proc Natl Acad Sci U S A* **104**, 12282-12287 (2007).

19. Xia, S. & Robertus, J. D. Effect of divalent ions on the minimal relaxase domain of MobA. *Arch Biochem Biophys* **488**, 42-47, doi:10.1016/j.abb.2009.06.004 (2009).

20. Climo, M. W., Sharma, V. K. & Archer, G. L. Identification and characterization of the origin of conjugative transfer (oriT) and a gene (nes) encoding a single-stranded endonuclease on the staphylococcal plasmid pGO1. *J Bacteriol* **178**, 4975-4983 (1996).

21. Finney, L. A. & O'Halloran, T. V. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. *Science* **300**, 931-936, doi:10.1126/science.1085049 (2003).

22. Harding, M. M. Geometry of metal-ligand interactions in proteins. *Acta Crystallogr D Biol Crystallogr* **57**, 401-411 (2001).

23. Grandoso, G. *et al.* Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J Mol Biol* **295**, 1163-1172, doi:10.1006/jmbi.1999.3425 (2000).

24. Nash, R. P., Niblock, F. C. & Redinbo, M. R. Tyrosine partners coordinate DNA nicking by the Salmonella typhimurium plasmid pCU1 relaxase enzyme. *FEBS Lett* **585**, 1216-1222, doi:10.1016/j.febslet.2011.03.043 (2011).

Chapter 4

Structure of C-term Domain and SAXS Envelope

4.1 Introduction

The N-terminal domain of NES contains the relaxase fold and is required for successful CPT of pLW1043. After studying the relaxase domain we were curious to investigate the role of the C-terminal residues 221_665. The relaxases from TrwC, F plasmid, and pCU1 all contain C-terminal helicase domains while R1162 contains a primase [1-4]. These domains play a role in processing the DNA during CPT. We assumed NES would have a similar domain organization to the other relaxases but were surprised when sequence alignments and structure predictions did not provide any significant matches. This left us wondering what the functional role of the ~400 amino acids on the C-term of NES was. We set out to determine the structure of the C-terminal domain to in hopes of shedding light on this question. The experimental procedures and results are presented in the following sections.

4.2 Construct and Oligonucleotide Design, Mutagenesis, Expression and Purification

The C-Terminal Domain of NES, NES 221-665, was cloned and purified as previously mentioned in Chapter 2. Briefly, all constructs in the following work were cloned and expressed into the CPD-Lasso expression vector using the *NdeI* (5') and *BamHI* (3') restriction endonuclease sites. Truncation constructs were made by adding stop codons through site directed mutagenesis and sequences were again confirmed by Eton Biosciences, Inc. The resultant cloned proteins were purified in the same fashion as used in crystallization trials. After HisTrap and size exclusion chromatography the pure fractions were combined, flash frozen in liquid nitrogen, and frozen in stocks stored at -80 °C for use in experiments.

4.3 Limited Proteolysis and N-terminal Sequencing

Limited proteolysis and N-terminal sequencing were used to determine the optimal start position for the C-term constructs. For the limited proteolysis trypsin and proteinase K were prepared in suitable buffers and diluted to a working concentration of 0.01 mg/mL. The trypsin buffer consisted of 20 mM CaCl and 1mM HCl at pH 7.5 and the proteinase K buffer consisted of 5 mM CaCl and 10 mM NaCl at pH 7.5. The reaction buffer was 20 mM Tris pH 7.5 and 300 mM NaCl. The reaction consisted of NES 665 (+DNA) at 0.6 mg/mL in 39 µL of reaction buffer. 5 µL of this solution were removed for an initial time point. 1 µL of protease was added to the remaining 34 µL to initiate the reaction. 5 µL were removed at time points of 1, 5, 15, 30, 60, and 120 minutes. The aliquots were added to tubes that contained 2 µL of 150 mM phenylmethanesulfonylfluoride (PMSF) and 3 µL of 5x laemelli loading dye. Following the quench the samples were boiled for 2 minutes to prepare for sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) electrophoresis.

The limited proteolysis with both trypsin and proteinase K resulted in the formation of a distinct band that corresponded to a MW of ~50 kDa (Figure 4.1). In both cases the addition of the *oriT* DNA did not seem to affect the protease activity. Given the relaxase domain of NES has a MW of ~26 kDa and the full-length protein has a MW of ~79 kDa we hypothesized that this fragment represented to the C-terminal domain of NES. The next step was to determine the N-terminal sequence of this fragment to confirm our hypothesis.

To determine this sequence we sent samples to the W.M. Keck Sequencing Facility at Yale University for analysis. Edman sequencing was carried out on an Applied Biosystems protein sequencer equipped with an on-line HPLC system. The samples purified by SDS-PAGE were electroblotted onto polyvinylidene fluoride (PVDF) membrane. PVDF bands were stained with Coomassie Blue and destained extensively with 4 changes of destaining solvent, and washed with 3-4 changes of ultra-pure water to lower the very high concentrations of Tris, glycine, and other gel and transfer buffers that otherwise will interfere with sequencing. After air drying, the

bands or spots of interest were individually excised and placed in 1.5 ml Eppendorf tubes for shipment to the Keck Facility. Figure 4.2 shows the data generated from the Keck facility for the samples provided. The sequence was compared to NES 665 sequence to identify any residues that were in the same order. We were in fact able to find this sequence and the first base was K228. The molecular weight of NES 228_665 is ~52 kDa, which corresponded well with our data form the limited proteolysis. This portion of the protein sequence was analyzed using the Phyre server [5] to determine if the sequence or the structure was similar to any known enzymes. Unfortunately, from this analysis were not able to find any homologues that had previously been studied leaving us with questions about the function of the C-terminal domain of NES. Using the secondary structure predictions generated from the server, we were able to design multiple truncation constructs by removing C-terminal portions at predicted loop regions. 11 constructs were designed that incorporated these truncations into the C-terminal domain and full-length protein (Figure 4.3)

4.4 Circular Dichroism and Dynamic Light Scattering of NES C-term

Circular dichroism (CD) and dynamic light scattering (DLS) were used in conjunction to investigate the size distribution profile and secondary structure of the C-terminal domain. CD measurements were made on a Chirascan CD Spectrometer (AppliedPhotophysics) with the protein (NES 228_594) diluted into CD buffer (200 mM KF, 10 mM potasim phosphate buffer pH 7.4) to a concentration of 5 uM. A wavelength scan from 190-260 nm revealed a protein that was folded containing α-Helices. This was determined by the presence of a positive peak at 190 nm and a double negative peak at 208 and 225 nm (Figure 4.4.A). The resulting ensemble of peaks made us believe that the protein was in fact folded and primarily α-Helical in nature which was in agreement with the secondary structure predictions. Thermal denaturation was performed while monitoring the 208 nm and 222 nm wavelengths that corresponded to the two negative peaks from the α-Helix signal. The melting temperature (T_m) was determined to be approximately

33 °C by observing where half to the protein was in a folded versus unfolded state from the data (Figure 4.4.B). This low T_m is characteristic of a protein that is not very stable.

Dynamic light scattering was performed on a Dynapro Dynamic Light Scattering Plate Reader (Wyatt Technology Corp.). Prior to measurement the samples were spun at 13,000 *xg* to ensure no dust or sediments would affect the measurement. The sample of NES 228_594 was measured at 10 mg/mL and showed the presence of multiple species. The major peak had a polydispersity of 17.6% and a radius of hydration of 5.4 nm (Figure 4.5). The molecular weight calculated from this reading was 172 kDa, corresponding to roughly a trimer or a molecule that is not spherical in shape. The other two peaks represented only 1% of the signal. Taken together these data are representative of a well folded, but unstable and in multiple states. CD and DLS measurements were taken on the other C-terminal constructs that incorporated both C-terminal and N-terminal truncations. All of the samples tested had a low T_m had similar size profile characteristics (data not shown).

4.5 Crystallization, Data Collection, and Data Processing

Selenomethionine-substituted protein was used in crystal trials to obtain phase information upon data collection. NES 243_594 (10 mg/mL) was crystallized in a solution containing 16% PEG 4,000, 5% PEG 400, 500 mM lithium chloride, 100 mM Tris (pH 8.0), and 12.5% glycerol. The trays were set with drop ratios of 1:1, 1:2, and 2:1 of protein to mother liquor respectively. Vapor diffusion hanging-drop trays were set at 20 °C to obtain well formed crystals. The crystallization solution was found to be an adequate cryoprotectant in subsequent diffraction experiments performed at the UNC Biomolecular X-ray Crystallography Facility. The crystals were flash frozen in liquid nitrogen for data collection at 100K.

Data sets were obtained on the 23_ID_B beamline maintained by The National Institute of General Medical Sciences and National Cancer Institute Collaborative Access Team (GM/CA-CAT) at the Advanced Photon Source (APS) part of the Biosciences Division (BIO) at Argonne National Laboratory (ANL). A fluorescence scan was performed at the selenium absorption edge

to confirm that the seleno-methione substitution was successful. The collection strategy used included 1-second exposures with a 1° rotation between frames. Three data sets were collected at the peak, inflection, and remote wavelengths for selenium. The data sets were indexed and scaled using HKL-2000 processing software [6]. Indexing resulted in resulted in a space group of $P4_32_12$ with one NES C-terminal domain per asymmetric unit. The best data was obtained when the resolution was cut to 3.0 Å (Table 4.1). Initial phases were determined using the Phenix software suite with the AutoSol function [7]. Four of the six heavy atom sites were successfully located. The first two heavy metal sites in the protein were disordered resulting in the structure for residues 255_594. The model was built using Coot [8] by placing the seleno-methiones and iteratively building the protein from these starting points. Refinements were performed using the Phenix refine application [7].

4.6 Analysis of the C-terminal Domain Structure

Only one C-terminal domain was observed per crystallographic asymmetric unit, indicating that this domain functions as a monomer, this conclusion was not in agreement with the data from DLS (Figure 4.5) or the size exclusion chromatography experiments (SEC) (data not shown) that seem to indicate the presence of an oligomer. This finding can be explained by the overall shape of the protein which is contains 15 α-Helices in a convex arrangement (Figure 4.6). This would make the molecule run like a larger species during SEC experiments and give the molecule a larger radius of hydration. The residues 255_294 could be built into the density and several loops were disordered. Multiple residues were left as alanine due to the low resolution and difficulty in placing the side chains in the difference density maps.

The overall fold of this domain contains 15 α-Helices arranged in a convex formation (Figure 4.6). The structure appears to be a domain swap in the secondary structure where the secondary structure of α1-4 are in a similar arrangement to α5-10. Observing the structure from above (Figure 4.6) it is clear that the helices are arranged in a convex shape. Since we are currently not able to assign function to this domain we are currently operating under the

assumption that is a docking domain that interacts with other members of the relaxosome or transportation machinery.

4.7 SAXS Data Collection and Analysis

In order to better understand the arrangement of the two domains of NES we employed small-angle X-ray scattering (SAXS) to determine the molecular envelope. A sample of full-length NES in complex with DNA was sent to Rigaku (The Woodlands, TX) for analysis. Through determining what the molecular envelope for the full-length protein looked like we hoped to be able to place both domains in the correct orientation and perhaps gain some insights into the C-term function.

The samples were studied using a BioSAXS-1000 system connected to a Rigaku FR-E+ X-ray generator. Data were collected for silver behenate (AgBeh), a common powder compound often used to calibrate both powder diffraction and SAXS instruments. AgBeh data collected on this instrument for 10 seconds were used to determine the beam center location and to calibrate the pixel-to-q conversion. Calibration of q and the direct beam position was carried out in the program SAXSLab automatically. For each sample, the original concentration (100%) and two dilutions corresponding to 50% and 25% of the original were prepared. Prior to SAXS measurements, ~30 µl of each sample dilution was centrifuged for 10 minutes using a microcentrifuge. then, 23 µl of each sample or buffer was pipetted inside a sample cell (i.e., a quartz capillary in a special holder), and all 3 sample cells were put inside the BioSAXS-1000. For each sample and buffer measurement, duplicate scans of 5 min and 15 min were collected. All data were collected at room temperature (20°C, in a temperature controlled sample holder). Sample and buffer images were converted from pixel space to q-space and averaged to produce 1D plots of 'I vs. q' using SAXSLab which was calibrated with a AgBeh standard. Duplicate 'I vs. q' plots were averaged together and the averaged buffer plots were subtracted from the averaged sample plots in the SAXSLab software.

Consecutive, duplicate scans of 5 and 15 minutes were collected for the highest

concentration of NES 665 + DNA. The scattering profiles for these initial exposures, shown in Figure 4.7, were scaled for display purposes. The averaged 5 minute scan demonstrates that the original solution was not aggregated, while the averaged 15 min exposure shows that X-rays did not cause any sample aggregation. This is confirmed by the radius of gyration (R_g) calculated from Guinier plots. The values listed in Table 4.2 are consistent at about 39.5 Å for both exposures. Lastly, note that the values for I_0 match from both exposure times. Figure 4.8 shows that the Guinier plots for all three concentrations are linear, indicating that the protein molecules are mono-dispersed, and the R_g remains consistent with an average value of 40 Å (Table 4.2). Also, note that the I_0 values for the 50% and 25% dilutions are half of the 100% and 50% concentrations, respectively, as expected.

Scattering data were collected on three concentrations of NES Figure 4.9.A shows the scattering data for all three solutions at 20°C; the scattering plots are superimposable, but they are scaled for display purposes. The scattering profile for the 100% solution shows the excellent quality of the data with little noise for the entire range of q (Figure 4.9.A). For $q > 0.2$, the noise increases with the sample dilution. A prominent peak in a Kratky plot is an indicator of a compact, or properly folded, sample. All Kratky plots show a single peak (Figure 4.9.B), so NES 665 + DNA appears to be folded.

The structural parameters calculated for NES 665 + DNA are listed in Table 4.3. R_g values were determined manually, using Guinier plots shown in Figure 4.8, and using the real-space Pair-Distance Distribution Function (PDDF) plots calculated by GNOM (Figure 4.10). The molecular weights (MW) were estimated from the Porod volume by using the program datPOROD and the webtool SAXSMoW–which is provided by the University of São Carlos, Brazil (http://www.ifsc.usp.br/~saxs/saxsmow.html) [9]. Both programs use GNOM files as input. For SAXSMoW, data up to a q value of 0.15 $Å^{-1}$ were used. The R_g and Dmax values estimated by GNOM remain consistent for all three concentrations at about 41 Å and 140 Å, respectively. Moreover, there is nice agreement between the Guinier and real-space R_g values. The average

MW values estimated from datPOROD (87 kDa) and SAXSMoW (110 kDa) indicate that NES 665 + DNA is a monomer. Figure 4.10 shows the PDDF plot and the corresponding fit to the experimental scattering data for the 100% concentration of NES 665 + DNA. Figure 4.10.A shows visually how closely the idealized curves fit the data. Figure 4.10.B, the three PDDF shapes remain mostly unchanged and superimposable, with a prominent peak position at 33 Å. These results are considered just short of excellent according to the GNOM solution quality scores (87-88%) in Table 4.3.

Using the data generated form the initial analysis by Rigaku we set about to calculate the molecular envelope for the molecule. DAMMIF [10] was first used for rapid *ab initio* shape determination by simulated annealing using a single phase dummy atom model. This program works to construct the molecular envelope through adjusting an arbitrary starting volume of particles to match the GNOM data. The program was run 30 times to generate a sufficient amount of models for accurate averaging. The DAMAVER [11] suite was programs was used to superimpose and average the bead models to remove any features that are not representative of the true envelope. Lastly, SUPCOMB [12] was used to attempt to place both models. Unfortunately, this was not possible due to the fact we had two separate domains with no insights as to how they fit together. The averaged envelope generated from DAMAVER was manually fit with the relaxase and C-terminal domain (Figure 4.11). Future studies will be conducted in order to get SAXS data for the separate domains as well as full length NES without DNA to facilitate our understand of the domain arrangement in NES.

4.8 Conclusions

Although prediction servers were not able to identify any similar sequences or secondary structures in the literature to date, we set out to determine the structure of the C-terminal domain in hopes of identifying a possible function for this domain. By using limited proteolysis and N-terminal sequencing we were able to identify a well folded, but unstable, C-terminal domain of NES. This was confirmed through CD and DLS, which indicated a protein that was highly α-

Helical in nature but fairly unstable due to the low T_m and polydispersity. Multiple constructs were designed and through X-ray crystallography we were able to determine the structure of the C-terminal domain of NES (residues 243_594) but not assign any biological function. This is unusual for a relaxase as the other enzymes studied to date show some sort of DNA processing domain at the C-terminus. We hypothesize that this domain is a scaffold, responsible for binding to another enzyme either in the relaxosome or part of the transfer proteins. To further investigate the domain arrangement we employed SAXS to determine the molecular envelope. The initial envelope showed an elongated orientation of the two domains indicating that the C-term does not play a role in processing the *oriT* DNA. Thus, we conclude that the C-term is not involved in DNA processing and is most like a structural domain. Further work will be conducted to identify the binding partner and assign a function to the C-term.

