CADE, CHRISTINE ELIZABETH. Allosteric Inactivation of Caspase-3 Without Major Loop Rearrangements: Properties of the Inactive Ensemble. (Under the direction of Dr. A. Clay Clark.)

Allosteric regulation of caspase-3 can occur through binding of small molecules or through mutagenesis. The binding of drugs FICA and DICA to the dimer interface of active caspase-3 induces large conformational changes which convert the enzyme to a zymogen-like form. The V266H mutation in the dimer interface of caspase-3, on the other hand, inactivates the enzyme through subtle changes. Previously, E124A and Y197C mutations were introduced into WT and V266H caspase-3 to study the role of these residues in mediating inhibition of the enzyme. Based on a comparison of WT and V266H crystal structures, additional residues which may also be involved in inhibition were selected for the current study. These residues were located both near the dimer interface and in the active site.

One set of mutations in the context of the caspase-3 V266H enzyme attempts to restore activity to the V266H mutant by reducing steric clashes. It was determined that activity can be restored, but amount of activity regained is highest when the restorative mutation is closest to V266H. Control mutants containing only the restorative mutations without V266H were also produced. Another set of mutations were created which attempt to mimic the V266H mutant by introducing steric clashes. In general, these did have lower activity than WT, but a surprising result was seen with the T140F mutant, having only a 4-fold lower activity than WT despite introducing a bulky amino acid. Mutations were
introduced to probe the significance of the K137-E190 salt bridge, which was found to be transient and unnecessary for activity. The importance of hydrogen bonding between T140 and Y195 was also probed with a T140V mutation, which was found to have 25-fold lower activity than WT. Overall, 17 mutants were assayed for their activity, and x-ray data sets and MD simulations both with and without inhibitor were obtained for 16 of those mutants.
Allosteric Inactivation of Caspase-3 Without Major Loop Rearrangements: Properties of the Inactive Ensemble

by
Christine Elizabeth Cade

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Biochemistry

Raleigh, North Carolina

2013

APPROVED BY:

__________________________________________
Dr. Dennis Brown

__________________________________________
Dr. Robert Rose

__________________________________________
Dr. Robert M. Kelly

Chair of Advisory Committee

__________________________________________
Dr. A. Clay Clark
DEDICATION

I am dedicating my thesis to the man who is most dedicated to me: my husband, Lars. It is his strength and encouragement that has enabled me to make it this far. He left his family, friends, and job in Maryland to move down to North Carolina to marry me and support me in my studies. I love him more than I could possibly express.
BIOGRAPHY

Christine Elizabeth (Fairchild) Cade was born in Kalamazoo, MI. Her family moved to Virginia, Hawaii, Florida, and North Carolina with her dad who was in the Air Force. From 2001 to 2005 she studied chemistry at Cedarville University in Ohio, and upon graduation moved to Beijing, China, to teach English and chemistry to Chinese high school students. In 2006 she returned to the United States to pursue a graduate degree in chemistry at North Carolina State University. She met the love of her life, Lars Cade, and married him on August 8, 2009. They have a 1-year-old son Theodore (Teddy or T-Rex) and are expecting their second child in February. She obtained her Master’s of Science in Chemistry in December 2009 and continued her studies at NC State to receive her Doctorate in Biochemistry in December 2013.
ACKNOWLEDGEMENTS

I would like to thank my parents for giving me a love for science and for supporting me through all my schooling in many different ways. I would also like to thank my husband for all his support including moving to NC when we got married so I could continue my graduate studies. It shows me that you think very highly of me.

I would like to thank Dr. Robert Kelly for encouraging me to keep my fellowship and find a new doctoral program at NC State when my first one fell through. Just those few words of encouragement at the right time have made a huge difference in my life. I would like to thank Dr. A. Clay Clark for taking me on as a student when I thought my dreams of getting a PhD were over. Thank you for believing in me and for being a wise and caring advisor. Working in your lab was like a breath of fresh air.

I would like to thank my other committee members, Dr. Robert Rose and Dr. Dennis Brown. Thank you for your advice and for making time in your busy schedules to have meetings with me. I would like to thank my labmates for being good friends and for helping me out with everything from checking the OD of my bacteria to making sure my column didn’t run dry to bouncing ideas off each other. You are a huge part of what made working in the Clark lab so wonderful. I especially would like to thank Dr. Sarah MacKenzie for all of the advice, encouragement, and just lending a listening ear. Thanks for being my advocate and mentor.
TABLE OF CONTENTS

LIST OF TABLES........................................................................................................................................... ix
LIST OF FIGURES......................................................................................................................................... x
CHAPTER 1: OVERVIEW OF CASPASES........................................................................................................ 1
  I. Structure.................................................................................................................................................... 1
  II. Classification.......................................................................................................................................... 3
    A. Apoptotic Caspases.......................................................................................................................... 3
      1. Initiator Caspases.......................................................................................................................... 3
      2. Effector Caspases.......................................................................................................................... 5
    B. Inflammatory Caspases..................................................................................................................... 5
    C. Other or Unclassified Caspases........................................................................................................ 6
  III. Mechanisms.......................................................................................................................................... 6
    A. Activation........................................................................................................................................... 6
    B. Catalysis............................................................................................................................................ 8
  IV. Functions.............................................................................................................................................. 10
    A. Apoptosis.......................................................................................................................................... 10
    B. Inflammatory Response.................................................................................................................... 11
    C. Other Functions................................................................................................................................ 12
  V. Substrates and Inhibitors....................................................................................................................... 13
    A. Synthetic Substrates and Substrate Specificity............................................................................. 13
    B. Endogenous Substrates.................................................................................................................... 15
      1. Substrates Involved in the Apoptotic Phenotype......................................................................... 15
      2. Substrates Involved in Other Aspects of Apoptosis.................................................................... 15
    C. Synthetic Inhibitors.......................................................................................................................... 16
      1. Active Site Inhibitors...................................................................................................................... 16
      2. Allosteric Inhibitors....................................................................................................................... 17
    D. Endogenous Inhibitors..................................................................................................................... 19
REFERENCES................................................................................................................................................. 23
CHAPTER 2: OVERVIEW OF PROTEASES...................................................................................................... 39
  I. Classification and Structure..................................................................................................................... 40
    A. Serine Proteases.............................................................................................................................. 40
    B. Threonine Proteases.......................................................................................................................... 40
    C. Cysteine Proteases............................................................................................................................ 40
    D. Aspartate Proteases.......................................................................................................................... 41
    E. Metallo Proteases.............................................................................................................................. 41
  II. Mechanisms.......................................................................................................................................... 41
    A. Activation.......................................................................................................................................... 41
      1. Serine Proteases............................................................................................................................ 41
      2. Threonine Proteases....................................................................................................................... 43
CHAPTER 4: PHAGE DISPLAY TO FIND PEPTIDES WHICH BIND TO WT AND T140G CASPASE-3

I. Introduction.................................................................................................................. 159
II. Materials and Methods.................................................................................................. 160
  A. Preparation of DNA....................................................................................................... 161
  B. Small Scale DNA Electroporation.................................................................................. 161
  C. PCR Amplification of Positive Phage Samples............................................................... 162
  D. Large Scale DNA Electroporation.................................................................................. 162
  E. PEG Precipitation.......................................................................................................... 163
  F. Biotinylation of Target Protein....................................................................................... 164
  G. Determining Saturation Point for Biotinylated Protein.................................................. 165
  H. Selections..................................................................................................................... 166
  I. ELISAs of Phage Pools from Selections......................................................................... 167
  J. Preparation for Individual ELISAs................................................................................. 168
III. Results.......................................................................................................................... 169
  A. Synthesis of Covalently Closed Circular DNA.............................................................. 170
  B. Random Isolates for Assessment of Library Quality...................................................... 170
  C. Calculation of Diversity and Titer................................................................................ 171
  D. Plate Saturation Assay and Selections.......................................................................... 171
  E. Selection ELISAs.......................................................................................................... 172
  F. Individual ELISAs......................................................................................................... 172
IV. Discussion..................................................................................................................... 173
REFERENCES.................................................................................................................... 175
APPENDICES....................................................................................................................... 176

Appendix A: Designing Allosteric Activators of Caspase-3 Based on Maybridge-42 using Docking........................................................................................................ 213
  I. Introduction.................................................................................................................... 213
  II. Materials and Methods............................................................................................... 214
  III. Results......................................................................................................................... 215
REFERENCES.................................................................................................................... 216
LIST OF TABLES

CHAPTER 1

Table 1. Abbreviation Table.......................................................................................... 22

CHAPTER 3

Table 1. Highlighted mutants sorted by category and location........................................ 96
Table 2. Mutant Activity Data........................................................................................ 96
Table S1. Mutagenic Primers........................................................................................ 98
Table S2. RMSD for crystal structures aligned with WT............................................... 98
Table S3. Summary of Data Collection and Refinement Statistics............................... 99

CHAPTER 4

Table 1. Diversity and Titer of 4 Libraries..................................................................... 176
Table 2. Sequences of Individual Phage....................................................................... 177
LIST OF FIGURES

CHAPTER 1

Figure 1. Crystal Structure of Caspase-3................................................................. 34
Figure 2. Domain Arrangement of Mammalian Caspases........................................ 34
Figure 3. Two proposed mechanisms of caspase catalysis adapted from Miscione et. al 2010 ................................................................................................................... 35
Figure 4. Caspase substrate binding site adapted from Stennicke et. al 1999............ 38

CHAPTER 2

Figure 1. Variation in Folds of Serine Proteases...................................................... 60
Figure 2. 20S Proteasome from T. acidophilum: A Threonine Protease............... 61
Figure 3. Papain: A Cysteine Protease................................................................. 62
Figure 4. Pepsin: An Aspartate Protease............................................................... 63
Figure 5. Variation in Folds of Metallo Proteases................................................... 64
Figure 6. Activation Mechanism of Pepsinogen..................................................... 65
Figure 7. Catalytic Mechanism of Chymotrypsin, a Typical Serine Protease......... 66
Figure 8. Catalytic Mechanism of Papain, a Typical Cysteine Protease................. 69
Figure 9. Catalytic Mechanism of Pepsin, a Typical Aspartate Protease.............. 72
Figure 10. Catalytic Mechanism of Thermolysin, a Typical Metallo Protease....... 73

CHAPTER 3

Figure 1. Caspase Activation and Inactivation....................................................... 101
Figure 2. Comparison of WT (green residues) and V266H (cyan residues) structures highlighting featured mutated residues................................................................. 102
Figure 3. Activity of WT Caspase-3 and Mutants.................................................. 103
Figure 4. Return of Activity vs. Distance of Rescuing Mutation from V266H....... 104
Figure 5. Conformations of H266 in Crystal Structures and Molecular Dynamics of Double Mutants................................................................................................. 105
Figure 6. Stabilizing interactions and waters in interface mutants compared to WT and V266H................................................................................................................... 106
Figure 7. Comparison of F55 Mutants with WT Caspase-3................................... 109
Figure S1. Gel of Partially Processed Mutants....................................................... 110
Figure S2. Helix 3 of V266H and Restorative Mutants........................................... 111
Figure S3. Helix Destabilization in T140G/V266H and T140G.............................. 112
| Figure S4. Distances between Selected Residues over the Course of 50 ns Molecular Dynamics Simulations | 117 |
| Figure S5. Helix 3 at 0 and 50 ns for Mutants with and without Active Site Inhibitor | 139 |

**CHAPTER 4**

| Figure 1. Oligo Design | 178 |
| Figure 2. DNA Gel | 181 |
| Figure 3. Random Isolates for Diversity | 182 |
| Figure 4. Saturation Assay | 187 |
| Figure 5. Rounds 1-4 Selection ELISA | 188 |
| Figure 6. Rounds 5-6 Selection ELISA | 192 |
| Figure 7. Rounds 5-6 Selection ELISA Repeat | 196 |
| Figure 8. T140G + CMK D ELISA | 200 |
| Figure 9. WT - CMK D ELISA | 203 |
| Figure 10. WT + CMK D ELISA | 206 |
| Figure 11. T140G + CMK X-GG ELISA | 209 |

**APPENDIX A**

| Figure 1. Docking Results of Maybridge-42 and Select Compounds From JLS 1-12, CEC Compounds 13-95, and AD 1-63 with WT Caspase-3 | 217 |
| Figure 2. Compounds 58 and 58-OH | 223 |
| Figure 3. Dose-Response with Compounds 58 and 58-OH | 224 |

**APPENDIX B**

| Figure 1. Helix Structure | 231 |
| Figure 2. Typical Capping Box | 232 |
| Figure 3. Hydrophobic Staple in WT caspase-3 | 233 |
| Figure 4. Typical Schellman Motif | 234 |
| Figure 5. Typical $\alpha_L$ Motif | 235 |
| Figure 6. Destabilizing helix 3 by removing one residue of a hydrophobic staple | 236 |
| Figure 7. Destabilizing helix 3 by removing three residues of a hydrophobic staple | 237 |
| Figure 8. D135 is a better N-cap in V266H inactive conformation | 238 |
| Figure 9. Destabilizing the inactive conformation by removing a potential end cap | 239 |
| Figure 10. Destabilizing the inactive conformation by removing interactions with K138 | 240 |
CHAPTER 1

OVERVIEW OF CASPASES

Christine E. Cade and A. Clay Clark

Department of Molecular and Structural Biochemistry, North Carolina State University,
Raleigh, NC 27695-7622

To Be Published
Caspases (cysteinal aspartate-specific proteases) are enzymes which utilize a catalytic cysteine to cleave their peptide substrates after specific aspartate residues. The first caspase was discovered in 1992 and because of its function was named interleukin-1-β converting enzyme (ICE) but was later renamed to caspase-1. In 1993, Ced-3 from C. elegans was found to be homologous to ICE and the corresponding human protein CPP32 (later named caspase-3) was found in 1994. The official caspase nomenclature was decided on in 1996 to alleviate the confusion that went along with discovery of ten different caspases, some with multiple names.

I. Structure

Caspases are expressed as proenzymes (zymogens) called procaspases, which then become activated to the mature caspase form. Procaspase structure can be divided into three domains: an N-terminal prodomain, a large subunit, and a small subunit. The first step in maturation is dimerization. Then, proteolytic processing removes the prodomain and cleaves a loop called the intersubunit linker between the large and small subunits.

The secondary structure of mature caspases consists of six core β-strands in a slightly twisted sheet in each monomer, with two main helices on one face (the “front”) of the protein and three helices on the other face (the “back”) of the protein (Figure 1). The first four core β-strands and helices 1-3 form the large subunit, whereas the last two core β-strands and helices 4-5 form the small subunit.
The dimer interface consists of the final $\beta$-strand from each monomer, side-by-side in an antiparallel manner. The two monomers are related through a C2 axis of symmetry such that one monomer is “upside-down” compared to the other monomer.

Five loops are important for the formation of the active site. Once the intersubunit linker is cleaved, the two halves of the cleaved linker are called L2 and L2’. Active site loops L1, L2, L3, and L4 come from one monomer, and loop L2’ comes from the other. The catalytic cysteine is part of loop L2, and the catalytic histidine is part of a loop extending from the C terminal end of $\beta3$.

II. Classification

Caspases are divided into two main categories based on their function: apoptotic caspases and inflammatory caspases. The apoptotic caspases are further divided into two categories based on time of entry into the apoptotic cascade: initiator caspases and effector caspases.

A. Apoptotic Caspases

1. Initiator Caspases

Initiator caspases are stable monomers in the cell until they are activated by dimerization. Once dimerized, initiator caspases have sufficient activity to autoprocess, cleaving their prodomain and intersubunit linker. An induced proximity model for dimerization was first invoked for caspases -8 and -10 but now seems to be generalizable to initiator caspases as a whole. This model says that activation complexes increase the local concentration of the initiator caspases, enabling them to dimerize. The prodomains of
initiator caspases contain either a CARD (caspase activation and recruitment domain) or DED (death effector domain), which allow initiator caspases to bind to activation complexes (Figure 2).

The initiator caspases -2 and -9 are involved in the intrinsic pathway, which is activated by mitochondrial damage, cytotoxic stress, chemotherapeutic drugs or certain developmental cues. Activation of caspase-2 leads to release of cytochrome c from the mitochondria, which then binds to Apaf-1 and forms the heptameric apoptosome. The apoptosome binds procaspase-9 to dimerize and therefore activate it. Once active, caspase-9 activates downstream effector caspases.

The initiator caspases -8 and -10 are activated by the extrinsic pathway: in order to eliminate excess cells created during development or remove cells with tumorigenic properties, a molecule binds to a death receptor at the membrane which is part of the tumor necrosis factor receptor (TNFR) superfamily. One such ligand/receptor pair is FasL (Fas ligand) and CD95(APO-1/Fas). The cytosolic death domains (DD) of the receptor recruit an adaptor molecule such as FADD (Fas-associated death domain), allowing the complex to bind initiator procaspases -8 or -10 to form a death-inducing signaling complex (DISC). Once the procaspases are part of the DISC, they are able to dimerize and therefore become active. The active caspase-8 or -10 then activates downstream effector caspases such as caspase-3.
2. Effector Caspases

The effector caspases -3, -6, and -7, are found as inactive dimers in the cell. They are activated once an initiator caspase cleaves their intersubunit linkers. Because they do not require death scaffolds for dimer formation\textsuperscript{10, 11}, their prodomains are short and lack the CARD and DED domains typical of initiator caspases. Their prodomains are, however, likely to be involved in targeting within the cell.\textsuperscript{12-15}

B. Inflammatory Caspases

Similarly to the initiator caspases, the inflammatory caspases -1, -4, -5, -11, -12, and -13 are activated by dimerization. Their prodomains contain a CARD which allows them to bind to activation complexes. Similarly to apoptosome formation, a multiprotein complex called the inflammasome consists of a NOD-like receptor such as NLRP1, an adaptor protein such as ASC (apoptosis-associated speck-like protein containing a CARD), and the inflammatory procaspase, particularly procaspase-1.\textsuperscript{16} In some cases, the procaspase can also be recruited to CARD domains in the receptor directly, without the aid of an adaptor molecule.\textsuperscript{17}

Once the inflammatory caspases become active, they are activators of cytokines through cleavage of their preforms. In monocytes and macrophages, caspase-1 activates interleukin-1β (IL-1β)\textsuperscript{3} and interleukin-18 (IL-18). These cytokines mediate innate immunity and inflammation.\textsuperscript{18}

The mouse caspase-11 is a homolog of human caspase-4.\textsuperscript{19} In humans, caspase-12 is generally truncated due to a premature stop codon, but in some people of African descent, a
read-through mutation causes expression of the full-length protein, causing increased risk of sepsis due to decreased inflammatory and immune response to endotoxins.\textsuperscript{20} Caspase-13 is a bovine ortholog of human caspase-4.\textsuperscript{21}

**C. Other or Unclassified Caspases**

Caspase-14 expression is restricted to epidermal keratinocytes and is involved in differentiation.\textsuperscript{22} Like the effector caspases, it has a short prodomain with no adaptor regions.

Several caspases are not yet classified: 15, 16, and 17.\textsuperscript{23} Caspase-15 is expressed in several mammalian species including pigs, dogs, and cattle.\textsuperscript{24} It contains a pyrin-like region in its prodomain similar to that found in zebrafish caspases caspy and caspy2.\textsuperscript{25} Caspase-16 is found in marsupials and placental mammals and contains a short prodomain with no adaptor regions.\textsuperscript{23} Caspase-17 is found in vertebrates except for marsupials and placental mammals and also does not contain adaptor regions in its prodomain. Caspase-18 is found in opossums and chickens and, like caspases -8 and -10, contains two DED regions in its prodomain, so it is likely also an initiator apoptotic caspase.\textsuperscript{23}

**III. Mechanisms**

**A. Activation**

Activation of caspases generally requires two events: they must be a dimer and the intersubunit linker must be cleaved. Removal of the prodomain is not necessary for activation; in fact, the prodomain may serve to stabilize the active enzyme.\textsuperscript{26}

After dimerization, cleavage of the intersubunit linker occurs first, followed by cleavage of the prodomain. Prior to cleavage, the intersubunit linker from one monomer occupies the
dimer interface. Upon cleavage of the intersubunit linker, the C-terminal portion of the linker, L2’, vacates the central cavity and rotates about 180 degrees toward the active site, forming contacts with L2, L3, and L4 from the opposite monomer. These loop bundle contacts stabilize the active site. The movement of L2’ out of the dimer interface allows L3 to slide in towards the interface and form the substrate binding pocket. Rotation of a key arginine on L2 from a solvent-exposed position into the interface allows its neighboring residue, the catalytic cysteine, to assume its proper position for catalysis.

For effector caspases, equilibrium favors the inactive dimer. For initiator caspases, however, dimerization is the main challenge to be overcome for activation. Addition of kosmotropes such as sodium citrate causes caspase-8 to dimerize and become activated. At least partly because the initiator caspases have longer intersubunit linkers than effector caspases, cleavage of the intersubunit linker is not necessary for activation, but rather, stabilizes the active conformation.

Effector caspase mutants, particularly procaspase-3 V266E, can also be activated without cleavage of the intersubunit linker. This mutant is even more effective at inducing apoptosis than the wild-type (WT) enzyme. The enhancement of activation caused by the mutation is predicted to occur because the mutation keeps the intersubunit linker from binding to the dimer interface. In general, when the intersubunit linker is in the dimer interface, the protein is inactive, whereas when it is out of the interface it can become active. The conformational ensemble of effector procaspases includes both active and inactive conformers. Although the inactive ensemble is favored, binding of allosteric activators could
shift the equilibrium to the active ensemble. On the other hand, binding of allosteric inhibitors to the active caspase could inactivate it. Manipulating the position of the intersubunit linker could lead to allosteric activation or inhibition. A drug which binds at the dimer interface and holds the intersubunit linker in place could inactivate the enzyme. Conversely, a drug which binds at the dimer interface and keeps the intersubunit linker from binding could activate the procaspase. In fact, a small molecule has been suggested to activate procaspase-3 by this mechanism.\textsuperscript{30}

Additionally, Wells and coworkers have found a small molecule termed 1541 which forms nanofibrils that act as a scaffold for (pro)caspase-3 binding and increase activation of the procaspase.\textsuperscript{31} They suggest that the procaspase is activated through induced proximity, similar to the activation of initiator caspases. \textit{In vitro}, amyloid-β (residues 1-40) fibrils were also shown to activate procaspase-3. The activation of caspases by fibrils may play a role in neurodegenerative diseases.\textsuperscript{32}

**B. Catalysis**

Proteases all have some mechanistic features in common. The trigonal planar peptide bond of the substrate is forced into a tetrahedral intermediate.\textsuperscript{33} As this tetrahedral intermediate forms, a nucleophile attacks the carbonyl carbon of the peptide bond. Then, the amino nitrogen of the leaving group is protonated. Caspases contain a catalytic dyad consisting of a cysteine and a histidine.\textsuperscript{33} Based on the catalytic mechanism accepted for cysteine proteases, the mechanism for caspases has been thought to be as follows (See Figure 3A): First, the catalytic histidine abstracts a proton from
the catalytic cysteine. The catalytic cysteine acts as the nucleophile to form a covalent tetrahedral intermediate with the peptide substrate. Once the cysteine has bound, the histidine donates the proton to the amino moiety of the peptide leaving group. The peptide bond is cleaved, with the N-terminal part of the peptide remaining covalently attached to the cysteine while the C-terminal part of the peptide leaves. Finally, hydrolysis frees the N-terminal part of the peptide and re-protonates the catalytic histidine.

An oxanion hole, a pocket in the enzyme that hydrogen bonds to the carbonyl oxygen of the substrate, is also thought to be key for catalysis.\textsuperscript{33} It is formed by the backbone nitrogens of a conserved glycine (238 in caspase-1) and the catalytic cysteine (285 in caspase-1). The oxanion hole is thought to be important for polarizing and stabilizing the scissile carbonyl group.\textsuperscript{34}

However, there are some problems with the proposed mechanism. The 6-7 Å distance between the two catalytic residues is larger than found in most proteases, and makes direct hydrogen transfer unlikely.\textsuperscript{33} Molecular dynamics simulations have shown that the catalytic residues cannot exist as a charged pair prior to catalysis.\textsuperscript{35} Therefore, the deprotonation of the cysteine likely occurs during catalysis. Also, the histidine residue is not in an optimal location for protonating the amino leaving group.\textsuperscript{36}

A DFT study of the first part of the catalytic process (Figure 3B, part 1) has been carried out for caspase-7.\textsuperscript{34} Miscione and coworkers found that first, a proton is transferred from the backbone nitrogen of the P1 aspartate to the carboxylate group of the P1 aspartate. In the second step, a proton is transferred from the aspartate to a water molecule, and from
that water to the catalytic histidine. In the third step, a proton is transferred from the catalytic cysteine to the backbone nitrogen of the P1 aspartate. The overall result of these first three steps is the protonation of the catalytic histidine and the deprotonation of the catalytic cysteine. In a fourth step, the catalytic cysteine nucleophile attacks the carbonyl carbon of the substrate to form a tetrahedral intermediate, the peptide bond is cleaved, and a proton is transferred from the catalytic histidine to a second water, which transfers a proton to the amino nitrogen of the leaving group.

