ABSTRACT

MACIAG, JOSEPH JOHN. The Role of Conserved Water Molecules in Allosteric Modulation of Caspase-3. (Under the direction of Dr. A. Clay Clark).

Caspases mediate apoptosis, a programmed cell death cascade which is a vital component of homeostatic maintenance among metazoans. The activation of the apoptotic caspases results in the irreversible obligation to cell suicide. Additionally, caspases demonstrate regulatory roles in a multitude of processes including: cellular proliferation, cellular differentiation, cytokine processing, macrophage infectious response, neuronal pruning and tissue development, referred to as the adaptive responses. Due to the involvement of caspases in disease it is important to understand how they are regulated. Investigation into the specific mechanisms governing caspase inhibition and activation may provide novel insights during design and rationale for potential regulators.

Cells utilize two general mechanisms to modify caspase activity, through modulating levels of active caspase or through allosteric mechanisms that change the distribution of conformations in the native ensemble. Caspase activity is also affected by post-translational modifications (PTMs), most notably by phosphorylation, where several sites on mature caspases are modified. Provided herein is the examination of the role of conserved water molecules in the allosteric regulation of conformational states in the native ensemble. The positioning of the conserved waters correlates with regions of the protein that are also modified, suggesting that PTMs may fine-tune activity by disrupting the conserved water networks. We used saturation mutagenesis in an allosteric site to test this hypothesis, and we show that the activities of the mutants vary over four orders of magnitude. Our database of 37 high-resolution structures of caspase-3 variants shows that de-solvating the enzyme decreases activity by repartitioning states within the native

ensemble. Overall, the database establishes a means to fine-tune caspase-3 activity over several orders of magnitude through conformational selection that is facilitated by globally dehydrating the native state ensemble. Importantly, the ensemble may provide the cell with the means to reversibly fine-tune caspase activity by utilizing combinations of common and/or unique sites of PTMs.

Recent innovations in genomic editing tools provides a reliable and efficient means to mutate caspaes-3 genes in a cellular or animal model context. These tools provide researchers with the ability to determine how cells modulate a threshold level of activity. Regulation of caspase-3 provides the ability to discern specific caspase involvement in disease and is vital when striving for limited off target effects during disease treatment. It is vital to understand caspase regulation so we may rebalance caspase activity.

The Role of Conserved Water Molecules in Allosteric Modulation of Caspase-3

by

Joseph John Maciag

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

2016

APPROVED BY:

Dr. Robert Rose Committee Chair	Dr. E.S. Maxwell
Dr. Robert M. Kelly	Dr. Michael B. Goshe

DEDICATION

To my mom, Debbie, and my siblings, Mike and Ash for your support and unconditional love throughout my life.

BIOGRAPHY

Joseph J. Maciag was born in Los Angeles, California. He moved to Winton-Salem, North Carolina at the age of seven. He has always been fascinated as well as curious about the mechanisms governing the physical world. He is a life-long Boston Red Sox fan. After high school he decided to pursue a degree in Biochemistry at North Carolina State University. In the fall of 2006, as a freshman, he began working as an undergraduate research assistant in Dr. A. Clay Clark's laboratory. He worked with the Clark lab for the entirety of his undergraduate career. Upon completion of his Bachelor of Science degree in Biochemistry he decided to pursue his Ph.D. in Structural and Molecular Biochemistry under the tutelage of Dr. A. Clay Clark. In the spring of 2015 when the laboratory moved to Arlington, Texas, Joe followed to complete the remainder of his work. After the completion of his Ph.D. he wishes to pursue a post-doctoral position and continue the fight against disease.

ACKNOWLEDGMENTS

I would like to thank Dr. A Clay Clark for his diligent nature and his compassion for his students. Your mentorship has allowed me to become the scientist I am today. Through your leadership I have gained a greater understanding of the physical world. I also would like to thank my doctoral committee Dr. Kelly, Dr. Maxwell, and Dr. Rose. Each of you have promoted my growth as a scientist by driving my research with thought provoking questions and direction during uncertainty. I view each of you as mentors and friends. Dr. MacKenzie you were a great mentor, I would not be the same person I am today without your support and instruction. Countless times you have taught me invaluable lessons in both professional and personal circumstances. I cannot thank you enough. Dr. Swartz thank you for sharing your expertise and answering thousands of my questions.

I would like to thank my family for their support throughout my life. My mother Debbie provided an environment in which intellectual growth was mandatory and worked diligently to provide for myself and siblings. I love you. To Mike thank you for pushing me to be a better person through competition and instruction. Ash, you are the rock of the family, everyone loves you even though we give you a tough time, stay golden.

Additionally I would like to thank the family I chose, my great friends. To Matt, John, Jared, Alex, Will, Sam, Bets, Tab, Mallory, Patrick, Berg, Brent, Meredith, Tony, Ryan, Chad, Ross, Mitch2 and Corey. We are all spread throughout the world but we share an unbreakable bond. Thanks for all the laughs and great times. You are all truly great people.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Chapter 2

CHAPTER 1:

CASPASE REGULATION AND ROLES IN DISEASE

CASPASE REGULATION

Initiator and Effector Caspases

Caspase is an acronym used to describe a family of cysteine-dependent aspartate-specific proteases. Caspases mediate apoptosis, a programmed cell death cascade which is a vital component of homeostatic maintenance among metazoans. The activation of the apoptotic caspases results in the irreversible obligation to cell suicide. Phenotypical features of apoptotic induced cell death include chromosome condensation, membrane blebbing, nuclear fragmentation and the condensation of cytoplasmic organelles (1,2,3,4). The family encompassing caspase enzymes is subdivided into two extensively studied apoptotic subgroups, the initiator (-2,-8,-9 and -10) and effector (-3, -6 and - 7) caspases (Fig. 2). The initiator/effector designation is determined by the entry of an enzyme into the apoptotic signaling cascade and the size of their N-terminal regions. Every caspase member is translated and folded into inactive zymogens which are reliant on activation through various stimuli. Initiator and effector caspases are located in the cytosol but differ in their oligomeric states prior to activation. Initiators exist as inactive monomers while effectors are maintained as inactive dimers. This difference in oligomeric states can be attributed to the hydrophobicity of their respective dimer interfaces. Initiator caspase-8, for example, is observed to have a 1,000 fold less affinity for its homodimer binding partner when compared to effector caspase-3 (K_d of 50μM and 50 nM respectively). Additional variance is observed between the subgroups as initiators possess large N-

terminal prodomain regions which are essential to their activation while effector caspases possess short (≈20-30 residues) prodomains (1,4,5).

Commonalities of all the caspases exist in their protomer structural motifs as well as their shared substrate preference. Caspases adopt a caspase-hemoglobinase fold which is characterized by the folding of the large (p20) and small (p10) subunits into a single functional protomer. This fold is composed of a central beta sheet core consisting of six beta sheets (β1-β6) surrounded by five α-helices (H1-H5) and a short beta strand proximal to the active site (β^1-β^3) (Fig. 3) (75). Every member also displays a preference for the recognition and carboxy terminal cleavage of aspartic acid residues, where the scissle bond is represented by the bond connecting P_1 and $P_{1'}$ (P_4-P_3-P_2-P_1-$P_{1'}$) and P_1 is an aspartic acid residue (Fig. 4). Overlap in substrate preference is so pronounced that the identification of an Inhibitor/Substrate specific to a single caspase still evades researchers (1, 22).

Caspases: Regulation and Activation

Apoptotic caspases are activated through both intrinsic and extrinsic apoptotic pathways. Each pathway relies on the activation of initiator caspases (signal transducers) and culminate with effector caspase activation (signal amplification) (22). Initiators exist as inactive monomers and require proximity induced dimerization via activation platforms. Mammalian recruitment platforms include the death-inducing signaling complex (DISC), apoptosome, inflammasome and the p53-induced protein with a death domain (PIDDosome). Recruitment of initiator caspases to the activation complexes is achieved through their association with adaptor molecules of the recruitment platforms (Fig. 1). Adaptor molecules include Fas-associated protein with death domain (FADD)

which associates with the DISC (extrinsic apoptosis) and Apaf-1 which is an integral component of the apoptosome (intrinsic apoptosis). The initiators associate to recruitment domains via homotypic interactions of their N-terminal prodomains. Initiator caspases of the extrinsic apoptotic signaling cascade (caspases -8 and -10) utilize death effector domains (DEDs) while intrinsic signaling (caspase-9) relies on caspase recruitment domains (CARDs) (Fig. 2). Activation platforms facilitate aggregation of initiators with recruitment domains (24, 25, 26). Upon recruitment to the platform, initiator caspases activate via homodimeric interactions and complete the procession of maturation through cleavage of their N-terminal domains. These "apical" caspases are then free to process the "downstream" effector caspases through the proteolytic cleavage of an aspartic acid connecting the large (p20) and small (p10) subunits of each effector caspase protomer (dimer of protomers) (1,2,3,20,21).

Effector caspases exist as inactive dimers reliant on initiator caspase-dependent proteolytic cleavage and subsequent active site stabilization for activation, they then mature through the removal of N-terminal prodomains (additional cleavage event). The general mechanism for effector caspases-3 and -7 activation is described as an ordering of active site loops. The zymogen possesses an uncoordinated active site in which the catalytic cysteine and active site loops are not well positioned for catalysis. The linker region connecting the p20/p10 (intersubunit linker) subunits contains active site loop2/loop2' (L2/L2'). The intersubunit linker of one protomer is sequestered in the dimer interface and must undergo cleavage to form the fully ordered active site (6,7,8). The cleavage of the intersubunit linker generates L2 (C-terminal end of the p20 subunit) and L2' (N-terminal of the p10 subunit) (6). Upon cleavage the active site loops L2, L2',

L3 and L4 undergo rearrangement. L2' rotates away from the dimer interface to form new contacts with the active site loops L2, L3 and L4 of the opposite monomer. Consecutively, L3 is able to transition from an extended orientation (zymogen) toward the dimer interface where it contributes to the formation of the substrate binding groove. L3 is unable to undergo this transition in the zymogen due to the occupancy of L2/L2' in the dimer interface prior to cleavage (6,7,8). L4 undergoes a 60° rotation simultaneously with movement in L3 (7). A conserved arginine (caspase-3 R164, caspsase-7 R187) adjacent to the catalytic cysteine (L2) is able to intercalate between L3 and β-7 thus stabilizing the orientation of the catalytic cysteine. The stabilized catalytic cysteine is now positioned in the S1 subsite where it is properly positioned for nucleophilic attack of the P1 substrate site. In combination, the ordering of active site loops and catalytic cysteine generate an enzyme which is competent to process substrates (Fig. 5) (6,7,8). It shall be noted that both initiator and effector caspases are fully capable of cleaving substrate upon activation and do not require maturation in order to carry out their physiological roles. Maturation is posited to be an additional post-activation step which further stabilizes the active form of the enzyme. Both initiator and effector caspases, however, require the stabilization of active site loops during their activation, whether by dimerization or intermolecular cleavage (1,2).

In response to apoptotic stimuli effector caspases become activated and ultimately lead to the dismantlement of cellular components though the activation of DNases, inactivation of PARP and the disruption of nuclear laminas (1, 22, 23). Estimations of the numerous substrate targets of the effector caspases (-3,-6, and -7) vary as researchers are unable to clearly distinguish proteins that are non-specifically targeted

between those which are required for the fulfillment of apoptosis. These non-specific targets are commonly referred to as “bystanders”(1, 23). Programed cell death is a tightly regulated process and therefore the activation of caspase enzymes is accordingly. Signaling via the intrinsic and extrinsic pathways of apoptosis is a highly coordinated process in which the terminal objective is to activate effector caspases and the commencement of cellular disassembly.