Data Collection			
X-ray source	GM/CA-CAT	GM/CA-CAT	GM/CA-CAT
Space Group	P 4_3 2_1 2	P 4_3 2_1 2	P 4_3 2_1 2
Unit cell: a,b,c (Å); α,β,γ (°)	75.7, 75.7, 181; 90, 90, 90	75.7, 75.7, 181; 90, 90, 90	75.7, 75.7, 181; 90, 90, 90
Data Set	NES 243-594 **PEAK**	NES 243-594 **INFLECTION**	NES 243-594 **REMOTE**
Wavelength (Å)	0.97926	0.97949	0.94928
Resolution (Å) (highest shell)	100.0-3.00 (2.90-2.95)	100 – 3.0 (2.90-2.95)	100 – 3.0 (2.90-2.95)
I/σ	26 (3.6)	29 (4.8)	31 (3.8)
Completeness (%)	100 (99.7)	100 (99.7)	100 (99.7)
Redundancy	9	9	9
Refinement			
Resolution (Å)		20.0 – 3.0	
No. reflections		10,985	
R_{work}		25.5	
R_{free}		32.7	
Molecules per asymmetric unit (AU)		1	
No. of amino acids per AU		339	
No. of waters per AU		0	
Average *B*-factors		79.8	
R.M.S. deviations			
Bond lengths (Å)		0.003	
Bond angles (°)		0.739	
Ramachandran (%)			
Preferred		80.1	
Allowed		10.0	
Outliers		10.0	

Table 4.1 Crystallographic Statistics

Conc.	Exposure Time (min.)	R_g (Å) (error) Guinier	Data Points Used	I_0 (error)
100%	5	39.4 (0.5)	3-17	3.6 (0.03)
	10	39.8 (0.3)	1-17	3.7 (0.02)
50%	5	40.4 (0.7)	3-17	1.9 (0.03)
	10	40.1 (0.5)	1-17	1.9 (0.02)
25%	5	39.8 (1.2)	4-17	0.88 (0.02)
	10	39.3 (0.7)	4-17	0.89 (0.01)

Table 4.2 Radius of Gyration for NES 665 + DNA at 2 Exposure Times

Conc.	Gunier Rg(Å) (error)	GNOM Rg (Å) (error)	Dmax (Å) GNOM	Pts used by GNOM	GNOM solution quality (%)	MW (kDa) datPOROD*	MW (kDa) SAXSMoW
100%	39.8 (0.3)	41.2(0.2)	142	1-175	86.6	86	113
50%	40.1 (0.5)	41.4(0.3)	140	1-151	87.7	86	112
25%	39.3 (0.7)	41.1(0.5)	139	4-105	88.2	90	106

Table 4.3 Structural Parameters Obtained for NES 665 + DNA

These data were obtained using the 15 minute exposure time at 20° C. *Calculated by dividing the Porod volume by 1.7

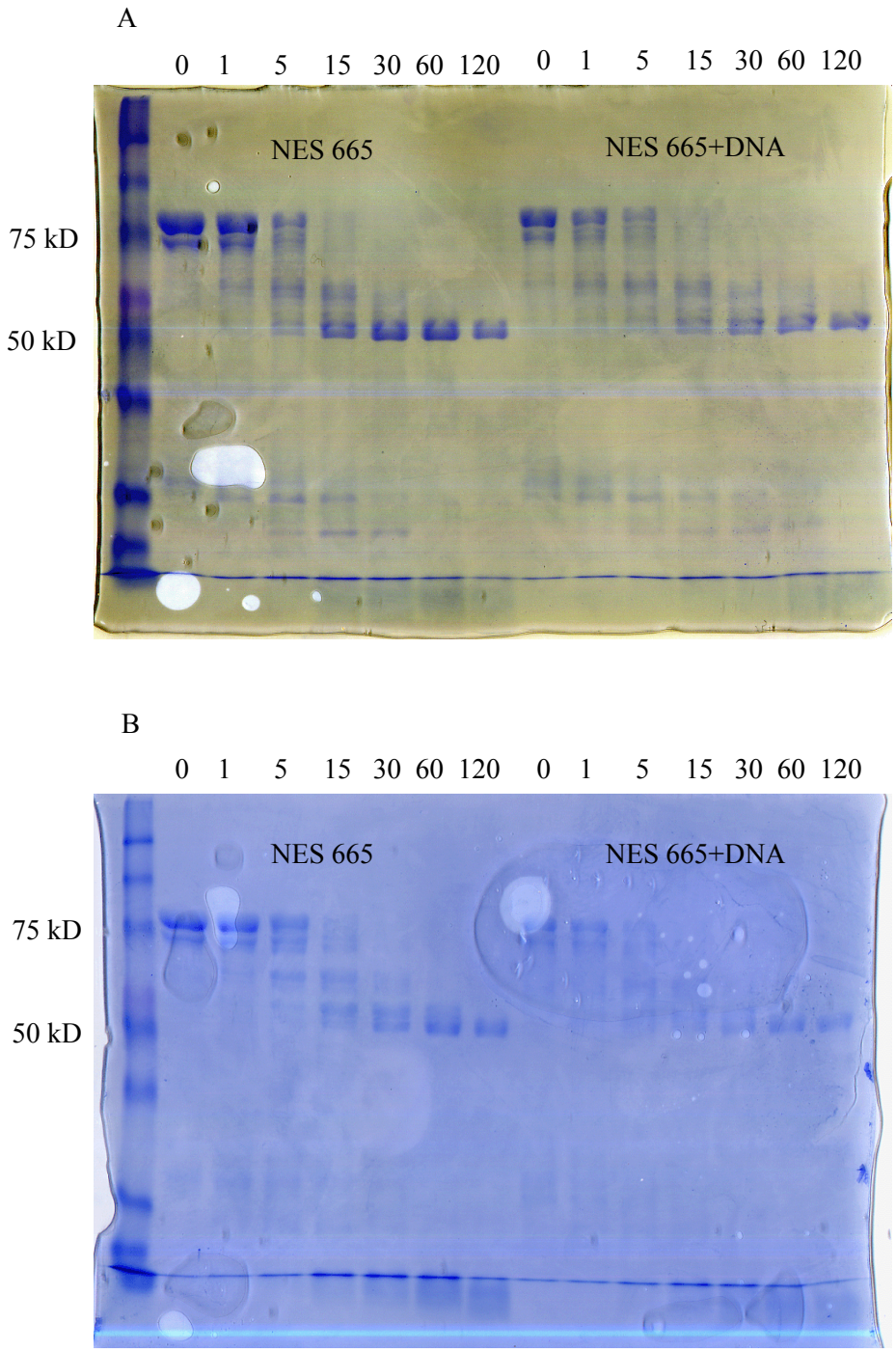


Figure 4.1 Limited Proteolysis of NES 665

SDS-PAGE gels of limited proteolysis performed on NES 665 and NES 665 + DNA. A. Limited proteolysis performed with Trypsin. B. Limited proteolysis performed with proteinase K.

```
                LYS
                THR
                ALA          (PHE)  GLN
                SER          (GLU)  VAL  (ILE)                GLY
SEQUENCE:        X   -   X   - ASN - GLU - SER - ILE - LYS - LYS -   X
CYCLE #:         1       2      3     4     5     6     7     8      9


                MET    PHE
SEQUENCE:       ASN - GLN
CYCLE #:         10     11
```

NES Sequence - k h n e s i k k r n q
228 238

Figure 4.2 N-Terminal Sequencing Results on Proteolysis Fragment

Yellow residues represent the highest percentage in sequencing results. An x signifies no data. The NES sequence that corresponds is shown below.

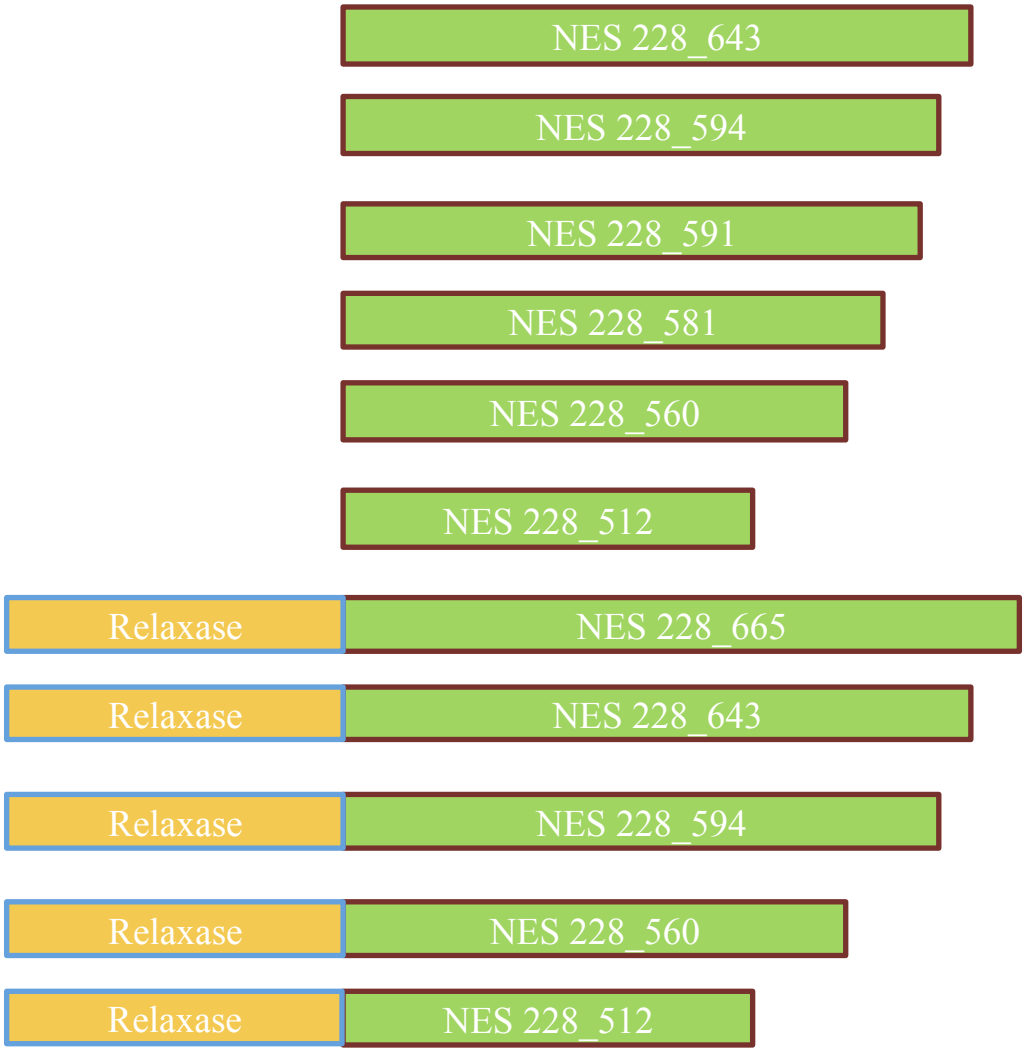
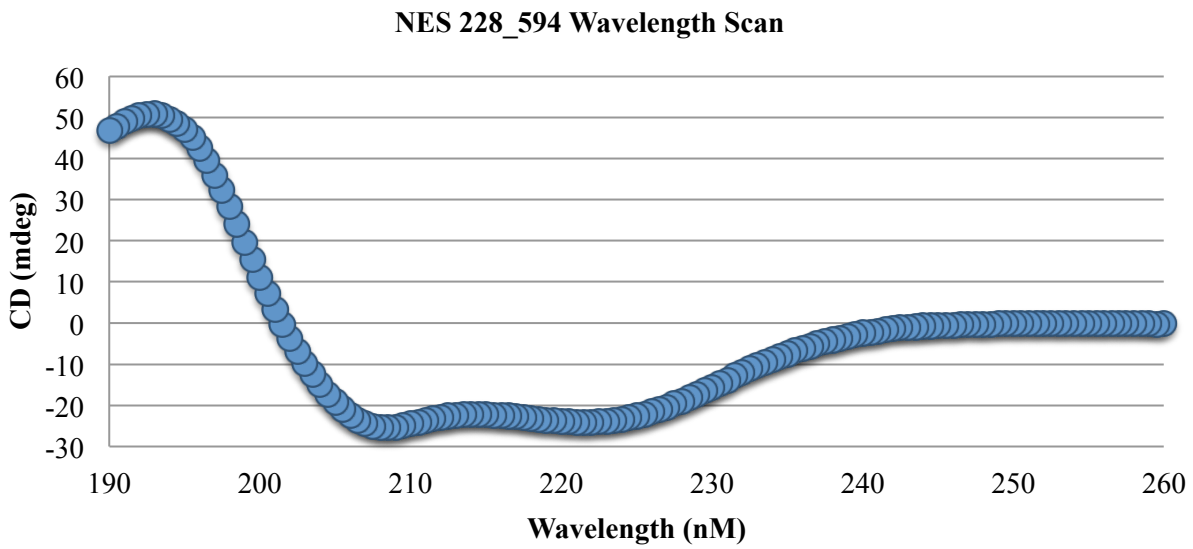


Figure 4.3 NES Crystallization Constructs



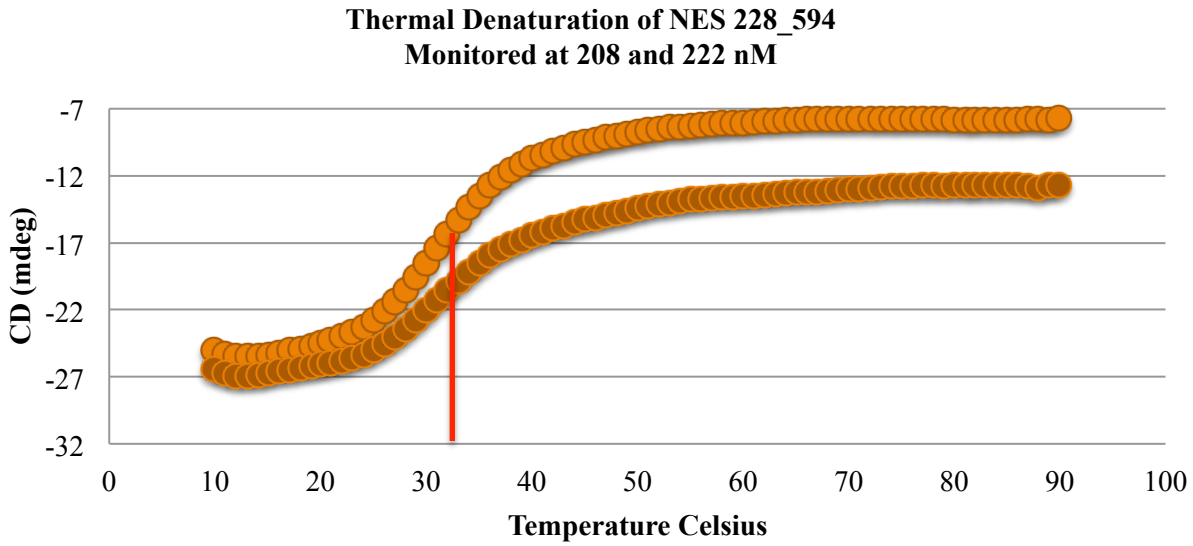
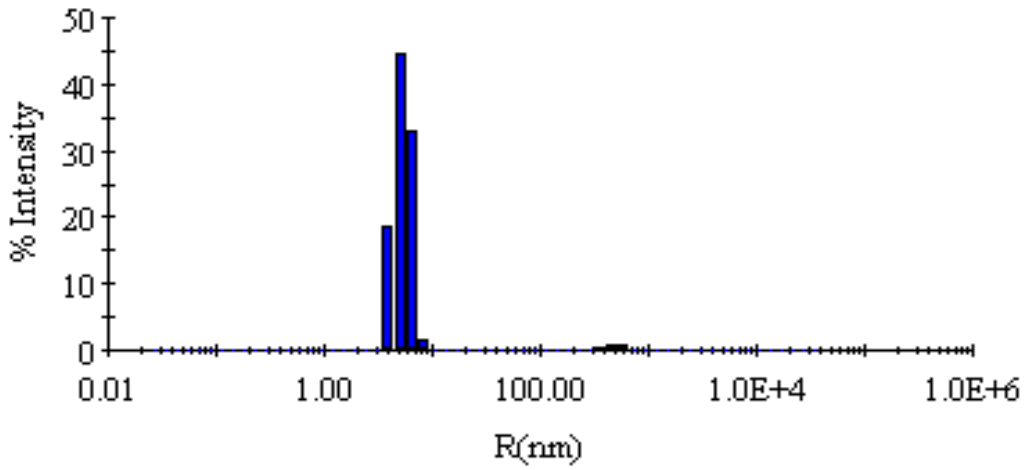


Figure 4.4 Circular Dichroism Analysis of NES 228_594

A. CD wavelength scan of NES 228_594 at 5 µM. B. Thermal denaturation of NES 220 monitoring the change in CD signal (mdeg) at 208 and 222 nM.



Item	R (nm)	%Pd	MW-R (kDa)	%Int	%Mass
Peak 1	5.4	17.6	172	97.6	99.0
Peak 2	530.2	20.5	7977510	1.7	0.1
Peak 3	21753.3	36.0	47437100000	0.7	1.0

Figure 4.5 Dynamic Light Scattering of NES 228_594

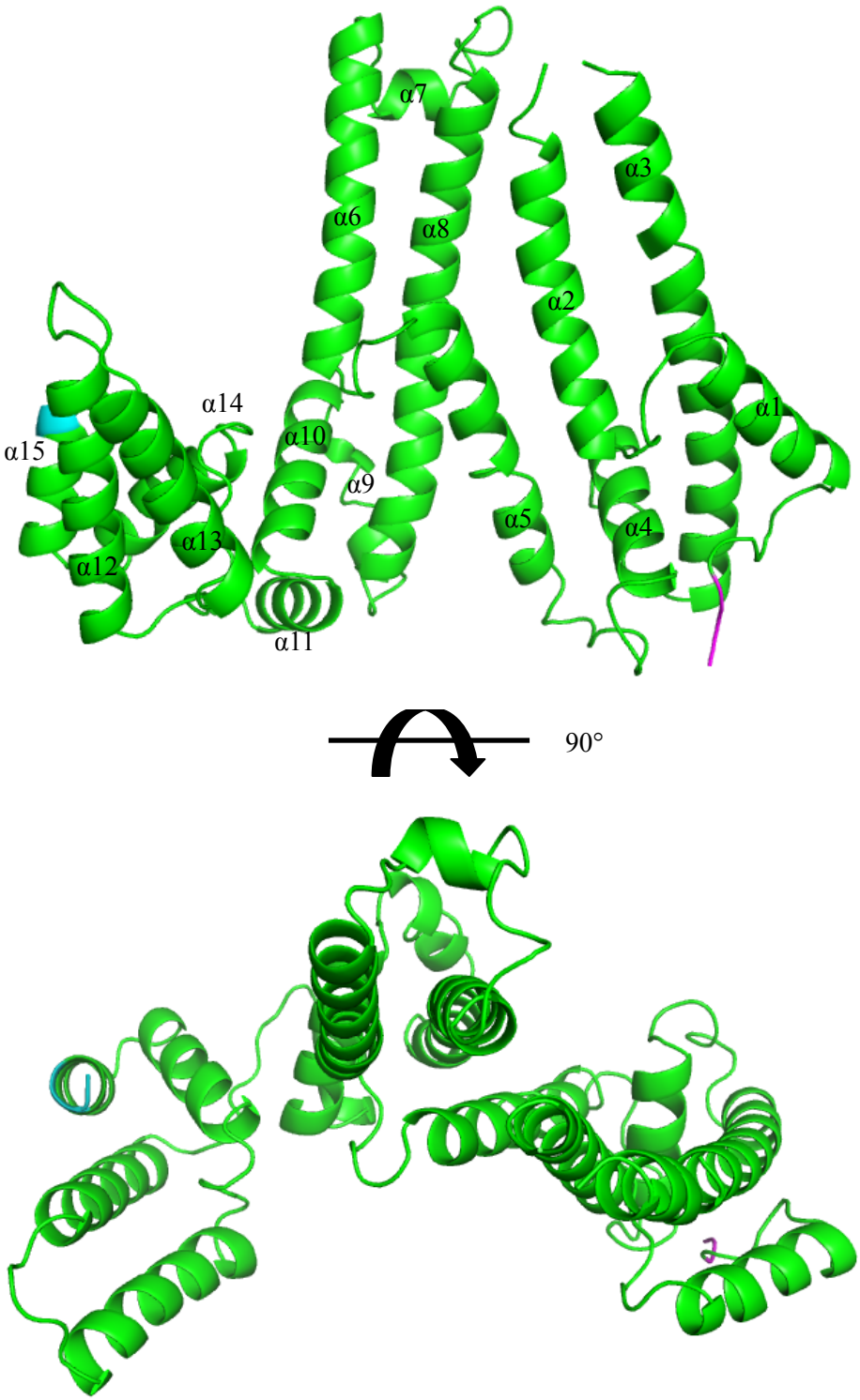


Figure 4.6 X-ray Crystal Structure of the C-terminal Domain of NES

The structure was determined to 3.0 Å. The magenta residues represent the N-term and the cyan residues represent the C-term.