A QM/MM simulation focused on the hydrolysis of the covalent adduct (Figure 3B, part 2). In the reaction scheme proposed by Sulpizi and coworkers, the catalytic histidine deprotonates a water molecule, which attacks the scissile carbonyl carbon (as in the original proposed mechanism). Then the proton from the catalytic histidine is abstracted by the now negatively-charged carbonyl oxygen, such that a diol is formed. A second water molecule interacts with the catalytic histidine and one of the diol hydroxy groups. Finally, a proton is transferred from that diol hydroxyl group to the P1 aspartate residue, causing cleavage of the covalent adduct. If this is true, it could more cogently explain the specificity for a P1 aspartate residue.

IV. Functions

A. Apoptosis

The activation of caspases commits the cell to apoptosis. The main hallmarks of apoptosis include rounding of cells and retraction from neighbors, membrane blebbing to form vesicles called apoptotic bodies, nuclear fragmentation, chromatin condensation,
hydrolysis of genomic DNA to approximately 200 bp fragments, and translocation of phosphatidylserine (PS) to the external surface of cells as an “eat me” signal to phagocytes. The apoptotic caspases are necessary for conferring all of these phenotypes.

In addition to the systematic dismantling of the cell, caspases are also involved in producing “find-me” signals to cause chemotaxis of phagocytes to apoptotic cells. The recruitment of phagocytes keeps cells from releasing their contents into extracellular space and activating an immune response which could be harmful to the tissue.

When the number of apoptotic cells is too great for consumption by phagocytes, secondary necrosis can occur. When this happens, the cell releases its contents into extracellular space. However, immune cells are somehow able to recognize the cells undergoing apoptosis (and secondary necrosis) differently from necrotic cells. This is likely due to the actions of caspases. Caspases keep danger-associated molecular patterns (DAMPs) and alarmins from being activated. This can be thought of as a “tolerate me” signal.

Caspases are also involved in turning off transcription and translation. This keeps any infecting viral particles from replicating using the host’s machinery. They also fragment the Golgi, ER, and mitochondria.

B. Inflammatory Response

In contrast to the actions of apoptotic caspases, which systematically dismantle the cell to avoid an immune response, the actions of inflammatory caspases lead to cell lysis and activation of the immune response in a process called pyroptosis. In order to activate an immune response, caspases cleave cytokine IL-1β and IL-18 to produce the mature form.
In addition to activation of cytokines, procaspase-1 is also able to activate the pro-inflammatory transcription factor NF-κB. Rather than using its catalytic activity, the CARD domain of procaspase-1 binds to a CARD domain in the kinase RIP2, which is involved in NF-κB activation.

C. Other Functions

Caspase expression is kept below a certain threshold required for apoptosis by IAPs (inhibitor of apoptosis proteins). At these subthreshold levels they are able to play roles that are neither apoptotic nor inflammatory. Caspase-3 activity was found to be important for differentiation of erythroblasts, skeletal muscle, bone marrow stromal stem cells, and neural stem cells.

Caspase-3 has several other non-apoptotic functions in nerve cells. In addition to neural cell differentiation, caspase-3 has also been implicated in neuronal migration and plasticity, axon pruning, and synapse elimination.

Caspases have been shown to play a role in cell migration and invasion under certain circumstances. They can also induce neighboring cells to proliferate to replace dying cells in a process called apoptosis-induced proliferation. These roles for caspases have implications for cancer: moderate activation of caspases could, in fact, cause cancer to progress rather than regress.

In addition to its apoptotic function, caspase-8 has an anti-apoptotic function when it forms a heterodimer with FLIP$_L$ (a protein similar to caspase-8 but lacking a catalytic site). This protein complex is able to activate the NF-κB signaling pathway leading to
proliferation.\textsuperscript{57} In another pro-survival capacity, the caspase-8/FLIP\textsubscript{L} complex is also able to inhibit RIPK3-dependent necrosis.\textsuperscript{56}

V. Substrates and Inhibitors

A. Synthetic Substrates and Substrate Specificity

Caspase substrate specificity is determined by a series of 4-5 substrate residues which bind to the active site of the caspase. These residues are named P1-P4 or P5, with P1 always being an aspartate residue (Figure 4). The P4 residue is especially important in determining specificity for a given caspase.\textsuperscript{58}

Because of this 4-5 residue contribution to specificity, substrates used for measuring caspase activity typically have a tetrapeptide preceded by an acetyl group (Ac) on the N terminus and followed by a fluorophore on the C terminus: for example, Ac-DEVD-AFC. When the peptide is cleaved by the caspase, the fluorophore is released and activity can be determined by fluorescence. Some typical fluorophores include AMC (7-aminomethylcoumarin) and AFC (7-amino-4-trifluoromethylcoumarin). Addition of p-nitroanilide (pNA) instead of a fluorophore to the C-terminus allows caspase activity to be determined colorimetrically.

A positional scanning combinatorial library approach has been used with these synthetic substrates to determine the substrate specificity for most of the mammalian caspases.\textsuperscript{58, 59} Caspases-3 and -7 share the same substrate specificity: DEVD. The optimal sequence for caspase-1 is WEHD, and the optimal sequence for both caspase-4 and caspase-5 is (W/L)EHD. The optimal sequence for caspase-2 is DEHD, for caspase-6 is VEHD, for
caspase-9 is LEHD, for caspase-8 is LETD, and for caspase-10 is LE(Nle)D (Nle = norleucine).

The P1-P4 residues fit into the S1-S4 pockets in the active site of the caspase. The S1 pocket, consisting of R179, R341, and Q283 (caspase-1 numbering), is nearly 100% conserved; its basicity and its size make it ideally suited for binding an aspartate residue. The S2 pocket of caspases-3 and -7 is formed by aromatic residues and accommodates small aliphatic amino acids. A substitution of a valine or alanine in place of a tyrosine opens up the S2 subsite to larger residues in the initiator and inflammatory caspases. The S3 pocket consists of main-chain interactions with R341 (caspase-1 numbering). In caspases-8, and -9, nearby basic residues enhance the binding of glutamic acid residues to the S3 subsite.

The S4 subsite of inflammatory caspases is long, shallow, and hydrophobic, accommodating bulky aromatic side chains such as a tryptophan. On the other hand, in apoptotic caspases, a tryptophan (214 in caspase-3) reduces the size of the subsite, causing a preference for an aspartate or a small aliphatic residue in the S4 pocket. An asparagine in caspases-2 and -3 or a glutamine in caspase-7 enhances interaction with a P4 aspartate. Caspase-2 requires a P5 residue to occupy a S5 subsite. The reason for this specificity may be that binding of a small hydrophobic residue to this subsite may enhance the burial of a P4 aspartate.
B. Endogenous substrates

To date more than 700 substrates of caspases have been catalogued. A searchable database can be found at http://bioinf.gen.tcd.ie/casbah/. Caspase substrates are involved in conferring an apoptotic phenotype to cells. They are also involved in producing “find-me” and “tolerate-me” signals during apoptosis.

1. Substrates involved in the apoptotic phenotype

The following are some of the substrates of caspases which are involved in producing the apoptotic phenotype. The rounding of cells is likely in part due to caspase cleavage of components of actin microfilaments and microtubular proteins. Retraction of cells from their neighbors likely facilitates phagocytosis and is caused in part by caspase cleavage of components of focal adhesion sites, components of cell-cell adherens junctions, cadherins, and desmosome-associated proteins. Caspase cleavage of Rho effector ROCK1 which regulates movement of the actin cytoskeleton is a factor in blebbing and nuclear fragmentation. Nuclear fragmentation also involves caspase cleavage of lamins A, B, and C. Chromatin condensation is caused by caspase cleavage of Mst1 kinase. Hydrolysis of genomic DNA to small fragments is carried out by caspase-activated DNase (CAD/DFF). Translocation of PS to the external surface of the cell is also caspase-dependent, but not fully understood.

2. Substrates involved in other aspects of apoptosis

Caspases are also involved in producing “find-me” signals to cause chemotaxis of phagocytes to apoptotic cells. Caspase-3 cleaves calcium-independent phospholipase A2,
causing phosphatidylcholine in the membrane to become hydrolyzed to produce lysophatidylcholine (LPC).\textsuperscript{38} The C-terminal fragment of endothelial monocyte-activating polypeptide II (EMAPII) is produced by caspase-dependent proteolysis and acts as a “find-me” signal to attract monocytes.\textsuperscript{39} Caspase-dependent cleavage of the membrane channel pannexin-1 causes release of modest amounts of ATP, which may also act as a “find-me” signal.\textsuperscript{40}

Caspases also function to keep danger-associated molecular patterns (DAMPs) and alarmins from being activated. This function can be thought of as a “tolerate-me” signal, and is important for avoiding autoimmunity.\textsuperscript{41} As mentioned above, caspase activation leads to hydrolysis of genomic DNA (which acts as a DAMP) into short fragments.\textsuperscript{68} Additionally, the alarmin IL-33 is inactivated by caspase-3/-7-dependent proteolysis.\textsuperscript{70}

C. Synthetic Inhibitors

1. Active Site Inhibitors

Active-site inhibitors bind in the place of substrate and are therefore competitive inhibitors. These inhibitors can be peptidic or nonpeptidic and can bind reversibly or irreversibly.

Peptidic inhibitors can have as few as one amino acid (for example, Boc-Asp-FMK), but typically have four (for example, Ac-DEVD-FMK).\textsuperscript{71} Peptides linked to leaving groups such as halomethylketones [for example, chloromethylketone (CMK) and fluoromethylketone (FMK)], acylomethylketones, and (phosphinoxy) methyl ketones bind irreversibly, whereas peptides linked to non-leaving groups such as aldehyde (CHO) bind
reversibly. The electrophilic carbonyl of the aldehyde or ketone binds to the catalytic cysteine, inhibiting it.

Several different peptidomimetics have been designed as inhibitors for caspases. These include urazolopyrazine-based β-strand peptidomimetics designed as inhibitors for caspase-3 and caspase-8,\textsuperscript{72} hydantoin-based peptidomimetics as inhibitors of caspase-3,\textsuperscript{72} dipeptidyl aspartyl fluormethylketones with unnatural amino acids,\textsuperscript{74} 1-(2-acylhydrazinocarbonyl)-cycloalkyl carboxamides,\textsuperscript{75} 8,5-fused bicyclic compounds,\textsuperscript{76} and peptidomimetics containing a caprolactam ring.\textsuperscript{77} Non-peptide inhibitors have also been discovered. These include isatins,\textsuperscript{78, 79} indole aspartyl ketones, fuchsone derivatives, and pyrrolo[3,4-c]quinolone-1,3-diones.\textsuperscript{80}

2. Allosteric Inhibitors

Caspases -3 and -7 were found to contain an allosteric site at the dimer interface.\textsuperscript{81} The drugs FICA and DICA form disulfide bonds with cysteines in the dimer interface of those caspases and inactivate the protein. The structural changes brought about by binding of these drugs involves massive loop rearrangements to a structure very similar to that of the proenzyme.

Mutation of valine 266 to a histidine at the dimer interface of caspase-3 also caused allosteric inactivation of the protein\textsuperscript{28}; however, the structural changes brought about by the mutation were much more subtle than those that occurred upon binding of FICA or DICA.\textsuperscript{82} Instead of conversion to a structure like that of the proenzyme, inactivation may be caused by a series of steric clashes, disordering of loop L1, and/or destabilization of helix 3.
A drug called compound 34 was found to bind to cysteines near the dimer interface of caspase-1.\textsuperscript{53} Similarly to the binding of FICA and DICA, the inactive structure was like that of the proenzyme.

Another set of allosteric inhibitors was found to inhibit caspases -3, -7, -8, and -9.\textsuperscript{54} A crystal structure with caspase-7 indicates that one and likely all of these compounds binds to the dimer interface. One of the compounds, Comp-A, inhibits dimerization of caspase-8; however, caspase-7 remained a dimer upon binding of the drug. As with FICA and DICA inhibition, the inhibited form was similar to that of the zymogen. However, these new compounds are reversible inhibitors, unlike FICA and DICA.

One of the urazole ring peptomimetic inhibitors which bind at the active site was also found to bind near the dimer interface of caspase-8.\textsuperscript{72} Some of the interacting residues of caspase-8 are Tyr334, Thr337, Glu396, and Phe399.

Caspase-2 was allosterically inhibited through binding of a designed ankyrin repeat protein (DARPin).\textsuperscript{85} Binding causes the caspase to be fixed in an inactive conformation different from that of the proenzyme.

A novel allosteric site was found on caspase-6.\textsuperscript{86} Phage display produced a peptide pep419 which binds near helix 2 and causes tetramerization and therefore inactivation of caspase-6. Interestingly, it was found that at pH 8, the zymogen of caspase-6 is a tetramer in solution, whereas at pH 5.5, the zymogen is a dimer, but can be induced to form a tetramer through the binding of pep419 or a related peptide pep420. The pH changes in the cell...
brought about by apoptosis could potentially lead to dissociation of caspase-6 tetramers to the dimeric form, leading to activation of the protein.

**D. Endogenous Inhibitors**

Both viral and endogenous inhibitors can block caspase activity by competing for binding to activation complexes. Viral inhibitors target caspase activity of their host cells in order to counter an immune response. Several γ-herpesviruses and molluscipoxvirus use v-FLIPs to block caspase access to the DISC. Similarly, endogenous FLIPs blocks procaspase-8 recruitment to DISC, and ICEBERG blocks caspase-1 recruitment to form the inflammasome.

Most protease inhibitors bind to the protease and block substrate access. Suicide inhibitors are cleaved and cause a conformational change to occur in either the inhibitor or the protease. Although it is typically a serine protease inhibitor, the serpin CrmA is also able to inhibit caspsases -1, -8, and -9, likely by forming a covalent attachment with the caspase and undergoing a conformational change upon cleavage of the scissile P1-P1’ bond to place the caspase on the “bottom” of the inhibitor. Similarly, the baculovirus protein p35 becomes covalently attached to the catalytic cysteine, the scissile bond is cleaved, but the protein is not liberated because it blocks the hydrolytic water from gaining access to the active site.

Another category of inhibitors are IAPs (inhibitor of apoptosis proteins). They were first discovered using baculovirus lacking a functional p35 gene. They contain a 70-80
residue Zn$^{2+}$ binding module named BIR. The most well-studied is X-linked inhibitor of apoptosis protein (XIAP).$^{96}$

XIAP targets caspases in two different ways. A linker to the BIR1 domain and The BIR2 domain of XIAP target effector caspases-3 and -7.$^{97}$ Residues in the active site, particularly in loop L1 make critical contacts with the inhibitor. Loop L2’ also makes contacts with XIAP. The necessity of ordered active site loops and cleaved intersubunit linker to form L2’ mean that XIAP only binds the active caspase rather than the inactive procaspase.

Unlike XIAP binding to the effector caspases, BIR3 and RING of XIAP target initiator caspase-9.$^{97, 98}$ Also, instead of binding to the active site, it binds to the dimer interface of the monomer and blocks dimer formation. Loop L2’ of caspase-9 binds BIR3 in a similar manner to how loop L2’ of caspase-3 binds BIR2. The pocket where loop L2’ of caspase-9 binds BIR3 is called the Smac (second mitochondrial activator of caspases) pocket because Smac can also bind there to derepress caspase activation.

Caspase activity is also controlled endogenously through the use of post-translational modifications. The RING domain of XIAP acts as an E3 ubiquitin ligase to ubiquitylate effector caspases-3 and -7, leading to proteasomal degradation.$^{99, 100}$ Sumoylation of pro-caspase-2$^{101}$ and caspase-8$^{102}$ likely leads to localization of the protein in the nucleus. Phosphorylation is a third post-translational modification which affects caspase activity. p38-MAPK phosphorylates S150 of caspase-3, inhibiting it.$^{103}$ Phosphorylation by PKC-δ at an as yet unknown site, on the other hand, enhances caspase-3 activity.$^{104}$ PAK2 phosphorylates
caspase-7 at three sites, decreasing its activity.\textsuperscript{105} For caspase-9, ERK phosphorylates T125,\textsuperscript{106} c-Abl phosphorylates Y153,\textsuperscript{107} and Akt phosphorylates S196,\textsuperscript{108} leading to decreased activity of the protein.
### TABLES

#### Table 1: Abbreviation Table

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC</td>
<td>Apoptosis-associated speck-like protein containing a CARD</td>
</tr>
<tr>
<td>CARD</td>
<td>Caspase activation and recruitment domain</td>
</tr>
<tr>
<td>Caspase</td>
<td>Cysteinal aspartate-specific protease</td>
</tr>
<tr>
<td>DAMPs</td>
<td>Danger-associated molecular patterns</td>
</tr>
<tr>
<td>DD</td>
<td>Death domain</td>
</tr>
<tr>
<td>DISC</td>
<td>Death inducing signaling complex</td>
</tr>
<tr>
<td>FADD</td>
<td>Fas-associated death domain</td>
</tr>
<tr>
<td>FasL</td>
<td>Fas ligand</td>
</tr>
<tr>
<td>FLICE</td>
<td>FADD-like interleukin 1β-converting enzyme</td>
</tr>
<tr>
<td>FLIP</td>
<td>FLICE-like inhibitory protein</td>
</tr>
<tr>
<td>FLIP&lt;sub&gt;L&lt;/sub&gt;</td>
<td>Long splice variant of FLIP which forms a heterodimer with caspase-8</td>
</tr>
<tr>
<td>FLIP&lt;sub&gt;S&lt;/sub&gt;</td>
<td>Short splice variant of FLIP which blocks caspase-8 from binding death receptor</td>
</tr>
<tr>
<td>ICE</td>
<td>Interleukin 1β-converting enzyme</td>
</tr>
<tr>
<td>IL-1β</td>
<td>Interleukin 1β</td>
</tr>
<tr>
<td>IL-18</td>
<td>Interleukin 18</td>
</tr>
<tr>
<td>PS</td>
<td>Phosphatidylserine</td>
</tr>
<tr>
<td>Smac</td>
<td>Second mitochondrial activator of caspases</td>
</tr>
<tr>
<td>TNFR</td>
<td>Tumor necrosis factor receptor</td>
</tr>
<tr>
<td>XIAP</td>
<td>X-linked inhibitor of apoptosis protein</td>
</tr>
</tbody>
</table>
REFERENCES


FIGURES

A. Pro caspase-3 Model  B. Caspase-3

Figure 1: Pro caspase-3 model and crystal structure of caspase-3
Active site loop coloring: yellow = L1, red = L2, cyan = L2', blue = L3, tan = L4

Figure 2: Domain arrangement of mammalian caspases
Figure 3: Two proposed mechanisms of caspase catalysis adapted from Miscione et. Al 2010
A: Typical cysteine protease mechanism

1. Formation of covalent adduct

2. Hydrolysis of covalent adduct
B: Simulated caspase mechanism

1. Formation of covalent adduct

2. Hydrolysis of covalent adduct
Figure 4: Caspase substrate binding site adapted from Stennicke et. al 1999
CHAPTER 2

OVERVIEW OF PROTEASES

Christine E. Cade and A. Clay Clark

Department of Molecular and Structural Biochemistry, North Carolina State University,
Raleigh, NC 27695-7622
I. Classification and Structure

Proteases are enzymes which catalyze the hydrolysis of peptide bonds. There are five major classes characterized by their active site residues: serine, threonine, cysteine, aspartate, and metallo proteases.¹ Within these five classes, the proteases can be further divided using the MEROPS protease classification system.² In this system, proteins are divided into clans and families.

A. Serine Proteases

Each of the clans of serine proteases has a different tertiary structure (Figure 1). Although the relative positions of the catalytic residues mostly remain constant, the order of those residues in the linear sequence differs. The S1A family of the PA clan of serine proteases is the largest family and contains many well-known enzymes including thrombin, trypsin, chymotrypsin, and Factor VIIa. The structure of both S1A and S1B proteases is arranged into two six-stranded beta-barrels with the active site in between.³ Some serine proteases contain regulatory domains in addition to the catalytic domain.⁴

B. Threonine Proteases

Threonine proteases have a completely different structure from serine proteases (Figure 2). They have a four-layered α-β-β-α structure.⁵ One famous member of the threonine proteases is the 20S proteasome.⁶

C. Cysteine Proteases

The most extensively studied cysteine protease, papain, has a globular structure with the active site found at the interface of two domains (Figure 3).⁷ Caspases, on the other hand,
are functional dimers and each monomer consists of six central roughly-planar \( \beta \)-sheets surrounded by five alpha helices (two on one face of the plane and three on the other). The active site is formed from several loops away from the dimer interface.

**D. Aspartate Proteases**

Like many cysteine proteases, the active site of aspartate proteases is found in a cleft between two similar domains (Figure 4).\(^8,^9\) Each of the domains contains one of the two catalytic aspartate residues.

**E. Metallo Proteases**

Metallo proteases contain either one or two metal ions, most commonly zinc, in the active site.\(^10\) Tertiary structure varies significantly between members of this type of protease (Figure 5). Similarly to the cysteine and aspartate proteases, single-metal ion metallo proteases contain two domains, with the active site formed in a cleft between the two domains. The active site of a clan of Zn\(^{2+}\)-dependent proteases called Metzincins contains three conserved histidines in a HEXXHXXGXXH motif as well as a conserved methionine.\(^11,^12\)

**II. Mechanisms**

**A. Activation**

**1. Serine Proteases**

The S1A family of serine proteases, for example, chymotrypsin, is activated through cleavage of a propeptide between residues 15 and 16.\(^13\) The newly-formed N-terminus creates an ion pair with a conserved aspartate (194 in chymotrypsin) which is part of the
substrate binding site and oxyanion hole. This ion pair formation induces a conformational change in the protein, causing it to become catalytically active. Although the catalytic residues are in the proper position in the zymogen, the substrate binding site is occluded and the oxyanion hole is not fully formed until the enzyme is activated. Additionally, autocatalytic cleavage releasing S14-R15 and T147-N148 occurs, but the rest of the prodomain (1-13) remains bound through a disulfide bond between C1 and C122.

Other proteases, on the other hand, such as tissue-type plasminogen activator, form an ion pair with the same aspartate using a lysine and do not require proteolytic cleavage for activation. In contrast, N-terminal cleavage of the serine protease complement factor D is insufficient to confer catalytic activity. The catalytic triad is not in proper position and the substrate binding pocket is narrow. Binding of its substrate factor B is necessary to induce conformational changes necessary to activate it.

Yet another method of activation is seen in the serine carboxypeptidase human protective protein (HPP, cathepsin A). In the zymogen, the catalytic residues, substrate binding pocket, and oxyanion hole are all in their proper positions. However, substrates are unable to access the active site because it is occluded by part of the protein. Cleavage of an internal (not N-terminal) activation segment causes conformational changes that free the active site for acceptance of substrates.
2. Threonine Proteases

The threonine proteases, like many serine proteases, are activated through N-terminal cleavage. However, the threonine proteases are capable of autoactivation. For the threonine proteases, removal of the propeptide reveals the catalytic threonine at the new N-terminus.

3. Cysteine Proteases

Papain-like cysteine proteases are activated through cleavage of an N-terminal propeptide. Part of the propeptide resides in the active site of the zymogen (though in the opposite direction relative to substrate binding), blocking substrate access. The zymogen is targeted to a lysosome, where the low pH facilitates autocatalysis.

Caspases are activated through dimerization and/or cleavage of a loop called the intersubunit linker. An N-terminal peptide is generally cleaved during activation; however, cleavage of the prodomain is unnecessary for activity.