Extrinsic Apoptosis

Extrinsic apoptotic signaling is initiated by the recognition of ligands and antibody agonists by members of the TNF receptor (TNFR) superfamily (Fig. 1). Apoptotic members of this family include the receptors tumor necrosis factor-α receptor 1 (TNFR1), Fas/CD95/APO-1/TNFRSF6, DR4 and DR5. Stimulation of these transmembrane receptors by tumor necrosis factor-α (TNF-α), FasL or TNF-related apoptosis-inducing ligand (TRAIL (Binding both DR -4 and -5)), respectively, results in receptor oligomerization and subsequently transduces the apoptotic signal across the cytoplasmic membrane (27, 28). All members of the TNFR superfamily contain a cytoplasmic death domain (DD), upon TNFR stimulation Fas-associated protein with death domain (FADD) is recruited to the cytoplasmic tail of the transmembrane receptor. FADD contains a C-terminal DD, DDs from TNFR and FADD interact via homotypic interactions. The complex of death receptor and FADD is known as the DISC. FADD also contains an N-terminal death effector domain (DED) which is used for the recruitment and co-localization of initiator caspases -8 and -10 (also containing N-terminal DEDs) (24, 25, 26, 27, 28). This proximity induction allows for the dimerization of initiator caspase by increasing the localized concentration. The initiator caspases then

become activated and proceed to propagate the signal via cleavage of effector caspases (25,27,28).

Intrinsic Apoptosis

Intrinsic apoptosis or the mitochrondrial/apoptosome dependent pathway induces cell death upon the recognition of metabolic stresses including; ionizing radiation, chemotherapeutic/cytotoxic drug treatment, mitochondrial damage, oxididative stress and excessive DNA damage (Fig. 1) (1, 9, 43, 44). Upon stimulation by the varying extra- and intracellular stimuli the mitochondria mediates a pro-apoptotic signaling cascade. Permeability of the mitochondria is required for the transmittal of an apoptotic signal (1,2 43, 44). Mitochondria permeability is regulated by the Bcl-2 (B-cell lymphoma 2) family of proteins which depending on their expression levels can alter the cellular sensitivity to apoptosis (43, 44). Members of the Bcl-2 family of proteins exhibit pro- and anti-apoptotic influences. Pro-apoptotic proteins Bax and Bak aid in mitochondrial permeability by generating pores in the mitochondrial membrane thus releasing pro-apoptotic proteins. Anti-apoptotic members of the Bcl-2 family; Bcl-2 and Bcl-xl prevent mitochondrial permeability through their inhibitory effects on Bax/Bak (43). Upon mitochondrial membrane pore formation pro-apoptotic mitochondrial proteins, which are otherwise secluded, are released into the cytoplasm. Membrane pore formation releases second mitochondrial activator of caspases (Smac), direct inhibitor of apoptosis binding protein with low pI (Diablo), apoptosis inducing factor (AIF), endonuclease g, high temperature recruitment protein A2 (HTRA2) and cytochrome c (2, 43, 44). Smac/Diablo/HTRA2 offset the inhibitory effects of natural inhibitors of apoptosis (IAPs) while AIF and endonuclease g demonstrate DNA damaging effects (43). The

release of Cytochrome c is required for the activation of the apoptosome (intrinsic recruitment scaffold). Cytochrome c aids in the structural rearrangement of the apoptotic protease activating factor-1 (Apaf-1) which contains an N-terminal CARD domain. Apaf-1 assembles, in a dATP dependent manner, into a multimeric complex called the apoptosome. The structural rearrangement and organization of Apaf-1 recruits caspase-9, providing a scaffold for the dimerization and activation of caspase-9. Activated caspase-9 is then able to process downstream effector caspases thus propagating the apoptotic signal (1,2,9, 43, 44, 51). Additionally, it should be noted that the extrinsic and intrinsic pathways of apoptosis do not exist in a mutually exclusive manner. Cross-talk between the two cascades is evident through the cleavage of BID by caspase-8 (extrinsic) which in turn causes mitochondrial permeability (2).

Non-apoptotic Roles of Caspases

While caspases are most routinely described by their roles during apoptosis they have additionally been shown to regulate a multitude of processes including: cellular proliferation, cellular differentiation, cytokine processing, macrophage infectious response, neuronal pruning and tissue development (1,2,10,12,13,14). The first discovered caspase, ICE (interleukin 1β converting enzyme), hereafter referred to as caspase-1, processes proIL-1β (interleukin 1β) through proteolytic cleavage thereby producing an active Il-1β. While caspase-1 activation and Il-1β excretion have been shown to coincide with hall marks of apoptosis, caspase-1 activation does not commit cells to apoptosis. Additionally caspase-1 aids in macrophage infectious response by regulating phagosome-lysome fusion through the processing of downstream effector caspase-7. In the absence of proper phagosome-lysosome fusion organisms are prone to

sustained infection. Caspase-1, caspase-7 and caspase-8 are vital in the clearance of bacterial infectious agents during macrophage response (15,16,17).

Early nervous system development is dependent on the well documented apoptotic roles of caspases, which involves the reduction of total neuron cells through apoptosis, as well as the recently described non-apoptotic roles. Proper intercellular neuronal signal propagation is maintained through the pruning of dendritic portions of neuron cells. While caspase activity is shown within dendritic sections of neuronal cells it is excluded from the axon and soma. The compartmentalization of caspases allows the cellular utilization of caspases without the procession of apoptosis. Through the use of caspase inhibitors it has been ascertained that caspases display regulatory effects on neural plasticity in regards to learning and the memory of the behaviors learned. The non-apoptotic roles of the caspase family are shown to be intimately involved in the development and maintenance of the nervous system (5).

Caspase-3, regarded as the "primary executioner" of apoptosis, is specifically shown to be activated during cellular proliferation despite the host cell indicating no commitment to apoptosis (10). Proliferative T lymphocytes have not only demonstrated caspase-3 activity in the absence of apoptosis but have been shown to be reliant on its activation for successful proliferation (10,11,13,14). Similarly, erythrocyte differentiation depends on the activation of caspase-3 while resisting apoptosis mediated cell death. It should be noted that caspases also mediate neuronal stem cell differentiation. During erythrocyte maturation nuclear membrane condensation and nuclear extrusion are necessary for the development of terminally differentiated erythrocytes. Inhibition and

reduction of active caspase-3, through either pancaspase inhibitors or siRNA knockdowns coincides with the arrest of erythropoiesis(5,18,19).

As activated caspase-3 has been determined to induce apoptosis yet be able to modulate a variety of imperative cellular processes there must exist a threshold level of caspase-3 activity to which a commitment to apoptosis is irreversible. Therefore, cellular control of caspases is essential in the regulation of homeostasis. The methods employed by the cell to modulate caspase activity include expression of endogenous caspase inhibitors, post-translation modifications and control of the activation and maturation of procaspase through signaling cascades. Recently, emerging evidence has suggested that active caspase levels may be controlled through allosteric modulation resulting in a conformational shift within the native ensemble of caspases, specifically caspase-3 (75).

THE ROLE OF CASPASES IN DISEASE

Neurodegeneration

Caspase activation within neuronal tissue has been implicated as a contributing factor in a number of diseases, namely: Huntington's disease (HD), Parkinson's disease (PD), Amyotrophic Lateral Sclerosis (ALS) and Alzheimer's disease (AD).

Huntington's disease results from the mutation of the huntingtin protein (htt), specifically polygluatamine expansions (29,31,32). The cleavage of polyglutamine expanded htt is known to proceed neural degradation and drive pathogenesis. Caspase-3 is shown to cleave htt in both normal and disease tissue models, suggesting that aberrant caspase-3 activity is not driving the disease state but rather normal caspase-3 activity is generating cytotoxic htt fragments which are detrimental to cell survivability (29,31). Wild type htt functions as a pro-survival protein and is demonstrated to inhibit the

activation of caspase-3. Conversely, mutant htt shows a decreased inhibitory effect with lower binding affinity to caspase-3. Cells undergoing stress are also more susceptible to caspase-3 activation and cell death when htt expression was decreased (32).

Parkinson's disease (PD) is attributed to decreased cognitive ability and motor dysfunction due to the loss of dopamine expressing neurons (dopaminergic cells) (29,33,34). Apoptosis induced by caspase-3 in the substantia nigra is linked with the progression of PD. It has been proposed that the PD neuronal cells expressing caspase-3 have a higher degenerative susceptibility when compared with those of normal tissue (33). PD mouse models indicated that a disruption of caspase-3 sufficiently prevents Parkinson related apoptosis (34).

ALS is a fatal neurodegenerative disease characterized by paralysis and loss of motor neural function(35,36,37). Mouse ALS models indicate that both caspase-1 and caspase-3 are activated in spinal cord tissue while activation of both proteases is lacking in wild-type samples. Elevated levels of caspase-1 activity is also detected in human ALS spinal tissue samples. Additionally pan-caspase inhibitors increase the life expectancy of mice in ALS models (35). Cellular mechanisms which strive to reduce caspase mediated cell death have also been observed in ALS models. BCL-A1, an anti-apoptotic protein which limits the activation of procaspase-3 to caspase-3, is over expressed in ALS mouse models (8). Further studies imply that ALS cells are highly sensitive to stress and therefore activate cell death pathways more readily than that of normal tissue. A hallmark of ALS is the cytoplasmic aggregation of cleaved RNA binding protein TDP-43, TDP-43 can be cleaved by caspase-3. Caspase-3 activation of TDP-43 is hypothesized to be concurrent with the polyglutamine expansion of ataxin 2. Polyglutamine expansion of

ataxin 2 is a risk factor for ALS and is shown to be positively correlated with the amount of activated caspase-3 in motor neurons of ALS patients with polyglutamine expansion when compared to ALS patients expressing wild-type ataxin 2 (37). Caspase-3 activation is therefore consistent with the pathogenesis of ALS (35,36,37).

Alzheimer disease is characterized by the loss of neuronal cells in the basal forebrain, hallmarks include the aggregation of neurofibrillary tangles (NFTs) composed of tau proteins as well as accumulation of amyloid-beta (Aβ). Both NFT and Aβ accrual are consistent with dementia and a decrease in cognition associated with AD. Studies investigating the genesis of both NFTs and Aβs demonstrate activated caspases -3 nad-6 as a driving force in their production while activated caspase-7 possesses the ability to mediate NFT accumulation (38,39,40,41,78). Chung demonstrates that the cleavage of tau by caspase-3 results in neuronal apoptosis (40). Cleavage of the carboxy terminus of tau is a necessity for aggregation into NFTs. Generation of truncated tau is mediated by caspase-3 cleavage at Asp421. Truncated tau is a cytotoxic polymer in neuronal cells (41). Rat models and cultured human cells of AD indicate that Aβ induces neuronal apoptosis and that degeneration does not occur in the absence of tau (41,42). Additional evidence also supports the claims that caspase-3 produces increased levels of Aβ by cleaving amyloid-β precursor protein (APP). Caspase-3 is able to process APP at three distinct Asp residues (D-197, -219 and 720) while D720 cleavage is predominantly detected in neuronal cell culture (42). Caspase-6 is also posited to cleave APP at aspartate 664 which produces a peptide (APP-C31) that induces neuronal death (78). Increased levels of APP are present in motor neuron exhibiting degeneration which leads to an increase in Aβ upon caspase-3 cleavage. It is also hypothesized that caspase initiation

may not be a requirement due to the basal levels of activity of the precursor procaspase-3. The effects of Aβ accumulation are cytotoxic and increase the susceptibility of motor neurons to degradation (42).