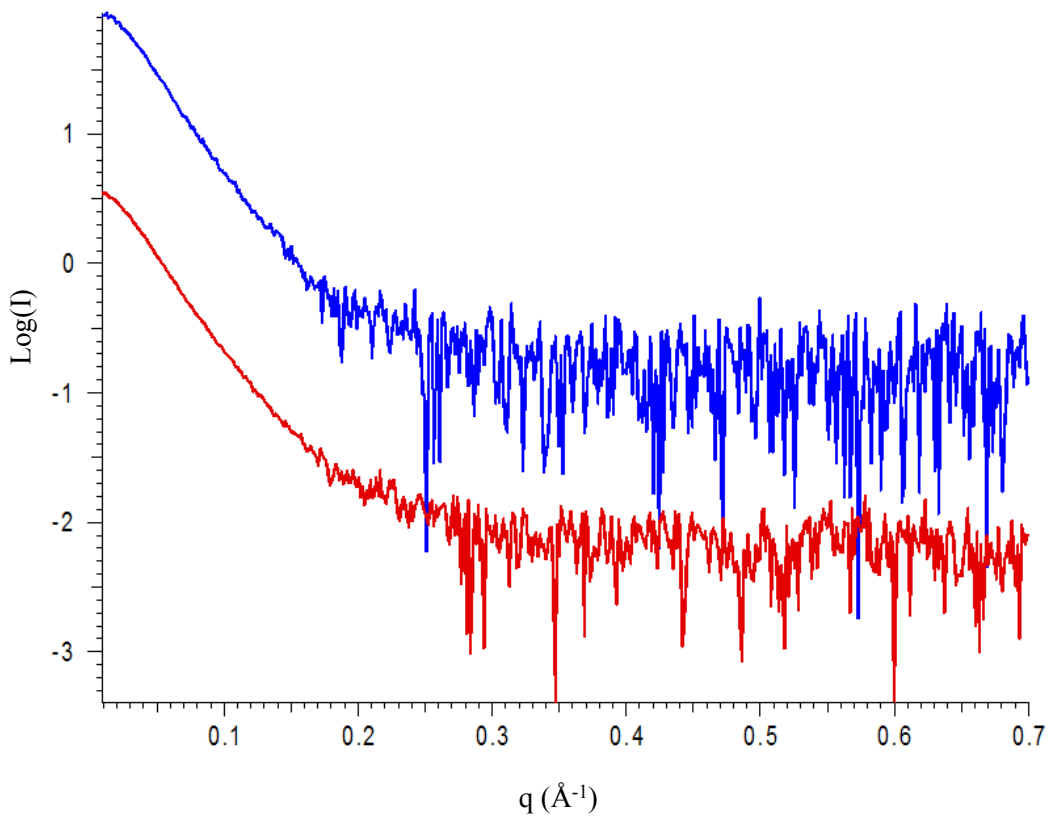


Figure 4.7 SAXS Scattering Profiles

Scattering profiles for 10.6 mg/ml (100%) NES 665 + DNA at 5 and 15 minutes at 20°C.

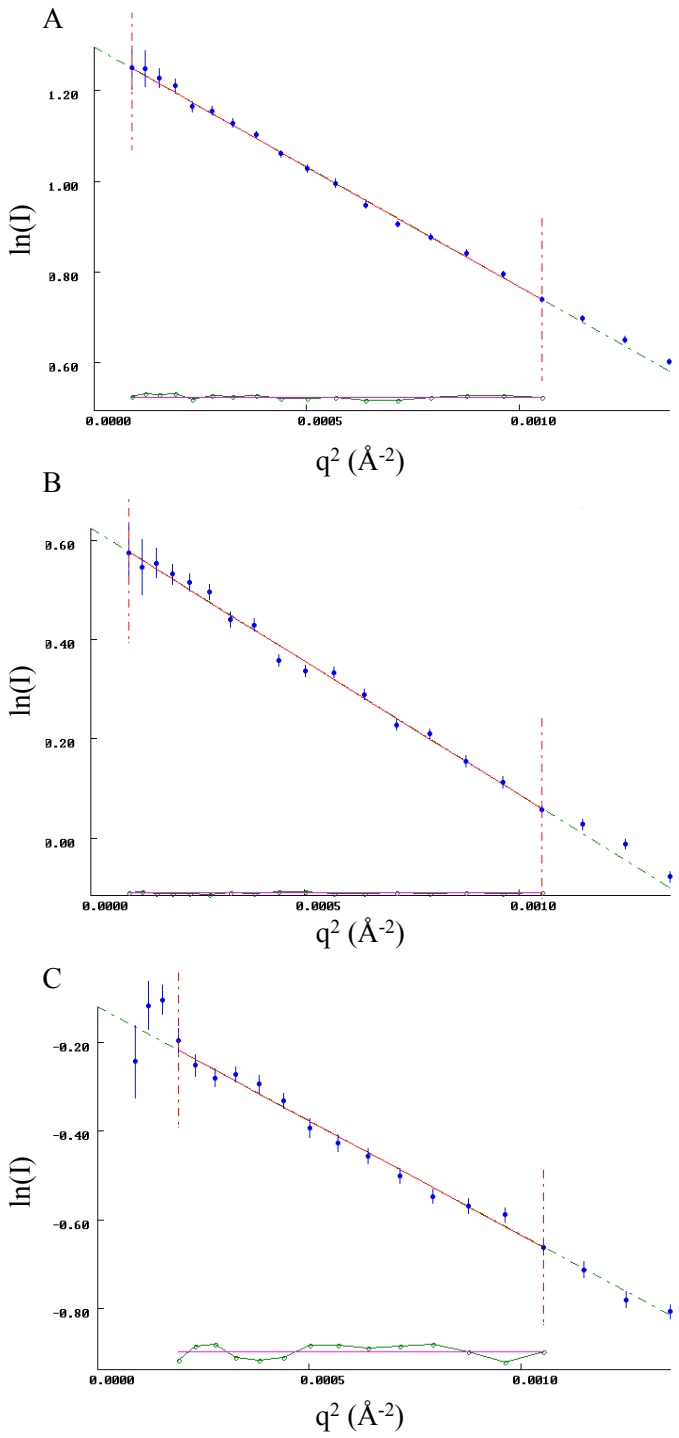


Figure 4.8 Guinier plots for all concentrations of NES 665 + DNA

Data is from a 15 min. exposure at 20° C. A. Data from 100% concentration. B. Data from 50% concentration. C. Data from 25% concentration.

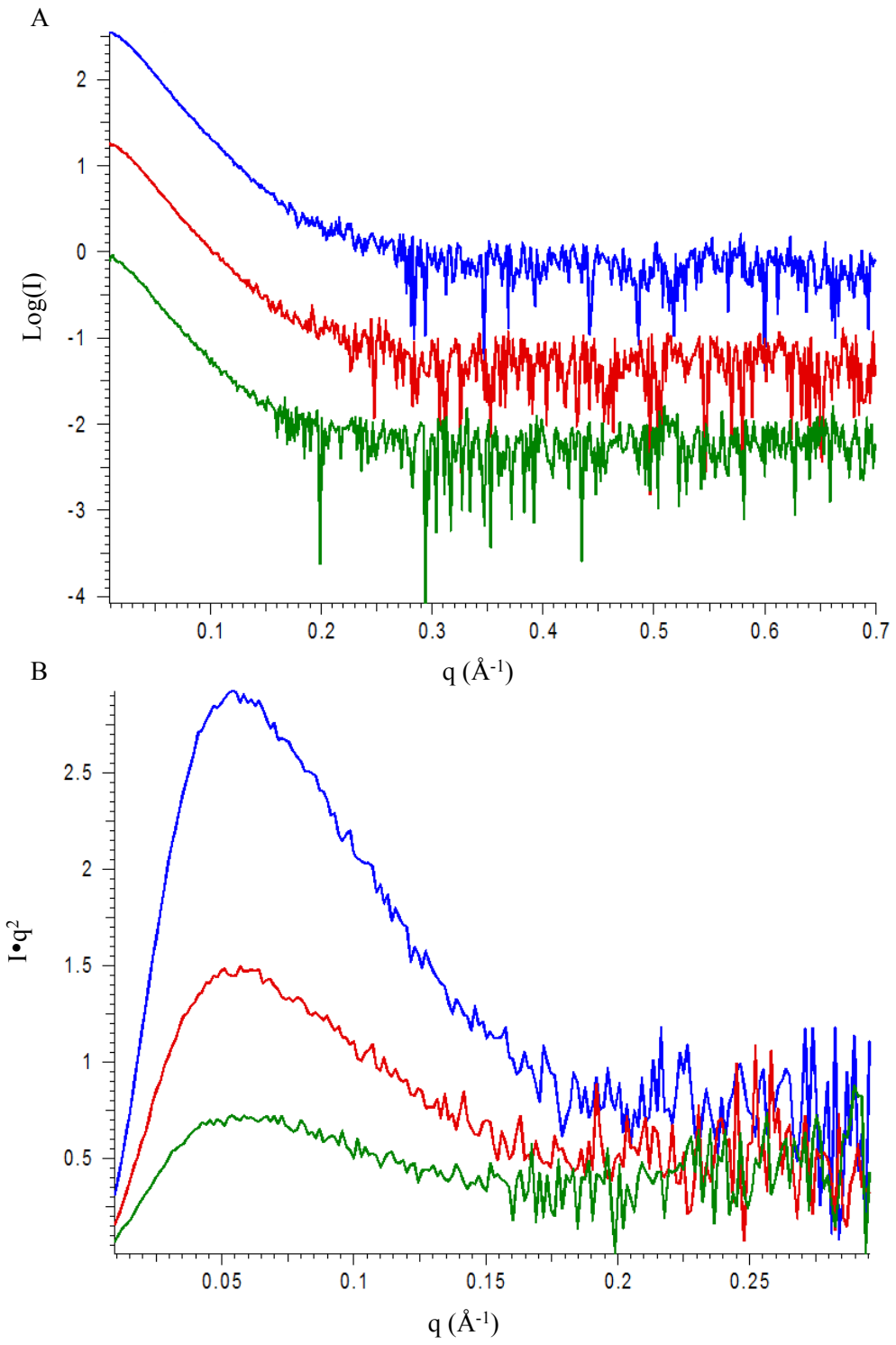


Figure 4.9 Scattering Profiles and Kratky Plots for NES 665 + DNA

Data from 3 concentrations of NES 665 + DNA with a 15 min. exposure time at 20°C. A. Scaled scattering profiles. B. Kratky plots for each concentration.

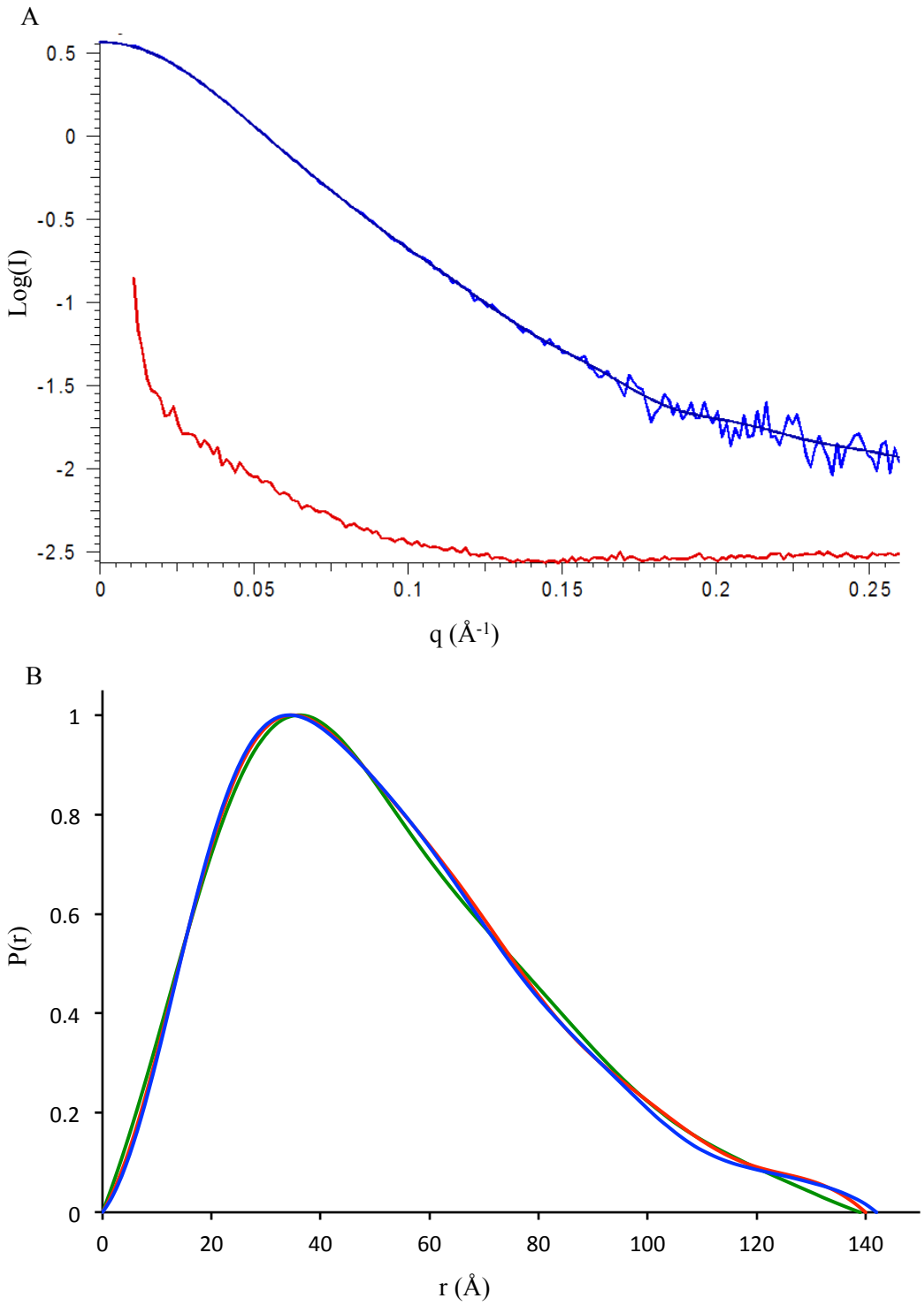


Figure 4.10 Fits to experimental data from GNOM and PDDF plots for NES 665 + DNA

A . GNOM fit to experimental data at the 100% concentration. B. PDDF plots for all concentrations with the Y-axis normalized for visual comparison.

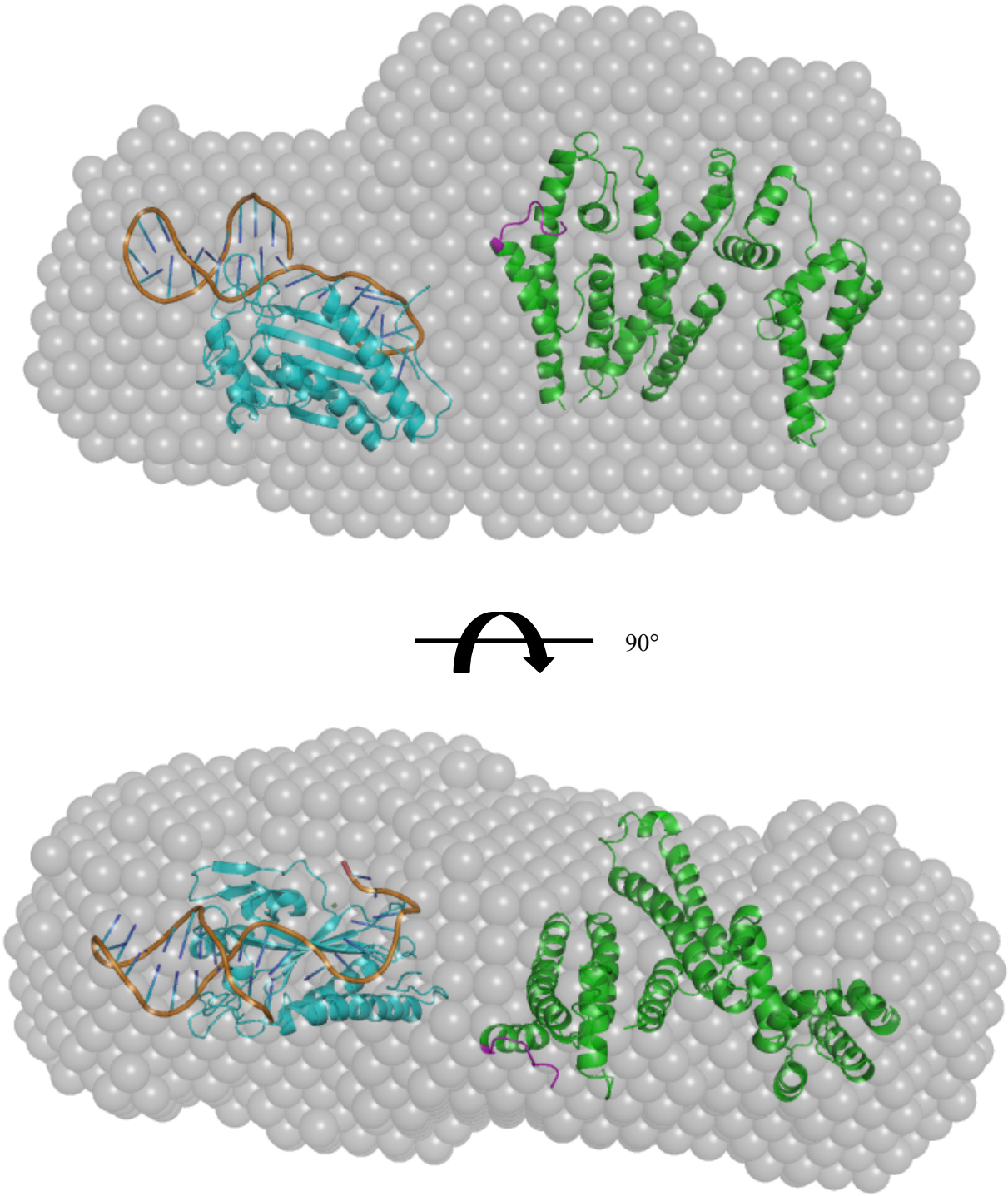


Figure 4.11 Averaged SAXS Molecular Envelope for NES 665 + DNA

The averaged molecular envelope with the structures of both domains of NES manually fit into the shell.