4. Aspartate Proteases

Aspartate proteases are mostly expressed as zymogens with an N-terminal propeptide. The propeptide of pepsinogen, for example, contains 44 amino acids, and resides in the active site cleft between the domains of the protein. It is highly basic and forms contacts with carboxylate side chains in the cleft of pepsinogen. Lowering the pH protonates these carboxyl groups, releasing the propeptide for autocatalysis and activation of the protein.

Several intermediates form in the activation of the protein (Figure 6). First, upon lowering the pH, helical regions of the prodomain unravel and are pushed away from the active site. Second, cleavage occurs between F26p and L27p of the prodomain, followed
by cleavage between the prodomain and the mature enzyme (L43p and S1).\textsuperscript{26} The last 16 residues of the prodomain are removed, but the first 26 remain temporarily noncovalently attached. These residues form the first beta sheet on the back of the protein. Upon cleavage of the propeptide, the new N-terminus of the protease shifts away from its location near the active site. The N-terminus of the propeptide leaves, and the new N-terminus of the enzyme takes its place, changing conformation from a loop and $3_{10}$ helix to becoming the new first beta strand in the back of the protein.

5. Metallo Proteases

Pancreatic Zn$^{2+}$-carboxypeptidases contain an N-terminal inhibitory prosegment in their zymogen form. Although the active site is fully formed, the prosegment blocks substrate access and thus inactivates the enzyme.\textsuperscript{28}

Matrix metalloproteinases (MMPs) contain a conserved cysteine in a consensus motif (PRCGXPDV) of the prodomain.\textsuperscript{11} This cysteine interacts with the zinc in the catalytic domain, preventing the enzyme from becoming activated.\textsuperscript{29} Similarly to the papain-like cysteine proteases, the prodomain is oriented in the opposite direction compared to substrate binding. Disruption of this interaction followed by proteolytic cleavage C-terminal to the consensus motif leads to activation of the protein.

B. Catalysis

The mechanism of catalysis for each of the classes of protease shares some common features. First, a general base activates the nucleophile. Then, nucleophilic attack occurs either through a catalytic amino acid or through water. A tetrahedral intermediate is
stabilized. Finally, a general acid donates its proton to the amino nitrogen of the peptide leaving group, leading to decomposition of the tetrahedral intermediate to form products.

1. Serine Proteases

The first four clans discovered of serine proteases all contain a catalytic triad consisting of serine, histidine, and aspartate (S195, H56, D99 in chymotrypsin). This catalytic triad is often called the charge-relay system. Several novel serine proteases have since been discovered which contain Ser-Lys, Ser-Lys/His, His-Ser-His, Ser-His-Glu, and N-terminal Ser catalytic dyads or triads.

In the generally accepted mechanism, the catalytic histidine acts as a general base to activate the catalytic serine (Figure 7). The catalytic serine acts as a nucleophile to attack the carbonyl carbon of the scissile peptide bond, forming a tetrahedral intermediate. The negative charge which forms on the carbonyl oxygen is stabilized by main chain N-H groups of residues that make up the oxyanion hole. The C-terminal portion of the peptide which acts as the leaving group is released, taking with it a proton from the catalytic histidine, and leaving a trigonal-planar intermediate still attached to the catalytic serine. The catalytic histidine then removes a proton from water, which attacks the carbonyl carbon of the trigonal-planar intermediate and forms a new tetrahedral intermediate. The N-terminal portion of the peptide now acts as the leaving group, resetting the catalytic serine and histidine for another round of catalysis.
2. Threonine Proteases

Threonine proteases are generally considered to have a catalytic tetrad.\(^6,1^0\) This tetrad consists of the hydroxyl group of the catalytic threonine, the amino terminus of the catalytic threonine, and side chains of lysine 33 and glutamic acid 17 (S20 proteasome numbering).\(^6\)

Similarly to the serine proteases, the catalytic threonine of threonine proteases acts as a nucleophile. The amino group of the catalytic threonine and/or the side chain of lysine 33 abstract the proton from the hydroxyl group of the threonine. The backbone N-H group of glycine 47 provides oxyanion stabilization, similarly to the oxyanion hole seen in serine proteases.

3. Cysteine Proteases

Papain and most of its family members contain a catalytic triad consisting of cysteine, histidine, and asparagine. However, other members exist which only contain a cysteine/histidine catalytic dyad.\(^36,3^7\) The apoptotic caspases also contain only a catalytic dyad consisting of cysteine and histidine.

Similarly to serine and threonine proteases, the catalytic cysteine acts as the nucleophile (Figure 8).\(^3^8\) For papain, Cys 25 attacks the carbonyl carbon of the peptide, forming a tetrahedral intermediate. This intermediate is broken down by transfer of a proton from H159 to the leaving group. H159 then removes a proton from water, activating it to hydrolyze the acyl-enzyme intermediate.
4. Aspartate Proteases

Aspartate proteases contain two aspartate residues which are necessary for catalysis. The protonation state of these aspartate residues is important for activity; most of these proteins are active from pH 2-4, and are therefore sometimes called acid peptidases.

The reaction mechanism of aspartate proteases is different from that of other proteases in that the catalytic residue is not the attacking nucleophile and no covalent acyl intermediate is formed (Figure 9). Instead, water acts as the nucleophile. One of the aspartate residues is protonated and acts as a general acid and the other is ionized and acts as a general base.

The acidic aspartate (32 in pepsin) polarizes the carbonyl of the substrate by sharing a proton. The basic aspartate (215 in pepsin) abstracts a proton from the nucleophilic water, which then attacks the substrate, forming a tetrahedral diol intermediate. The proton is transferred from Asp 215 to the amine nitrogen of the leaving group and Asp 32 regains its proton from the hydroxyl group, breaking down the tetrahedral intermediate and releasing the products. Tyr 75 also likely plays a role in catalysis by stabilizing the tetrahedral intermediate.

5. Metallo Proteases

Similarly to that of the aspartate proteases, catalysis by metallo proteases proceeds by a noncovalent intermediate (Figure 10). For the single-metal ion protease thermolysin, zinc polarizes the scissile carbonyl of the substrate. Glutamate 143 acts as a general base to remove a proton from water which acts as a nucleophile. This proton then is donated to the
amino nitrogen of the leaving group. The side chains of His 231 and Y157 stabilize the oxyanion of the tetrahedral intermediate.

III. Functions

A. Serine Proteases

Serine proteases are involved in a variety of functions. Trypsin, chymotrypsin, and pancreatic elastase are digestive enzymes.\textsuperscript{10, 31} Tryptase, complement factors B, C, and D, complement component 2, mast cell protease, cathepsin G, and neutrophil elastase are involved in the immune response.\textsuperscript{31} Coagulation factors VIIa, IXa, Xa, and XIIa, as well as thrombin and protein c are involved in formation of blood clots. Urokinase, tissue plasminogen activator, plasmin, and kallikrein are involved in fibrinolysis. Serine proteases involved in reproduction include prostate specific antigen and acrosein. Subtilisin is secreted outside the cell to scavenge nutrients.

B. Threonine Proteases

The best understood threonine protease is the proteasome.\textsuperscript{10} This complex is involved in degrading proteins in the cell outside of lysosomes. Another threonine protease is gamma-glutamyltranspeptidase which is necessary for maintaining homeostasis of glutathione (GSH).\textsuperscript{39}

C. Cysteine Proteases

The plant enzyme papain is the most well-studied cysteine protease. It can cleave a variety of substrates and is used medically as a treatment for dental caries\textsuperscript{40} and sports
injuries. Mammalian members of the CA1 (papain-like) cysteine proteases are lysosomal enzymes. Caspases are involved predominantly in inflammation and apoptosis.

**D. Aspartate Proteases**

Aspartate proteases generally function best at low pH. Pepsin is a digestive enzyme which works in the stomach. Renin is involved in regulation of blood pressure.

**E. Metallo Proteases**

Thermolysin catalyzes hydrolysis of peptide bonds containing hydrophobic amino acids. It is used commercially in the production of aspartame. Carboxypeptidase A is a pancreatic enzyme. Metzincins including matrix metalloproteinases (MMPs) are involved in many processes including digestion of intake proteins; tissue development, maintenance, and remodeling; and activation of themselves or other zymogens.

**IV. Regulation**

**A. Production as zymogens**

One important level of regulation of proteases is their production as zymogens. Proteases are generally produced in an inactive form which becomes activated upon proteolytic cleavage of an “activation segment.” Often, this activation segment is the N-terminal portion of the protein and termed the “prosegment” or “prodomain.” This prodomain can be cleaved by external factors or autocatalytically upon a change in pH. For details, see the above section on activation mechanisms of proteases.

This regulatory mechanism ensures that the proteases are activated at the right place and the right time. For example, pepsin can be activated by acid in the stomach where it is
needed. The act of proteolytic cleavage to convert the zymogen to the active form is essentially irreversible, although binding of some inhibitors can revert the protease to a zymogen-like form.\textsuperscript{46}

\textbf{B. Segregation from substrates}

Cellular localization can also play a role in regulation of proteases. For example, the cellular localization of CtrA, a regulator of \textit{Caulobacter} cell cycle and substrate of ClpXP protease, determines when and if it is degraded.\textsuperscript{47} An additional protein, RcdA, forms a complex with CtrA and ClpX at the cell pole and is necessary for the localization and degradation of CtrA.

Cellular localization is also important for degradation of amyloid-β (Aβ), a peptide thought to be involved in Alzheimer’s disease.\textsuperscript{48} Aβ accumulates in lipid rafts. The protease neprilysin (NEP) is localized to lipid rafts after becoming glycosylated, and this localization enables it to degrade Aβ.

\textbf{C. Association with inhibitors}

Serine proteases have two main families of inhibitors: substrate-like canonical inhibitors\textsuperscript{49, 50} and “serine proteinase inhibitors” (“serpins”).\textsuperscript{51, 52} Canonical inhibitors contain a “reactive centre loop” (RCL) which is short and rigid. Serine proteases are able to cleave this loop slowly, but the products are not released, the cleaved bond can be reformed, and no conformational change occurs in either enzyme or inhibitor, resulting in inhibition of the protease.\textsuperscript{53, 54} Serpins, on the other hand, contain a long RCL which is able to form an acylenzyme complex with the catalytic serine, but this complex cannot be cleaved because
the serpin pulls the serine and oxyanion hole residues of the serine protease out of position.\textsuperscript{55,56} The salt-bridge between Ile16 and D194 is also broken, so the inhibited protease is essentially converted back to a zymogen form.

The papain-like cysteine proteases are inhibited by proteins called cystatins. These inhibitors share a common tertiary structure and bind to the cysteine proteases to block their substrate binding site.\textsuperscript{57,58} Other cysteine protease inhibitory proteins also exist. For example, inhibitors of apoptosis proteins (IAPs) inhibit caspases. XIAP inhibits caspase-9 by preventing dimerization, whereas it inhibits caspase-3 and -7 by occluding the active site. Additionally, XIAP serves as an E3-ubiquitin ligase, targeting the caspase for degradation.

Most aspartic peptidases are inhibited by the peptidomimetic pepstatin. This hexapeptide molecule binds to the active site and is thought to mimic the transition state of the substrate. When it binds, a conformational change occurs which allows the protease to bind pepstatin more tightly.\textsuperscript{59}

**D. Regulation by cofactors (ions, proteins or sugars)**

Proteases can be regulated by ions. For example, sodium acts as a cofactor for the trypsin-like serine protease thrombin.\textsuperscript{60} It determines which substrates the enzyme prefers. The sodium-free “fast” form preferentially cleaves procoagulant substrates.\textsuperscript{61} The sodium-bound “slow” form, on the other hand, activates an anticoagulation pathway by cleaving protein C.
Regulation can also occur through other proteins. For example, as mentioned above, RcdA forms a complex with protease ClpXP and its substrate CtrA, regulating the localization and degradation of CtrA.47

Sugars can also regulate proteases. As mentioned above, glycosylation is necessary for localization of the protease neprilysin (NEP) to lipid rafts.48 This localization is necessary for degradation of amyloid-β (Aβ), a peptide which accumulates in lipid rafts and is thought to be involved in Alzheimer’s disease.

E. Allostery

Allostery is a mechanism of protein regulation by which one protein site interacts with a different site over a distance. This can occur through changing the protein’s dynamics62 or, more often, through effecting a conformational change.1 Often, a ligand called an allosteric effector is involved.

Most allosteric effectors of proteases bind reversibly. Inositol hexakisphosphate acts as a noncovalent activator of C80 cysteine proteases by binding to a basic site and affecting the active site through a nearby beta turn.63, 64 For the trypsin-like serine protease thrombin, sodium acts as an allosteric regulator.60 As mentioned above, the sodium-free “fast” form and sodium-bound “slow” form cleave different substrates.

A few allosteric regulators of proteases bind irreversibly. For example Wells and coworkers found small molecules which covalently bind to cysteines in the dimer interface of caspases -1, -3, and -7.46 These inactivate the enzyme by shifting its conformation to a form similar to that of the inactive proenzyme.
Some allosteric effectors work by affecting the oligomeric state of their target protein. The compound DD2 binds to the dimer interface of Karposi’s sarcoma herpesvirus protease (KSHV Pr) to disrupt dimerization and prevent ordering of the active site. Comp-A was found to prevent dimerization of caspase-8. Pep419 was found to cause inactivation of caspase-6 through tetramerization.
REFERENCES


the zymogen and characterization of its mechanism of processing, *J Biol Chem* 266, 21451-21457.


60. Wells, C. M., and Di Cera, E. (1992) Thrombin is a Na(+) -activated enzyme, Biochemistry 31, 11721-11730.


Figure 1. Variation in Folds of Serine Proteases
A. Trypsin (PDB code 2PTN). B. Subtilisin (PDB code 1JE0). C. Carboxypeptidase Y (PDB code 1WPX). D. ClpP Subunit (PDB code 3QWD)
Figure 2. 20S Proteasome from *T. acidophilum*: A Threonine Protease
A. 20S Proteasome composed of α and β subunits (PDB code 1PMA). B. Single catalytic β subunit
Figure 3. Papain: A Cysteine Protease
(PDB code 1PPN)
Figure 4. Pepsin: An Aspartate Protease
(PDB code 3UTL)
Figure 5. Variation in Folds of Metallo Proteases
A. Thermolysin (PDB code 3T73). B. Carboxypeptidase A (PDB code 1YME).
Figure 6. Activation Mechanism of Pepsinogen
Figure 7. Catalytic Mechanism of Chymotrypsin, a Typical Serine Protease
Figure 8. Catalytic Mechanism of Papain, a Typical Cysteine Protease
Figure 9. Catalytic Mechanism of Pepsin, a Typical Aspartate Protease
Figure 10. Catalytic Mechanism of Thermolysin, a Typical Metallo Protease
CHAPTER 3

ALLOSTERIC MUTANTS OF CASPASE-3 FROM THE DIMER INTERFACE TO THE ACTIVE SITE

Christine E. Cade and A. Clay Clark

Department of Molecular and Structural Biochemistry, North Carolina State University,
Raleigh, NC 27695-7622

To Be Published
I. Introduction

Too much cell death occurs in neurodegenerative disorders such as Alzheimer's and Parkinson's. Therapeutics for these diseases could include inhibitors for caspases; the activation of these proteases commits the cell to apoptosis. Effector caspases, such as caspase-3 are expressed as monomeric zymogens which quickly dimerize, but remain inactive. Cleavage of a loop called the intersubunit linker allows conformational changes to occur which enable the active site to fully form. These changes include rotation of loops L2 and L2’ (formerly the N- and C-terminal halves of the intersubunit linker) out of the dimer interface and movement of loop L3 into the active site to form part of the substrate binding pocket, and rotation of R164 away from a solvent-exposed position to intercalate between Y197 and P201’ and enable proper positioning of the neighboring catalytic residue C163.\textsuperscript{1-4}

For caspases, allosteric inhibitors are preferable to active-site inhibitors because the active sites of caspases are highly conserved, and therefore specificity toward a single caspase has been unable to be achieved at the active site. In contrast to the conserved active site, the dimer interface is not highly conserved and has been shown to be an allosteric site.\textsuperscript{5} The allosteric inhibitors FICA and DICA which bind at the dimer interface of several caspases including caspase-3 induced large conformational changes; loop L2’ was locked into the dimer interface, loop L2 rotated out of the active site, loop L3 was extended, loop L4 was disordered, leading to a conformation which was very similar to that of the inactive procaspase.\textsuperscript{5}
Until recently, it has been thought that there are essentially just two important conformations of caspase-3: one which is zymogen-like and inactive, and one which is mature and active. A common allosteric mechanism among the caspases, mediated through short-range interactions, switches between these two conformations.⁶ This allosteric switch can be mediated either through small molecules such as FICA and DICA⁵ or through mutagenesis.⁷

However, the V266H mutation at the dimer interface of caspase-3 produced very subtle conformational changes that nonetheless completely inactivated the protein (See Figure 1).⁸ These subtle changes lead us to believe that the allostery of caspases is much more complicated than previously thought. Rather than a single active conformation and a single inactive conformation, it is more appropriate to think in terms of active and inactive ensembles.

Conceivably, there could be other allosteric sites on the protein which could also be exploited for inhibition. For example, an allosteric site has been found for caspase-6 near helix 2, which can be exploited for inactivation of the enzyme through tetramerization.⁹ Comparison of the wild-type (WT) and V266H structures led us to become interested in several residues which shifted position. We made single, double, and triple mutations at E124, Y197, and V266H.⁸ Based on activity data, crystal structures, and MD simulations, we gained an initial understanding of subtle changes that lead to inactivation of caspase-3. However, we wanted to broaden our understanding by mutating some of the other residues which shifted in the V266H mutant relative to the WT enzyme (Figure 2).
We have made two main groups of mutations: a set of single mutants (T140F, T140M, F55Y, and F55W) which aim to introduce steric clashes to mimic the V266H mutant and therefore decrease activity (“steric clash mutants”), and a set of double mutants (Y195A/V266H, Y195F/V266H, T140G/V266H, M61A/V266H, and F128A/V266H) which aim to relieve steric clashes in order to restore activity in the V266H mutant (“restorative mutants”). We have also made single mutants for each of the double mutants (Y195A, Y195F, T140G, M61A, and F128A) to determine the effect that mutation alone has on activity and structure (“control mutants”). Finally, we have looked at a salt bridge in the dimer interface with K137A and E190A mutations, and have used the T140V mutant to probe the importance of hydrogen bonding between T140 and Y195 (“other mutants”).

As expected, the control single mutants had little effect on activity, whereas in conjunction with V266H those same mutations could have a significant restorative effect. The greatest restorative effect is achieved when the mutation is closest to V266H; a Y195A mutation enabled the V266H mutant to regain full activity. We have also been able to mimic the V266H mutant by introducing steric clashes, though surprising results were seen with the T140F mutant, which accommodated the larger mutation and gained a new way of stabilizing helix 3. We have also determined that the K137-E190 salt bridge is transient and not important for activity, even though loss of this salt bridge was a major change between the WT and V266H structures. Finally, we have looked at hydrogen-bonding interactions in the dimer interface, particularly with a T140V mutant, and have seen that a T140-Y195
interaction has some effect on activity, but other hydrogen-bonding interactions also play an important role.

II. Materials and Methods

A. Cloning, expression and purification

Mutagenesis was performed as described in the Supplementary information. *Escherichia coli* BL21(DE3)pLysS cells were transformed with each of the plasmids and proteins were expressed as previously described.\textsuperscript{10-13}

B. Crystallization and data collection

Crystallization was carried out as previously described.\textsuperscript{12} Most crystals appeared within about 3 days, although some took as long as 3 weeks to grow. Cryo-protectants included 20% PEG 400/80% reservoir solution, 10% MPD/90% reservoir solution, 20% MPD/80% reservoir solution, and 20% glycerol/80% reservoir solution. 20% MPD/80% reservoir solution was the most successful. Data sets were collected at 100 K at the SER-CAT synchrotron beamline (Advance Photon Source, Argonne National Laboratory, Argonne IL, U.S.A.) The X-rays had a wavelength of 1 Å (where 1 Å = 0.1 nm), and 180 of data were collected for each protein at 1 intervals. Those mutants that crystallized with the symmetry of the orthorhombic space group I222 were phased with the caspase 3 (V266L) crystal structure, as described. Caspase 3 mutants that crystallized with the symmetry of the monoclinic space group C2 were phased with the caspase-3 (V266H) crystal structure (PDB entry 4EHA). A summary of the data collection and refinement statistics is shown in Supplementary Table S3. The atomic co-ordinates and structure factors for caspase 3
(F55W), caspase 3 (F55Y), caspase 3 (M61A), caspase 3 (M61A/V266H), caspase 3 (F128A), caspase 3 (F128A/V266H), caspase 3 (K137A), caspase 3 (T140G), caspase 3 (T140G/V266H), caspase 3 (T140F), caspase 3 (T140V), caspase 3 (E190A), caspase 3 (Y195A), caspase 3 (Y195A/V266H), caspase 3 (Y195F), and caspase 3 (Y195FV266H) have been deposited in the PDB under accession codes (TBD) respectively.

C. Enzyme activity assay

Initial velocity was measured in enzyme activity buffer (150mM Tris, 50mM NaCl, 1% sucrose, pH 7.5) at 25°C in the presence of varying concentrations of Ac-DEVD-AFC (acetyl-Asp-Glu-Val-Asp-7-amino-4-trifluoromethylcoumarin) substrate using a BioTek Synergy 2 plate reader. Substrate and buffer were mixed in a 96 well plate prior to addition of 20 µL protein by the plate reader. Fluorescence was monitored at 505 nm for 60 s.

D. Molecular dynamics simulations

Molecular dynamics simulations were performed as previously described. Briefly, 50 ns simulations were performed with GROMACS 4.5, using the Amber99 force field and the TIP3P water model.

RESULTS

A. Activity of potentially allosteric caspase-3 mutants

1. Control Mutants

The single mutants M61A, F128A, T140G, Y195A, and Y195F were designed as controls for the double mutants which have those same mutations in the background of
V266H. For the most part, they have activity comparable to that of WT (Figure 3, Table 2). The activity of F128A is the lowest of the control mutants, being reduced 15-fold from WT.

2. Restorative Mutants

The M61A/V266H, F128A/V266H, T140G/V266H, Y195A/V266H, and Y195F/V266H mutants were designed to relieve steric clashes and restore activity compared to the V266H single mutant. The Y195A mutation was the most successful in restoring activity, with the Y195A/V266H double mutant having essentially the same activity as the control Y195A single mutant. The M61A/V266H, F128A/V266H, T140G/V266H, and Y195F/V266H mutants were partially active, whereas the V266H single mutant had no activity whatsoever, indicating that some activity was restored in all of the double mutants.

The Y195A/V266H mutant had over 500-fold higher activity than the Y195F/V266H mutant. The higher activity of the Y195A/V266H mutant relative to the Y195F/V266H mutant indicates that loss of the bulky aromatic ring increases the activity. The fact that all of the double mutants gained activity compared to the single mutant and the fact that the bulkier Y195F has lower activity than the smaller Y195A point to steric clashes being one of the causes of allosteric inactivation in the V266H mutant.

In general, the restoration of activity increased with increasing proximity to the V266H mutation (Figure 4). For the three alanine double-mutants (Y195A/V266H, F128A/V266H, and M61A/V266H), activity decreases linearly with decreasing distance from V266H. The Y195F mutation, as mentioned above, is more bulky than the Y195A mutation and so its restorative effect is diminished. The T140G/V266H mutation has lower
activity than expected from the trend. It is possible that its location of T140 in helix 3 could be playing a role in its low activity.

It is important to note that the four double mutants with low measured activity were not fully cleaved during purification (See Supplementary Figure S1). Addition of Granzyme B was not able to fully process them either (data not shown). The measured activity is likely to be coming mostly from the portion of protein which is cleaved, so the $k_{cat}$ for the cleaved form, which is dependent on protein concentration, is likely to be higher for these five mutants than what is listed in Table 2. An estimated percent cleaved protein is given in Figure S1 based on our gel-doc software (Quantity One by BioRad), but estimation by eye seems to give a lower percent cleaved protein. We wish to leave it up to the reader to decide which they favor.

3. Steric Clash Mutants

The T140F and T140M mutations in the dimer interface and F55Y and F55W mutations in the active site were made in order to introduce steric clashes in a manner similar to that seen in the V266H crystal structure. These mutations were predicted to decrease activity of the enzyme. T140F decreased activity 4-fold, T140M decreased activity 200-fold, F55Y decreased activity 25-fold, and F55W decreased activity 500-fold compared to WT.