Caspases and Cancer

The pathogenesis of cancer is facilitated by cellular deregulation of both differentiation and proliferation (43-49). The most common abnormalities leading to altered differentiation and proliferation are variations in tumor suppressor genes, notably p53 (49). Current cancer therapeutics are aimed at stimulating apoptotic cascades mediated by the caspase family, specifically through intrinsic apoptotic activation. The progression of cancer derives from a cell's ability to escape canonical apoptotic stimuli (43,44,45). Mutations in tumor suppressor genes, both pro- and anti- apoptotic regulatory proteins, the translational preference of cell survival proteins, promoter hypermethylation and mutations within the caspase family directly influence the progression of tumorigenesis (43,44,45).

The altered expression of caspases, inhibitors and activators leads to morphological changes resulting in decreased sensitivity and higher activation threshold responses to apoptotic stimuli (44). The up-regulation of anti-apoptotic regulatory proteins such as c-FLIP, Bcl-2, XIAP and Survivin aid in the cellular capability to resist apoptotic signaling. Similarly, the down-regulation of pro-apoptotic elements such as SMAC/DIABLO, APAF-1, BAX and FADD aid in apoptotic resistance (43,44,46,48,49). Expression levels of specific anti- and pro-apoptotic regulatory proteins vary among different cancer types as well as in the patients participating in the expression profile studies (43).

Numerous cancers exhibit mutations and silencing of initiator caspase-8 and caspase-9 including but not limited to colorectal, neuroblastoma, retinoblastoma, lung and gastric cancers (43, 44). Mutation of caspase-8 results in the premature attenuation of polypeptide synthesis, stop codon alteration, overall decrease in enzymatic activity and the inability of recruitment to the DISC. Caspase-9 mutations similarly are posited to decrease recruitment abilities to the apoptosome. Gastric cancer has been shown to express a mutant form of caspase-9 which lacks a catalytic domain (44). Caspase-9 alterations have also been described as risk factors for lung cancer among tobacco smokers (43). The deregulation of apical caspases bestows an apoptotic resistance to tumor cells by limiting their sensitivity to cytotoxic stimuli as well as limiting their ability to propagate pro apoptotic signaling (43,44).

Activation of effector caspases is the primary goal in many modern cancer therapeutics. While many treatments seek to enhance apoptotic signaling through upstream activation, the ultimate and terminal activation event is that of the effector caspases. The direct activation of effector caspases, specifically caspase-3 are of particular interest due to the elevated levels of procaspae-3 in lung, melanoma, renal and breast cancers. The small molecule activator PAC-1 was identified as an activator of procaspase-3. Pac-1 reverses the regulatory effects elicited by zinc mediated inhibition of caspase-3(48, 76). Putt and colleagues suggest that small molecule activators such as PAC-1 have potential benefits in personalized cancer treatments (48). Interestingly and counterintuitive to current cancer treatments is the observation that apoptotic signaling alone may not be sufficient in the already complex battle against cancer. Reports indicate that the activation of caspase-3 in breast and colorectal cancers coincides with the release

of the mitogen prostaglandin E2 thus facilitating proliferation in nearby tissue. This finding suggests that cells which are sensitive to apoptotic stimulation, through treatment, aid in the proliferation of neighboring, resistant, cells (47). Mutation of caspase-3 has been documented in a variety of cancer types including breast, renal, colorectal, lung, bladder and stomach cancers (43, 50). While caspase-3 mutations have been sporadically detected in tumor samples they are not indicated as the sole contributors of apoptosis resistance. Mutations within the apoptotic cascades often occur simultaneously and cohesively contribute to the overall disturbance in cellular sensitivity to apoptosis (43,44,50).

The data supporting caspase involvement in neural degeneration and cancer described in the literature supports the notion that regulation of caspase activity is a viable target for the longevity of life and attenuation of pathogenesis. Investigation into the specific mechanisms governing caspase inhibition and activation may provide novel insights during design and rationale for potential regulators. Additionally, the ability to regulate specific caspase family members as well as the ability to discern specific caspase involvement in disease is vital when striving for limited off target effects during disease treatment. Taken together, the data support that a comprehensive understanding of the mechanisms of caspase regulation as well as the non-apoptotic consequences of caspase activation have the potential to play pivotal roles in disease control and personalized medications (29-51).

ALLOSTERIC MODULATION OF CASPASES

Introduction to Allostery

Allostery is the propagation of regulatory effects from an effector site to a distal, functional site. Perturbation, resulting from effector site ligand binding, is transmitted throughout a macromolecule by conformational changes. Macromolecules are thus able to communicate between functional domains through allosteric regulation regardless of non-overlapping molecular structure. Allosteric conformational modulation is often expressed as "action at a distance" (1, 57, 59, 60).

Two conflicting models to describe allostery were published over fifty years ago, yet still dominate the field today. Both models strive to describe the allosteric transitions observed between distant yet stereospecific sites. The well categorized haemoglobin serves as the archetype for both models. At the time two states of heamoglobin were known, a low (T) and high (R) affinity state. Each model relies on similar assumptions. Firstly, ligand binding causes a conformational shift in the population of oligomeric proteins and secondly that oligomeric proteins consist of two distinct states, a Tensed (low affinity) and Relaxed (high affinity) state (52, 53, 54, 59, 60, 61). The two models describe the transition between each state in distinct manner yet recent observations reveal that the two models may not be mutually exclusive (56).

MWC

The first model aimed at describing allostery was published by Monod, Wyman and Changeux and has been referred to as the MWC, concerted or pre-existing equilibria model. The MWC model posits that the two states, R and T, exist in equilibrium. The

native oligomeric structure therefore is an ensemble of states rather than one static, low energy structure. Presence of a ligand has the ability to shift the equilibria of the ensemble in favor of either state. As ligand binds the high affinity state (R), the ensemble redistributes and causes a shift in equilibrium toward the high affinity state. This model has been discussed in succeeding literature as a "Lock and Key" mechanism, where ligand binding does not occur until an exact match is encountered (52, 54).

KNF

The second model to describe allosteric modulation and the transition between oligomeric states, each with varying affinity for ligand, was proposed by Koshland, Néméthy and Filmer in 1966. This model is referred to throughout the literature as the KNF, induced fit or sequential model. KNF asserts that the two states do not exist in pre-equilibrium but rather the binding of ligand causes a "sequential" transition, among protomers within the oligomeric protein complex, to a high affinity state. It also assumes that ligand and effector site are not complementary until binding occurs. One can visualize a glove which closes upon close contact with an object. In a similar way the KNF predicts that as ligand/oligomer proximity increases, the effector site is induced to fit the ligand resulting in a competent binding region of the oligomer. Each model describes a method for why an allosteric transition is occurring yet neither describe how the conformational perturbation is being propagated (52, 54, 59, 60, 61).

Structural Contributions in Allostery

Following the proposed methods of allostery by the MWC and KNF models came the contention that allosteric transitions could be analyzed through a purely structural

interpretation. It was observed that caspase-1 displays cooperativity, where substrate binding in one active site leads to improved catalysis in the active site of its' homodimer binding partner. The communication between active sites is mediated through the rearrangement of twenty one hydrogen bonds and nine amino acids per monomer. This pathway connects both active sites through the dimer interface, a known allosteric site amongst the caspase family (71,75). In late 1970, Perutz showed that the T to R state transition of hemoglobin was proposed by a step by step relaxation of constraints governing the T state and thereby shifting the equilibrium towards a relaxed state in which numerous salt bridge bonds were broken (55, 82). It is now understood that a transition from Tensed to Relaxed states is coordinated by disrupting subunit connectivity in haemoglobin (55). This disruption, as Parsegian realized, resulted in a macromolecule which has become more solvent exposed. In the early 1990's it was observed that water acts as an allosteric modulator but since then there has been a dearth of literature focused on allosteric enzymes and their regulation through solvent modulation. It was observed that the T to R transition was accompanied by 60 extra water molecules binding to haemoglobin (62). Furthermore, it is posited that the additional water molecules in the R state was not only essential to oxygen affinity but to the stabilization of the R state, citing the postulation that the states could not exist without solvation. This surveillance of a solvent and its ability to modulate conformational change justifies a thermodynamic contribution to allostery that the purely structural allosteric pathway approaches lack. Parsegian also suggested that more emphasis and experimentation be focused on the ability of ligands to alter the solvation of an oligomer, a suggestion that has until recently been all but ignored (62). Structurally based allosteric paradigms have proved

incomplete in their ability to accurately define conformational transition between states (54,75).

Recently the field of allostery has seen many publications addressing how conformational change is transferred intra-molecularly but lacking in these proposals is one unifying theory capable of predicting the outcome of an allosteric binding event (54, 56). While numerous protein structures are now readily available through the protein data base (PDB), many proteins are represented by either an unbound (apo), bound (holo) forms or both. Intermediate states and sub-states are not abundantly represented. This lack of representation lends to the inability of structural perturbations to be accurately mapped as they are propagated through a macromolecule. (75). While allosteric pathways, contiguous residues which facilitate the transmission of a signal, have been discovered there is also evidence to the contrary. Evidence that include mutations which cause no observable structural changes despite eliciting an allosteric response, the fact the reciprocity is not seen between coupled functional domains and that the same ligand can exhibit both agonism and antagonism within in the same macromolecule. (54, 71). These observations have led to additional thermodynamic modeling to explain allosteric modulation on the macroscopic scale.

EAM

The Ensemble Allosteric Model (EAM) views allosteric propagation dependence on how the ensemble is poised prior to ligand binding, governed by free energy conformational changes, and how strongly two functional domains are coupled. Coupling of functional domains are therefore proposed to “sense” one another, where an energetic change in one facilitates the change via coupling in the other. The EAM provides

thermodynamic rules which dictate how disparate functional sites can effect one another. Additionally EAM proposes an explanation for the agonistic and agonism effects elicited by the same ligand. The question of agonsim vs antagonism for a given ligand is dependent on the sequential ordering of a macromolecules' binding partners as well as the amount of coupled functional domains i.e. if cell receptor H possesses three disparate binding sites labeled A, B and C the order in which they bind to H is important in the stabilization of a given state. So too is how the ensemble is poised prior to binding. If the binding of C is known to cause an agonistic effect at site A when site B is unoccupied, the binding of B prior to that of ligand recognition at C can cause an antagonistic effect. It should be noted that the example above is a modest explanation of the fine tuning of a receptors ability to propagate a signal but illustrates fundamental concepts that must be accounted for prior to drug design and rationale (54,**57**).

It is currently accepted that proteins exists in dynamic fluctuation around an average structure (Fig. 6). The proteins maintain an equilibrium consisting of numerous sub states, where some states are populated more frequently. Nuances in local side chain motions and less subtle secondary structural alterations produce transitions among sub states (75). The dynamic ability of macromolecules to exist simultaneously in several states allows for the "trapping" of a given state by a ligand. Allosteric modulation can therefore stabilize a given state and conformationally shift an ensemble to favor activation or inactivation depending on the ligand and the context in which it binds. (64, 75).