4.10 References

1. Cheng, Y., McNamara, D. E., Miley, M. J., Nash, R. P. & Redinbo, M. R. Functional characterization of the multidomain F plasmid TraI relaxase-helicase. *J Biol Chem* **286**, 12670-12682, doi:10.1074/jbc.M110.207563 (2011).

2. Henderson, D. & Meyer, R. The MobA-linked primase is the only replication protein of R1162 required for conjugal mobilization. *J Bacteriol* **181**, 2973-2978 (1999).

3. Llosa, M., Grandoso, G., Hernando, M. A. & de la Cruz, F. Functional domains in protein TrwC of plasmid R388: dissected DNA strand transferase and DNA helicase activities reconstitute protein function. *J Mol Biol* **264**, 56-67, doi:10.1006/jmbi.1996.0623 (1996).

4. Nash, R. P., Habibi, S., Cheng, Y., Lujan, S. A. & Redinbo, M. R. The mechanism and control of DNA transfer by the conjugative relaxase of resistance plasmid pCU1. *Nucleic Acids Res* **38**, 5929-5943, doi:10.1093/nar/gkq303 (2010).

5. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J. & Kelley, L. A. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **70**, 611-625 (2008).

6. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol* **276**, 307-326 (1997).

7. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

8. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

9. Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I., Craievich, A. F.,. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J Appl Crystallogr*, 101-109, doi:doi:10.1107/S0021889809043076 (2010).

10. Franke, D., Svergun, Dmitri I., . DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342-346, doi:doi:10.1107/S0021889809000338 (2009).

11. Volkov, V. V., Svergun, Dmitri I.,. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* **36**, 860-864, doi:doi:10.1107/S0021889803000268 (2003).

12. Kozin, M. B., Svergun, D. I., . Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **34**, 33-41, doi:doi:10.1107/S0021889800014126 (2001).

Chapter 5

Immobilization of Active Human Carboxylesterase 1 in Biomimetic Silica Nanoparticles*

5.1 Introduction

Sol-gel and biomimetic reactions have facilitated the use of immobilized enzymes in a range of applications [1]. For example, several recent studies have employed the biomineralization of silica to encapsulate a variety of enzymes. This reaction occurs rapidly at room temperature using tetramethylorthosilicate (TMOS) and a polycationic catalyst to promote silica formation around the target enzyme. Proteins encapsulated using this approach can be used to catalyze highly specific reactions, to screen for enzyme inhibitors, or to detect the presence of chemical compounds. Indeed, silica particles containing enzymes have been employed in biosensing applications, including microfluidic devices and aerosol detectors [2-3].

There are several advantages to using silica particles for encapsulation, including the fact that the mild reaction conditions promote the capture of active enzymes, and, in some cases, increase enzyme stability [4]. The simplicity and speed of the reaction also allow for rapid encapsulation and the cost-effective production of the particles for various needs. For example, encapsulated enzymes could be used in a column format to monitor for chemical contaminants in water sources. However, encapsulation efficiency varies with the catalyst-enzyme pair [5]. Thus, optimal encapsulation conditions must be identified for each enzyme before they can be applied in a useful setting.

One promising application of encapsulated protein is the detection of chemical compounds, such as organophosphates (OP). OPs are employed as agricultural pesticides, but

---

*Edwards, J. S. *et al.* Immobilization of active human carboxylesterase 1 in biomimetic silica nanoparticles. *Biotechnol Prog* **27**, 863-869, doi:10.1002/btpr.604 (2011).

have also been weaponized for use as highly toxic nerve agents like sarin, soman, and cyclosarin [6-7]. OP toxicity arises from the noncompetitive inhibition of acetylcholinsterase (AChE), which produces sustained acetylcholine stimulus of the muscarinic and nicotinic receptors and may lead to diaphragm incapacitation resulting in suffocation. Biosensors containing enzymes inhibited by OPs have recently been explored for military and civilian applications [2,8,9].

The human liver carboxylesterase 1 (hCE1) is a promiscuous serine hydrolase that catalyzes the hydrolysis of endogenous molecules and xenobiotics containing ester, amide, or thioester bonds [10-17]. This structural AChE homolog is currently being explored as a prophylactic bioscavenger to decontaminate OP toxicity [18]. Based upon the ability of hCE1 to bind Ops [15,18], we sought to examine the silica encapsulation of hCE1 and the ability of the resultant nanoparticles to detect the presence of OP toxicants.

5.2 Methodology

Expression and purification of hCE1. A secreted form of hCE1 was expressed using baculovirus infection of *Spodptera frugiperda* Sf21 cells and purified as described [19]. hCE1 was concentrated in 50 mM HEPES (pH 7.4) and stored at 4 ^{o}C. Protein concentration was determined spectrophotometrically using an extinction coefficient 17700 M^{-1} cm^{-1} (NanoDrop, ND-1000 Spectrophotometer).

Preparation of silica nanoparticles with hCE1. All materials were purchased from Sigma-Aldrich unless otherwise noted. A stock solution of lysozyme at 100 mg/mL was prepared in 0.1 M potassium phosphate buffer (pH 7.4). A solution of hydrolyzed tetramethyl orthosilicate (TMOS) was freshly prepared by diluting into 1 mM hydrochloric acid to a final concentration of 1 M. A typical reaction mixture consisted of 800 µL of 0.1 M potassium phosphate buffer containing 5 µg of hCE1, 100 µL of lysozyme or R5 solution (100 mg/mL), and 100 µL of 1 M hydrolyzed TMOS. The mixture was incubated at room temperature (22° C) for 10 minutes with gentle shaking. The slurry was centrifuged for 60 s (13,000 rpm) to remove the particles and washed in ddH_2O three times. Control particles were produced in a similar manner using a

reaction mixture without hCE1 by combining 800 μL of 0.1 M potassium phosphate buffer, 100 μl of lysozyme solution, and 100 μL of hydrolyzed TMOS. Combination particles were created using ratios of polyethyleneimine (1%) to lysozyme (100 mg/mL) 1:4, 1:1, 4:1 as the catalyst to compare the efficiency of enzyme incorporation. The total volume of catalyst used was always equal to 10% of the silica reaction.

Scanning and transmission electron microscopy. The morphology of silica-encapsulated hCE1 with various catalyst was observed with a Hitachi 4700 Scanning Electron Microscope (SEM) and a JEOL 2010F Transmission Electron Microscope (TEM). SEM samples were prepared by depositing dry powdered samples on a double-stick carbon tape on an aluminium SEM sample holder. The samples were coated with a 3nm Au-Pd conductive coating in order to prevent charging effect during imaging. TEM samples were prepared by air-drying a drop of concentrated reaction mixture on a carbon-stabilized copper grid (Ted Pella Inc.).

Determination of hCE1 enzyme activity. Activity of silica-encapsulated or solution hCE1 were monitored using para-nitrophenyl butyrate (pNPB) as a substrate and measuring the absorbance of the para-nitrophenol product at 410 nm using a Pherastar spectrophotometer (BMG Labtech). The substrate was dissolved in ethanol to a concentration of 1 M, and then diluted to 10 mM in 0.1 M phosphate buffer pH 7.4. hCE1 silica particles were mixed with substrate at a final concentration of 5 mM. Supernatants from the washes were saved and monitored in the same fashion to determine the incorporation efficiency of the silica reactions.

Kinetics of para-nitrophenyl butyrate hydrolysis. K_m, V_{max}, and k_{cat} values were determined by monitoring the conversion of pNPB to para-nitrophenol at 410 nm using a Pherastar spectrophotometer (BMG Labtech). Substrate was prepared as described above. Data was collected over 4 minutes and the linear portions were used for the linear regressions to calculate the reaction velocities. Kinetic parameters were determined by plotting the reaction velocities versus substrate concentration. An extinction coefficient of 17700 was used for para-nitrophenol in calculating the Vmax. The molecular weight for hCE1 that was used to calculate

the k_{cat} was 61.68 kDa. The Prism software packaged was used to fit the data (GraphPad, San Diego, CA).

Column-based experiments. Column activity of silica-encapsulated hCE1 was determined using a stainless steel microbore column (2 cm x 2 mm, Upchurch Scientific, Oak Harbor, WA) with 0.5 µm frits. The silica nanoparticle reaction mixture was scaled up to a total of 10 mL. The particles were resuspended in 0.1 M phosphate buffer pH 7.4 and pumped at a flow rate of 10 µL/min to pack the column. The column was washed with 10 column volumes of phosphate buffer prior to use. All column experiments were performed at room temperature (22 ^{o}C) at a flow rate of 100 µL/min using a NE-1000 syringe pump (New Era Pump Systems, Wantagh, NY). Phosphate buffer containing 5 mM substrate was flowed through the column. Samples were collected and analyzed in a Pherastar spectrophotometer to measure absorbance at 410 nm until the substrate hydrolysis had reached a maximum.

Organophosphate pesticide detection. Using a column constructed in the manner described above, phosphate buffer containing 5 mM substrate was applied and absorbance was measured until the substrate hydrolysis reached its maximum value. Phosphate buffer containing 5 mM substrate and paraoxon or DMPNPP (1mM) was then flowed through the column, and absorbance was measured at 410 nm to determine if OP was detected. A drop in absorbance at 410 nm correlates with enzyme inhibition because the product of the enzyme assay, para-nitrophenol, absorbs at 410 nm. The pesticides were diluted 1:10 in methanol and further diluted in 0.1 M phosphate buffer pH 7.4 to their final concentration. Inhibition was confirmed by flowing phosphate buffer containing 5 mM substrate through the column and monitoring absorbance at 410 nm.

5.3 Results

5.3.1 Silica Encapsulation of Active hCE1

Initial experiments employed lysozyme as the polycation catalyst for silica nanoparticle formation in reaction conditions containing hydrolyzed tetramethylorthosilicate (TMOS),

lysozyme (100 mg/mL), and potassium phosphate buffer at pH 7.4 with 5 µg of hCE1 enzyme. hCE1 was incorporated into the particles with high efficiency, leaving no enzyme activity in the supernatant (Figure 5.1.A). In addition, no hCE1 activity was detected in solutions used to wash the particles after formation (Figure 5.1.A). Thus, while the process may render some hCE1 inactive, all detectable enzyme activity localized solely to the particles. hCE1 activity in the silica nanoparticles, as measured by para-nitrophenyl butyrate (pNPB) hydrolysis, reached the same maximal velocity as 5 µg of hCE1 in solution; however, the encapsulated enzyme exhibited a slightly reduced initial rate of product formation (Figure 5.1.A, Table 5.1). Multiple washes of silica particles with phosphate buffer did not release active enzyme, indicating that the hCE1 activity associated with the particles did not arise from enzyme adhered relatively loosely to particle surfaces. Similarly, several washes of particles in 1% Tween 20 did not dissociate hCE1 activity from the particles (data not shown). When washed in detergent, the hCE1 activity in the particles was comparable to the activity measured from particles washed in buffer alone (Table 5.1).

We examined the effect of entrapping hCE1 in silica on the enzyme's ability to hydrolyze pNPB. Kinetic experiments were performed maintaining a constant amount of enzyme with varied substrate concentrations to determine initial reaction velocities. Michaelis-Menten kinetics was assumed and the kinetic parameters were determined (Table 5.2). hCE1 in silica had a K_m that was approximately three times higher than enzyme in solution. Furthermore, hCE1 in silica exhibited a slightly reduced efficiency ($k_{cat}/K_m = 3.1 \times 10^4\ s^{-1}M^{-1}$) relative to free enzyme ($k_{cat}/K_m = 1.9 \times 10^4\ s^{-1}M^{-1}$). These solution values are in agreement with previously reported kinetic parameters for hCE1 with para-nitrophenyl butyrate (pNPB) [20]. Taken together, these results demonstrate that active hCE1 can be efficiently incorporated into lysozyme-catalyzed silica nanoparticles.

We also investigated the ability of other cationic catalysts to generate hCE1-encapsulated silica nanoparticles. We created silica particles in the presence of hCE1using the following

catalysts: cationic peptide R5 (NH_2-SSKKSGSYSGSKGSKRRIL-CO_2H)[21] and the polymer polyethyleneimine (PEI). We found that silica particles formed with the R5 peptide exhibited similar incorporation efficiency and enzyme activity as the lysozyme hCE1-silica particles (Figure 5.1.B, Table 5.1). In contrast, PEI catalyzed particles, which formed immediately upon mixing (compared to ~5 minutes with the lysozyme or R5 silica particles) excluded hCE1 into the supernatant and thus had no associated enzyme activity with the particles. (Figure 5.1.C; Table 5.1).

We hypothesized that the lack of entrapped hCE1 during PEI-mediated particle creation was caused by either the rapid particle formation or catalyst size. That hypothesis was tested by creating beads using different ratios of lysozyme and PEI in the presence of the same amount of hCE1 (80% lysozyme:20% PEI, 50% lysozyme:50% PEI, and 20% lysozyme:80% PEI). It was determined that increasing the amount of lysozme relative to PEI improved the encapsulation of hCE1, but that no silica particles formed in the presence of PEI exhibited higher levels of entrapped hCE1 activity than those formed without PEI (Figure 5.1.D-F, Table 5.1). We also concluded that hCE1 alone could not facilitate the precipitation of silica particles without either lysozyme or the R5 peptide (data not shown). The reaction was cooled to 4° C to see if the silica condensation could be slowed in the presence of PEI. The reaction appeared to occur just as fast and no hCE1 was incorporated into the particles (data not shown). Thus, we conclude that macromolecular catalysts as small as the R5 peptide and as large as hen egg white lysozyme are capable of generating silica nanoparticles that entrap active hCE1.

5.3.2 Electron Microscopy of hCE1 Nanoparticles

We employed scanning electron microscopy (SEM) and transmission electron microscopy (TEM) to examine the size and morphology of silica particles produced in the presence of hCE1. Silica hCE1 particles using lysozyme as the catalyst were larger than the particles obtained using either R5 or PEI as catalysts. Lysozyme-catalyzed particles were 0.8-1 μm in diameter and appeared to fuse to form large aggregates on the order of ~ 2 μm (Figure

5.2.A). In contrast, silica particles obtained in the presence of hCE1 using R5 and PEI as catalysts formed individual particles 0.3-0.8 μm in diameter and an aggregate size of ~ 1 μm (Figure 5.2.B-C). Dynamic light scattering (DLS) performed on the lysozyme catalyzed silica particles confirmed the sizes measured from SEM analysis (data not shown). The silica particles formed using the combinations of lysozyme and PEI were smaller (~ 0.5 μm) and appear monodisperse as compared to the particles using only lysozyme or PEI as catalysts (Figure 5.2.D-F). Increasing amounts of PEI in the reaction lead to the formation of smaller particles and smaller aggregate size.

TEM images corroborated the SEM data (Figure 5.3.A-B). For example, particles obtained in the presence of lysozyme were larger in size compared to that of the PEI-catalyzed silica particles. Further, the TEM images revealed that the lysozyme-catalyzed silica particles exhibited a smoother surface than that of the PEI-catalyzed particles. In contrast, the particles obtained using ratios of lysozyme and PEI as catalysts, appeared to have rougher surface features similar to that seen for the PEI catalyzed particles. This was presumably due to the presence of PEI on the particle surface (Figure 5.3.C-E).

5.3.3 Solution Detection Studies of Pesticides with hCE1 Nanoparticles

We examined whether silica nanoparticles containing active hCE1 could function in a column chromatography format. Biosilica particles formed with lysozyme and hCE1 (Figure 5.1.A, Figure 5.2.A) were packed using buffer flowed into a stainless steel microbore column (2 cm x 2 mm) with 0.5 μm frits controlled by a New Era single syringe pump (New Era Pump Systems, Wantagh, NY). Particles catalyzed with lysozyme in the presence of hCE1 were chosen because they were cost-effective and displayed robust entrapment efficiency and enzyme activity. After extensive washing of the column, 5 mM pNPB was applied to the column and para-nitrophenol formation was observed, indicating that hCE1 retains its enzyme activity in the silica particles packed within the column (Figure 5.4). No hCE1 activity was detected in the solutions

used to wash the column, supporting the conclusion that the enzyme remained encapsulated in the column particles.