The T140F substitution is bulkier than the T140M substitution. Phenylalanine has a volume of 193.5 Å³ whereas methionine has a volume of 167.7 Å³. Therefore, based on size alone the T140F mutant would be expected to have lower activity than T140M; however, this was not the case.
In the case of the F55 mutants, the bulkier substitution led to greater reduction of activity. Although F55Y was able to become fully cleaved during purification, F55W was not. In general, the mutants which are not able to fully auto-activate have lower measured activities. This gave us a rough estimate that the larger F55W mutation decreased enzymatic activity to a greater extent than the smaller F55Y mutation. This was confirmed with the activity data.

4. Other Mutants

The E190A and K137A mutants were designed to test the importance of a salt bridge between K137 in helix 3 and E190 in the dimer interface. Although this salt bridge is present in the WT crystal structure, it is no longer present in the crystal structure of the inactive V266H mutant. It was suggested that the movement of K137 could affect residues across the interface to destabilize the L2-L2’ loop bundle. Unexpectedly, both the E190A and K137A mutants had full enzymatic activity. This demonstrates that the salt bridge between E190 and K137 is unnecessary for catalysis.

The purpose of the T140V mutation was to probe the importance of hydrogen-bonding by T140 in the activity of the enzyme since the two residues only differ in that valine has a methyl group in place of the hydroxyl group of threonine. The volumes of the two amino acids only differ slightly (139.1 Å³ for valine vs. 120.0 Å³ for threonine). The T140V mutation decreased activity 25-fold. This indicates that the hydroxyl group on T140 is playing a role in activity of the protein.
B. Crystal structures of caspase-3 mutants

1. Control Mutants

The crystal structures of all control mutants were very similar to the crystal structure of wild-type caspase-3. The crystal structure of F128A gives no clear indication as to why it has somewhat lower activity than the other control mutants and WT. Alignment of each of the crystal structures with WT shows that F128A actually aligns better than all four of the other control mutants (Table S2). The positions of the catalytic residues remain unchanged.

2. H266 Conformation in Restorative Mutants

Like the V266H single mutant, the double mutants only showed subtle differences from the WT crystal structure. The crystal structure of the inactive V266H single mutant shows a single conformation of H266: one in which the two H266 residues are facing each other across the interface, the $\varepsilon_2$-nitrogen atoms of the two imidazole rings are 3.7 Å apart, and each $\varepsilon_2$-nitrogen faces R164 from the same monomer.\(^7\) In contrast, the high activity Y195A/V266H mutant shows a different single conformation: one in which the two H266 side chains are farther apart (7 Å between the two $\varepsilon_2$-nitrogen atoms) as the $\varepsilon_2$-nitrogen atom of each imidazole ring faces the C terminus of β8.

The crystal structures of the low-activity double mutants M61A/V266H, F128A/V266H, T140G/V266H, and Y195F/V266H all show two different conformations of H266 (Figure 5). The electron density for the Y195F/V266H mutant best supports a single conformation in monomer A in which the $\varepsilon_2$-nitrogen faces R164, whereas both
conformations are seen in monomer B. M61A/V266H, F128A/V266H, and T140G/V266H show both rotamers in both monomers.

3. T140 Series

The T140V mutant is missing a hydroxyl group found in the WT protein. In WT caspase-3, this hydroxyl group hydrogen bonds to a water that connects it to K137, N141, and E190 (Figure 6A). It also hydrogen bonds to a water that connects it to Y195. In the T140V mutant, the first water is present, but only connects N141 and E190 (Figure 6E). The second water is also present and hydrogen bonds with Y195 only. Therefore, the direct connection between T140 and Y195 is lost, which may contribute to loss of activity in the T140V mutant.

Based on the crystal structure (Figure 6F), the relatively high activity of the T140F mutant can be attributed to a couple of things. First, the Phe rings at position 140 are able to rotate such that steric clashes are minimized. Second, they participate in pi-stacking interactions which stabilize helix 3.

The T140G single mutant retains full activity, despite losing the connection between T140 and Y195 (Figure 6G). The lack of a side chain at residue 140 allows Y195 to swing toward helix 3 unobstructed (the hydroxyl of Y195 moves 2.3 Å). This movement enables more water molecules to fill the interface and create a more intricate hydrogen-bonding network. These water molecules may be responsible for the retention of full activity.

The T140G/V266H mutant, in contrast, has low activity: slightly higher than the V266H single mutant (it is measurable) but much lower than WT or the T140G single
mutant. Y195 also swings 2.3 Å towards helix 3 in the double mutant; however, the resultant hydrogen bonding network is less intricate (Figure 6H). K137 is no longer connected through a water molecule to Y195, and with the connection between Y195 and T140 gone, there is no longer a direct connection between Y195 and helix 3. The waters connected to the backbone carbonyl of P201’ are no longer connected to the network of waters connecting K137, E190, and Y195, so inter-molecular communication is at least partially lost.

4. F55 Steric Clash Mutants

The F55W mutation appears from the crystal structure to be accommodated fairly well. The indole ring of the tryptophan in the mutant lies in the same plane and same approximate location as the phenylalanine ring in WT (Figure 7). However, loop L1 is slightly distorted and the M61 side chain shifts. This shift does not, however, affect the position of H121 or C163, at least in the static crystal structure.

The ring of Y55 in the F55Y structure overlays with the ring of F55 in the WT structure (Figure 7). However, no shift in loop L1 occurs and M61 from the mutant essentially overlays with that of WT. As with the F55W mutant, the catalytic residues remain unchanged.

5. K137-E190 Salt Bridge

The E190A and K137A structures differ very little from WT, which likely contributes to their high activity. Although in these two structures helix 3 is no longer connected to residue 190 through K137, a hydrogen-bonding network remains which connects N141 of
helix 3 to E190. Additionally, interactions in the hydrogen bonding network continue to connect helix 3 to P201’ across the interface.

**C. Molecular dynamics simulations**

1. **Helix 3 Destabilization**

The distance between K138 and P201’ was measured through the course of each molecular dynamics simulation to determine extent of helix 3 destabilization. Previous studies showed that the inactive V266H mutant exhibited unraveling of helix 3 in monomer B. During helix destabilization, the N-terminal half of the helix rotates towards the interface, whereas the C-terminal half stays in place. As unraveling occurs, T140 rotates away from Y195, K137 rotates away from E190 and forms interactions with P201’ and E124’. K138 rotates into the interface about 4 Å closer to P201’ and forms interactions with E190.

F55W monomer B with no inhibitor, T140G/V266H monomers A and B with no inhibitor, and T140V monomer B with inhibitor briefly exhibited unraveling (See Figure S2 for helix 3 structures of four mutants at 0 and 50 ns). T140G monomer B with no inhibitor and M61A/V266H monomer A with inhibitor exhibited unraveling for most of the simulation. Destabilization caused by T140G/V266H and to a lesser extent the T140G single mutant was seen not only in helix 3, but in the other helices as well, particularly helix 2, which is on the same face of the enzyme (Figure S3).

Since unraveling was seen in the V266H single mutant, but little to no unraveling was seen in the F128A/V266H, Y195A/V266H, and Y195F/V266H double mutants, those three restorative mutations can be considered to have a stabilizing effect on helix 3. The F55W,
T140V, and especially T140G mutations, on the other hand, can be considered to have a destabilizing effect on helix 3. It makes sense that loss of stabilizing hydrogen-bonding interactions provided by T140 (which is part of helix 3) leads to loss of stability in helix 3. The fact that the T140F mutant does not exhibit helix unraveling is likely due to stabilizing pi-stacking interactions involving the phenylalanine.

2. H266 Conformation

The conformation of the H266 residues can be seen by measuring the distance between each ε2-nitrogen and either R164 of the same monomer or R164’ in the opposite monomer. Another way to look at the H266 conformation is that of each of the histidine residues relative to each other and can be measured using the distance between the two ε2-nitrogen atoms. Using this method, one conformation is seen in which the histidines are 8-10 Å apart. A second, distinct conformation is seen in which the histidines are only 4 Å apart.

As seen in the crystal structures, two distinct conformations for H266 were seen for M61A/V266H, F128A/V266H, T140G/V266H, and Y195F/V266H (Figure 5). When the simulation is run without inhibitor, M61A/V266H spends most of its time in the conformation in which the two residues are facing each other, but when the simulation is run with inhibitor, most of its time is spent in the opposite conformation. In both simulations, the F128A/V266H mutant spends the beginning of the simulation with the residues facing each other, but moves within about 10 ns to the opposite conformation. The T140G/V266H also spends only the beginning of the simulation in the conformation in which the histidines are close together, but spends a longer time in that conformation before switching when the
simulation is run without inhibitor. The Y195F/V266H mutant spends the beginning and part of the middle of the simulation in the conformation in which the histidine residues are facing each other, but only in the simulation run with inhibitor. When run without inhibitor, the Y195F/V266H mutation only exhibits the conformation in which the histidines are farther apart. Both with and without inhibitor, only a single conformation in which the histidines were farther apart was seen for the Y195A/V266H mutant over the course of the 50ns simulation.

3. H121 Conformation

H121 alternates between two different conformations: one which is 8 Å from G122 and one which is 5 Å from G122. As it points towards G122, H121 is closer to the catalytic cysteine C163. As it points away from G122, H121 is closer to loop L1.

Taking into account monomers A and B with and without inhibitor, in the K137A, T140G, T140V, Y195A, Y195A/V266H mutants, H121 spends most of its time away from G122 and C163. In the T140F, T140G/V266H, E190A, Y195F, Y195F/V266H mutants, H121 spends more time facing away from G122 and C163, but still a significant amount of time facing towards those residues. In the F55Y mutant, H121 is split about half and half between conformations. In the F55W, M61A, M61A/V266H, F128A, F128A/V266H mutants, H121 spends most of its time facing G122 and C163.

Therefore, in general, mutations in the active site increase the amount of time H121 spends facing C163. These mutations both increase (F55W and F55Y) and decrease (M61, M61A/V266H, F128A, F128A/V266H) steric clashes. The mutants in which H121 spends
most of its time away from C163 are among the most active of those studied. However, some of the mutants in which H121 spends most of its time facing C163 (for example, M61A) also have full activity, so there is no direct correlation between conformational dwell time and activity.

The $\varepsilon_2$-nitrogen of H121, when facing away from C163, can form a polar contact with the backbone carbonyl of T62 in loop L1. Although loop L1 is relatively dynamic and in close proximity to H121, the driving force for H121 flipping does not seem to be loop L1 movement.

4. **K137-E190 Salt Bridge**

The molecular dynamics simulations show that the K137-E190 salt bridge is transient. K137 in particular is very dynamic, with the K137-E190 distance ranging from about 3 Å to about 16 Å. The dynamic nature of this salt bridge is consistent with the fact that mutating either of the involved residues does not decrease activity of the enzyme.

5. **T140-Y195 Hydrogen Bonding**

The T140-Y195 distance ranges from about 2.5 to about 8 Å but for the most part stays between 3 and 5 Å. The distance between the gamma methyl of valine 140 and the hydroxyl of Y195 was measured for the T140V mutant. Although missing a hydroxyl group normally found in T140 which connects the residue to Y195 through hydrogen-bonding, the V140-Y195 distance mostly falls within the normal 3-5 Å range. The other T140 mutants, which had a different sized side chain from threonine, were not measured for their 140-195 distance. The K137A mutant, which also affects helix 3, has a normal T140-Y195 distance.
IV. Discussion

One goal of this research was to determine the role of several residues adjacent to the activity-modulating V266 residue in the dimer interface as well as several residues adjacent to the catalytic residues in the active site. In particular, we were interested in finding out if steric clashes involving these residues lead to inactivation of the V266H mutant. To this end, we created double mutants which introduced smaller amino acids at these positions in conjunction with the V266H mutant to determine if activity could be restored. We found that a Y195A mutation, which is 4.6 Å from the V266H mutation, was able to completely restore activity of the mutant. A F128A mutation, which is 20.1 Å from the V266H mutation, had the second-highest gain in activity. An M61A mutation, which is 26.8 Å from the V266H mutation, had the third-highest gain in activity. Therefore, in general, mutations closest to the offending mutation (V266H) had the greatest effect.

One of the changes between the inactive V266H and the active Y195A/V266H mutant was that helix 3 was stabilized (Figure S2). The helix is also stabilized in the second-highest activity restorative mutant F128A/V266H. The lower activity M61A/V266H and T140G/V266H mutants continued to have an unstable helix 3. Although it too had low activity, the Y195F/V266H mutant exhibited stabilization of helix 3. This indicates that stabilization of the helix may be necessary, but insufficient for restoration of activity.

The crystal structure of the V266H single mutant shows a single conformation of H266: one in which the histidine side chains face each other across the interface and the ε2-nitrogen atom of each imidazole ring faces R164 from the same monomer. However,
molecular dynamics simulations show that within 5 nanoseconds, it has flipped to the opposite conformation: one in which the ε2-nitrogen atom faces the C terminus of β8 as well as R164’. The electron density for inactive E124A/V266H supports both conformations for monomer B, with the MD simulation showing that monomer A stays in the R164-facing conformation and monomer B is predominantly also in the R164-facing conformation with a brief switch to the R164’-facing conformation.8 The low-activity double mutants M61A/V266H, F128A/V266H, T140G/V266H, and Y195F/V266H show both conformations in the crystal structure as well as the MD simulations. The crystal structure of the medium-activity double-mutant Y197C/V266H shows only the R164-facing conformation; the MD simulation shows that it flips briefly to the R164’-facing conformation, then back to the R164-facing conformation.8 The crystal structure and MD simulation of the medium-activity E124A/Y197C/V266H triple mutant show both conformations.8 The high activity Y195A/V266H mutant only shows the R164’-facing conformation in the crystal structure and MD simulation. This is the only double mutant to have regained full activity and the only mutant to maintain the C-terminus-facing conformation. Therefore, this conformation may be involved in restoration of activity to V266H mutants.

The molecular dynamics simulations showed that H121 spends most of its time facing away from C163 in many of the active mutants. This occurs whether active site inhibitor which mimics substrate is present or absent. The fact that H121 does not spend most of its time facing C163 is further evidence that the catalytic mechanism does not involve direct
proton transfer between the two catalytic residues. Other studies have shown that water is likely involved in proton transfer.$^{18, 19}$

T140F which has the stabilizing pi-stacking interactions between F140 and Y195 shows no unraveling of helix 3. The other three T140 mutants for which MD simulations were done showed at least brief unraveling in at least one condition (with or without inhibitor and monomer A or B). This indicates that the stabilization conferred by the hydroxyl group of T140 is lost in those 3 mutants but regained in T140F, likely due to the added stabilization of ring stacking.

Based on the activity data for the K137A and E190A mutants, the K137-E190 salt bridge appears to be unnecessary for activity. Additionally, based on the molecular dynamics simulations, the salt bridge appears to be transient. The crystal structures show that an extensive hydrogen-bonding network connects helix 3 to the other side of the interface, particularly P201’. This connection to the opposite monomer may be important for activity.

Taken together, these data suggest that allosteric modulation of caspase-3 is complicated. Steric clashes, destabilization of helix 3, and disorder in loop L1 are among the possible factors that can shift the enzyme population to the inactive ensemble. Caspase-3 inhibition up to this point has focused on either active site inhibitors or allosteric inhibitors which bind to the dimer interface and cause a conformational shift to a structure similar to that of the inactive zymogen with active site loops out of position. We have found that there are other inactive conformations in which the active site remains intact, which could also be exploited for allosteric inhibitor binding. Although the mutations we have looked at are in the
active site or dimer interface, for which inhibitors already exist, our understanding gained from these mutants will pave the way for finding new drug target sites for treatment of diseases such as neurodegenerative diseases, heart disease, AIDS, or diabetes.
REFERENCES


11. Bose, K., Pop, C., Feeney, B., and Clark, A. C. (2003) An uncleavable procaspase-3 mutant has a lower catalytic efficiency but an active site similar to that of mature caspase-3, Biochemistry 42, 12298-12310.


### TABLES

**Table 1: Highlighted mutants sorted by category and location**

<table>
<thead>
<tr>
<th></th>
<th>Steric Clash</th>
<th>Restorative</th>
<th>Control</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dimer Interface</strong></td>
<td>T140F</td>
<td>T140GV266H</td>
<td>T140G</td>
<td>K137A</td>
</tr>
<tr>
<td></td>
<td>T140M</td>
<td>Y195A/V266H</td>
<td>Y195A</td>
<td>E190A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Y195F/V266H</td>
<td>Y195F</td>
<td>T140V</td>
</tr>
<tr>
<td><strong>Active Site</strong></td>
<td>F55W</td>
<td>M61A/V266H</td>
<td>M61A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F55Y</td>
<td>F128A/V266H</td>
<td>F128A</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Mutant Activity Data**

<table>
<thead>
<tr>
<th></th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}/K_M$ (M$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T140G</td>
<td>1.15±0.04</td>
<td>16±2</td>
<td>7.0E+04</td>
</tr>
<tr>
<td>E190A</td>
<td>1.24±0.03</td>
<td>18±1</td>
<td>6.9E+04</td>
</tr>
<tr>
<td>Wild-type*</td>
<td>0.8</td>
<td>14</td>
<td>5.7E+04</td>
</tr>
<tr>
<td>K137A</td>
<td>0.98±0.04</td>
<td>17±2</td>
<td>5.6E+04</td>
</tr>
<tr>
<td>Y195F</td>
<td>1.18±0.08</td>
<td>28±5</td>
<td>4.2E+04</td>
</tr>
<tr>
<td>M61A</td>
<td>1.04±0.05</td>
<td>27±4</td>
<td>3.8E+04</td>
</tr>
<tr>
<td>Y195A/V266H</td>
<td>0.33±0.02</td>
<td>23±3</td>
<td>1.5E+04</td>
</tr>
<tr>
<td>T140F</td>
<td>0.49±0.03</td>
<td>39±6</td>
<td>1.3E+04</td>
</tr>
<tr>
<td>Y195A</td>
<td>0.208±0.007</td>
<td>16±2</td>
<td>1.3E+04</td>
</tr>
<tr>
<td>F128A</td>
<td>0.054±0.002</td>
<td>15±2</td>
<td>3.7E+03</td>
</tr>
<tr>
<td>F55Y</td>
<td>0.030±0.001</td>
<td>13±2</td>
<td>2.3E+03</td>
</tr>
<tr>
<td>T140V</td>
<td>0.027±0.002</td>
<td>12±3</td>
<td>2.3E+03</td>
</tr>
<tr>
<td>T140M</td>
<td>0.0086±0.0001</td>
<td>32±1</td>
<td>2.7E+02</td>
</tr>
<tr>
<td>F128A/V266H</td>
<td>0.00252±0.00004</td>
<td>12±1</td>
<td>2.1E+02</td>
</tr>
<tr>
<td>F55W</td>
<td>0.00388±0.00006</td>
<td>36±3</td>
<td>1.1E+02</td>
</tr>
<tr>
<td>M61A/V266H</td>
<td>0.00102±0.00003</td>
<td>19±3</td>
<td>5.4E+01</td>
</tr>
<tr>
<td>Y195F/V266H</td>
<td>0.00089±0.00007</td>
<td>41±4</td>
<td>2.2E+01</td>
</tr>
<tr>
<td>T140G/V266H</td>
<td>0.00022±0.00005</td>
<td>36±4</td>
<td>6.1E+00</td>
</tr>
</tbody>
</table>
SUPPLEMENTARY DATA

Materials and Methods

Mutagenic primer sequences are listed in Table S1. The single mutants T140F, T140M, T140G, T140V, Y19A, Y195F, E190A, K137A, M61A, F128A, F55Y, F55W, were created through site-directed mutagenesis with plasmid pHC332 containing WT caspase-3 as a template. The double mutants T140G/V266H, Y195F/V266H, M61A/V266H, and F128A/V266H were created through site-directed mutagenesis with pHC33203 containing caspase-3 V266H as a template. The double mutant Y195A/V266H was created through site-directed mutagenesis with pHC332125 containing caspase-3 Y195A as a template and using the V266H primers listed in the Supplementary information of Walters, et. al (2012).
**SUPPLEMENTARY TABLES**

**Table S1: Mutagenic Primers**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Restriction Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>F55W For</td>
<td>5’-GTATAATAATTAATTAAGAAATGGCATAAAAAGTACTGGAATGACATCTCGG-3’</td>
<td>Scal</td>
</tr>
<tr>
<td>F55W Rev</td>
<td>5’-CCGAGATGTATCCAGTACTTTTTATGGCAATTTTATTATTAATTATTAC-3’</td>
<td>Scal</td>
</tr>
<tr>
<td>F55Y For</td>
<td>5’-GTATAATAATTAATTAAGAAATGGCATAAAAAGTACTGGAATGACATCTCGG-3’</td>
<td>Scal</td>
</tr>
<tr>
<td>F55Y Rev</td>
<td>5’-CCGAGATGTATCCAGTACTTTTTATGGCAATTTTATTATTAATTATTAC-3’</td>
<td>Scal</td>
</tr>
<tr>
<td>M61A For</td>
<td>5’-CATAAAAGCATGGAGCCGACATCGTGGCAAAAAATACAACAC-3’</td>
<td>AatI</td>
</tr>
<tr>
<td>M61A Rev</td>
<td>5’-CTGTACGACGACGACGTCCACAGTGTTTTTATG-3’</td>
<td>AatI</td>
</tr>
<tr>
<td>F128A For</td>
<td>5’-GGTGAAGAAGGAATAATTGCTGGAAACAATGGACCGGTGTTGA-CC-3’</td>
<td>AgeI</td>
</tr>
<tr>
<td>F128A Rev</td>
<td>5’-GGTCAGCGGTGGCTATTGCTGCCAGAAATTTTTCTC TTC-CC-3’</td>
<td>AgeI</td>
</tr>
<tr>
<td>K137A For</td>
<td>5’-GGAACAAAAAGTGGACCTGTGCACATGGCAAAAAATAACACAC-3’</td>
<td>SalI</td>
</tr>
<tr>
<td>K137A Rev</td>
<td>5’-GTTTGTTTATTTTTGGCGAGTGACAGTGCCTTTGTTCC-3’</td>
<td>SalI</td>
</tr>
<tr>
<td>T140F For</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>SspI</td>
</tr>
<tr>
<td>T140F Rev</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>SspI</td>
</tr>
<tr>
<td>T140G For</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>XmnI</td>
</tr>
<tr>
<td>T140G Rev</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>XmnI</td>
</tr>
<tr>
<td>T140M For</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>XmnI</td>
</tr>
<tr>
<td>T140M Rev</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>XmnI</td>
</tr>
<tr>
<td>T140V For</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>XmnI</td>
</tr>
<tr>
<td>T140V Rev</td>
<td>5’-CCCTGGCTGAAAAAGTTTCAACCTTTTTTCCAGAGGG-3’</td>
<td>XmnI</td>
</tr>
<tr>
<td>E190A For</td>
<td>5’-GGCGGTCATGATAATAATCCGGTGGCGGCGACTTC-3’</td>
<td>BsgI</td>
</tr>
<tr>
<td>E190A Rev</td>
<td>5’-GAAGTCCGGCCGCCACCGGTATTTTTATGACACGCCC-3’</td>
<td>BsgI</td>
</tr>
<tr>
<td>Y195A For</td>
<td>5’-GCGGACTTCCTGCGTGCATACCTCAGCAGCTCGTGG-3’</td>
<td>BsgI</td>
</tr>
<tr>
<td>Y195A Rev</td>
<td>5’-CGAGGTGACCTGCGATACGTCGACGCAAAGAGTCCGC-3’</td>
<td>BsgI</td>
</tr>
<tr>
<td>Y195F For</td>
<td>5’-GCGGACTTCCTGCGTGCATACCTCAGCAGCTCGTGG-3’</td>
<td>BsgI</td>
</tr>
<tr>
<td>Y195F Rev</td>
<td>5’-CGAGGTGACCTGCGATACGTCGACGCAAAGAGTCCGC-3’</td>
<td>BsgI</td>
</tr>
</tbody>
</table>

