STRUCTURAL INSIGHTS INTO XENOBIOTIC AND ORGANOPHOSPHATE BINDING BY HUMAN CARBOXYLESTERASE 1 AND EFFORTS MADE TOWARDS THE CHARACTERIZATION OF THE ANDROGEN RECEPTOR MODULATOR MAGE-11

By
Christopher Daniel Fleming

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
2007

Approved by

Advisor: Professor Matthew Redinbo, Ph. D.

Reader: Professor Gary Pielak, Ph. D.

Reader: Professor Richard Wolfenden, Ph. D.

Reader: Professor Brian Kuhlman, Ph. D.

Reader: Professor John Sondek, Ph. D.

ABSTRACT

Christopher D. Fleming: Structural Insights into Xenobiotic and Organophosphate Binding by Human Carboxylesterase 1 and Efforts Made Towards the Characterization of the Androgen Receptor MAGE-11
(Under the direction of Matthew R. Redinbo)

The processing and elimination of harmful exogenous compounds is important for the successful survival of an organism in its environment. Several proteins classified as drug metabolism enzymes have evolved to provide this protection by catalyzing reactions that increase the polarity of lipophilic molecules, facilitating excretion. The drug metabolism enzyme human carboxylesterase hCE1 works to cleave ester, thioester, and amide linkages in many structurally distinct compounds. The crystal structures of hCE1 bound to tamoxifen, mevastatin, ethyl acetate, and benzil are presented here. These complexes show that hCE1 binds and metabolizes these ligands differently, highlighting its substrate promiscuity. We have additionally sought out to utilize this promiscuity in developing hCE1 as a protein based therapeutic for exposure to chemical warfare agents. Organophosphate nerve agents work by permanently inhibiting human acetylcholinesterase, an enzyme responsible for processing the neurotransmitter acetylcholine and thus terminating cholinergic nerve impulses. Current treatments for nerve agent exposure are limited, and must be administered quickly to be effective. Therefore, developing an enzyme towards the prophylactic treatment of nerve agent exposure is essential. Crystal structures of hCE1 covalently bound to the chemical weapons soman and tabun are presented here. These structures show that hCE1 is

stereoselective towards these nerve agents, and its active site architecture may provide it resistance to permanent inhibition by these compounds.

The Androgen Receptor (AR) is an intracellular transcription factor responsible for the regulation of androgen-responsive genes. AR activity is modulated by co-activators that bind to the activation function (AF-2) region of its ligand binding domain through an LxxLL motif. AR, however, exhibits a novel N/C-terminal self-association between the AF-2 and an FxxLF N-terminal motif that precludes the recruitment of these co-activators. The melanoma antigen protein MAGE-11 disrupts this interaction, facilitating AR mediated transactivation. The efforts made towards the purification and biophysical characterization of MAGE-11 are presented here. MAGE-11 was found to bind the FxxLF motif of AR an order of magnitude stronger than the AR LBD, providing an explanation for its role in AR mediated gene expression.

DEDICATION

I would like to dedicate this work to my wife, who has supported me over the nine years we have known each other, the eight years we have been together, and the three years we have been married. I couldn’t have asked for a better person to spend my life with.

ACKNOWLEDGEMENTS

To begin, I would like to thank my family for all of their love and encouragement over the years I have spent away at school. My mom, for the unwavering support and desire to let me pursue my career, always telling me that I am destined to do well at whatever I try. My dad, for instilling in me the love of science and always being a voice of reason when I step off track. I would probably be in a completely different place if I hadn't had your guidance. My sister, for always showing me that you can be both a professional and a goofball without losing anything. It is impossible for me to consider what life would have been like if I had a different family. I love you all.

Secondly, I must acknowledge members of the Redinbo laboratory both past and present for all of their help. Most importantly, the other three bionerds that I started with in the lab: Jill Orans (a.k.a. "The Sandwich"), Denise Teotico (a.k.a. "Yellow Bear"), and Scott Lujan (a.k.a. "Whole Wheat"). We have kept each other sane, and it would not have been as much fun in graduate school without good people to work with. I also have to thank Sompop Bencharit, for giving me guidance at the start of my project (and also for fixing a few teeth along the way). I also thank Mike Miley, Eric Ortlund, Laura Guogas, Kim Lane, Sarah Kennedy, Rebekkah Potts, Andy Hemmert, Joe Lomino, Yuan Cheng, Yu Xue, and Holly Cheeseman, for answering all of my inane questions over the years, I appreciate it.

None of the work I have accomplished here at Carolina could have been accomplished without the help of some top-notch collaborators: Phil Potter, Doug Cerasoli, and Elizabeth Wilson. I have appreciated the opportunity to work with you, and for trusting me to help in reaching a mutual goal.

I also must acknowledge the members of my committee for their unbiased guidance on my projects and endeavors over the past five years. Gary Pielak, Dick Wolfenden, John Sondek, and Brian Kuhlman have really shown me what it takes to be a successful researcher and academic, and I thank them for that.

Finally, I must thank my advisor Matt Redinbo, for giving me the opportunity to work on such an exciting set of projects. He has taught me how to approach and solve a scientific problem, how to be a proficient scientist, and how to utilize the skills I have learned to be successful. I wholeheartedly believe that I have worked for one of the best professors on campus, and I thank him for the opportunity to work under him (even if he thinks my dancing is ridiculous).

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2-PAM	Pralidoxime Chloride
Å	Angstrom (10^{-10} meters)
AF	Activation Function
AIB1	Amplified in Breast Cancer 1
AR	Androgen Receptor
ARE	Androgen Response Element
ACAT	Acyl-Coenzyme A Cholesterol Acyltransferase
AcChE	Human Acetylcholinesterase
BuChE	Human Butyrylcholinesterase
BSA	Bovine Serum Albumin
β-Gluc	β-Glucuonidase
°C	Degrees Centigrade
CAR	Constitutive Androstane Receptor
CE	Carboxylesterase
CEH	Cholesterol Ester Hydrolase
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CNS	Crystallography and Nuclear Magnetic Resonance Software
COV	Covalent Product
CPT-11	Camptosar® (Irinotecan)
CRP	C-Reactive Protein
CTE	C-Terminal Extension
CYP	Cytochrome P450

DBD	DNA Binding Domain
DHT	Dihydrotestosterone
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EA	Ethyl Acetate
EDTA	Ethylenediamine tetraacetic acid
ER	Estrogen Receptor
ER	Endoplasmic Reticulum
FAEE	Fatty Acyl Ethyl Ester
F_{calc}	Calculated Structure Factor
F_{obs}	Observed Structure Factor
FP	Fluorescence Polarization
GR	Glucocorticoid Receptor
GST	Gluathione-S-Transferase
hCE	Human Carboxylesterase
hiCE	Human Intestinal Carboxylesterase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HI-6	1-(2-hydroxy-iminomethylpyridinium)-1-(4-carboxyamino)-pyridinium dimethylether dichloride
HMG-CoA	3-hydroxy-3-methylglutaryl Coenzyme A
HNF	Hepatocyte Nuclear Factor
HPLC	High Pressure Liquid Chromatography

hr	hour(s)
HRE	Hormone Response Elements
IPTG	isopropyl-b-D-1-thiogalatopyranoside
kDa	Kilodalton
K_D	Dissociation Constant
K_i	Inhibition Constant
LBD	Liver Binding Domain
LXR	Liver X Receptor
MAGE	Melanoma Antigen Gene Protein
Mb	Mega bases
MBP	Maltose Binding Protein
MHC-1	Major Histocompatability Class 1
MHD	MAGE Homology Domain
min	Minute(s)
mL	Mililiter(s)
mM	milimolar (milimoles/liter)
μM	micromolar (micromoles/liter)
MR	Mineralcorticoid Receptor
*Mu*AcChE	*Mus musculus* Acetylcholinesterase
NAG	N-acetylglucosamine
N/C	N-terminus/C-terminus Interaction
NCP	Non Covalent Product
N-CoR	Nuclear Receptor Co-Repressor

nM	nanomolar (nanomoles/liter)
NMR	Nuclear Magnetic Resonance
NR	Nuclear Receptor
NTD	N-Terminal Domain
OP	Organophosphate
Ortho-7	1,7-heptylene-bis-*N*,*N*'-2-pyridiniumaldoxime dichloride
PEG	Polyethylene Glycol
PDB	Protein Data Bank
PON1	Human Serum Paraoxonase 1
PONDR	Predictor of Naturally Disordered Regions
PPAR	Peroxisome Proliferator-Activated Receptor
PR	Progesterone Receptor
PXR	Pregnane X Receptor
RAR	Retinoid Acid Receptor
rCE	Rabbit Liver Carboxylesterase
RCSB	Research Collaboratory for Structural Bioinformatics
rmsd	root mean square deviation
RXR	Retinoid X Receptor
s	second(s)
Sarin	methylethyl methylphosphonofluoridate
SEC	Size Exclusion Chromatography
SIA	Sialic Acid
SMRT	Silencing Mediator for Retinoic acid and Thyroid Hormone Receptors

Soman	*O*-pinacolyl methylphosphonofluoridate
SRC	Steroid Receptor Co-activator
Tabun	ethyl *N*,*N*-diemthylphosphoramidocyanidate
*Tc*AcChE	*Torpedo californica* Acetylcholinesterase
TEV	Tobacco Etch Virus Protease
TIF2	Transcriptional Intermediary Factor 2
TR	Thyroid Receptors
VDR	Vitamin D Receptor
VX	*O*-ethyl *S*-(2-diisopropylaminoethyl) methylphosphonothiolate
w/v	weight by volume
w/w	weight by weight

Chapter 1:

Mammalian Carboxylesterases and their role in Xenobiotic Metabolism

Introduction

The processing and elimination of potentially harmful exogenous compounds by an organism is important for successful survival in its environment. By developing systems to remove these compounds, cells are allowed to function normally in response to such environmental stresses. The enzyme class known as esterases plays an important role in this process by catalyzing reactions to increase the polarity of otherwise lipophilic molecules, facilitating excretion. Carboxylesterases (CEs), a subset of the esterase super-family, cleave ester, thioester, and amide linkages in a number of structurally-distinct substrates. Understanding the role CEs play in drug and xenobiotic metabolism is important for tuning the bioavailability of many clinical drugs. This chapter provides a background into carboxylesterases and the role of these promiscuous enzymes in xenobiotic metabolism.

The Phases of Xenobiotic Metabolism

Upon ingestion of many clinical drugs, narcotics, and potentially dangerous xenobiotics, the body responds by signaling the transcription of genes responsible for processing these compounds for elimination. The primary mechanism for detection of

exogenous compounds is through xenobiotic nuclear receptors [1-5]. These receptors, such as PXR (pregnane X receptor) and CAR (constitutively activated receptor), can detect numerous ligands with their large, flexible active site pockets. Upon drug binding, these nuclear receptors localize to response elements on the cellular DNA and up-regulate genes involved in xenobiotic elimination. The subsequent proteins produced include the major drug metabolism enzymes classified as either phase I or phase II, and phase III conjugate-specific drug transporters (Figure 1.1) [4, 6]. Phase I enzymes modify lipophilic drugs by either oxidation-reduction or hydrolytic reactions, increasing the water solubility of the small molecule. Phase I enzymes include the cytochrome P450s (CYPs), which work by catalyzing the addition of oxygen into non-activated C-H bonds, effectively adding a polar side group where none existed [7]. The CYP isoform 3A4 alone is responsible for the metabolism of greater than 50% of all prescription drugs on the market [8]. Hydrolysis reactions are primarily carried out by esterases, which catalyze transformation of an ester linkage into alcohol and carboxylic acid products. Once these functional groups are exposed, further processing can occur by way of the phase II enzymes. These enzymes are generally transferases that conjugate hydrophilic moieties such as glucuronic acid, sulfate, or acetyl groups to the xenobiotics [9]. Enzymes from both phases work in concert to increase the polarity of the compound for eventual excretion by transmembrane proteins that are specific for the conjugate molecule, termed Phase III transporters [10]. Many multidrug resistance-associated proteins are phase III transporters, such as the ABC transporter system and P-glycoprotein [11]. Whereas all three processes are important for xenobiotic clearance, this work focuses on the phase I enzyme carboxylesterase.

General Characteristics of Carboxylesterases

Carboxylesterases (CEs, EC 3.1.1.1) are proteins involved in the processing and elimination of many compounds. CEs are adept catalysts of ester-, amide-, and thioester-containing xenobiotics, including many clinical drugs and environmental toxins. CEs also catalyze the biotransformation of endogenous compounds such as cholesterol esters and fatty acyl coenzymeA molecules [12]. CEs are a member of the α/β-hydrolase-fold family of enzymes that contains cholinesterases, such as acetyl- and butyrylcholinesterase, and lipases such as cutinase and lipase B [13, 14]. While found in many different species (e.g. bacteria, plants, and higher vertebrates), the principal endogenous function of CEs is not clear. Concordantly, the naming and categorization of CEs have predominately been based on their substrate specificity, tissue localization, and molecular characteristics, which vary widely. Recently, a new method of nomenclature has been proposed by Satoh & Hosokawa, which groups CEs by their sequence identity [15]. This method groups mammalian CEs into 5 subfamilies, with the CES1 members representing the primary CE of each organism (60% identity or higher to the primary human carboxylesterase, hCE1). Enzymes are further delineated by subgroups A-H that allow additional information about organism, importance, and identity to be conveyed. This classification method should allow for more clear definitions about the function of each carboxylesterase in the family.

The expression of mammalian CEs has been confirmed in a broad range of tissues, including the heart, plasma, lung, intestines, testes, and kidney, with the highest level of esterase activity found in the microsomal and lysosomal fractions of liver [12, 15]. The

localization of carboxylesterases, however, can differ between species and isoforms of the enzymes. The human carboxylesterases are an example of this diversity. The primary human carboxylesterase, hCE1, can be found primarily in the liver, macrophages, and lung epithelia, whereas human intestinal CE (hiCE) is found in the small intestine, kidney, heart, skeletal muscle, and liver [16-18]. A third human CE, hCE3, is found primarily in the brain [19]. When expressed in the liver, carboxylesterases exhibit a C-terminal ER retention sequence (e.g. HXEL-COOH) that directs its cellular localization to the membrane KDEL receptors [12]. If this tetrapeptide sequence is proteolytically removed by the protein processing machinery, or is simply absent from the primary sequence, the resultant enzyme is excreted from the cell by the Golgi apparatus [12, 20]. This likely allows for CE activity to be found in circulating plasma of rats, mice, horses, and cats; humans however, do not exhibit such activity [21].

CEs work by catalyzing the breakdown of ester, amide, or thioester linkages using a two-step hydrolysis mechanism that results in the corresponding alcohol and carboxylic acid products (Figure 1.2). In more detail, a charge relay system is present between three catalytic residues (Glu-His-Ser) that works to increase the nucleophilicity of the γ-oxygen on the serine residue. This oxygen then attacks the carbonyl carbon of the linkage, forming a charged tetrahedral intermediate. The charge, stabilized in the “oxyanion hole” by the main-chain nitrogens of two neighboring glycine residues, reverts back and forces the release of the alcohol product. This first hydrolysis step results in a covalent, acyl-enzyme complex. Once the alcohol product diffuses out of the active site, a water molecule acts as the second nucleophile, attacking the carbonyl carbon of the acyl-enzyme complex. A similar charged

tetrahedral intermediate is formed, but results in the release of free enzyme and the carboxylic acid product. The reaction is considered identical for ester, amide, or thioester linkages, which is reflected in the products released. CEs are also capable of transesterification reactions, where the second nucleophile water is replaced with an alcohol molecule. This type of reaction can lead to side products that are more dangerous then the initial compounds, which will be explained further in context of the human carboxylesterases below.

The structural characteristics of carboxylesterases are similar across species, as designated by their conserved α/β-hydrolase fold. The ESTHER database (http://bioweb.ensam.inra.fr/ESTHER/general?what=index) compiles information about proteins exhibiting this type of fold, and shows several crystal structures of carboxylesterases that have been solved to date. The first structure of a mammalian carboxylesterase from rabbit liver (rCE) was reported in 2002 [22]. This structure revealed that rCE consisted of 3 domains: a catalytic domain, a αβ domain, and a regulatory domain (Figure 1.3). The catalytic domain consists of a central anti-parallel beta sheet surrounded by several alpha-helices, and contains the catalytic triad. The αβ domain is a small domain adjacent to both the catalytic and regulatory domains. The regulatory domain is primarily alpha helical, and contains two disordered loops that are thought to close partially over the active site entrance. The position of two residues in the catalytic triad (Glu353 and His467) appears to be affected by these loops being disordered, which might alter the enzymes catalytic function. The active site is a deep cavity at the interface of the three domains, and exhibits some structural flexibility that may allow for larger substrates to enter and undergo catalysis. The protein

also has high mannose type glycosylation sites at residues Asn79 and Asn389. All of these structural features are likely conserved across other mammalian carboxylesterases, the only other structure solved to date being hCE1, which is examined more closely below.

Human Carboxylesterases and hCE1

There are several human isoforms of carboxylesterase identified to date by primary sequence, however, only 3 are considered to be true carboxylesterases. hCE1, or human liver carboxylesterase, is the primary isoform and was the first identified [23]. hiCE, or human intestinal carboxylesterase, shares 49% sequence identity with hCE1. A related liver isoform, hCE2, shares 99% sequence identity with hiCE, and are generally considered to be interchangeable. The third carboxylesterase, hCE3 or human brain carboxylesterase, shares 77% and 49.5% sequence identity to hCE1 and hiCE, respectively. Less is known about hCE3 in terms of localization and its substrate preferences than the other major isoforms. Other human carboxylesterases include the human AcylCoA enzyme and the lesser known human AADAC, with 37% and 31% sequence identity to hCE1, respectively [15].

Human carboxylesterase 1 (a.k.a. Egasyn, CES1A1, hCE1) is a 62 kDa glycoprotein that plays an important role in the biotransformation of numerous medically relevant drugs. Generally, any compound that contains an ester linkage is a potential target of hCE1 activity, including the analgesics lidocaine and meperidine, the illicit drugs cocaine and heroin, and the angiotensin-converting enzyme inhibitors temocapril and delapril [24-27]. Most of these compounds are inactivated by hCE1 by removal of the ester moiety, effectively altering the

bioavailability of these drugs. Alternatively, introducing hCE1-cleavable ester linkages can be an effective method at creating pro-drugs with increased solubility. For example, the cancer therapeutic capecitabine was developed with an ester linkage which upon removal renders it effective against metastatic colorectal cancer [16, 28]. A second example of this is CPT-11 (Irinotecan), where hiCE removes an ester-linked dipiperidino moiety resulting in SN-38, a potent human topoisomerase inhibitor [20]. hCE1, however, does not hydrolyze the conversion of CPT-11 into SN-38. When comparing the general substrate specificity of hCE1 and hiCE, hCE1 is found to prefer smaller ester groups, with hiCE selecting for larger ester groups.

While hCE1 hydrolytic activity is generally advantageous to humans, the enzyme is also capable of transesterification reactions that can lead to deleterious products, with cocaine being a prime example. The transesterification of cocaine with ethanol can generate cocaethylene, which is more a more potent stimulator of the central nervous system, and exhibits a longer serum half life than cocaine, increasing the potential for overdose [29]. hCE1 also is responsible for the production of endogenous compounds that exhibit toxic side effects. The production of fatty acyl ethyl esters (FAEEs) from the transesterification of ethanol with fatty acyl-CoA molecules by hCE1 can result from chronic alcohol abuse and lead to necrotic liver decay. FAEEs work to disrupt cellular function by uncoupling oxidative phosphorylation in the inner membrane of mitochondria [30]. Transesterification activity is not limited to small molecules, with hCE1 additionally capable of acyl-CoA:acyl transferase activity, where fatty acyl-CoA is transesterified with cholesterol, forming cholesterol esters [31].

Additionally, hCE1 plays an important role in retention and trafficking of other proteins within the cell. One example of this is the interaction between hCE1 and the protein known as C-reactive protein (CRP) [32]. CRP is released into the serum upon tissue inflammation as part of the systemic acute phase response to injury. It is therefore a sensitive marker of heart muscle damage from myocardial infarction and other inflammatory events. CRP is retained within the ER of hepatocytes by hCE1 (and hCE2, both of which are subsequently bound to the membrane by way of the KDEL receptors) by interacting with a loop near the active site known as the "side door" [32]. The release of CRP by hCE1 could be the initial dose of CRP in the acute phase response. hCE1 (a.k.a. egasyn) has also been shown to interact with another protein, β-glucuronidase (β-gluc). β-gluc enzymatically removes conjugated glucuronic acid groups from xenobiotics that have been targeted for excretion by way of phase II and III drug metabolism proteins. This interaction between hCE1 and β-gluc involves the active site of hCE1 and a domain on β-gluc known as the serine protease inhibitor domain [33]. The interaction appears to enhance the activation of compounds that are targets of both enzymes, and may alter the half life of some drugs by "reactivating" molecules already bound for elimination.

hCE1 has also been proposed to be a target for development into a potent nerve agent hydrolase [16, 34, 35]. Nerve agents are classified as organophosphates (OPs), some of the deadliest synthetic chemicals developed by man [36]. The nerve agents sarin, soman, VX, and tabun, work by permanently inhibiting acetylcholinesterase (AcChE), the enzyme responsible for the processing and elimination of the neurotransmitter acetylcholine at

neuromuscular junctions [37]. This inhibition is potentially permanent through a process known as aging, where AcChE is covalently modified at the active site, killing activity (Figure 1.4) [37]. Currently, treatment options for nerve agent exposure are limited to a few small molecule drugs. Atropine, a competitive inhibitor of the muscarinic acetylcholine receptors, works to block parasympathetic neurons in an effort to maintain respiration [38]. Oximes, such as pradlidoxime chloride, work to reactivate inhibited AcChE before aging occurs [39]. However, both compounds must be administered quickly after exposure for proper protection. These limited treatment options have led researchers to look into using enzymes as prophylactics to clear nerve agents from the body. Extensive work has been done in the development of butyrylcholinesterase (BuChE, 55% sequence identity to AcChE), where a mutant variant (G117H) has been shown to have some catalytic activity towards nerve agents [40]. However, BuChE remains subject to aging reactions (like AcChE), and is therefore effective only as a bioscavenger of nerve agents [41]. For that reason, other related enzymes are being examined for development as well, with the newest focus on CEs. Rat serum CE has been shown to contain natural OP hydrolase activity on the order of that seen with the G117H mutant of BuChE [34, 35]. Because of hCE1's high sequence identity to rat serum CE (68%), hCE1 seems to be a logical human candidate for development as nerve agent hydrolase.

Extensive structural work has been done on hCE1 to date, with the first reported structures determined in 2003 [29]. hCE1 was found to exist as either a trimer or hexamer, with each monomer comprised of three domains; the catalytic, αβ, and regulatory domains (Figure 1.5). The catalytic domain consists of the canonical α/β serine hydrolase fold, with a

long 15 strand beta sheet surrounded by numerous helices. This domain, as its name suggests, also contains the catalytic triad of residues (Ser221, His468, and Glu354) that work in concert to hydrolyze ester, amide, and thioester bonds. The small αβ domain, typical of hydrolases, makes numerous contacts between monomers of the trimer, and the regulatory domain consists of 4 helices present in rCE, and two loops (Ω1 and Ω2) that were disordered in the rCE structure. hCE1, unlike rCE, contains only one N-linked, high-mannose type glycosylation site at residue Asn79, removal of which abrogates activity by destabilization of the protein [42]. hCE1 has two ligand binding sites, the active site and a surface ligand binding site called the Z-site. The active site is a primarily hydrophobic pocket with the catalytic triad found ~15 Å from the entrance. It consists of two major cavities, a large and structurally flexible promiscuous pocket, and a smaller, rigid pocket. These different areas appear to select for small (rigid) or large (promiscuous) moieties of ligands, as was seen in the homatropine-hCE1 co-crystal structure [29]. The Z-site, bordered by the Ω1 and Ω2 loops, is a low affinity ligand binding site that self-associates when hCE1 forms a hexamer. Ligand binding at the Z-site was found to modulate the trimer-hexamer equilibrium by shifting it towards trimer upon introduction of a ligand specific for that site [29]. Additional ligand co-crystal structures of hCE1 have been solved, including hCE1 bound to the Alzheimer's drug tacrine, and with numerous endogenous ligands such as Coenzyme A and the bile acid taurocholate [31, 43]. Taken together, these structures exemplify the substrate promiscuity of hCE1, and highlight hCE1's role in the metabolism of exogenous and endogenous compounds.

Despite the extensive information in the literature about carboxylesterases such as hCE1, additional questions remain. What more can be learned about the structural basis for hCE1's drug binding and metabolic activity? Is there a structural explanation for how and why hCE1 binds diverse ligands such as the breast cancer therapeutic tamoxifen, the prototypical cholesterol lowering drug mevastatin, the FAEE analogue ethyl acetate, and the hCE1 specific inhibitor benzil? What happens when hCE1 is exposed to organophosphate nerve agents? Are there any structural characteristics of hCE1 that might make it a potential target for development as an organophosphate hydrolase? The work outlined in the following two chapters was done in an effort to answer these questions, and to allow further understanding of this promiscuous phase I drug metabolism enzyme.

REFERENCES

1. Handschin, C. and U.A. Meyer, *Induction of drug metabolism: the role of nuclear receptors.* Pharmacol Rev, 2003. **55**(4): p. 649-73.

2. Reddy, J.K., *Peroxisome proliferators and peroxisome proliferator-activated receptor alpha: biotic and xenobiotic sensing.* Am J Pathol, 2004. **164**(6): p. 2305-21.

3. Traber, M.G., *Vitamin E, nuclear receptors and xenobiotic metabolism.* Arch Biochem Biophys, 2004. **423**(1): p. 6-11.

4. Wang, H. and E.L. LeCluyse, *Role of orphan nuclear receptors in the regulation of drug-metabolising enzymes.* Clin Pharmacokinet, 2003. **42**(15): p. 1331-57.

5. Xie, W., et al., *Orphan nuclear receptor-mediated xenobiotic regulation in drug metabolism.* Drug Discov Today, 2004. **9**(10): p. 442-9.

6. Xu, C., C.Y. Li, and A.N. Kong, *Induction of phase I, II and III drug metabolism/transport by xenobiotics.* Arch Pharm Res, 2005. **28**(3): p. 249-68.

7. Pylypenko, O. and I. Schlichting, *Structural aspects of ligand binding to and electron transfer in bacterial and fungal P450s.* Annu Rev Biochem, 2004. **73**: p. 991-1018.

8. Guengerich, F.P., *Cytochrome P-450 3A4: regulation and role in drug metabolism.* Annu Rev Pharmacol Toxicol, 1999. **39**: p. 1-17.

9. Jaeschke, H., et al., *Mechanisms of hepatotoxicity.* Toxicol Sci, 2002. **65**(2): p. 166-76.

10. Ishikawa, T., *The ATP-dependent glutathione S-conjugate export pump.* Trends Biochem Sci, 1992. **17**(11): p. 463-8.

11. Suzuki, T., K. Nishio, and S. Tanabe, *The MRP family and anticancer drug metabolism.* Curr Drug Metab, 2001. **2**(4): p. 367-77.

12. Satoh, T. and M. Hosokawa, *The mammalian carboxylesterases: from molecules to functions.* Annu Rev Pharmacol Toxicol, 1998. **38**: p. 257-88.

13. Holmquist, M., *Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms.* Curr Protein Pept Sci, 2000. **1**(2): p. 209-35.

14. Ollis, D.L., et al., *The alpha/beta hydrolase fold.* Protein Eng, 1992. **5**(3): p. 197-211.

15. Satoh, T. and M. Hosokawa, *Structure, function and regulation of carboxylesterases.* Chem Biol Interact, 2006. **162**(3): p. 195-211.

16. Redinbo, M.R. and P.M. Potter, *Mammalian carboxylesterases: from drug targets to protein therapeutics.* Drug Discov Today, 2005. **10**(5): p. 313-25.

17. Redinbo, M.R., S. Bencharit, and P.M. Potter, *Human carboxylesterase 1: from drug metabolism to drug discovery.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 620-4.

18. Schwer, H., et al., *Molecular cloning and characterization of a novel putative carboxylesterase, present in human intestine and liver.* Biochem Biophys Res Commun, 1997. **233**(1): p. 117-20.

19. Mori, M., et al., *cDNA cloning, characterization and stable expression of novel human brain carboxylesterase.* FEBS Lett, 1999. **458**(1): p. 17-22.

20. Potter, P.M., et al., *Cellular localization domains of a rabbit and a human carboxylesterase: influence on irinotecan (CPT-11) metabolism by the rabbit enzyme.* Cancer Res, 1998. **58**(16): p. 3627-32.

21. Li, B., et al., *Butyrylcholinesterase, paraoxonase, and albumin esterase, but not carboxylesterase, are present in human plasma.* Biochem Pharmacol, 2005. **70**(11): p. 1673-84.

22. Bencharit, S., et al., *Structural insights into CPT-11 activation by mammalian carboxylesterases.* Nat Struct Biol, 2002. **9**(5): p. 337-42.

23. Junge, W., et al., *Human liver carboxylesterase. Purification and molecular properties.* Arch Biochem Biophys, 1974. **165**(2): p. 749-63.

24. Alexson, S.E., et al., *Involvement of liver carboxylesterases in the in vitro metabolism of lidocaine.* Drug Metab Dispos, 2002. **30**(6): p. 643-7.

25. Kamendulis, L.M., et al., *Metabolism of cocaine and heroin is catalyzed by the same human liver carboxylesterases.* J Pharmacol Exp Ther, 1996. **279**(2): p. 713-7.

26. Takai, S., et al., *Hydrolytic profile for ester- or amide-linkage by carboxylesterases pI 5.3 and 4.5 from human liver.* Biol Pharm Bull, 1997. **20**(8): p. 869-73.

27. Zhang, J., et al., *Binding and hydrolysis of meperidine by human liver carboxylesterase hCE-1.* J Pharmacol Exp Ther, 1999. **290**(1): p. 314-8.

28. Rooseboom, M., J.N. Commandeur, and N.P. Vermeulen, *Enzyme-catalyzed activation of anticancer prodrugs.* Pharmacol Rev, 2004. **56**(1): p. 53-102.

29. Bencharit, S., et al., *Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme.* Nat Struct Biol, 2003. **10**(5): p. 349-56.

30. Lange, L.G. and B.E. Sobel, *Mitochondrial dysfunction induced by fatty acid ethyl esters, myocardial metabolites of ethanol.* J Clin Invest, 1983. **72**(2): p. 724-31.

31. Bencharit, S., et al., *Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1.* J Mol Biol, 2006. **363**(1): p. 201-14.

32. Yue, C.C., et al., *Identification of a C-reactive protein binding site in two hepatic carboxylesterases capable of retaining C-reactive protein within the endoplasmic reticulum.* J Biol Chem, 1996. **271**(36): p. 22245-50.

33. Novak, E.K., et al., *Expression of egasyn-esterase in mammalian cells. Sequestration in the endoplasmic reticulum and complexation with beta-glucuronidase.* J Biol Chem, 1991. **266**(10): p. 6377-80.

34. Maxwell, D.M. and K.M. Brecht, *Carboxylesterase: specificity and spontaneous reactivation of an endogenous scavenger for organophosphorus compounds.* J Appl Toxicol, 2001. **21 Suppl 1**: p. S103-7.

35. Maxwell, D.M., et al., *Comparison of Cholinesterases and Carboxylesterase as Bioscavengers for Organophosphorus Compounds*, in *Structure and Function of Cholinesterases and Related Proteins*, B.P. Doctor, Editor. 1998, Plenum Press: New York. p. 387-392.

36. Wiener, S.W. and R.S. Hoffman, *Nerve agents: a comprehensive review.* J Intensive Care Med, 2004. **19**(1): p. 22-37.

37. Casida, J.E. and G.B. Quistad, *Organophosphate toxicology: safety aspects of nonacetylcholinesterase secondary targets.* Chem Res Toxicol, 2004. **17**(8): p. 983-98.

38. Bajgar, J., *Organophosphates/nerve agent poisoning: mechanism of action, diagnosis, prophylaxis, and treatment.* Adv Clin Chem, 2004. **38**: p. 151-216.

39. Eyer, P., *The role of oximes in the management of organophosphorus pesticide poisoning.* Toxicol Rev, 2003. **22**(3): p. 165-90.

40. Millard, C.B., O. Lockridge, and C.A. Broomfield, *Design and expression of organophosphorus acid anhydride hydrolase activity in human butyrylcholinesterase.* Biochemistry, 1995. **34**(49): p. 15925-33.

41. Cerasoli, D.M., et al., *In vitro and in vivo characterization of recombinant human butyrylcholinesterase (Protexia) as a potential nerve agent bioscavenger.* Chem Biol Interact, 2005. **157-158**: p. 363-5.

42. Kroetz, D.L., O.W. McBride, and F.J. Gonzalez, *Glycosylation-dependent activity of baculovirus-expressed human liver carboxylesterases: cDNA cloning and characterization of two highly similar enzyme forms.* Biochemistry, 1993. **32**(43): p. 11606-17.

43. Bencharit, S., et al., *Crystal structure of human carboxylesterase 1 complexed with the Alzheimer's drug tacrine: from binding promiscuity to selective inhibition.* Chem Biol, 2003. **10**(4): p. 341-9.

Figure Legends

Figure 1.1: Overview of the three phases of drug metabolism. Endo- and xenobiotics are targeted by phase I drug metabolism enzymes that introduce functional groups for phase II transferases to attach large, polar molecules to. These phases work together to increase the polarity of the metabolite for eventual excretion by phase III transporters. Figure courtesy of Dr. Michael Miley.

Figure 1.2: Standard two-step hydrolysis mechanism of carboxylesterases.

Figure 1.3: Structure of rabbit liver carboxylesterase, rCE (RCSB PDB Accession Code 1K4Y). The catalytic domain is shown in blue, the regulatory domain in red, and the αβ domain in green. Glycosylation sites at residues Asn79 and Asn389 are shown in yellow ball-and-stick. The catalytic triad of residues (Glu353, His467, and Ser221) are also shown in yellow.

Figure 1.4: Mechanism of aging of cholinesterases by nerve agents. Sarin is shown as the representative nerve agent.

Figure 1.5: Trimeric structure of human carboxylesterase 1. Two of the three monomers are shown in blue and green, with the third monomer split into the three domains, catalytic (red), regulatory (purple), and αβ (pink). Glycosylations at residue Asn79 of each monomer are shown as yellow sticks, as are the six sulfate molecules that help to stabilize the trimer interface.