To determine whether columns containing hCE1 biosilica nanoparticles were capable of detecting organophosphate pesticides, we added the pesticides and monitored substrate conversion to para-nitrophenol. Columns were operated in the manner described above, and equilibrated such that a stable amount of para-nitrophenol product was being formed over time (e.g. 10 μmol/min.). At that point, 1 mM diethyl para-nitrophenyl phosphate (paraoxon) was applied to a column, which produced a dramatic drop in enzyme activity resulting in decreased substrate conversion (Figure 5.4.A). Detection of the presence of this hCE1 inhibitor occurs rapidly, likely due to the small column volume of 16.2 mL. hCE1 remains permanently inhibited by paraoxon even 50 minutes (~70 column volumes) after this inhibitor is removed from the substrate-containing solution applied to the column (Figure 5.4.A). In a second study, dimethyl-*para*-nitrophenylphosphate (DMPNPP) was applied to a column instead of paraoxon. In this case, detection of the hCE1 inhibitor was rapid, but once the inhibitor was removed, hCE1 activity returned over time (Figure 5.4.B). Paraoxon inhibits hCE1 activity 3-fold more robustly in solution than DMPNPP (with bimolecular rates of inhibition [k_i] values of $9.1 \pm 0.5 \times 10^6$ min^{-1} M^{-1} and $2.9 \pm 0.3 \times 10^6$ $min^{-1}M^{-1}$, respectively) [18]. Taken together, these results indicate that OP pesticides can be detected using hCE1 biosilica nanoparticles within a column chromatography format.

5.4 Conclusions

Silica nanoparticles have the potential to generate a range of platforms for biodetection. We have shown that hCE1 can be effectively encapsulated into silica nanoparticles catalyzed by lysozyme, and that these particles maintain overall enzyme activity comparable to hCE1 in solution. In our studies, however, lysozyme-encapsulated hCE1 exhibited a slightly reduced initial rate of product formation, although activity reached that of free hCE1 over time. We hypothesize that this effect is caused by impeded substrate access to the active site of the enzyme

in the silica matrix. We also found that hCE1 is more efficiently incorporated into silica nanoparticles using lysozyme as a catalyst in comparison to PEI catalyzed silica particles. Similar results have been observed with other enzymes during encapsulation [5].

We focused on three catalysts – lysozyme, the R5 peptide, and PEI – which have been demonstrated to be successful in previous studies with other target enzymes [22-24]. Based on our results, we hypothesize that the lack of entrapped hCE1 during PEI-mediated particle creation was caused by the rapid particle formation and the small size of PEI. For example, when lysozyme was used as a catalyst, the reaction proceeded slowly taking approximately 2 minutes until particles can be visualized in solution; in contrast, when PEI was employed, the particles formed immediately. This observation, combined with the smaller sized PEI-catalyzed particles relative to those created with lysozyme suggested that PEI-generated silica matrixes excluded hCE1 under the reaction conditions employed. The lysozyme-catalyzed particles underwent more complete condensation similar to that seen for pure sol-gel silica reactions with other enzymes[25].

Finally, we demonstrated that silica nanoparticles containing hCE1 are able to detect OP pesticides in a column chromatography format. The drop in hCE1 activity correlated with the strength of the inhibitor employed; for example, paraoxon, a more potent hCE1 inhibitor, did not allow enzyme activity to recover during the time course of the study. It is possible that such an apparatus could be employed in the detection of OP pesticides or chemical weapons agents using silica nanoparticles containing hCE1. Both liquid- and gas-based detection have already been reported involving other enzymes, including the cholinesterases [2]. Our data indicate that another member of the serine hydrolase family of enzymes may also be employed for this or related technologies.

Sample	**hCE1 Activity** (moles of nitrophenol/second)	R^2
hCE1 in solution	3.42×10^{-7}	0.99
Lysozyme	2.74×10^{-7}	0.99
R5	2.90×10^{-7}	0.99
PEI	NA	
80% Lysozyme, 20% PEI	2.52×10^{-7}	0.99
50% Lysozyme, 50% PEI	3.03×10^{-8}	0.99
20% Lysozyme, 80% PEI	NA	
Lysozyme washed with 1% Tween 20	1.50×10^{-7}	0.99
hCE1 Solution with 1% Tween 20	6.04×10^{-8}	0.99

Table 5.1 hCE1 rate of pNBP hydrolysis in solution compared to silica particles created with different catalysts. NA: not applicable, no enzyme activity contained in the particles.

K_m (mM)	3.30 ± 0.44	10.13 ± 0.85
V_{max} (µmole/min/mg)	97.71 ± 4.84	189.72 ± 7.41
Curve Fit (r^2)	0.99	0.99
k_{cat}/K_m (s^{-1} M^{-1})	6.09 x 10^4	3.85 x 10^4

Table 5.2 Kinetic parameters for the metabolism of pNPB by hCE1 in solution and in lysozyme catalyzed silica particles. N=3, S.E. for K_m and V_{max} values.

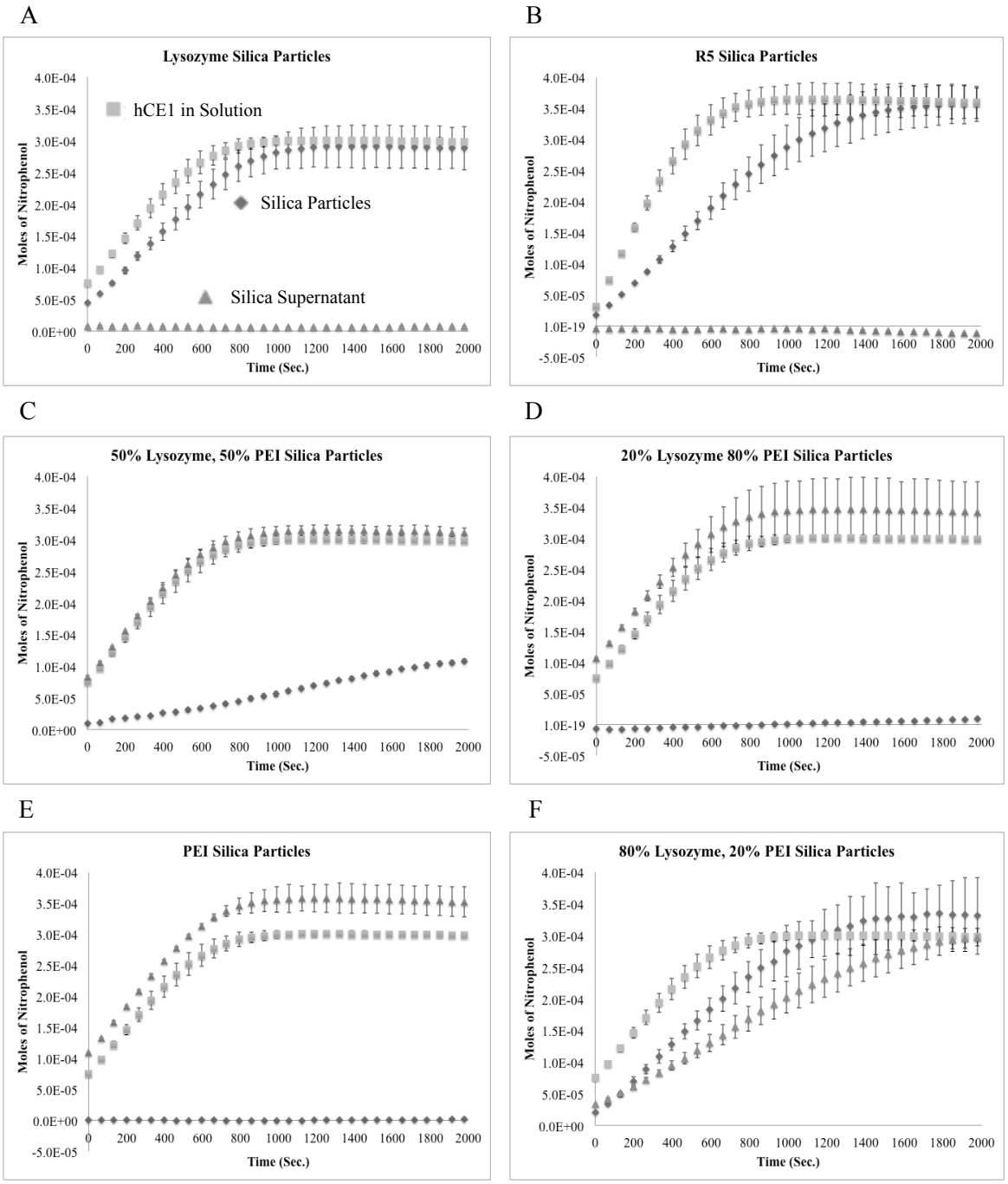


Figure 5.1 hCE1 Activity in Silica Nanoparticles

Activity assays measuring the formation of para-nitrophenol at 410 nm. Silica particles were formed using the respective catalysts and the absorbance was measured over time. Activity was measured for the silica particles and the supernatant from the reaction, then compared to hCE1 in solution. All data points had an N = 3 and S.E. is reported.

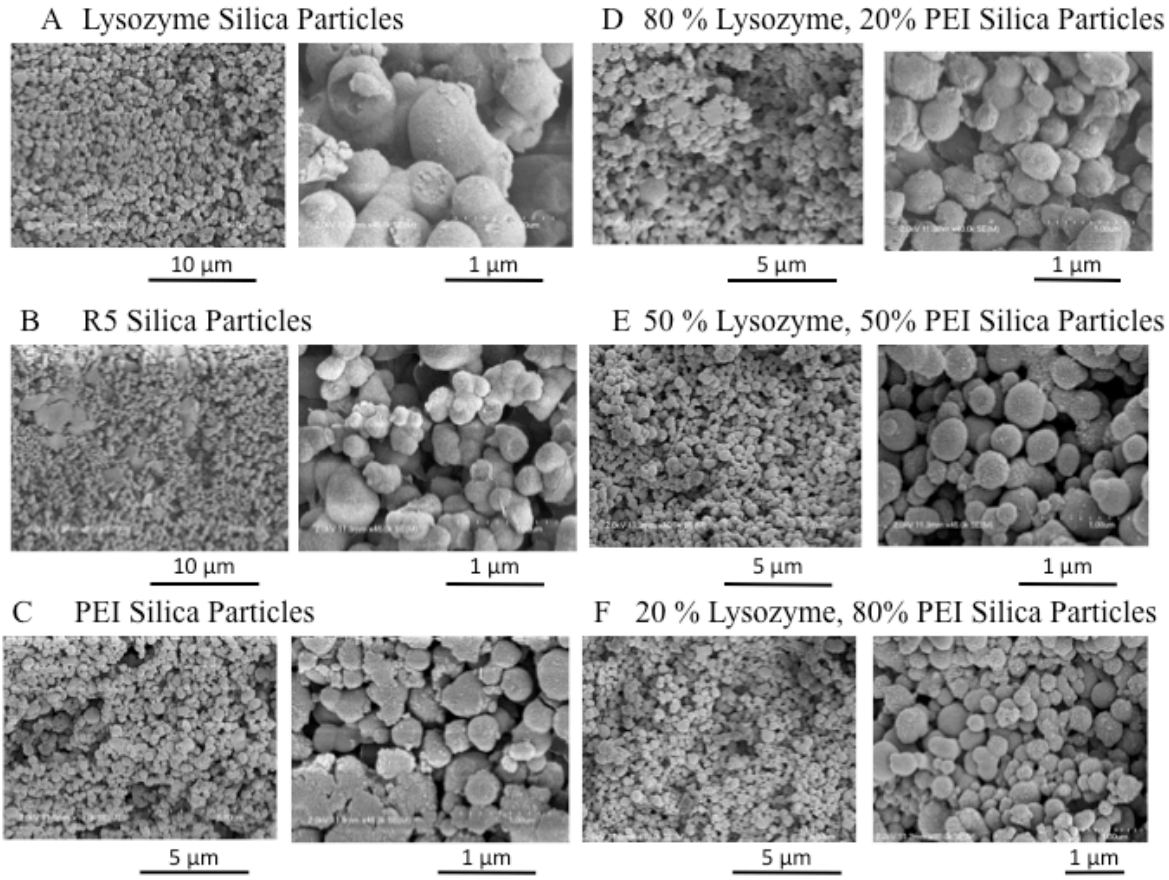


Figure 5.2 Kinetic Parameters for the Metabolism of pNPB by hCE1

SEM images of the silica particles containing hCE1 at two magnifications. The samples were sputtered with Au-Pd conductive coating prior to analyses.

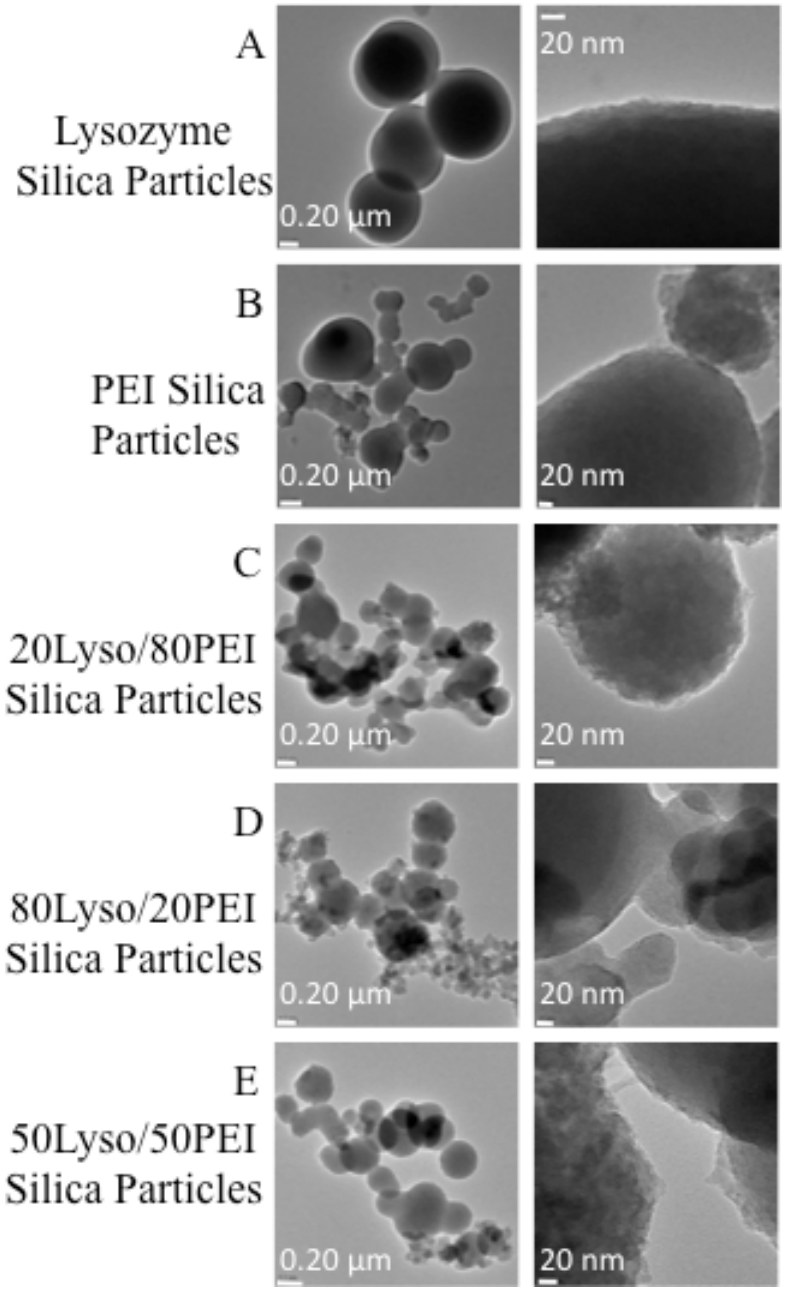
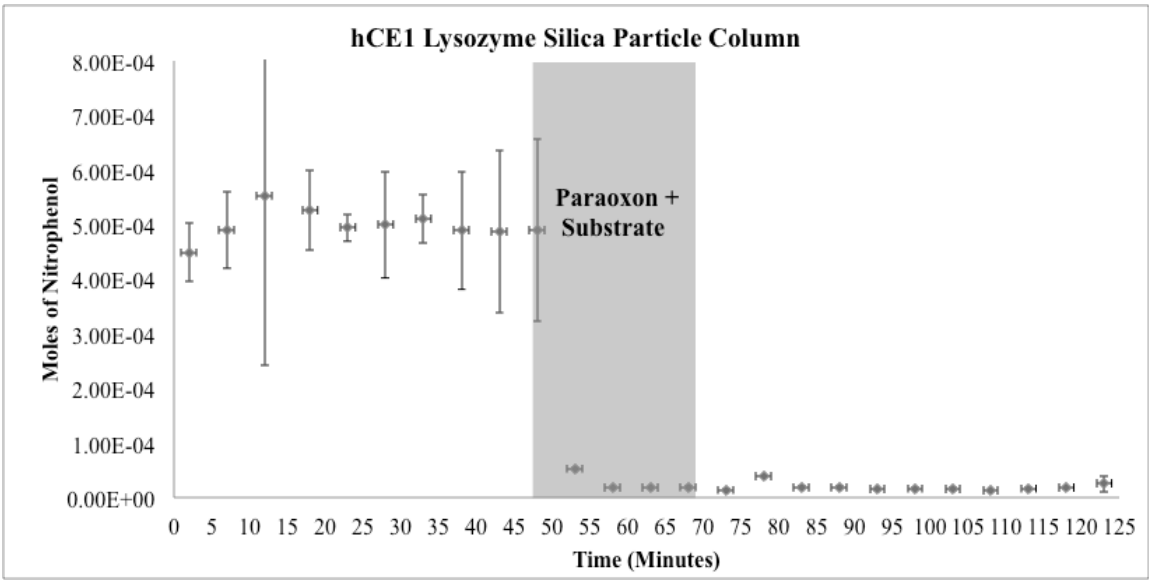


Figure 5.3 hCE1 TEM Images of Silica Nanoparticles

TEM images of the hCE1 encapsulated silica particles at two magnifications. Particles were air dried on a carbon stabilized copper grid prior to imaging.