**Table S2: RMSD for crystal structures aligned with WT**

<table>
<thead>
<tr>
<th>Mutant</th>
<th>F55W</th>
<th>F55Y</th>
<th>M61A</th>
<th>M61A/V266H</th>
<th>F128A</th>
<th>F128A/V266H</th>
<th>K137A</th>
<th>T140F</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD</td>
<td>0.188</td>
<td>0.132</td>
<td>0.183</td>
<td>0.106</td>
<td>0.065</td>
<td>0.098</td>
<td>0.094</td>
<td>0.241</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mutant</th>
<th>T140G</th>
<th>T140G/V266H</th>
<th>T140V</th>
<th>E190A</th>
<th>Y195A</th>
<th>Y195A/V266H</th>
<th>Y195F</th>
<th>Y195F/V266H</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD</td>
<td>0.290</td>
<td>0.088</td>
<td>0.109</td>
<td>0.097</td>
<td>0.092</td>
<td>0.142</td>
<td>0.103</td>
<td>0.223</td>
</tr>
</tbody>
</table>

98
Table S3: Summary of Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WT*</th>
<th>V266H*</th>
<th>F55W</th>
<th>F55Y</th>
<th>M61A</th>
<th>M61A/V266H</th>
<th>F128A</th>
<th>F128A/V266H</th>
<th>K137A</th>
<th>T140F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>I222</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I222</td>
<td></td>
<td>I222</td>
<td></td>
</tr>
<tr>
<td>Unit cell</td>
<td>C2</td>
<td>C2</td>
<td>C121</td>
<td>C121</td>
<td>C121</td>
<td>I222</td>
<td>I222</td>
<td></td>
<td>I222</td>
<td>C121</td>
</tr>
<tr>
<td>a (Å)</td>
<td>68.73</td>
<td>109.91</td>
<td>108.508</td>
<td>127.936</td>
<td>108.227</td>
<td>67.648</td>
<td>68.793</td>
<td></td>
<td>118.285</td>
<td>68.728</td>
</tr>
<tr>
<td>b (Å)</td>
<td>84.40</td>
<td>96.77</td>
<td>96.752</td>
<td>68.311</td>
<td>96.387</td>
<td>84.256</td>
<td>84.633</td>
<td></td>
<td>85.017</td>
<td>84.423</td>
</tr>
<tr>
<td>c (Å)</td>
<td>96.35</td>
<td>69.78</td>
<td>68.417</td>
<td>84.217</td>
<td>68.794</td>
<td>96.054</td>
<td>96.313</td>
<td></td>
<td>68.994</td>
<td>96.33</td>
</tr>
<tr>
<td>α (◦)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td></td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>β (◦)</td>
<td>90</td>
<td>127.19</td>
<td>126.34</td>
<td>131.265</td>
<td>126.81</td>
<td>90</td>
<td>90</td>
<td></td>
<td>125.768</td>
<td>90</td>
</tr>
<tr>
<td>γ (◦)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td></td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.40</td>
<td>1.70</td>
<td>1.898</td>
<td>2.495</td>
<td>1.917</td>
<td>1.715</td>
<td>1.561</td>
<td></td>
<td>1.758</td>
<td>1.689</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>54279</td>
<td>54485</td>
<td>43965</td>
<td>42522</td>
<td>29405</td>
<td></td>
<td>40279</td>
<td></td>
<td>53266</td>
<td>31121</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.9</td>
<td>84.7</td>
<td>98.2</td>
<td>99.02</td>
<td>98.04</td>
<td>98.79</td>
<td>99.96</td>
<td></td>
<td>96.59</td>
<td>97.87</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.8</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average I/σ (1.34)</td>
<td>46.2</td>
<td>36.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R_work (%)</td>
<td>19.6</td>
<td>19.8</td>
<td>22.01</td>
<td>17.15</td>
<td>15.6</td>
<td>16.32</td>
<td>14.74</td>
<td></td>
<td>15.24</td>
<td>15.92</td>
</tr>
<tr>
<td>R_free (%)</td>
<td>20.7</td>
<td>23.3</td>
<td>25.94</td>
<td>24.63</td>
<td>18.76</td>
<td>19.12</td>
<td>17.06</td>
<td></td>
<td>19.44</td>
<td>19.04</td>
</tr>
<tr>
<td>R_merger (%)†</td>
<td>5.9</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSD for bond lengths (Å)</td>
<td>0.005</td>
<td>0.007</td>
<td>0.007</td>
<td>0.008</td>
<td>0.007</td>
<td>0.007</td>
<td>0.006</td>
<td></td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>RMSD for bond angles (◦)</td>
<td>1.30</td>
<td>1.15</td>
<td>1.167</td>
<td>1.125</td>
<td>1.073</td>
<td>1.123</td>
<td>1.127</td>
<td></td>
<td>1.122</td>
<td>1.126</td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>1974</td>
<td>3905</td>
<td>3880</td>
<td>3938</td>
<td>3890</td>
<td>2029</td>
<td>2027</td>
<td></td>
<td>4328</td>
<td>2057</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>290</td>
<td>386</td>
<td>206</td>
<td>219</td>
<td>405</td>
<td>209</td>
<td>323</td>
<td></td>
<td>461</td>
<td>286</td>
</tr>
</tbody>
</table>

99
Table S3 Continued

<table>
<thead>
<tr>
<th>Parameter</th>
<th>T140G</th>
<th>T140G/V266H</th>
<th>T140V</th>
<th>E190A</th>
<th>Y195A</th>
<th>Y195A/V266H</th>
<th>Y195F</th>
<th>Y195F/V266H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>C 1 2 1</td>
<td>I 2 2 2</td>
<td>I222</td>
<td>I222</td>
<td>I222</td>
<td>I 2 2 2</td>
<td>I222</td>
<td>C 1 2 1</td>
</tr>
<tr>
<td>Unit cell</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a (Å)</td>
<td>113.973</td>
<td>68.363</td>
<td>68.563</td>
<td>69.215</td>
<td>68.644</td>
<td>68.62</td>
<td>68.32</td>
<td>109.125</td>
</tr>
<tr>
<td>b (Å)</td>
<td>96.845</td>
<td>84.936</td>
<td>84.43</td>
<td>84.83</td>
<td>84.493</td>
<td>84.496</td>
<td>84.081</td>
<td>96.362</td>
</tr>
<tr>
<td>c (Å)</td>
<td>68.866</td>
<td>96.304</td>
<td>96.212</td>
<td>96.349</td>
<td>96.487</td>
<td>96.31</td>
<td>96.164</td>
<td>68.833</td>
</tr>
<tr>
<td>α (°)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>β (°)</td>
<td>129.95</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>127.1</td>
</tr>
<tr>
<td>γ (°)</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.759</td>
<td>1.498</td>
<td>1.908</td>
<td>1.602</td>
<td>1.974</td>
<td>1.954</td>
<td>1.892</td>
<td>1.843</td>
</tr>
<tr>
<td>Number of reflections</td>
<td>56387</td>
<td>44913</td>
<td>21189</td>
<td>35203</td>
<td>18917</td>
<td>20602</td>
<td>21432</td>
<td>47770</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.16</td>
<td>98.99</td>
<td>95.93</td>
<td>93.67</td>
<td>94.21</td>
<td>99.67</td>
<td>95.31</td>
<td>97.57</td>
</tr>
<tr>
<td>Redundancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average I/σ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R_work (%)</td>
<td>15.83</td>
<td>16.36</td>
<td>19.48</td>
<td>17.28</td>
<td>19.89</td>
<td>14.97</td>
<td>20.89</td>
<td>18.57</td>
</tr>
<tr>
<td>R_free (%)</td>
<td>18.87</td>
<td>18.55</td>
<td>25.28</td>
<td>19.86</td>
<td>25.89</td>
<td>18.6</td>
<td>25.38</td>
<td>22.37</td>
</tr>
<tr>
<td>R_merge (%)†</td>
<td>0.007</td>
<td>0.006</td>
<td>0.007</td>
<td>0.006</td>
<td>0.007</td>
<td>0.006</td>
<td>0.007</td>
<td>0.007</td>
</tr>
<tr>
<td>RMSD for bond lengths (Å)</td>
<td>1.106</td>
<td>1.127</td>
<td>1.141</td>
<td>1.148</td>
<td>1.095</td>
<td>1.084</td>
<td>1.095</td>
<td>1.092</td>
</tr>
<tr>
<td>RMSD for bond angles (°)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of protein atoms</td>
<td>3992</td>
<td>2126</td>
<td>2022</td>
<td>2027</td>
<td>1982</td>
<td>2052</td>
<td>1987</td>
<td>4028</td>
</tr>
<tr>
<td>No. of water molecules</td>
<td>419</td>
<td>228</td>
<td>166</td>
<td>257</td>
<td>142</td>
<td>208</td>
<td>213</td>
<td>288</td>
</tr>
</tbody>
</table>
Figure 1: Caspase Activation and Inactivation
Figure 2: Comparison of WT (green residues) and V266H (cyan residues) structures highlighting featured mutated residues.
Figure 3: Activity of WT Caspase-3 and Mutants
**Figure 4: Return of Activity vs. Distance of Rescuing Mutation from V266H**

Measured from alpha carbon to alpha carbon. Orange dotted line indicates trend among alanine mutants.
Figure 5: Conformations of H266 in Crystal Structures and Molecular Dynamics of Double Mutants

Figure 6: Stabilizing interactions and waters in interface mutants compared to WT and V266H

A) WT  B) V266H  C) K137A  D) E190A  E) T140V  F) T140F  G) T140G  H) T140G/V266H
Figure 7: Comparison of F55 Mutants with WT Caspase-3
WT = Green, F55W = Cyan, F55Y = Magenta
SUPPLEMENTARY FIGURES

Figure S1: Gel of Partially Processed Mutants
Lane 1: MW Ladder; Lane 2: F55W; Lane 3: M61A/V266H; Lane 4: F128A/V266H;
Lane 5: T140G/V266H; Lane 6: Y195F/V266H
Figure S2: Helix 3 of V266H and Restorative Mutants
A) V266H Monomer B + inhibitor  B) Y195A/V266H Monomer A - inhibitor  C) M61A/V266H Monomer A + inhibitor  D) T140G/V266H Monomer A - inhibitor
Figure S3: Helix Destabilization in T140G/V266H and T140G
T140GV266H Helix 3 at 0 and 50 ns
T140GV266H Helix 3 at 0 and 50 ns
T140G Helix 3 at 0 and 50 ns
T140G Helix 3 at 0 and 50 ns

Figure S4: Distances between Selected Residues over the Course of 50 ns Molecular Dynamics Simulations
No Inhibitor

With Inhibitor
No Inhibitor

With Inhibitor
No Inhibitor

With Inhibitor
No Inhibitor

With Inhibitor
No Inhibitor

With Inhibitor
No Inhibitor

With Inhibitor
Figure S5: Helix 3 at 0 and 50 ns for Mutants with and without Active Site Inhibitor
WT and V266H Helix 3 at 0 and 50 ns
M61AV266H Helix 3 at 0 and 50 ns
F128AV266H Helix 3 at 0 and 50 ns

T140GV266H Helix 3 at 0 and 50 ns
T140GV266H Helix 3 at 0 and 50 ns
Y195AV266H Helix 3 at 0 and 50 ns
Y195FV266H Helix 3 at 0 and 50 ns
M61A Helix 3 at 0 and 50 ns

F128A Helix 3 at 0 and 50 ns
T140G Helix 3 at 0 and 50 ns
T140G Helix 3 at 0 and 50 ns

Y195A Helix 3 at 0 and 50 ns
Y195F Helix 3 at 0 and 50 ns
K137A Helix 3 at 0 and 50 ns
E190A Helix 3 at 0 and 50 ns
T140V Helix 3 at 0 and 50 ns
T140F Helix 3 at 0 and 50 ns
F55W Helix 3 at 0 and 50 ns
F55Y Helix 3 at 0 and 50 ns
CHAPTER 4

PHAGE DISPLAY TO FIND PEPTIDES WHICH BIND TO WT AND T140G CASPASE-3

Christine E. Cade and A. Clay Clark

Department of Molecular and Structural Biochemistry, North Carolina State University,
Raleigh, NC 27695-7622
I. Introduction:

Phage display is a technique in which a large library of peptides or proteins is displayed on filamentous phage to select polypeptides with desired properties such as binding affinity to a given target.\(^1\) DNA encoding the peptide library is fused to the DNA for phage coat protein such that an individual sequence from the library is contained in the phage and the corresponding peptide is displayed on the surface of the phage, providing a genotype-phenotype link which allows for selection to be carried out.

This technique allows for determination of an ideal peptide sequence to bind to a target protein. In recent crystal structures of several mutants of caspase-3, the “safety catch” of one dimer (part of loop L2’ consisting of three aspartates in a row)\(^2\) was seen binding to helix 2 of another dimer. This binding location is similar to that of pep419 which binds to caspase-6 and mediates tetramerization.\(^3\) In the following experiments, phage display was used to probe whether the binding of the safety catch is simply an artifact of crystallography or whether it might have biological relevance.

Four different libraries were inserted into the gene for the pIII coat protein. They were selected for binding to wild-type (WT) caspase-3 both with and without active-site inhibitor and T140G caspase-3 with inhibitor. The D and DXC libraries held a central aspartate constant while varying most or all of the other residues of the peptide to see if the safety catch DDD would be selected. T140G caspase-3 was chosen as a target because it was one of the mutants to crystallize with the safety catch bound to helix 2. Experiments were also carried out with WT caspase-3 because similar binding might occur even though it was
not seen in the crystal structure. Active-site inhibitor was added to increase the chances of
finding a peptide that binds remotely to the active site.

A separate project which will not be discussed here except for the results was
designed with the X-GG and NTK-GG libraries. This project involved using WT caspase-3
with and without active site inhibitor. Since both of these projects had similar targets, all
four libraries were added to WT +/- inhibitor and T140G + inhibitor.

II. Materials and Methods:

A. Preparation of DNA:

Oligonucleotides were designed and synthesized by Integrated DNA Technologies as
in Figure 1. The D library (Figure 1B) contains an aspartate in the middle with five random
amino acids on either side. The DXC library (Figure 1C) contains five random amino acids,
an aspartate in the middle, four random amino acids and a cysteine at the end. If
circularization through disulfide bond formation is favored, another cysteine should be
selected for on the other side. The NTK-GG library (Figure 1D) contains two random amino
acids followed by four amino acids encoded by NTK (I, L, V, M or F), two central glycines,
four more amino acids encoded by NTK, and two more random residues. The X-GG library
(Figure 1E) contains six random residues, two glycines in the middle, and six more random
residues.

Then, the oligonucleotides were phosphorylated. For each library, 13 µL of H2O, 2
µL of 10x TM buffer, 2 µL of 10 mM ATP, 1 µL of 100 mM DTT, 1uL of oligo (1 µg; 40
pmole), and 1 µL of T4 polynucleotide kinase were mixed for a total of 20 µL. The reaction was incubated for 1 hour at 37 °C.

The phosphorylated oligonucleotides were annealed to ssM13 SAM33 DNA. For each library, 178 µL of H₂O, 27 µL of ssM13 SAM33 DNA (20 µg), 20 µL of phosphorylated oligonucleotide (1 µg), and 25 µL of 10x TM Buffer were mixed for a total of 250 µL. The reaction was placed into a 90 °C heat block for 2 minutes. The heat block was then removed and placed on the bench and allowed to cool slowly to 30 °C.

After annealing, covalently closed circular DNA was synthesized. For each library, 15 µL 10 mM ATP, 15 µL 100 mM DTT, 25 µL 10 mM dNTP mix (10 mM of each), 1 µL T4 DNA Ligase, and 3 µL T7 DNA polymerase were mixed for a total of 69 µL. The reaction was incubated at room temperature for 3 hours.

The covalently closed circular DNA was then purified. Each sample was added to a Microcon spin column (YM100) and spun at 10,000 rpm in a microcentrifuge for 10 minutes. Then, each sample was washed once with 250 µL of H₂O. 50 µL was added to the column, the column was inverted and the DNA collected.

B. Small Scale DNA Electroporation:

Small scale DNA electroporation was carried out prior to large scale electroporation to determine that the DNA and cells were working properly. For each library, a 1.5 mL microcentrifuge tube and a 1 mm electroporation cuvette were pre-cooled on ice. 25 µL electrocompetent SS320 cells were mixed with 1 µL purified phage library in the microcentrifuge tube. SS320 + DNA was distributed into the pre-cooled electroporation
cuvette. Samples were electroporated at 1.6 kV, 200 Ohms, 25 µF. 1 mL of pre-warmed SOC was added to the electroporation cuvette and the sample transferred to a 15 mL culture tube and incubated at 37 °C with shaking for 30 min.

To examine the diversity of the library by sequencing random isolates, a 10-fold dilution series of the library (100 µL into 900 µL of PBS) was generated out to $10^{-5}$. 100 µL of each dilution was plated with 200 µL of ER2738 cells, 40 µL 2% X-gal, 40 µL of 2% IPTG and 3 mL of molten top agar onto a 2xYT plate.

C. PCR Amplification of Positive Phage Samples:

Individual plaques from the small scale electroporation plates were picked. To 30 mL of 2xYT medium were added 500 µL of ER2738 cells. These diluted cells were added to 15 mL culture tubes (2 mL cells/tube and about 15 tubes per library). Using a pipet tip on a P200 pipetman, a blue plaque was cored from a plate. The cored plaque was added to the culture tube containing diluted cells (the media was plunged up & down in the pipet to ensure transfer of the core to the tube). After changing pipet tips, coring was repeated for all of the desired samples—one plaque per tube. Tubes were placed in the 37 °C incubator with shaking overnight. Samples were centrifuged down to pellet out the cells, leaving the phage in the supernatant.

Phage samples were PCR amplified. For each sample, 18 µL of water, 5 µL of phage supernatant, 2 µL of -96 primer + gene III primer mix (10 µM), and 25 µL of 2x Taq Master Mix were combined for a total volume of 50 µL. Samples were placed in a thermocycler: 95 °C – 50 °C – 72 °C for 30 cycles.
Samples were cleaned up using the QIAquick PCR purification kit as directed, eluting with buffer EB. Concentration of DNA was determined using a nanodrop and sequencing was performed by Eton Bioscience to assess the quality of each library.

D. Large Scale DNA Electroporation:

After DNA electroporation was carried out successfully on a small scale, it was performed on a large scale. For each library, a 1.5 mL microcentrifuge tube and three 0.2 cm electroporation cuvettes were pre-cooled on ice. To a 250 mL flask was added 30 mL of SOC medium (30 mL 2xYT + 75 µL 1M KCl + 300 µL 1M MgCl₂ + 600 µL 1M glucose) and pre-warmed to 37 °C. Electrocompetent SS320 cells (900 µL) were mixed with the purified phage library DNA (~60 µL) in the microcentrifuge tube. SS320 + DNA was distributed into the pre-cooled electroporation cuvettes (~300 µL per cuvette). Samples were electroporated at 2.5 kV, 200 Ohms, 25 µF. To the electroporation cuvette was added 1 mL of pre-warmed SOC and the sample transferred to the 250 mL flask containing 30 mL of pre-warmed SOC. Electroporation was repeated for the remaining 2 samples. The 250 mL flask containing electroporated samples was incubated at 37 °C with shaking for 25 min. To 990 µL of PBS was added 10 µL of the electroporated sample (used to calculate diversity of the library). The remaining culture was added to 1L of 2xYT + tetracycline in a 2L flask and incubated overnight at 37 °C with shaking.

To determine diversity of the library, a 10-fold dilution series of the library (100 µL into 900 µL of PBS) was generated out to $10^{-7}$. 100 µL of each dilution was plated with 200
μL of ER2738 cells, 40 μL 2% X-gal, 40 μL of 2% IPTG and 3 mL of molten top agar onto a 2xYT plate.

E. PEG Precipitation:

The overnight culture from large scale electroporation was centrifuged by dividing into sterile centrifuge bottles and spinning down in a Sorval RC-5C centrifuge using the GS3 rotor at 8000 rpm for 15 minutes. The supernatant (containing the M13 particles) was transferred to a sterile 2L flask. To the supernatant was added ¼ of the total volume of 30% PEG/1.5M NaCl (250 mL for 1L of culture) and mixed by swirling many times. This mixture was incubated on ice for 1 hour or O/N at 4°C.

The PEG/phage suspension was transferred to sterile Sorval bottles and centrifuged for 30 mintues at 8000 rpm to sediment the M13 particles. The supernatant was aspirated and discarded. The tube was inverted to drain any excess liquid, and any remaining traces of supernatant around the rim were removed with a kim wipe.

Each pellet was resuspended in 1 mL sterile PBS (4 mL total for 4 centrifuge bottles). The suspension was transferred to microcentrifuge tubes. The tubes were centrifuged for a few minutes to pellet out any bacteria and the supernatant was transferred to clean microcentrifuge tubes. Samples were placed in a 65°C water bath and heat pasteurized for 5 minutes. Sterile glycerol was added to 10% final concentration (about 1.4 mL of 40%). Phage was titered (serial dilutions to 10^{-13}) to calculate pfu/mL. Phage was kept at 4°C during QC then aliquotted (aseptically in increments of 120 or 500 μL) and frozen at -80°C.
F. Biotinylation of Target Protein:

Biotinylation of WT caspase-3 was previously carried out by Dr. Sarah MacKenzie. Biotinylation of T140G caspase-3 was carried out for this experiment. Target proteins were biotinylated with EZlink-sulfo-NHS-LC-LC-biotin (mw = 669 g/mol) and excess modifying reagent was removed by gel filtration. The total number of moles protein in the sample was calculated. Biotin reagent was removed from freezer equilibrated to room temperature and 2-5 mg were weighed out in a microcentrifuge tube. When all samples were ready dH₂O was added to biotin reagent to make it 10 mM (149 µl dH₂O per 1 mg of biotinylation reagent; I used 4.4 mg biotin and 655.6 µL dH₂O). Immediately, ten-fold molar excess was added to the protein sample. 1.25 µmol (125 µL of 10 mM) biotin was added to 0.0625 µmol (2.5 mL of 25 µM) protein (20:1). This reaction was left for 20 minutes at room temperature. (Dependent on protein, optimal conditions can vary, typically from 5-20 fold molar excess biotin reagent and 5 to 30 minutes incubation.)

Immediately thereafter, excess reagent was removed from the biotinylated protein by gel filtration on the desalting column equilibrated in storage buffer (max volume applied is 2.5 ml; the smaller the volume the better). Fractions (500 µl) were collected and protein was identified using a BioRad protein assay (5 µL protein + 195 µL Coomassie reagent). Protein-containing fractions were pooled and total protein was determined. The pool was aliquotted and frozen at -80°C.
G. Determining Saturation Point for Biotinylated Protein:

A 1.5X dilution series was made in a clean 96-well plate with biotinylated protein (WT and T140G), starting with 200 pmol target in TBST for a total volume of 450 µL in the first well of the column. Of this solution, 300 µL was added to 150 µL TBST in the second well of the column. Similarly 300 µL was taken from well 2 and mixed with 150 µL TBST in the third well of the column, and so on for a total of 16 wells (2 columns) for each of the two proteins.

For each dilution, 100 µL was transferred to the corresponding well in a 96 well plate coated in streptavidin. As controls, 0.5 mM biotin was added to the first four wells in the fifth column and TBST only was added to the last four wells in the fifth column. The plate was incubated for 30 minutes at room temperature with shaking and washed 5X with TBST.

To each well was added 20 pmol of biotinylated alkaline phosphatase in 100 µL TBST. The plate was incubated for 30 minutes at room temperature with shaking and washed 5X with TBST. Color was developed with 100 µL PNPP per well (Sigma fast tablets) and end point fluorescence was determined at 405 nm on the plate reader.

H. Selections:

Plates were prepared ahead of time. An Immulon4 plate was coated with streptavidin (enough wells to run both positive and negative controls for each phage or phage pool to be tested). The plate was blocked with 0.5% BSA in PBST and stored at 4 °C.
Streptavidin-coated plates were washed 5X with PBST. Biotinylated-target protein (WT and T140G) was added to wells in 100 μL 0.5% BSA in PBS at the concentration required to saturate the well (50 pmol protein was used for both WT and T140G). WT + Caspase-3 inhibitor Ac-DEVD-CMK was added to four wells (one for each library), WT without inhibitor was added to four wells, and T140G + inhibitor was added to four wells. Plates were incubated at room temperature for 30 minutes with shaking and washed 5X with PBST.

PBST + 0.5 mM biotin was added to the wells (150 μL for round 1 and 100 μL for subsequent rounds). For round 1, 50 μL library stock was added to the wells, and for subsequent rounds, 100 μL amplified phage supernatant from the previous round was added to the wells. Plates were incubated at room temperature for 1 hour with shaking and washed 5X with PBST.