Role of PTMS in Allostery

Cellular maintenance of homeostasis is dependent on the function of signaling cascades which provide the cell a means of adaption to alterations in the environment and varying extracellular stimuli. The ability of a cell to quickly respond to the continually changing extracellular environment is reliant on the adjustment of these signaling pathways. Regulation of signal propagation relies on the allosteric modulation of signal transducers in order to properly augment cellular function (63, 67, 68). One method readily employed by cells to alter a signaling response is through the use of post translational modifications (PTMs). The addition of PTMs to an amino acid side chain provides the cell with a means to alter protein function thereby broadening the repertoire in which the protein is utilized. Protein modification occurs in two distinct manners: Orthosterically and Allosterically. Orthosteric modifications are defined by the addition of PTMs to the active site thereby precluding substrate recognition, as demonstrated by the addition glutathione or nitrosylation to the catalytic cysteine among caspases (63, 67, 70, 75). Allosteric modifications are described as modifications at sites distal from the active site, for example the binding of phosphate at varying non active site residues amongst caspases (discussed further below) (63, 70, 76). PTMs provide a reversible tool to alter cellular response without the necessity of additional protein translation or genomic alteration. It is estimated that proteins have on average five PTM sites and that every protein in the human genome is susceptible to modification (63, 64, 67). The addition of PTMs and the combinatorial effects of multiple modifications can therefore increase the intricacy of a limited amount of protein coding genes.

PTMs elicit both local and long-range structural alterations of the protein creating unique binding sites for protein effector molecules as well as modifying protein function. Modulation of function occurs through the ability of a modification to conformationally shift the population within the native ensemble and increase/decrease the concentration of active vs. inactive states. A redistribution of active vs. inactive states provides the cell with the capability to alter protein function through the disruption of the native state by modifying naturally occurring energy minima (65, 66, 67). For example, Xin postulates that phosphorylation may increase the concentration of low energy conformations by reducing the dynamic complexity of the ensemble. PTMs therefore are able to stabilize a given conformation which causes a shift in the equilibrium of the protein ensemble, favoring the conformation which is a competent binder of a modification (64, 75). Depending on the modification site and the type of modification, proteins can exhibit a myriad of function (64, 65). The function of a protein is a representation of the relative distribution of states within the ensemble and the addition of various combinations of PTMs have the ability to shift the equilibrium of an ensemble through the population of specific states (65, 75).

Allosteric Modulation of Caspases

Allosterically modulating the caspase family of enzymes provides a unique approach in the identification and characterization of novel therapeutics (72). The caspase family displays overlapping substrate specificity which limits the ability of orthosterically (active site targeting) controlling selective member (1, 22, 63, 67, 75). Conversely, allosterically targeting caspases provides the potential for an increase in specificity, a broader range of targeting sites and lower off target effects of therapeutic candidates

(75,76). Prospective uses of allosteric caspase modulators include personalizing medications for cancers, neurodegenerative and auto immune disorders. Understanding how cellular mechanisms elicit allosteric control over the caspases as well as the identification of allosteric sites amongst the caspases is crucial for the design of novel therapeutics.

The Role of the Dimer Interface as an Allosteric Site Caspases-3-7

Small molecule binding and mutational analysis have revealed the dimer interface of effector caspases as an allosteric site. (72, 73, 74, 75). Two small molecules 2-(2,4-dichlorophenoxy)-*N*-(2-mercaptoethyl) acetamide (DICA) and 5-fluoro-1*H*-indole-2-carboxylic acid (2-mercaptoethyl) amide (FICA) are shown to preferentially bind a cysteine within the dimer interface and inhibit substrate turnover (cysteine 264 of caspase-3 and cysteine 290 of caspase-7) (72, 74). The compounds utilize a mercaptoethyl functional group in order to form a covalent attachment to the protein via disulfide bonding (72, 74). DICA and FICA binding to the dimer interface of the effector caspases -3 and-7 alter the catalytic orientation and thereby prohibit substrate binding. Inhibition of caspase-3 is also achieved through dimer interface mutation (β strand 6, β-6) at position 266 (valine 266 or V266). Mutation of V266 to histidine (V266H) results in complete inhibition of caspase-3. The presence of histidine in the dimer interface leads to a rotation of helix-3 (H3) towards the dimer interface which alters the positioning of catalytic residues histidine 121 and cysteine 163. Beta strands β^1-β^3 are perturbed upon H3 rotation and subsequently cause the displacement of the catalytic residues. Catalytic residue and beta strand movements caused by mutation lead to steric clashes in the active site and the preclusion of substrate binding. (73, 75). Point mutations at V266 to all other

amino acids generates a library of caspase-3 activity. Mutants range from complete loss of activity to near wild-type levels of caspase-3. Structural analysis suggests that through dimer interface mutation the enzymes undergo global desolvation resulting in the trapping of a high energy inactive conformation (discussed in detail; chapter2). Conversely, mutation of V266 to glutamine (V266E) in the context of the zymogen, procapase-3, leads to enzyme activation independent of chain cleavage. Annexin V staining of cultured Human Embryonic Kidney (HEK-293A) cells indicate that the procaspase-3 V266E variant effectively induces apoptosis. Activation is hypothesized to occur through the expulsion of the intersubunit linker from the dimer interface thereby permitting active site loop ordering (77). Caspase-7 structural studies also identify the dimer interface as an allosteric site. During activation of caspase-7 by intersubunit cleavage (cite caspase activation portion here) loop L2' can remain bound in the interface resulting in an active site incapable of properly orienting (75).

Zinc Inhibition of Caspases

The regulatory effects elicited by zinc binding are vital to cell survival and apoptosis initiation. It is reported that alterations in intra-cellular zinc concentrations compel a cellular survival or cell death response (76). Zinc binding has an inhibitory effect on the caspases and is shown to bind caspases -3, -6, -7 and -9 (Fig. 7). Hardy shows that Zinc can bind and inhibit members of both the initiator and effector caspases. While caspases -3 , -6, -7 and -9 exhibit the potential for active site binding competency, with respect to zinc, so far only caspases-6 and -9 have characterized exosites which establish a link between zinc and enzyme inhibition. Lysine 36, glutamate 244, histidine 287 and a water molecule were identified as the contributors of zinc coordination in

caspase-6. The exosite is located at the base of helix-5 (H5) and stabilizes an extended helix conformation of helix-3 (H3). The H3 extended form disrupts the short surface beta sheets, β^1-β^3, located N-terminally to H3 and alters the positioning of the catalytic residues. This drastic transition inhibits caspase-6 by conformationally shifting the ensemble to an extended/inactive form (78). The zinc exosite of caspase-9 coordinates zinc with the residues; histidine 224, cysteine 229, cysteine 230, and cysteine 272 at the C-terminal base of H3. It remains unclear whether or not zinc binding of the exosite is sufficient to inhibit caspase-9 in the absence of active site zinc occupation. However, Hardy does remark that the exosite location is at the base of H3 which is shown to undergo drastic transitions in the allosteric inhibition of caspase-6 (78, 79). Caspases-3, and -7 also demonstrate the ability to bind zinc albeit hypothesized to be in an orthorsteric manner as the additional binding sites are not well investigated. It should be noted that caspase-6 is also shown to bind zinc in the active site upon exosite mutation but zinc binding to caspase-6 is preferential to the exosite. The evidence summarized here indicates that the cell uses zinc as an allosteric inhibitor of caspases-6 and -9 and that more information is needed to establish the allosteric role, if any, mediated by zinc in effector caspases -3 and -7(76, 78, 79).

The Role of Phosphorylation in Caspase Allostery

Phosphorylation is the most commonly characterized PTM and not unlike most modifications we know little about how their regulatory effects are transmitted. Caspases have been thoroughly established as kinase substrates yet it is unclear how many of these PTMs contribute to and regulate cellular processes. However, evidence of the

conformational outcome of phosphorylation can be discerned through structural experimentation and phosphomimetics (75, 76, 80, 81).

The structural effects of phosphorylation on caspase-6 are particularly pronounced. Modifications demonstrate the ability to select for both inactive unliganded and "closed" conformations of caspase-6, each displaying unique structural dispositions. As discussed previously zinc can conformationally select for an inactive extended helix (H3) form of caspase-6. The addition of phosphorylation also can trap this unliganded state of the enzyme maintaining the extended helix. The extended helix conformation is thought to be favored prior to ligand binding and that the modification establishes an ensemble unequally shifted towards helix extension and catalytic residue displacement (75, 76). This conformation is characterized by the extension of H-3 by four turns, comprising the surface beta sheet, β^1-β^3. Helix 1 (H1) also undergoes extension thereby displacing the canonical active conformation of loop (L1). Additionally, L2' is bound in the dimer interface and hydrogen bonds to H3 through the interaction of tyrosine 198 (L2') and Q137 (H3). These concerted structural motifs aid in the stabilization of an enzyme which has a compromised ability to process substrate due to the maintenance of catalytic residues (histidine 121 and cysteine 163) in an altered orientation that does not permit proton transfer (75). The 'closed conformation" of caspase-6, which is stabilized through the phosphorylation of serine 257 by ARK5 also exhibits L2' binding in the interface (75, 76, 80). When L2' occupies the dimer interface it precludes the intercalation of active site loop L3. The correct organization of the active site loops is therefore unachievable. While caspase-6 is known to possess more kinase targets, it is

unclear at this time what their individual or collective effects may be on caspase-6 regulation (75).

While the selection of inactive conformations through phosphorylation have been moderately explored for the caspase-6 ensemble, the regulatory effects they elicit to caspase-3 is predominantly unknown. Currently, only speculation into the nature of caspase-3 phosphorylation is available, however, some modification sites exist at regions which are known to regulate other caspase members. For example, caspase-3 is phosphorylated on serine 150, S150 resides in between helices -2 and -3 in a region denoted as turn 9. As discussed previously, caspase family members exhibit extensive structural rearrangements in these helices. H2/H3 alteration is consistent with a conformational shift towards inactivity mediated by modification and mutation. S150 modification implies that caspase-3 may also be controlled through the selection of inactive states. Additional phosphorylation sites have been identified in caspase-3 including: serine 26, threonine 77, threonine 152, threonine 174, serine 176 and threonine 245. While the regulation set forth by theses modification sites are not well established, one can make assumptions based on their location. For instance phosphorylation at sites S26, D174, and D176 may aid in maintaining an increased population of the zymogen procaspae-3 due to the proximity of cleavage sites. Processing of caspase-3 occurs at D9, D28 and D175 so the modification at residues near these cleavage sites may alter the ability of caspase to activate and subsequently mature. Lastly, T152 resides in close proximity to S150 and may modulate caspase activity in a similar manner to that of proposed S150 regulatory effects. It remains unclear as to the importance of either

phosphorylation event or the concerted influence these two residues have on the population shifts of active vs. inactive states within the ensemble.

CONCLUSIONS

Caspases are highly modified enzymes which mediate numerous signaling cascades. They are involved in the progression of disease states and cellular development pathways. Understanding the procession of caspase activation as well as the regulatory mechanisms utilized by the cell to alter their function is imperative in the ablation of disease and the re-establishment of homeostasis. It is vital to determine where caspases are allosterically modulated, how modulation adjusts caspase function and which modification sites are unique to individual caspase members. These criteria could allow for increased therapeutic specificity, decreased off target effects and the expansion of targeting sites. .