Figure 1.1

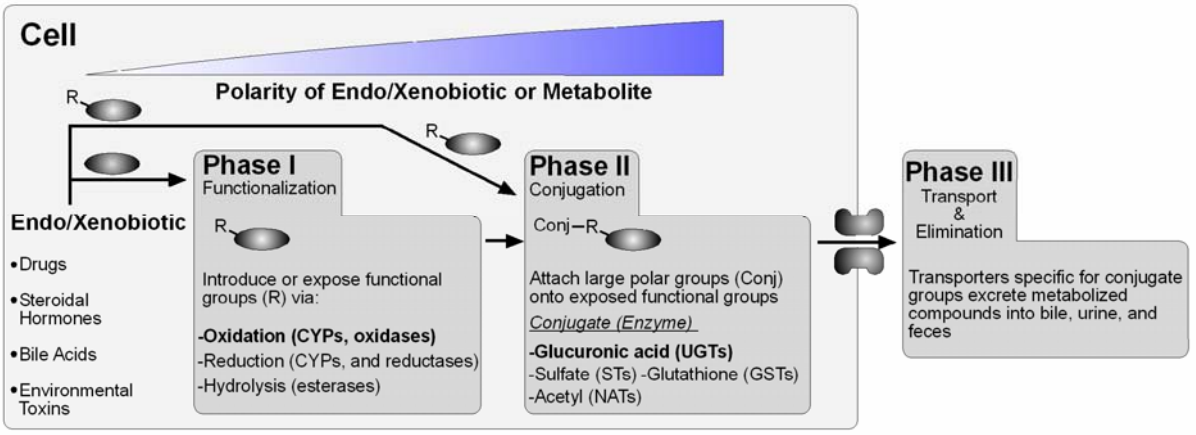

Cell
Polarity of Endo/Xenobiotic or Metabolite
R
R
Endo/Xenobiotic
•Drugs
•Steroidal Hormones
•Bile Acids
•Environmental Toxins
Phase I
Functionalization
R
Introduce or expose functional groups (R) via:
-Oxidation (CYPs, oxidases)
-Reduction (CYPs, and reductases)
-Hydrolysis (esterases)
Phase II
Conjugation
Conj–R
Attach large polar groups (Conj) onto exposed functional groups
Conjugate (Enzyme)
-Glucuronic acid (UGTs)
-Sulfate (STs) -Glutathione (GSTs)
-Acetyl (NATs)
Phase III
Transport & Elimination
Transporters specific for conjugate groups excrete metabolized compounds into bile, urine, and feces

Figure 1.2

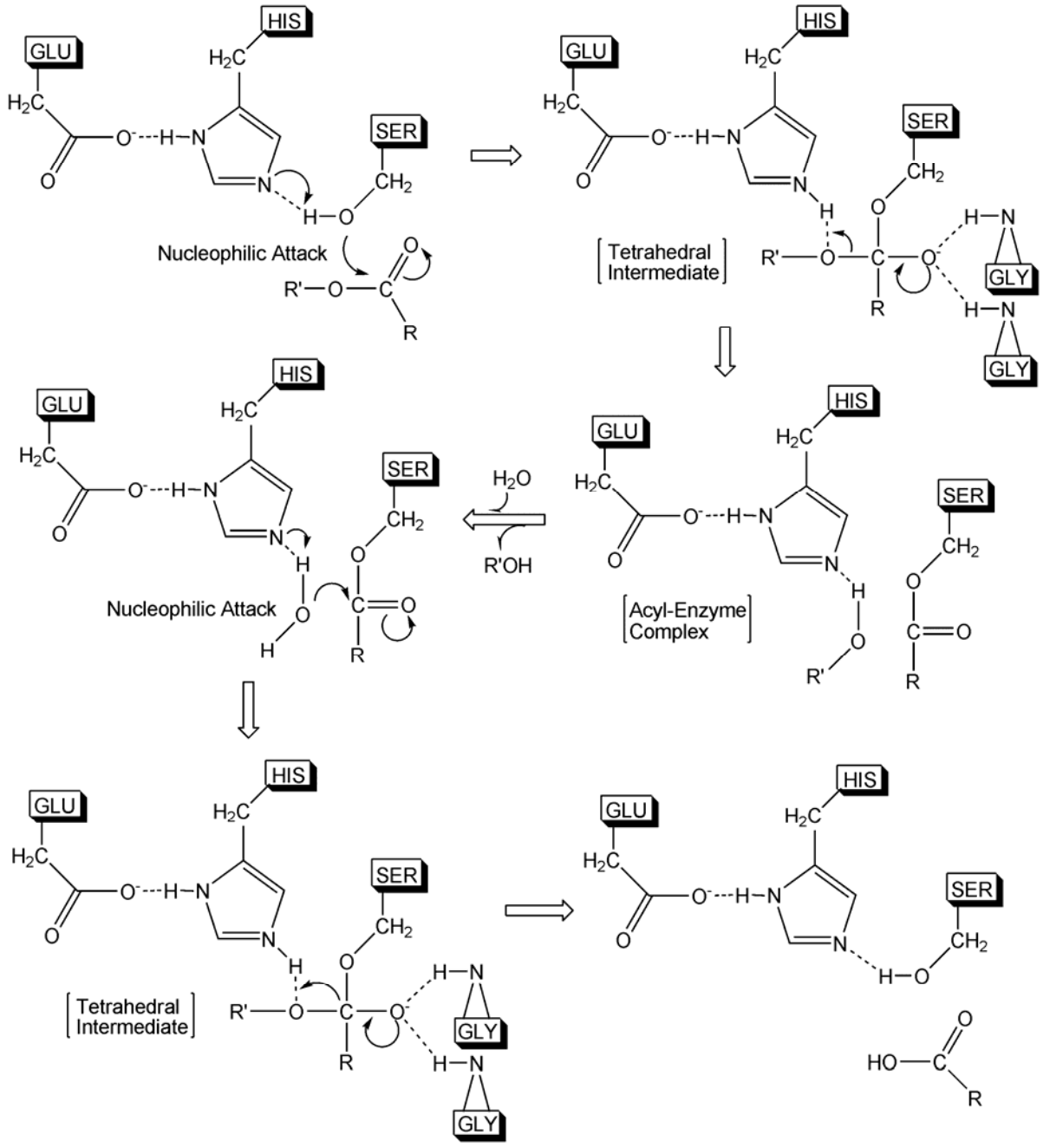

GLU
HIS
SER
Nucleophilic Attack
Tetrahedral Intermediate
GLY
GLY
Nucleophilic Attack
H2O
R'OH
Acyl-Enzyme Complex
Tetrahedral Intermediate
GLY
GLY

Figure 1.3

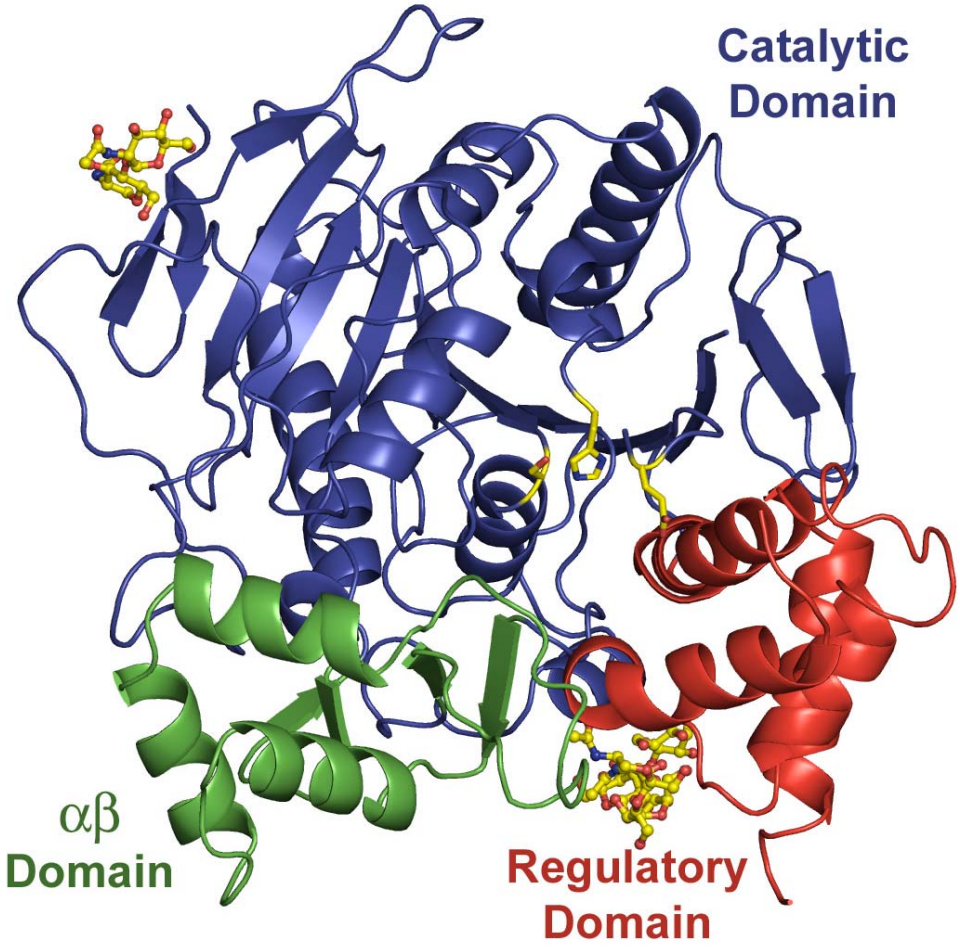


Figure 1.4

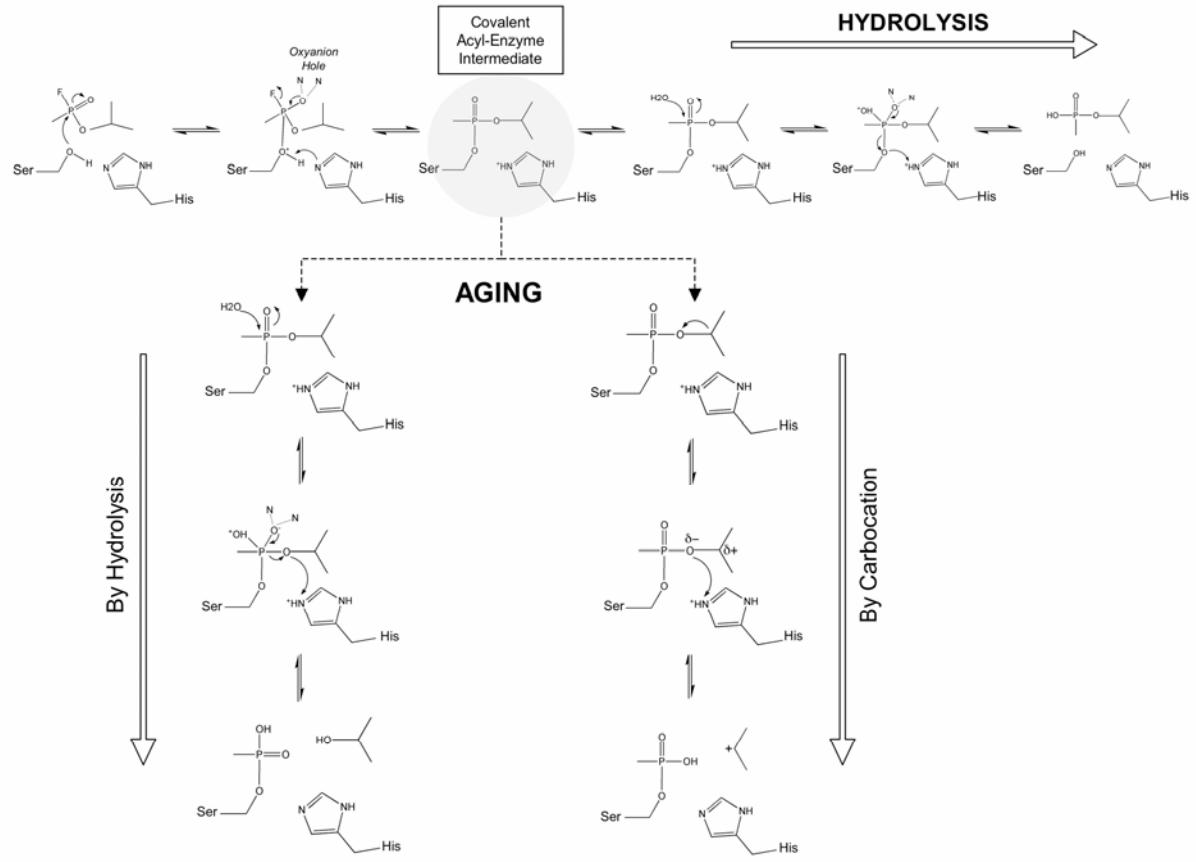
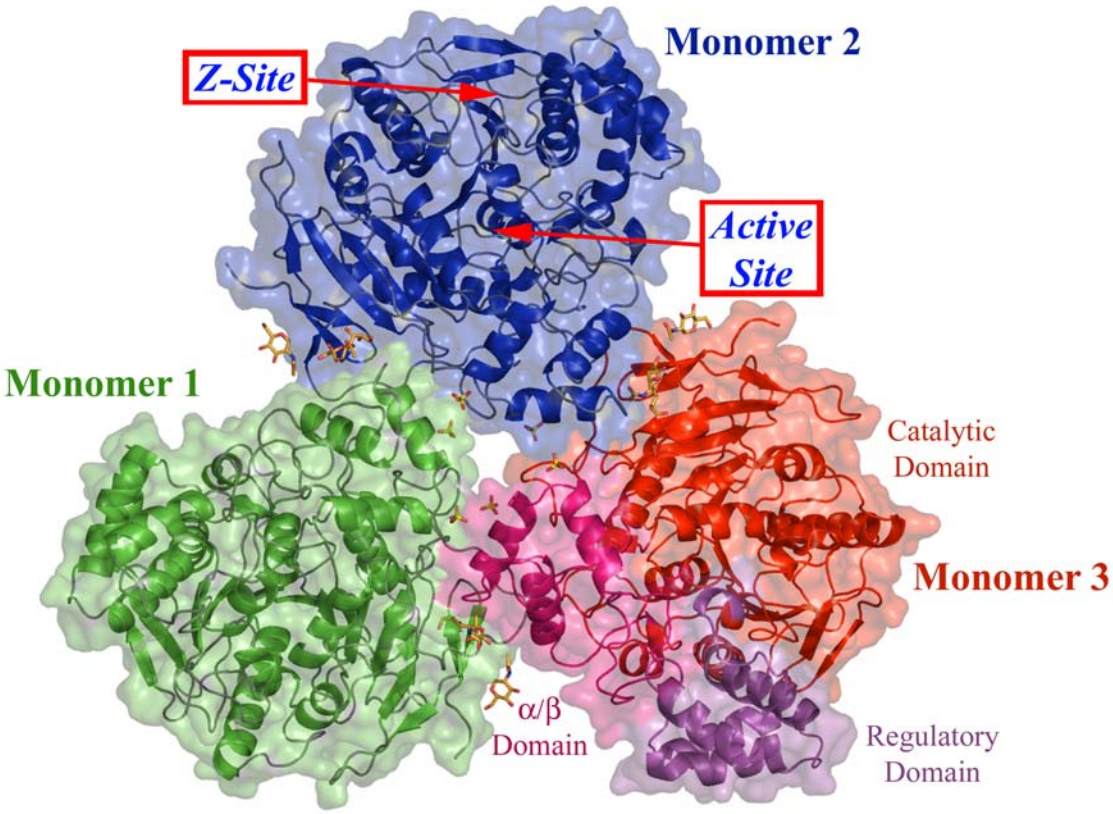

Covalent
Acyl-Enzyme
Intermediate
HYDROLYSIS
Oxyanion
Hole
AGING
By Hydrolysis
By Carbocation
Ser
His

Figure 1.5


Monomer 2
Z-Site
Active Site
Monomer 1
Catalytic Domain
Monomer 3
α/β Domain
Regulatory Domain

Chapter 2:

Structural Insights into Drug Processing by Human Carboxylesterase 1: Tamoxifen, Mevastatin, and Inhibition by Benzil

Christopher D. Fleming[1,2], Sompop Bencharit[1,3], Carol C. Edwards[5], Janice L. Hyatt[5], Lyudmila Tsurkan[5], Feng Bai[6], Charles Fraga[6], Christopher L. Morton[5], Escher L. Howard-Williams[1], Philip M. Potter[5], and Matthew R. Redinbo[1,2,4,*]

[1]Department of Chemistry,[2]Department of Biochemistry and Biophysics, [3]School of Dentistry, and [4]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [5]Department of Molecular Pharmacology and [6]Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN 38105, USA.

Published in *Journal of Molecular Biology*. 352, 165-177 (2005)

ABSTRACT

Human carboxylesterase 1 (hCE1) exhibits broad substrate specificity and is involved in xenobiotic processing and endobiotic metabolism. We present and analyze crystal structures of hCE1 in complexes with the cholesterol-lowering drug mevastatin, the breast cancer drug tamoxifen, the fatty acyl ethyl ester (FAEE) analogue ethyl acetate, and the novel hCE1 inhibitor benzil. We find that mevastatin does not appear to be a substrate for hCE1, and instead acts as a partially non-competitive inhibitor of the enzyme. Similarly, we show that tamoxifen is a low micromolar, partially non-competitive inhibitor of hCE1. Further, we describe the structural basis for the inhibition of hCE1 by the nanomolar-affinity dione benzil, which acts by forming both covalent and non-covalent complexes with the enzyme. Our results provide detailed insights into the catalytic and non-catalytic processing of small molecules by hCE1, and suggest that the efficacy of clinical drugs may be modulated by targeted hCE1 inhibitors.

INTRODUCTION

Human carboxylesterase 1 (hCE1) is a broad-spectrum serine hydrolase (E.C. 3.1.1.1) that cleaves small ester, thioester, and amide linkages in a variety of structurally-distinct compounds [1]. The enzyme is present in numerous tissues, including heart, intestines, kidney, lung, testes, monocytes and macrophages, but is most abundantly expressed in the liver, where it plays a major role in the detoxification of potentially harmful xenobiotics and endobiotics [2, 3]. hCE1 is known to metabolize a variety of drugs including cocaine, meperidine, and lidocaine [4-6], and to activate a variety of prodrugs, including heroin, the chemotherapeutic capecitabine, and the angiotensin-converting enzyme inhibitors temocapril, imidipril, and delapril [6-8]. The enzyme utilizes a two-step serine hydrolase mechanism involving the formation of a covalent acyl-intermediate on the active site serine, which is subsequently removed by hydrolysis. In addition, the enzyme is able to perform transesterification reactions to generate secondary metabolites of abundant endogenous and xenobiotic compounds [9, 10].

Three crystal structures of hCE1 and one structure of a rabbit liver CE (rCE) have been reported to date [11-13]. hCE1 exists in a trimer-hexamer equilibrium [12], and the overall structure of the hCE1 trimer is shown in Figure 2.1A. Each monomer is comprised of a catalytic domain, an α/β domain typical of hydrolases, and a regulatory domain that contains the low affinity surface ligand binding Z-site [11, 12]. The regulatory domain was identified in the first crystal structure of a mammalian CE (that of rCE, which shares 81% sequence identity and 0.68 Å root-mean-square deviation with hCE1 [12]), and was so-

termed because it was proposed to regulate access to the catalytic gorge [13]. The enzyme's active site containing the catalytic triad Ser221-His468-Glu354 is a 10-15 Å deep hydrophobic pocket at the interface of the three domains. A secondary pore into the active site of mammalian CEs, termed the "side door", has also been proposed based on the observed binding of a product of the activation of the anticancer drug CPT-11 to rCE [13]. Each hCE1 monomer also contains two disulfide linkages and one high mannose, N-linked glycosylation site on Asn79, where only the first N-acetylglucosamine (NAG) and terminal sialic acid (SIA) could be properly placed in the electron density maps for all complexes. A surface ligand binding site (the above-mentioned "Z-site") was found to control the trimer-hexamer equilibrium of hCE1; this site has been observed either to contain bound ligand, or to be involved in the formation of the hCE1 hexamer [12].

The statins, currently some of the most widely prescribed drugs in the U.S., lower serum cholesterol levels by inhibiting HMG-CoA (3-hydroxy-3-methylglutaryl-coenzyme A) reductase, which catalyzes the rate-limiting step in cholesterol biosynthesis [14]. Mevastatin (Compactin), a class 1 statin similar to simvastatin (Zocor®, Merck) and lovastatin (Mevacor®, Merck), is believed to be activated by hCE1 in the liver by conversion of the lactone form of the drug to the carboxylate form [5, 15]. Newer class 2 statins, such as rosuvastatin (Crestor®, AstraZeneca) and atorvastatin (Lipitor®, Pfizer), already exist in the carboxylate form, which removes the requirement for activation [16]. Although mevastatin was never used clinically, it provided the basis for the development of this family of drugs [14], and provides an effective model to examine the potential activation of clinical drugs by hCE1.

Tamoxifen, a nonsteroidal estrogen receptor antagonist used to treat breast cancer, limits the estrogen-dependent proliferation of cancerous cells [17]. It is the most widely used cancer drug in the world [18], and the ability of this compound to pass successfully through the liver is central to its efficacy [19]. hCE1 was not thought to be involved in tamoxifen's trafficking through the body until it was found that only four proteins in rat liver bound a labeled form of tamoxifen, and one of these proteins was the rat homologue of hCE1 [20]. Thus, it is possible that the binding of tamoxifen to hCE1 may limit the bioavailability of the drug, or may lead to drug-drug interactions and/or side effects associated with tamoxifen use, including impact on cholesterol levels [21].

In addition to its role in xenobiotic metabolism, hCE1 has been shown to catalyze reactions involved in cholesterol homeostasis and fatty acid metabolism. For example, hCE1 has been reported to possess acyl-coenzyme A: cholesterol acyltransferase (ACAT) activity, which generates cholesterol esters from fatty-acyl coenzyme A and free cholesterol [22]. hCE1 has also been found to catalyze the reverse reaction under certain conditions and in that way to act as a cholesterol ester hydrolase (CEH) [23]. In addition, the enzyme can produce fatty acyl ethyl esters (FAEEs) via the transesterification of short- and long-chain fatty acids with ethanol [24]. FAEEs are toxic byproducts associated with long-term alcohol abuse, and these compounds have been implicated in the necrotic decay of the liver and other tissues attendant with such abuse [25]. The FAEEs generated by the FAEE synthases disrupt cellular function by uncoupling oxidative phosphorylation in the inner membrane of mitochondria [26, 27]. There are four enzymes known to act as FAEE synthases, with hCE1 labeled as FAEE-synthase IV [25, 28].

The novel and potent hCE1-selective inhibitor benzil was identified during screens to discover compounds that would inhibit CE activity to aid in anticancer drug efficacy. The diphenyl-diketone compound benzil was found to inhibit hCE1-mediated hydrolysis of o-nitrophenyl acetate with a K_i value of 45 nM [29]. It was not clear, however, whether this compound functioned by binding non-covalently or covalently to the active site of hCE1.

We sought to examine at the structural level how hCE1 processes a variety of therapeutic and non-therapeutic compounds. To that end, we determined crystal structures of hCE1 in complexes with mevastatin, tamoxifen, ethyl acetate (a fatty acyl ethyl ester analogue), and benzil, and we examined the inhibitory action of mevastatin and tamoxifen on the enzyme. Our results suggest that certain classes of compounds are capable of inhibiting this drug metabolism enzyme, and that selective inhibitors of hCE1 may impact the bioavailability of clinical drugs.

RESULTS

Structural Features of hCE1

The hCE1-ligand complex crystal structures reported here were refined with the Crystallography and NMR System (CNS) [30] utilizing torsion angle dynamics throughout all steps, and with simulated annealing and non-crystallographic symmetry restraints in the early steps of refinement (Table 2.1). Despite relatively high R_{sym} values for the diffraction data employed, likely due to the low overall diffraction intensity, the maps produced after refinement were of good quality and allowed the placement of both protein and non-protein atoms. The asymmetric unit of the mevastatin, ethyl acetate, and tamoxifen complexes contained one hCE1 trimer, while the asymmetric unit of the benzil complex contained four hCE1 trimers.

Because a detailed examination of the trimer interfaces in CEs had not been reported previously, an analysis of the role of carbohydrate groups in the formation of the hCE1 trimer is presented here. A sialic acid from the glycosylation site on one monomer stacks adjacent to α7 and Glu183 in the adjacent monomer within the trimer (Figure 2.1A). In some cases, the sialic acid is within 3 Å of the main chain nitrogen at Thr279, which appears to stabilize the N-terminal, positively charged dipole of helix α7. Two charge clamps across the trimer interface, Arg186 and Glu183 of one monomer to Glu72 and Lys78, respectively, of the next monomer. Six bound sulfate molecules near His284 are also observed. While the monomers within the hCE1 trimer are related by non-crystallographic symmetry, the monomers in the structure of the related rabbit enzyme rCE (RCSB accession code 1KY4;[13]) form a trimer

related by crystallographic symmetry. In this case, although rCE also contains a similar glycosylation site at Asn79, a carbohydrate group does not appear to be involved in the rCE trimer interface. This N-linked carbohydrate is important for enzyme function, however, as the mutation of Asn79 to alanine in rCE reduces esterase activity by ~50% using o-nitrophenol acetate as a substrate (Espinosa and Potter, unpublished results). An alignment of several mammalian CEs (from mouse, rat, rabbit, human, guinea pig, pig, and monkey) revealed that residues involved in the trimer interface observed in hCE1 and rCE are identical or conserved across species [31]. In addition, both rat and porcine CE isoforms related in sequence to hCE1 have been shown biochemically to form trimers [32-34]. Thus, the known mammalian homologues of hCE1 appear to be trimers.

Drug Processing: Mevastatin

Since previous reports in the literature implicated hCE1 in the activation of mevastatin [5, 14, 35], we sought to determine the structural basis for this hydrolysis event. hCE1 is potentially capable of metabolizing mevastatin via two pathways (Figure 2.2A). α–cleavage opens the lactone ring, yielding the active, carboxylate form of the drug. β–cleavage, the alternative pathway, involves the removal of an isopentoic acid (2-methylbutyric acid) moiety from the compound, reducing the inhibitory effect of the drug on HMG-CoA reductase by 200-fold [36-38]. We crystallized the enzyme in the presence of 100-fold molar excess mevastatin, and determined and refined the structure to 3.0 Å resolution. While the α-cleavage product was expected, the initial 3.0 σ positive difference electron density in the active sites was too small to accommodate the α-cleavage product. However, simulated annealing omit electron contoured at 3.0 σ appeared appropriate for an

isopentoic acid, the small product of mevastatin β-cleavage (Figure 2.2A). Isopentoic acid refined well in this position, produced no negative electron density, and forms a hydrogen bond between one of its carboxylate oxygens and a water molecule bound within the enzyme's oxyanion hole.

We also observed 3.0 σ positive difference density within each monomer's Z-site, the non-specific, low affinity cleft that can be occupied by a wide variety of structurally-distinct ligands [11, 12]. The Z-site is formed by two loops Ω1 and Ω2, which interdigitate during hexamer formation, providing primary stacking interactions between the two trimers [12]. In one Z-site within the trimer, the difference density was appropriate for an isopentoic acid stabilized via an electrostatic contact between its carboxylate group and the ζ-nitrogen of Lys414. In the remaining two monomers, the electron density was interpreted to be the other product of mevastatin β-cleavage (the larger, decalin-lactone alcohol; Figure 2.2C); this compound refined well in this position and produced no negative electron density. It docks with its lactone ring facing down in Van der Waals contact with Leu368 and Trp357, and forms a hydrogen bond with Lys414. The α-cleavage product of mevastatin, in contrast, did not satisfy this electron density at the Z-site.

Because the crystallographic data indicated the presence of β-cleavage products bound to hCE1, we sought to validate our results biochemically utilizing an HPLC-based *in vitro* assay. After a 25 day incubation of purified hCE1 with 10-fold molar excess mevastatin at 37°C in 50 mM HEPES pH7.4 (the conditions used for crystallization of this complex), we were unable to detect enzyme-mediated production of either the α-cleavage or

β-cleavage products (data not shown). Since these results were contrary to what we observed structurally, we assessed whether mevastatin could act as an inhibitor towards hCE1. These experiments confirmed that the drug was a weak, partially non-competitive inhibitor of o-nitrophenyl acetate hydrolysis by hCE1 with a K_i value of 20.8 ± 8.8 μM. Previous studies have shown that similar statins (lovastatin, simvastatin) are able to inhibit human butyrylcholinesterase, a structural homologue of hCE1, with K_i values on the same order as that observed for hCE1 inhibition by mevastatin (12 and 4.5 μM for lovastatin and simvastatin, respectively) [39]. Thus, mevastatin appears *in vitro* to act as a weak inhibitior of hCE1.

Substrate Conjugation: FAEEs

Fatty acyl ethyl esters (FAEEs), toxic byproducts of alcohol abuse, are generated by hCE1 via the transesterification of fatty acids with ethanol [24]. To elucidate the structural basis of hCE1's action as an FAEE synthase, we attempted to crystallize the enzyme in the presence ethanol and one of a variety of poorly soluble fatty acids and related compounds. Crystals were only obtained in the presence of ethanol and Acetyl-CoA, the most soluble of the compounds tested. We found that hCE1 had transesterified Acetyl-CoA with ethanol to form ethyl acetate (EA), a small FAEE mimic (Figure 2.3A). Strong electron density for EA was observed in 3.0 Å simulated annealing omit maps at both the active site and Z-site of the enzyme (Figure 2.3B). The EA bound at the active site forms Van der Waals contacts with only two amino acids within the pocket, Ser221 and Leu304, and thus does not appear to bind with a high level of specificity (Figure 2.3C). Non-specific binding is also observed at the Z-site, where EA forms only a hydrogen bond with Lys414 (Figure 2.3D). These data

indicate that hCE1 can accommodate Acetyl-CoA within its active site and utilize ethanol to transesterify the acetyl product of Acetyl-CoA cleavage to generate EA.

Drug Binding: Tamoxifen

The rat homologue of hCE1 (ES-10) was found to be one of only four liver proteins that bound to tamoxifen with high affinity [20]. We sought to elucidate the structural basis of hCE1's ability to interact with this widely-used anticancer drug, and to determine if tamoxifen inhibited the enzyme. Tamoxifen was observed bound at both the active site and Z-site in our 3.2 Å structure of the drug-hCE1 complex (Figure 2.4A). Tamoxifen is well ordered within the catalytic pocket, making hydrophobic contacts with eight active site residues that line the pocket (Figure 2.4B). Indeed, tamoxifen's triphenyl structure fits remarkable well in the binding cavity, with two of the rings filling pockets adjacent to Phe101 and Leu388, and the drug's tail protruding toward the entrance to the cavity. Leu304, which sits near the top of the pocket, undergoes a rotamer shift relative to other hCE1 structures to provide room for drug binding. Helices α1 and α10' clamp over the top of the active site, closing down onto the tamoxifen compound. It has been proposed that α10' has the ability to shift in position to allow bulky molecules to enter the active site (Figure 2.4C) [12]. Once the position of α10' becomes relatively stable, it serves as the "floor" of the Z-site. Tamoxifen binds with less specificity at the Z-site, making only four Van der Waals contacts with amino acid side chains and exhibits higher thermal displacement parameters in this position (69-79 $Å^2$) relative to that observed at the active site (30-43 $Å^2$).

We found that tamoxifen was a micromolar-affinity inhibitor of both hCE1 and rCE, with K_i values of 15.2 ± 2.8 and 23.3 ± 14.3 μM, respectively (Table 2.2). The K_i values agree with previous work showing tamoxifen as a CE inhibitor in rat [20]. We also found that both hCE1 and rCE are inhibited by tamoxifen in a partially noncompetitive fashion. In the case of hCE1, these results indicate that the Z-site may act as an allosteric site. The Z-site is located directly "above" the active site, separated from it by α10' (Figure 2.4C), and may control ligand access to the active site [12]. The analogous loops are disordered in the rCE structure, so it is not known whether this enzyme contains a surface Z-site. A surface binding site proposed to act in a similar manner, as an allosteric or substrate recognition site, was recently observed in the structure of the human drug metabolism enzyme cytochrome P450-3A4 [40].

Inhibition of hCE1: Benzil

Finally, we determined the crystal structure of hCE1 in complex with benzil, a potent CE inhibitor with a K_i value of 45 nM for hCE1 [29]. The structure contained four trimers in the P1 asymmetric unit, which will be referred to as trimers A-D with each monomer denoted 1-3 (*e.g.,* monomer A3 or D1). Positive (2.0-3.0 σ) difference density in the active site in all twelve monomers was not satisfied by placement of intact benzil. After further examination and refinement, it appeared that hCE1 had cleaved benzil to generate a covalent product and either benzaldehyde and/or benzoic acid. Based on our structural results, we propose that the following elements may be involved in benzil inhibition of hCE1 (Figure 2.5A). First, a cycling reaction may occur in which attack of the catalytic Ser221 residue on one of the benzil's carbonyl groups to form the covalent intermediate reverses to generate benzil and

free enzyme (Figure 2.5A). Such a repeating cycle would occupy the active site and efficiently inhibit the enzyme. This is consistent with the previously published kinetic analyses [29].

Second, benzil could be hydrolyzed to the two benzyl ring products we have named NCP1 (benzaldehyde, or non-covalent product 1) and NCP2 (benzoic acid). Such a cleavage reaction would require a retro-aldol condensation followed by a hydrolysis event (Figure 2.5A). For benzaldehyde formation, a transfer of electrons from the first hydroxyl group to the second carbonyl carbon would occur, releasing NCP1. This movement of electrons would also result in the covalent acyl product found in 3 of the active sites (COV). A hydrolysis event, analogous to the second step of the standard two-step serine hydrolase mechanism, would then be required to remove the benzoyl ester intermediate and release the second benzyl ring product benzoic acid (NCP2), as well as the free enzyme (Figure 2.5A). Of the twelve monomers in the asymmetric unit, three of the active sites contain the covalent modification, (A3, B3, and D1) and the remaining nine contain the NCP products. In one of the monomers that contains the covalent modification (A3; Figure 2.5B), the free carboxylic oxygen is within hydrogen bonding distance of the oxyanion hole, and replaces a water typically observed in that position (see Figure 2.2B).

Analysis of hCE1 by mass spectrometry was attempted to detect covalently bound product, but was not successful due to our inability to remove the high mannose glycosylation sites using the enzymes PNGase F and/or Endo-F1; these efforts produced extensive mass-charge heterogeneity (data not shown). Simulated-annealing omit difference

density maps, however, indicate the presence of covalent modification and of non-covalent product (NCP) binding (Figure 2.5D). This is the first structure of hCE1 in which the enzyme is observed to be covalently modified. Benzoic acid (NCP) binding is similar to other smaller ligands, including the isopentoic acid product of mevastatin (Figure 2.5C; see also Figure 2.2B). It is held in place primarily by a direct hydrogen bond to a water molecule, which exhibits clear tetrahedral geometry in the oxyanion hole. This geometry is similar to that seen for the free carboxylic oxygen in the covalent product (see Figure 2.5B). NCP products are also found in all 12 Z-sites with binding modalities similar to nonspecific binding seen in other hCE1 complexes. The observation of covalent and non-covalent complexes at the hCE1 active site helps to explain how the dione structure of benzil inhibits hCE1, and may provide an avenue to design more effective hCE1 inhibitors for clinical use.

DISCUSSION

Mevastatin

Analysis of hCE1 crystallized in the presence of mevastatin indicated that isopentoic acid was present within the active site of the protein. Because we assumed that mevastatin would be a substrate for CEs [5, 15], we hypothesized that the observed isopentoic acid resulted from hCE1-mediated β-activation of the drug (Figure 2.2A-C; the other product of this activation would be a decalin-lactone compound). At the Z-site, electron density was interpreted as the decalin-lactone compound, although it is possible that the product of α-activation may also be present (Figure 2.2A,C). *In vitro* biochemical assays using purified hCE1 and mevastatin as a substrate failed to detect the presence of isopentoic acid or the larger, decalin-lactone alcohol product. Subsequent analyses indicated that mevastatin was a partially non-competitive inhibitor of hCE1. The inhibition of hCE1 occurs at moderately low mevastatin concentrations (~20 μM), which may have biological consequences. Thus, it is possible that hCE1 does not activate mevastatin *in vivo*. Human serum paraoxonase (PON1) may instead perform this catalytic processing, as it has been previously shown to contain lactonase activity toward the statin class of compounds [41].

It is possible that the complexes of hCE1 with portions of mevastatin that we observe in our structure arose from a mechanism similar to that seen with benzil, i.e. that under prolonged incubation of enzyme with drug, very low levels of hydrolysis may occur resulting in the generation of a small molecule (isopentoic acid for mevastatin) that stabilizes protein

structure. This would favor formation of the crystal containing this moiety and hence may 'select' this structure due to increased thermodynamic stability. We would expect that crystallization trials with isopentoic acid (2-methylbutyric acid) alone would give similar results. It is possible, however, that hCE1 does play a role *in vivo* in mevastatin activation, but that our *in vitro* conditions failed to reproduce the conditions found in human tissues that enable catalytic action. For example, the processing of mevastatin by hCE1 may require the presence of other proteins necessary to traffic the substrate into or the products away from the active site of the enzyme.

Production of Ethyl Acetate

The mechanism of transesterification of Acetyl CoA with ethanol to form EA is likely a simple modification of the standard two-step serine hydrolase mechanism, with ethanol replacing water in second step that releases the bound acyl-enzyme intermediate. An open question, however, is how ethanol gains access to this intermediate, which is formed from the cleavage of the thioester bond in Acetyl CoA (Figure 2.3A). Acetyl CoA fills the active site cavity of hCE1 [42]. Thus, we suggest that ethanol accesses the active site gorge via the "side door" secondary pore first identified for the mammalian carboxylesterases in the structure of a rabbit liver CE (rCE) [13]. This pore, which lies adjacent to Thr252 at the base of the active site gorge, appears capable of allowing ethanol or water to enter the active site without having to negotiate past the bound CoA molecule. We have also seen that this pore is used by the enzyme to allow long fatty acyl regions of endogenous hCE1 substrates to extrude from the enzyme's active site [42]. Thus, the catalytic action of hCE1 on a variety of structurally-distinct substrates appears to be enhanced by the presence of an additional

channel into and out of the enzyme's active site, which enables larger and more varied substrate molecules access to the buried catalytic residues.

Sequestering of Tamoxifen

The binding of tamoxifen to hCE1 could affect the efficacy of this chemotherapeutic agent. The drug is administered in 20 mg doses once or twice daily and, upon trafficking through the liver, a fraction of it likely binds to hCE1. This may delay the passage of the drug out of the liver, or enhance its clearance by other hepatic enzymes. The use of selective inhibitors of hCE1 could improve tamoxifen efficacy by blocking tamoxifen binding by the enzyme. The non-productive binding of tamoxifen to hCE1 could also contribute to the side effects of the drug. hCE1 has been reported to contain both cholesterol ester hydrolase (CEH) and acyl-CoA cholesterol transferase activities (ACAT) [22, 23], and thus even weak inhibition of hCE1 by tamoxifen could impact cholesterol metabolism [43].

Degradation of Benzil

It was not possible to distinguish between the two NCP products of benzil hydrolysis in our 3.2 Å resolution crystal structure. By electron density considerations alone, two orientations of benzaldehyde (NCP1) could appear to be a benzoic acid molecule (NCP2). Thus, we conclude that either NCP1 or NCP2 could be present in the nine enzyme active sites where no covalent modification is observed (Figure 2.5A). We chose to place benzoic acid (NCP2) in the active sites because this single molecule satisfied the electron density; however, multiple orientations of small planar molecules are possible and have been seen previously in other hCE1-ligand structures[11]. The presence of these NCP products within

the active site of hCE1 likely results from the extended time frame required for crystal growth (two months). Indeed, similar to the structures derived from the mevastatin incubations described above, the crystallization of the hCE1-benzil complex appears to require the stabilizing effects of these small ligand products of slow substrate hydrolysis events.

Overall, these studies provide insights into the mechanism of hCE1, and into inhibition of the enzyme by a variety of molecules. Our results indicate that esters can act either as substrates or as inhibitors for hCE1; thus, predicting whether a compound will be hydrolyzed by the enzyme is difficult. In addition, our studies suggest that the use of hCE1 inhibitors of CEs could modulate drug disposition *in vivo*, as has been achieved successfully in the clinics with inhibitors of other drug metabolism enzymes [44].

EXPERIMENTAL PROCEDURES

Crystallization and Crystal Handling

A secreted form of hCE1 was expressed using baculovirus in *Spodptera frugiperda* Sf21 cells and purified as described [45, 46]. All ligands were purchased from Sigma. hCE1 was concentrated to 8-10 mg ml^{-1} in 50 mM HEPES pH 7.4, and kept at 4°C prior to crystallization. Since the solubility of the ligands varied, different methods were utilized for crystallization. Crystals of hCE1 complexes with tamoxifen and benzil were grown using the "dry-drop" method. This protocol involved dissolving the ligand in methanol at a set concentration (1 mM tamoxifen or 10 mM benzil), placing 1 μl of it in the sitting drop well, and allowing the methanol to evaporate, leaving the dried substance. A protein-mother liquor crystallization drop was then applied to the top of this dehydrated compound. For the hCE1-mevastatin complex, 1 μl of 10 mM mevastatin was added to a drop containing 2 μL of protein plus 2 μL of mother liquor. For the ethyl acetate product complex, 10 mM Acetyl-CoA in 5% ethanol was added to a drop containing 2 μL of protein plus 2 μL of mother liquor. Long, plate like crystals (up to 800x100x50 μM) were grown by sitting drop vapor diffusion at 22˚C in 8-10% (w/v) PEG 3350, 0.1 M citrate pH 5.5, 0.3 M Li_2SO_4, 0.1 M LiCl, 0.1 M NaCl, and 5% glycerol (v/v) in 1 week (*e.g.,* for the tamoxifen complex) to 2 months (*e.g.,* for the benzil complex). Crystals were cryo-protected in 40% sucrose (v/v) plus mother liquor before flash-cooling in liquid nitrogen.