To elute the phage, 200 μL mid-log ER2738 E. coli cells were added to each well. Plates were incubated for 30 minutes at 37 °C. Contents of each well were transferred to individual 15 mL culture tubes containing 3 mL 2xYT and incubated at 37 °C for 4 hours or overnight. After this time period, cultures were spun down to begin a new round of selection. The remaining supernatant was stored in microcentrifuge tubes for ELISAs.

I. ELISAs of Phage Pools from Selections:

A streptavidin-coated plate was washed 5X with PBST. Biotinylated target protein was added in 100 μl PBST at a concentration required to saturate the well. The plate was incubated at room temperature for 30 minutes then washed 5X with 1X PBST. Overnight
phage cultures were spun down and kept at 4 °C until ready to use. To each well was added 75 µl of 1X PBST and 25 µl phage. The plate was incubated at room temperature for 1 hour then washed 5X with 1X PBST.

To each well was added 100 µl HRP-conjugated αM13 antibody (1:5000 dilution in 0.5% BSA in PBST; 1 µl αM13 antibody per 5ml 0.5% BSA in PBST). The plate was incubated at room temperature for 30 minutes then washed 5X with 1X PBST. To each well was added 100 µl ABTS. The color was developed for 10-20 minutes at room temperature. Reactions were stopped with 20 µl 10% SDS solution. The plate was read at 405nm in a plate reader.

J. Preparation for Individual ELISAs:

For selection rounds that had a higher than 3 signal:noise ratio from the ELISAs, individual phage were isolated for ELISAs. A 10-fold serial dilution series was made of the phage pools, preparing one row per phage pool. To 10 wells of a microtiter plate were added 180 µL of PBS. To the first well was added 20 µL of the phage pool and mixed. After changing pipet tips, 20 µL were transferred from column #1 to column #2 and mixed. This dilution was repeated through column 10.

Dilutions 7, 8 & 9 were plated on a lawn of E. coli. Tubes of molten top agar (3 mL of 2xYT top agar/tube) were prepared and placed in a 50 °C heat block. To each agar tube was added 40 µL of 2% IPTG + 40 µL of 2% X-gal + 100 µL of diluted phage (tubes #7, #8, or #9 or no phage added) + 100 µL of ER2738 cells. The tube was capped and inverted three
times. The tube was poured onto a 2xYT agar plate and the plate was tilted to spread the agar evenly across the surface. The plate was allowed to sit undisturbed for 5 minutes until the top agar solidified. The plates were inverted and placed in a 37 °C incubator overnight.

ER2738 cells were diluted 1:50 in 2xYT medium. To each well of a 96 well 2 mL block was added 1 mL of diluted cells. A toothpick was used to pick individual blue plaques into the 1 mL of 2xYT + cells. Cultures were grown overnight at 37 °C with shaking. Cells were pelleted and phage-containing supernatant was used in individual ELISAs as above.

III. Results:

A. Synthesis of Covalently Closed Circular DNA:

Oligonucleotides for each library were phosphorylated, annealed with ssM13 DNA, then extended with DNA polymerase and joined to themselves with DNA ligase. A gel was run to check that the DNA was properly synthesized (Figure 2). Then the DNA was purified.

B. Random Isolates for Assessment of Library Quality:

Small scale electroporation was carried out using the purified DNA and electocompetent *E. coli* cells to make sure cells were viable and library DNA was working properly. Plaques were picked and the DNA was purified and sequenced to examine the diversity of each library (Figure 3). The D, DXC, and X-GG libraries were all unique sequences, showing 100% recombination, whereas NTK-GG was only 75% (two of the eight sequenced isolates were the SAM33 vector sequence).
C. Calculation of Diversity and Titer:

Large scale electroporation was carried out and phage was plated to calculate the diversity of the library (Table 1). D, DXC, and X-GG libraries were very diverse, whereas NTK-GG was only just barely sufficiently diverse to continue with selections. After PEG precipitation, phage was plated to calculate the titer of the library (Table 1). All four libraries had a high titer.

D. Plate Saturation Assay and Selections:

WT and T140G caspase-3 were biotinylated. A plate saturation assay was performed to determine amount needed to saturate a streptavidin plate (Figure 4). For WT, the plate was saturated through approximately the fifth dilution, corresponding to 13.2 pmol. For T140G, the plate was saturated through about the eighth dilution, corresponding to 3.9 pmol.

For selections, biotinylated protein (WT +/- inhibitor and T140G + inhibitor) was bound to streptavidin-coated plates. Phage was added and, after washing, bound phage was eluted with *E. coli* cells to be used in further rounds of selection and/or ELISAs. Selections were carried out with 50 pmol protein per well to ensure complete saturation of both WT and T140G.

E. Selection ELISAs:

After selections were carried out, phage pools from each round of selection were assayed by ELISA. As in selections, biotinylated protein (WT +/- inhibitor and T140G + inhibitor) was bound to streptavidin-coated plates. Phage was added and, after washing, antibody conjugated to horseradish peroxidase (HRP) was added to bind to phage. ABTS, a
substrate of HRP, was added to each well and the developing green color was monitored at 405 nm.

The first set of ELISAs was done on rounds 1-4, but no wells had a signal:noise ratio above 3 (Figure 5). Two additional rounds of selection were carried out and assayed by ELISA (Figures 6 and 7). The second ELISA for rounds 5-6 was performed because in the first ELISA, round 5 selections carried out against WT + inhibitor were assayed against WT – inhibitor, and round 6 selections carried out against WT – inhibitor were assayed against WT + inhibitor. Based on signal:noise ratio of at least 2, four phage pools from round 6 were chosen to be plated for use in individual ELISAs: WT – inhibitor D library, WT + inhibitor D library, T140G + inhibitor D library, and T140G + inhibitor X-GG library.

F. Individual ELISAs:

Four phage pools were plated and 48 individual plaques were picked from each, added to E. coli, and centrifuged to isolate phage-containing supernatant. Individual ELISAs were carried out using this isolated phage. Most of the isolates from the T140G + inhibitor D library (Figure 8) and WT – inhibitor D (Figure 9) library were hits (signal:noise > 3). Seven isolates from each of those two phage pools were chosen for sequencing (Table 2). Six of the WT + inhibitor D library (Figure 10) and three of the T140G + inhibitor X-GG library (Figure 11) were hits. All of those isolates were sequenced (Table 2).

All of the hits had similar sequences. The hits from the X-GG library were lacking the central two glycines and instead had a central aspartate, so it’s possible they were due to
contamination somehow. Neither the intensity of the signal nor whether the phage was raised against protein with or without inhibitor affected the sequence.

IV. Discussion:

The diversity of each library corresponded to its performance in the ELISAs after selections. The more diverse D and X-GG libraries had hits after 6 rounds of selection, whereas the DXC and NTK-GG signals were not as strong.

In the first ELISA of rounds 5-6 of selections, round six of phage raised against WT without inhibitor was assayed against WT with inhibitor and the D library was a strong hit. This indicates that at least some of this selection pool is binding away from the active site, even though it was raised against protein without active site inhibitor. This is especially surprising since the preferred substrate for caspase-3 is DEVD, with the P1 aspartate being essential.

Another indication that the phage raised against WT without inhibitor is binding away from the active site is that the sequences of the individual phage were similar if not identical to the sequences of phage raised against WT with active site inhibitor.

It is currently unclear where the phage peptides are binding. It is possible that they are binding to the dimer interface, which is an allosteric site and has been seen to bind a peptide of a different sequence (not yet published). It is also possible that the phage peptides are binding to helix 2, mimicking the binding of the “safety catch” seen in several of the crystal structures.
One possible situation in which the intermolecular binding of the safety catch could be biologically relevant is when β-amyloid fibrils activate procaspase-3. Zymogens co-localize to the fibrils and are activated through induced proximity. The interaction between the safety catch of one dimer and helix 2 of another dimer could also occur when caspase-3 binds to the apoptosome. Caspase-3 can bind in multiple locations on the apoptosome, so multiple dimers could be localized together. The active caspase-3 form binds to the apoptosome whereas the inactive procaspase-3 form does not bind strongly. This indicates that the apoptosome is able to recognize the cleaved form of the enzyme. The location of the safety catch in the activated enzyme is different from in the inactive enzyme; it could be that the safety catch plays a role in the recognition of the active enzyme by the apoptosome.
REFERENCES


### TABLES

**Table 1: Diversity and Titer of 4 Libraries**

<table>
<thead>
<tr>
<th>Library</th>
<th>Diversity (phage)</th>
<th>Titer (phage/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>$3.3 \times 10^{10}$</td>
<td>$3.8 \times 10^{10}$</td>
</tr>
<tr>
<td>DXC</td>
<td>$1.5 \times 10^{9}$</td>
<td>$2.2 \times 10^{10}$</td>
</tr>
<tr>
<td>X-GG</td>
<td>$1.3 \times 10^{10}$</td>
<td>$1.2 \times 10^{10}$</td>
</tr>
<tr>
<td>NTK-GG</td>
<td>$3.0 \times 10^{7}$</td>
<td>$1.2 \times 10^{10}$</td>
</tr>
<tr>
<td>Isolate</td>
<td>Signal/Noise</td>
<td>Sequence</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>T140G+CMK D A2</td>
<td>12.40</td>
<td>R P S V K F F D I P P E T S R AGGCCTCTGTGTAAGTTTGGATATCCGCGAGACTTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK D F4</td>
<td>9.34</td>
<td>R P S V K F F D I P P E T S R AGGCCTCTGTGTAAGTTTGGATATCCGCGAGACTTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK D C2</td>
<td>9.16</td>
<td>R P H F Y M F D P S T I H F R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK D D2</td>
<td>8.99</td>
<td>R P S V K F F D I P P E T S R AGGCCTCTGTGTAAGTTTGGATATCCGCGAGACTTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK D B3</td>
<td>5.87</td>
<td>R P H F H M Y D L Y N R H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK D G1</td>
<td>5.19</td>
<td>R P S V K F F D I P P D T S R AGGCCTCTGTGTAAGTTTGGATATCCGCGAGACTTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK D F3</td>
<td>5.70</td>
<td>R P T F K L F D V P S Y S R AGGCCTACTTTTAAAGTTGAGTGTCCGCTCTTTATCTAGA</td>
</tr>
<tr>
<td>T140G+CMK X-GG B3</td>
<td>7.88</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>T140G+CMK X-GG G5</td>
<td>8.14</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT+CMK D A1</td>
<td>7.86</td>
<td>R P H F H K Y D P H T S H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT+CMK D C1</td>
<td>10.51</td>
<td>R P H F K Y D I P P D V S R AGGCCTACTTTTAAAGTTGATTTCTCTGATGTCCTAGA</td>
</tr>
<tr>
<td>WT+CMK D C6</td>
<td>3.94</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT+CMK D F3</td>
<td>5.56</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT+CMK D H4</td>
<td>5.04</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT+CMK D H6</td>
<td>7.27</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT-CMK D G3</td>
<td>19.01</td>
<td>R P H T K Y Y D I P P V V S R AGGCCTCATACTAAGTATTGATATCCCTCTGATGTCCTAGA</td>
</tr>
<tr>
<td>WT-CMK D F3</td>
<td>17.86</td>
<td>R P H T K Y Y D I P P D V S R AGGCCTCATACTAAGTATTGATATCCCTCTGATGTCCTAGA</td>
</tr>
<tr>
<td>WT-CMK D H5</td>
<td>18.21</td>
<td>R P H T K Y Y D I P P D V S R AGGCCTCATACTAAGTATTGATATCCCTCTGATGTCCTAGA</td>
</tr>
<tr>
<td>WT-CMK D F1</td>
<td>17.49</td>
<td>R P H T K Y Y D I P P D V S R AGGCCTCATACTAAGTATTGATATCCCTCTGATGTCCTAGA</td>
</tr>
<tr>
<td>WT-CMK D A3</td>
<td>9.39</td>
<td>R P H F H M Y D P Y N N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
<tr>
<td>WT-CMK D C6</td>
<td>9.23</td>
<td>R P G F H K Y D T P N Q S S R AGGCCTGSTTTTCTATAAGTATCGATATCCCTCAATCTAGA</td>
</tr>
<tr>
<td>WT-CMK D C4</td>
<td>9.30</td>
<td>R P H F H M Y D P Y T N H S R AGGCCTCATTTTATATGGTCCATCTCTACCATTCTAGA</td>
</tr>
</tbody>
</table>
FIGURES

Figure 1: Oligo Design
A

SAM 33 vector

| V K K L L F A I P L V V S F Y S H |
| ATTTCAACGTGAACAAATATTATTATTCGCAATTATCTTTTAGTTGTTCTTTCTATTCTCAC |
| TAAAAGTTGCACTTTTTTTAAAAATAAAGCGTTAAGGAATCAACAAAGAAGATAAGAGTG |
| c-myc epitope |
| S S T R P E - K L I S E E D L N G S R P |
| TCCTCGACTAGGGCTGAAATAGAAAACTCATCTCAGAAGAGGCATTGAAATCTAGAGCT |
| AGGAGCTGATCCCCGA AGATCTGGA |
| Stu I |
| Xba I |

SR T V E S C L A K S H T E N S F T N V
TCGAGAACCTGATGAAATGGTGGTACGAAGAAAATCCCATACAGAAAAATCCATTTACTAAGCTC
AGCTCTGACAACCTTTCAACAAATCGTTTTAGGGTATGCTTTTAAGTTAAATGATTGCAG

The region in SAM 33 encoding the c-myc epitope is replaced with an oligo to encode the new library.

B

Library design:
(X) 5–D–(X) 5

Library inserted in M13 gIII

| H S S T R P (X) 5 D (X) 5 S R P S R T |
| CAC TCC TCG ACT AGG CCT (NNK) 5 GAT (NNK) 5 TCT AGA CCT TCG AGA ACT |
| GTG AGG AGC TGA TCC GGA (NNM) 5 CTA (NNM) 5 AGA TCT GGA AGC TCT TGA |

Oligo for phage construction:
5’ AGTTCCTCGAAGGTCTAGAMNNMNNMNNMNMMMNMNATCMNNMNNMNMMNNMNAGGCTAGTCGAGGAGTG
3’
C

**DXC Library design:**

(X) 5-D-(X) 4-C

Library inserted in M13 gIII

H S S T R P (X) 5 D (X) 4 C S R P S R T
CAC TCC TCG ACT AGG CCT (NNK) 5 GAT (NNK) 4 TGT TCT AGA CCT TCG AGA ACT
GTG AGG AGC TGA TCC GGA (NNM) 5 CTA (NNM) 4 ACA AGA TCT GGA AGC TCT TGA

Oligo for phage construction:

5’ AGTTCTCGAAGGTCATAGACMNMMNNNNNNNATCMNNNNNNNNNNNNNNAGGCCTAGTCGAGGAGTG 3’

D

**NTK-GG Library design:**

(X) 2-(NTK) 4-G-G-(NTK) 4-(X) 2

H S S T R P (X) 2 (ILVFM) 4 G G (ILVFM) 4 (X) S R P
S CAC TCC TCG ACT AGG CCT (NNK) 2 (NTK) 4 GGC GGT (NTK) 4 (NNK) 2 TCT AGA
CCT TCG
R I
AGA ACT

NTK - {I,L,V,M,F}

Oligo for phage construction:

5’ AGTTCTCGAAGGTCATAGACMNMMNNMMNNNNNAMANANNANMANAMNANNANMNNNNNNNNAPGCCTAGTC
GAGGAGTG 3’

E

**X-GG Library design:**

(X) 6-G-G-(X) 6

H S S T R P (X) 6 G G (X) 6 S R P S R T
CAC TCC TCG ACT AGG CCT (NNK) 6 GGC GGT (NNK) 6 TCT AGA CCT TCG AGA ACT

Oligo for phage construction:

5’ AGTTCTCGAAGGTCATAGACMNMMNNMMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAAGGCCTAGTC
GAGGAGTG 3’
Figure 2: DNA Gel
Lanes: 1) Ladder 2) ssM13 DNA control 3) X-GG 4) NTK-GG 5) D 6) DXC
Figure 3: Random Isolates for Diversity
D3
R P A S L A G D W L N G D S R
AGGCCTGCTAGTCCTGCTGATGGCTGAAATGGATCTCTAGA

D4
R P D T A A M D N D A L L S R
AGGCCTGATACCTGCTATGGGATAATGATGCTCTGCTCTAGA

D5
R P G C N D F D C W I F T S R
AGGCCTGATGGAATGAATTTTGGATTGTGATTTACTCTCTAGA

D11
R P E Y W F Y D A S A S Q S R
AGGCCTGATATTGGTTTATGATCGGCTCTTGTCCTAGTCTAGA

D13
R P V L W Y M D N S S D L S R
AGGCCTGCTGCTTTGGAATAATGGGATAATTCTGGGATCTGCTAGA

D14
R P S T S A N D L A Q R S R
AGGCCTAGTACTAGTCAATGATGATCTTCAACCGGCTCTAGA

D15
R P N D L Y V D G A F H I S R
AGGCCTAATGATCTTTATGATGATGCTGCTGCTTTTCATATTCTAGA
B

DXC1
RPLCTYVDRNISCR
AGGCCCTGTTACATTGTGGGATCATGGAATTACTATTTGTTCTAGA

DXC2
RPSIAATDYTDNCSCR
AGGCCCTAGTTATGCTACTGATTATACGGATAATTTGTTCTAGA

DXC4
RPVNCMDSHNICSRR
AGGCCCTGTTAATTGTCTTTATGGATTAGTCTAATAATATTGTTCTAGA

DXC5
RPFGCDSASAECSRR
AGGCCCTTTTGGGTGTGATAGTGTGATGCTAGTGGCTAGTGTGTTCTAGA

DXC6
RPSEITCDITVCSCR
AGGCCCTAGTGAGATTACTTTGATGATATTACTGTTTTGTTCTAGA

DXC8
RPDLAVDRIAGCSRR
AGGCCCTGATGCTCTTATTGTTGATCTATTGCCTGGGTGTTCTAGA

DXC9
RPVDNHRIHHPCSRR
AGGCCCTGTTGATAATCATCGTATATCATTACATTATCTGTTCTAGA

DXC10
RPGSLYDITTCCSR
AGGCCCTGCTGAGTCTTTATTAGATATTACGCATTGTGTTCTAGA

DXC11
RPELSNGDNHSTCSCR
AGGCCCTGAGCTTATGTAATGTTGATATTATCTACTTCTTGTCTAGA

DXC12
RPFCCYPGDLSRLSCSR
AGGCCCTTTTGTATTCTGTTGATTCTTTGCGGTCTTCTTCTAGA

DXC14
RPVNCESDPWAICSRR
AGGCCCTGTAATTGTTGAGTCGTCGTACCTGATTGGGTATTGTTCTAGA

DXC15
RPGNPVTDDSYRCSRR
AGGCCCTGGTAATCCTGTTGACTGATTTATAGTTGTTCTAGA
C

X-GG2
RPWIQQIAAGGVISVGNRS
AGG CCTTGGCATCGAGATGCCTGGGCCTGAATTTCGTTAGCTAGA

X-GG3
RPITHDGGVITSGPSR
AGG CCTACGATTACGTACAAGGGGGGTTGAGGCATGATTGCTCTCTAGA

X-GG4
RPLSMYRNNGGVYLTHERS
AGG CCTCTTCGATGATCGGAATGGCCGATTTATTGACTCATGAGCTAGA

X-GG5
RPYIGNYTGGDIVAANMR
AGG CCTATATTGGTAATTACGGGGCGGATGTTGCTGCGGTAATTCTAGA

X-GG6
RPSPAIIDDGGISLQPTS
AGG CCTCCTTCCGGCTATTTGATGAGCCGGTTAGTGCTGACCCGACGTCTAGA

X-GG7
RPNTAPESGGFQILYSSR
AGG CCTACGAATCGCTCTGAGTCGGCCGGTCTCTGAGCTAGCTCTCTTAGA

X-GG11
RPDVLPAARGGMPIQSSR
AGG CCTGATGTGTCTTCTGCTAGGGCCGTTGATGATTCACTGAGCTCTAGA

X-GG13
RPDSLTYTSGGPAKHEQSSR
AGG CCTATTTCTTTCATCTGGGCCTGCTGCTAAGCATGACGACTCTAGA

X-GG14
RPSTHPAGGTAYSGKS
AGG CCTTCGACGATCTGCGGCTGCGGCTACTGCTTATTCCGGGAAAGCTAGA
DNK2
RPSPLGGVMVDSSR
AGGCCATCTGTTCTGTGCTGGGCTTTTTGTATGTTGGATCTTAGA

DNK4
RPEKLISEEDLNSR (SAM33 Vector sequence)
AGGCCCTGATAGAAACTCATCTCATAGAAGAGATCTGATGGATCTAGA

DNK5
RPDCLSISR
AGGCCCTGATATGCTTTATATTCTAGA

DNK8
RPTSIYDWADIASR
AGGCCCTACTAGTATCTGATTAGGGCTGATATTGCCCTAGA

DNK9
RPAGLVGMGFLMIEGSR
AGGCCCTGCTGTCTTGATTTGGGCGGTGTTGTTATGATGGACATTCTAGA

DNK10
RPSPLILGGVMMEHCSR
AGGCCCTTCTCCTATGCTTTATTTTTGGGCGGTGTTGTTATGATGGACATTCTAGA

DNK11
RPLPMLILGGVVMMEHCSR
AGGCCCTCTCTGCTTATTTTTGGGCGGTGTTGTTATGATGGACATTCTAGA

DNK12
RPEKLISEEDLNSR (SAM33 Vector sequence)
AGGCCCTGATAGAAACTCATCTCATAGAAGAGATCTGATGGATCTAGA
Figure 4: Saturation Assay
Figure 5: Rounds 1-4 Selection ELISA