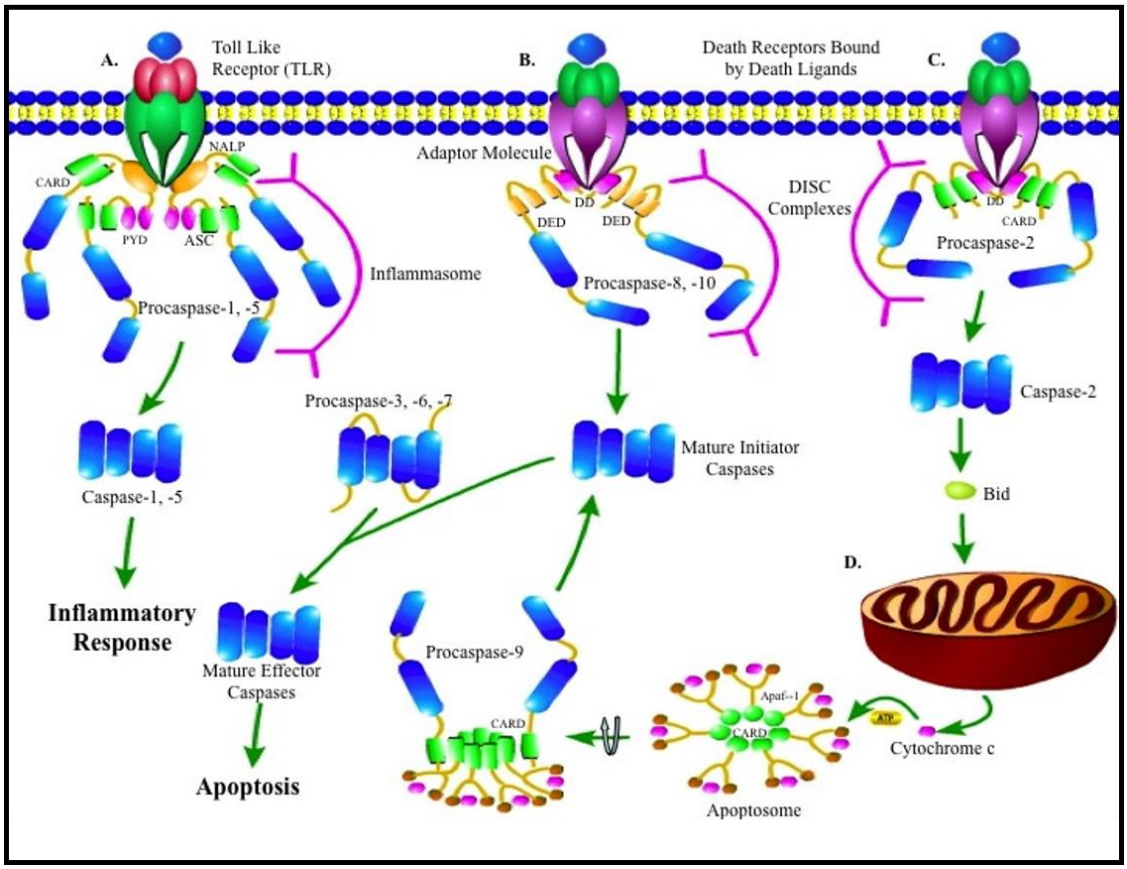


Fig.1 (84) Caspase Pathways

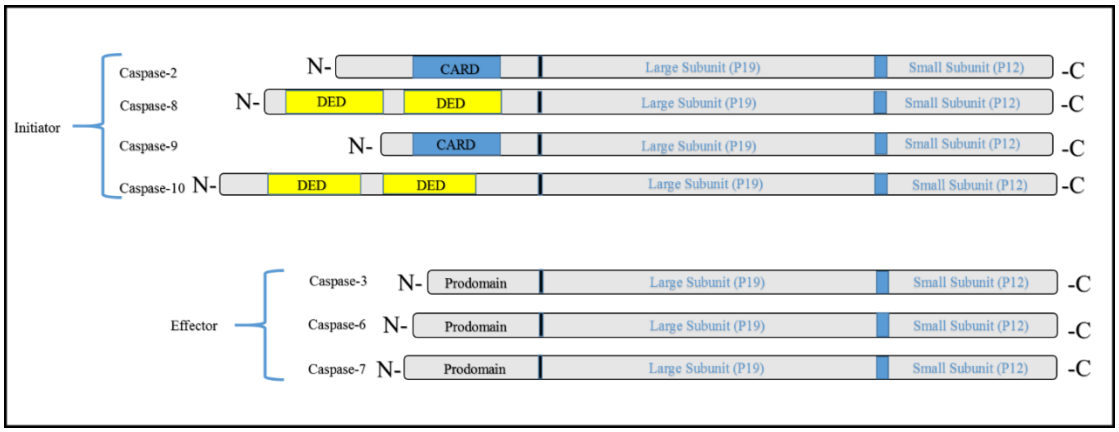


Fig. 2 Apoptotic Caspase Organization

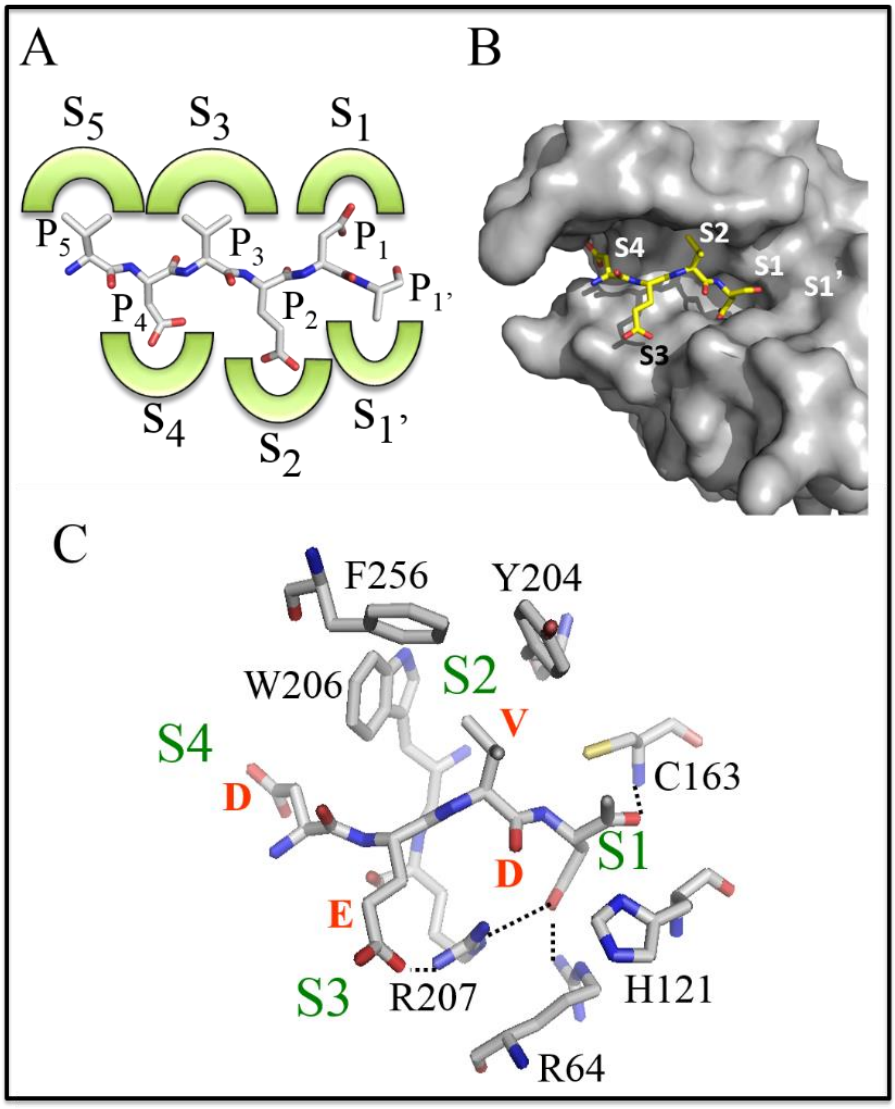


Fig. 3 (75) Active site binding of caspase substrate.