Structure Determination and Refinement

Diffraction data were collected at the Advanced Photon Source at Argonne National Laboratory (Argonne, IL) on beam line 22-ID (SER-CAT), or at the Stanford Synchrotron Radiation Laboratory (Palo Alto, CA) on beam lines 9-1 and 9-2. Data were indexed and scaled using HKL2000 or using MOSFLM and SCALEPACK. The structures were determined by molecular replacement with AMoRe [47] using the hCE1-Tacrine complex structure (RCSB accession code 1MX1) as a search model. Refinement was accomplished using simulated annealing and torsion angle dynamics in CNS [30] with the maximum likelihood function target, and included an overall anisotropic B factor and bulk solvent correction. Prior to any refinement, 7% of the data was removed and set aside for cross validation using the R_{free} statistic. Noncrystallographic symmetry restraints were employed for the first round of refinement for each complex, but removed for all subsequent rounds to allow for each monomer to be refined independently. Model adjustments and manual rebuilding were accomplished using O [48] and σ_a-weighted [49] electron density maps. Asparagine-linked glycosylation sites, ligands, sulfates, water, and ions were added in the final stages of refinement to achieve R-factors listed in Table 2.1. Final structures exhibit good geometry with no Ramachandran outliers using PROCHECK [50]. Figures were constructed using Dino (www.dino3d.org), Bobscript [51], Raster3D [52], POV-RAY (www.povray.org), and PyMol [53].

Analysis of mevastatin hydrolysis

The interaction of mevastatin with hCE1 was examined by HPLC followed by LC-MS/MS. Briefly, reactions containing 125μM of drug were incubated for up to 25 days in the presence

of hCE1 at 37°C in 50mM HEPES pH7.4. The reactions were terminated by adding an equal volume of methanol and following centrifugation, the products were analyzed by reverse phase HPLC as previously described (ref 41). Peaks obtained from the chromatographic analyses were then subjected to LC-MS/MS using a Shimadzu HPLC connected to a PE SCIEX API 365 LC/MS/MS with Turbo Ion Spray and a heated nebulizer. Data were collected using Analyst version 1.4 software (API). Routinely, results obtained for the different drug metabolites were with 0.5AMU of their calculated value.

Inhibition Assays

Inhibition of hCE1 and rCE-mediated hydrolysis of 3mM o-nitrophenyl acetate was performed as described previously [54]. Briefly, a spectrophotometric assay was employed and inhibitor concentrations ranged from 128 μM to 1 nM. All assays were performed in duplicate and data was analyzed using a multifactorial equation that also assigned the mode of enzyme inhibition [55]. The equation shown below is the equation used to determine the K_i values.

$$i = \frac{[I]\{[s](1-\beta) + Ks(\alpha-\beta)\}}{[I]\{[s] + \alpha Ks\} + Ki\{\alpha[s] + \alpha Ks\}}$$

i is the fractional inhibition, [I] is inhibitor concentration, [s] is the substrate concentration, α is the change in affinity of substrate for the enzyme, β the change in rate of the enzyme-substrate (ES) decomposition, K_s is the dissociation constant of the ES complex, and K_i is the inhibitor constant. The equation can be subdivided into six smaller equations that account for the different types of inhibition. Computer analysis gives the best curve fit (r^2) value that indicates the mode of inhibition based on previously defined assumptions. Definitive

assignment of the correct mode of enzyme inhibition was determined using Akaike's Information Criteria [56, 57].

Coordinates

The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 1YA8, 1YAH, 1YA4, 1YAJ, for the mevastatin, ethyl acetate, tamoxifen, and benzil complexes, respectively.

REFERENCES

1. Williams, F.M., *Clinical significance of esterases in man.* Clin Pharmacokinet, 1985. **10**(5): p. 392-403.

2. Satoh, T. and M. Hosokawa, *The mammalian carboxylesterases: from molecules to functions.* Annu Rev Pharmacol Toxicol, 1998. **38**: p. 257-88.

3. Redinbo, M.R., S. Bencharit, and P.M. Potter, *Human carboxylesterase 1: from drug metabolism to drug discovery.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 620-4.

4. Zhang, J., et al., *Binding and hydrolysis of meperidine by human liver carboxylesterase hCE-1.* J Pharmacol Exp Ther, 1999. **290**(1): p. 314-8.

5. Tang, B.K. and W. Kalow, *Variable activation of lovastatin by hydrolytic enzymes in human plasma and liver. 4.* Eur J Clin Pharmacol, 1995. **47**(5): p. 449-51.

6. Kamendulis, L.M., et al., *Metabolism of cocaine and heroin is catalyzed by the same human liver carboxylesterases.* J Pharmacol Exp Ther, 1996. **279**(2): p. 713-7.

7. Tabata, T., et al., *Bioactivation of capecitabine in human liver: involvement of the cytosolic enzyme on 5'-deoxy-5-fluorocytidine formation.* Drug Metab Dispos, 2004. **32**(7): p. 762-7.

8. Takai, S., et al., *Hydrolytic profile for ester- or amide-linkage by carboxylesterases pI 5.3 and 4.5 from human liver.* Biol Pharm Bull, 1997. **20**(8): p. 869-73.

9. Brzezinski, M.R., et al., *Purification and characterization of a human liver cocaine carboxylesterase that catalyzes the production of benzoylecgonine and the formation of cocaethylene from alcohol and cocaine.* Biochem Pharmacol, 1994. **48**(9): p. 1747-55.

10. Bourland, J.A., D.K. Martin, and M. Mayersohn, *Carboxylesterase-mediated transesterification of meperidine (Demerol) and methylphenidate (Ritalin) in the presence of [2H6]ethanol: preliminary in vitro findings using a rat liver preparation.* J Pharm Sci, 1997. **86**(12): p. 1494-6.

11. Bencharit, S., et al., *Crystal Structure of Human Carboxylesterase 1 Complexed with the Alzheimer's Drug Tacrine: from Binding Promiscuity to Selective Inhibition.* Chemistry & Biology, 2003. **10**(April): p. 341-349.

12. Bencharit, S., et al., *Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme.* Nat Struct Biol, 2003. **10**(5): p. 349-56.

13. Bencharit, S., et al., *Structural insights into CPT-11 activation by mammalian carboxylesterases.* Nat Struct Biol, 2002. **9**(5): p. 337-42.

14. Evans, M., et al., *Medical lipid-regulating therapy: current evidence, ongoing trials and future developments.* Drugs, 2004. **64**(11): p. 1181-96.

15. Mauro, V.F. and J.L. MacDonald, *Simvastatin: a review of its pharmacology and clinical use.* Dicp, 1991. **25**(3): p. 257-64.

16. Istvan, E.S. and J. Deisenhofer, *Structural mechanism for statin inhibition of HMG-CoA reductase.* Science, 2001. **292**(5519): p. 1160-4.

17. Buzdar, A.U., *Tamoxifen's clinical applications: old and new.* Arch Fam Med, 2000. **9**(9): p. 906-12.

18. Jones, K.L. and A.U. Buzdar, *A review of adjuvant hormonal therapy in breast cancer.* Endocr Relat Cancer, 2004. **11**(3): p. 391-406.

19. Etienne, M.C., et al., *Tamoxifen metabolism: pharmacokinetic and in vitro study.* Br J Cancer, 1989. **60**(1): p. 30-5.

20. Mesange, F., et al., *Identification of two tamoxifen target proteins by photolabeling with 4-(2-morpholinoethoxy)benzophenone.* Bioconjug Chem, 2002. **13**(4): p. 766-72.

21. Morello, K.C., G.T. Wurz, and M.W. DeGregorio, *SERMs: current status and future trends.* Crit Rev Oncol Hematol, 2002. **43**(1): p. 63-76.

22. Becker, A., et al., *Purification, cloning, and expression of a human enzyme with acyl coenzyme A: cholesterol acyltransferase activity, which is identical to liver carboxylesterase.* Arterioscler Thromb, 1994. **14**(8): p. 1346-55.

23. Ghosh, S., *Cholesteryl ester hydrolase in human monocyte/macrophage: cloning, sequencing, and expression of full-length cDNA.* Physiol Genomics, 2000. **2**(1): p. 1-8.

24. Diczfalusy, M.A., et al., *Characterization of enzymes involved in formation of ethyl esters of long-chain fatty acids in humans.* J Lipid Res, 2001. **42**(7): p. 1025-32.

25. Beckemeier, M.E. and P.S. Bora, *Fatty acid ethyl esters: potentially toxic products of myocardial ethanol metabolism.* J Mol Cell Cardiol, 1998. **30**(11): p. 2487-94.

26. Lange, L.G. and B.E. Sobel, *Mitochondrial dysfunction induced by fatty acid ethyl esters, myocardial metabolites of ethanol.* J Clin Invest, 1983. **72**(2): p. 724-31.

27. Bora, P.S., et al., *Myocardial cell damage by fatty acid ethyl esters.* J Cardiovasc Pharmacol, 1996. **27**(1): p. 1-6.

28. Bora, P.S., et al., *Purification and characterization of human heart fatty acid ethyl ester synthase/carboxylesterase.* J Mol Cell Cardiol, 1996. **28**(9): p. 2027-32.

29. Wadkins, R.M., et al., *Identification and characterization of novel benzil (diphenylethane-1,2-dione) analogues as inhibitors of mammalian carboxylesterases.* J Med Chem, 2005. **48**(8): p. 2906-15.

30. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr D Biol Crystallogr, 1998. **54 (Pt 5)**: p. 905-21.

31. Thompson, J.D., D.G. Higgins, and T.J. Gibson, *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.* Nucleic Acids Res, 1994. **22**(22): p. 4673-80.

32. Kaphalia, B.S. and G.A. Ansari, *Purification and characterization of rat hepatic microsomal low molecular weight fatty acid ethyl ester synthase and its relationship to carboxylesterases.* J Biochem Mol Toxicol, 2001. **15**(3): p. 165-71.

33. Hosokawa, M., T. Maki, and T. Satoh, *Multiplicity and regulation of hepatic microsomal carboxylesterases in rats.* Mol Pharmacol, 1987. **31**(6): p. 579-84.

34. Musidlowska-Persson, A. and U.T. Bornscheuer, *Recombinant porcine intestinal carboxylesterase: cloning from the pig liver esterase gene by site-directed mutagenesis, functional expression and characterization.* Protein Eng, 2003. **16**(12): p. 1139-45.

35. Corsini, A., F.M. Maggi, and A.L. Catapano, *Pharmacology of competitive inhibitors of HMG-CoA reductase.* Pharmacol Res, 1995. **31**(1): p. 9-27.

36. Brown, M.S., J.R. Faust, and J.L. Goldstein, *Induction of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity in human fibroblasts incubated with compactin (ML-236B), a competitive inhibitor of the reductase.* J Biol Chem, 1978. **253**(4): p. 1121-8.

37. Chakravarti, R. and V. Sahai, *Compactin-a review.* Appl Microbiol Biotechnol, 2004. **64**(5): p. 618-24.

38. Endo, A., M. Kuroda, and K. Tanzawa, *Competative inhibition of 3-hydroxy-3-methylglutaryl coenzyme A reductase by ML-236A and ML-236B fungal metabolites, having hypocholesterolemic activity.* FEBS Lett., 1976. **72**(2): p. 323-6.

39. Darvesh, S., et al., *Differential effects of lipid-lowering agents on human cholinesterases.* Clin Biochem, 2004. **37**(1): p. 42-9.

40. Williams, P.A., et al., *Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone.* Science, 2004. **305**(5684): p. 683-6.

41. Billecke, S., et al., *Human serum paraoxonase (PON1) isozymes Q and R hydrolyze lactones and cyclic carbonate esters.* Drug Metab Dispos, 2000. **28**(11): p. 1335-42.

42. Bencharit, S., et al., *Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1.* J Mol Biol, 2006. **363**(1): p. 201-14.

43. Kusama, M., et al., *Effects of toremifene (TOR) and tamoxifen (TAM) on serum lipids in postmenopausal patients with breast cancer.* Breast Cancer Res Treat, 2004. **88**(1): p. 1-8.

44. Barry, M., et al., *Pharmacokinetics and potential interactions amongst antiretroviral agents used to treat patients with HIV infection.* Clin Pharmacokinet, 1999. **36**(4): p. 289-304.

45. Danks, M.K., et al., *Comparison of activation of CPT-11 by rabbit and human carboxylesterases for use in enzyme/prodrug therapy.* Clin Cancer Res, 1999. **5**(4): p. 917-24.

46. Morton, C.L. and P.M. Potter, *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): p. 193-202.

47. Navaza, J., *AmoRe: an automated package for molecular replacement.* Acta Crysallogr. A, 1994. **50**: p. 157-163.

48. Jones, T.A., et al., *Improved methods for building protein models in electron density maps and the location of errors in these models.* Acta Crystallogr A, 1991. **47 (Pt 2)**: p. 110-9.

49. Read, R.J., *Improved Fourier Coefficients for maps using phases from partial structures with errors.* Acta Crysallogr. A, 1986. **42**: p. 140-149.

50. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* J. Appl. Cryst., 1993. **26**: p. 283-291.

51. Esnouf, R.M., *Further additions to MolScript version 1.4, including reading and contouring of electron-density maps.* Acta Crystallogr D Biol Crystallogr, 1999. **55 (Pt 4)**: p. 938-40.

52. Merritt, E.A. and M.E.P. Murphy, *Raster3D version 2.0. A program for photorealistic molecular graphics.* Acta Crysallogr. D, 1994. **50**: p. 869-873.

53. DeLano, W.L., *The PyMOL Molecular Graphics System.* 2002, DeLano Scientific: San Carlos, CA, USA.

54. Wadkins, R.M., et al., *Discovery of novel selective inhibitors of human intestinal carboxylesterase for the amelioration of irinotecan-induced diarrhea: synthesis, quantitative structure-activity relationship analysis, and biological activity.* Mol Pharmacol, 2004. **65**(6): p. 1336-43.

55. Webb, J.L., *Enzyme and Metabolic Inhibitors. General Principles of Inhibition*. Vol. 1. 1963, New York, NY: Academic Press Inc.

56. Akaike, H., *Information Theory and an extension of the maximum likelihood principle*, in *Second International Symposium on Information Theory*. 1973, Akademiai kiado: Budapest.

57. Akaike, H., *A mew look at the statistical model indentification. IEEE Trans Automatic Control.* AC-19, 1974: p. 716-723.

FIGURE LEGENDS

Figure 2.1: Trimeric structure of hCE1. (**A**) Overall structure of hCE1 in complex with mevastatin. Sulfates are indicated in yellow, carbohydrates are shown in orange, and mevastatin products are depicted as yellow and red space-filling models. Monomer three is colored separately by domain with the catalytic domain in red, the α/β domain in pink, and the regulatory domain in magenta. (**B**) Expanded view of trimer interface indicated by a box in (A). The distances in angstroms are represented by black dotted lines, with the gold dotted line representing the high-mannose glycosylation chain between the initial N-acetylglucosamine and terminal the sialic acid.

Figure 2.2: Mevastatin metabolism by hCE1. (**A**) Potential mechanisms of mevastatin cleavage by hCE1. The locations of the α– and β-cleavage sites are shown in blue and red, respectively. (**B**) Stereo view of the active site of hCE1 showing the isopentoic acid product of mevastatin β-cleavage (gold) bound to the catalytic residues. Simulated annealing omit maps, at 3.0 Å resolution, are contoured to 3.0 σ (blue) and 2.0 σ (magenta). Ser221 and His468 are shown in cyan, with the catalytic water molecule in red. The oxyanion hole is represented in stick format (green), and all distances are displayed as blue dotted lines. Protein secondary structure is colored by domain according to figure 2.1A, monomer 3, in both (B) and (C). (**C**) The β-cleavage decalin product of mevastatin bound to the non-specific Z-site of hCE1. Simulated annealing omit maps are shown as in part (B). The hydrogen bonding main chain carbonyls are shown in green. Water-mediated hydrogen bonding is represented as blue lines.

Figure 2.3: FAEE-type conjugation by hCE1. (**A**) Mechanism of hCE1-dependant ethyl acetate (EA) production by transesterification of acetyl-CoA with ethanol. The transferred acetyl group is indicated in red. (**B**) Stereo view of 3.0 Å simulated annealing omit maps of EA bound at both active and Z-sites. Maps are contoured to 4.0 and 2.0 σ (blue and magenta, respectively) for the active site with 6.0 and 2.0 σ (blue and magenta, respectively) for the Z-site. (**C**) EA bound to active site of hCE1. The molecular surface of the hCE1 active site cavity is represented in green, with the catalytic triad labeled in red. (**D**) EA bound to Z-site, located slightly above the EA bound within the active site. The molecular surface of the protein is represented in green, with the Ω loops shown in grey, and the helices in dark green.

Figure 2.4: Binding of tamoxifen by hCE1. (**A**) Stereo view of simulated annealing omit map of tamoxifen (gold) bound to the active site of hCE1.Maps are at 3.2 Å resolution and contoured to 3.0 σ (blue) and 2.0 σ (magenta). (**B**) Tamoxifen bound within the active site gorge. The molecular surface of tamoxifen (green) and the surface of the hCE1 active site (gold) indicate that the drug fits in a complementary manner. Amino acid residues making hydrophobic contacts with tamoxifen are boxed in black. (**C**) Relationship of active site to the Z-site in the hCE1-tamoxifen complex. The secondary structure is labeled and tamoxifen is shown in gold.

Figure 2.5: hCE1 inhibition by benzil. (**A**) Proposed mechanism of benzil inhibition and processing by hCE1. Non covalent products NCP1 (benzaldehyde, in purple) or NCP2, (benzoic acid, in blue) are found in 9 out of 12 active sites, with the covalent product (COV)

found in 3 of 12 active sites, depicted in red. (**B**) Covalent modification of the catalytic serine by benzil. The catalytic triad is labeled in grey with the oxyanion hole in blue. (**C**) NCP2 bound to the active site of hCE1. The catalytic water molecule is represented as a red sphere and amino acids are labeled as in Figure 2.5B. (**B**) Stereo views of the simulated annealing omit maps at 3.2 Å resolution of the active site COV and NCP2 complexed with hCE1, contoured to 2.0 σ (magenta) and 4.0 σ (blue).

Figure 2.1

(a)

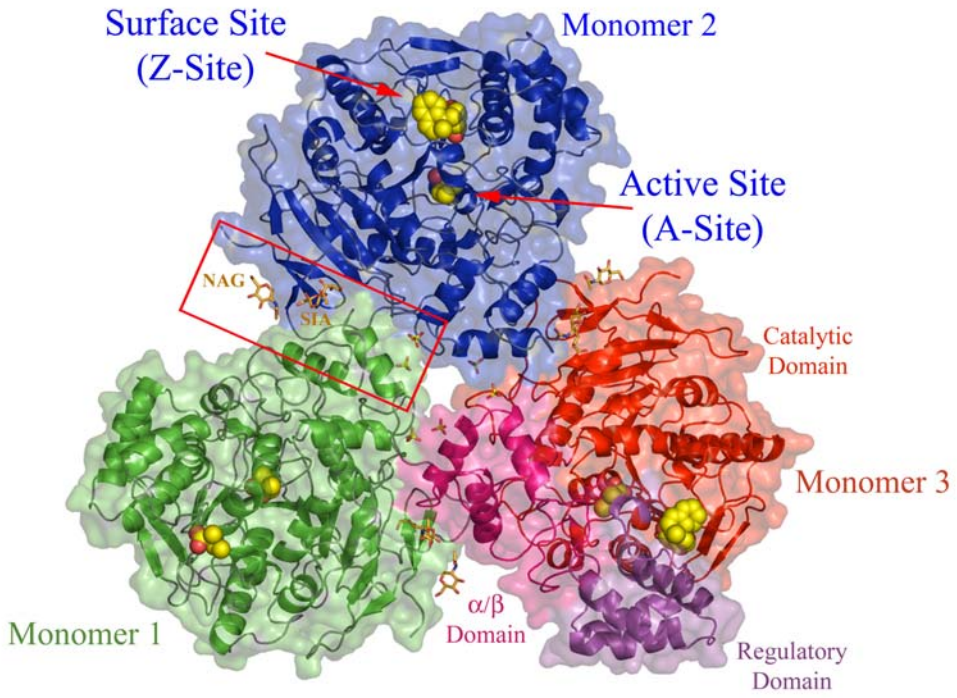

Surface Site
(Z-Site)
Monomer 2
Active Site
(A-Site)
NAG
SIA
Catalytic
Domain
Monomer 3
Monomer 1
α/β
Domain
Regulatory
Domain

Figure 2.1

(b)

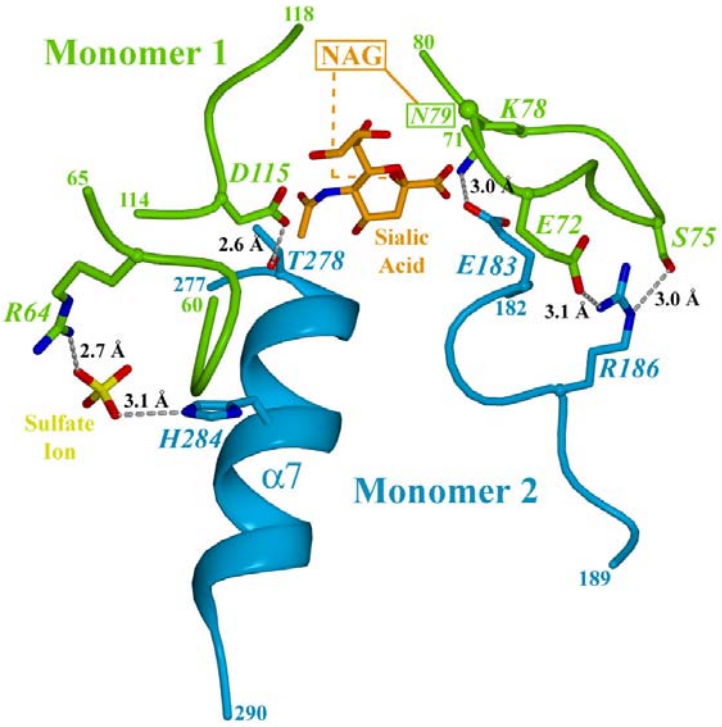

Monomer 1
118
NAG
80
N79
K78
71
D115
65
114
3.0 Å
2.6 Å
T278
Sialic
Acid
E183
E72
S75
277
60
182
3.1 Å
3.0 Å
R64
2.7 Å
R186
3.1 Å
Sulfate
Ion
H284
α7
Monomer 2
189
290

Figure 2.2

(a)

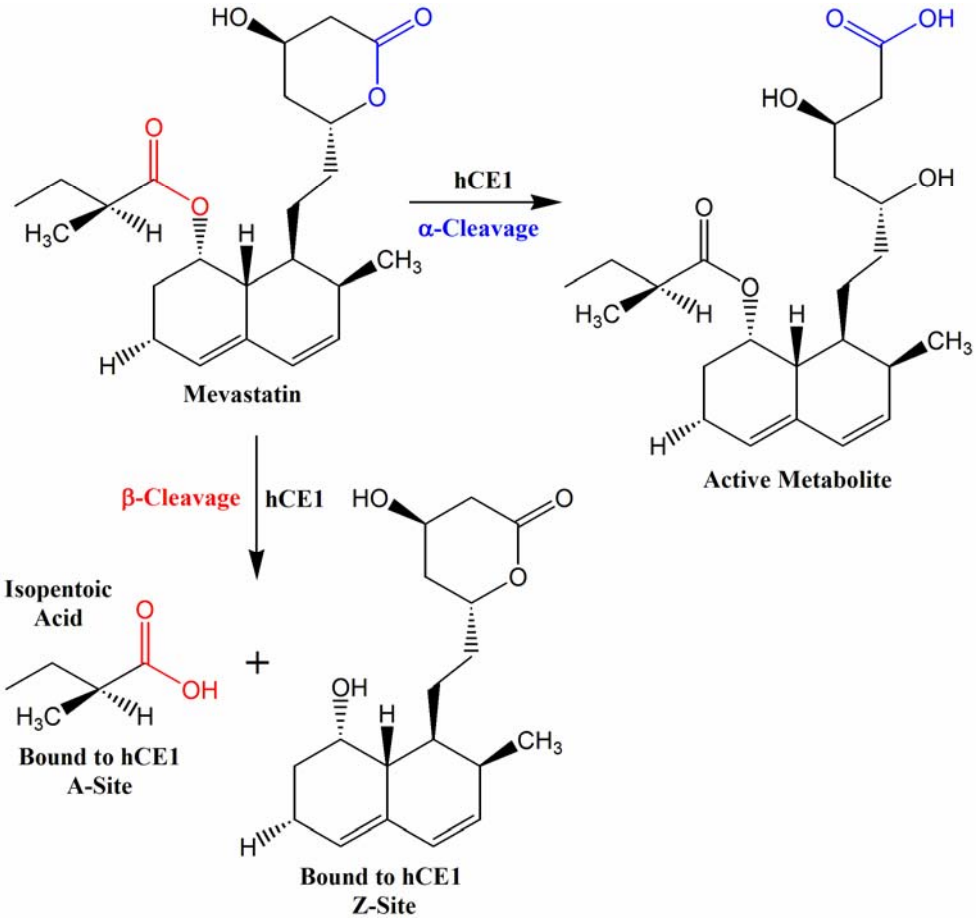

HO
O
O
O
H3C
H
H
CH3
H
Mevastatin
hCE1
α-Cleavage
O
OH
HO
OH
O
H3C
H
O
H
CH3
H
Active Metabolite
β-Cleavage
hCE1
HO
O
O
Isopentoic
Acid
O
OH
H3C
H
+
OH
H
CH3
H
Bound to hCE1
A-Site
Bound to hCE1
Z-Site

Figure 2.2

(b)

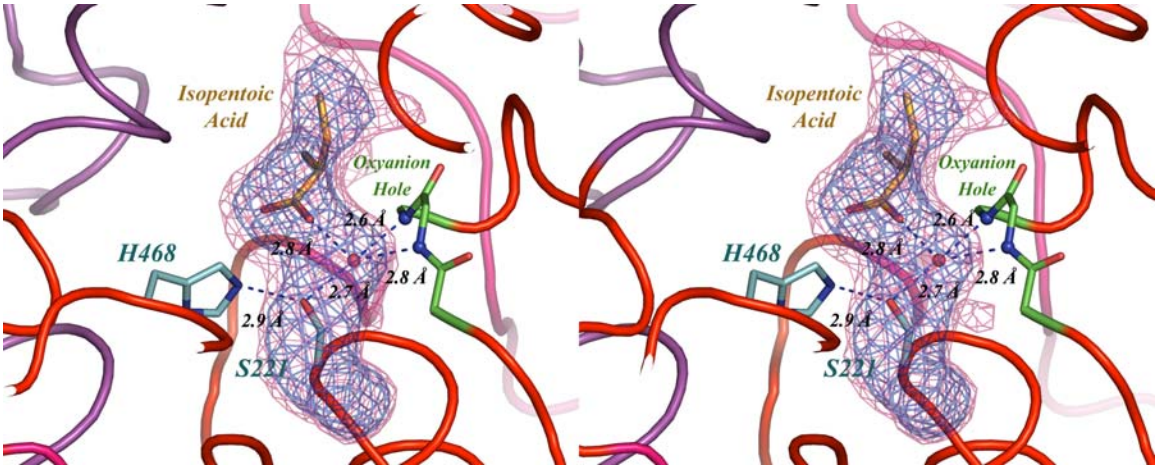

Isopentoic Acid
Oxyanion Hole
H468
S221
2.6 Å
2.8 Å
2.8 Å
2.7 Å
2.9 Å
Isopentoic Acid
Oxyanion Hole
H468
S221
2.6 Å
2.8 Å
2.8 Å
2.7 Å
2.9 Å

Figure 2.2

(c)

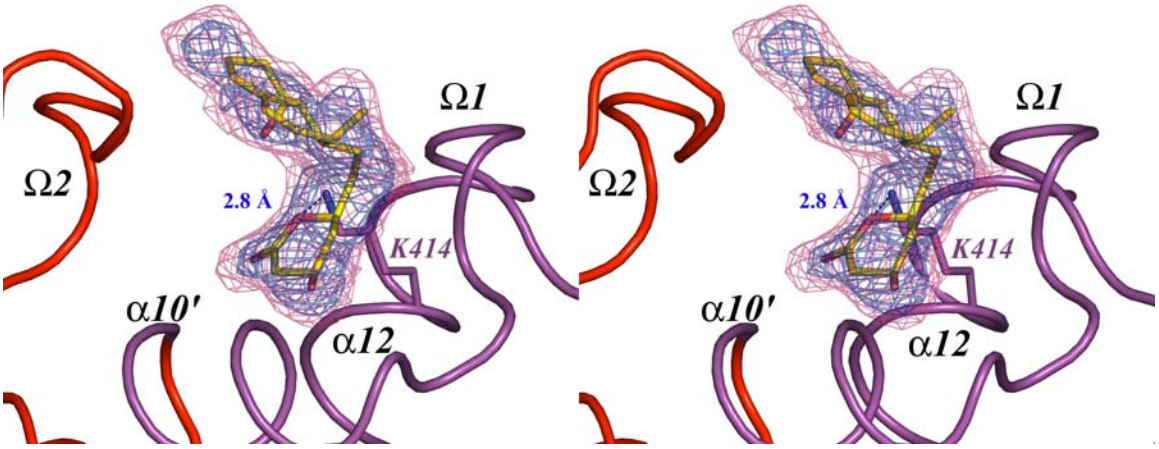


Figure 2.3

(a)

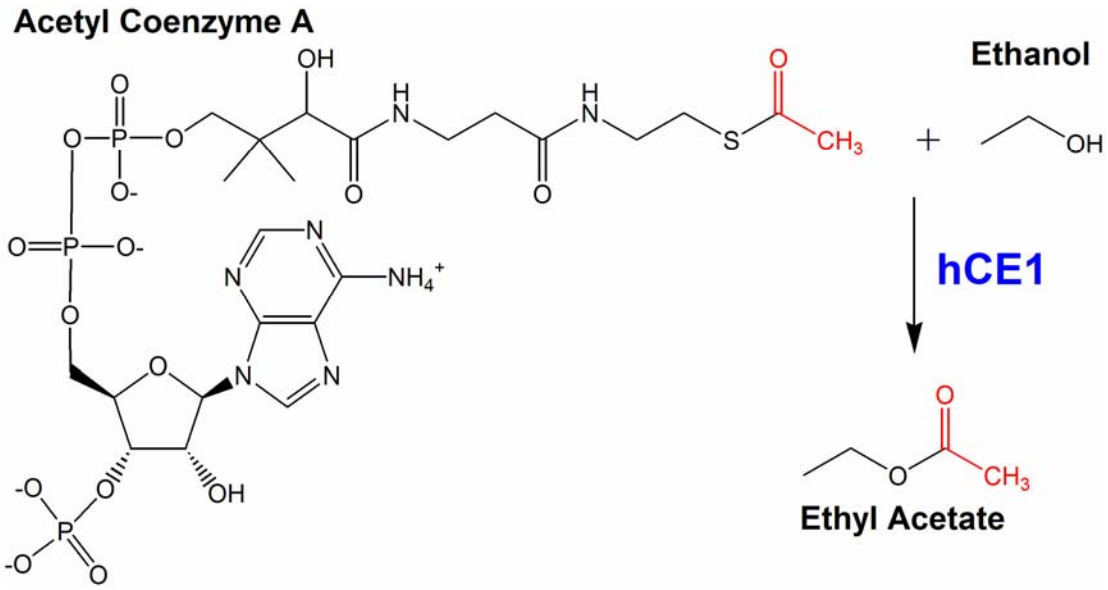


Figure 2.3

(b)

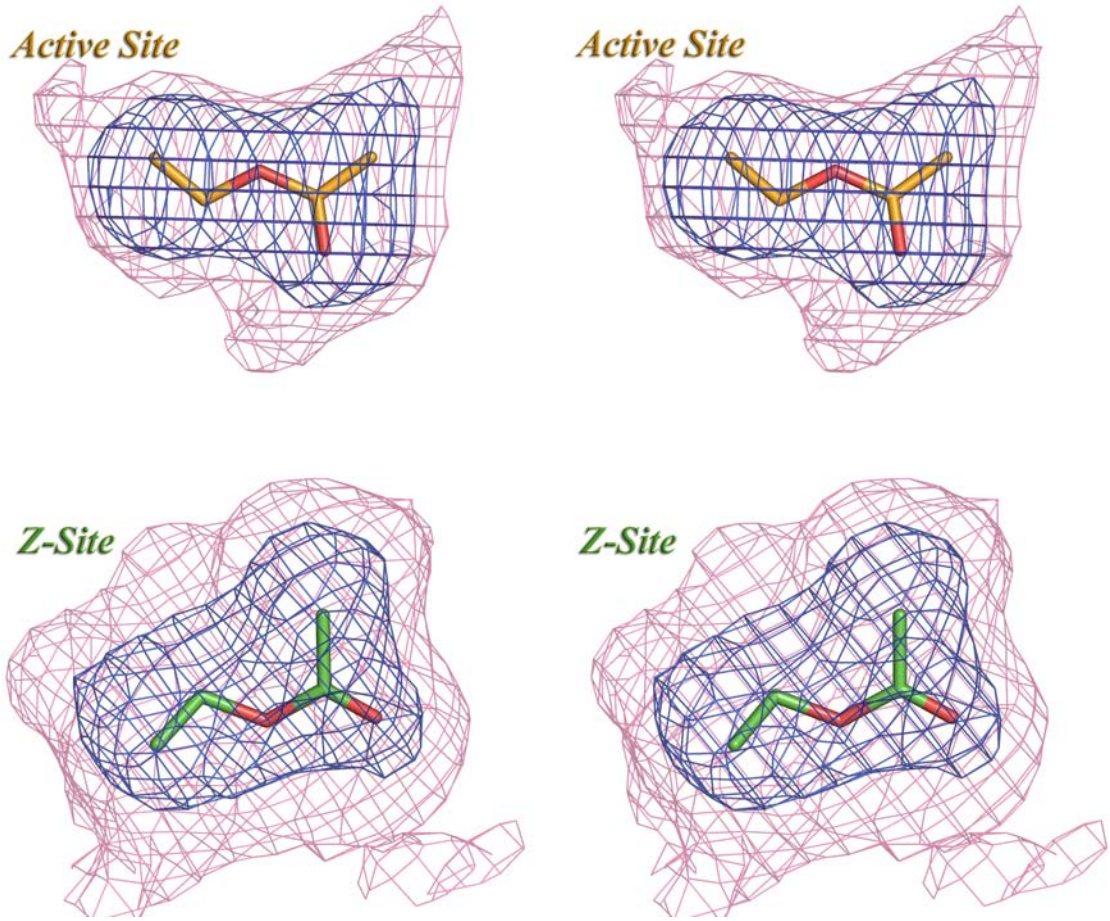


Figure 2.3

(c)

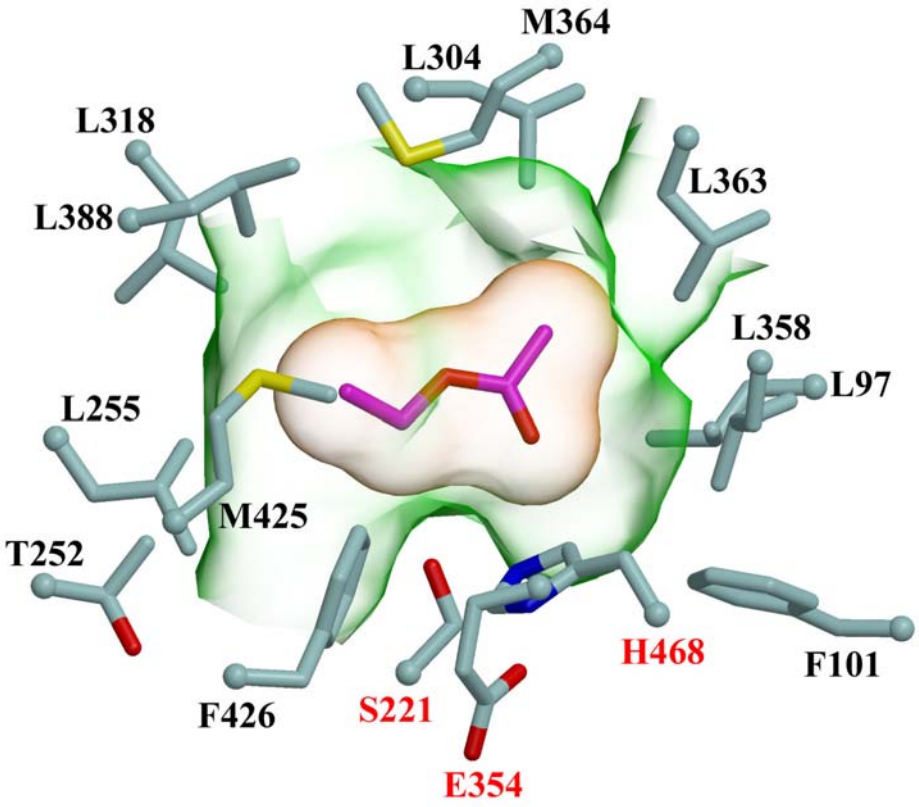

M364
L304
L318
L388
L363
L358
L97
L255
M425
T252
H468
F101
F426
S221
E354

Figure 2.3

(d)

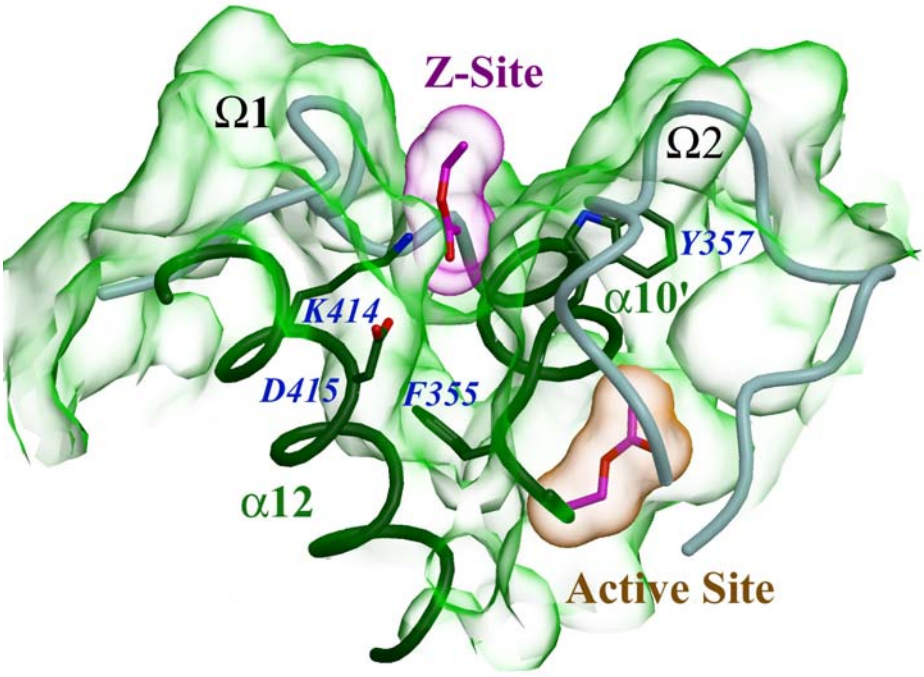


Figure 2.4

(a)