A: Plate layout: indicated protein was added to columns 1-6. No protein was added to columns 7-12. B: Raw Data. C: Signal:Noise calculated by taking Abs of well containing protein and dividing by Abs of well not containing protein.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WT-CMK</td>
<td>WT-CMK</td>
<td>WT-CMK</td>
<td>WT-CMK</td>
<td>T140G+CMK</td>
<td>T140G+CMK</td>
</tr>
<tr>
<td></td>
<td>NTK1</td>
<td>NTK3</td>
<td>NTK1</td>
<td>NTK3</td>
<td>NTK1</td>
<td>NTK3</td>
</tr>
<tr>
<td>B</td>
<td>WT-CMKX-GG1</td>
<td>WT-CMKX-GG3</td>
<td>WT-CMKX-GG1</td>
<td>WT-CMKX-GG3</td>
<td>T140G+CMKX-GG1</td>
<td>T140G+CMKX-GG3</td>
</tr>
<tr>
<td>C</td>
<td>WT-CMKDXC1</td>
<td>WT-CMKDXC3</td>
<td>WT-CMKDXC1</td>
<td>WT-CMKDXC3</td>
<td>T140G+CMKDXC1</td>
<td>T140G+CMKDXC3</td>
</tr>
<tr>
<td>D</td>
<td>WT-CMKD1</td>
<td>WT-CMKD3</td>
<td>WT-CMKD1</td>
<td>WT-CMKD3</td>
<td>T140G+CMKD1</td>
<td>T140G+CMKD3</td>
</tr>
<tr>
<td>E</td>
<td>WT-CMKNTK2</td>
<td>WT-CMKNTK4</td>
<td>WT-CMKNTK2</td>
<td>WT-CMKNTK4</td>
<td>T140G+CMKNTK2</td>
<td>T140G+CMKNTK4</td>
</tr>
<tr>
<td>F</td>
<td>WT-CMKX-GG2</td>
<td>WT-CMKX-GG4</td>
<td>WT-CMKX-GG2</td>
<td>WT-CMKX-GG4</td>
<td>T140G+CMKX-GG2</td>
<td>T140G+CMKX-GG4</td>
</tr>
<tr>
<td>G</td>
<td>WT-CMKDXC2</td>
<td>WT-CMKDXC4</td>
<td>WT-CMKDXC2</td>
<td>WT-CMKDXC4</td>
<td>T140G+CMKDXC2</td>
<td>T140G+CMKDXC4</td>
</tr>
<tr>
<td>H</td>
<td>WT-CMKD2</td>
<td>WT-CMKD4</td>
<td>WT-CMKD2</td>
<td>WT-CMKD4</td>
<td>T140G+CMKD2</td>
<td>T140G+CMKD4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WT-CMK</td>
<td>WT-CMK</td>
<td>WT-CMK</td>
<td>WT-CMK</td>
<td>T140G+CMK</td>
<td>T140G+CMK</td>
</tr>
<tr>
<td></td>
<td>NTK1</td>
<td>NTK3</td>
<td>NTK1</td>
<td>NTK3</td>
<td>NTK1</td>
<td>NTK3</td>
</tr>
<tr>
<td>B</td>
<td>WT-CMKX-GG1</td>
<td>WT-CMKX-GG3</td>
<td>WT-CMKX-GG1</td>
<td>WT-CMKX-GG3</td>
<td>T140G+CMKX-GG1</td>
<td>T140G+CMKX-GG3</td>
</tr>
<tr>
<td>C</td>
<td>WT-CMKDXC1</td>
<td>WT-CMKDXC3</td>
<td>WT-CMKDXC1</td>
<td>WT-CMKDXC3</td>
<td>T140G+CMKDXC1</td>
<td>T140G+CMKDXC3</td>
</tr>
<tr>
<td>D</td>
<td>WT-CMKD1</td>
<td>WT-CMKD3</td>
<td>WT-CMKD1</td>
<td>WT-CMKD3</td>
<td>T140G+CMKD1</td>
<td>T140G+CMKD3</td>
</tr>
<tr>
<td>E</td>
<td>WT-CMKNTK2</td>
<td>WT-CMKNTK4</td>
<td>WT-CMKNTK2</td>
<td>WT-CMKNTK4</td>
<td>T140G+CMKNTK2</td>
<td>T140G+CMKNTK4</td>
</tr>
<tr>
<td>F</td>
<td>WT-CMKX-GG2</td>
<td>WT-CMKX-GG4</td>
<td>WT-CMKX-GG2</td>
<td>WT-CMKX-GG4</td>
<td>T140G+CMKX-GG2</td>
<td>T140G+CMKX-GG4</td>
</tr>
<tr>
<td>G</td>
<td>WT-CMKDXC2</td>
<td>WT-CMKDXC4</td>
<td>WT-CMKDXC2</td>
<td>WT-CMKDXC4</td>
<td>T140G+CMKDXC2</td>
<td>T140G+CMKDXC4</td>
</tr>
<tr>
<td>H</td>
<td>WT-CMKD2</td>
<td>WT-CMKD4</td>
<td>WT-CMKD2</td>
<td>WT-CMKD4</td>
<td>T140G+CMKD2</td>
<td>T140G+CMKD4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>A</td>
<td>0.11</td>
<td>0.108</td>
<td>0.101</td>
<td>0.087</td>
<td>0.087</td>
<td>0.099</td>
</tr>
<tr>
<td>B</td>
<td>0.119</td>
<td>0.088</td>
<td>0.095</td>
<td>0.102</td>
<td>0.093</td>
<td>0.094</td>
</tr>
<tr>
<td>C</td>
<td>0.138</td>
<td>0.081</td>
<td>0.097</td>
<td>0.096</td>
<td>0.1</td>
<td>0.096</td>
</tr>
<tr>
<td>D</td>
<td>0.18</td>
<td>0.111</td>
<td>0.112</td>
<td>0.146</td>
<td>0.118</td>
<td>0.113</td>
</tr>
<tr>
<td>E</td>
<td>0.112</td>
<td>0.106</td>
<td>0.136</td>
<td>0.121</td>
<td>0.101</td>
<td>0.104</td>
</tr>
<tr>
<td>F</td>
<td>0.083</td>
<td>0.114</td>
<td>0.091</td>
<td>0.129</td>
<td>0.08</td>
<td>0.147</td>
</tr>
<tr>
<td>G</td>
<td>0.093</td>
<td>0.093</td>
<td>0.096</td>
<td>0.14</td>
<td>0.093</td>
<td>0.1</td>
</tr>
<tr>
<td>H</td>
<td>0.106</td>
<td>0.228</td>
<td>0.106</td>
<td>0.13</td>
<td>0.077</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.102</td>
<td>0.094</td>
<td>0.108</td>
<td>0.094</td>
<td>0.1</td>
<td>0.099</td>
</tr>
<tr>
<td>B</td>
<td>0.102</td>
<td>0.1</td>
<td>0.095</td>
<td>0.089</td>
<td>0.099</td>
<td>0.098</td>
</tr>
<tr>
<td>C</td>
<td>0.129</td>
<td>0.093</td>
<td>0.104</td>
<td>0.095</td>
<td>0.112</td>
<td>0.114</td>
</tr>
<tr>
<td>D</td>
<td>0.101</td>
<td>0.086</td>
<td>0.096</td>
<td>0.085</td>
<td>0.112</td>
<td>0.101</td>
</tr>
<tr>
<td>E</td>
<td>0.08</td>
<td>0.098</td>
<td>0.088</td>
<td>0.106</td>
<td>0.085</td>
<td>0.104</td>
</tr>
<tr>
<td>F</td>
<td>0.092</td>
<td>0.108</td>
<td>0.079</td>
<td>0.124</td>
<td>0.076</td>
<td>0.104</td>
</tr>
<tr>
<td>G</td>
<td>0.084</td>
<td>0.098</td>
<td>0.083</td>
<td>0.091</td>
<td>0.089</td>
<td>0.106</td>
</tr>
<tr>
<td>H</td>
<td>0.113</td>
<td>0.109</td>
<td>0.075</td>
<td>0.099</td>
<td>0.079</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>A</td>
<td>1.08</td>
<td>1.15</td>
<td>0.94</td>
<td>0.93</td>
<td>0.87</td>
<td>1.00</td>
</tr>
<tr>
<td>B</td>
<td>1.17</td>
<td>0.88</td>
<td>1.00</td>
<td>1.15</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td>C</td>
<td>1.07</td>
<td>0.87</td>
<td>0.93</td>
<td>1.01</td>
<td>0.89</td>
<td>0.84</td>
</tr>
<tr>
<td>D</td>
<td>1.78</td>
<td>1.29</td>
<td>1.17</td>
<td>1.72</td>
<td>1.05</td>
<td>1.12</td>
</tr>
<tr>
<td>E</td>
<td>1.40</td>
<td>1.08</td>
<td>1.55</td>
<td>1.14</td>
<td>1.19</td>
<td>1.00</td>
</tr>
<tr>
<td>F</td>
<td>0.90</td>
<td>1.06</td>
<td>1.15</td>
<td>1.04</td>
<td>1.05</td>
<td>1.41</td>
</tr>
<tr>
<td>G</td>
<td>1.11</td>
<td>0.95</td>
<td>1.16</td>
<td>1.54</td>
<td>1.04</td>
<td>0.94</td>
</tr>
<tr>
<td>H</td>
<td>0.94</td>
<td>2.09</td>
<td>1.41</td>
<td>1.31</td>
<td>0.97</td>
<td>1.55</td>
</tr>
</tbody>
</table>
Figure 6: Rounds 5-6 Selection ELISA
A: Plate layout: description designates phage that was added. WT - CMK was added to column 1, WT + CMK was added to column 2, and T140G + CMK was added to column 3. No protein was added to columns 7-9. B: Raw Data. C: Signal:Noise calculated by taking Abs of well containing protein and dividing by Abs of well not containing protein. D: Picture of ELISA plate.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>WT +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NTK 5</td>
<td>NTK 6</td>
<td>NTK 5</td>
<td>A10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>WT +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X-GG 5</td>
<td>X-GG 6</td>
<td>X-GG 5</td>
<td>H2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>WT +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXC 5</td>
<td>DXC 6</td>
<td>DXC 5</td>
<td>A10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>WT +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D 5</td>
<td>D 6</td>
<td>D 5</td>
<td>H2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>T140G +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NTK 5</td>
<td>NTK 6</td>
<td>NTK 6</td>
<td>A10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>T140G +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X-GG 5</td>
<td>X-GG 6</td>
<td>X-GG 6</td>
<td>H2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>T140G +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXC 5</td>
<td>DXC 6</td>
<td>DXC 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>T140G +CMK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D 5</td>
<td>D 6</td>
<td>D 6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>NTK 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NTK 5</td>
<td>NTK 6</td>
<td>NTK 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>X-GG 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X-GG 5</td>
<td>X-GG 6</td>
<td>X-GG 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>DXC 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXC 5</td>
<td>DXC 6</td>
<td>DXC 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>D 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D 5</td>
<td>D 6</td>
<td>D 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>NTK 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NTK 5</td>
<td>NTK 6</td>
<td>NTK 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>X-GG 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X-GG 5</td>
<td>X-GG 6</td>
<td>X-GG 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>DXC 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DXC 5</td>
<td>DXC 6</td>
<td>DXC 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>WT +CMK</td>
<td>WT +CMK</td>
<td>T140G +CMK</td>
<td>D 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D 5</td>
<td>D 6</td>
<td>D 6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>A</td>
<td>0.097</td>
<td>0.109</td>
<td>0.082</td>
<td>0.047</td>
<td>0.075</td>
<td>0.048</td>
</tr>
<tr>
<td>B</td>
<td>0.092</td>
<td>0.102</td>
<td>0.08</td>
<td>0.053</td>
<td>0.086</td>
<td>0.057</td>
</tr>
<tr>
<td>C</td>
<td>0.088</td>
<td>0.106</td>
<td>0.096</td>
<td>0.071</td>
<td>0.077</td>
<td>0.063</td>
</tr>
<tr>
<td>D</td>
<td>0.271</td>
<td>0.926</td>
<td>0.362</td>
<td>0.048</td>
<td>0.098</td>
<td>0.07</td>
</tr>
<tr>
<td>E</td>
<td>0.084</td>
<td>0.091</td>
<td>0.316</td>
<td>0.078</td>
<td>0.079</td>
<td>0.066</td>
</tr>
<tr>
<td>F</td>
<td>0.075</td>
<td>0.092</td>
<td>0.235</td>
<td>0.053</td>
<td>0.149</td>
<td>0.061</td>
</tr>
<tr>
<td>G</td>
<td>0.079</td>
<td>0.089</td>
<td>0.391</td>
<td>0.062</td>
<td>0.074</td>
<td>0.087</td>
</tr>
<tr>
<td>H</td>
<td>0.096</td>
<td>0.165</td>
<td>1.73</td>
<td>0.046</td>
<td>0.07</td>
<td>0.052</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.186</td>
<td>0.121</td>
<td>1.222</td>
<td>0.048</td>
<td>0.048</td>
<td>0.091</td>
</tr>
<tr>
<td>B</td>
<td>0.083</td>
<td>0.076</td>
<td>0.529</td>
<td>0.049</td>
<td>0.057</td>
<td>0.056</td>
</tr>
<tr>
<td>C</td>
<td>0.132</td>
<td>0.078</td>
<td>0.779</td>
<td>0.056</td>
<td>0.051</td>
<td>0.059</td>
</tr>
<tr>
<td>D</td>
<td>0.134</td>
<td>0.082</td>
<td>1.03</td>
<td>0.052</td>
<td>0.062</td>
<td>0.063</td>
</tr>
<tr>
<td>E</td>
<td>0.181</td>
<td>0.078</td>
<td>0.82</td>
<td>0.05</td>
<td>0.055</td>
<td>0.056</td>
</tr>
<tr>
<td>F</td>
<td>0.084</td>
<td>0.07</td>
<td>0.08</td>
<td>0.052</td>
<td>0.055</td>
<td>0.058</td>
</tr>
<tr>
<td>G</td>
<td>0.089</td>
<td>0.069</td>
<td>0.429</td>
<td>0.048</td>
<td>0.061</td>
<td>0.06</td>
</tr>
<tr>
<td>H</td>
<td>0.878</td>
<td>0.077</td>
<td>0.069</td>
<td>0.039</td>
<td>0.045</td>
<td>0.056</td>
</tr>
</tbody>
</table>
### C

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.52</td>
<td>0.90</td>
<td>0.07</td>
<td>0.98</td>
<td>1.56</td>
<td>0.53</td>
</tr>
<tr>
<td>B</td>
<td>1.11</td>
<td>1.34</td>
<td>0.15</td>
<td>1.08</td>
<td>1.51</td>
<td>1.02</td>
</tr>
<tr>
<td>C</td>
<td>0.67</td>
<td>1.36</td>
<td>0.12</td>
<td>1.27</td>
<td>1.51</td>
<td>1.07</td>
</tr>
<tr>
<td>D</td>
<td>2.02</td>
<td>11.29</td>
<td>0.35</td>
<td>0.92</td>
<td>1.58</td>
<td>1.11</td>
</tr>
<tr>
<td>E</td>
<td>0.46</td>
<td>1.17</td>
<td>0.39</td>
<td>1.56</td>
<td>1.44</td>
<td>1.18</td>
</tr>
<tr>
<td>F</td>
<td>0.89</td>
<td>1.31</td>
<td>2.94</td>
<td>1.02</td>
<td>2.71</td>
<td>1.05</td>
</tr>
<tr>
<td>G</td>
<td>0.89</td>
<td>1.29</td>
<td>0.91</td>
<td>1.29</td>
<td>1.21</td>
<td>1.45</td>
</tr>
<tr>
<td>H</td>
<td>0.11</td>
<td>2.14</td>
<td>25.07</td>
<td>1.18</td>
<td>1.56</td>
<td>0.93</td>
</tr>
</tbody>
</table>

### D

[Image of a 96-well plate with two wells highlighted in green]
Figure 7: Rounds 5-6 Selection ELISA Repeat
A: Plate layout: indicated protein was added to columns 9-10. No protein was added to columns 11-12. B: Raw Data. C: Signal:Noise calculated by taking Abs of well containing protein and dividing by Abs of well not containing protein. D: Picture of ELISA plate.
<table>
<thead>
<tr>
<th></th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WT-CMK NTK 5</td>
<td>WT+CMK NTK 5</td>
<td>WT-CMK NTK 5</td>
<td>WT+CMK NTK 5</td>
</tr>
<tr>
<td>B</td>
<td>WT-CMK X-GG 5</td>
<td>WT+CMK X-GG 5</td>
<td>WT-CMK X-GG 5</td>
<td>WT+CMK X-GG 5</td>
</tr>
<tr>
<td>C</td>
<td>WT-CMK D XC 5</td>
<td>WT+CMK D XC 5</td>
<td>WT-CMK D XC 5</td>
<td>WT+CMK D XC 5</td>
</tr>
<tr>
<td>D</td>
<td>WT-CMK D 5</td>
<td>WT+CMK D 5</td>
<td>WT-CMK D 5</td>
<td>WT+CMK D 5</td>
</tr>
<tr>
<td>E</td>
<td>WT-CMK NTK 6</td>
<td>WT+CMK NTK 6</td>
<td>WT-CMK NTK 6</td>
<td>WT+CMK NTK 6</td>
</tr>
<tr>
<td>F</td>
<td>WT-CMK X-GG 6</td>
<td>WT+CMK X-GG 6</td>
<td>WT-CMK X-GG 6</td>
<td>WT+CMK X-GG 6</td>
</tr>
<tr>
<td>G</td>
<td>WT-CMK D XC 6</td>
<td>WT+CMK D XC 6</td>
<td>WT-CMK D XC 6</td>
<td>WT+CMK D XC 6</td>
</tr>
<tr>
<td>H</td>
<td>WT-CMK D 6</td>
<td>WT+CMK D 6</td>
<td>WT-CMK D 6</td>
<td>WT+CMK D 6</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>---</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>A</td>
<td>0.097</td>
<td>0.123</td>
<td>0.087</td>
<td>0.086</td>
</tr>
<tr>
<td>B</td>
<td>0.116</td>
<td>0.104</td>
<td>0.091</td>
<td>0.088</td>
</tr>
<tr>
<td>C</td>
<td>0.105</td>
<td>0.105</td>
<td>0.114</td>
<td>0.094</td>
</tr>
<tr>
<td>D</td>
<td>0.341</td>
<td>0.111</td>
<td>0.11</td>
<td>0.096</td>
</tr>
<tr>
<td>E</td>
<td>0.116</td>
<td>0.146</td>
<td>0.101</td>
<td>0.118</td>
</tr>
<tr>
<td>F</td>
<td>0.122</td>
<td>0.116</td>
<td>0.102</td>
<td>0.114</td>
</tr>
<tr>
<td>G</td>
<td>0.15</td>
<td>0.121</td>
<td>0.104</td>
<td>0.122</td>
</tr>
<tr>
<td>H</td>
<td>0.499</td>
<td>0.327</td>
<td>0.133</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1.11</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>1.27</td>
<td>1.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.92</td>
<td>1.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td><strong>3.10</strong></td>
<td>1.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1.15</td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1.20</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1.44</td>
<td>0.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td><strong>3.75</strong></td>
<td>2.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 8: T140G + CMK D ELISA
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.971</td>
<td>1.327</td>
<td>0.724</td>
<td>0.487</td>
<td>0.456</td>
<td>0.354</td>
</tr>
<tr>
<td>B</td>
<td>0.648</td>
<td>0.587</td>
<td>0.54</td>
<td>0.647</td>
<td>0.455</td>
<td>0.585</td>
</tr>
<tr>
<td>C</td>
<td>0.183</td>
<td>0.705</td>
<td>0.685</td>
<td>0.265</td>
<td>0.513</td>
<td>0.677</td>
</tr>
<tr>
<td>D</td>
<td>0.57</td>
<td>0.809</td>
<td>0.702</td>
<td>0.316</td>
<td>0.244</td>
<td>0.477</td>
</tr>
<tr>
<td>E</td>
<td>0.656</td>
<td>0.424</td>
<td>0.595</td>
<td>0.652</td>
<td>0.568</td>
<td>0.69</td>
</tr>
<tr>
<td>F</td>
<td>0.463</td>
<td>0.547</td>
<td>0.479</td>
<td>0.775</td>
<td>0.391</td>
<td>0.44</td>
</tr>
<tr>
<td>G</td>
<td>0.727</td>
<td>0.622</td>
<td>0.512</td>
<td>0.56</td>
<td>0.74</td>
<td>0.583</td>
</tr>
<tr>
<td>H</td>
<td>0.653</td>
<td>0.325</td>
<td>0.484</td>
<td>0.658</td>
<td>0.626</td>
<td>0.315</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.316</td>
<td>0.107</td>
<td>0.083</td>
<td>0.17</td>
<td>0.219</td>
<td>0.094</td>
</tr>
<tr>
<td>B</td>
<td>0.085</td>
<td>0.073</td>
<td>0.092</td>
<td>0.092</td>
<td>0.12</td>
<td>0.082</td>
</tr>
<tr>
<td>C</td>
<td>0.073</td>
<td>0.077</td>
<td>0.095</td>
<td>0.101</td>
<td>0.125</td>
<td>0.078</td>
</tr>
<tr>
<td>D</td>
<td>0.086</td>
<td>0.09</td>
<td>0.081</td>
<td>0.096</td>
<td>0.132</td>
<td>0.082</td>
</tr>
<tr>
<td>E</td>
<td>0.092</td>
<td>0.102</td>
<td>0.086</td>
<td>0.093</td>
<td>0.156</td>
<td>0.086</td>
</tr>
<tr>
<td>F</td>
<td>0.075</td>
<td>0.076</td>
<td>0.084</td>
<td>0.083</td>
<td>0.108</td>
<td>0.08</td>
</tr>
<tr>
<td>G</td>
<td>0.14</td>
<td>0.084</td>
<td>0.091</td>
<td>0.083</td>
<td>0.09</td>
<td>0.091</td>
</tr>
<tr>
<td>H</td>
<td>0.084</td>
<td>0.089</td>
<td>0.076</td>
<td>0.082</td>
<td>0.086</td>
<td>0.086</td>
</tr>
</tbody>
</table>
### B

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3.07</td>
<td>12.40</td>
<td>8.72</td>
<td>2.86</td>
<td>2.08</td>
<td>3.77</td>
</tr>
<tr>
<td>B</td>
<td>7.62</td>
<td>8.04</td>
<td>5.87</td>
<td>7.03</td>
<td>3.79</td>
<td>7.13</td>
</tr>
<tr>
<td>C</td>
<td>2.51</td>
<td>9.16</td>
<td>7.21</td>
<td>2.62</td>
<td>4.10</td>
<td>8.68</td>
</tr>
<tr>
<td>D</td>
<td>6.63</td>
<td>8.99</td>
<td>8.67</td>
<td>3.29</td>
<td>1.85</td>
<td>5.82</td>
</tr>
<tr>
<td>E</td>
<td>7.13</td>
<td>4.16</td>
<td>6.92</td>
<td>7.01</td>
<td>3.64</td>
<td>8.02</td>
</tr>
<tr>
<td>F</td>
<td>6.17</td>
<td>7.20</td>
<td>5.70</td>
<td>9.34</td>
<td>3.62</td>
<td>5.50</td>
</tr>
<tr>
<td>G</td>
<td>5.19</td>
<td>7.40</td>
<td>5.63</td>
<td>6.75</td>
<td>8.22</td>
<td>6.41</td>
</tr>
<tr>
<td>H</td>
<td>7.77</td>
<td>3.65</td>
<td>6.37</td>
<td>8.02</td>
<td>7.28</td>
<td>3.66</td>
</tr>
</tbody>
</table>

### C

![Image of a 96-well plate with green solutions in each well.](image-url)
Figure 9: WT - CMK D ELISA
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.042</td>
<td>0.689</td>
<td>1.052</td>
<td>0.589</td>
<td>0.234</td>
<td>0.992</td>
</tr>
<tr>
<td>B</td>
<td>1.374</td>
<td>1.369</td>
<td>0.476</td>
<td>1.203</td>
<td>0.363</td>
<td>1.407</td>
</tr>
<tr>
<td>C</td>
<td>0.238</td>
<td>0.191</td>
<td>1.331</td>
<td>0.865</td>
<td>1.238</td>
<td>0.914</td>
</tr>
<tr>
<td>D</td>
<td>1.337</td>
<td>0.606</td>
<td>0.405</td>
<td>0.35</td>
<td>1.277</td>
<td>1.402</td>
</tr>
<tr>
<td>E</td>
<td>0.916</td>
<td>0.876</td>
<td>1.025</td>
<td>1.236</td>
<td>0.565</td>
<td>1.251</td>
</tr>
<tr>
<td>F</td>
<td>1.487</td>
<td>0.517</td>
<td>1.572</td>
<td>0.936</td>
<td>1.323</td>
<td>0.416</td>
</tr>
<tr>
<td>G</td>
<td>0.311</td>
<td>0.174</td>
<td>1.502</td>
<td>0.148</td>
<td>1.063</td>
<td>0.707</td>
</tr>
<tr>
<td>H</td>
<td>1.067</td>
<td>1.322</td>
<td>0.34</td>
<td>0.329</td>
<td>1.402</td>
<td>0.719</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.107</td>
<td>0.101</td>
<td>0.112</td>
<td>0.089</td>
<td>0.112</td>
<td>0.102</td>
</tr>
<tr>
<td>B</td>
<td>0.079</td>
<td>0.086</td>
<td>0.113</td>
<td>0.088</td>
<td>0.115</td>
<td>0.085</td>
</tr>
<tr>
<td>C</td>
<td>0.096</td>
<td>0.105</td>
<td>0.084</td>
<td>0.093</td>
<td>0.083</td>
<td>0.099</td>
</tr>
<tr>
<td>D</td>
<td>0.095</td>
<td>0.107</td>
<td>0.109</td>
<td>0.096</td>
<td>0.098</td>
<td>0.085</td>
</tr>
<tr>
<td>E</td>
<td>0.109</td>
<td>0.088</td>
<td>0.103</td>
<td>0.097</td>
<td>0.111</td>
<td>0.088</td>
</tr>
<tr>
<td>F</td>
<td>0.085</td>
<td>0.114</td>
<td>0.088</td>
<td>0.106</td>
<td>0.093</td>
<td>0.1</td>
</tr>
<tr>
<td>G</td>
<td>0.102</td>
<td>0.098</td>
<td>0.079</td>
<td>0.096</td>
<td>0.089</td>
<td>0.09</td>
</tr>
<tr>
<td>H</td>
<td>0.096</td>
<td>0.086</td>
<td>0.089</td>
<td>0.072</td>
<td>0.077</td>
<td>0.097</td>
</tr>
</tbody>
</table>
### Table B