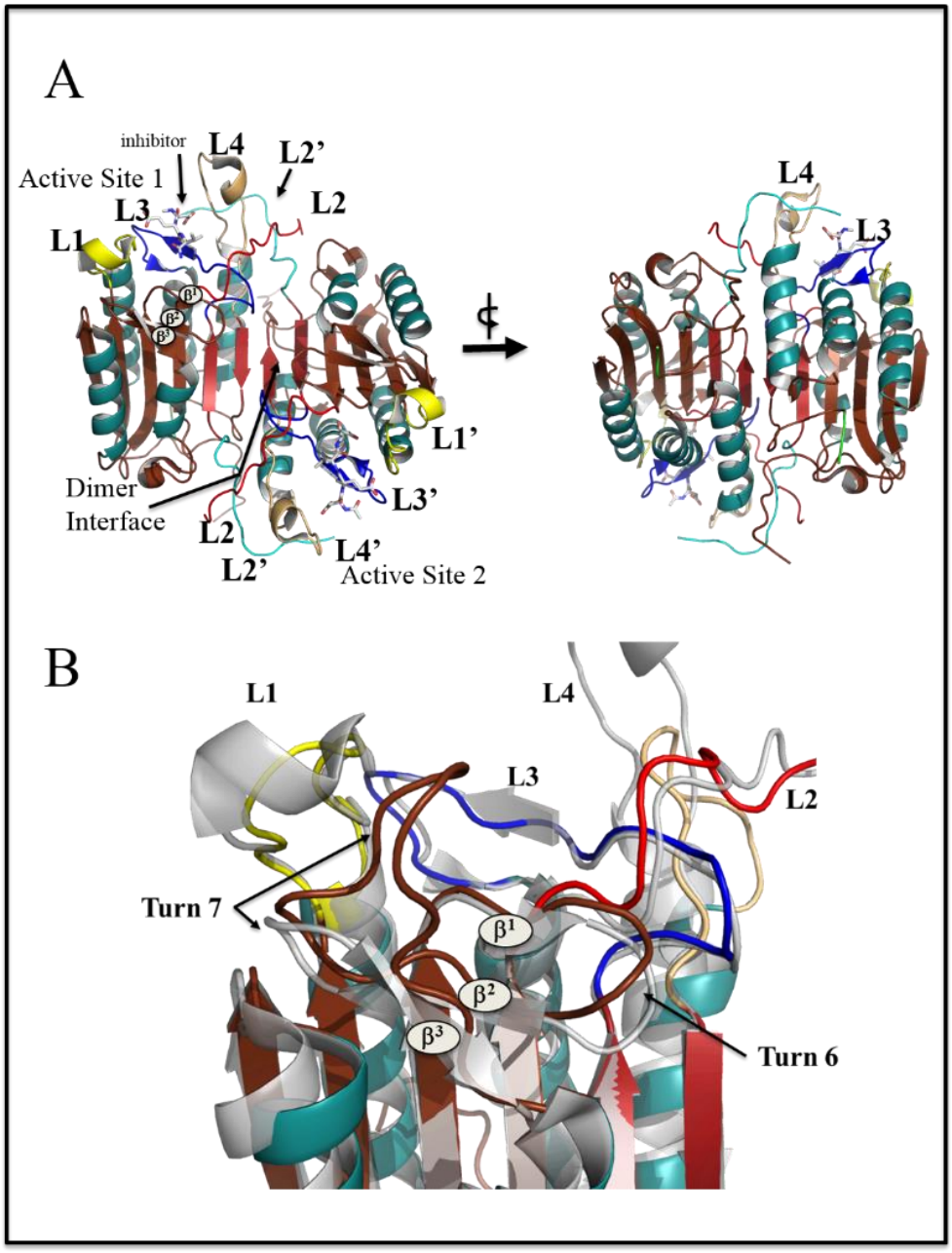


Fig. 4 (75) Structural Representation of Mature Caspase-3

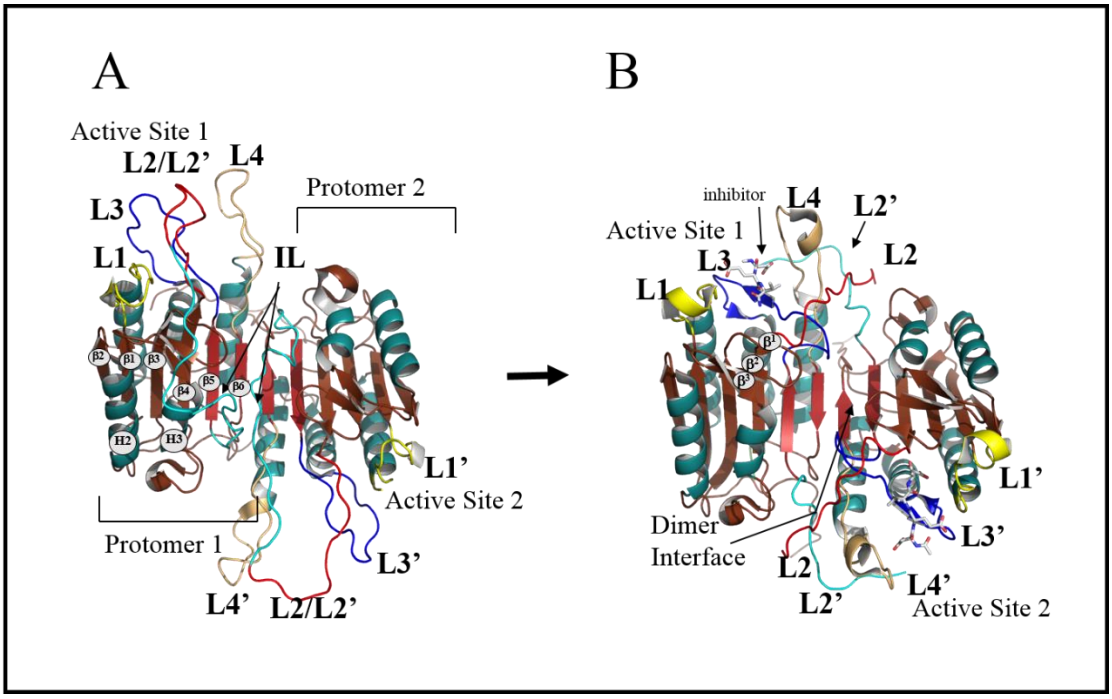


Fig. 5 (75) Structural Representation of the Maturation from Zymogen to Mature Caspase-3

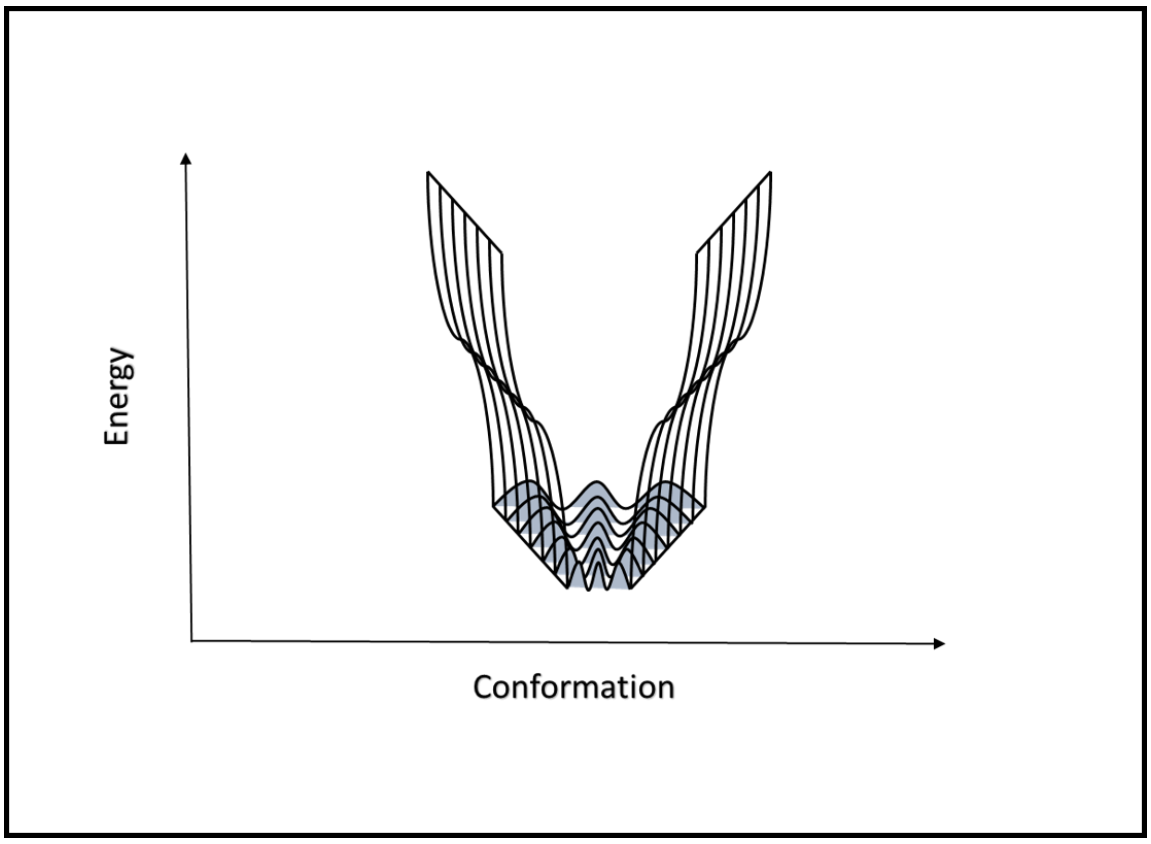


Fig. 6. Energy Landscape of Protein Conformation

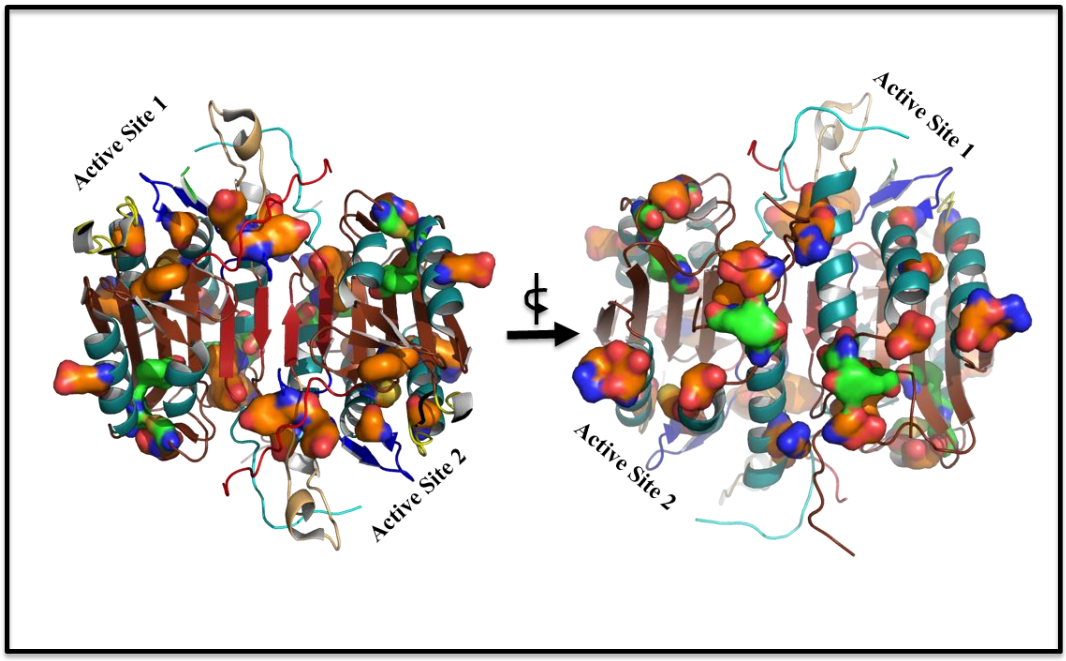


Fig. 7 (75) Post Translational Modifications of Caspases

REFERENCES

1. Pop, C., & Salvesen, G. S. (2009). Human Caspases: Activation, Specificity, and Regulation. The Journal of Biological Chemistry, 284(33), 21777–21781. http://doi.org/10.1074/jbc.R800084200Mechanisms of caspase activation Salvesen Boatright
2. YI, C. H., & JUNYING, Y. (2009). The Jekyll and Hyde Functions of Caspases. Developmental cell, 16(1), 21-34.Molecular mechanisms of caspase regulation during apoptosis Stefan J. Riedl[1] & Yigong Shi[1]
3. Earnshaw, W. C., Martins, L. M., & Kaufmann, S. H. (1999). Mammalian caspases: structure, activation, substrates, and functions during apoptosis. Annual review of biochemistry, 68(1), 383-424.
4. Cohen, G. M. (1997). Caspases: the executioners of apoptosis. Biochemical Journal, 326(1), 1-16.
5. D'amelio M, Cavallucci V, Cecconi F (2010) Neuronal caspase-3 signaling: not only cell death. *Cell Death Differ* 17:1104–1114.
6. Witkowski, W. A., & Hardy, J. A. (2009). L2′ loop is critical for caspase-7 active site formation. *Protein Science : A Publication of the Protein Society*, *18*(7), 1459–1468. http://doi.org/10.1002/pro.151
7. Chai, J., Wu, Q., Shiozaki, E., Srinivasula, S. M., Alnemri, E. S., & Shi, Y. (2001). Crystal structure of a procaspase-7 zymogen: mechanisms of activation and substrate binding. *Cell*, *107*(3), 399-407.

8. Riedl, S. J., Fuentes-Prior, P., Renatus, M., Kairies, N., Krapp, S., Huber, R., ... & Bode, W. (2001). Structural basis for the activation of human procaspase-7. *Proceedings of the National Academy of Sciences*, *98*(26), 14790-14795.
9. Miura, M. (2012). Apoptotic and nonapoptotic caspase functions in animal development. Cold Spring Harbor perspectives in biology, 4(10), a008664.
10. Zeuner, A., Eramo, A., Peschle, C., & De Maria, R. (1999). Caspase activation without death. Cell Death & Differentiation, 6(11).
11. Wilhelm, S., Wagner, H., & Häcker, G. (1998). Activation of caspase-3-like enzymes in non-apoptotic T cells. European journal of immunology, 28(3), 891-900.
12. Wang, J. Y., & Luo, Z. G. (2014). Non-apoptotic role of caspase-3 in synapse refinement. Neuroscience bulletin, 30(4), 667-670.
13. Saligrama, P. T., Fortner, K. A., Secinaro, M. A., Collins, C. C., Russell, J. Q., & Budd, R. C. (2014). IL-15 maintains T-cell survival via S-nitrosylation-mediated inhibition of caspase-3. Cell Death & Differentiation, 21(6), 904-914.
14. Kennedy, N. J., Kataoka, T., Tschopp, J., & Budd, R. C. (1999). Caspase activation is required for T cell proliferation. The Journal of experimental medicine, 190(12), 1891-1896.
15. Akhter, A., Gavrilin, M. A., Frantz, L., Washington, S., Ditty, C., Limoli, D., ... & Marsh, C. B. (2009). Caspase-7 activation by the Nlrc4/Ipaf inflammasome restricts Legionella pneumophila infection. PLoS Pathog, 5(4), e1000361.
16. Vandenabeele, P. (1999). The proteolytic procaspase activation network: an in vitro analysis. Cell death and differentiation, 6, 1117-1124.