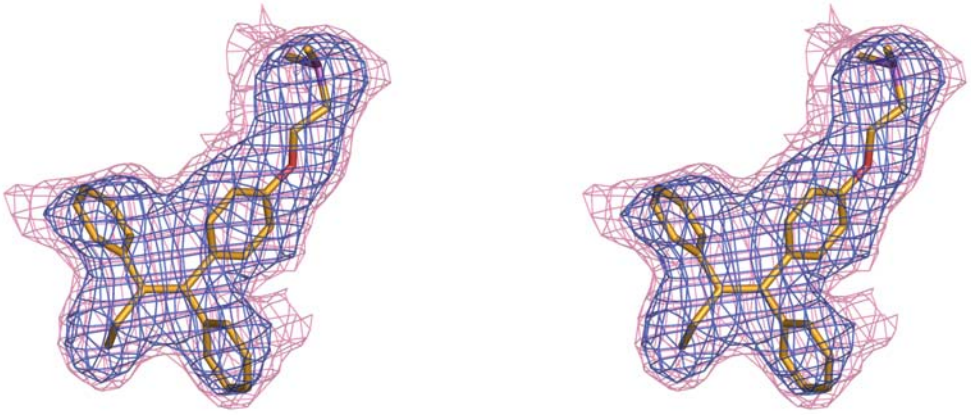


Figure 2.4

(b)

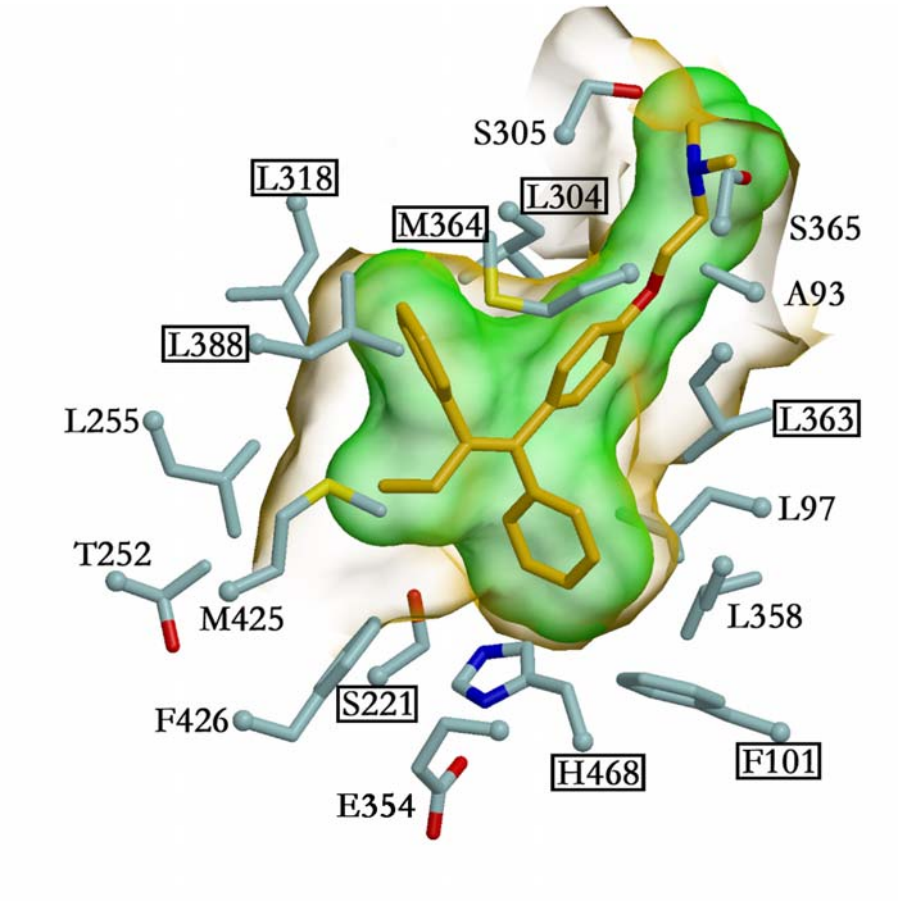

S305
L318
L304
M364
S365
A93
L388
L255
L363
L97
T252
M425
L358
S221
F426
H468
F101
E354

Figure 2.4

(c)

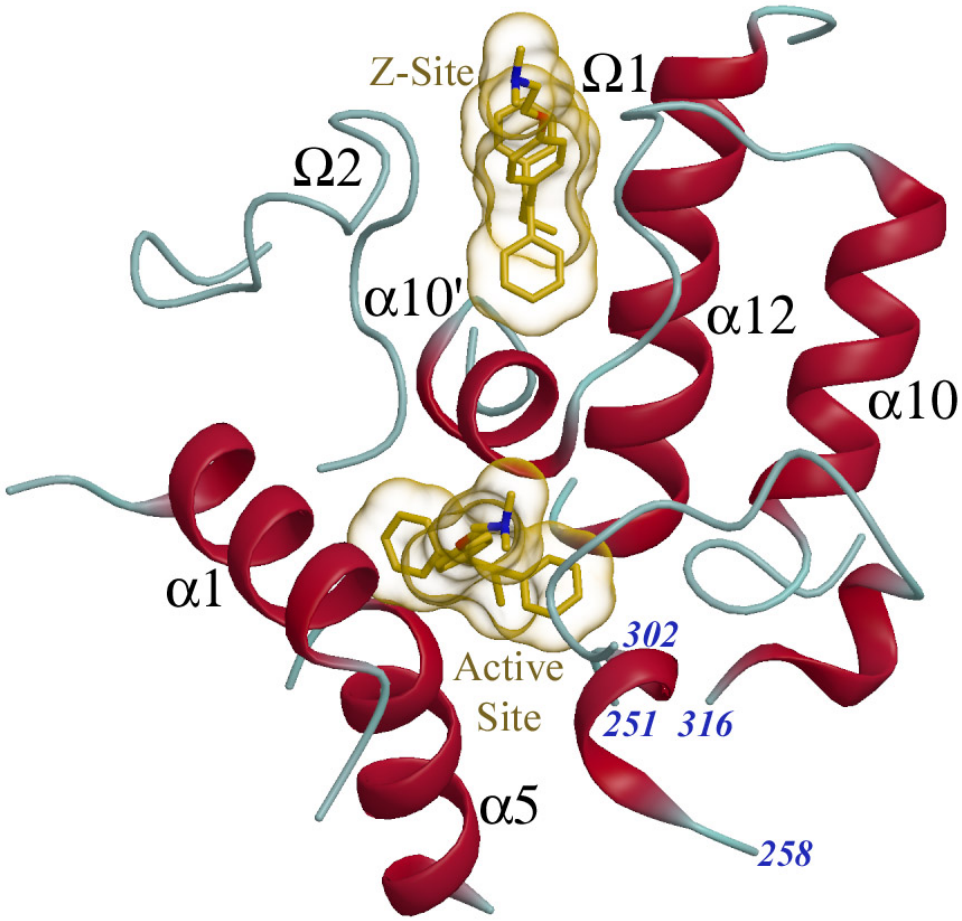


Figure 2.5

(a)

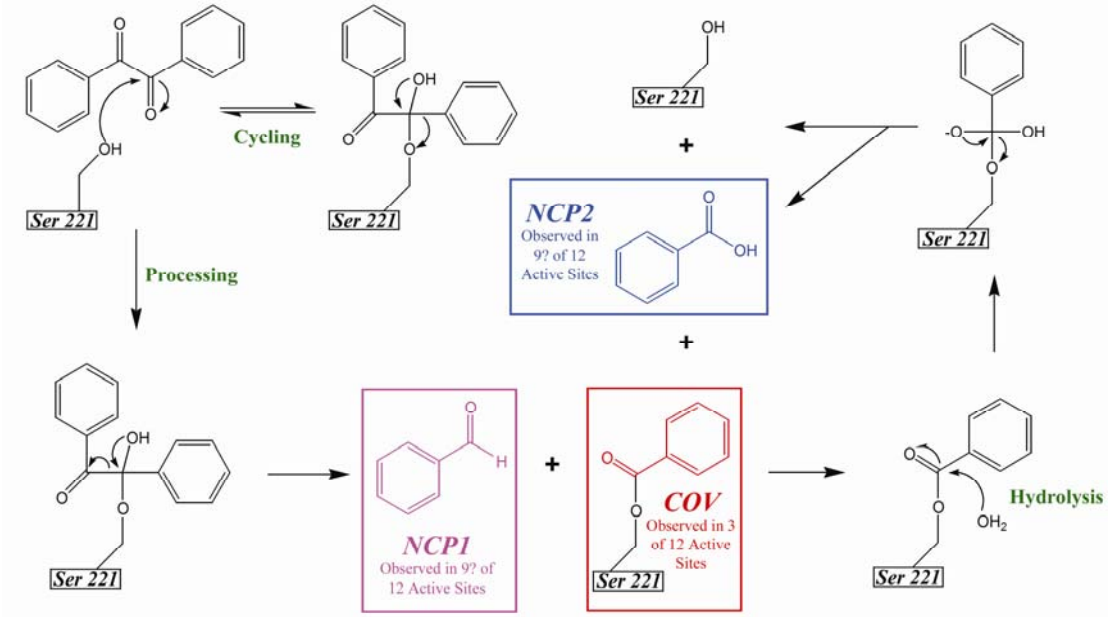

Cycling
Processing
Ser 221
NCP2
Observed in 9? of 12 Active Sites
NCP1
Observed in 9? of 12 Active Sites
COV
Observed in 3 of 12 Active Sites
Hydrolysis

Figure 2.5

(b)

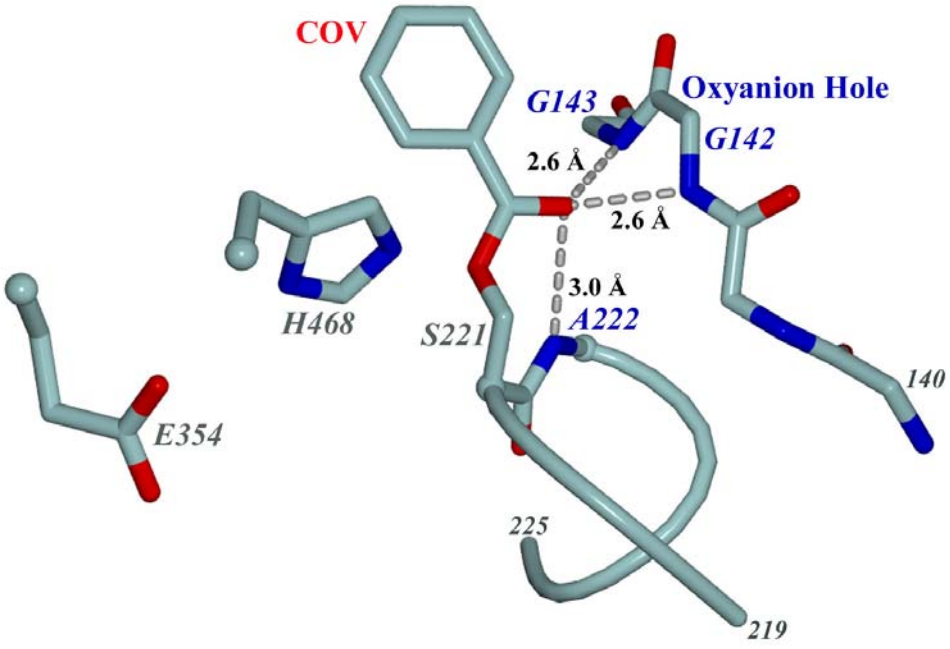

COV
G143
Oxyanion Hole
G142
2.6 Å
2.6 Å
3.0 Å
A222
S221
H468
E354
140
225
219

Figure 2.5

(c)

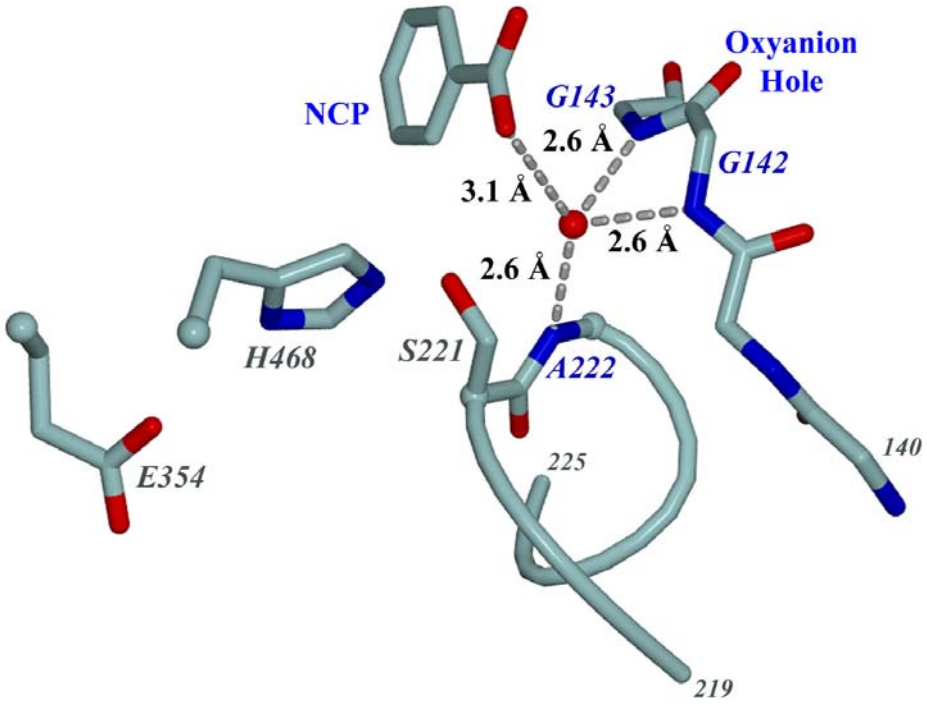

NCP
Oxyanion
Hole
G143
2.6 Å
G142
3.1 Å
2.6 Å
2.6 Å
H468
S221
A222
E354
225
140
219

Figure 2.5

(d)

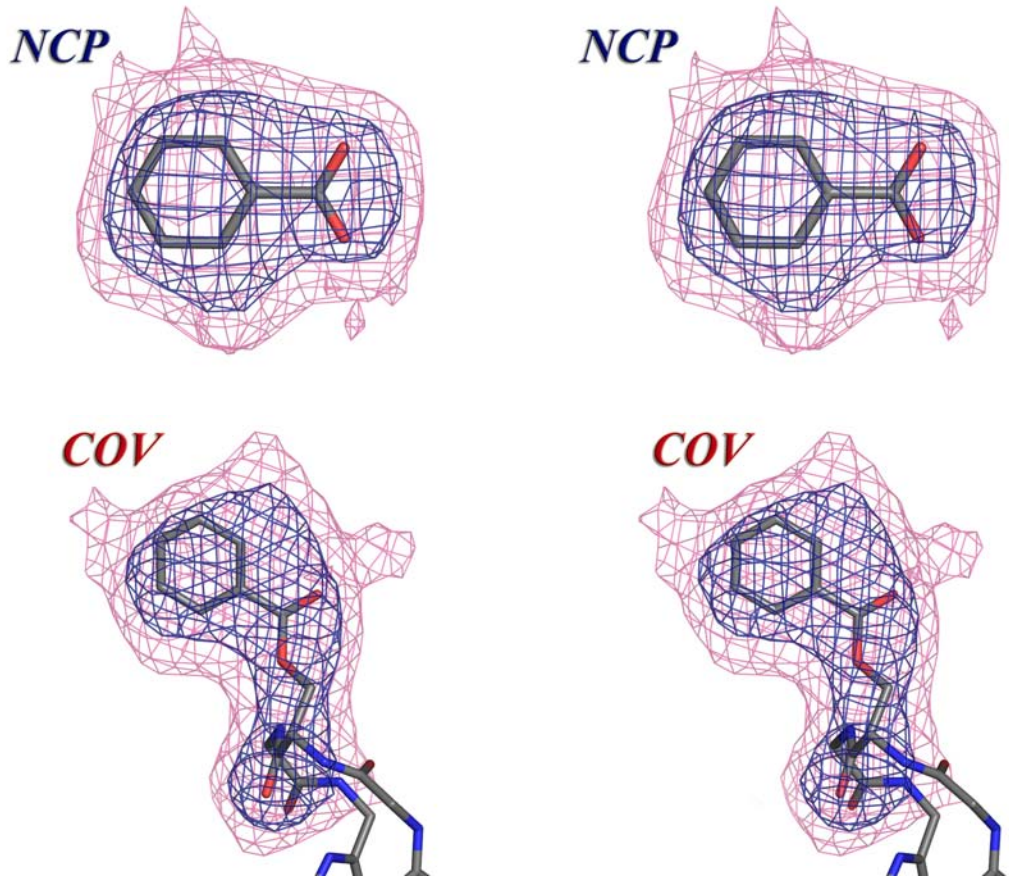


TABLE 2.1: Crystallographic Statistics

	Mevastatin	***Ethyl Acetate***	***Tamoxifen***	***Benzil***
Resolution (Å)[1]	50-3.0 (3.11-3.0)	50-3.0 (3.2-3.0)	50-3.2 (3.31-3.2)	50-3.2 (3.31-3.2)
Space Group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	P1
Asymmetric Unit	One Trimer	One Trimer	One Trimer	Four Trimers
Cell Constants (Å, °)	a=55.78 b=181.59 c=202.87 α=β=γ=90	a=55.50 b=181.15 c=203.01 α=β=γ=90	a=55.37 b=179.59 c=201.58 α=β=γ=90	a=54.56 b=181.49 c=202.71 α=90.12 β=89.93 γ=89.72
Total Reflections	462,117	289,419	690,386	681, 289
Unique Reflections	42,843	41,952	34,502	125, 020
Mean Redundancy	10.8	6.9	20.0	5.5
R_{sym} (%)[1,2]	12.8 (41.5)	13.5 (46.7)	13.5 (32.6)	15.6 (44.6)
Wilson B Factor (Å^2)	29.7	45.3	37.3	32.2
Completeness (%)[1]	98.1 (97.5)	99.7 (98.6)	98.5 (99.6)	97.2 (96.9)
Mean I/σ[1]	11.3 (3.7)	10.5 (2.1)	7.0 (2.2)	3.9 (1.8)
R_{cryst} (%)[1,3]	18.7	19.0	20.3	20.7
R_{free} (%)[1,4]	24.7	22.9	25.2	28.7
Number of Atoms:				
Protein	12,391	12,390	12,385	50,078
Solvent	347	420	193	468
Carbohydrate	105	105	105	420
Ligand	74	36	168	227
Ion	30	30	30	120

[1] Number in parentheses is for the highest shell.

[2] $R_{sym} = \Sigma|I - \langle I \rangle| / \Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity of multiple symmetry-related observations of that reflection.

[3] $R_{cryst} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

[4] $R_{free} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$ for 7% of the data not used at any stage of structural refinement.

TABLE 2.2: Kinetic Data for Tamoxifen Inhibition

	r^2 for data values using indicated mode of enzyme inhibition							
Enzyme	GE	C	PC	NC	**PNC**	M	UC	K_i (μM)
hCE1	0.987	0.956	0.956	0.956	**0.987**	0.956	0.956	15.2
rCE	0.987	0.923	0.923	0.923	**0.984**	0.923	0.923	23.4

General equation versus partially non-competitive equation						
	Akaike's Information Criteria	Favored Model	ΔAIC	Probability favored model is correct	Probability GE is correct	Ratio of Prob.
hCE1	-55.59 vs. -61.59	**PNC**	-6.00	95.26 %	4.74 %	20.09
rCE	-36.53 vs. -52.21	**PNC**	-15.68	99.06 %	0.04 %	2532

General equation (GE):

$$i = \frac{[I]\{[s](1-\beta) + Ks(\alpha-\beta)\}}{[I]\{[s] + \alpha Ks\} + Ki\{\alpha[s] + \alpha Ks\}}$$

Competitive inhibition (C):	Assume $\alpha = \infty$
Partially competitive (PC):	Assume $1 < \alpha < \infty$; $\beta = 1$
Non competitive (NC):	Assume $\alpha = 1$; $\beta = 0$
Partially noncompetitive (PNC):	Assume $\alpha = 1$; $0 < \beta < 1$
Mixed (M):	Assume $1 < \alpha < \infty$; $\beta = 0$
Uncompetitive (UC):	Assume $\alpha < 1$; $\beta < 1$

Chapter 3:

Crystal Structures of Human Carboxylesterase 1 in Covalent Complexes with the Chemical Warfare Agents Soman and Tabun

Christopher D. Fleming[#], Carol C. Edwards[||], Stephen D. Kirby[§], Donald M. Maxwell[§], Philip M. Potter[||], Douglas M. Cerasoli[§], and Matthew R. Redinbo[#,*]

[#]Department of Chemistry and Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599; [||]Department of Molecular Pharmacology, St. Jude Children's Research Hospital, Memphis, TN, 38105; and [§]U.S. Army Medical Research Institute of Chemical Defense, Aberdeen Proving Ground, MD, 21010.

Published in *Biochemistry*, 46(17), 5063-5071 (2007)

ABSTRACT

The organophosphorus nerve agents sarin, soman, tabun, and VX exert their toxic effects by inhibiting the action of human acetylcholinesterase, a member of the serine hydrolase superfamily of enzymes. The current treatments for nerve agent exposure must be administered quickly to be effective and they often do not eliminate long-term toxic side effects associated with organophosphate poisoning. Thus, there is significant need for effective prophylactic methods to protect at-risk personnel from nerve agent exposure, and protein-based approaches have emerged as promising candidates. We present the 2.7 Å resolution crystal structures of the serine hydrolase human carboxylesterase 1 (hCE1), a broad-spectrum drug metabolism enzyme, in covalent acyl-enzyme intermediate complexes with the chemical weapons soman and tabun. The structures reveal that hCE1 binds stereoselectively to these nerve agents; for example, hCE1 appears to react preferentially with the 10^4-fold more lethal P_S stereoisomer of soman relative to the P_R form. In addition, structural features of the hCE1 active site indicate that the enzyme may be resistant to dead-end organophosphate aging reactions that permanently inactivate other serine hydrolases. Taken together, these data provide important structural details toward the goal of engineering hCE1 into an organophosphate hydrolase and protein-based therapeutic for nerve agent exposure.

INTRODUCTION

The organophosphorus (OP) nerve agents sarin, soman, tabun, and VX are among the deadliest man-made chemicals [1]. While the military use of OP nerve agents is widely banned, these compounds have been employed in recent decades by rogue states and terrorist groups. In 1988, Iraq employed weaponized sarin against its own Kurdish citizens in Halabja, a town adjacent to the Iranian border, killing an estimated 5,000 [2]. Coordinated attacks on the Tokyo subway system by the Japanese Aum Shinrikyo cult in 1995 also employed sarin, killing 12 and injuring nearly a thousand [3]. The level of OP toxicity is striking, with LD_{50} values for percutaneous exposure ranging from 1 gm/person (tabun) to 10 mg/person (VX) [1]. For these reasons, there is considerable interest in developing methods to detect OP nerve agents, to treat exposed patients, and to detoxify contaminated surfaces.

OP toxicity is thought to be mediated through the inhibition of human acetylcholinesterase (AcChE), a serine hydrolase responsible for processing the neurotransmitter acetylcholine and thus terminating cholinergic nerve impulses [4]. The serine residue in the active site of AcChE is activated by adjacent histidine and glutamic acid residues, which together constitute the enzyme's catalytic triad [5]. Nucleophilic attack on OPs by the active site serine displaces a strong leaving group from the chiral phosphate and creates a covalent acyl-enzyme intermediate [6]. This reaction step has been reported to be S_N2, which would result in the stereoinversion of the chiral phosphate upon binding [6]. Typically, acyl-enzyme intermediates in serine hydrolases are removed via hydrolysis, releasing an alcohol product; however, OPs have been observed to undergo an "aging"

reaction that leads to a permanent alkyl-phosphate adduct bound to the catalytic serine. Aging, which involves dealkylation or, in the case of tabun, deamidation of the acyl-enzyme intermediate, has been hypothesized to proceed either through a hydrolytic (P-O or P-N bond scission) or a carbocation (C-O bond scission) pathway [7]. Prior to aging, the activity of the OP-inhibited serine hydrolases may be restored through reactivation of the catalytic serine with strong nucleophiles such as oximes; once aging occurs, however, reactivation is impossible with such methods [1]. Aging of AcChE with nerve agents can occur within minutes (e.g., for soman) [8, 9] or can take hours (e.g., for VX) [7].

The existing treatments for nerve agent exposure must be administered quickly to be effective, and they often do not eliminate long-term toxic side effects [10]. Atropine, a muscarinic receptor antagonist, blocks cholinergic parasympathetic neurons to maintain capacity for respiration, and is typically administered as soon as possible after an OP exposure [11]. Oxime reactivators like pralidoxime (2-PAM) or the Hagedom oxime HI-6 are also typically provided in an attempt to recover AcChE activity [12]. Finally, an anticonvulsant like diazepam is commonly administered to treat seizures that can develop, albeit via unknown mechanisms [13]. If this series of compounds is not given within minutes, however, victims quickly succumb [10]. In addition, even if rapidly treated, many patients exposed to OP pesticides experience long-term brain damage, permanent electrocardiogram changes, or an “intermediate syndrome” associated with persistent muscle weakness [14].

There has been considerable interest in utilizing enzymes to detoxify OP nerve agents *in vivo*. Human butyrylcholinesterase (BuChE) shares 55% sequence identity and 0.8 Å root

mean square deviation (rmsd) in Cα positions with human AcChE. While exhibiting no natural hydrolytic activity towards nerve agents (with k_i values between 2.2 and 3.0×10^5 M^{-1} s^{-1} for sarin and VX, respectively), BuChE functions well as a bioscavenger, in that it binds and sequesters nerve agents as OP-enzyme complexes [15, 16]. Mutagenesis of residues adjacent to the BuChE active site led to the identification of variants that retain esterase activity and increase rates of reactivation 100-fold over the wild-type enzyme [15, 17]. Transgenic mice expressing one such mutant of human BuChE, Gly-117-His, have been shown to possess modestly increased resistance to intoxication by OP pesticides [18]. In spite of these advances, an efficient prophylactic enzyme-based treatment for nerve agent exposure has remained elusive. Most mutant and wild-type forms of BuChE are subject to the same irreversible aging reactions as AcChE. Furthermore, relatively large doses of these stoichiometric scavenger enzymes must be administered for effective protection [19]. For these reasons, additional enzymes have been examined for potential conversion into efficient OP hydrolases.

Human carboxylesterase 1 (hCE1) has been proposed as a candidate for development into an OP hydrolase [20-22]. Carboxylesterases (CEs; EC 3.1.1.1) are members of the same serine hydrolase superfamily that contains AcChE and BuChE, and some CEs exhibit natural OP hydrolysis activity. For example, rat serum CE has been shown to metabolize sarin at least as efficiently as the Gly-117-His mutant form of human BuChE [20]. hCE1 (78% identity to rat serum CE) metabolizes carboxylester, amine ester, and thioester linkages in a variety of endogenous and xenobiotic compounds [22-24]. This glycoprotein shares 34% identity and 1.2 Å rmsd over Cα positions with AcChE, and utilizes the same two-step serine

hydrolase catalytic mechanism. It has been shown that AcChE, BuChE and mammalian CE exhibit a similar range of inhibition rate constants ($\log_{10} k_i$) of between 6.7 and 7.9 $M^{-1}min^{-1}$ for soman [21]. Several crystal structures of hCE1 have been determined, including complexes with heroin and cocaine analogues, the clinical therapeutics tamoxifen, tacrine, and mevastatin, and the endobiotics taurocholate, cholate, and coenzyme A [25-28]. These structures highlight the promiscuous ability of the enzyme to bind a broad spectrum of structurally diverse compounds, a characteristic that would be important for an OP hydrolase with specificity for a variety of OP compounds. Here, we present the crystal structures of hCE1 in covalent acyl-enzyme intermediate complexes with the nerve agents soman and tabun. These structures reveal that hCE1 exhibits key structural differences in relation to the AcChE [6, 29] and BuChE [7] complexes with OPs elucidated previously, features that may be exploited in the rational design of hCE1 into an efficient, broad-spectrum OP hydrolase.

MATERIALS AND METHODS

Nerve Agent Treatment and Crystallization.

A secreted form of hCE1 was expressed using baculovirus in *Spodoptera frugiperda* Sf21 cells and purified as previously described [30, 31]. Milligram quantities of purified enzyme were subsequently incubated with ~10-fold molar excesses of racemic mixtures of soman (GD) and tabun (GA) (obtained from the Research Development and Engineering Command, Aberdeen Proving Ground, MD) for 1 hour at room temperature. Excess agent was removed using size exclusion chromatography by passing the enzyme over PD-10 sephadex G-25 columns (Amersham Biosciences, Uppsala, Sweden). The treated enzyme was tested to confirm the absence of both carboxylesterase activity (indicating complete inhibition of the hCE1) and the capacity to subsequently inhibit BuChE (indicating removal of unbound OP). Samples were then concentrated to 3-5 mg/ml using Amicon Ultra-15 (Millipore) spin concentrators. Long, plate-like crystals (up to 500 μm × 100 μm × 40 μm) were grown by sitting drop vapor-diffusion methods in 9-13% PEG 3350, 0.1-0.4 M Li_2SO_4, 0.1 M Citrate (pH 5.5), 0.1 M NaCl, 0.1 M LiCl, and 5% Glycerol over a period of 1-4 weeks. Crystals were cryoprotected step-wise fashion into 30% (w/v) sucrose plus mother-liquor before cooling to 100K in a liquid nitrogen cryo-stream.

Structure Determination and Refinement.

Diffraction data were collected at 100K at the Advanced Photon Source at Argonne National Laboratory (Argonne, IL) at beamline 22-ID (SER-CAT) for the hCE1-tabun complex, and at 23-IDD (GM/CA-CAT) for the hCE1-soman complex. Data were indexed

and scaled using HKL-2000 [32]. Molecular replacement was conducted using the program MolRep in the CCP4i suite [33] (v6.0) using one trimer of the hCE1-Tacrine structure (RCSB PDB Accession code 1MX1, [26]) as a search model. Refinement was accomplished using simulated annealing and torsion angle dynamics in CNS [34], and included an overall anisotropic B-factor and bulk solvent corrections as well as initial NCS restraints. A subset (7%) of the data for each structure were set aside for cross-validation by R_{free} calculation prior to any structural refinement. Manual rebuilding was conducted using O [35] and σ_a-weighed electron density maps [36]. Data collection and refinement statistics are outlined fully below (Table I). Final structures were validated using PROCHECK [37], and all figures were generated in PyMol (http://pymol.sourceforge.org, [38]).

RESULTS

Crystallographic Analysis.

The hCE1-soman and hCE1-tabun complex structures were determined by molecular replacement, and refined to 2.7 Å resolution in space group $P2_1$ (hCE1-soman) or $P2_12_12_1$ (hCE1-tabun). These space groups have been observed in previous hCE1 structures [25-28], and correspond either to two ($P2_1$, hCE1-soman) or one ($P2_12_12_1$, hCE1-tabun) trimer per asymmetric unit. While the cell constants of the hCE1-soman structure appear to be capable of adapting $P2_12_12_1$ space group symmetry, R_{sym} values establish $P2_1$ as correct for this complex. The initial maps were of high quality, and were easily traceable for any deviations that occurred between the search model and the new structures. The final R factors were 17.0% (R_{cryst}) and 22.5% (R_{free}) for the hCE1-soman complex and 16.7% (R_{cryst}) and 23.0% (R_{free}) for the hCE1-tabun complex (Table I).

Domain Architecture of hCE1.

The 62 kDa hCE1 monomer is comprised of three domains and two ligand binding sites (Figure 3.1). The catalytic domain exhibits the canonical α/β-hydrolase fold and contains a high-mannose glycosylation site at Asn79 that is critical to protein stability and function [27]. The αβ and regulatory domains sit adjacent to the active site, with the latter flanking the surface ligand-binding site termed the Z-site [27]. The binding of various small molecules to the Z-site has been shown to regulate a trimer-hexamer equilibrium of hCE1. In both structures presented here, the Z-site contains bound sucrose, the crystallographic cryoprotectant utilized in these studies (Figure 3.1). The active site of the enzyme is located

at the interface of the three domains and is composed of the catalytic triad Ser221, His468, and Glu354. The hCE1 trimer exhibits C3 symmetry and is formed largely by contacts between the αβ domains of each monomer (not shown) [25, 27, 28]. A hexamer has also been observed for hCE1 in which two trimers are stacked with their active sites facing in and with the Z-site loops interdigitating [26, 27]; neither of the structures described here form a hexamer, however.

Structure of the hCE1-Soman Complex.

Initial difference density (F_o-F_c) maps in the hCE1-soman structure indicated a 7σ electron density peak 1.6 Å away from the Oγ of Ser221 in all six active site cavities, suggesting the presence of a covalently-attached phosphoryl adduct. At lower sigma levels (*e.g.,* 3-5 σ), the maps revealed additional density to suggest the presence of the acyl-enzyme intermediate, in which the fluoride ion has been displaced by the attack of the enzyme's catalytic serine (Figure 3.2A). This intermediate was built into the model using this electron density and refined well, yielding a covalent adduct at the active site that exhibits P_R stereochemistry for the chiral phosphorus atom (Figure 3.2B). The covalent adduct is stabilized by two hydrogen bonds (2.8 Å and 2.5 Å) between the phosphoryl oxygen and the main-chain nitrogens of Gly141 and Gly143 in the enzyme's oxyanion hole. The large, methylpinacolyl alkoxy group of the soman adduct is directed towards the spacious region of the active site cavity adjacent to Met364, while the methyl group occupies the smaller, rigid pocket near Phe101. This rigid pocket has been previously shown to select for the small acetyl and methyl ester linkages in heroin and cocaine [27].

The P_R stereochemistry observed for the covalent adduct in this structure suggests that hCE1 selectively reacted with the P_S stereoisomer of soman, which is 10,000-fold more effective than P_R at inhibiting AcChE [39], based on an S_N2 stereoinversion [40, 41]. A 2.7 Å resolution simulated annealing omit map confirmed the stereochemistry of the covalent adduct (Figure 3.3A). We also placed the presumably incorrect P_S stereoisomer into the original model (before refinement or any ligands were added) and performed one round of structural refinement. Distinct positive and negative difference density ($|F_{obs}-F_{calc}|$, ϕ_{calc}) peaks support the conclusion that the acyl-enzyme intermediate observed in the structure is in the R conformation (Figure 3.3B). Structural constraints in the hCE1 active site also indicate that P_S-soman is unlikely to be processed by the enzyme because this stereoisomer's alkoxy group is too large to be accommodated in the small, rigid pocket of the catalytic gorge. Soman also contains a second stereocenter located at the C5 atom in the methylpinacolyl alkoxy group. While it is observed in the C_R orientation in this structure, there appear to be no structural constraints that would prevent the C_S orientation from being accommodated in the hCE1 active site. Taken together, these data suggest that hCE1 is selective for the lethal P_S stereoisomer of soman.

Structure of the hCE1-Tabun Complex.

Similar to the hCE1-soman structure, initial difference density in the hCE1-tabun structure indicated that the active site Ser221 in this structure had been covalently modified. Subsequent refinement confirmed the presence in the hCE1 active site of the acyl-enzyme intermediate of tabun, which is generated by the replacement of the cyano group with the enzyme's catalytic serine (Figure 3.4A, B). In this covalent complex, the alkoxy group is

positioned in the large pocket adjacent to Met364, while the *N*,*N*-dimethylamine group is located in the smaller, rigid pocket; these are orientations similar to those observed in the hCE1-soman complex. The covalent adduct in this structure exhibits P_S stereochemistry, which indicates that the protein molecules in the crystal reacted with the P_R isoform of tabun. As with soman, simulated annealing omit map was used to confirm the stereochemistry of the adduct (Figure 3.5A). This revealed that the correct adduct stereoisomer was placed in the active site cavity. We also placed the presumably incorrect P_R stereoisomer into the original model (prior to any refinement or the addition of ligands) and conducted one round of structural refinement. While the resultant difference density maps were not as dramatic as those obtained for the incorrect P_S adduct of soman (due to the similarity in size between the two R groups of tabun), they support the conclusion that the stereochemistry of the hCE1-tabun adduct is P_S, and indicate that the hCE1 molecules present in the crystal examined reacted with P_R-tabun (Figure 3.5B). Finally, the P_S adduct satisfies all difference density after several rounds of refinement (Figure 3.6).

Comparison to Cholinesterase-Nerve Agent Structures.

We next compared the hCE1 complexes reported here to two covalent acyl-enzyme intermediate complexes of AcChE with nerve agents: the *M. musculus (Mu)* AcChE with tabun [42] and the *T. californica (Tc)* AcChE with VX [29]. The *Mu*AcChE-tabun complex exhibits an acyl-enzyme intermediate in the P_R conformation, indicating that the enzyme reacted with the P_S form of tabun (Figure 3.7A). This observation is in agreement with the work of Degenhart, et al. that shows AcChE is inhibited 6-fold more readily by P_S-tabun relative to the P_R isoform [43]. Note, however, that hCE1 and AcChE exhibit acyl-enzyme

intermediates with the opposite stereochemistry: P_S for hCE1 and P_R for AcChE (Figure 3.7A). The structure of *Tc*AcChE in covalent complex with VX reveals a P_R covalent adduct, indicating that the enzyme reacted with the S-isomer of VX (VX_S) (Figure 3.7B). VX_S is known to be 115-fold more potent than VX_R towards AcChE [44]. Relative to hCE1, the *Tc*AcChE-VX complex also exhibits alternative organization of its active site with respect to the acyl-enzyme intermediate. The *Tc*AcChE-VX complex places its larger ethoxy-group adjacent to H400, while hCE1 positions the larger methylpinacolyl alkoxy group of its soman complex away from its catalytic H468 (Figure 3.7B). Although limited biochemical work has been conducted with CEs and their stereoselectivity towards nerve agents, the guinea pig serum CE is known to be 100-fold more susceptible to P_S-soman [45], in accordance with the structural data presented above for hCE1. Taken together, these observations support the conclusion that the hCE1-acyl-enzyme intermediates observed are formed via an S_N2-like stereoinversion, and generate distinct covalent complexes relative to the cholinesterases.