A



B

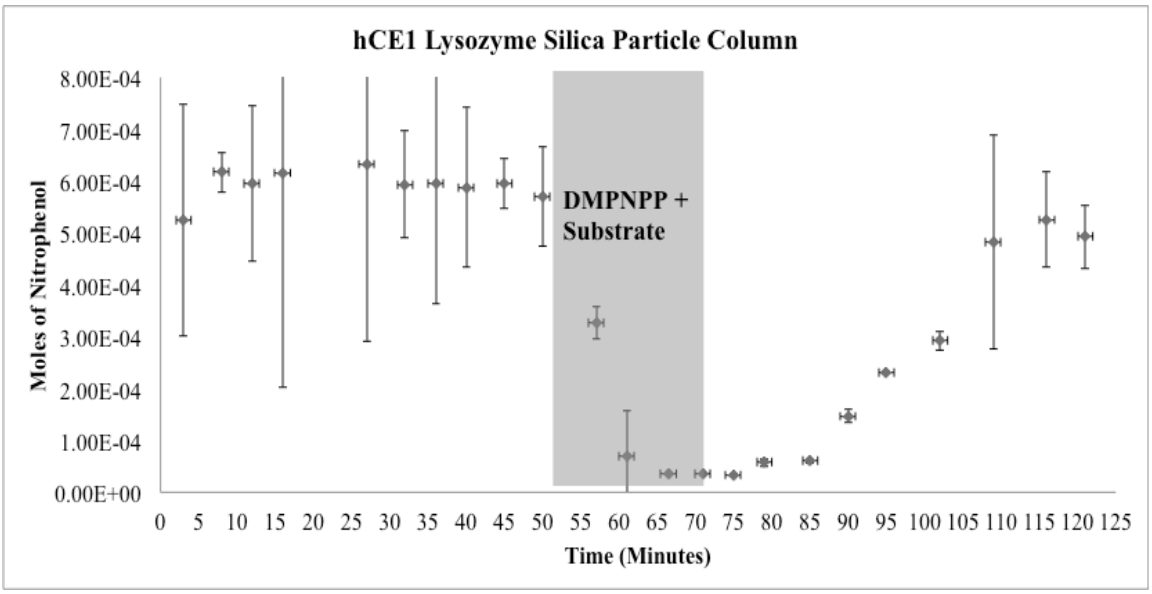


Figure 5.4 Microbore Column Detection of Pesticides

Microbore columns packed with lysozyme catalyzed silica particles containing hCE1 at a concentration of 8.065 x 10^{-2} nmol used to detect pesticides. The columns were equilibrated at 25° C to maximum activity with 5 mM substrate before 1 mM pesticide was applied (grey box). The eluent from the column was collected and measured in triplicate at the respective time points. All data points had an N = 3 and S.E. is reported.

5.5 References

1. Avnir D, Coradin T, Lev O, Livage J. Recent bio-applications of sol-gel materials. *J of Mat Chem* 2006;16(11):1013-1030.

2. Luckarift HR, Greenwald R, Bergin MH, Spain JC, Johnson GR. Biosensor system for continuous monitoring of organophosphate aerosols. *Biosens Bioelectron* 2007;23(3):400-6.

3. Luckarift HR, Ku BS, Dordick JS, Spain JC. Silica-immobilized enzymes for multi-step synthesis in microfluidic devices. *Biotechnol Bioeng* 2007;98(3):701-5.

4. Luckarift HR, Spain JC, Naik RR, Stone MO. Enzyme immobilization in a biomimetic silica support. *Nat Biotechnol* 2004;22(2):211-3.

5. Betancor L, Luckarift HR. Bioinspired enzyme encapsulation for biocatalysis. *Trends Biotechnol* 2008;26(10):566-72.

6. Aardema H, Meertens JH, Ligtenberg JJ, Peters-Polman OM, Tulleken JE, Zijlstra JG. Organophosphorus pesticide poisoning: cases and developments. *Neth J Med* 2008;66(4):149-53.

7. Newmark J. Nerve agents. *Neurologist* 2007;13(1):20-32.

8. Luckarift HR, Johnson GR, Spain JC. Silica-immobilized enzyme reactors; application to cholinesterase-inhibition studies. *J Chromatogr B Analyt Technol Biomed Life Sci* 2006;843(2):310-6.

9. Ramanathan M, Luckarift HR, Sarsenova A, Wild JR, Ramanculov EK, Olsen EV, Simonian AL. Lysozyme-mediated formation of protein-silica nano-composites for biosensing applications. *Colloids Surf B Biointerfaces* 2009;73(1):58-64.

10. Bencharit S, Edwards CC, Morton CL, Howard-Williams EL, Kuhn P, Potter PM, Redinbo MR. Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1. *J Mol Biol* 2006;363(1):201-14.

11. Bencharit S, Morton CL, Howard-Williams EL, Danks MK, Potter PM, Redinbo MR. Structural insights into CPT-11 activation by mammalian carboxylesterases. *Nat Struct Biol* 2002;9(5):337-42.

12. Bencharit S, Morton CL, Hyatt JL, Kuhn P, Danks MK, Potter PM, Redinbo MR. Crystal structure of human carboxylesterase 1 complexed with the Alzheimer's drug tacrine: from binding promiscuity to selective inhibition. *Chem Biol* 2003;10(4):341-9.

13. Bencharit S, Morton CL, Xue Y, Potter PM, Redinbo MR. Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme. *Nat Struct Biol* 2003;10(5):349-56.

14. Fleming CD, Bencharit S, Edwards CC, Hyatt JL, Tsurkan L, Bai F, Fraga C, Morton CL, Howard-Williams EL, Potter PM and others. Structural insights into drug processing by human carboxylesterase 1: tamoxifen, mevastatin, and inhibition by benzil. *J Mol Biol* 2005;352(1):165-77.

15. Fleming CD, Edwards CC, Kirby SD, Maxwell DM, Potter PM, Cerasoli DM, Redinbo MR. Crystal structures of human carboxylesterase 1 in covalent complexes with the chemical warfare agents soman and tabun. *Biochemistry* 2007;46(17):5063-71.

16. Redinbo MR, Bencharit S, Potter PM. Human carboxylesterase 1: from drug metabolism to drug discovery. *Biochem Soc Trans* 2003;31(Pt 3):620-4.

17. Redinbo MR, Potter PM. Mammalian carboxylesterases: from drug targets to protein therapeutics. *Drug Discov Today* 2005;10(5):313-25.

18. Hemmert AC, Otto TC, Wierdl M, Edwards CC, Fleming CD, MacDonald M, Cashman JR, Potter PM, Cerasoli DM, Redinbo MR. Human carboxylesterase 1 stereoselectively binds the nerve agent cyclosarin and spontaneously hydrolyzes the nerve agent sarin. *Mol Pharmacol* 2010;77(4):508-16.

19. Morton CL, Potter PM. Comparison of Escherichia coli, Saccharomyces cerevisiae, Pichia pastoris, Spodoptera frugiperda, and COS7 cells for recombinant gene expression. Application to a rabbit liver carboxylesterase. *Mol Biotechnol* 2000;16(3):193-202.

20. Wadkins RM, Morton CL, Weeks JK, Oliver L, Wierdl M, Danks MK, Potter PM. Structural constraints affect the metabolism of 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin (CPT-11) by carboxylesterases. *Mol Pharmacol* 2001;60(2):355-62.

21. Nam DH, Won K, Kim YH, Sang BI. A novel route for immobilization of proteins to silica particles incorporating silaffin domains. *Biotechnol Prog* 2009;25(6):1643-9.

22. Betancor L, Berne C, Luckarift HR, Spain JC. Coimmobilization of a redox enzyme and a cofactor regeneration system. *Chem Commun (Camb)* 2006;(34):3640-2.

23. Luckarift HR, Dickerson MB, Sandhage KH, Spain JC. Rapid, room-temperature synthesis of antibacterial bionanocomposites of lysozyme with amorphous silica or titania. *Small* 2006;2(5):640-3.

24. Naik RR, Tomczak MM, Luckarift HR, Spain JC, Stone MO. Entrapment of enzymes and nanoparticles using biomimetically synthesized silica. *Chem Commun (Camb)* 2004; (15):1684-5.

25. Gao Y, Heinemann A, Knott R, Bartlett J. Encapsulation of protein in silica matrices: structural evolution on the molecular and nanoscales. *Langmuir* 2010;26(2):1239-46.

Chapter 6

Nerve Agent Hydrolysis Activity Designed into a Human Drug Metabolism Enzyme*

6.1 Introduction

Organophosphorus (OP) nerve agents are some of the most poisonous chemicals known (Figure 6.1.A) [1]. By covalently phosphonylating the catalytic serine of the neurotransmitter-regulating enzyme acetylcholinesterase (AChE), OPs produce a cholinergic crisis that can cause respiratory failure and death [2]. OP nerve agents have been used as chemical weapons by both established governments and terrorist groups [2,3]. In addition, there are an estimated 250,000 deaths worldwide each year associated with OP pesticides [4]. Thus, there is significant interest in developing novel approaches to detoxify these compounds for military, security and clinical applications [5].

Human carboxylesterase 1 (hCE1) is primarily expressed in the liver and metabolizes a variety of chemically distinct substrates, from potentially harmful endogenous compounds to exogenous compounds including drugs and environmental toxins [6]. hCE1 is homologous in structure and catalytic mechanism to AChE and exhibits features that make it attractive as a potential nerve agent hydrolase [6]. Both hCE1 and AChE employ a catalytic triad consisting of serine, histidine, and glutamic acid residues. In addition, for both enzymes, OP nerve agents act as hemisubstrates – compounds that can start but not complete a reaction cycle to regenerate free active enzyme [7-9]. Indeed, hCE1 is potently inhibited by nerve agents, forming long-lived covalent phosphonyl-enzyme complexes (Figure 6.1.B) [8]. We have previously shown that hCE1 does not undergo the dead-end "aging" process from this intermediate that is observed for the

---

*Hemmert, A. C. *et al.* Nerve agent hydrolysis activity designed into a human drug metabolism enzyme. *PLoS One* **6**, e17441, doi:10.1371/journal.pone.0017441 (2011).

cholinesterases, AChE or butyrylcholinesterase (BChE) [10]. In addition, wild-type (wt) hCE1 exhibits a slow hydrolysis-based reactivation after exposure to sarin, but no reactivation is observed when hCE1 is exposed to soman or cyclosarin [8].

Thus, we sought to design an hCE1 variant capable of hydrolyzing OP nerve agents. We employed molecular modeling techniques based on crystal structures of hCE1 in trapped covalent complexes with agents to engineer amino acids capable of promoting dephosphonylation. Herein, we report the placement of a pair of residues in the hCE1 active site that significantly increases the rate of nerve agent hydrolysis relative to native hCE1, while still retaining high affinity and selective functional substrate activity towards these compounds. Thus, using only two amino acid changes, we have created one of the most efficient enzymes engineered to date with respect to converting hemisubstrates into substrates.

6.2 Methodology

hCE1 Mutagenesis, Expression, and Purification. Site-directed mutagenesis was done using PCR to introduce the desired amino acid changes into hCE1 (Genbank accession M73499). Briefly, 100 pM of both sense and anti-sense primers were mixed with 50 ng hCE1 cDNA in a pUC9 vector, 200 μM dNTPs, 1X *Pfu*-BSA buffer, and 2 u *Pfu* (New England Biolabs, Ipswich, MA). Mutations were produced through 15 cycles of PCR, where each cycle consisted of 95° C for 1 minute, 58° C for 30 seconds, and 70° C for 10 minutes. The reaction was digested for 1 hour with *Dpn1* (Fermentas, Burlington, Ontario) at 37° C and transformed into chemically competent DH5α cells (Invitrogen, Carlsbad, California). pUC9 plasmids were isolated with a GeneJET Plasmid Miniprep kit (Fermentas) and mutations were confirmed by DNA sequencing. Once the mutations were successfully incorporated, the 1.7 kB hCE1 cDNA was ligated into pCIneo for mammalian cell expression (Promega, Madison, WI). Non-secreted forms of wt and mutant hCE1 proteins were expressed in COS-7 cells (American Type Culture Collection, Rockville, Maryland) as previously described [11]. Presence of active hCE1 was determined by measuring the rate of nitrophenol production in the presence of 3 mM *o*-nitrophenol acetate

through a change in absorbance at 420 nm. Specific activities were determined by normalizing hCE1 activity relative to protein expression as visualized through western blot analysis with an anti-hCE1 antibody [11]. To facilitate measurement of additional kinetic constants, secreted forms of wt and V146H/L363E hCE1 were expressed in *Spodoptera frugiperda* Sf21 insect cells (Clontech, Palo Alto, California) using baculovirus-expression vectors and purified as previously described [24].

Examination of Kinetics of hCE1 Variants. Experiments with all racemic OP nerve agents were conducted at the United States Army Medical Research Institute of Chemical Defense (USAMRICD), Aberdeen Proving Ground, MD. Nerve agents were obtained from the U.S. Army Edgewood Chemical Biological Center (ECBC, Aberdeen Proving Ground, MD). Analysis by nuclear magnetic resonance spectroscopy showed them to be > 95% pure. Dilute nerve agent was handled according to safety guidelines in place at USAMRICD. All kinetic assays were conducted at 25° C. In separate experiments, 50 mL of whole cell COS lysates expressing an hCE1 mutant was inhibited with a ~1000-fold molar excess of racemic sarin, soman, or cyclosarin for 10 minutes. Unbound agent was removed by passing inhibited samples over a CENTRI-SEP size exclusion column (Princeton Separations, Adelphia, NJ). The column eluate was diluted 10-fold in 0.1 M potassium phosphate buffer, pH 7.4. Aliquots were tested for hCE1 carboxylesterase functional activity using 5 mM pNPB and compared to an uninhibited sample. Measurements were taken over 60 hours, and the observed rate of enzyme reactiviation (k_{obs}), or the ability of enzyme to hydrolyze pNPB, and maximal percent recovery (A_{max}) were determined by fitting the collected data to equation 1:

$$A = A_O + A_{\max}(1 - e^{-k_{obs}t}) \quad (1)$$

where A was the percent activity at time t, and A_0 was the initial activity at t=0. Reactivation experiments were conducted in triplicate and independently replicated. It should be noted that

these experiments gauge reactivation via a second substrate and do not measure OP turnover or its products directly.

Phosphonylated esterases may either spontaneously reactivate or reside in an inactive state, which is the aged state [8]. However, for wt hCE1 no aging has ever been measured. As shown in Figure 6.1.B these two pathways compete as first-order reactions. Therefore, the true rate of enzyme dephosphonylation (k_3) was determined by equation 2 [25]:

$$k_3 = (A_{\max})(k_{obs})/100 \quad (2)$$

where k_{obs} and $A_{\max}$ were defined in eq. 1. The denominator normalizes for $A_{\max}$. All data were analyzed in KaleidaGraph 4 (Synergy Software, Reading, PA). Because the reactivation step represented by k_3 was the slow step in the OP hydrolysis reaction, k_3 effectively represented the rate of OP turnover.

To determine the bimolecular rate of inhibition (k_i) with racemic cyclosarin, 100 nM of purified V146H/L363E hCE1 or wt enzyme was exposed to 100-fold excess of cyclosarin and the decrease in pNPB hydrolysis was measured over time. Data were fit to equation 3 and the k_i was calculated by dividing the rate of phosphonylation (k_2) by the cyclosarin concentration:

$$A = A_O e^{k_2 t} \quad (3)$$

where A is the CE activity at a given time t, and A_0 is the enzyme activity prior to cyclosarin inhibition.

Further inhibition kinetics were determined with stereoisomers of OP model compounds, where 100 nM of V146H/L363E hCE1 was incubated at room temperature with increasing concentrations of inhibitor. The thiomethylated OPs examined (Figure 6.1.A) have previously been shown to form identical adducts to *bona fide* nerve agents [8,13]. These compounds were generously provided by Dr. John Cashman at the Human BioMolecular Research Institute and their synthesis is described by Gilley et al. [13]. Aliquots of enzyme in the presence of sarin, soman, and cyclosarin model compounds were removed at various time points (up to 1 hour) and the

level of remaining enzyme activity was determined by comparing 4-methylumbelliferyl acetate (4-MUA) hydrolysis relative to an uninhibited sample. 4-Methylumbelliferone fluorescence emission, measured at 450 nm following excitation at 350 nm, was monitored at 37° C on a Pherastar microplate reader (BMG Labtech, Cary, NC) and the data were fit to equation 4 [26]:

$$\frac{\Delta t}{\Delta \ln v} = \frac{K_m}{k_2} \times \frac{1}{[OP](1-\alpha)} + \frac{1}{k_2} \tag{4}$$

where K_m was the nerve agent model Michaelis-Menten constant, k_2 the unimolecular phosphonylation rate constant, v the remaining percent enzyme activity, and [OP] the OP concentration. The term α was defined as $[S]/(K_m+[S])$, in which [S] was the substrate concentration and K_m was the 4-MUA Michaelis-Menten constant. For these experiments α was 0.91. All experiments were performed in triplicate and data were analyzed to determine k_i, where $k_i=k_2/K_m$ [16]. Again, all data were analyzed in KaleidaGraph 4.

Computational Modeling of hCE1 Mutants. To simulate the hCE1-OP crystal structure for computational design, a new residue type was created for S221 covalently inhibited by sarin, soman, or cyclosarin. Following removal of all water molecules, saccharides, ligands, and extra monomers from the hCE1-cyclosarin homotrimer crystal structure (RSCB PDB 3K9B [8]), protein hydrogen atoms were added into one monomer with Molprobity [27]. OP-inhibited S221 hydrogen atoms were added using Accelrys Discovery Studio Visualizer 2.5 (Accelrys, San Diego, CA). Any strain or steric interactions in the resulting structure were then removed by performing 50 steps of conjugate gradient energy minimization.

Residues with side chains pointing towards and within 5 Å of the OP, V146, or L363 residues (using 3K9B numbering, residues 89, 90, 93, 96, 97, 100, 101, 145, 220, 222, 252, 254, 255, 304, 318, 358, 359, 361, 364, 388, 425, 426, and 468,) were allowed to sample alternative conformations during the design phase of the simulation, but their identities were not modified. At positions 146 and 363, only histidine or glutamic acid was allowed, respectively. A standard backbone-dependent side chain rotamer library with expansions by one standard deviation about

$\chi 1$ and $\chi 2$ was used [28]. The crystallographic conformer at each designed position was also allowed. Prior to the modeling procedure, rotamer libraries for the various OPs were generated. The OP rotamer libraries introduced torsions in the phosphorus-alkoxy oxygen and alkoxy oxygen-carbon bonds of each OP-bound serine residue, in 5° increments. The rotamers with conformational energies lower than a specified value (usually 0) were included in the OP rotamer library. Computational design was done using the PHOENIX protein design software. The energy function used was based on the DREIDING force field [29] and included a scaled van der Waals term, hydrogen bonding and electrostatic terms, as well as terms for implicit solvation and phi-psi propensities [29]. Implicit solvation energies were evaluated using a model based on occluded volume. Amino acid phi-psi propensities were derived and applied following the method of Shortle [30]. The final model consisted of the lowest energy rotamers for the OP, V146H, L363E, and surrounding active site residues.