17. Edelmann, B., Bertsch, U., Tchikov, V., Winoto-Morbach, S., Perrotta, C., Jakob, M., ... & Schütze, S. (2011). Caspase-8 and caspase-7 sequentially mediate proteolytic activation of acid sphingomyelinase in TNF-R1 receptosomes. The EMBO journal, 30(2), 379-394.

18. Carlile, G. W., Smith, D. H., & Wiedmann, M. (2004). Caspase-3 has a nonapoptotic function in erythroid maturation. Blood, 103(11), 4310-4316.

19. Keerthivasan, G., Wickrema, A., & Crispino, J. D. (2011). Erythroblast enucleation. Stem cells international, 2011.

20. Slee, E. A., Adrain, C., & Martin, S. J. (2001). Executioner caspase-3,-6, and-7 perform distinct, non-redundant roles during the demolition phase of apoptosis. Journal of biological Chemistry, 276(10), 7320-7326.

21. Fulda, S., & Debatin, K. M. (2006). Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy. Oncogene, 25(34), 4798-4811.

22. Shahzidi, S., Brech, A., Sioud, M., Li, X., Suo, Z., Nesland, J. M., & Peng, Q. (2013). Lamin A/C cleavage by caspase-6 activation is crucial for apoptotic induction by photodynamic therapy with hexaminolevulinate in human B-cell lymphoma cells. Cancer letters, 339(1), 25-32.

23. Crawford, E. D., & Wells, J. A. (2011). Caspase substrates and cellular remodeling. Annual review of biochemistry, 80, 1055-1087.

24. Park, H. H., Lo, Y. C., Lin, S. C., Wang, L., Yang, J. K., & Wu, H. (2007). The death domain superfamily in intracellular signaling of apoptosis and inflammation. Annual review of immunology, 25, 561.

25. Wolf, B. B., & Green, D. R. (1999). Suicidal tendencies: apoptotic cell death by caspase family proteinases. Journal of Biological Chemistry, 274(29), 20049-20052.

26. Kaufmann, S. H., & Hengartner, M. O. (2001). Programmed cell death: alive and well in the new millennium. Trends in cell biology, 11(12), 526-534.

27. Lavrik, I. N., & Krammer, P. H. (2012). Regulation of CD95/Fas signaling at the DISC. Cell Death & Differentiation, 19(1), 36-41.

28. Ren, Y. G., Wagner, K. W., Knee, D. A., Aza-Blanc, P., Nasoff, M., & Deveraux, Q. L. (2004). Differential regulation of the TRAIL death receptors DR4 and DR5 by the signal recognition particle. Molecular biology of the cell, 15(11), 5064-5074.

29. D'amelio M, Cavallucci V, Cecconi F (2010) Neuronal caspase-3 signaling: not only cell death. *Cell Death Differ* 17:1104–1114.

30. McIlwain, D. R., Berger, T., & Mak, T. W. (2013). Caspase functions in cell death and disease. *Cold Spring Harbor perspectives in biology*, *5*(4), a008656.

31. Wellington, C. L., Ellerby, L. M., Gutekunst, C. A., Rogers, D., Warby, S., Graham, R. K., ... & Gafni, J. (2002). Caspase cleavage of mutant huntingtin precedes neurodegeneration in Huntington's disease. *The journal of neuroscience*, *22*(18), 7862-7872.

32. Zhang Y, Leavitt BR, van Raamsdonk JM, et al. Huntingtin inhibits caspase-3 activation. *The EMBO Journal*. 2006;25(24):5896-5906. doi:10.1038/sj.emboj.7601445

33. Hartmann, A., Hunot, S., Michel, P. P., Muriel, M. P., Vyas, S., Faucheux, B. A., ... & Evan, G. I. (2000). Caspase-3: a vulnerability factor and final effector in apoptotic death of dopaminergic neurons in Parkinson's disease. *Proceedings of the National Academy of Sciences*, *97*(6), 2875-2880.

34. Yamada, M., Kida, K., Amutuhaire, W., Ichinose, F., & Kaneki, M. (2010). Gene disruption of caspase-3 prevents MPTP-induced Parkinson's disease in mice. *Biochemical and biophysical research communications*, *402*(2), 312-318.

35. Li, M., Ona, V., Guégan, C., Chen, M., Jackson-Lewis, V., Andrews, L., . . . Friedlander, R. (2000). Functional Role of Caspase-1 and Caspase-3 in an ALS Transgenic Mouse Model. *Science, 288*(5464), 335-339. Retrieved from

36. Iaccarino, C., Mura, M. E., Esposito, S., Carta, F., Sanna, G., Turrini, F., ... & Crosio, C. (2011). Bcl2-A1 interacts with pro-caspase-3: implications for amyotrophic lateral sclerosis. *Neurobiology of disease*, *43*(3), 642-650.

37. Hart, M. P., & Gitler, A. D. (2012). ALS-associated ataxin 2 polyQ expansions enhance stress-induced caspase 3 activation and increase TDP-43 pathological modifications. *The Journal of Neuroscience*, *32*(27), 9133-9142.

38. Chu, J., Li, J. G., Joshi, Y. B., Giannopoulos, P. F., Hoffman, N. E., Madesh, M., & Praticò, D. (2015). Gamma secretase-activating protein is a substrate for caspase-3: implications for Alzheimer's disease. *Biological psychiatry*, *77*(8), 720-728.

39. Jarero-Basulto, J. J., Luna-Muñoz, J., Mena, R., Kristofikova, Z., Ripova, D., Perry, G., ... & Garcia-Sierra, F. (2013). Proteolytic cleavage of polymeric tau

protein by caspase-3: implications for Alzheimer disease. *Journal of Neuropathology & Experimental Neurology*, *72*(12), 1145-1161.

40. Chung, C. W., Song, Y. H., Kim, I. K., Yoon, W. J., Ryu, B. R., Jo, D. G., ... & Mook-Jung, I. H. (2001). Proapoptotic effects of tau cleavage product generated by caspase-3. *Neurobiology of disease*, *8*(1), 162-172.
41. Binder, L. I., Guillozet-Bongaarts, A. L., Garcia-Sierra, F., & Berry, R. W. (2005). Tau, tangles, and Alzheimer's disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, *1739*(2), 216-223.
42. Gervais, F. G., Xu, D., Robertson, G. S., Vaillancourt, J. P., Zhu, Y., Huang, J., ... & Clarke, E. E. (1999). Involvement of caspases in proteolytic cleavage of Alzheimer's amyloid-β precursor protein and amyloidogenic Aβ peptide formation. *Cell*, *97*(3), 395-406.
43. Ghavami, S., Hashemi, M., Ande, S. R., Yeganeh, B., Xiao, W., Eshraghi, M., ... & Los, M. (2009). Apoptosis and cancer: mutations within caspase genes. *Journal of medical genetics*, *46*(8), 497-510.
44. Philchenkov, A., Zavelevich, M., Kroczak, T. J., & Los, M. J. (2004). Caspases and cancer: mechanisms of inactivation and new treatment modalities. *Experimental oncology*, *26*(2), 82-97.
45. Wiita AP, Ziv E, Wiita PJ, et al. Global cellular response to chemotherapy-induced apoptosis. Cravatt B, ed. *eLife*. 2013;2:e01236. doi:10.7554/eLife.01236.
46. Kania, J., Konturek, S. J., Marlicz, K., Hahn, E. G., & Konturek, P. C. (2003). Expression of survivin and caspase-3 in gastric cancer. *Digestive diseases and sciences*, *48*(2), 266-271.

47. Galluzzi, L., Kepp, O., & Kroemer, G. (2012). Caspase-3 and prostaglandins signal for tumor regrowth in cancer therapy. *Oncogene*, *31*(23), 2805-2808.

48. Putt, K. S., Chen, G. W., Pearson, J. M., Sandhorst, J. S., Hoagland, M. S., Kwon, J. T., ... & Doerge, D. R. (2006). Small-molecule activation of procaspase-3 to caspase-3 as a personalized anticancer strategy. *Nature Chemical Biology*, *2*(10), 543-550.

49. Olsson M, Zhivotovsky B. Caspases and cancer. *Cell Death and Differentiation*. 2011;18(9):1441-1449. doi:10.1038/cdd.2011.30.

50. Soung, Y. H., Lee, J. W., Kim, S. Y., Park, W. S., Nam, S. W., Lee, J. Y., ... & Lee, S. H. (2004). Somatic mutations of CASP3 gene in human cancers. *Human genetics*, *115*(2), 112-115.

51. Newton, K., & Strasser, A. (2000). Cell death control in lymphocytes. *Advances in immunology*, *76*, 179-226.

52. Monod, J., Wyman, J., & Changeux, J. P. (1965). On the nature of allosteric transitions: a plausible model. *Journal of molecular biology*, *12*(1), 88-118.

53. Koshland Jr, D. E., Nemethy, G., & Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits*. *Biochemistry*, *5*(1), 365-385.

54. Structural basis of allostery Hilser, V. J., Wrabl, J. O., & Motlagh, H. N. (2012). Structural and energetic basis of allostery. *Annual review of biophysics*, *41*, 585.

55. Perutz, M. F. (1970). Stereochemistry of cooperative effects in haemoglobin: haem–haem interaction and the problem of allostery.

56. Silva, D. A., Bowman, G. R., Sosa-Peinado, A., & Huang, X. (2011). A role for both conformational selection and induced fit in ligand binding by the LAO protein. *PLoS Comput Biol*, *7*(5), e1002054.

57. Hilser, V. J. (2010). An ensemble view of allostery. *Science*, *327*(5966), 653-654.

58. Amor, B., Yaliraki, S. N., Woscholski, R., & Barahona, M. (2014). Uncovering allosteric pathways in caspase-1 using Markov transient analysis and multiscale community detection. *Molecular BioSystems*, *10*(8), 2247-2258.

59. Nussinov, R., & Tsai, C. J. (2013). Allostery in disease and in drug discovery. *Cell*, *153*(2), 293-305.

60. Motlagh, H. N., Wrabl, J. O., Li, J., & Hilser, V. J. (2014). The ensemble nature of allostery. *Nature*, *508*(7496), 331-339.

61. Nussinov, R., Ma, B., & Tsai, C. J. (2014). Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophysical chemistry*, *186*, 22-30.

62. Colombo, M. F., Rau, D. C., & Parsegian, V. A. (1992). Protein solvation in allosteric regulation: a water effect on hemoglobin. *Science*, *256*(5057), 655.