The two covalent hCE1-complexes presented here also indicate that hCE1 does not undergo aging (dealkylation) when inhibited by soman or tabun. While the volumes of the catalytic gorges of hCE1, AcChE, and BuChE are similar (924, 993, 982 $Å^3$, respectively) [46], their active site architectures are significantly distinct. These differences are largely generated by the position of the so-called “acyl” loop, which has been previously suggested to be involved in the aging process of AcChE [6]. The acyl loops of the cholinesterases place large phenylalanine or leucine residues adjacent to the catalytic serine, producing a small pocket that can accommodate the acyl group of an OP but not the larger alkoxy groups of

nerve agents like soman (Figure 3.8). As a consequence, the alkoxy moieties in these cholinesterases structures are within ~3.5 Å of the catalytic triad His (Figure 3.7), which has been proposed to facilitate aging [6, 29, 42]. In contrast, in hCE1 the corresponding loop is shifted by ~7 Å away from the adduct and is stabilized by two rigid proline residues, which generate the large pocket that accommodates the alkoxy group in the soman- and tabun-bound structures presented here (Figures 3.2, 3.4, 3.8). These binding modes position the alkoxy oxygens outside hydrogen-bonding range (4.7-5.1 Å) with respect to His468 of the hCE1 catalytic triad. Note that while this histidine is static in position in the hCE1 structures determined to date, it has been observed to shift in other cholinesterase structures [29, 42]. If one imposes this shift on the hCE1 histidine, however, it would only come within 4.4 Å of the alkoxy oxygen of the hCE1 acyl-enzyme intermediates. Thus, the distinct architecture of the hCE1 active site appears unfavorable to the process of aging because the alkoxy group moiety is placed relatively distant from the catalytic histidine. Indeed, the >1 week time-frame necessary for the hCE1-nerve agent crystals to grow would presumably provide sufficient time for aging to occur. The fact that a secondary dealkylation reaction is not observed in the structures presented here supports the conclusion that hCE1 may be resistant to this process. However, it remains possible (though unlikely) that the crystallization process may have selected only for non-aged hCE1 complexes. As discussed below, lack of aging could not be confirmed by the standard experimental method of reactivation using an oxime.

DISCUSSION

The hCE1-nerve agent complex structures presented here are of covalent acyl-enzyme intermediates rather than the fully aged products. This result was unexpected because the trapping of the cholinesterases (AcChE, BuChE) in non-aged, acyl-enzyme intermediate complexes required the exposure of pre-grown crystals to nerve agents immediately prior to x-ray data collection [7, 29, 42]. In contrast, the hCE1-nerve agent complex crystals used to produce the structures reported here were generated after the enzyme was exposed to nerve agents in solution, and then crystallized over 1-4 weeks. This timeframe would be expected to allow for sufficient time for aging to occur [7].

Aging is defined experimentally by the inability of the OP-treated enzyme to be reactivated by a strong nucleophile like an oxime, which are able to displace a covalent acyl-enzyme intermediate but not a covalent aged complex [1, 11]. We examined the ability of the oximes diacetylmonoxime and monoisonitrosoacetone, which are effective with both cholinesterases and rat serum carboxylesterase [21], to reactivate tabun- and soman-inhibited hCE1. However, no enzyme reactivation was detected (data not shown), perhaps because hCE1, like other intracellular liver CEs from guinea pig [47] and rat (D. Maxwell, unpublished data), is not reactivated by these oximes. Indeed, analysis of the active site of *Mu*AcChE bound to the oximes HI-6, Ortho-7, and obidoxime, reveals that a series of aromatic residues are required for stabilization and proper alignment of the oxime nucleophiles for reactivation [48]; these aromatic residues are not present in the active site of hCE1. In the absence of oxime reactivation of the inhibited hCE1, it is impossible to

experimentally confirm a lack of aging after inhibition by soman or tabun. Additional work to identify an oxime that can reactivate hCE1 might allow this protein to function as a "pseudo-catalytic" scavenger [49].

Several structural features of the hCE1-nerve agent complexes presented here indicate that hCE1 may be resistant to aging. Detailed structural studies of nerve agent-cholinesterase complexes reported previously have highlighted the importance of the catalytic histidine in aging, either via hydrolysis or via carbocation stabilization [6, 29, 42]. As shown above, the active site of hCE1 is distinct from the cholinesterases in terms of interactions with nerve agents (Figures 3.7-3.8). In hCE1, the alkoxyl substituent is positioned away from the catalytic histidine in the active site gorge, while the chlolinesterases place this group in close proximity to their catalytic histidines. Controversy still exists over the mechanism of aging for some nerve agents. For example, mass-spectrometry and x-ray crystallographic studies of the AcChE-tabun complex indicate that aging occurs via the loss of the dimethylamine group through P-N bond scission rather than by loss of the alkoxy group [42, 50]. The mechanism proposed involves the protonation of the dimethylamine nitrogen by the catalytic hisitdine [42]. Both aged and non-aged *Mu*AcChE-tabun complex structures appear to contradict this mechanism, however, because the dimethylamine group is pointing away from the catalytic histidine [42] (Figure 3.7A). In addition, we found no evidence for aging of the hCE1-tabun structure (Figure 3.5), even though the dimethylamine group is proximal to His468. Regardless however, new methods must be developed to experimentally examine the ability of hCE1 to age in the presence of organophosphorus nerve agents.

Inspired by the identification of cholinesterase mutants with favorable features in relation to organophosphate binding and hydrolysis [15, 17, 51, 52], the structures of hCE1 in complexes with nerve agents presented here may be useful in developing this enzyme into a protective nerve agent hydrolase. Use of enzymes as a prophylaxis for OP exposure carries significant benefits beyond current chemical treatments. A prophylactic treatment for nerve agent exposure would be designed to protect at-risk military personnel and civilian first responders. As such, the use of a human enzyme for such a purpose would be expected to avoid potentially harmful immune responses that may arise from protein-based therapies derived from non-human sources. Extensive work has been conducted toward this goal with a recombinant form of BuChE purified from the milk of transgenic goats [49, 53]. While this enzyme has been shown to confer a level of protection up to 5.5 × LD_{50} for VX and soman in guinea pigs [49], it acts purely as a stoichiometric bioscavenger and binds to, but does not catalytically inactivate, nerve agents. Thus, protection levels are limited stoichiometrically to the dose given. The most promising catalytic bioscavenger identified to date is the organophosphorous hydrolase (also known as phosphotriseterase) from the bacterium *Pseudomonas diminuta*; this enzyme exhibits moderate activity (k_{cat}=56 s^{-1} for sarin, 77 s^{-1} for tabun) for breaking down nerve agents [54, 55]. However, the bacterial origin of this organophosphorous hydrolase may limit its effectiveness in humans [49]. The human serum enzyme paraoxonase 1 (PON1) has been examined for development, but no structural data and low catalytic efficiency for nerve agents have plagued this process [49]. Therefore, a human enzyme with catalytic ability to hydrolyze nerve agents efficiently would be a welcome addition to this growing field of protein-based potential therapeutics.

Several features of hCE1 make it a promising candidate for development into a catalytic prophylactic for nerve agent exposure. As a human enzyme, it is unlikely to elicit an immune response; indeed, cross-species injections of CEs have established that these enzymes can circulate stably within mammalian serum for weeks [20, 21]. hCE1 is also a naturally broad-spectrum enzyme that acts on a wide range of structurally-distinct substrates; thus, protection against several nerve agents may be possible. Finally, as described above, the active site of hCE1 may be uniquely resistant to the secondary dealkylation reaction that can result in aging after inhibition by an OP nerve agent. Taken together, these observations suggest that if hCE1 were converted into an efficient OP hydrolase, it may be an effective protein-based therapeutic that detoxifies organophosphate chemical weapons.

REFERENCES

1. Wiener, S.W. and R.S. Hoffman, *Nerve agents: a comprehensive review.* J Intensive Care Med, 2004. **19**(1): p. 22-37.

2. Newmark, J., *The birth of nerve agent warfare: lessons from Syed Abbas Foroutan.* Neurology, 2004. **62**(9): p. 1590-6.

3. Lee, E.C., *Clinical manifestations of sarin nerve gas exposure.* Jama, 2003. **290**(5): p. 659-62.

4. Casida, J.E. and G.B. Quistad, *Organophosphate toxicology: safety aspects of nonacetylcholinesterase secondary targets.* Chem Res Toxicol, 2004. **17**(8): p. 983-98.

5. Kraut, J., *Serine proteases: structure and mechanism of catalysis.* Annu Rev Biochem, 1977. **46**: p. 331-58.

6. Millard, C.B., et al., *Crystal structures of aged phosphonylated acetylcholinesterase: nerve agent reaction products at the atomic level.* Biochemistry, 1999. **38**(22): p. 7032-9.

7. Nachon, F., et al., *Role of water in aging of human butyrylcholinesterase inhibited by echothiophate: the crystal structure suggests two alternative mechanisms of aging.* Biochemistry, 2005. **44**(4): p. 1154-62.

8. Saxena, A., et al., *The pH dependence of dealkylation in soman-inhibited cholinesterases and their mutants: further evidence for a push-pull mechanism.* Biochemistry, 1998. **37**(43): p. 15086-96.

9. Shafferman, A., et al., *Aging of phosphylated human acetylcholinesterase: catalytic processes mediated by aromatic and polar residues of the active centre.* Biochem J, 1996. **318 (Pt 3)**: p. 833-40.

10. Newmark, J., *Therapy for nerve agent poisoning.* Arch Neurol, 2004. **61**(5): p. 649-52.

11. Bajgar, J., *Organophosphates/nerve agent poisoning: mechanism of action, diagnosis, prophylaxis, and treatment.* Adv Clin Chem, 2004. **38**: p. 151-216.

12. Eyer, P., *The role of oximes in the management of organophosphorus pesticide poisoning.* Toxicol Rev, 2003. **22**(3): p. 165-90.

13. Marrs, T.C., *The role of diazepam in the treatment of nerve agent poisoning in a civilian population.* Toxicol Rev, 2004. **23**(3): p. 145-57.

14. Brown, M.A. and K.A. Brix, *Review of health consequences from high-, intermediate- and low-level exposure to organophosphorus nerve agents.* J Appl Toxicol, 1998. **18**(6): p. 393-408.

15. Millard, C.B., O. Lockridge, and C.A. Broomfield, *Design and expression of organophosphorus acid anhydride hydrolase activity in human butyrylcholinesterase.* Biochemistry, 1995. **34**(49): p. 15925-33.

16. Raveh, L., et al., *Human butyrylcholinesterase as a general prophylactic antidote for nerve agent toxicity. In vitro and in vivo quantitative characterization.* Biochem Pharmacol, 1993. **45**(12): p. 2465-74.

17. Millard, C.B., O. Lockridge, and C.A. Broomfield, *Organophosphorus acid anhydride hydrolase activity in human butyrylcholinesterase: synergy results in a somanase.* Biochemistry, 1998. **37**(1): p. 237-47.

18. Wang, Y., et al., *Resistance to organophosphorus agent toxicity in transgenic mice expressing the G117H mutant of human butyrylcholinesterase.* Toxicol Appl Pharmacol, 2004. **196**(3): p. 356-66.

19. Raveh, L., et al., *The stoichiometry of protection against soman and VX toxicity in monkeys pretreated with human butyrylcholinesterase.* Toxicol Appl Pharmacol, 1997. **145**(1): p. 43-53.

20. Maxwell, D.M. and K.M. Brecht, *Carboxylesterase: specificity and spontaneous reactivation of an endogenous scavenger for organophosphorus compounds.* J Appl Toxicol, 2001. **21 Suppl 1**: p. S103-7.

21. Maxwell, D.M., et al., *Comparison of Cholinesterases and Carboxylesterase as Bioscavengers for Organophosphorus Compounds*, in *Structure and Function of Cholinesterases and Related Proteins*, B.P. Doctor, Editor. 1998, Plenum Press: New York. p. 387-392.

22. Redinbo, M.R. and P.M. Potter, *Mammalian carboxylesterases: from drug targets to protein therapeutics.* Drug Discov Today, 2005. **10**(5): p. 313-25.

23. Satoh, T. and M. Hosokawa, *The mammalian carboxylesterases: from molecules to functions.* Annu Rev Pharmacol Toxicol, 1998. **38**: p. 257-88.

24. Satoh, T. and M. Hosokawa, *Structure, function and regulation of carboxylesterases.* Chem Biol Interact, 2006. **162**(3): p. 195-211.

25. Bencharit, S., et al., *Multisite promiscuity in the processing of endogenous substrates by human carboxylesterase 1.* J Mol Biol, 2006. **363**(1): p. 201-14.

26. Bencharit, S., et al., *Crystal structure of human carboxylesterase 1 complexed with the Alzheimer's drug tacrine: from binding promiscuity to selective inhibition.* Chem Biol, 2003. **10**(4): p. 341-9.

27. Bencharit, S., et al., *Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme.* Nat Struct Biol, 2003. **10**(5): p. 349-56.

28. Fleming, C.D., et al., *Structural insights into drug processing by human carboxylesterase 1: tamoxifen, mevastatin, and inhibition by benzil.* J Mol Biol, 2005. **352**(1): p. 165-77.

29. Millard, C.B., et al., *Reaction products of acetylcholinesterase and VX reveal a mobile histidine in the catalytic triad.* J. Am. Chem. Soc., 1999. **121**: p. 9883-9884.

30. Danks, M.K., et al., *Comparison of activation of CPT-11 by rabbit and human carboxylesterases for use in enzyme/prodrug therapy.* Clin Cancer Res, 1999. **5**(4): p. 917-24.

31. Morton, C.L. and P.M. Potter, *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): p. 193-202.

32. Otwinowski, Z. and W. Minor, *Processing of X-ray Diffraction Data Collected in Oscillation Mode*, in *Macromolecular Crystallography, part A*, C. Carter Jr. and R. Sweet, Editors. 1997, Academic Press: New York. p. 307-326.

33. *The CCP4 suite: programs for protein crystallography.* Acta Crystallogr D Biol Crystallogr, 1994. **50**(Pt 5): p. 760-3.

34. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination.* Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-21.

35. Jones, T.A., et al., *Improved methods for building protein models in electron density maps and the location of errors in these models.* Acta Crystallogr A, 1991. **47 (Pt 2)**: p. 110-9.

36. Read, R.J., *Improved Fourier Coefficients for maps using phases from partial structures with errors.* Acta Crysallogr. A, 1986. **42**: p. 140-149.

37. Laskowski, R.A., et al., *PROCHECK: a program to check the stereochemical quality of protein structures.* J. Appl. Cryst., 1993. **26**: p. 283-291.

38. DeLano, W.L., *The PyMOL Molecular Graphics System*. 2002, DeLano Scientific: Palo Alto, CA.

39. Benschop, H.P., et al., *Isolation, anticholinesterase properties, and acute toxicity in mice of the four stereoisomers of the nerve agent soman.* Toxicol Appl Pharmacol, 1984. **72**(1): p. 61-74.

40. Kim, D.H., *Design of protease inhibitors on the basis of substrate stereospecificity.* Biopolymers, 1999. **51**(1): p. 3-8.

41. Komiyama, M. and M.L. Bender, *Do cleavages of amides by serine proteases occur through a stepwise pathway involving tetrahedral intermediates?* Proc Natl Acad Sci U S A, 1979. **76**(2): p. 557-60.

42. Ekstrom, F., et al., *Structural changes of phenylalanine 338 and histidine 447 revealed by the crystal structures of tabun-inhibited murine acetylcholinesterase.* Biochemistry, 2006. **45**(1): p. 74-81.

43. Degenhardt, C.E.A.M., et al., *Enantiospecific Complexation Gas Chromotography of Nerve Agents. Isolation and Properties of the Enatiomers of Ethyl N,N-Dimethylphosphoramidocyanidate (tabun).* J. Am. Chem. Soc., 1986. **108**: p. 8290-8291.

44. Ordentlich, A., et al., *The role of AChE active site gorge in determining stereoselectivity of charged and noncharged VX enantiomers.* Chem Biol Interact, 2005. **157-158**: p. 191-8.

45. Langenberg, J.P., et al., *Development of a physiologically based model for the toxicokinetics of C(+/-)P(+/-)-soman in the atropinized guinea pig.* Arch Toxicol, 1997. **71**(5): p. 320-31.

46. Binkowski, T.A., S. Naghibzadeh, and J. Liang, *CASTp: Computed Atlas of Surface Topography of proteins.* Nucleic Acids Res, 2003. **31**(13): p. 3352-5.

47. Sterri, S.H. and F. Fonnum, *Carboxylesterases in guinea-pig plasma and liver. Tissue specific reactivation by diacetylmonoxime after soman inhibition in vitro.* Biochem Pharmacol, 1987. **36**(22): p. 3937-42.

48. Ekstrom, F., et al., *Crystal structures of acetylcholinesterase in complex with HI-6, Ortho-7 and obidoxime: structural basis for differences in the ability to reactivate tabun conjugates.* Biochem Pharmacol, 2006. **72**(5): p. 597-607.

49. Lenz, D.E., et al., *Stoichiometric and catalytic scavengers as protection against nerve agent toxicity: A mini review.* Toxicology, 2006.

50. Elhanany, E., et al., *Resolving pathways of interaction of covalent inhibitors with the active site of acetylcholinesterases: MALDI-TOF/MS analysis of various nerve agent phosphyl adducts.* Chem Res Toxicol, 2001. **14**(7): p. 912-8.

51. Gopal, S., et al., *Mutagenesis of organophosphorus hydrolase to enhance hydrolysis of the nerve agent VX.* Biochem Biophys Res Commun, 2000. **279**(2): p. 516-9.

52. Lockridge, O., et al., *A single amino acid substitution, Gly117His, confers phosphotriesterase (organophosphorus acid anhydride hydrolase) activity on human butyrylcholinesterase.* Biochemistry, 1997. **36**(4): p. 786-95.

53. Cerasoli, D.M., et al., *In vitro and in vivo characterization of recombinant human butyrylcholinesterase (Protexia) as a potential nerve agent bioscavenger.* Chem Biol Interact, 2005. **157-158**: p. 363-5.

54. Grimsley, J.K., et al., *Structural and mutational studies of organophosphorus hydrolase reveal a cryptic and functional allosteric-binding site.* Arch Biochem Biophys, 2005. **442**(2): p. 169-79.

55. Hill, C., et al., *Enhanced Degredation of Chemical Warfare Agents through Molecular Engineering of the Phosphotriesterase Active Site.* J. Am. Chem. Soc., 2003. **125**: p. 8990-8991.

FIGURE LEGENDS

Figure 3.1: The overall monomeric structure of hCE1 bound to soman. The catalytic domain is shown in blue, the regulatory domain in green, and the αβ domain in purple. Active site residues are highlighted in yellow, with the catalytic triad and soman adduct in red. The Z-site is occupied by sucrose, modeled as ball-and-stick in orange. The carbohydrates attached to residue Asn79 are shown ball-and-stick in magenta, of which only the initial *N*-Acetylglucosamine (NAG) and terminal sialic acid (SIA) are observed.

Figure 3.2: hCE1-Soman Complex. (A) Chemical scheme of soman reacting with hCE1. Soman is attacked at the chiral phosphate by the Oγ of Ser221, resulting in a covalent acyl-enzyme intermediate and a free fluoride ion. (B) Cut-away view of the active site of hCE1 bound covalently to soman. Surrounding residues are shown in purple, the catalytic triad and soman adduct in green, and the oxyanion hole in white. Hydrogen bonds between the phosphoryl oxygen and the oxyanion hole are shown in red.

Figure 3.3: Stereochemistry of the hCE1-soman adduct. The catalytic triad of amino acids (Ser221, His468 and Glu 354) and the adduct are shown in green, with the oxyanion hole in white. (A) Stereo view of a 2.7 Å resolution F_o-F_c simulated annealing omit map (purple, contoured to 3σ) calculated for the P_R-soman adduct. (B) Stereo view of initial difference density maps for the incorrect P_S-soman adduct (cyan for positive, shown at 3σ; red for negative, shown at -3σ).

Figure 3.4: hCE1-Tabun Complex. (A) Chemical scheme of tabun reacting with hCE1. Tabun is attacked at the chiral phosphate by the Oγ of Ser221, resulting in a covalent acyl-enzyme intermediate and free cyanide ion. (B) Cut away view of the hCE1 active site covalently bound to tabun. Surrounding residues are shown in blue, the catalytic triad and adduct in yellow, and the oxyanion hole in white. Hydrogen bonds between the adduct and the oxyanion hole (dashed lines) are shown in red.

Figure 3.5: Stereochemistry of the hCE1-tabun adduct. The catalytic triad and adduct are shown in yellow, with the oxyanion hole in white. (A) Stereo view of a 2.7 Å resolution F_o-F_c simulated annealing omit map (purple, contoured to 3σ) calculated for the P_S-tabun adduct. (B) Stereo view of initial difference density maps for the incorrect P_R-tabun adduct (cyan for positive, shown at 2σ; red for negative, shown at -2σ).

Figure 3.6: Final 2.7 Å resolution difference density (F_o-F_c) at 2.0, 2.5, and 3.0 σ shows that the P_S-Tabun adduct satisfies all pertinent active site density.

Figure 3.7: Comparison between hCE1-nerve agent complexes and non-aged nerve agent-cholinesterase complexes (the oxyanion hole in each structure is rendered in white). (A) Superposition of hCE1-tabun complex (yellow) with the non-aged complex of murine AcChE [42] structure with tabun (purple). (B) The hCE1-Soman complex (green) is on the left, with the non-aged *Tc*AcChE-VX complex [29] on the right (blue).

Figure 3.8: The acyl-loop regions of the aged AcChE-soman [6] complex (green) and the aged BuChE-echothiophate [7] complex (salmon) superimposed on that of the non-aged hCE1-soman complex (purple). Chemical representation of pocket locations highlights differences between structures.

Figure 3.1

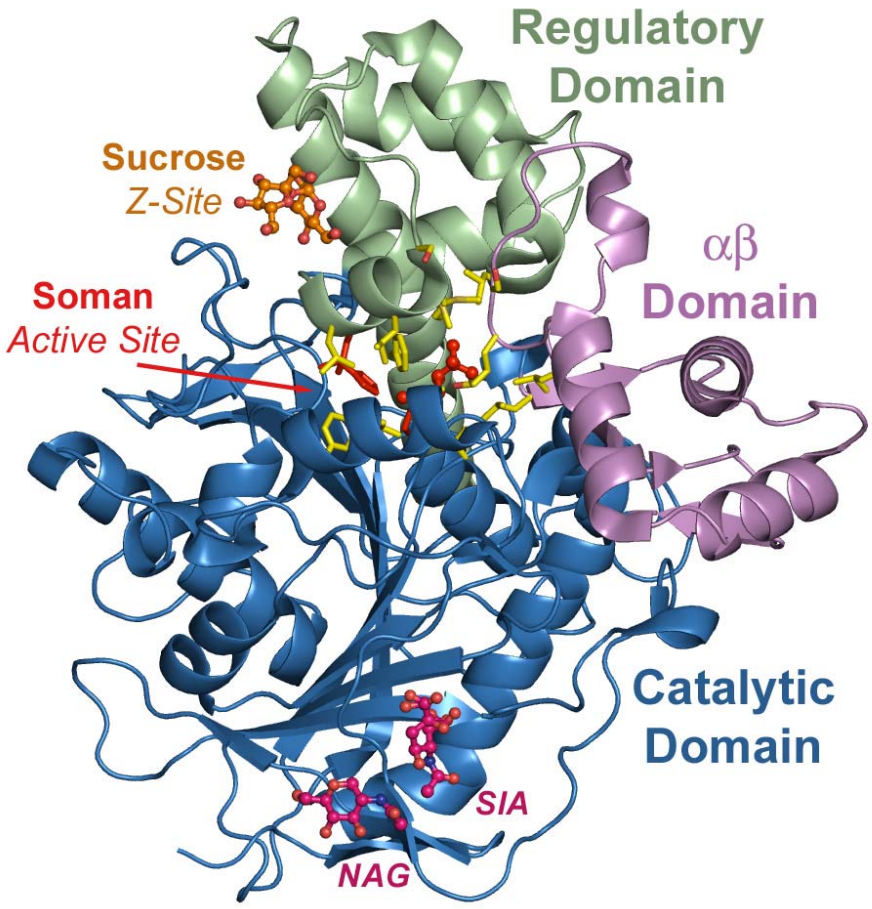

Regulatory
Domain
Sucrose
Z-Site
αβ
Domain
Soman
Active Site
Catalytic
Domain
SIA
NAG

Figure 3.2

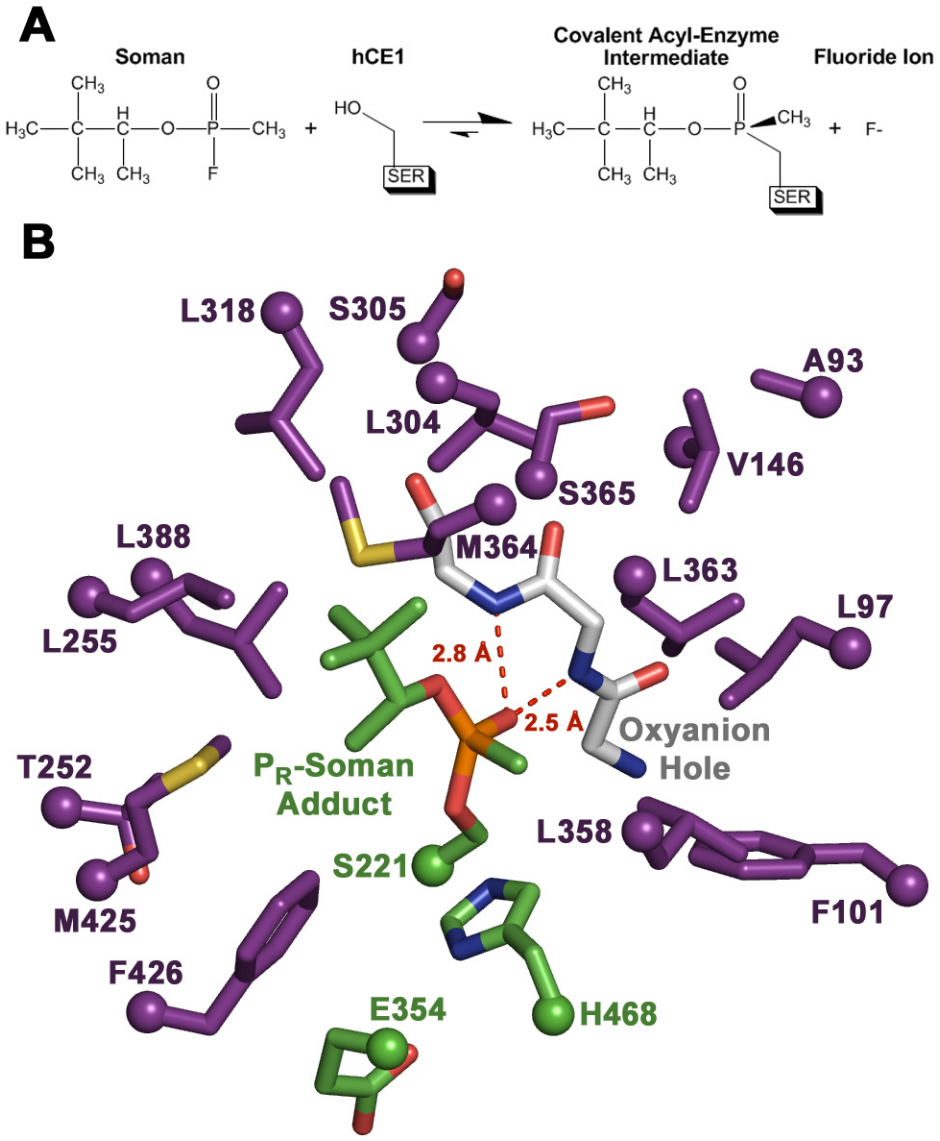

A
Soman
hCE1
Covalent Acyl-Enzyme Intermediate
Fluoride Ion
SER
SER
B
L318
S305
A93
L304
V146
S365
L388
M364
L363
L97
L255
2.8 Å
2.5 Å
Oxyanion Hole
T252
PR-Soman Adduct
L358
S221
F101
M425
F426
E354
H468

Figure 3.3

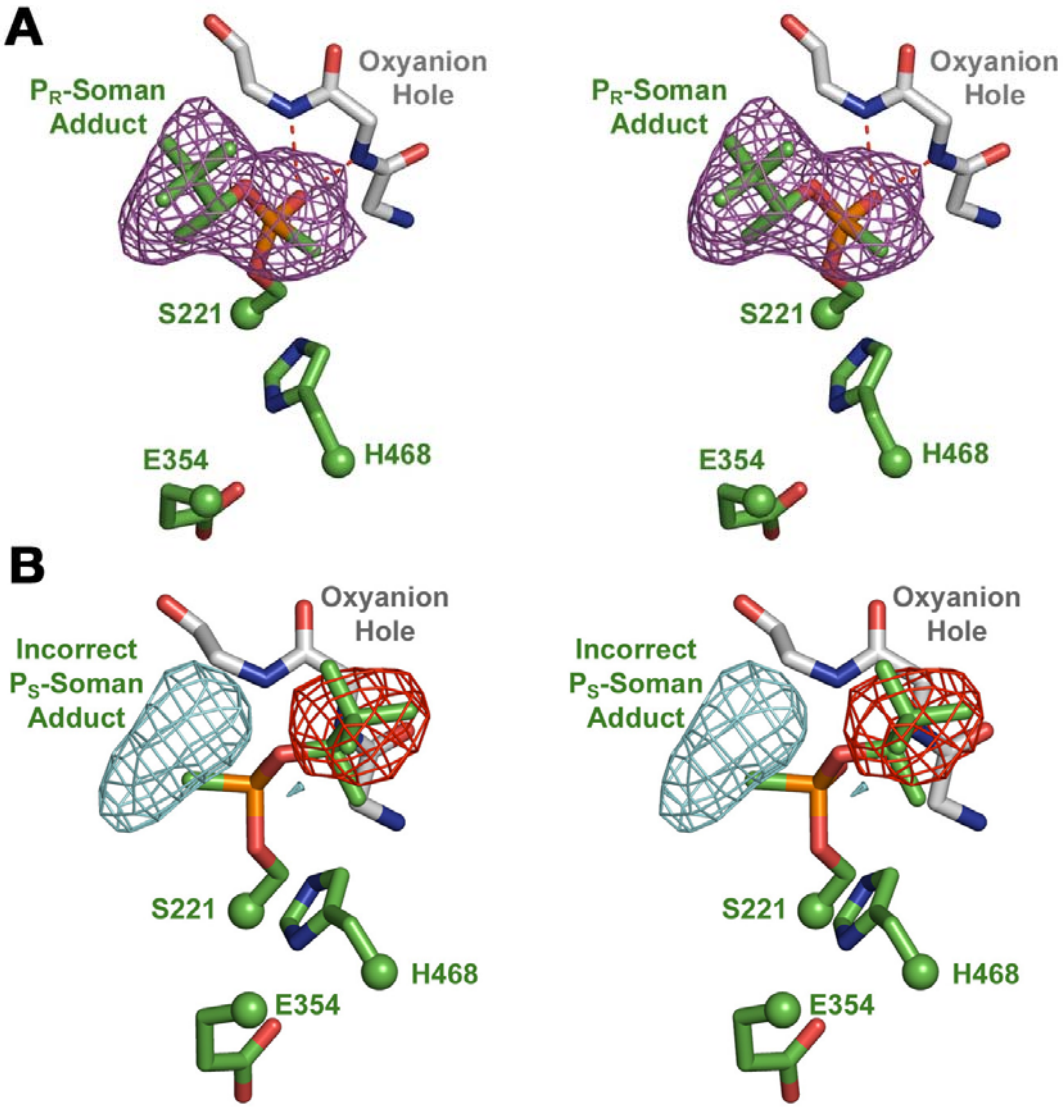

A
Oxyanion Hole
P_R-Soman Adduct
S221
H468
E354
Oxyanion Hole
P_R-Soman Adduct
S221
H468
E354
B
Oxyanion Hole
Incorrect P_S-Soman Adduct
S221
H468
E354
Oxyanion Hole
Incorrect P_S-Soman Adduct
S221
H468
E354

Figure 3.4

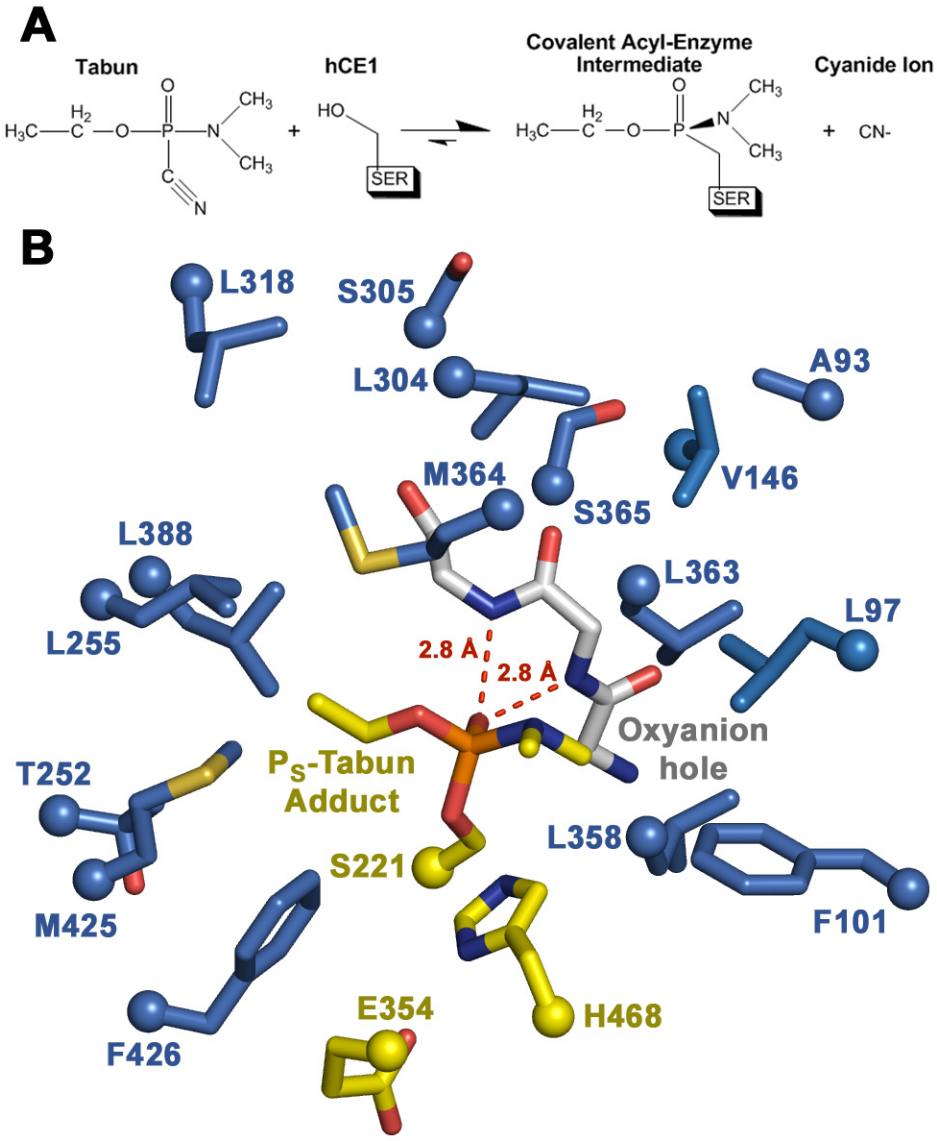

A
Tabun
hCE1
Covalent Acyl-Enzyme Intermediate
Cyanide Ion
SER
SER
B
L318
S305
L304
A93
V146
M364
S365
L388
L363
L97
L255
2.8 Å
2.8 Å
Oxyanion hole
T252
PS-Tabun Adduct
L358
S221
F101
M425
E354
H468
F426

Figure 3.5

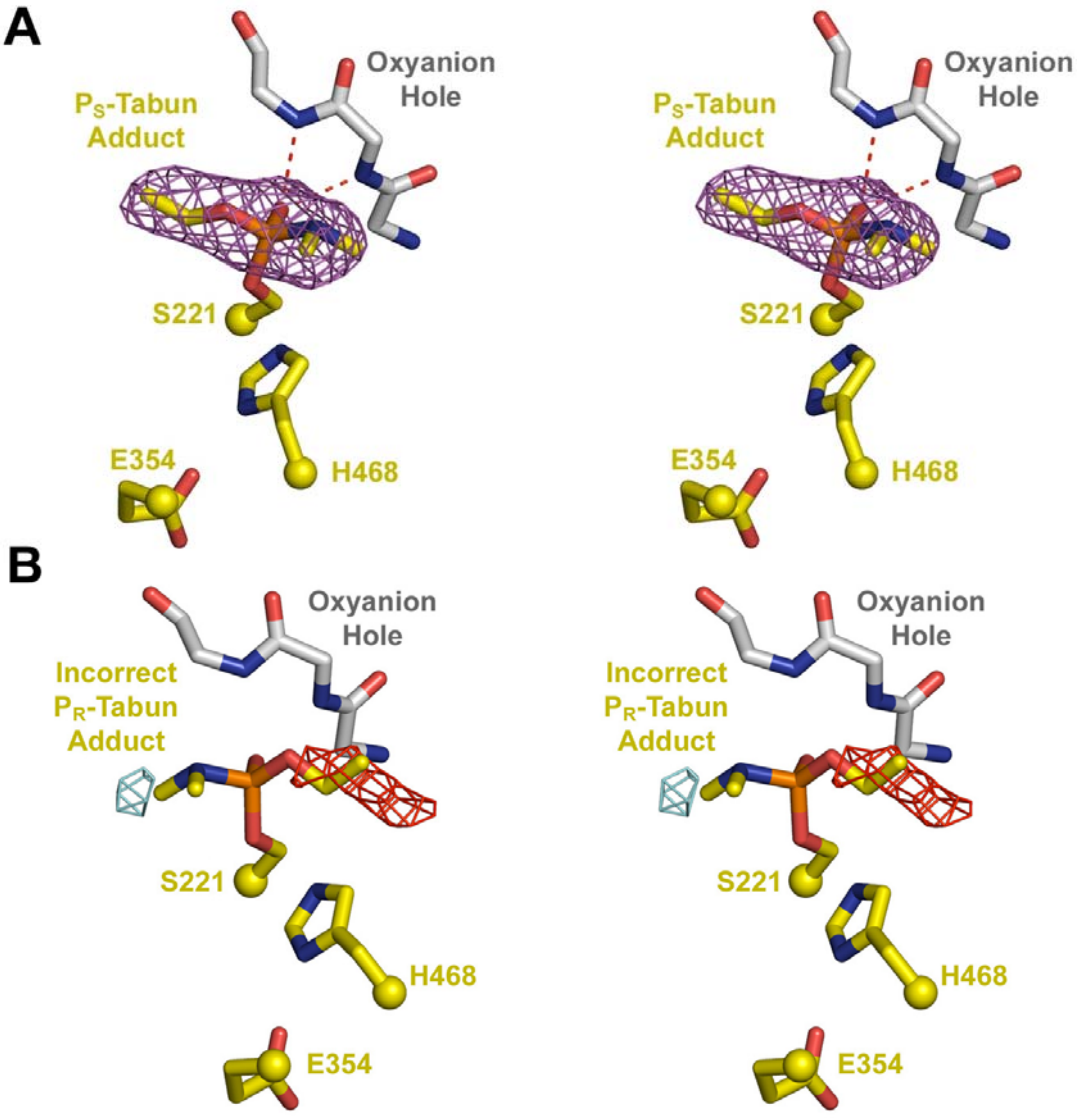

A
Oxyanion Hole
P_S-Tabun Adduct
S221
E354
H468
Oxyanion Hole
P_S-Tabun Adduct
S221
E354
H468
B
Oxyanion Hole
Incorrect P_R-Tabun Adduct
S221
H468
E354
Oxyanion Hole
Incorrect P_R-Tabun Adduct
S221
H468
E354