6.3 Results

6.3.1 Structure-Guided Dyad Design

To hydrolyze the covalent phosphonyl-enzyme OP-bound species of hCE1, we introduced a histidine (H) at position 146 and a glutamic acid (E) at position 363. These residues were chosen because they appeared ideally positioned to order a water molecule for hydrolysis. The corresponding wild-type residues V146 and L363 are positioned on either side of the active site pocket, adjacent to the OP-serine bond formed during the two-step catalytic mechanism of the enzyme (Figure 6.2). V146 is part of an extended loop between b7 and a2 that also contains the oxyanion hole formed by residues G142 and G143. L363 caps a12, at the edge of the hCE1 active site. As outlined below, the structurally similar residue glutamine (Q) was used as a control mutation for each position. hCE1 mutants were successfully expressed, purified, and shown to exhibit enzyme activity using a standard esterase substrate para-nitrophenylbutyrate (pNPB) [11], although wildtype hCE1 was ~30-fold faster toward pNPB than the V146H/L363E mutant (660 nmol/min/mg vs. 21 nmol/min/mg).

6.3.2 Enhanced Nerve Agent Hydrolysis

The V146H/L363E hCE1 mutant exhibited a significant increase in enzyme reactivation following nerve agent inhibition (Figure 6.3.A). We had shown previously that wild-type (wt) hCE1 reactivates following sarin exposure at a rate (k_3) of 2.4 ± 0.50 × 10^{-4} min^{-1}, or a half-time of reactivation ($t_{1/2}$) of 46 hrs, and that wt enzyme was not able to recover following soman or cyclosarin inhibition [8]. In contrast, however, the V146H/L363E hCE1 mutant spontaneously reactivated following sarin and soman inhibition with rates of 12 ± 2.0 × 10^{-4} min^{-1} ($t_{1/2}$ of 9.5 hrs) and 10 ± 1.0 × 10^{-4} min^{-1} ($t_{1/2}$ of 11.5 hrs), respectively. The reactivation rates of V146H/L363E hCE1 after sarin and soman inhibition were 5- and 20-fold faster, respectively, than the wt enzyme. Most strikingly, the designed V146H/L363E dyad exhibited a rate of reactivation of 67 ± 6.0 × 10^{-4} min^{-1} ($t_{1/2}$ of 1.2 hr) following cyclosarin exposure, an increase in reactivation of ~33,000-fold relative to wt enzyme (where the wt rate was estimated to be 1 x 10^{-5} min^{-1}, the lower limit of rate measurement). Recall that no reactivation was observed with wt hCE1 in the presence of this OP. Thus, efficient recovery of carboxylesterase activity following nerve agent exposure was achieved by introducing a pair of residues adjacent to the hCE1 active site.

We next tested a series of single and double mutant controls containing Q residues at positions V146 and L363 to examine the effect of each mutant on reactivation rate (k_3) with respect to sarin, cyclosarin, and soman hydrolysis (Figure 6.3.B, Table 6.1). For sarin hydrolysis, introduction of a polar side chain (H or Q) in either position alone or in combination produced similar rates of reactivation, ranging from 3.2 × 10^{-4} to 26 × 10^{-4} min^{-1}, up to a 11-fold increase relative to wt hCE1. Following soman or cyclosarin inhibition, L363E appeared critical to increasing k_3 from sub-detection limits in wt hCE1 to 4.0 × 10^{-4} and 19 × 10^{-4} min^{-1}, up 8- and 9,300-fold, respectively (Table 6.1). For soman, polar residues at position 146 (H or Q) weakly stimulated reactivation; however, when V146H was combined with L363E, reactivation increased 20-fold relative to wt hCE1 (Table 6.1).

However, the largest k_3 enhancement was for cyclosarin and the V146H/L363E mutant. This designed enzyme exhibited an increased rate of enzyme-mediated cyclosarin dephosphonylation of >33,000-fold relative to the wt enzyme (Table 6.1). A second double-mutant in which a Q at position 146 was combined with an E at position 363 was also found to exhibit robust enzyme reactivation following cyclosarin exposure, with an rate increase of >21,000-fold relative to wt hCE1 (Table 6.1). Thus, a glutamic acid at position 363 appeared to function as a catalytic base to enhance cyclosarin reactivation; a glutamine in that position failed to stimulate enzyme activity in the presence of cyclosarin. Taken together, these data demonstrate that targeted mutagenesis within hCE1 can significantly increase nerve agent processing by this serine hydrolase.

6.3.3 Structural Mechanism for Enhanced Cyclosarin Hydrolysis

The crystal structure of hCE1 in covalent complex with cyclosarin was then used to generate an energy-minimized model of V146H/L363E hCE1. In the resultant model, the side chains of L363E and V146H were found to be 3.3 Å apart. In addition, the orientation of L363E suggested that a water molecule could be positioned 3.2 Å from L363E and 3.2 Å from the phosphorous atom on the OP-enzyme adduct (Figure 6.4.A). Based on this model and the biochemical data presented above, we propose that L363E is positioned for water activation and is stabilized by H146 (Figure 6.4.B). Indeed, pH rate dependence studies showed that reactivation proceeded most effectively below pH 6.2, when V146H is likely to be positively charged (Figure 6.4.C).

6.3.4 Bimolecular Rates of Inhibition

To ensure that the enhanced rates of reactivation of V146H/L363E were not artifacts of altered substrate binding, we determined the bimolecular rates of inhibition (k_i) of the mutant protein with racemic cyclosarin or stereoisomers of OP model compounds similar in structure to cyclosarin, and compared those rates to wt enzyme. V146H/L363E hCE1 exhibited a similar k_i to wt hCE1, confirming that the ability of the mutant protein to bind cyclosarin and related OP

compounds was not disrupted (Table 6.2, Table 6.3). Indeed, as measured with stereoisomers of the model compounds, V146H/L363E exhibited an improved binding constant (K_m) compared to wt hCE1 with either stereospecific compound [8]. Therefore, the addition of an ion pair across the active site entrance was not detrimental to cyclosarin binding.

6.4 Conclusions

OP nerve agents serve as hemisubstrates of the wild-type AChE, BChE, and hCE1 enzymes because these chemicals fail to complete the typical two-step serine hydrolase mechanism. Rather, OP nerve agents produce relatively long-lived covalently phosphonylated active site serines [12,13]. Histidine (H468 in wt hCE1; see Figure 6.2.B) activates the serine (S221 in hCE1) for nucleophilic attack on the OP phosphonyl center. The resultant pentahedral intermediate is stabilized by hCE1 backbone nitrogen atoms in the oxyanion hole (G142 and G143); the fluoride leaving group dissociates upon collapse of the transition state, resulting in a covalently modified enzyme [14]. The resulting tetrahedral phosphonyl adduct adversely affects the ability of H468 to facilitate base mediated C-O bond cleavage [15], which leaves this residue in a state where it cannot activate the water molecule required for hydrolytic desphosphonylation [16]. Thus, for serine hydrolases like AChE and hCE1 that are inhibited by OP hemisubstrates, the addition of a strong oxime is often required to complete the catalytic cycle and regenerate active enzyme. Current treatments following OP exposure include administration of strongly-nucleophilic oximes to dephosphonylate AChE. However, such compounds, such as 2-pralidoxime obidoxime, do not offer broad-spectrum protection against all agents and must be administered quickly [17].

Based on previous crystal structures of hCE1 in covalent complexes with nerve agents [7,8], we hypothesized that introducing a general base catalyst would facilitate the activation of a properly located water molecule to hydrolyze an OP adduct. The V146H/L363E amino acid substitutions dramatically increased the spontaneous rates of enzyme reactivation following sarin, soman, and cyclosarin inhibition compared to native enzyme (Figure 6.3). A panel of single and

double mutant controls indicated that V146H or L363Q facilitated increased sarin hydrolysis, while L363E acted as general base catalyst for cyclosarin and soman hydrolysis. Because V146H, L363Q, or their combination, resulted in similar rates of reactivation, these mutations may enhance dephosphonylation by interacting with the sarin *O*-isopropyl group, thereby affording deprotonation of the catalytic H468 for water activation and hydrolysis. A G117H mutation in BChE resulted in increased rates of dephosphonylation through an analogous mechanism [16,18].

In contrast, the L363E mutation positions an anionic carboxylate adjacent to the phosphonyl-enzyme intermediate to activate a water molecule for nucleophilic attack. Glutamic acid and aspartic acid residues act as general acid-base catalysts in several established enzyme mechanisms, including those of lysozyme and protein tyrosine phosphatases [19,20], and apparently support the same role in hCE1. The V146H addition synergistically increased base activity of L363E (Figure 6.4.A). The pH profile for cyclosarin hydrolysis by V146H/L363E showed that the fastest rates of hCE1 reactivation was when H146 was likely to be in a charged state, below pH 6.2 and where a cationic H146 can stabilize the anionic base (Figure 6.4.C). Based on these data, we postulate that E363 Oe1 is stabilized by H146 and Oe2 on E363 is available to deprotonate a water molecule to hydrolyze the covalent cyclohexyl phosphonyl group (Figure 6.4.B). This hypothesis is supported by the observation that the V146Q/L363E variant also exhibited a significant increase in cyclosarin hydrolysis (Table 6.1).

Previous attempts to introduce general base catalysts in OP-inhibited serine hydrolases have resulted in enzymes that also showed enhanced rates of reactivations against soman, sarin, and VX, but the ability of these mutant enzymes to bind the respective OPs was greatly diminished [10,16]. For example a G117H/E197Q BChE mutant enhanced the rate of soman hydrolysis up to 2,500-fold faster than wt BChE, but the soman bimolecular rate of inhibition for this mutant was decreased by a similar magnitude [10]. In contrast to the G117H/E197Q BChE mutant, V146H/L363E hCE1 exhibits a greater second-order rate of inhibition (k_i) than the wt hCE1 enzyme (Table 6.2). Further, the K_m values for the cyclosarin model compounds reported

in Table II demonstrate that the affinity for V146H/L363E hCE1 is greater than the wt enzyme for these analogs. Taken together, these data show that V146H/L363E hCE1 retains greater than wt affinity for cyclosarin, and that the engineered hCE1 mutants enhance rates of reactivation via dephosphonylation rather than decreasing inhibition.

In conclusion, we have rationally designed a variant form of the liver detoxifying enzyme hCE1 that spontaneously dephosphonylates after inhibition by sarin, soman, and cyclosarin up to 33,000-fold faster than wt enzyme. Wild type hCE1 has a catalytic efficiency (k_{cat}/K_m) [21] towards the standard esterase substrate para-nitrophenylbutyrate (pNPB) of 1.2×10^5 $M^{-1}s^{-1}$ [22]. AChE, one of the most efficient enzymes known, exhibits a k_{cat}/K_m close to the diffusion control limit (10^8 $M^{-1}s^{-1}$) for acetylcholine hydrolysis [23]. Combining the nanomolar K_m of V146H/L363E towards the P_R cyclosarin-like compound along with rate of reactivation after inhibition by this compound, this double mutant shows a catalytic efficiency of 8.8×10^2 $M^{-1}s^{-1}$. Thus, the redesigned hCE1 compares favorably to other mammalian enzymes that have been rationally engineered to improve hemi-substrate metabolism (Table 6.4). This enzyme will likely require substantial increases in catalytic efficiency for OP compounds in order to provide *in vivo* protection, but nonetheless can serve as a lead candidate for further development of novel countermeasures for nerve agent or pesticide poisoning.

hCE1	Sarin		Soman		Cyclosarin	
	k_3 ($\times 10^{-4}$ min^{-1})	Fold Increase	k_3 ($\times 10^{-4}$ min^{-1})	Fold Increase	k_3 ($\times 10^{-4}$ min^{-1})	Fold Increase*
WT	2.4 ± 0.5	-	0.5 ± 0.3	-	< 0.01	-
V146H	20 ± 4	8	2 ± 1	4	0.032 ± 0.006	16
V146Q	3.2 ± 0.9	1	2 ± 1	4	0.34 ± 0.09	171
L363E	4.6 ± 0.5	2	4 ± 1	8	19 ± 5	9,300
L363Q	14 ± 1	6	< 0.01	-	< 0.01	-
V146Q/L363E	5.7 ± 0.7	2	2.6 ± 0.5	5	43 ± 7	21,500
V146H/L363Q	26 ± 3	11	< 0.01	-	< 0.01	-
V146Q/L363Q	12 ± 1	5	< 0.01	-	< 0.01	-
V146H/L363E	12 ± 2	5	10 ± 1	20	67 ± 6	33,355

Table 6.1 Rates of dephosphonylation for hCE1 mutants. N.R. is no reactivation with a detection limit < 1×10^{-5} min^{-1}. pH 7.4, 25°C, N=3, s.d.*k_3 ~ 1×10^{-7} min^{-1}, i.e. (1×10^{-5} min^{-1})*(1% activity)/100.

V146H/L363E	k_i (× 10^3 $M^{-1}s^{-1}$)	K_m (× 10^{-6} M)	k_3 (× 10^{-4} min^{-1})
Racemic[a]	79 ± 2	N.D.	67 ± 6
P_R[b]	5.4 ± 0.4	0.095 ± 0.006	50 ± 10
P_S[b]	0.12 ± 0.02	5.2 ± 0.7	40 ± 10
WT			
Racemic[a]	11 ± 5	N.D.	< 0.01[c]
P_R[b]	15 ± 1[c]	0.32 ± 0.01[c]	< 0.01[c]
P_S[b]	0.0053 ± 0.0007[c]	1700 ± 200[c]	< 0.01[c]

Table 6.2 Bimolecular rates of inhibition, Michaelis-Menten constants, and rates of reactivation for wild-type and V146H/L363E hCE1 against racemic cyclosarin and stereoisomers of cyclosarin model compounds. N=3, s.d., N.D. is not determined, N.R. is no reactivation, pH 7.4, 25°C [a]Racemic *bona fide* cyclosarin, [b]stereoisomers of cyclosarin model compounds, [c](8).

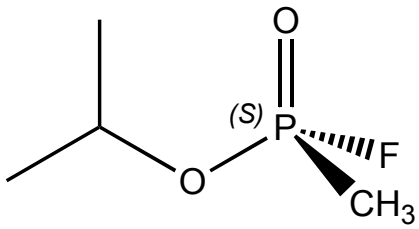
hCE1	Model Compound	k_i (× 10^3 $M^{-1}s^{-1}$)		K_m (× 10^{-6} M)	
		P_R	P_S	P_R	P_S
WT[a]	Sarin	0.77 ± 0.05	4.1 ± 0.1	1.4 ± 0.2	0.37 ± 0.02
	Soman	2.5 ± 0.5	0.0015 ± 0.0007	0.22 ± 0.02	800 ± 200
V146H/L363E	Sarin	3.40 ± 0.07	4.7 ± 0.2	3.5 ± 0.8	0.9 ± 0.1
	Soman	10.8 ± 0.4	0.67 ± 0.06	0.101 ± 0.005	8 ± 2

Table 6.3 Inhibition and Michaelis-Menten constants for wild-type and V146H/L363E hCE1 against stereoisomers of sarin and soman model compounds. (N=3, s.d.) [a](7).

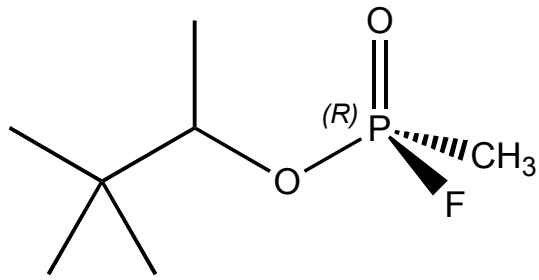
Enzyme	Engineered Activity	Mutation(s)	Catalytic Efficiency ($M^{-1}s^{-1}$)
Butyrylcholinesterase	Sarin hydrolysis[a]	G117H	0.79
Glutathione Transferase A1-1	Thioester hydrolysis[b]	A216H	2.60
4-Oxalocrotonate Tautomerase	Oxaloacetate decarboxylase[c]	P1A	114
Phospho(serine/threonine/tyrosine) Binding Protein	Phosphatase activity[c]	G120C	490
Human Carboxylesterase 1	**Cyclosarin analog hydrolysis**	**V146H/L363E**	**877**
Cyclophilin	Protease activity[c]	A91S/F104H /N106D	1675

Table 6.4 Catalytic efficiencies (k_{cat}/K_m) of engineered enzymes towards hemisubstrates. [a](16) [b](32) [c](32)

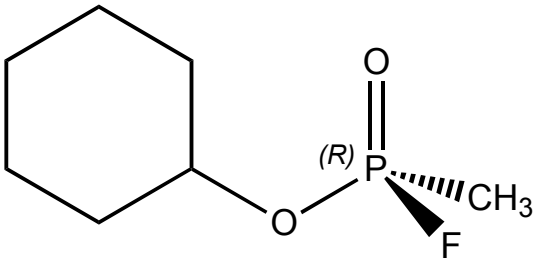
A



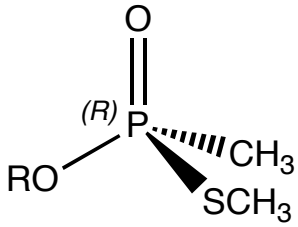
Sarin (GB)



Soman (GD)



Cyclosarin (GF)



OP Model Compound

B

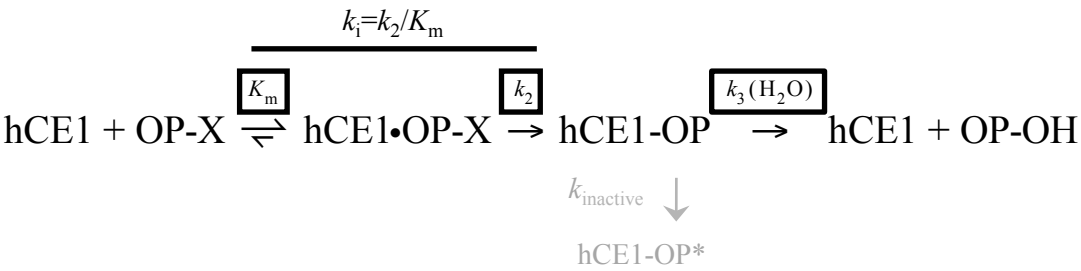


Figure 6.1 Organophosphate (OP) Inhibition of Human Carboxylesterase 1 (hCE1)

A. Three G-type OP nerve agents and OP model compound (R represents respective *O*-alkoxy groups). Wild-type hCE1 preferentially binds the stereoisomers shown [7].
B. Schematic mechanism of OP hydrolysis by hCE1. X represents the leaving group and * denotes a non-reactive state.