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9.74</td>
<td>6.82</td>
<td>9.39</td>
<td>6.62</td>
<td>2.09</td>
<td>9.73</td>
</tr>
<tr>
<td>B</td>
<td>17.39</td>
<td>15.92</td>
<td>4.21</td>
<td>13.67</td>
<td>3.16</td>
<td>16.55</td>
</tr>
<tr>
<td>C</td>
<td>2.48</td>
<td>1.82</td>
<td>15.85</td>
<td>9.30</td>
<td>14.92</td>
<td>9.23</td>
</tr>
<tr>
<td>D</td>
<td>14.07</td>
<td>5.66</td>
<td>3.72</td>
<td>3.65</td>
<td>13.03</td>
<td>16.49</td>
</tr>
<tr>
<td>E</td>
<td>8.40</td>
<td>9.95</td>
<td>9.95</td>
<td>12.74</td>
<td>5.09</td>
<td>14.22</td>
</tr>
<tr>
<td>F</td>
<td>17.49</td>
<td>4.54</td>
<td>17.86</td>
<td>8.83</td>
<td>14.23</td>
<td>4.16</td>
</tr>
<tr>
<td>G</td>
<td>3.05</td>
<td>1.78</td>
<td>19.01</td>
<td>1.54</td>
<td>11.94</td>
<td>7.86</td>
</tr>
<tr>
<td>H</td>
<td>11.11</td>
<td>15.37</td>
<td>3.82</td>
<td>4.57</td>
<td>18.21</td>
<td>7.41</td>
</tr>
</tbody>
</table>

### Table C

![Image of a 96-well plate with green and blue solutions]
Figure 10: WT + CMK D ELISA
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.982</td>
<td>0.137</td>
<td>0.163</td>
<td>0.177</td>
<td>0.148</td>
<td>0.156</td>
</tr>
<tr>
<td>B</td>
<td>0.092</td>
<td>0.086</td>
<td>0.095</td>
<td>0.118</td>
<td>0.147</td>
<td>0.131</td>
</tr>
<tr>
<td>C</td>
<td>1.461</td>
<td>0.101</td>
<td>0.124</td>
<td>0.114</td>
<td>0.102</td>
<td>0.449</td>
</tr>
<tr>
<td>D</td>
<td>0.113</td>
<td>0.149</td>
<td>0.11</td>
<td>0.123</td>
<td>0.138</td>
<td>0.128</td>
</tr>
<tr>
<td>E</td>
<td>0.112</td>
<td>0.102</td>
<td>0.113</td>
<td>0.106</td>
<td>0.106</td>
<td>0.223</td>
</tr>
<tr>
<td>F</td>
<td>0.107</td>
<td>0.119</td>
<td>0.656</td>
<td>0.117</td>
<td>0.106</td>
<td>0.114</td>
</tr>
<tr>
<td>G</td>
<td>0.088</td>
<td>0.092</td>
<td>0.101</td>
<td>0.157</td>
<td>0.219</td>
<td>0.109</td>
</tr>
<tr>
<td>H</td>
<td>0.105</td>
<td>0.104</td>
<td>0.087</td>
<td>0.756</td>
<td>0.108</td>
<td>0.953</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.125</td>
<td>0.103</td>
<td>0.134</td>
<td>0.125</td>
<td>0.11</td>
<td>0.095</td>
</tr>
<tr>
<td>B</td>
<td>0.094</td>
<td>0.101</td>
<td>0.098</td>
<td>0.112</td>
<td>0.113</td>
<td>0.109</td>
</tr>
<tr>
<td>C</td>
<td>0.139</td>
<td>0.106</td>
<td>0.104</td>
<td>0.115</td>
<td>0.117</td>
<td>0.114</td>
</tr>
<tr>
<td>D</td>
<td>0.127</td>
<td>0.136</td>
<td>0.113</td>
<td>0.117</td>
<td>0.121</td>
<td>0.149</td>
</tr>
<tr>
<td>E</td>
<td>0.112</td>
<td>0.106</td>
<td>0.127</td>
<td>0.112</td>
<td>0.09</td>
<td>0.1</td>
</tr>
<tr>
<td>F</td>
<td>0.118</td>
<td>0.112</td>
<td>0.118</td>
<td>0.09</td>
<td>0.098</td>
<td>0.103</td>
</tr>
<tr>
<td>G</td>
<td>0.098</td>
<td>0.107</td>
<td>0.119</td>
<td>0.082</td>
<td>0.091</td>
<td>0.114</td>
</tr>
<tr>
<td>H</td>
<td>0.117</td>
<td>0.12</td>
<td>0.117</td>
<td>0.15</td>
<td>0.13</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>A</td>
<td>7.86</td>
<td>1.33</td>
<td>1.22</td>
<td>1.42</td>
<td>1.35</td>
<td>1.64</td>
</tr>
<tr>
<td>B</td>
<td>0.98</td>
<td>0.85</td>
<td>0.97</td>
<td>1.05</td>
<td>1.30</td>
<td>1.20</td>
</tr>
<tr>
<td>C</td>
<td>10.51</td>
<td>0.95</td>
<td>1.19</td>
<td>0.99</td>
<td>0.87</td>
<td>3.94</td>
</tr>
<tr>
<td>D</td>
<td>0.89</td>
<td>1.10</td>
<td>0.97</td>
<td>1.05</td>
<td>1.14</td>
<td>0.86</td>
</tr>
<tr>
<td>E</td>
<td>1.00</td>
<td>0.96</td>
<td>0.89</td>
<td>0.95</td>
<td>1.18</td>
<td>2.23</td>
</tr>
<tr>
<td>F</td>
<td>0.91</td>
<td>1.06</td>
<td>5.56</td>
<td>1.30</td>
<td>1.08</td>
<td>1.11</td>
</tr>
<tr>
<td>G</td>
<td>0.90</td>
<td>0.86</td>
<td>0.85</td>
<td>1.91</td>
<td>2.41</td>
<td>0.96</td>
</tr>
<tr>
<td>H</td>
<td>0.90</td>
<td>0.87</td>
<td>0.74</td>
<td>5.04</td>
<td>0.83</td>
<td>7.27</td>
</tr>
</tbody>
</table>
Figure 11: T140G + CMK X-GG ELISA
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.212</td>
<td>0.082</td>
<td>0.094</td>
<td>0.076</td>
<td>0.076</td>
<td>0.079</td>
</tr>
<tr>
<td>B</td>
<td>0.124</td>
<td>0.086</td>
<td><strong>0.733</strong></td>
<td>0.075</td>
<td>0.075</td>
<td>0.085</td>
</tr>
<tr>
<td>C</td>
<td>0.072</td>
<td>0.096</td>
<td>0.078</td>
<td>0.075</td>
<td>0.088</td>
<td><strong>0.759</strong></td>
</tr>
<tr>
<td>D</td>
<td>0.083</td>
<td>0.088</td>
<td>0.085</td>
<td>0.078</td>
<td>0.146</td>
<td>0.123</td>
</tr>
<tr>
<td>E</td>
<td>0.076</td>
<td>0.079</td>
<td>0.083</td>
<td>0.091</td>
<td>0.131</td>
<td>0.092</td>
</tr>
<tr>
<td>F</td>
<td>0.078</td>
<td>0.077</td>
<td>0.079</td>
<td>0.079</td>
<td>0.096</td>
<td>0.087</td>
</tr>
<tr>
<td>G</td>
<td>0.085</td>
<td>0.078</td>
<td>0.111</td>
<td>0.082</td>
<td><strong>0.57</strong></td>
<td>0.076</td>
</tr>
<tr>
<td>H</td>
<td>0.085</td>
<td>0.1</td>
<td>0.095</td>
<td>0.076</td>
<td>0.082</td>
<td>0.086</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.084</td>
<td>0.078</td>
<td>0.073</td>
<td>0.07</td>
<td>0.072</td>
<td>0.108</td>
</tr>
<tr>
<td>B</td>
<td>0.083</td>
<td>0.071</td>
<td>0.093</td>
<td>0.07</td>
<td>0.068</td>
<td>0.075</td>
</tr>
<tr>
<td>C</td>
<td>0.078</td>
<td>0.073</td>
<td>0.075</td>
<td>0.071</td>
<td>0.075</td>
<td>0.297</td>
</tr>
<tr>
<td>D</td>
<td>0.08</td>
<td>0.086</td>
<td>0.086</td>
<td>0.068</td>
<td>0.078</td>
<td>0.074</td>
</tr>
<tr>
<td>E</td>
<td>0.106</td>
<td>0.077</td>
<td>0.077</td>
<td>0.069</td>
<td>0.074</td>
<td>0.077</td>
</tr>
<tr>
<td>F</td>
<td>0.074</td>
<td>0.095</td>
<td>0.076</td>
<td>0.071</td>
<td>0.077</td>
<td>0.07</td>
</tr>
<tr>
<td>G</td>
<td>0.077</td>
<td>0.094</td>
<td>0.07</td>
<td>0.075</td>
<td>0.07</td>
<td>0.064</td>
</tr>
<tr>
<td>H</td>
<td>0.073</td>
<td>0.073</td>
<td>0.071</td>
<td>0.067</td>
<td>0.065</td>
<td>0.09</td>
</tr>
</tbody>
</table>
### Table B

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2.52</td>
<td>1.05</td>
<td>1.29</td>
<td>1.09</td>
<td>1.06</td>
<td>0.73</td>
</tr>
<tr>
<td>B</td>
<td>1.49</td>
<td>1.21</td>
<td>7.88</td>
<td>1.07</td>
<td>1.10</td>
<td>1.13</td>
</tr>
<tr>
<td>C</td>
<td>0.92</td>
<td>1.32</td>
<td>1.04</td>
<td>1.06</td>
<td>1.17</td>
<td>2.56</td>
</tr>
<tr>
<td>D</td>
<td>1.04</td>
<td>1.02</td>
<td>0.99</td>
<td>1.15</td>
<td>1.87</td>
<td>1.66</td>
</tr>
<tr>
<td>E</td>
<td>0.72</td>
<td>1.03</td>
<td>1.08</td>
<td>1.32</td>
<td>1.77</td>
<td>1.19</td>
</tr>
<tr>
<td>F</td>
<td>1.05</td>
<td>0.81</td>
<td>1.04</td>
<td>1.11</td>
<td>1.25</td>
<td>1.24</td>
</tr>
<tr>
<td>G</td>
<td>1.10</td>
<td>0.83</td>
<td>1.59</td>
<td>1.09</td>
<td>8.14</td>
<td>1.19</td>
</tr>
<tr>
<td>H</td>
<td>1.16</td>
<td>1.37</td>
<td>1.34</td>
<td>1.13</td>
<td>1.26</td>
<td>0.96</td>
</tr>
</tbody>
</table>

### Table C

[Image of a table with columns and rows, possibly depicting some experiment data or analysis results.]
Appendix A: Designing Allosteric Activators of Caspase-3 Based on Maybridge-42 using Docking

I. Introduction

Effector caspases such as caspase-3 are the terminal enzymes which commit the cell to apoptosis. For diseases such as cancer, in which there is too little apoptosis, caspases are promising drug targets. However, the active sites of each of the caspases are very similar, making it difficult to achieve target specificity. Allosteric activation through binding of a drug to a location remote to the active site would achieve greater specificity since locations other than the active site are less conserved among the caspases.

Effector caspases are expressed as zymogens which, under normal circumstances, must become cleaved in a loop called the intersubunit linker to be activated. Upon cleavage, the loops L2 and L2’ formed from the N and C terminal halves of the intersubunit linker rotate out of the dimer interface and form critical interactions in the active site of the enzyme.

Mutation of a key residue at the dimer interface of caspase-3 can lead to activation or inactivation of the enzyme.1 Small molecules have been found which bind to the dimer interface of caspase-3 and inactivate the enzyme by inducing a conformational shift to a zymogen-like form.2 Recently, a small molecule called compound 42 from the Maybridge screening library was found to be an activator of caspase-3 and also binds to the dimer interface.3 In high doses, compound 42 causes a 27-fold increase in procaspase-3 activity. $K_D$ for the binding of the compounds about 4 $\mu$M.
Because of the poor binding affinity of compound 42 and relatively low level of activation, we are interested in using compound 42 as a basis for designing new small molecules that have the potential to be better activators of procaspase-3. We believe our small molecule activator could work by blocking the central cavity so the intersubunit linker cannot bind. Additionally, the small molecule could hold R164 in place in the dimer interface through hydrogen bonding, stabilizing the active site loops in an active conformation.

In-silico docking was used to assess potential compounds for their ability to bind to the dimer interface of caspase-3. One of the compounds that came out of the docking studies, AD58, was synthesized and dose-response was determined. This particular compound was unsuccessful in activating procaspase-3, but others from the docking studies could potentially be tested.

II. Materials and Methods

Twelve compounds designed by Joshua Schipper (JLS 1-12), eighty-three compounds designed by Christine Cade (CEC 13-95), and sixty-three compounds designed by Christine Cade and Alex Deiters (AD 1-63) were screened in-silico for docking to the dimer interface of caspase-3. Although many of the CEC compounds had high docking scores, they were not able to be synthesized by the Deiters group, so the AD compounds were designed in conjunction with Alex Deiters.

Compound 58 (2.9 mg, Figure 2A) was resuspended in 526.7 µL DMSO for a final concentration of 0.01M. Compound 58-OH (4.5 mg, Figure 2B) was resuspended in 805.5 µL DMSO for a final concentration of 0.01M.
Assay buffer contained 150 mM Tris pH 7.5, 50 mM NaCl, 1% sucrose, 10 mM DTT, and 0.1% CHAPS. Procaspace-3 concentration was 100 nM, substrate concentration was 20 µM, and compound concentration was 0, 1, 5, 10, 25, 50, and 100 µM. Both compound 58 and its alcohol precursor 58-OH (Figure 2) were assayed for ability to activate procaspace-3. A positive control with 10 nM caspace-3 and 20 µM substrate was carried out.

III. Results

Compound alone for 58 and 58-OH has little autofluorescence (Figure 3). An increase in fluorescence is seen when procaspace-3, substrate, and compound are mixed. However, the amount of cleaved substrate with and without procaspace-3 is approximately equal. This indicates that the compounds are able to cleave substrate independently of the presence of enzyme. Additionally, there is no discernible dose-response: increasing compound does not significantly increase activity.

The positive controls with 10 nM caspace-3 plus 20 µM substrate (no compound) were lower than expected; only 0.4 µM substrate was cleaved. It is possible that the protein had lost activity in the freezer. The 0 µM compound column acts as a negative control, containing procaspace + substrate without compound. As expected, this number is lower than for the positive control with active caspace-3. Although more controls could have been carried out, the fact that compounds 58 and 58-OH likely cleave substrate independently of procaspace-3 made them unsuitable for further testing.
REFERENCES


FIGURES

Figure 1: Docking Results of Maybridge-42 and Select Compounds From JLS 1-12, CEC Compounds 13-95, and AD 1-63 with WT Caspase-3

JLS compounds designed by Josh Schipper based on Maybridge-42; re-docked by Christine Cade; CEC compounds designed and docked by Christine Cade; AD compounds designed by Alex Deiters and Christine Cade and docked by Christine Cade
Maybridge 42

Y197

R164

3.6Å

- Docking Score: -8 kcal/mol
- Forms H-bonds with Y197 but not R164
• Docking Score: -8.5 kcal/mol
• Forms H-bonds with Y197 and R164
• Slightly shorter than Maybridge 42
• Docking Score: -11 kcal/mol
• Does not form H-bonds with Y197 and R164
• More rigid than and fills cavity as well as Maybridge 42
- Docking Score: -10.3 kcal/mol
- Forms H-bonds with Y197 and R164
- More rigid than and fills cavity almost as well as Maybridge 42
• Docking Score: -8.9 kcal/mol
Figure 2: Compounds 58 and 58-OH
A) Compound 58. B) Compound 58-OH
Figure 3: Dose-Response with Compounds 58 and 58-OH
A) After 1 hour incubation at 25 C (µM substrate cleaved). B) After substrate addition and 15 minutes incubation at 25 C (µM substrate cleaved). C) Graphical representation of B
Appendix B: Molecular Dynamics Simulations of Potential Helix Mutants

I. Introduction

Alpha helices were first described by Pauling et. al in 1950. They were confirmed in 1951 with the x-ray crystal structure of hemoglobin, which contains a mostly helical secondary structure. About 30% of protein residues in general are found in alpha helices.

In alpha helices, the backbone is twisted to allow hydrogen-bonding between the backbone C=O of residue i with the backbone N-H group of i+4. This gives these helices a linear rise of 5.4 Å/turn and 3.6 residues per turn (Figure 1). The side chains of i, i+3, i+4, and i+7 are generally on the same face of the helix. Alpha helices also tend to have an integral number of turns so that the N-capping and C-capping residues are on the same side of the helix.

Several factors affect helix stability: residue preferences for helix interiors and N- and C- terminal positions, capping motifs unique to helix termini, side-chain interactions, metal binding, helix length, and covalent crosslinks. Some potential capping motifs include a capping box, hydrophobic staple, Schellman motif, and \( \alpha_L \) motif. A typical capping box (Figure 2) has a double-reciprocal H-bond between N-cap and N3: the amide of the N-cap H-bonds to the side chain of N3 and the side chain of the N-cap H-bonds to the amide of N3. The best residues for the N-cap are D, N, S, and T. The best residues for N3 are D, E, and Q. A typical hydrophobic staple (Figure 3) is \( N' \) to N4. It often surrounds a capping box (N-cap to N3). In a typical Schellman motif (Figure 4), the amide of C” H-bonds to the carbonyl of C3 and the amide of C’ hydrogen-bonds to carbonyl of C2. C’ is usually a
glycine so it can adopt the left-handed helical conformation. In a typical αL motif (Figure 5), the amide of C’ hydrogen-bonds to the carbonyl of C3.

Residues on the ends of helices have a certain nomenclature: N”–N’–N-cap–N1–N2–N3–N4…C4–C3–C2–C1–C-cap–C’–C”’. The residues N1-C1 have helical φ, and ψ dihedral angles. The N-cap and C-cap transition to non-helical angles. For N1, N2, N3, etc. only C=O groups make (i, i+4) contacts. For C1, C2, C3, etc. only NH groups make (i, i+4) contacts. This leaves free H-bond donor/acceptor groups. Overall, helices have a dipole which is positive on the N-terminal end and negative on the C-terminal end.

Caspases are proteases which commit the cell to apoptosis. As such, they are potential drug targets for diseases such as cancer in which too little cell death occurs, or neurodegenerative disorders in which too much cell death occurs. Because the active site of caspases is highly conserved, it is ideal for drugs to bind to allosteric sites in the protein in order to achieve specificity for a given caspase.

Caspase-3 was found to contain an allosteric site in its dimer interface. The binding of the drugs FICA and DICA inactivated the protein through conversion of the enzyme to a zymogen-like conformation. Mutation of valine 266 in the dimer interface also inactivated the protein, but much more subtle conformational changes occurred. In a 50 ns molecular dynamics simulation of V266H, helix 3 was seen to become destabilized and rotate into the dimer interface, whereas no rotation was seen for the WT protein. Determination of the factors contributing to stability of helix 3 could potentially lead to design of allosteric activators or inactivators of caspase-3.
Caspase-3 helix 3 contains a hydrophobic staple. Although aspartate is a good capping residue, D135 does not function as such in WT caspase-3. I139A and I96A/V134A/I139A mutations were made in silico to remove the hydrophobic staple of helix 3 in attempts to destabilize that helix in the WT protein and mimic V266H. Molecular dynamics simulations (50 ns) were run to determine if the helix was destabilized by the mutations.

The V266H MD simulation showed that the rotated form of the helix actually gains a capping interaction between D135 and K137. Additionally, K138 takes over some of the interactions that had been formerly carried out by K137. Based on the V266H MD simulations, D135A/V266H and K138A/V266H mutations were made in silico to destabilize the inactive helix 3 conformation in the V266H protein and mimic that of the WT protein. Molecular dynamics simulations (50 ns) were run to determine if the helix was stabilized by the mutations since the rotated form would theoretically be less stable than before. In silico mutations and MD simulations were also run for D135P/V266H, D135S/K138E/V266H, F142E/V266H, I96G/V134G/I139G, I139G, K137A/V266H, and N141A, but will not be discussed here.

II. Materials and Methods

Crystals structures of wild-type (WT) or V266H caspase-3 were mutated in PYMOL. Molecular dynamics simulations were performed as previously described. Briefly, 50 ns simulations were performed with GROMACS 4.5, using the Amber99 force field and the TIP3P water model.
III. Results/Discussion

The I139A and I96A/V134A/I139A mutations removed the hydrophobic staple of helix 3 in caspase-3 in an attempt to destabilize the helix (Figures 6 and 7). After 50 ns MD simulations were performed, no significant rotation of helix 3 was seen, indicating that removal of the hydrophobic staple is insufficient to destabilize the helix.

In the V266H mutant, rotation of the helix leads to formation of hydrogen-bonding interactions between D135 and K137, allowing D135 to act as a capping residue (Figure 8). The D135A/V266H mutation removed those hydrogen-bonding interactions in the rotated helix, enabling the enzyme to spend more time in a conformation in which helix 3 is not rotated. No significant rotation was seen in the course of a 50 ns MD simulation (Figure 9).

In the V266H mutant, rotation of the helix also leads to K138 essentially taking the place of K137. In doing so, it assumes some of the interactions formerly carried out by K137. The K138A/V266H mutation removed these interactions, again destabilizing the rotated form of the helix and keeping the helix in an un-rotated form. No significant rotation was seen in the course of a 50 ns MD simulation (Figure 10).

These studies show that disruption of the hydrophobic staple in helix 3 of caspase-3 is not likely to be a good strategy for inactivation of the protein. Capping interactions also do not seem to play a significant role in stabilization of the helix. Disruption of interactions between helix 3 and the rest of the protein seems to be a more effective strategy for inactivation of the protein (for example, in the V266H mutant).
Although in general, rotation of helix 3 is destabilizing, the rotated conformation does form interactions not seen in the un-rotated conformation. In the WT protein D135 is not able to perform a capping role, but in the inactive V266H protein it is able to H-bond to K137. Additionally in the V266H mutant, K138 takes the place K137 once occupied. Disrupting these interactions makes the rotated form less stable and therefore keeps the helix in the un-rotated form. This could potentially increase the activity of the protein. Mutational studies in vitro are needed to test the activity of these mutants.
REFERENCES


**Figure 1: Helix Structure**

α-helices have a linear rise of 5.4 Å/turn. There are 3.6 residues/turn. Side chains of (i, i+3), (i, i+4), (i, i+7) are close together. Helices tend to have an integral # of turns so the N-cap and C-cap are on the same side of the helix.
Figure 2: Typical Capping Box

A typical capping box has a double-reciprocal H-bond between N-cap and N3: the amide of the N-cap H-bonds to the side chain of N3 and the side chain of the N-cap H-bonds to the amide of N3. The best residues for the N-cap are D, N, S, and T. The best residues for N3 are D, E, and Q.
Figure 3: Hydrophobic Staple in WT caspase-3
A typical hydrophobic staple is N to N4. It often surrounds a capping box (N-cap to N3).
Figure 4: Typical Schellman Motif
In a typical Schellman motif, the amide of C" H-bonds to the carbonyl of C3 and the amide of C' H-bonds to carbonyl of C2. C' is usually a glycine so it can adopt the left-handed helical conformation
Figure 5: Typical $\alpha_L$ Motif
In a typical $\alpha_L$ motif, the amide of $C'$ H-bonds to the carbonyl of $C_3$. 
Figure 6: Destabilizing helix 3 by removing one residue of a hydrophobic staple
I139A mutant removes the hydrophobic staple; however, helix 3 is not destabilized. A) WT. B) I139A. C) I139A helix 3 monomer A. D) I139A helix 3 monomer B.
Figure 7: Destabilizing helix 3 by removing three residues of a hydrophobic staple
Triple mutant removes the hydrophobic staple; however, helix 3 is not destabilized. A) WT. B) I96A/V134A/I139A triple mutant. C) Triple mutant helix 3 monomer A. D) Triple mutant helix 3 monomer B.
Figure 8: D135 is a better N-cap in V266H inactive conformation
D135 H-bonds with K137 in 50 ns MD simulation of V266H.
Figure 9: Destabilizing the inactive conformation by removing a potential end cap

The D135A/V266H mutant can no longer form capping interactions with K137. The helix is stabilized. A) D135A/V266H helix 3 monomer A. B) D135A/V266H helix 3 monomer B.
**Figure 10: Destabilizing the inactive conformation by removing interactions with K138**

The K138A/V266H mutant can no longer make interactions with K138 as helix 3 rotates. The helix is stabilized. A) K138A/V266H helix 3 monomer A. B) K138A/V266H helix 3 monomer B.