Figure 3.6

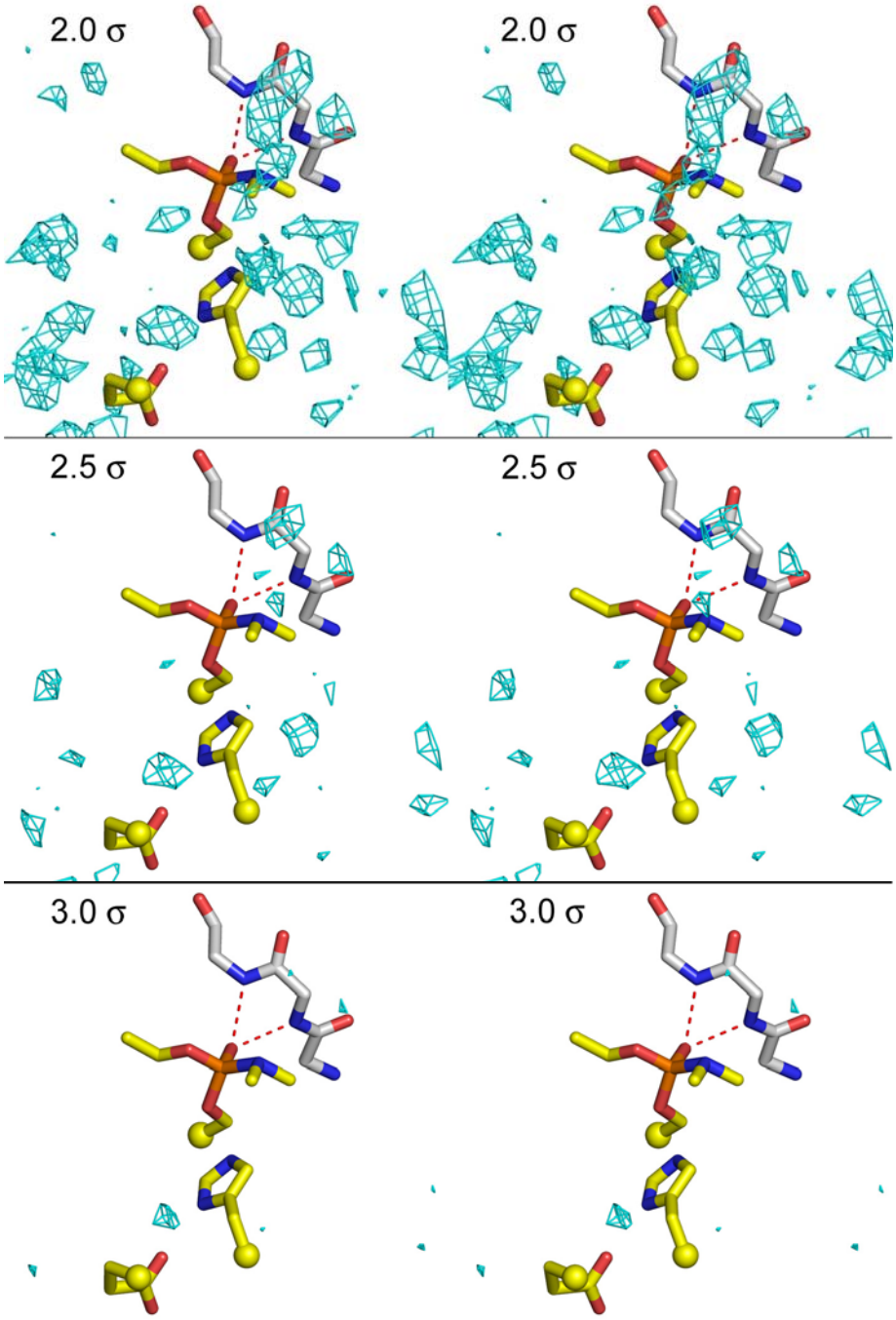


Figure 3.7

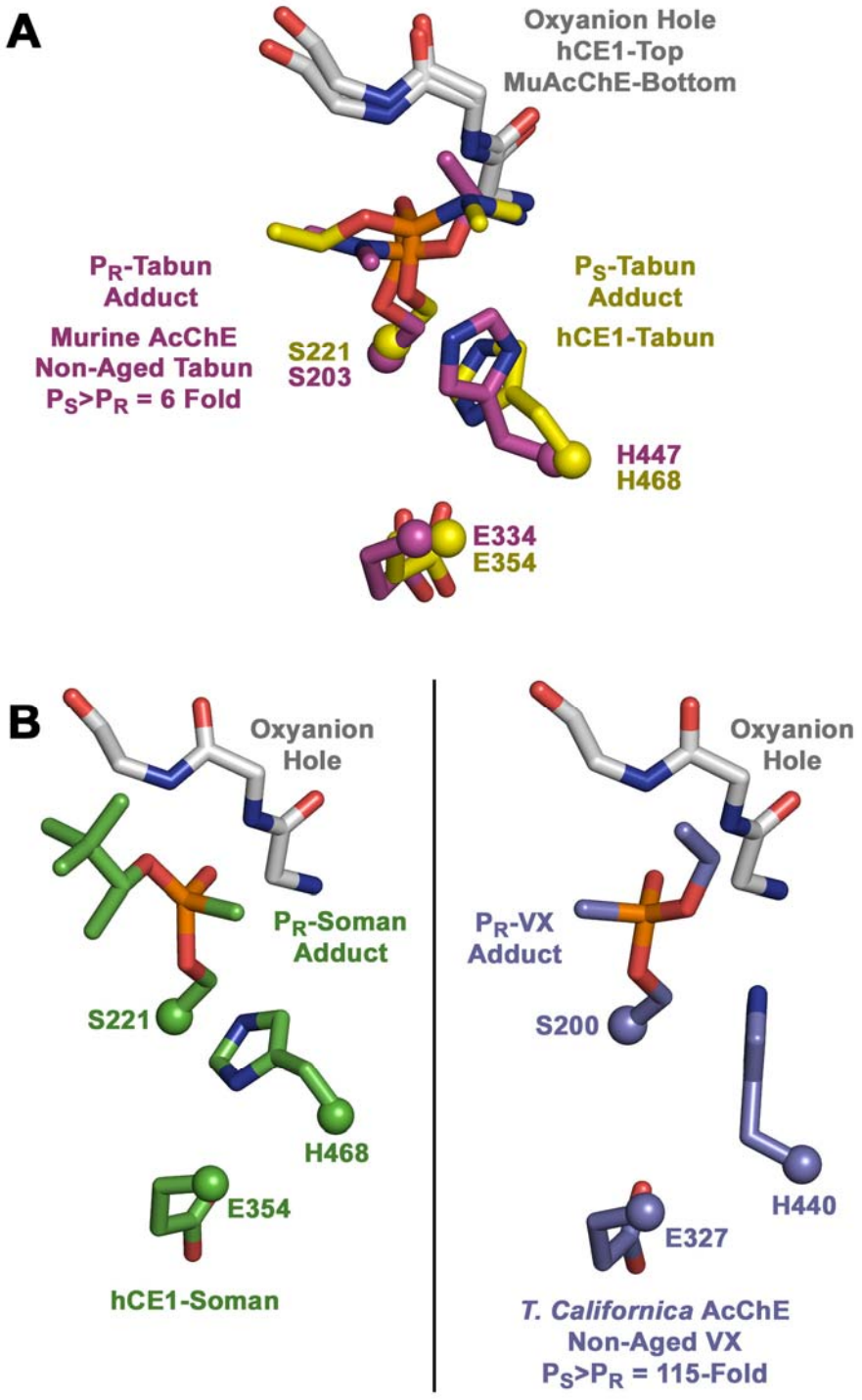

A
Oxyanion Hole
hCE1-Top
MuAcChE-Bottom
PR-Tabun
Adduct
Murine AcChE
Non-Aged Tabun
PS>PR = 6 Fold
S221
S203
PS-Tabun
Adduct
hCE1-Tabun
H447
H468
E334
E354
B
Oxyanion
Hole
PR-Soman
Adduct
S221
H468
E354
hCE1-Soman
Oxyanion
Hole
PR-VX
Adduct
S200
H440
E327
T. Californica AcChE
Non-Aged VX
PS>PR = 115-Fold

Figure 3.8

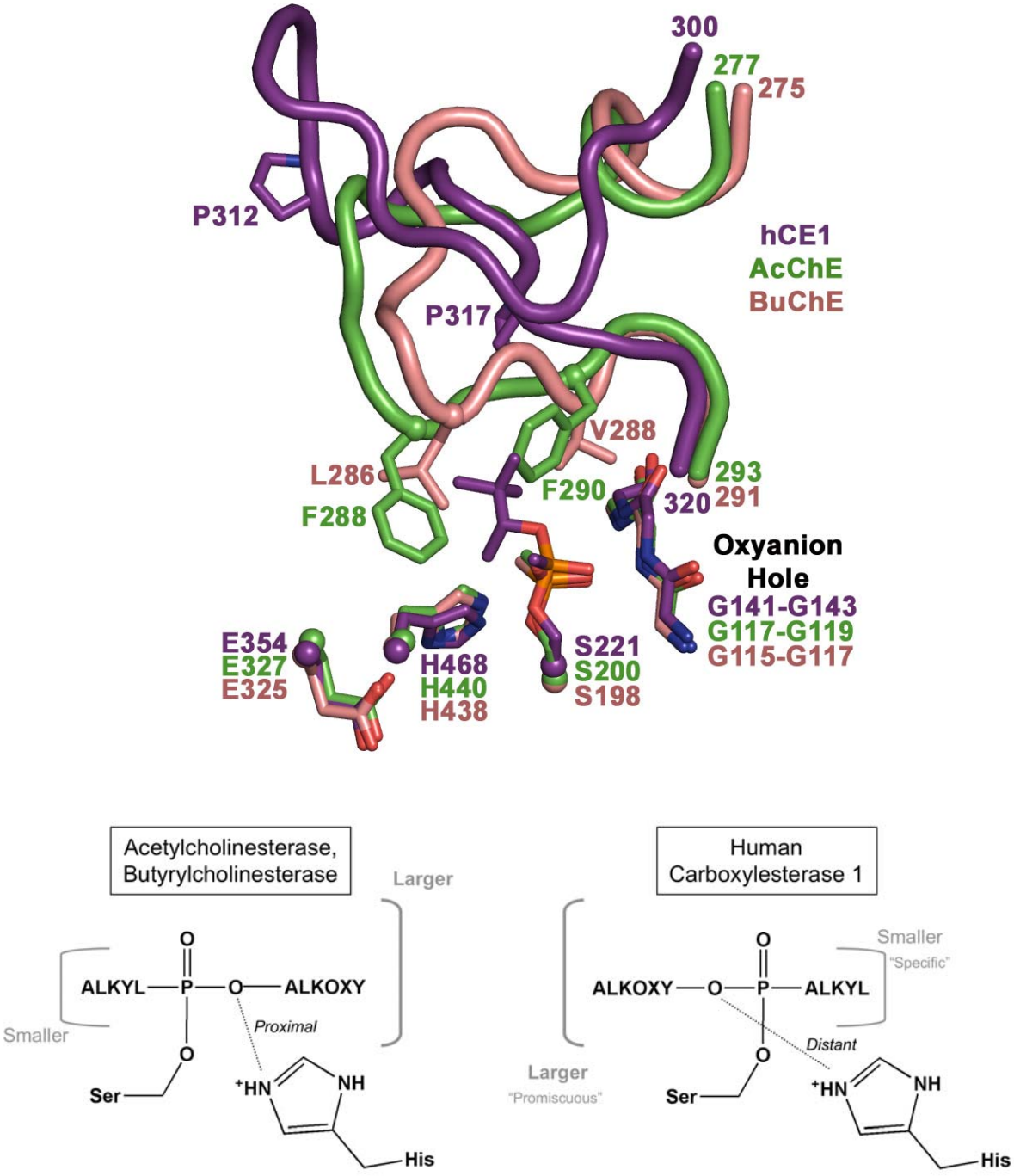

300
277
275
P312
hCE1
AcChE
BuChE
P317
V288
L286
F290
293
320
291
F288
Oxyanion
Hole
G141-G143
G117-G119
G115-G117
E354
E327
E325
H468
H440
H438
S221
S200
S198
Acetylcholinesterase,
Butyrylcholinesterase
Larger
Smaller
ALKYL
P
O
ALKOXY
Proximal
Ser
+HN
NH
His
Human
Carboxylesterase 1
Smaller
"Specific"
ALKOXY
ALKYL
Distant
Larger
"Promiscuous"

Table 3.1: Crystallographic Statistics

	hCE1-Soman	***hCE1-Tabun***
Resolution (Å)[1]	50-2.70 (2.87-2.70)	50-2.70 (2.80-2.70)
Space Group	$P2_1$	$P2_12_12_1$
Asymmetric Unit	Two Trimers	One Trimer
Cell Constants (Å, °)	a=55.5	a=55.6
	b=181.2	b=181.1
	c=203.1	c=202.9
	$\alpha=\gamma=90$	$\alpha=\beta=\gamma=90$
	$\beta=89.9$	
Total Reflections	108,029	53,127
Unique Reflections	29,197	8,301
Mean Redundancy[1]	3.7 (3.3)	6.4 (6.3)
R_{sym} (%)[1,2]	10.2 (21.6)	11.7 (29.2)
Wilson B Factor (Å^2)	34.6	28.8
Completeness (%)[1]	99.9 (99.9)	91.8 (93.8)
Mean I/σ[1]	19.7 (7.8)	21.8 (7.7)
R_{cryst} (%)[1,3]	17.0	16.7
R_{free} (%)[1,4]	22.5	23.2
RCSB Access Codes	2HRQ	2HRR

[1] Number in parentheses is for the highest shell.

[2] $R_{sym} = \Sigma|I - \langle I \rangle| / \Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity of multiple symmetry-related observations of that reflection.

[3] $R_{cryst} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

[4] $R_{free} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$ for 7% of the data not used at any stage of structural refinement.

Chapter 4:

Introduction to the Androgen Receptor Modulator MAGE-11

Introduction

The androgen receptor (AR) is a member of the steroid nuclear receptor family and is responsible for regulating gene expression in response to male steroid ligands such as testosterone. AR plays an important role in prostate cancer tumorigenesis and progression, and is therefore a target for treatment of metastatic forms of the disease. AR activity is regulated by associated co-activators and co-repressors, which facilitate or disrupt the expression of AR controlled proteins by recruiting transcriptional machinery. Co-activators and co-repressors bind to the activation function 2 (AF-2) region of the ligand binding domain (LBD), which is properly assembled upon androgen binding. AR, however, exhibits a self association between the AF-2 region and an FxxLF peptide of its N-terminal domain (NTD) that must be disrupted for co-activator or co-repressor binding. The melanoma antigen gene 11 (MAGE-11) protein was recently found to abrogate this interaction by competing for the FxxLF region of AR's NTD, facilitating co-activator/co-repressor binding to the AF-2 region.

General Characteristics of Nuclear Receptors

Nuclear receptors are proteins that generally respond to chemical stimuli to regulate gene expression involved in the growth, differentiation, morphology, and reproduction of many different cell types and tissues [1]. These ligand-inducible transcription factors are separated into several subfamilies that have been grouped by sequence identity and the chemicals to which they respond [2, 3]. Class 1 is a large subfamily that includes the first receptors with known cognate ligands, including the vitamin D receptor (VDR), the peroxisome proliferator activated receptors (PPAR), and thyroid receptors (TR). This subfamily also includes several NRs that are known as adopted orphan receptors. In adopted orphans, the endogenous ligands were not initially known but have since been assigned. These include xenobiotic detecting NRs such as the pregnane X receptor (PXR), constitutive androstane receptor (CAR), and the liver X receptor (LXR). Class II receptors include the retinoid X receptor (RXR), which is a common dimeric partner of other NRs in class I, and the hepatocyte nuclear factor (HNF) group of receptors. Class III encompasses all steroid receptors, including the estrogen receptor (ER), glucocorticoid and mineralcorticoid receptors (GR, MR), and finally AR. Further classes include nuclear receptor like proteins NGF-induced clone B (IV), the steroidogenic factor 1 or SF-1 (V), germ cell nuclear factor (VI), and the small heterodimeric partner SHP (O) [1].

Nuclear receptors typically share four common primary domains: the A/B or N-terminal domain, a DNA binding domain (DBD), a modular hinge region, and the LBD (Figure 4.1) [1, 4]. The modular A/B or NTD is a highly variable domain that can result in

multiple isoforms of the same nuclear receptor due to different transcriptional start sites, alternative splicing, and different promoters [5]. A prime example of this is the retinoic acid receptor (RAR), of which three genes result in eight separate receptors with different NTDs [6]. The NTD contains an activation function region (AF-1) which determines in part the ligand-independent activity of nuclear receptors. The signal for this basal activity is likely due to the interaction of nuclear receptors with cellular kinases that phosphorylate this region, enhancing the transcriptional activity of the NTD [7]. Whereas no crystal structure exists for an intact NTD, nuclear magnetic resonance (NMR) and circular dichroism studies have shown that the region lacks well-ordered structure in GR and both isoforms (α,β) of ER [8]. However, this may not be valid when additional domains or binding partners are present. The NTD of the progesterone receptor (PR) and GR were shown by limited proteolysis to be more structured when expressed with an intact DBD than without [9, 10].

Whereas the NTD exhibits extremely high levels of diversity across the nuclear receptor family, the DNA binding domain (DBD) is highly conserved. The small, helical DBD is primarily comprised of two zinc finger motifs that encompass approximately 70 residues and a C-terminal extension (CTE) sequence that adds an additional 25 residues. Several of these residues are critical for function, including two cystine clusters that coordinate the zinc ions and additional smaller subdomains known as the “P” and “D” boxes [1]. The zinc finger containing the “P” box is involved in the identification of hormone response elements (HREs) on the regulatory regions of nuclear DNA, while the “D” box on the other zinc finger is involved in dimerization. Additional DNA binding recognition and stabilization is facilitated by the CTE, which has been shown to recruit the coregulatory

proteins HMGB-1/-2 to class I and II NRs [11, 12]. HREs are configured as one or two half sites with the sequence AG(G/T)TCA for the majority of NRs, with the class III steroid receptors generally responding to an AGAACA sequence [1]. Depending on the receptor, NRs bind these DNA sequences as a monomer, homodimer, or heterodimer, with the aforementioned "D" box responsible for part of the dimer interaction. Steroid receptors nearly always bind DNA as homodimers, whereas those NRs that act as heterodimers (e.g. PXR, VDR, and PPARγ) partner with RXRα at response elements [1, 13]. The dimerization process may confer additional levels of regulation by allowing distal, inverted, or everted response elements to be activated [5]. The hinge region may facilitate this movement by allowing the DBD and LBD to move independent of each other [5, 14]. The hinge has also been shown to possess nuclear localization signals and binding sites for some co-repressors [15].

The ligand binding domain (LBD) is responsible for the majority of nuclear receptor functions, including the mediation of ligand-dependant transcriptional activity, dimerization, and interaction with heat-shock proteins [1, 5]. The LBD is a highly structured domain, with 10-13 α helices and a short antiparallel beta sheet (Figure 4.2) [16]. The LBD also contains the second activation function region (AF-2), a hydrophobic patch on the surface of the domain that is stabilized by agonist binding [8]. The conformational change that occurs to form the AF-2 region facilitates the recruitment of co-activators essential for transcriptional activity, such as SRC-1, a member of the SRC/p160 class of co-activators [17]. Co-repressors such as N-CoR (nuclear receptor co-repressor) and SMRT (silencing mediator for retinoic

acid and thyroid hormone receptors), alternatively, can bind the LBD in absence of ligand, when the AF-2 helix is in a non-productive conformation [15].

The Androgen Receptor

The androgen receptor (AR) is a member of the steroid nuclear receptor subclass that acts as an intracellular transcription factor against endogenous androgens. AR expression is primarily limited to androgen-specific tissues, namely the prostate, skeletal muscle, liver, and central nervous system, with the highest levels found in the prostate, epididymis, and adrenal gland [18]. AR is responsible for mediating male physiological changes throughout a lifetime, from sexual differentiation *in utero* and pubertal changes, to maintaining libido, muscle mass, and spermatogenesis in adults [19]. The principal male androgen testosterone, which is secreted by both the testes and adrenal gland, is the primary activator of AR in males. Testosterone can act by three different mechanisms, depending on whether or not it undergoes modification. Within some androgen-specific tissues, testosterone activity is naturally increased through its biotransformation into dihydrotestosterone (DHT) by 5α-reductase (Figure 4.3) [20]. DHT is a more potent activator of AR then testosterone (2-10x), and is hence responsible for specific effects in the prostate and hair follicles [19]. Unmodified testosterone is responsible for sperm production and muscle growth, while a small percentage is converted into estrogen that affects behavior and bone growth. DHT binding can also alter the cellular localization of AR. AR is normally localized throughout prostate epithelial cells bound to heat-shock proteins that preclude AR from binding DNA [21]. Upon DHT binding, however, AR dissociates from the heat-shock proteins, and is

subsequently localized to the nucleus for active transcription [22]. This localization is mediated by the nuclear localization sequence in the hinge region of AR. AR also homodimerizes, undergoes phosphorylation, and is directed to androgen response elements (AREs) that are located in the promoter or enhancer regions of androgen specific gene targets. Co-regulators are then recruited along with transcriptional machinery, further ensuring the transactivation of AR-regulated gene expression [23, 24].

Several families of co-activators have been shown to interact with AR in the nucleus. These co-activators upregulate hormone mediated gene expression through either direct modification of chromatin or by recruiting nuclear proteins that contain histone acyl- or methyl-transferase activity. Members of the p160 class of co-activators have been found to interact with AR, including SRC-1, transcriptional intermediary factor 2 (TIF2), and amplified in breast cancer-1 (AIB1) [23, 25, 26]. These co-activators bind to the AF-2 region in the LBD of AR through a conserved LxxLL motif, with the leucine residues fitting into a hydrophobic groove. While these interactions have been shown to increase steroid receptor activity, it has not yet been shown if this interaction is required in vivo. Loss of the SRC-1 gene in mice only resulted in minor physiologic changes in androgen specific tissues, and additional knockout experiments done with other p160 co-activators gave similar findings [27]. This led researchers to explore other aspects of AR, most notably its ability to self-activate transcription.

The androgen receptor is different from many other nuclear receptors in that it can activate transcription independent of co-activators. AR exhibits a novel N-terminus/C-

terminus (N/C) self-association that can be important for robust steroid-dependent transactivation. The interaction is mediated by the AF-2 region of the LBD and an FxxLF motif found within the NTD, with an additional WxxLF motif contributing to a lesser extent [28-30]. The FxxLF peptide is alpha helical, similar in size to the LxxLL box found in the p160 class of co-activators [31]. This N/C communication is androgen-dependant (since AF-2 formation is due to ligand binding) and has been shown to be required for the activation of some AR-regulated genes [29]. While the AF-2 region of other steroid receptors (e.g. ERα) is important for co-activator binding, the AF-2 of AR is different. AR prefers the internal FxxLF motif over the LxxLL motif of the SRC/p160 co-activators by approximately 10-fold, and hence competitively inhibits AF-2 dependent recruitment of SRC-1 [28, 32]. Structural determinants likely drive this process, since an FxxLF peptide promotes additional hydrogen bonding between the peptide and the charge clamp residue E897 of the LBD that are not found in the LxxLL/AF-2 interaction (Figure 4.2) [31].

Despite the preference for FxxLF over the LxxLL by the AF-2 of AR, co-activators are still important in AR regulated gene expression. For example, TIF2 has been shown to promote receptor stability by facilitating the N/C interaction [33]. While TIF2 expression is naturally low in prostate epithelial cells, recurrence of prostate tumors after androgen deprivation therapy has been tied to increased levels of co-activators of the SRC/p160 family [34, 35]. Furthermore, transient transfection assays have shown increased co-activator levels promote AR mediated expression of androgen-regulated gene products [29]. Taken together, these findings imply the existence of other co-regulators that modulate the N/C interaction in order to facilitate AF-2 binding of co-activator LxxLL motifs. Using a yeast two-hybrid

screen with the AR FxxLF peptide as bait, researchers were able to identify the cellular protein MAGE-11 as a potential regulator of this interaction [36]. The role of MAGE-11 as a novel co-regulator of AR is examined more fully below.

Melanoma Antigen Gene Protein MAGE-11

The melanoma antigen gene proteins (MAGE) consist of a large superfamily with a broad range of functional diversity and expression. The first MAGE protein discovered, MAGE-A1, was isolated in 1991 from a human melanoma cell line [37]. A proteolytic fragment of MAGE-A1 was presented as an cell surface antigen bound by major histocompatability class 1 (MHC-1), and in turn recognized by cytolytic T lymphocytes in melanomas [37]. Researchers further discovered eleven additional MAGE family members with high levels of sequence identity (64-85%) to MAGE-A1 [38]. The 12 MAGE-A genes are localized to the terminal 3.5-Mb q28 region of the X-chromosome, indicating a potential role in x-linked diseases [38, 39]. Unless indicated otherwise, all numbered MAGE genes discussed from here on are of the A family. Additional sequencing efforts directed towards other areas of the human X-chromosome isolated three more families of genes, MAGE-B, MAGE-C, and MAGE-D [40-42]. The majority of MAGE genes (-A, -B, -C) contain one open reading frame found in the last exon of the gene, which are subsequently classified as class I MAGE family genes. The major feature that comprises class I MAGE genes is the MAGE homology domain, or MHD, which is found in the A, B, and C families but not the MAGE-D family (Figure 4.4) [43].

An extensive study of normal, somatic cell lines showed that MAGE genes were expressed almost exclusively in the testes [44]. However, numerous cancer cell lines showed increased levels of these proteins, including metastatic melanomas, carcinomas of the head and neck, bladder tumors, non small cell lung cancer, paciltaxel resistant ovarian cancer, and synovial sarcomas [44, 45]. For example, RT-PCR experiments have shown that MAGE-3 is expressed in 76% of metastatic melanomas and 82% of testicular germ cell tumors [46]. Most positive tumors in one study found that greater than 50% of the tumor cells were immunopositive towards the MAGE antibody, and that higher levels of staining correlated with the progression of the disease [44]. MAGE family members have been found in hepatocellular carcinomas as well, but with different levels of expression. MAGE genes -1 and -3 were present at the highest levels, in approximately 68% of tumors, while MAGE-8 was in 46%, and MAGE-2, 6, 10, 11, and 12 at only 30% [47]. In terms of the intracellular localization of the MAGE family of proteins, little conclusive evidence is available. For example, MAGE-1 and -3 have been shown to localize primarily to the cytoplasmic fraction of tumor cell lines, with small amounts of MAGE-1 found in the nucleus [48, 49]. Alternatively, MAGE-10 and MAGE-11 were found to be exclusively in the nucleus [50, 51]. These results indicate that even though MAGE proteins a share high level of sequence identity, individual protein functions may vary significantly. Expression within cells of MAGE proteins in normal and tumor cell lines may also be methylation controlled. The promoter sequence for MAGE-1 contains a critical CpG bi-nucleotide site which precludes binding of the transcription factor Ets when methylated [52]. Accordantly, when cell lines are treated with a de-methylating agent, MAGE-1 expression was induced [53]. This may act as a trigger for MAGE presentation upon neoplastic growth.

While extensive research has been done on some MAGE family proteins, the focus of this study, MAGE-11, is relatively uncharacterized. As mentioned before, MAGE-11 was first identified in 1994 along with the other eleven members of the MAGE-A family, and was found to be a primarily nuclear protein expressed in testes and placenta [38, 50]. MAGE-11 contains an extended upstream coding region that adds roughly 112 amino acids longer than other MAGE-A family proteins [54]. More importantly, MAGE-11 was recently identified as an androgen receptor co-modulator [36]. The androgen receptor exhibits an N/C interaction that must be disrupted for proper co-activator recruitment to the AF-2 region of its LBD. Using a yeast two-hybrid screen of a human testis library, MAGE-11 was identified to bind the FxxLF motif in the androgen receptor NTD, working to allow co-activator binding at the AF-2. Mutations at the FxxLF site to FxxAA resulted in loss of the interaction, confirming that MAGE-11 recruitment was FxxLF dependent. Furthermore, the MAGE-11/AR interaction was lessened in the presence of the tight binding synthetic androgen R1881, which increases the N/C self-association of AR. An immunoprecipitation assay confirmed the intracellular interaction of MAGE-11 with AR in the absence of androgen, but it was again lost in presence of DHT. Immunocytochemical staining of several cell lines expressing both MAGE-11 and AR showed that they co-localized in the cytoplasm in the absence of androgen, but upon DHT addition were transported to the nucleus.

The area of AR that MAGE-11 interacts with is extended beyond just the FxxLF motif, and was found to encompass residues 16-36 of AR by mutational analysis. This sequence contains numerous residues important to the MAGE-11/AR interaction that were not required for AR N/C self-association. MAGE-11 appears specific for AR 16-36, since

other FxxLF containing AR co-regulators failed to interact with MAGE-11 by yeast two-hybrid. Finally, MAGE-11 was also shown to facilitate binding of AR to TIF2, a co-activator of the SRC/p160 family. Immunoprecipitation assays showed that MAGE-11 co-expression increased levels of TIF2 pulled down by AR. This finding backs up the claim that MAGE-11 disrupts the N/C self-association of AR, allowing the AF-2 of AR to be bound by the co-activator LxxLL motif. Taken together, these results propose a model for how MAGE-11 affects AR-dependant transactivation of androgen specific genes (Figure 4.5).

The identification of MAGE-11 as a novel co-regulator of AR, in combination with the lack of structural information for any MAGE family protein, make MAGE-11 an ideal candidate for a crystallographic analysis. We hypothesize that MAGE-11 contains a surface binding site with similar characteristics to the AF-2 region of AR, with additional structural determinants that explain the sequence requirements of MAGE-11 binding to AR. Solving the crystal structure of MAGE-11 in complex with the AR 16-36 peptide will be the primary experiment, with additional biophysical characterization to fully understand the interaction. The efforts made towards this goal are fully explained in the following chapter.

REFERENCES

1. Aranda, A. and A. Pascual, *Nuclear hormone receptors and gene expression.* Physiol Rev, 2001. **81**(3): p. 1269-304.

2. *A unified nomenclature system for the nuclear receptor superfamily.* Cell, 1999. **97**(2): p. 161-3.

3. Zhang, Z., et al., *Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome.* Genome Res, 2004. **14**(4): p. 580-90.

4. Mangelsdorf, D.J., et al., *The nuclear receptor superfamily: the second decade.* Cell, 1995. **83**(6): p. 835-9.

5. Giguere, V., *Orphan nuclear receptors: from gene to function.* Endocr Rev, 1999. **20**(5): p. 689-725.

6. Giguere, V., *Retinoic acid receptors and cellular retinoid binding proteins: complex interplay in retinoid signaling.* Endocr Rev, 1994. **15**(1): p. 61-79.

7. Shao, D. and M.A. Lazar, *Modulating nuclear receptor function: may the phos be with you.* J Clin Invest, 1999. **103**(12): p. 1617-8.

8. Warnmark, A., et al., *Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation.* Mol Endocrinol, 2003. **17**(10): p. 1901-9.

9. Bain, D.L., et al., *The N-terminal region of the human progesterone A-receptor. Structural analysis and the influence of the DNA binding domain.* J Biol Chem, 2000. **275**(10): p. 7313-20.

10. Kumar, R., et al., *Interdomain signaling in a two-domain fragment of the human glucocorticoid receptor.* J Biol Chem, 1999. **274**(35): p. 24737-41.

11. Melvin, V.S., et al., *The role of the C-terminal extension (CTE) of the estrogen receptor alpha and beta DNA binding domain in DNA binding and interaction with HMGB.* J Biol Chem, 2004. **279**(15): p. 14763-71.

12. Melvin, V.S., et al., *The C-terminal extension (CTE) of the nuclear hormone receptor DNA binding domain determines interactions and functional response to the HMGB-1/-2 co-regulatory proteins.* J Biol Chem, 2002. **277**(28): p. 25115-24.

13. Orans, J., D.G. Teotico, and M.R. Redinbo, *The nuclear xenobiotic receptor pregnane X receptor: recent insights and new challenges.* Mol Endocrinol, 2005. **19**(12): p. 2891-900.

14. Glass, C.K., *Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers.* Endocr Rev, 1994. **15**(3): p. 391-407.

15. Chen, J.D. and R.M. Evans, *A transcriptional co-repressor that interacts with nuclear hormone receptors.* Nature, 1995. **377**(6548): p. 454-7.

16. Wurtz, J.M., et al., *A canonical structure for the ligand-binding domain of nuclear receptors.* Nat Struct Biol, 1996. **3**(1): p. 87-94.

17. Leo, C. and J.D. Chen, *The SRC family of nuclear receptor coactivators.* Gene, 2000. **245**(1): p. 1-11.

18. Keller, E.T., W.B. Ershler, and C. Chang, *The androgen receptor: a mediator of diverse responses.* Front Biosci, 1996. **1**: p. d59-71.

19. Gao, W., J. Kim, and J.T. Dalton, *Pharmacokinetics and pharmacodynamics of nonsteroidal androgen receptor ligands.* Pharm Res, 2006. **23**(8): p. 1641-58.

20. Jenkins, E.P., et al., *Genetic and pharmacological evidence for more than one human steroid 5 alpha-reductase.* J Clin Invest, 1992. **89**(1): p. 293-300.

21. Marivoet, S., et al., *Interaction of the 90-kDa heat shock protein with native and in vitro translated androgen receptor and receptor fragments.* Mol Cell Endocrinol, 1992. **88**(1-3): p. 165-74.

22. Jenster, G., J. Trapman, and A.O. Brinkmann, *Nuclear import of the human androgen receptor.* Biochem J, 1993. **293 (Pt 3)**: p. 761-8.

23. McKenna, N.J., R.B. Lanz, and B.W. O'Malley, *Nuclear receptor coregulators: cellular and molecular biology.* Endocr Rev, 1999. **20**(3): p. 321-44.

24. McKenna, N.J., et al., *Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions.* J Steroid Biochem Mol Biol, 1999. **69**(1-6): p. 3-12.

25. Onate, S.A., et al., *Sequence and characterization of a coactivator for the steroid hormone receptor superfamily.* Science, 1995. **270**(5240): p. 1354-7.

26. Voegel, J.J., et al., *The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways.* Embo J, 1998. **17**(2): p. 507-19.

27. Xu, J., et al., *Partial hormone resistance in mice with disruption of the steroid receptor coactivator-1 (SRC-1) gene.* Science, 1998. **279**(5358): p. 1922-5.

28. He, B., J.A. Kemppainen, and E.M. Wilson, *FXXLF and WXXLF sequences mediate the NH2-terminal interaction with the ligand binding domain of the androgen receptor.* J Biol Chem, 2000. **275**(30): p. 22986-94.

29. He, B., et al., *Dependence of selective gene activation on the androgen receptor NH2- and COOH-terminal interaction.* J Biol Chem, 2002. **277**(28): p. 25631-9.

30. He, B., et al., *Activation function 2 in the human androgen receptor ligand binding domain mediates interdomain communication with the NH(2)-terminal domain.* J Biol Chem, 1999. **274**(52): p. 37219-25.

31. He, B., et al., *Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance.* Mol Cell, 2004. **16**(3): p. 425-38.

32. He, B., et al., *Androgen-induced NH2- and COOH-terminal Interaction Inhibits p160 coactivator recruitment by activation function 2.* J Biol Chem, 2001. **276**(45): p. 42293-301.

33. Berrevoets, C.A., et al., *Functional interactions of the AF-2 activation domain core region of the human androgen receptor with the amino-terminal domain and with the transcriptional coactivator TIF2 (transcriptional intermediary factor2).* Mol Endocrinol, 1998. **12**(8): p. 1172-83.

34. Gregory, C.W., et al., *Epidermal growth factor increases coactivation of the androgen receptor in recurrent prostate cancer.* J Biol Chem, 2004. **279**(8): p. 7119-30.

35. Gregory, C.W., et al., *A mechanism for androgen receptor-mediated prostate cancer recurrence after androgen deprivation therapy.* Cancer Res, 2001. **61**(11): p. 4315-9.

36. Bai, S., B. He, and E.M. Wilson, *Melanoma antigen gene protein MAGE-11 regulates androgen receptor function by modulating the interdomain interaction.* Mol Cell Biol, 2005. **25**(4): p. 1238-57.

37. van der Bruggen, P., et al., *A gene encoding an antigen recognized by cytolytic T lymphocytes on a human melanoma.* Science, 1991. **254**(5038): p. 1643-7.

38. De Plaen, E., et al., *Structure, chromosomal localization, and expression of 12 genes of the MAGE family.* Immunogenetics, 1994. **40**(5): p. 360-9.

39. Rogner, U.C., et al., *The melanoma antigen gene (MAGE) family is clustered in the chromosomal band Xq28.* Genomics, 1995. **29**(3): p. 725-31.

40. Lucas, S., et al., *Identification of a new MAGE gene with tumor-specific expression by representational difference analysis.* Cancer Res, 1998. **58**(4): p. 743-52.

41. Muscatelli, F., et al., *Isolation and characterization of a MAGE gene family in the Xp21.3 region.* Proc Natl Acad Sci U S A, 1995. **92**(11): p. 4987-91.