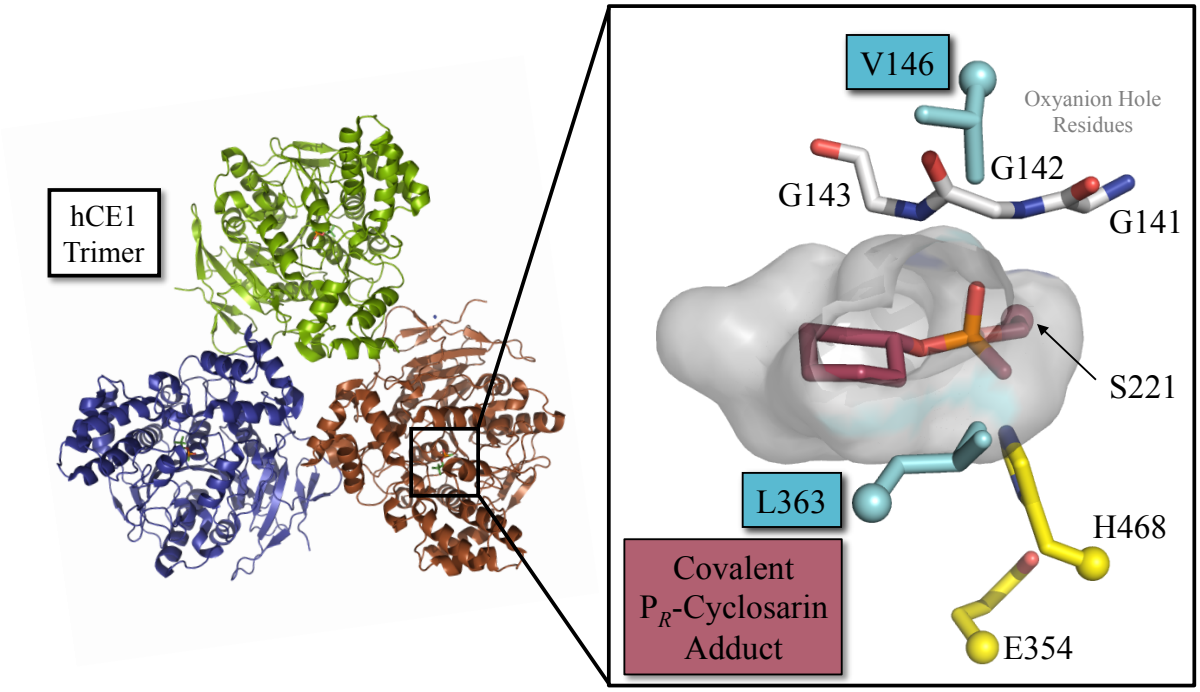


Figure 6.2 Human Carboxylesterase 1 Active Site Structure.

Active site of human carboxylesterase 1 covalently inhibited via S221 with cyclosarin (magenta) [8]. The other catalytic residues, in addition to S221, are H468 and E354 (yellow), and are surrounded by hydrophobic residues (grey surface) including V146 and L363 (light blue), as well as the oxyanion hole (white).

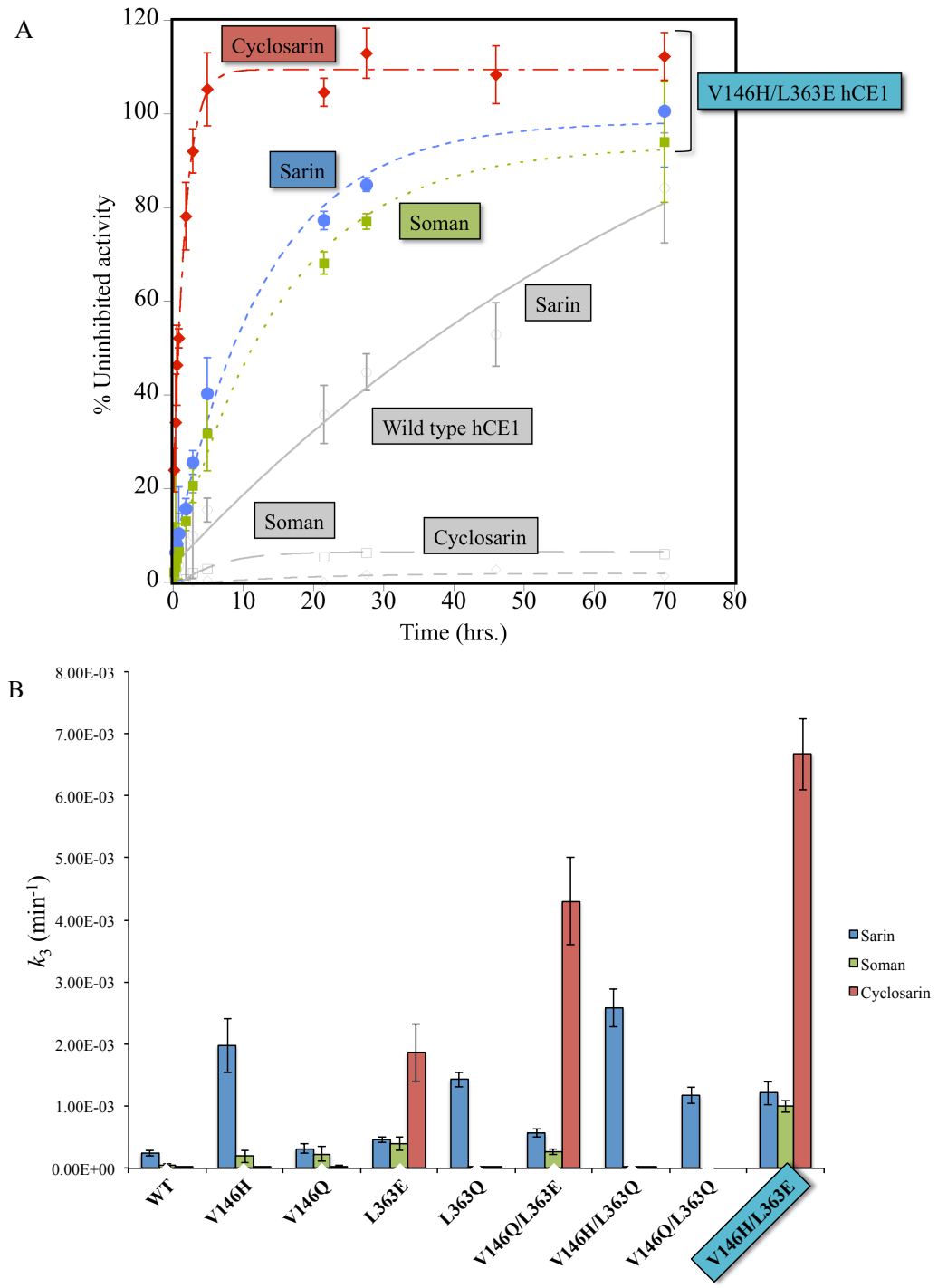


Figure 3. Reactivation of hCE1 following nerve agent exposure.

A. Spontaneous reactivation of V146H/L363E hCE1 following inhibition by racemic sarin (blue), soman (green), or cyclosarin (red). Wild type hCE1 (grey) only reactivates following sarin inhibition [7]. n=6, s.d.

B. Rates of dephosphonylation for hCE1 variants in the presence of sarin (blue), soman (green) and cyclosarin (red). n=3, s.d.

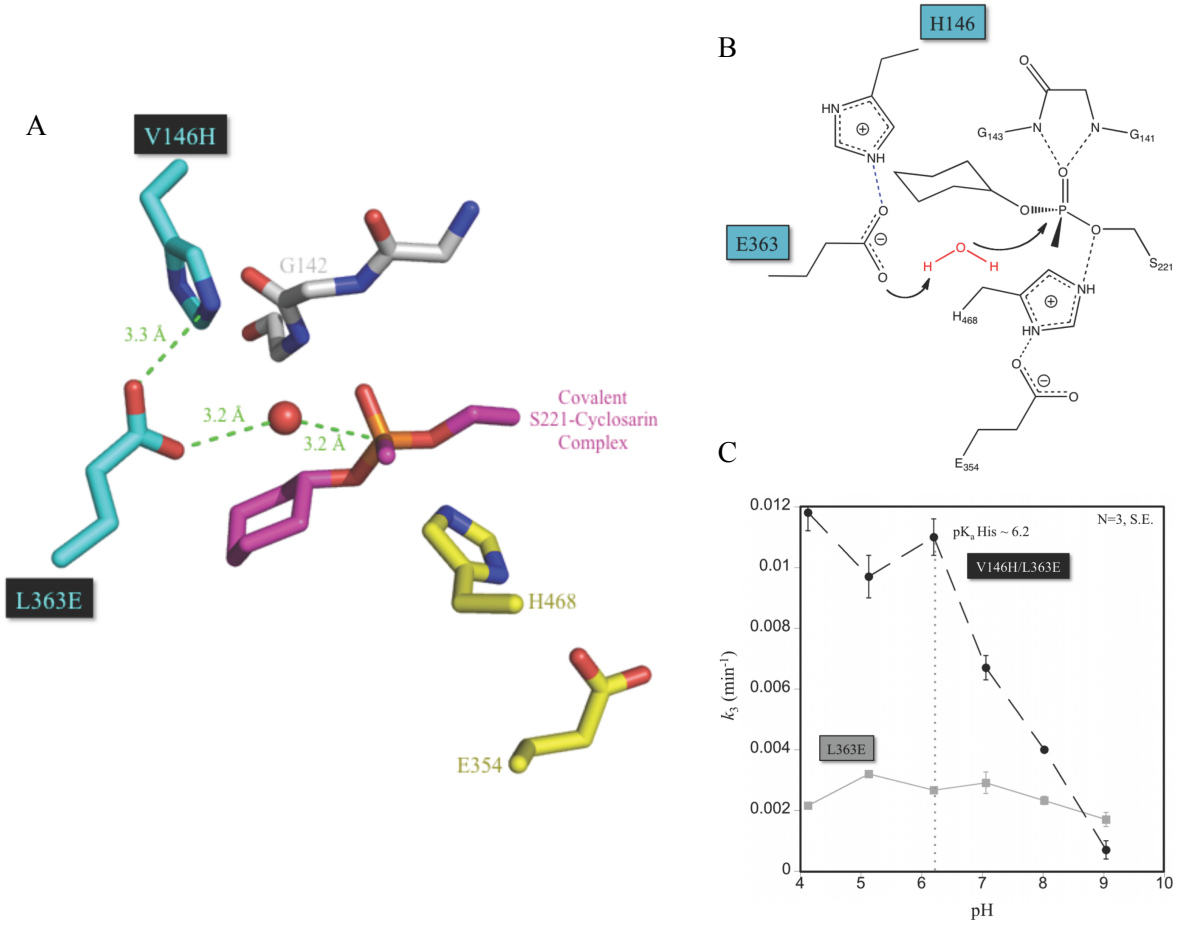


Figure 4. **Mechanism of reactivation by V146H/L363E hCE1 after cyclosarin binding**.

A. Model of V146H/L363E (cyan) hCE1 with P_R cyclosarin (magenta) including a water molecule (red) between E363 and the central phosphorus.
B. Proposed mechanism for enhanced reactivation following cyclosarin inhibition.
C. pH dependence of V146H/L363E (black) and L363E (grey) hCE1 dephosphonylation following cyclosarin inhibition.

6.5 References

1. Dixon M, Webb EC (1979) In: Boyer PD, editor. Enzymes. 3rd ed. New York: Academic Press. pp. 333.

2. Newmark J (2007) Nerve agents. Neurologist 13: 20-32.

3. Ohbu S, Yamashina A (1997) Sarin poisoning on Tokyo subway. Southern Medical Journal 90: 587-693.

4. Gunnell D, Eddleston M, Phillips MR, Konradsen F (2007) The global distribution of fatal pesticide self-poisoning: systematic review. BMC Public Health 7: 357.

5. Jett DA Finding new cures for neurological disorders: a possible fringe benefit of biodefense research? Sci Transl Med 2: 23ps12.

6. Redinbo MR, Potter PM (2005) Mammalian carboxylesterases: from drug targets to protein therapeutics. Drug Discov Today 10: 313-325.

7. Fleming CD, Edwards CC, Kirby SD, Maxwell DM, Potter PM, et al. (2007) Crystal structures of human carboxylesterase 1 in covalent complexes with the chemical warfare agents soman and tabun. Biochemistry 46: 5063-5071.

8. Hemmert AC, Otto TC, Wierdl M, Edwards CC, Fleming CD, et al. (2010) Human carboxylesterase 1 stereoselectively binds the nerve agent cyclosarin and spontaneously hydrolyzes the nerve agent sarin. Mol Pharmacol 77: 508-516.

9. Marcel V, Estrada-Mondaca S, Magne F, Stojan J, Klaebe A, et al. (2000) Exploration of the Drosophila acetylcholinesterase substrate activation site using a reversible inhibitor (Triton X-100) and mutated enzymes. J Biol Chem 275: 11603-11609.

10. Millard CB, Lockridge O, Broomfield CA (1998) Organophosphorus acid anhydride hydrolase activity in human butyrylcholinesterase: synergy results in a somanase. Biochemistry 37: 237-247.

11. Wierdl M, Tsurkan L, Hyatt JL, Edwards CC, Hatfield MJ, et al. (2008) An improved human carboxylesterase for enzyme/prodrug therapy with CPT-11. Cancer Gene Ther 15: 183-192.

12. Casida JE, Quistad GB (2005) Serine hydrolase targets of organophosphorus toxicants. Chem Biol Interact 157-158: 277-283.

13. Gilley C, MacDonald M, Nachon F, Schopfer LM, Zhang J, et al. (2009) Nerve agent analogues that produce authentic soman, sarin, tabun, and cyclohexyl methylphosphonate-modified human butyrylcholinesterase. Chem Res Toxicol 22: 1680-1688.

14. Lenz DE, Broomfield CA, Yeung D, Masson P, Maxwell DM, et al. (2008) Nerve Agent Bioscavengers: Progress in Development of a New Mode of Protection against Organophosphorus Exposure. In: Romano Jr. JA, Lukey BJ, Salem H, editors. Chemical Warfare Agents: Chemistry, Pharmacology, Toxicology, and Therapeutics. 2nd ed. Boca Raton, FL: CRC Press. pp. 175-202.

15. Qian N, Kovach IM (1993) Key active site residues in the inhibition of acetylcholinesterases by soman. FEBS Lett 336: 263-266.

16. Millard CB, Lockridge O, Broomfield CA (1995) Design and expression of organophosphorus acid anhydride hydrolase activity in human butyrylcholinesterase. Biochemistry 34: 15925-15933.

17. Kuca K, Jun D, Bajgar J (2007) Currently used cholinesterase reactivators against nerve agent intoxication: comparison of their effectivity in vitro. Drug Chem Toxicol 30: 31-40.

18. Lockridge O, Blong RM, Masson P, Froment MT, Millard CB, et al. (1997) A single amino acid substitution, Gly117His, confers phosphotriesterase (organophosphorus acid anhydride hydrolase) activity on human butyrylcholinesterase. Biochemistry 36: 786-795.

19. Vocadlo DJ, Davies GJ, Laine R, Withers SG (2001) Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. Nature 412: 835-838.

20. Denu JM, Dixon JE (1998) Protein tyrosine phosphatases: mechanisms of catalysis and regulation. Curr Opin Chem Biol 2: 633-641.

21. Wolfenden R (2006) Degrees of difficulty of water-consuming reactions in the absence of enzymes. Chem Rev 106: 3379-3396.

22. Wadkins RM, Morton CL, Weeks JK, Oliver L, Wierdl M, et al. (2001) Structural constraints affect the metabolism of 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin (CPT-11) by carboxylesterases. Mol Pharmacol 60: 355-362.

23. Elcock AH, Gabdoulline RR, Wade RC, McCammon JA (1999) Computer simulation of protein-protein association kinetics: acetylcholinesterase-fasciculin. J Mol Biol 291: 149-162.

24. Morton CL, Potter PM (2000) Comparison of Escherichia coli, Saccharomyces cerevisiae, Pichia pastoris, Spodoptera frugiperda, and COS7 cells for recombinant gene expression. Application to a rabbit liver carboxylesterase. Mol Biotechnol 16: 193-202.

25. Hovanec JW, Broomfield CA, Steinberg GM, Lanks KW, Lieske CN (1977) Spontaneous reactivation of acetylcholinesterase following organophosphate inhibition. I. An analysis of anomalous reactivation kinetics. Biochim Biophys Acta 483: 312-319.

26. Aurbek N, Thiermann H, Szinicz L, Eyer P, Worek F (2006) Analysis of inhibition, reactivation and aging kinetics of highly toxic organophosphorus compounds with human and pig acetylcholinesterase. Toxicology 224: 91-99.

27. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, et al. (2009) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66: 12-21.

28. Dunbrack RL, Jr., Cohen FE (1997) Bayesian statistical analysis of protein side-chain rotamer preferences. Protein Sci 6: 1661-1681.

29. Mayo S, Olafson B, Goddard W (1990) DREIDING: A Generic Force Field for Molecular Simulations. J Phys Chem 94: 8897-8909.

30. Shortle D (2003) Propensities, probabilities, and the Boltzmann hypothesis. Protein Sci 12: 1298-1302.