42. Pold, M., et al., *Identification of a new, unorthodox member of the MAGE gene family.* Genomics, 1999. **59**(2): p. 161-7.

43. Barker, P.A. and A. Salehi, *The MAGE proteins: emerging roles in cell cycle progression, apoptosis, and neurogenetic disease.* J Neurosci Res, 2002. **67**(6): p. 705-12.

44. Jungbluth, A.A., et al., *Expression of MAGE-antigens in normal tissues and cancer.* Int J Cancer, 2000. **85**(4): p. 460-5.

45. Duan, Z., et al., *Overexpression of MAGE/GAGE genes in paclitaxel/doxorubicin-resistant human cancer cell lines.* Clin Cancer Res, 2003. **9**(7): p. 2778-85.

46. Ohman Forslund, K. and K. Nordqvist, *The melanoma antigen genes--any clues to their functions in normal tissues?* Exp Cell Res, 2001. **265**(2): p. 185-94.

47. Tahara, K., et al., *Expression of the MAGE gene family in human hepatocellular carcinoma.* Cancer, 1999. **85**(6): p. 1234-40.

48. Kocher, T., et al., *Identification and intracellular location of MAGE-3 gene product.* Cancer Res, 1995. **55**(11): p. 2236-9.

49. Schultz-Thater, E., et al., *MAGE-1 gene product is a cytoplasmic protein.* Int J Cancer, 1994. **59**(3): p. 435-9.

50. Jurk, M., et al., *MAGE-11 protein is highly conserved in higher organisms and located predominantly in the nucleus.* Int J Cancer, 1998. **75**(5): p. 762-6.

51. Rimoldi, D., et al., *cDNA and protein characterization of human MAGE-10.* Int J Cancer, 1999. **82**(6): p. 901-7.

52. De Smet, C., et al., *DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter.* Mol Cell Biol, 1999. **19**(11): p. 7327-35.

53. De Smet, C., et al., *The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation.* Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7149-53.

54. Irvine, R.A. and G.A. Coetzee, *Additional upstream coding sequences of MAGE-11.* Immunogenetics, 1999. **49**(6): p. 585.

Figure Legends

Figure 4.1: Classical overall domain structure of nuclear receptors. (AF: Activation Function; DBD: DNA Binding Domain; LBD: Ligand Binding Domain)

Figure 4.2: Crystal structure of Androgen Receptor LBD (RCSB PDB accession code 1X0W) with inset of AR 21-30 FxxLF peptide bound at the AF-2 region. The highly helical LBD is shown in gray, with the AF-2 region highlighted in green. The FxxLF peptide is shown in purple, with charge clamp residues E897 and K720 shown in red and blue, respectively. Synthetic androgen R1881 is shown in space-filling models and sticks in yellow. Inset shows surface of AF-2 region with FxxLF peptide shown in tube form.

Figure 4.3: Chemical structures of the conversion of testosterone into dihydrotestoserone by 5α-reductase.

Figure 4.4: Sequence alignment of the MAGE-A family. Alignment was done using LALIGN and was diagrammed using BOXSHADE. Green represents complete identity, yellow near identity, and blue conserved. MAGE homology domain is highlighted by a red box. Notice the additional up-stream coding region of MAGE-11.

Figure 4.5: Model for the influence of MAGE-11 on AR transactivation. MAGE-11 disrupts the AR N/C interaction by binding the FxxLF motif of AR. Co-activators of the SRC/p160 class then bind to the free AF-2 region by way of the LxxLL box peptide, facilitating androgen dependent transcription.

Figure 4.1

Figure 4.2

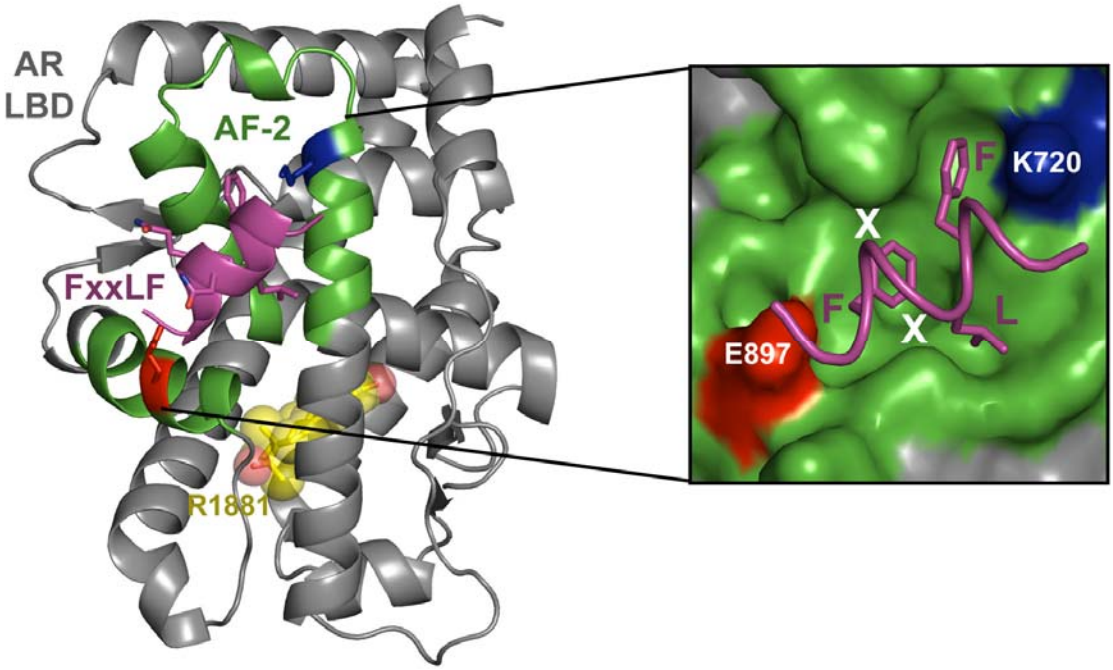


Figure 4.3

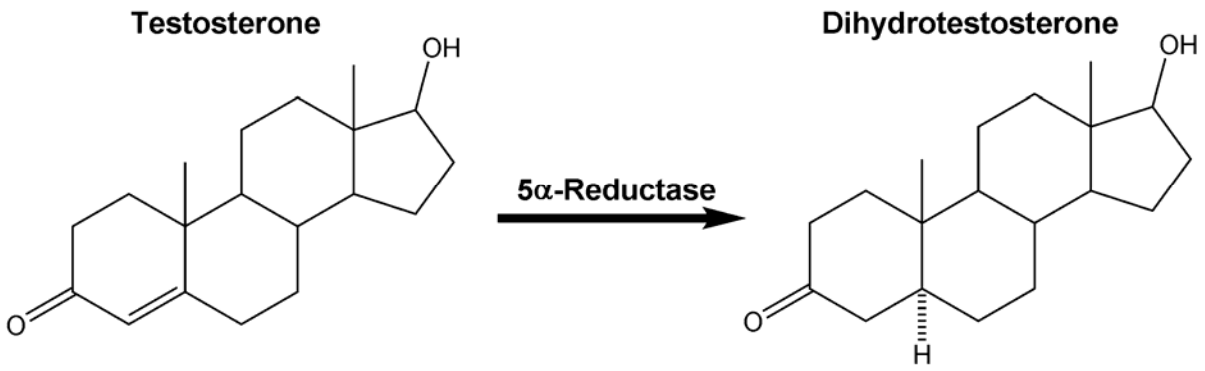


Figure 4.4

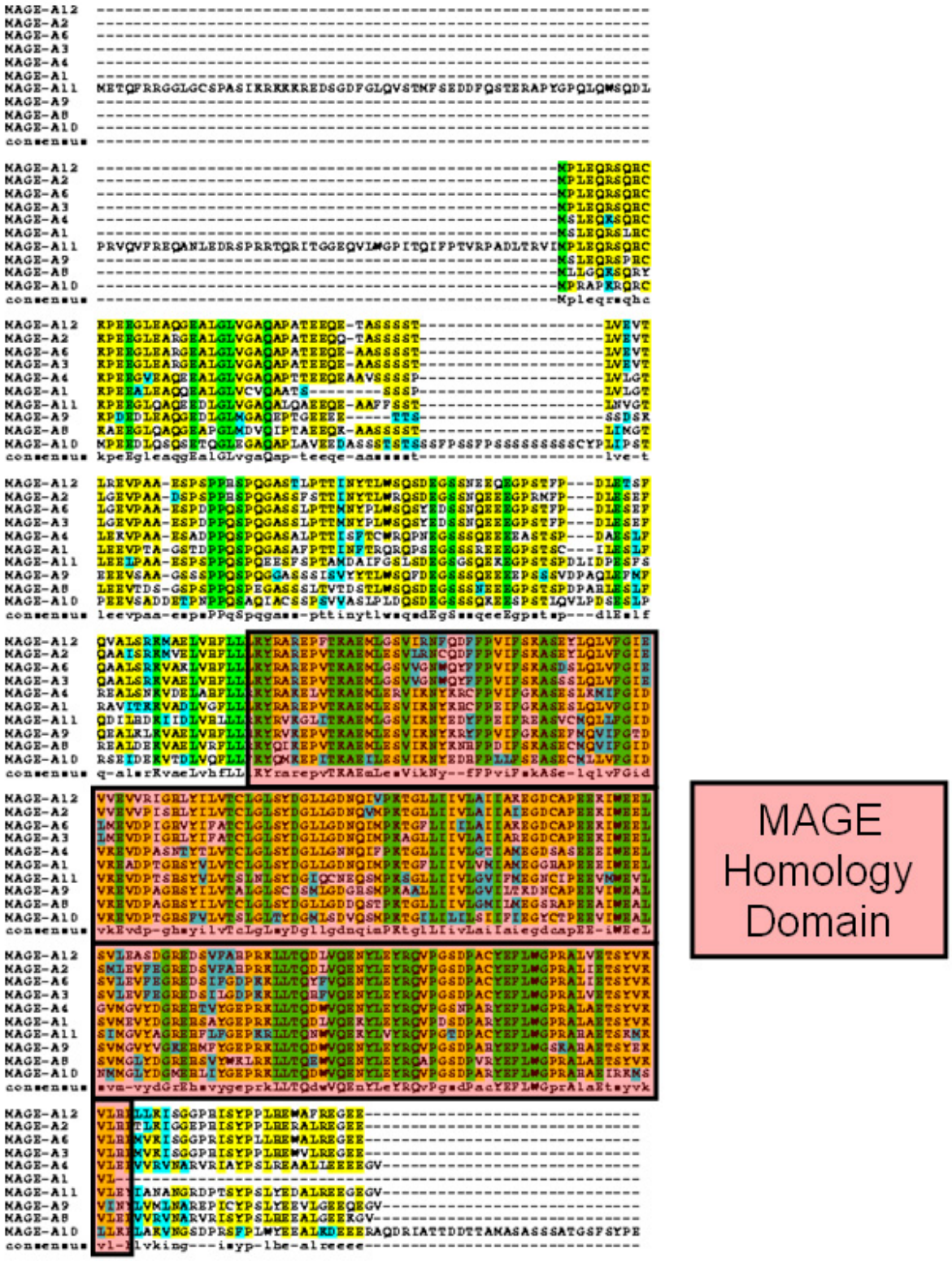
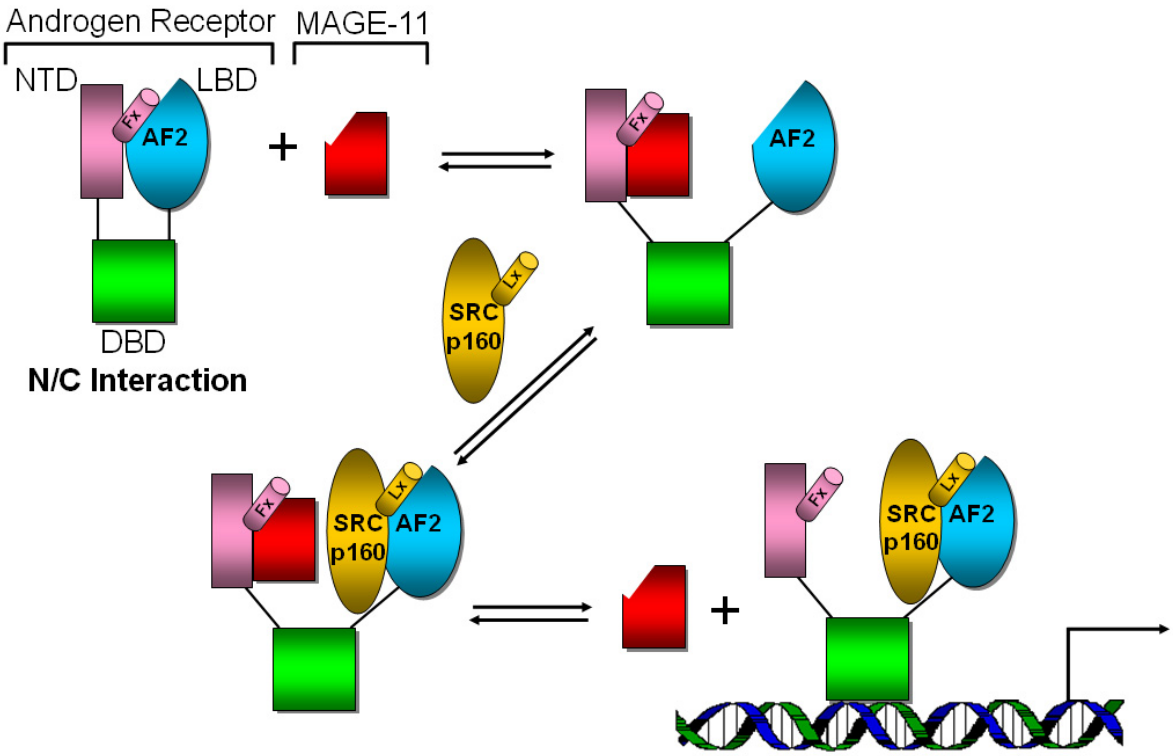


Figure 4.5


Androgen Receptor
MAGE-11
NTD
LBD
Fx
AF2
DBD
N/C Interaction
SRC
p160
Lx

Chapter 5:

Efforts towards the Purification and Biophysical Characterization of the Androgen Receptor Modulator MAGE-11

Introduction

The androgen receptor (AR) is a member of the steroid nuclear receptor family that is responsible for the regulation of androgen-specific genes [1, 2]. AR activity is modulated by co-activators that work to recruit transcriptional machinery to facilitate the expression of these gene products [3, 4]. Co-activators of the p160/SRC class bind to the activation function 2 (AF-2) region of the ligand binding domain (LBD) of AR through an LxxLL motif that forms an amphipathic helix [5]. Upon androgen binding, however, AR exhibits a self-association between its AF-2 region and an internal FxxLF motif found in its N-terminal domain (NTD) [6]. This N/C-terminus interaction prohibits co-activator recruitment by blocking the AF-2 site from being bound by the LxxLL motif [7]. The melanoma activated gene product MAGE-11 protein was found by yeast two hybrid screening to disrupt the N/C interaction by binding the FxxLF motif of the NTD, exposing the AF-2 region and allowing co-activators to bind and regulate AR transactivation [8]. The interaction between MAGE-11 and the FxxLF motif of AR therefore warrants examination and further characterization to fully understand the role MAGE-11 plays in AR mediated gene expression. This chapter outlines the efforts made towards the purification of MAGE-11, and further examines the interaction between MAGE-11 and AR by fluorescence polarization assays.

Materials and Methods

Construct and Expression Vector Production

The open reading frame of the MAGE-11 gene consists of 1287 bp encoding 429 amino acids and was amplified out of a testis cDNA library by PCR. Initial vectors for examination were kindly provided by the laboratory of Dr. Elizabeth Wilson. PCR products encoding residues 2-429, 112-429, 218-413, and 222-429 of MAGE-11 were cloned into the EcoR1-Xho1 site of pGEX-4T1 vectors to express proteins as fusions to a thrombin cleavable glutathione-S-transferase (GST) tag. MAGE-11 constructs were further cloned into additional vectors in order to test purifications utilizing nickel affinity. MAGE-11 222-429 was amplified out of pGEX-4T1-MAGE-11-2-429 using primers with Nde1 and Xho1 ends for placement into pET-15b (Novagen). Constructs 222-429, 112-429, and 2-429 were also placed into a variant of the pMAL-C2H10 vector using primers containing Sal1 and Pst1 ends. This vector creates a fusion of MAGE-11 to maltose binding protein (MBP), with a 10x His tag and tobacco etch virus (TEV) protease site to cleave the dual tag from MAGE-11. The pMAL-C2H10 vector was also used to create fusion protein to a fragment of the androgen receptor of residues 4-52. Any site directed mutagenesis was accomplished using the QuikChange Mutagenesis kit (Stratagene) with primers encoding individual mutations. All constructs and mutations were confirmed using automated DNA sequencing (UNC genome analysis center).

Bacterial Expression

Fusion proteins of MAGE-11 and AR were produced in bacteria using several methods, but a general protocol was optimized for overexpression. Vectors encoding MBP-tagged versions of MAGE-11 were transformed into the E. Coli cell line BL21(DE3) gold and plated onto LB agar plates containing 50 μg/ml ampicillin (Amp) for selection of transformed colonies. Single colonies were then transferred to 100 ml of fresh LB liquid media containing 100 μg/ml Amp and allowed to grow for 16 h at 37 °C. Larger cultures were inoculated with 1% of overnight and allowed to grow to mid-log phase at an optical density (OD_{600}) of 0.6-0.8. Protein expression was induced by the addition of 1 mM isopropyl-b-D-1-thiogalactopyranoside (IPTG) and allowed to grow for 20 h at 15 °C. Cells were then pelleted by centrifugation and stored at -80 °C.

General Purification of MBP-MAGE-11 Fusions

Cell pellets of *E.Coli* expressing MBP-MAGE-11 fusions were thawed at 4 °C overnight, and suspended in an appropriate volume of nickel buffer A (NiA) (50 mM sodium phosphate pH 7.5, 500 mM NaCl, 20 mM imidizole , 5% (w/v) glycerol). Cells were then sonicated for 5 minutes on ice in the presence of protease inhibitors (PMSF at 100 μg/mL, leupeptin and aprotein at 1 μg/mL each) and DNAse (10 μg/ml) for proper cell lysis. Cell debris was removed by centrifugation at 37,000g for 90 minutes. Cell lysate was then loaded directly onto a 10 ml Ni-Sepharose-6 (Amersham) gravity column and allowed to bind under agitation for 10 minutes. Lysate was then eluted off, followed by ten column volumes (CV) of buffer A to wash away unbound protein. Elution of bound MBP-MAGE-11 was accomplished with 5 CV of nickel buffer B (NiB) (50 mM sodium phosphate pH 7.5, 500

mM NaCl, 300 mM imidizole, 5% (w/v) glycerol) or until no more protein was detected by Bradford. Fractions were examined by SDS-PAGE, and those that contained MBP-MAGE-11 were pooled. DTT was added to 10 mM and 1-2% (w/w) of His-tagged TEV protease was added for MBP-10xHis removal and allowed to sit at room temperature for 4 h. Sample was subsequently dialyzed into 4 L of NiA overnight at 4 °C, and then loaded again onto a Ni-Speharose-6 column to trap cleaved MBP and TEV. Untagged MAGE-11 passed through, examined by SDS-PAGE, pooled and concentrated to an appropriate volume. Several additional column methods were attempted, and are fully explained below in the results.

FLAG-MAGE-11 Purification

Spodoptera frugiperda (Sf21) cell pellets expressing full length MAGE-11 (residues 1-429) tagged with a C-terminal FLAG peptide (DYKDDDDK) were kindly provided by the laboratory of Dr. Elizabeth Wilson. Pellets were resuspended in 50 ml buffer A (25 mM Tris-HCl pH 7.8, 300 mM NaCl, 1 mM EDTA, 50 mM NaF, 0.05% (v/v) IPEGAL, 10% (w/v) glycerol, 250 μM AEBSF, 2 μg/mL Aprotein & Leupeptin, 1 μg/mL PepstatinA, 0.8 mM DTT) and split into 4-15 mL falcon tubes. Cells were then treated to 3 freeze/thaw cycles for proper cell lysis. Cells debris was pelleted at 27,000g for two hours in an ultracentrifuge. Cell lysate was removed and incubated with 1.5 mL of pre-equilibrated ANTI-FLAG agarose beads (Sigma) for 3 hours at 4 °C with gentle agitation. Beads were pelleted at 3000 rpm for 3 min and washed in 13 mL buffer B (same components as buffer A, but with 500 mM NaCl) for 10 min under gentle agitation three times to allow for complete removal of non-specific proteins. After last wash, pelleted beads were then agitated in 5 CV buffer C (buffer A with 150 mM NaCl) containing 0.15 mg/ml FLAG peptide (UNC Peptide Facility) for 10 min to

elute FLAG-MAGE-11. Bead slurry was placed into a small gravity column and fractions were collected, with two additional 1 CV washes of buffer C plus peptide. Fractions were analyzed using SDS-PAGE for purity and pooled. Additional purification steps are outlined fully below.

Fluorescence Polarization Assays

A fluorescein-labeled peptide (residues 16-36) containing the FxxLF motif of AR (SKTYRGAFQNLFQSVREVIQN-[KFlc], SigmaGenosys) was used to determine the binding constant of MAGE-11 to AR. Fluoroscein absorbs light at 485 nanometers and emits at 520 nanometers. Serial dilutions of FLAG-MAGE-11 (3.7 μM starting concentration, 1-1000 nM sample concentrations) were incubated at 22 °C for 30 minutes in the presence of 100 nM peptide, with buffer conditions of protein and peptide corrected for. Bovine serum albumin (NEB) was used at the same concentrations as a negative protein control, and a fluorescein-labeled peptide of residues 681-696 of SRC-1 containing an LxxLL motif (SLTERHKILHRLLQE-[KFlc], SynPep) as the negative peptide control. Fluorescence readings were taken in 96-well format using a PHERAstar (BMG Labtech) plate reader with readings every 30 seconds for 5 minutes. All readings were done in triplicate, and averaged across time and samples. Data analysis was done using Microsoft Excel and SigmaPlot.

Limited Proteolysis Assays

Samples of FLAG-MAGE-11 were treated with proteases to determine the stability of protein domains. 15 μg of both trypsin (which cleaves C-terminal to basic Lys, Arg residues) and chymotrypsin (which cleaves C-terminal to aromatic Phe, Tyr, and Trp residues) were

added to 20 μg of full length FLAG-MAGE-11 in the absence and presence of 5 micromoles AR 16-36 peptide. Samples were taken at 15, 30, and 60 minutes and quenched with 100 μg/ml PMSF. Samples were subsequently evaluated by SDS-PAGE.

Results and Discussion

Rationale for MAGE-11 Construct & Vector Selection

Constructs of MAGE-11 for structural studies were determined by several bioinformatic methods. Initial constructs were based off of previous work classifying MAGE proteins [9, 10]. Full length MAGE-11 (1-429) encompassed the entire MAGE homology domain and the extended upstream coding region (1-111). Several of the MAGE-A family proteins are encoded by a single exon that encompasses residues 112-429 of MAGE-11, and therefore this construct (which is initiated at the third Met in the MAGE-11 sequence) was examined as well. Shorter, truncated versions (218-413, 222-429, 218-429) were based off of secondary structure predictors and a PONDR analysis, which indicated this region was a tight, compact domain (Figure 5.1). Therefore, this domain was also cloned into several vectors for expression.

Bacterial expression vectors were chosen for purification schemes and ability to produce soluble protein. Initial cloning of MAGE-11 constructs (1-429, 112-429, 218-413) into pGEX-4T1 vectors should have provided soluble protein since GST has been shown to increase solubility of unstable proteins [11]. However, several rounds of test expressions examining numerous cell lines and growth conditions of GST-MAGE-11 fusions failed to yield significant quantities of usable protein (data not shown). Only the shorter MAGE-11 construct 218-413 appeared to have the potential for additional testing. Therefore, cloning of the shorter MAGE-11 constructs (218-413, 222-429) into pET-15b and pMAL-C2H10 vectors was done using standard methods. pET-15b also failed to yield significant quantities

of soluble protein, but the pMAL system appeared to work the best (data not shown). This is likely due to the high level of solubility MBP exhibits, acting as a "life-raft" for MAGE-11. MBP fusions of all shorter constructs (222-429, 218-429, 112-429) produced high levels of soluble protein, while full length MAGE-11 appears to be permanently insoluble for bacterial expression. Therefore, the pMAL system appears to be the best vector for expression and purification of MAGE-11 constructs, which is outlined more fully below.

Expression and Purification of MAGE-11 Constructs

E. Coli expression conditions of MAGE-11 constructs were found by several iterative rounds of smaller test expressions. Temperature, cell line, IPTG concentration, and length of growth were all sampled. The conditions that yielded the highest level of soluble protein were under low temperature (15 °C) for an extended period (20 hours) after induction by a moderate amount of (1 mM) of IPTG. These conditions consistently gave high amounts (>100 mg/3L growth) of fusion for MAGE-11 constructs 218-413, 222-429, 218-429, and 112-429. This amount of protein allowed for additional purification conditions to be sampled throughout each individual protein preparation.

A typical nickel-affinity purification run of MBP-MAGE-11 (222-429, the first construct tested) is outlined in Figure 5.2. Cell lysate was loaded onto a Ni-speharose-6 column, which retained the MBP-MAGE-11 fusion by way of its 10xHis tag that is present between the two proteins. Once the protein was eluted from the column, fractions containing fusion were treated with DTT and TEV protease to remove the MBP-10xHis from the MAGE-11 protein. The sample was dialyzed back into NiA buffer overnight and

subsequently re-passed over the nickel column to allow for MAGE-11 to flow through and the His-tagged MBP to bind to the column. Optimizing the buffer conditions of the initial nickel affinity runs proved daunting, considering the troublesome nature of the protein. Buffers were altered from Tris based to sodium phosphate based, since Ni-sepharose-6 resin requires a strong charge to mask the extended sepharose chain holding the Ni^{2+} ions. Increasing sodium chloride from 150 mM to 500 mM also gave slightly better results. However, these increases in purity were only moderate, and therefore the buffer conditions stated in the materials and methods section were held after these optimization runs. Additional additives were also tested in an attempt to limit non-specific protein binding to the column, such as reducing agents (e.g. DTT), detergents (e.g. CHAPS), and mild denaturants (e.g. 1 M urea). These additives did not seem to increase purity, however, and therefore were generally left out. At this point, MAGE-11 appears roughly 70% pure by SDS-PAGE and ready for additional column chromatography steps to further remove contaminants and increase its purity.

Considering how MAGE-11 (222-429) looked to be the only protein at roughly 25 kDa on an SDS-PAGE gel, size exclusion chromatography (SEC) seemed to be the next logical step in the purification. Initial runs of a Superose-12 (GE Healthcare) column in minimal buffer conditions (20 mM HEPES, 150 mM NaCl, 5% (w/v) glycerol) showed that the protein was primarily running in the void volume (Figure 5.3). This result was surprising, considering that the void volume of the Superose-12 column is surpassed by anything over 200 kDa in size. Additional runs on a Superdex-75 and Superdex-200 (GE Healthcare) column also showed MAGE-11 primarily in the void volume. It was hypothesized at this

point that MAGE-11 was forming a soluble aggregate that was mediated by disulfide linkages between monomers. SDS-PAGE analysis with and without β-mercaptoethanol seemed to confirm this result (Figure 5.4). Therefore, the decision was made at this point to mutate the three Cys residues in this construct to serine (C270S, C304S, C328S) in an attempt to limit this association. Additionally, the construct was changed to 218-429 to encompass four residues that secondary structure predictions say are the terminal helix of this domain in MAGE-11.

Once the Cys→Ser mutant forms of MAGE-11 218-429 were produced, test purifications were conducted under standard methods. Single mutant C270S showed to limit higher order species previously seen by SDS-PAGE, but was also found in the void volume of a Superdex-75 size exclusion column (Figure 5.4). Other single and double mutation (C304S, C328S, C270S-C328S) versions gave similar results. This indicated that all cystines must be removed to facilitate prevention of soluble aggregate formation. However, once all cystines were mutated to serines (C270S-C304S-C328S), size exclusion chromatography still indicated that MAGE-11 218-429 continued to be in the soluble aggregate form (Figure 5.4). Secondary methods were tested on WT 218-429 to mimic the Cys→Ser mutations, such as exposing MAGE-11 to iodoacetamide (which blocks free cystines from forming disulfides) and running SEC with high levels (up to 50 mM DTT) of reducing agent. Additionally, CHAPS was also added to SEC buffers in the attempt to abrogate hydrophobic interactions. MAGE-11 persisted in the void volume of all SEC columns despite all efforts to prevent it. Consequently, it appeared that MAGE-11 aggregation was not disulfide dependent, and was either an unfolded peptide chain or extremely hydrophobic protein that cannot be dissociated

into monomeric form. The MAGE-11 218-429 construct was therefore abandoned from all further attempts at purification.

The lessons learned from MAGE-11 218-429 were applied to the purification of the longer MAGE-11 domain 112-429. It was hypothesized that the longer construct may allow for easier purification if the additional residues act to cover the hydrophobic patches found in the 218-429 construct. Additionally, adding a fragment of AR may also help in the purification of MAGE-11 by acting to control any sort of aggregation. MAGE-11 112-429 expression and nickel affinity purification conditions were nearly identical to those of the shorter construct (Figure 5.5), and yielded similar amounts of protein. Once SEC methods were attempted, however, MAGE-11 112-429 was again found primarily in the void volume of the Superdex-200 column. A smaller peak at the correct molecular weight of MAGE-11 112-429 turned out to be MBP that had been trapped in the soluble aggregate. This result was confirmed by passing the sample over an amylose column, of which MBP binds with high affinity (Figure 5.5). As a last resort, a small fragment of AR (residues 4-52, encompassing the FxxLF motif) was made and expressed as an MBP fusion protein, and was purified in parallel to MAGE-11 112-429 (Figure 5.6). These two proteins were then combined and cleaved by TEV together to see if this would promote a stable complex between AR and MAGE-11. Predictably, when passed over the Superdex-200 column, the majority of the protein remained in the void volume and other high molecular weight species, again indicating soluble aggregate. Therefore, the hopes of a bacterially expressed, soluble MAGE-11 were abandoned and all efforts were turned towards insect cell expression systems.

Insect Cell Expression and Purification of MAGE-11

Baculovirus mediated expression of full length MAGE-11 in insect cell lines proved to be the best way to achieve soluble protein. Utilizing FLAG chromatography has been successful in isolating the protein, albeit at limited amounts (Figure 5.7). Typical recovery from 100 mL of cells is between 1-2 mg of FL-MAGE-11. This amount may be limited by the type of tag used, and switching to a simple His tag may increase gains. Additional purification was accomplished using a Superdex-200 using buffer conditions (20 mM Tris pH 7.8, 350 mM NaCl, 10 % (w/v) glycerol, 8 mM CHAPS) that have been previously determined to yield monomeric protein. A large amount of protein is still found to be in the void volume, and running the column under low salt (150 mM) or without CHAPS from the buffer only increases the void volume peak. Despite these issues, the fractions containing monomeric MAGE-11 were concentrated to 2.5 mg/ml and run through one round of screening for preliminary crystallization conditions. However, no viable crystal hits were found, and therefore additional purifications must be done to test more conditions in the presence and absence of FxxLF peptide.

In the attempt to further characterize MAGE-11, limited proteolysis assays were done to see if any domains appear to be resistant to cleavage, and potentially yield additional constructs for future examination (Figure 5.8). The experiments were done with both chymotrypsin and trypsin in the presence and absence of FxxLF-containing AR 16-36 peptide. Cleavage by trypsin appeared to produce one band at roughly 20 kDa in size for both with and without peptide samples. The majority of this fragment likely consists of a stretch between residues 116-202 (which contains no trypsin cleavage sites) and further C-terminal

additional residues that are undetermined. Peptide mass spectroscopy (MS/MS) and N-terminal sequencing of this fragment needs to be done in order to absolutely confirm the sequence of MAGE-11. In terms of chymotrypsin, 90% of the protein has been cleaved after only 15 minutes time, with only one faint band running at ~27 kDa. Interestingly, however, the presence of AR peptide appears to stabilize this band enough to withstand 30 minutes with chymotrypsin. This band is fully removed at 30 minutes without peptide. This would indicate that MAGE-11 is partially stabilized by peptide binding, and may even gain structure upon association with AR. Predicting the sequence of this peptide is difficult, however, since chymotrypsin cleavage sites are spread equally throughout the primary sequence of MAGE-11. As with the trypsin results, these peptides need to be subjected to MS/MS and N-terminal sequencing analysis to confirm what residues this contains. One could assume that this band corresponds to part of the region of MAGE-11 previously identified to interact with AR [8]. However, experimental confirmation is required to accurately classify the residues involved.

Fluorescence Polarization Assays

Considering that MAGE-11 is known to bind the FxxLF motif found in the NTD of AR, and that peptide binding appears to confer some structural stability to full length MAGE-11, fluorescence polarization (FP) assays were done to determine a binding constant of MAGE-11 to the FxxLF motif of AR. As described in the materials and methods section of this chapter, assays were done using a synthesized peptide encoding residues 16-36 of AR covalently bound to a fluoroscein molecule in 96 well format. The level of binding was conferred from the polarization of the light emitted from the fluoroscein upon excitation. The

FP experiments showed that full length MAGE-11 bound to the FxxLF region of AR with an average K_D of 131 nM (Figure 5.9). Several peptide probe concentrations were tested, and all gave a K_D between 101 and 173 nM (Table 5.1). Multiple negative control runs were done, and no binding was seen between BSA and the AR peptide probe. MAGE-11 also did not bind a fluorescently labeled SHP Box1 peptide (that contains an LxxLL motif), confirming that MAGE-11 is selective for the FxxLF motif. Surprisingly, the data did not fit a standard one-site model with good R-square values. The data best fit the hill equation, which is usually used in the measure of cooperativity of binding. A strong level of cooperativity was evident in all of the binding data, with Hill constants found to be between 2.5 and 5. These results indicate that MAGE-11 forms higher order species, and facilitates the binding of the AR peptide by other monomers of MAGE-11. The ligand binding domain of AR has been previously shown to bind the FxxLF peptide (residues 20-30) with a K_D in the low micromolar range, depending on the androgen bound (Table 5.1). In the presence of the synthetic androgen R1881, AR was found to bind the FxxLF peptide at a K_D of 9.2 ± 0.4 μM, with testosterone and DHT giving lower K_D's of 5.5 ± 0.3 and 3.0 ± 0.4 μM, respectively ([12] and unpublished data). Binding to the box III LxxLL peptide of the co-activator TIF2 by AR exhibited much higher K_D values, ranging from 13.1 ± 1.5 μM for DHT to 78 ± 28 μM for R1881 ([12] and unpublished data). The fact that MAGE-11 binds the FxxLF motif an order of magnitude over the AR LBD supports the conclusion that MAGE-11 disrupts the AR N/C interaction by out competing the LBD for the FxxLF region. Taken together, these results clarify the role MAGE-11 plays as a modulator of AR activity.

Conclusions and Future Directions

The androgen receptor modulator MAGE-11 has been shown to play an important role in AR mediated transactivation of androgen-specific genes. Structural studies of MAGE-11 would provide the best insight into how AR is influenced by MAGE-11, and how it disrupts the N/C interaction of AR. Despite the several attempts at the bacterial expression and purification of MAGE-11 constructs, soluble, monomeric protein was unable to be produced in high quantities. The persistent soluble aggregate produced after cleavage from MBP complicated the process, since the protein did not indicate its instability by falling out of solution. Additionally, adding in a binding partner upon cleavage failed to confer a higher level of stability in solution. These findings indicate that MAGE-11 likely requires post-translational modifications to induce proper folding of the protein. Several phosphorylation sites have been identified by sequence mapping and MS analysis at residues 174 and 208 of MAGE-11 (E. Wilson, personal communication). While phosphorylation can be mimicked by mutagenesis, it is still unclear as to if any other post-translational modifications (glycosylation, acetylation, etc.) are present and required for stability. Regardless, it appears that the best way to produce MAGE-11 for any future attempts at crystallization should be done using insect cell expression systems. Additionally, MAGE-11 should be transferred into an expression system where purification can be done using nickel affinity, rather than FLAG chromatography. While FLAG purification is generally more selective, the lifetime of the resin (limited to four uses) and the low binding capacity (up to 3 mg protein/mL resin) do not make it a viable system for the purification high quantities of protein required for

crystallization. Therefore, other avenues of expression and purification should be attempted for future studies of this protein.

The most interesting and conclusive result found with MAGE-11 stems from the FP experiments. MAGE-11 was found to bind the FxxLF motif of AR an order of magnitude over the AF-2 region of AR. This result works to further understand the reasoning behind why MAGE-11 is so effective at modulating AR activity. By directly competing for the FxxLF motif, MAGE-11 is able to open the AF-2 region for co-activator binding. The tight binding of MAGE-11 to the FxxLF appears to be necessary, since the AR LBD only exhibits moderate affinity for the LxxLL peptide of co-activators (Table 5.1). This order of magnitude molecular "switch" between affinities of MAGE-11 to FxxLF, the AF-2 of AR for FxxLF, and the AF-2 of AR for LxxLL, is likely a secondary level of regulation for AR activity. Furthermore, the level of cooperativity seen in the binding of MAGE-11 to the AR 16-36 peptide implies that monomers of MAGE-11 work in concert with each other to completely tie up the FxxLF motif, and may provide more levels of regulation to the process.

Finally, while the interaction between MAGE-11 and AR is likely important for understand AR-mediated prostate cancer tumorigenesis, MAGE-11 does not appear to be a viable structural target. Perhaps other MAGE-A family proteins, with their high level of sequence identity to MAGE-11, would provide some insights into the structural characteristics of the overall MAGE fold. However, no other members have been implicated as interacting with AR, and therefore would not provide all of the information required to fully comprehend this important interaction.

REFERENCES

1. Aranda, A. and A. Pascual, *Nuclear hormone receptors and gene expression.* Physiol Rev, 2001. **81**(3): p. 1269-304.

2. Keller, E.T., W.B. Ershler, and C. Chang, *The androgen receptor: a mediator of diverse responses.* Front Biosci, 1996. **1**: p. d59-71.

3. McKenna, N.J., et al., *Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions.* J Steroid Biochem Mol Biol, 1999. **69**(1-6): p. 3-12.

4. Onate, S.A., et al., *Sequence and characterization of a coactivator for the steroid hormone receptor superfamily.* Science, 1995. **270**(5240): p. 1354-7.

5. Heery, D.M., et al., *A signature motif in transcriptional co-activators mediates binding to nuclear receptors.* Nature, 1997. **387**(6634): p. 733-6.

6. He, B. and E.M. Wilson, *The NH(2)-terminal and carboxyl-terminal interaction in the human androgen receptor.* Mol Genet Metab, 2002. **75**(4): p. 293-8.

7. He, B., et al., *Androgen-induced NH2- and COOH-terminal Interaction Inhibits p160 coactivator recruitment by activation function 2.* J Biol Chem, 2001. **276**(45): p. 42293-301.

8. Bai, S., B. He, and E.M. Wilson, *Melanoma antigen gene protein MAGE-11 regulates androgen receptor function by modulating the interdomain interaction.* Mol Cell Biol, 2005. **25**(4): p. 1238-57.

9. Barker, P.A. and A. Salehi, *The MAGE proteins: emerging roles in cell cycle progression, apoptosis, and neurogenetic disease.* J Neurosci Res, 2002. **67**(6): p. 705-12.

10. De Plaen, E., et al., *Structure, chromosomal localization, and expression of 12 genes of the MAGE family.* Immunogenetics, 1994. **40**(5): p. 360-9.

11. Smith, D.B. and K.S. Johnson, *Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase.* Gene, 1988. **67**(1): p. 31-40.

12. He, B., et al., *Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance.* Mol Cell, 2004. **16**(3): p. 425-38.

Figure Legends

Figure 5.1: Construct design of MAGE-11. Top graph is a PONDR (Predictor of naturally disordered regions) plot showing the level of disorder across the MAGE-11 primary sequence. Secondary structure prediction (GOR4, www.expasy.org) shows helix (Blue) and beta sheet (Red) locations. Bottom shows constructs used in MAGE-11 expression based off of the top two plots.

Figure 5.2: SDS-PAGE of initial nickel affinity steps in the purification of MAGE-11 222-429. MAGE-11 and MBP are highlighted in the second gel.

Figure 5.3: Chromatogram of MAGE-11 222-429 after size exclusion chromatography run using a Superose-12 column. SDS-PAGE gel below corresponds to the fractions from the chromatography run. MAGE-11 is the major band near 20 kDa on the gel.

Figure 5.4: Cys to Ser mutants of MAGE-11 218-429. Left panel shows WT MAGE-11 218-429 with and without β-mercaptoethanol (β-ME), indicating some level of intramolecular disulfide formation. Middle panel is chromatogram of MAGE-11 218-429 C270S over a Superdex-75 column and its accompanying SDS-PAGE gel. Right panel is a chromatogram of the triple mutant (C270S-C304S-C328S) of MAGE-11 218-429 over a Superdex-200 column. Both chromatograms indicate soluble aggregate formation.

Figure 5.5: Purification run of MBP fusion of MAGE-11 112-429. Top left and bottom left SDS-PAGE gels are of the nickel affinity runs. Lower right SDS-PAGE gel shows both results from amylose column and Superdex-200. Chromatogram is of the Superdex-200 run.

Figure 5.6: Dual purification of MAGE-11 112-429 and AR 4-52, both as MBP fusions. SDS-PAGE gels on the left indicated nickel affinity runs both before and after cleavage by TEV. Right panels are the following chromatogram and SDS-PAGE analysis of a Superdex-200 column. The location of the AR 4-52 fragment is indicated.

Figure 5.7: Purification of insect cell expressed full length MAGE-11. Left panel is of SDS-PAGE analysis of FLAG-antibody column purification step. Right panel is chromatogram and SDS-PAGE analysis of Superdex-200 column. Location of MAGE-11 monomer is indicated.

Figure 5.8: Limited proteolysis of full length MAGE-11. SDS-PAGE gel shows peptide fragments produced from treatment of MAGE-11 with proteases. Samples with AR 16-36 peptide are indicated.

Figure 5.9: Fluorescence polarization assays of full length MAGE-11 with the AR 16-36 peptide. Legend corresponds to peptide concentrations tested. Notice no curve is shown for the negative control, SHP box1 peptide. Data values are culled in Table 5.1.

Figure 5.1

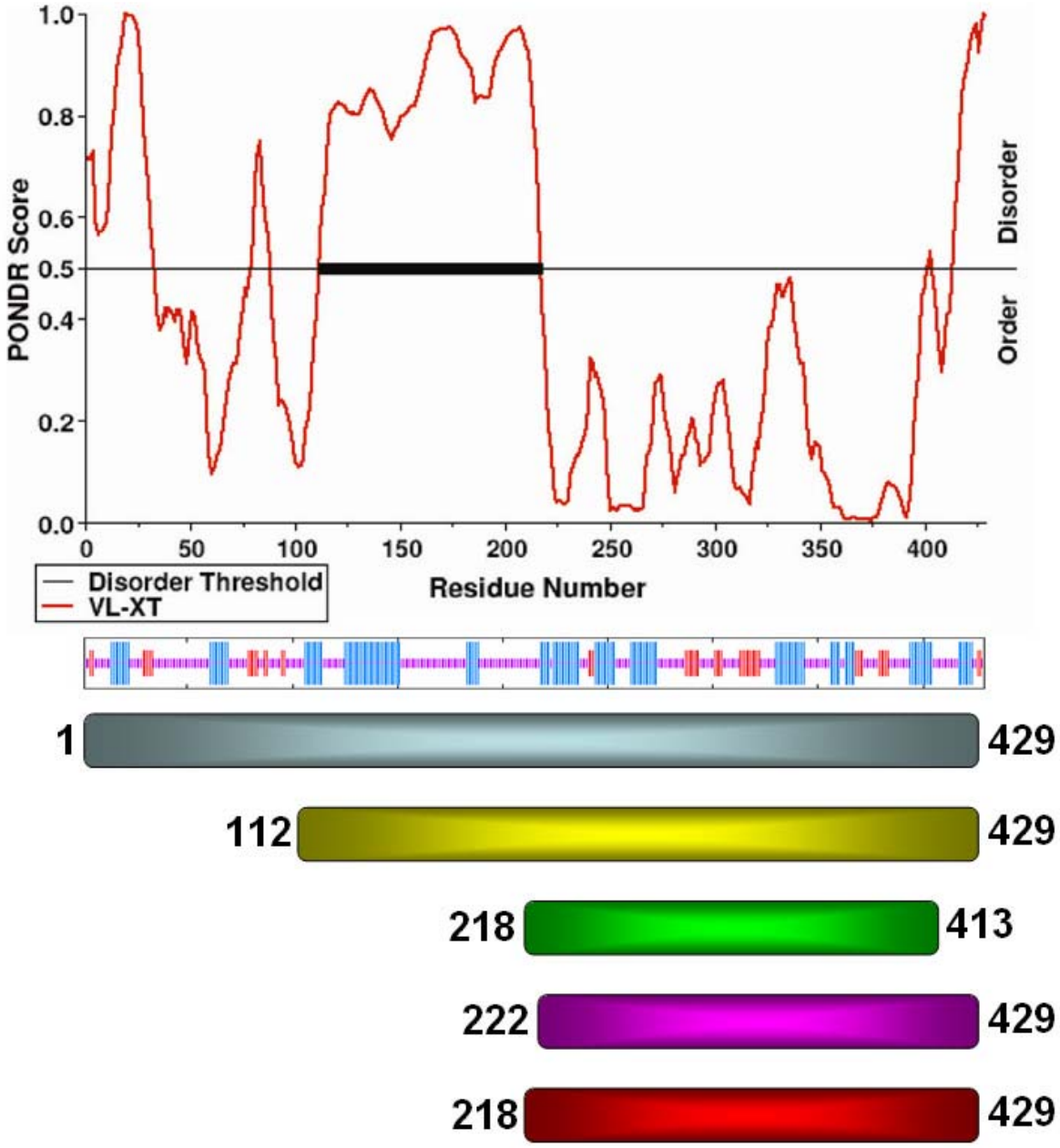

PONDR Score
1.0
0.8
0.6
0.5
0.4
0.2
0.0
Disorder
Order
0
50
100
150
200
250
300
350
400
Residue Number
Disorder Threshold
VL-XT
1
429
112
429
218
413
222
429
218
429

Figure 5.2

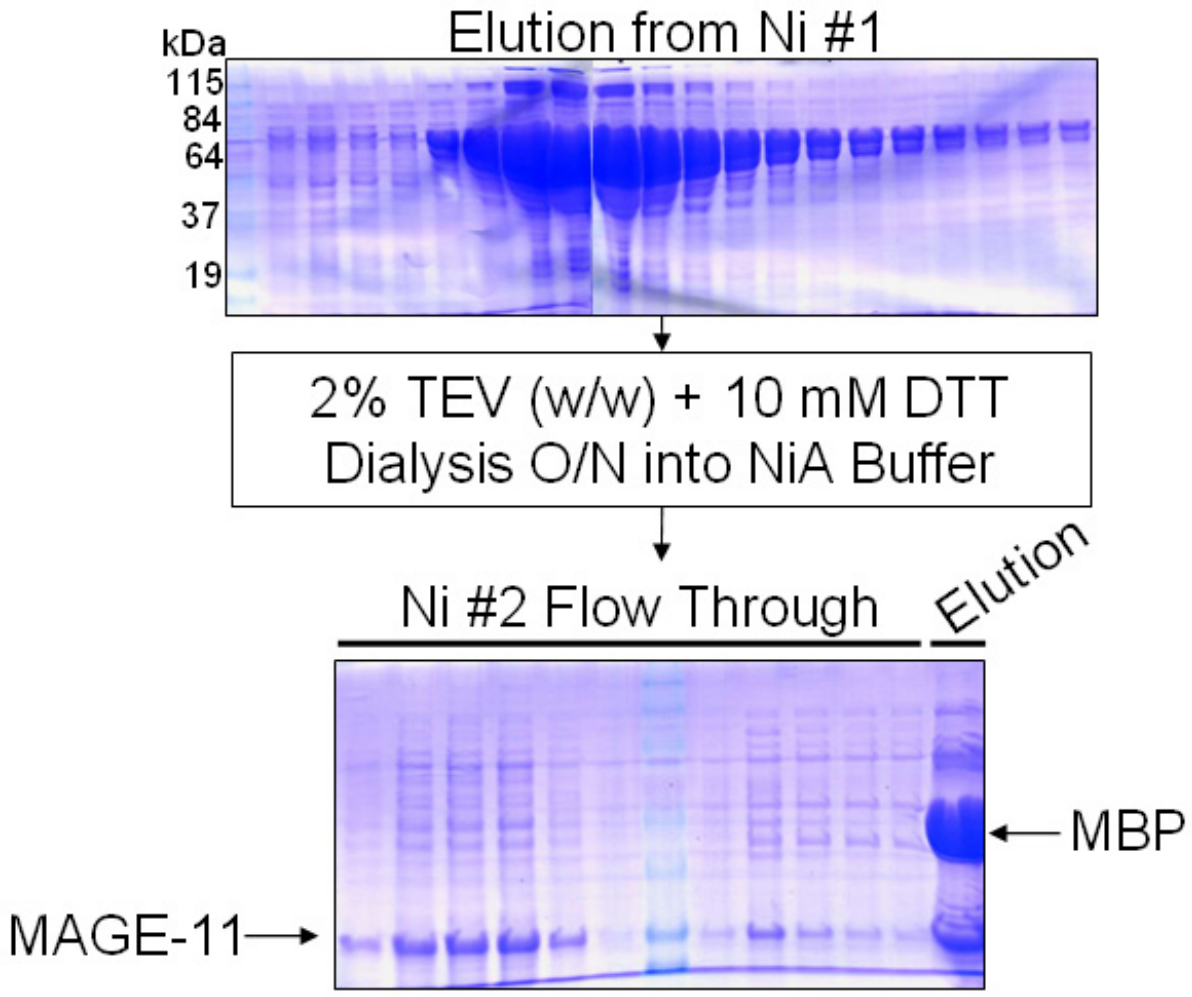

kDa
115
84
64
37
19
Elution from Ni #1
2% TEV (w/w) + 10 mM DTT
Dialysis O/N into NiA Buffer
Ni #2 Flow Through
Elution
MBP
MAGE-11

Figure 5.3

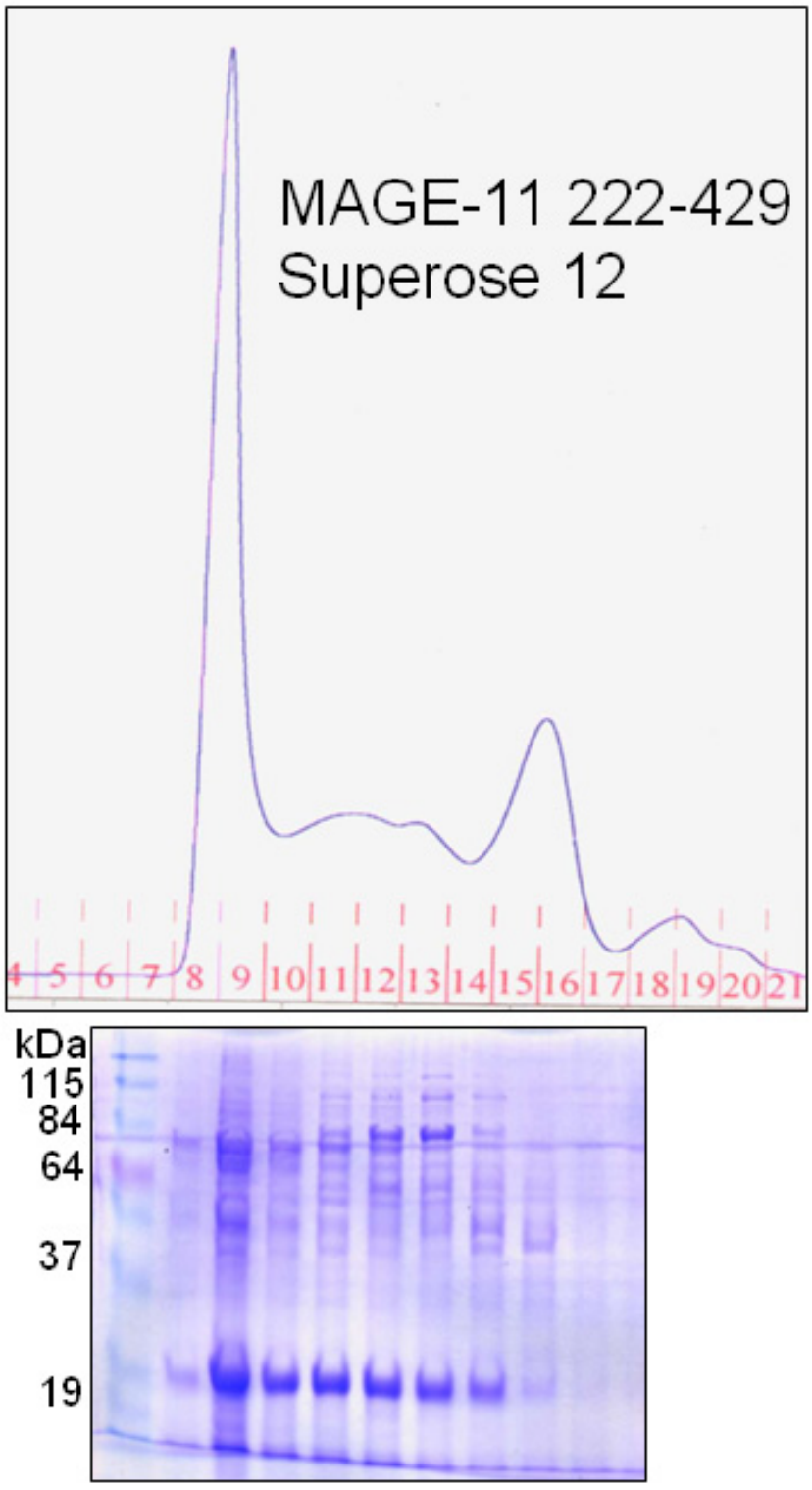

MAGE-11 222-429
Superose 12
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
kDa
115
84
64
37
19

Figure 5.4

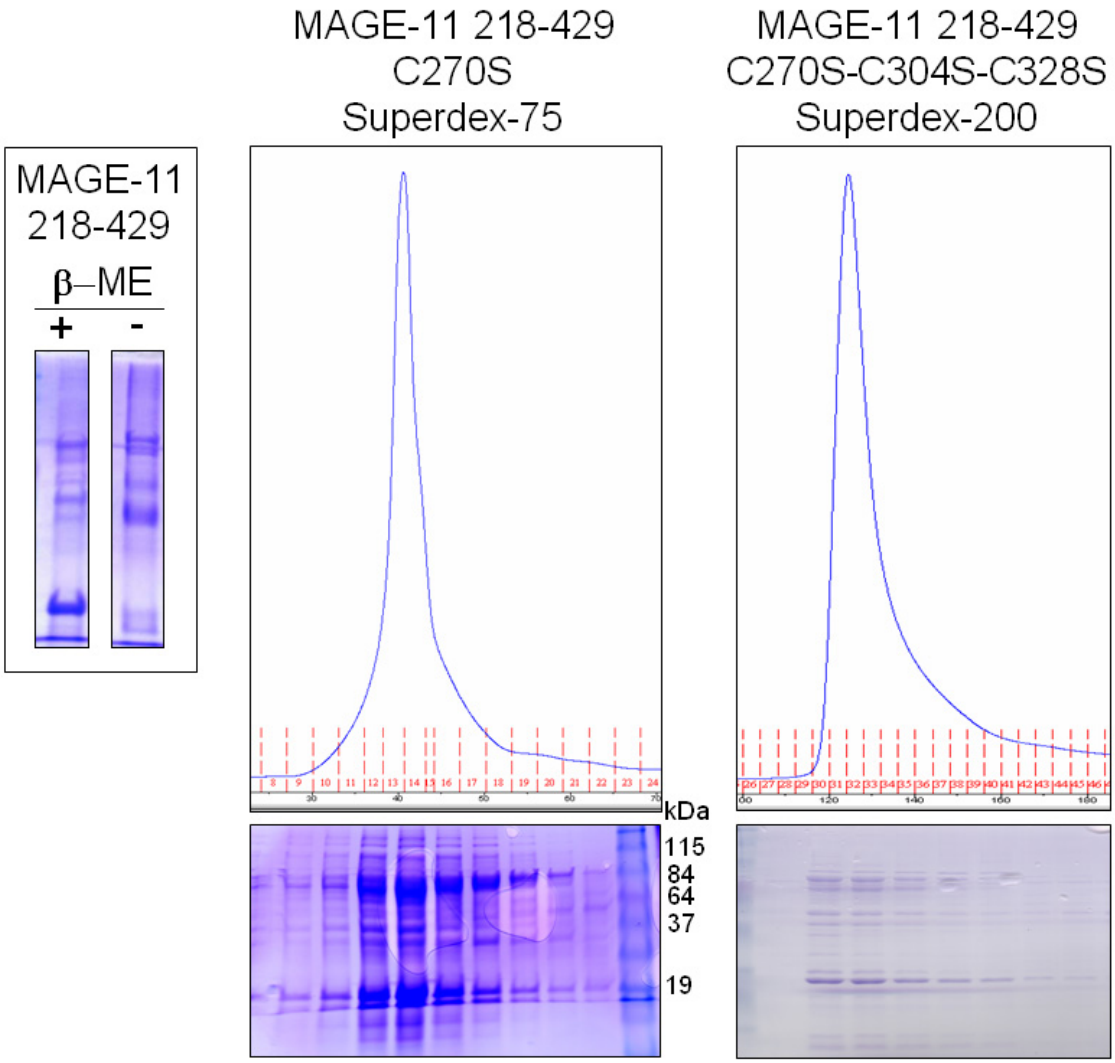

MAGE-11
218-429
β–ME
+
-
MAGE-11 218-429
C270S
Superdex-75
MAGE-11 218-429
C270S-C304S-C328S
Superdex-200
kDa
115
84
64
37
19

Figure 5.5

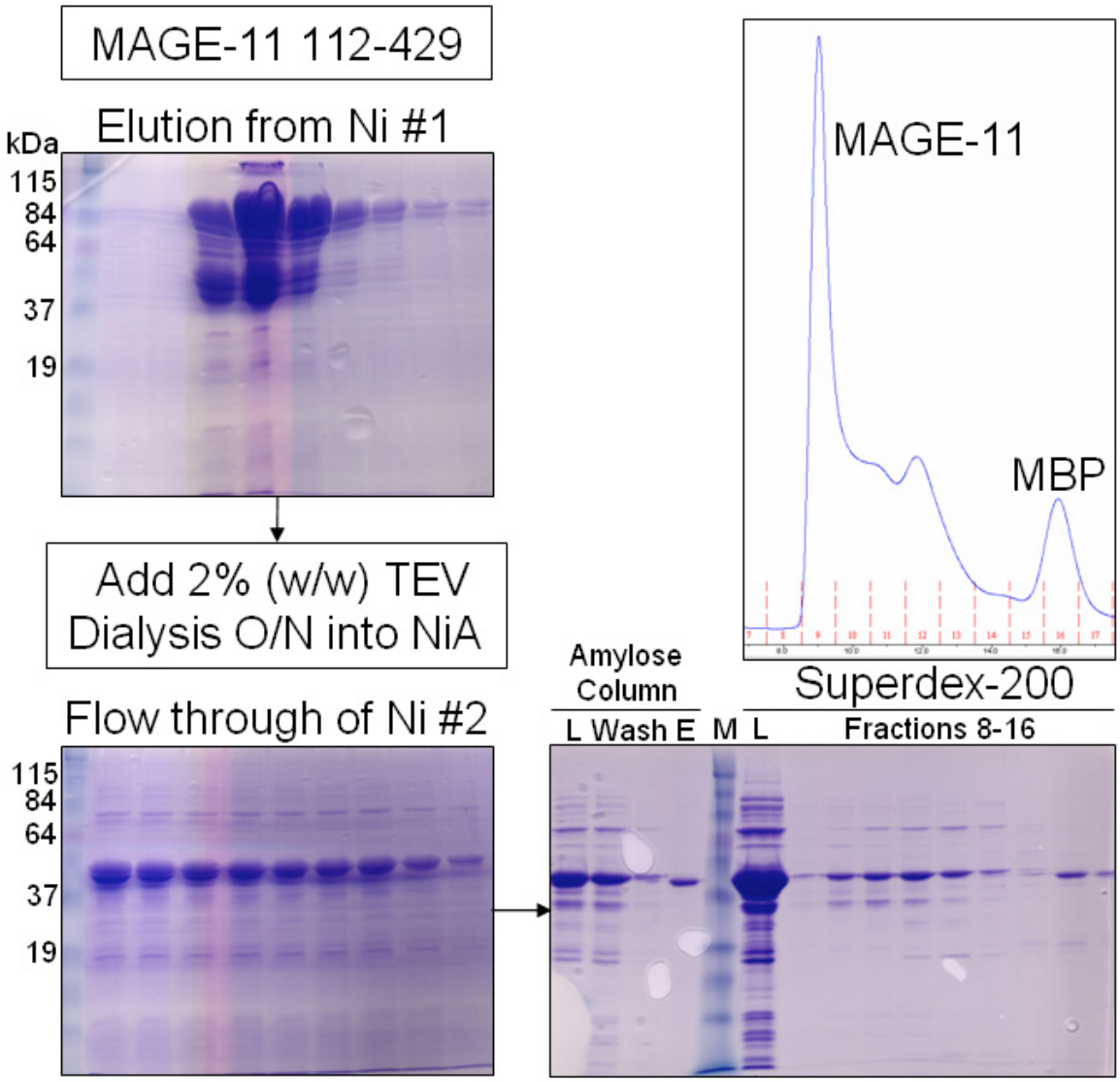

MAGE-11 112-429
Elution from Ni #1
kDa
115
84
64
37
19
Add 2% (w/w) TEV
Dialysis O/N into NiA
Flow through of Ni #2
115
84
64
37
19
MAGE-11
MBP
Amylose
Column
Superdex-200
L Wash E M L
Fractions 8-16

Figure 5.6

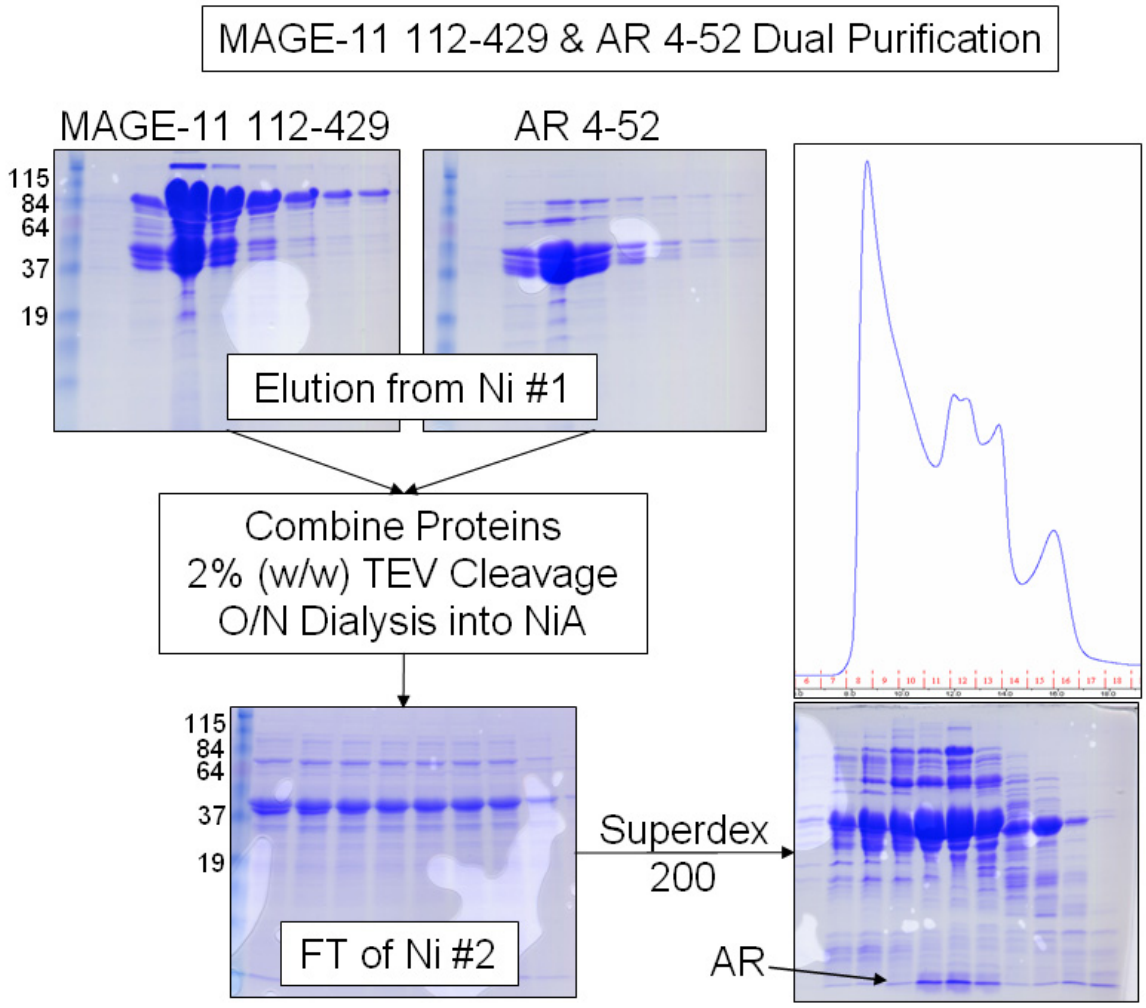

MAGE-11 112-429 & AR 4-52 Dual Purification
MAGE-11 112-429
AR 4-52
115
84
64
37
19
Elution from Ni #1
Combine Proteins
2% (w/w) TEV Cleavage
O/N Dialysis into NiA
115
84
64
37
19
FT of Ni #2
Superdex
200
AR

Figure 5.7

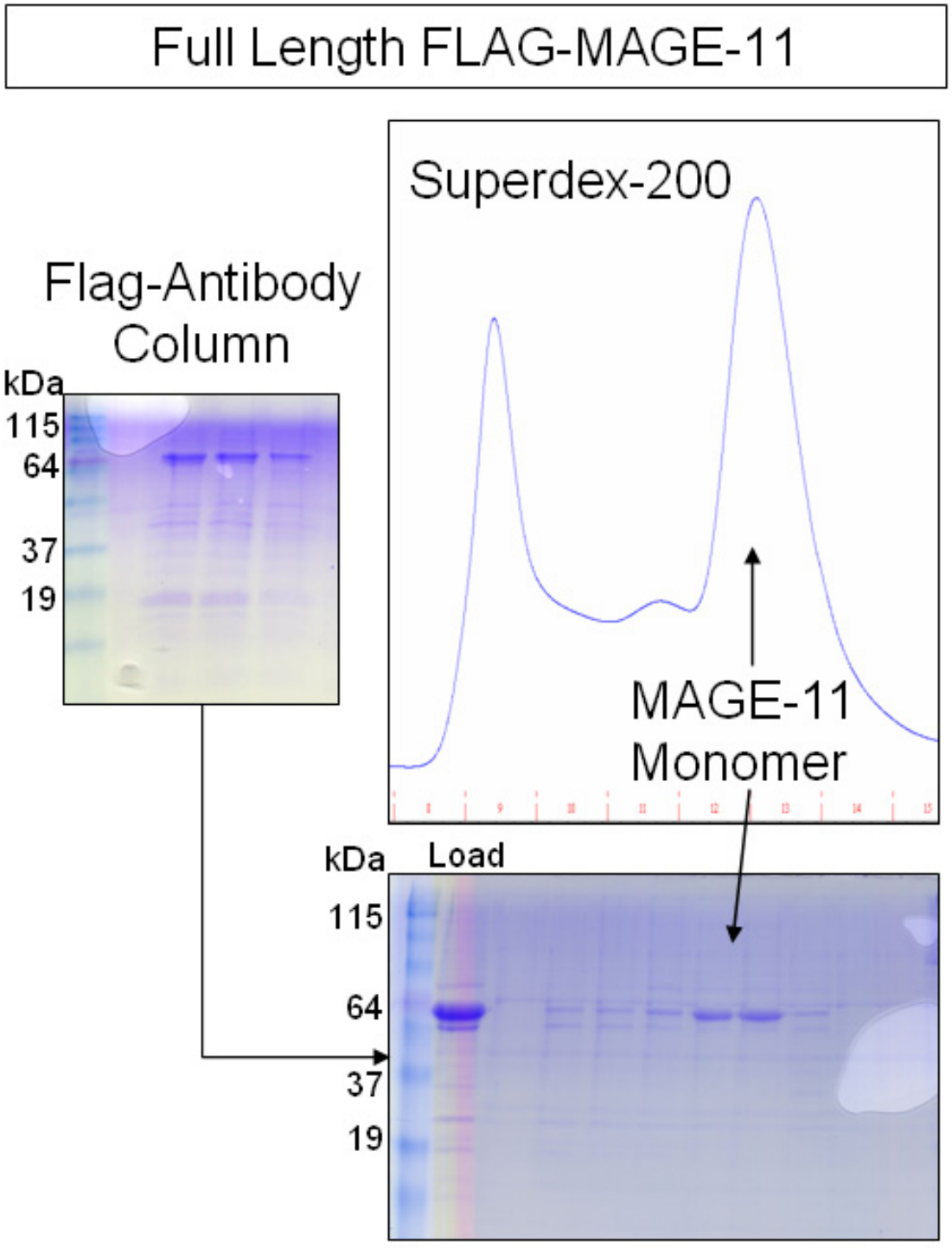

Full Length FLAG-MAGE-11
Superdex-200
Flag-Antibody Column
kDa
115
64
37
19
MAGE-11 Monomer
kDa
Load
115
64
37
19

Figure 5.8

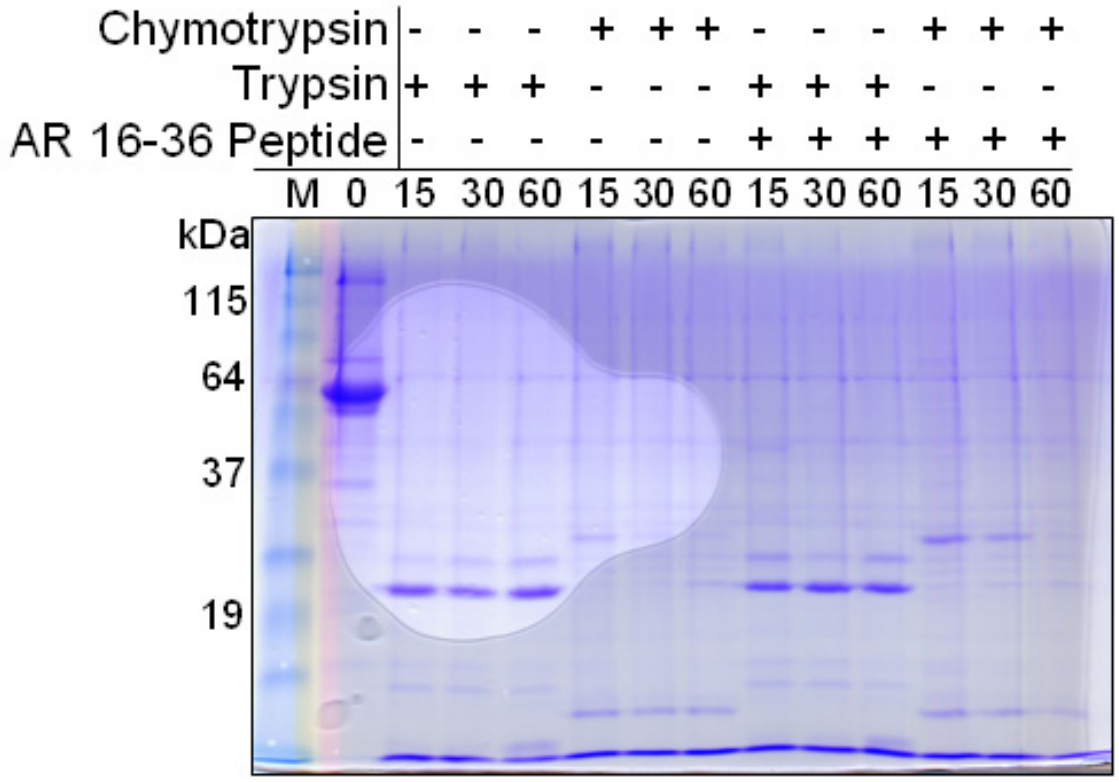

Chymotrypsin - - - + + + - - - + + +
Trypsin + + + - - - + + + - - -
AR 16-36 Peptide - - - - - - + + + + + +
M 0 15 30 60 15 30 60 15 30 60 15 30 60
kDa
115
64
37
19

Figure 5.9

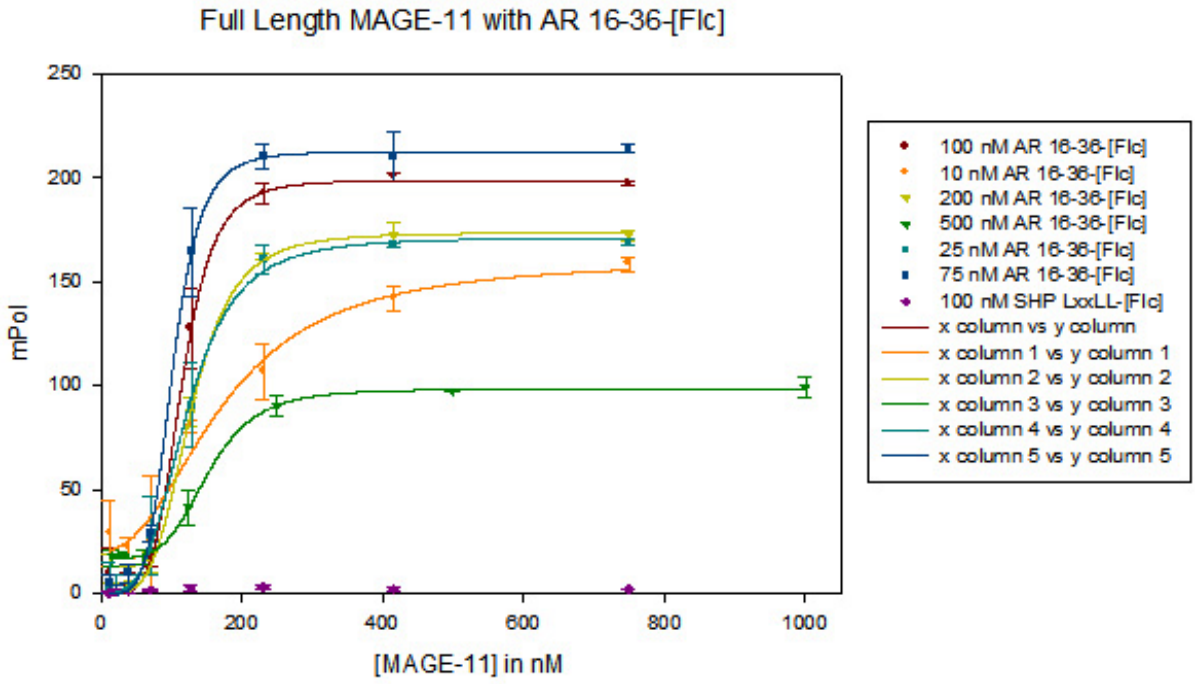

Full Length MAGE-11 with AR 16-36-[Flc]
mPol
[MAGE-11] in nM
0
50
100
150
200
250
0
200
400
600
800
1000
100 nM AR 16-36-[Flc]
10 nM AR 16-36-[Flc]
200 nM AR 16-36-[Flc]
500 nM AR 16-36-[Flc]
25 nM AR 16-36-[Flc]
75 nM AR 16-36-[Flc]
100 nM SHP LxxLL-[Flc]
x column vs y column
x column 1 vs y column 1
x column 2 vs y column 2
x column 3 vs y column 3
x column 4 vs y column 4
x column 5 vs y column 5

TABLE 5.1: Binding Constants of MAGE-11 and AR Determined by FP Assays

	[FxxLF] (nM)	K_D (nM)	Hill Constant	Averages
MAGE-11 Binding to AR 16-36-[Flc] Peptide	10	173	2.3	K_D = 131 nM
	25	121	3.5	
	75	101	5.2	
	100	115	5.1	Hill = 4.2
	200	128	4.2	
	500	152	4.5	
	LxxLL (100 nM)	NB	NB	NB

	Androgen	FxxLF K_D (μM)	LxxLL K_D (μM)
AR LBD Peptide Binding with Androgen	R1881	9.2 ± 0.4	78 ± 28
	Testosterone	5.5 ± 0.3	27 ± 4
	DHT	3.0 ± 0.4	13.1 ± 1.5