63. Nussinov, R., Tsai, C. J., Xin, F., & Radivojac, P. (2012). Allosteric post-translational modification codes. *Trends in biochemical sciences*, *37*(10), 447-455.

64. Xin, F., & Radivojac, P. (2012). Post-translational modifications induce significant yet not extreme changes to protein structure. *Bioinformatics*, *28*(22), 2905-2913.

65. Tsai, C. J., & Nussinov, R. (2014). A unified view of "how allostery works". *PLoS Comput Biol*, *10*(2), e1003394.

66. Radivojac, P., Iakoucheva, L. M., Oldfield, C. J., Obradovic, Z., Uversky, V. N., & Dunker, A. K. (2007). Intrinsic disorder and functional proteomics. *Biophysical journal*, *92*(5), 1439-1456.

67. Deribe, Y. L., Pawson, T., & Dikic, I. (2010). Post-translational modifications in signal integration. *Nature structural & molecular biology*, *17*(6), 666-672.

68. Nussinov, R., Ma, B., Tsai, C. J., & Csermely, P. (2013). Allosteric conformational barcodes direct signaling in the cell. *Structure*, *21*(9), 1509-1521.

69. Nussinov, R., & Tsai, C. J. (2013). Allostery in disease and in drug discovery. *Cell*, *153*(2), 293-305.

70. Pejaver, V., Hsu, W. L., Xin, F., Dunker, A. K., Uversky, V. N., & Radivojac, P. (2014). The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Science*, *23*(8), 1077-1093.

71. Datta, D., Scheer, J. M., Romanowski, M. J., & Wells, J. A. (2008). An allosteric circuit in caspase-1. *Journal of molecular biology*, *381*(5), 1157-1167.

72. Häcker, H. G., Sisay, M. T., & Gütschow, M. (2011). Allosteric modulation of caspases. *Pharmacology & therapeutics*, *132*(2), 180-195.

73. Walters, J., Schipper, J. L., Swartz, P., Mattos, C., & Clark, A. C. (2012). Allosteric modulation of caspase 3 through mutagenesis. *Bioscience reports*, *32*(4), 401-411.

74. Hardy, J. A., Lam, J., Nguyen, J. T., O'Brien, T., & Wells, J. A. (2004). Discovery of an allosteric site in the caspases. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(34), 12461-12466.

75. Clark, A. C. (2016). Caspase Allostery and Conformational Selection. *Chemical reviews*.

76. Dagbay, K., Eron, S. J., Serrano, B. P., Velazquez-Delgado, E. M., Zhao, Y., Lin, D., ... & Hardy, J. A. (2014). A multi-pronged approach for compiling a global map of allosteric regulation in the apoptotic caspases. *Methods in enzymology*, *544*, 215.

77. Walters, J., Pop, C., Scott, F. L., Drag, M., Swartz, P., Mattos, C., ... & Clark, A. C. (2009). A constitutively active and uninhibitable caspase-3 zymogen efficiently induces apoptosis. *Biochemical Journal*, *424*(3), 335-345.

78. Velázquez-Delgado, E. M., & Hardy, J. A. (2012). Zinc-mediated allosteric inhibition of caspase-6. *Journal of Biological Chemistry*, *287*(43), 36000-36011.

79. Huber, K. L., & Hardy, J. A. (2012). Mechanism of zinc-mediated inhibition of caspase-9. *Protein Science*, *21*(7), 1056-1065.

80. Velázquez-Delgado, E. M., & Hardy, J. A. (2012). Phosphorylation regulates assembly of the caspase-6 substrate-binding groove. *Structure*, *20*(4), 742-751.

81. Cao, Q., Wang, X. J., Liu, C. W., Liu, D. F., Li, L. F., Gao, Y. Q., & Su, X. D. (2012). Inhibitory mechanism of caspase-6 phosphorylation revealed by crystal structures, molecular dynamics simulations, and biochemical assays. *Journal of Biological Chemistry*, *287*(19), 15371-15379.

82. Silva, M. M., Rogers, P. H., & Arnone, A. (1992). A third quaternary structure of human hemoglobin A at 1.7-A resolution. *Journal of Biological Chemistry, 267*(24), 17248-17256.

83. MacKenzie, S. H., & Clark, A. C. (2012). Death by caspase dimerization. In *Protein Dimerization and Oligomerization in Biology* (pp. 55-73). Springer New York.

CHAPTER 2

A Tunable Allosteric Library of Caspase-3 Identifies Coupling Between Conserved Water Molecules and Conformational Selection

Joseph J. Maciag[1,2], Sarah H. MacKenzie[2], Matthew B. Tucker[2], Joshua L. Schipper[2,4], Paul Swartz[2] and A. Clay Clark[1,2,3*]

[1]Department of Biology, University of Texas at Arlington, Arlington, TX 76019, USA

[2]Department of Molecular and Structural Biochemistry and [3]Center for Comparative Medicine and Translational Research, North Carolina State University, Raleigh, NC 27695, USA

[4]Current address: Institute for Genome Science, Duke University, Durham, NC 27708, USA

ABSTRACT

The native ensemble of caspases is described globally by a complex energy landscape where the binding of substrate selects for the active conformation, while targeting an allosteric site in the dimer interface selects an inactive conformation that contains disordered active site loops. Mutations and post-translational modifications stabilize high-energy inactive conformations, with mostly formed, but distorted, active sites. In order to examine the interconversion of active and inactive states in the ensemble, we used DRoP to analyze 4,995 waters in 15 high-resolution (<2.0 Å) structures of wild-type caspase-3, resulting in 450 clusters with the most highly conserved set containing 145 water molecules. The data show that regions of the protein that contact the conserved waters also correspond to sites of post-translational modifications, suggesting that the conserved waters are an integral part of allosteric mechanisms. To test this hypothesis, we created a library of nineteen caspase-3 variants through saturation mutagenesis in a single position of the allosteric site of the dimer interface, and we show that the enzyme activity varies over four orders of magnitude. Altogether, our database consists of 37 highresolution structures of caspase-3 variants, and we demonstrate that each 10-fold decrease in activity correlates with a loss of 23 conserved water molecules. The data show that the activity of caspase-3 can be fine-tuned through globally de-solvating the active conformation within the native ensemble, providing a mechanism for cells to repartition the ensemble and thus fine-tune activity through conformational selection.

Keywords: Allostery, saturation mutagenesis, conformational selection, native ensemble, protein solvation, protein structure, protein dynamics

SIGNIFICANCE STATEMENT

The interconversion of states in the caspase-3 native ensemble is affected by binding of ligands that either stabilize or destabilize active site loops. It is not clear how the ensemble is regulated in cells aside from modulating levels of endogenous caspase inhibitors. We describe a library of caspase-3 variants with activities that vary over four orders of magnitude and show that removal of conserved water molecules may provide a strategy to design novel allosteric inhibitors that globally destabilize the active conformation within the ensemble. Our results suggest that posttranslational modifications fine-tune caspase activity by disrupting conserved water networks, and our database provides an approach to examine caspase signaling in cells by modifying caspase-3 activity while simultaneously maintaining endogenous enzyme levels.

Abbreviations: WT, wild-type; Ac-DEVD-AFC, acetyl-Asp-Glu-Val-Asp-7-amino-4trifluoromethylcoumarin; Ac-DEVD-CMK, acetyl-Asp-Glu-Val-Asp-chloromethyl ketone; MD molecular dynamics; H-bond, hydrogen bond; protomer, large and small subunit obtained by processing a monomer of procaspase-3; L1-L4, active site loops 1 through 4; PTM, posttranslational modification

INTRODUCTION

Caspase function in cell development and cell death results from a continuum of enzyme activity, where an as-yet-undefined activity threshold is required for cell death. At sub-threshold levels, caspase activity is important for a variety of physiological reactions (referred to as adaptive responses), from remodeling the cytoplasm (1), cell differentiation (2), neuron pruning (3), receptor endocytosis (4), macrophage function (5) to development of eye lens (6) and inner ear (7). The roles of caspases in apoptosis are well-known, but their roles in adaptive responses are less clear, particularly in regard to how cells set the threshold of caspase activity to limit apoptosis while ensuring sufficient activity for signaling and differentiation.

Cells utilize two general mechanisms to modify caspase activity, through modulating levels of active caspase or through allosteric mechanisms that change the distribution of conformations in the native ensemble, although the two are not mutually exclusive. Levels of caspase-3 are controlled by cleavage of the inactive zymogen to yield a dimer of protomers (Fig 1A) (8, 9), and this process is responsive to several signaling pathways, such as transient expression of the BadBax cascade (10) or phosphorylation of the zymogen (Fig 1B) (11). Alternatively, IAPs (inhibitor of apoptosis proteins) affect levels of active caspase-3 by direct inhibition through active-site binding or through ubiquitination, leading to proteasome degradation (12–15). The interplay between these signaling pathways modulates total activity by changing the amount of active caspase-3.

Caspase activity is also affected by post-translational modifications (PTMs), most notably by phosphorylation, where several sites on mature caspases are modified (11, 16). Depending on the site of modification, the PTM inactivates the caspase either by directly

interfering with substrate binding or through allosteric mechanisms, so controlling kinase or phosphatase activities allows for a level of control over caspases. Consequently, PTMs link caspase activity to cellular signaling reactions that are important in the adaptive responses (17–19).

The mature caspase energy landscape is an ensemble of states consisting of active and inactive conformations (Fig 1B, brackets), where the free energies of the inactive states relative to that of the active state vary among caspases. Both caspases-2 and -3, for example, contain high-energy inactive states (Fig 1B, state 2) while caspases-1, -6, and -7 contain low-energy, socalled "closed-loop" inactive states with disordered active site loops (Fig 1B, state 3) (20–25). Conformational selection by ligand binding to the active or allosteric sites affects the partitioning of active *versus* inactive conformations, so the activity of the caspase reports on the relative distribution of active and inactive states within the ensemble (23, 25), providing a mechanism to fine-tune activity. An allosteric site in the dimer interface (Fig 1A,B), for example, prevents active site loop formation when bound to small molecule allosteric inhibitors (26, 27), demonstrating a common allosteric mechanism through steric clashes with active site loops, and selecting the "disordered loop" conformation (Fig 1B, state 3). Although the active (Fig 1B, state 1) and disordered-loop (Fig 1B, state 3) conformations are well described structurally and appear to be well-populated thermodynamically, much less is known about the high-energy inactive state(s) within the caspase-3 ensemble, collectively referred to here as state 2 (Fig 1B). Higher energy inactive states have been observed in caspase-2 when inhibited by DARPins and demonstrate that small changes in positioning the catalytic groups are sufficient to allosterically inhibit the enzyme (28). In contrast, caspase-6

undergoes very large transitions when allosterically inhibited by zinc, where a short surface β-sheet (β^1-β^3, Fig 1A) undergoes a coil-tohelix transition that extends helix 3 and disrupts the catalytic histidine and cysteine (29). We showed previously that a high-energy inactive conformation in the caspase-3 ensemble results from transient rotation of helix 3 toward the dimer interface, reducing interface volume by ~800 $Å^3$ (Fig 1A, H3), and resulting in disruption of the catalytic dyad, H121 and C163, increased mobility of active site loop 1 (Fig 1A, L1), and a narrower S1' binding pocket (22, 23). Collectively, the data for caspases-2 and -3 suggest that conformational selection in the native ensemble decreases enzyme activity by populating high-energy states with mostly intact active sites but with disrupted catalytic groups. Importantly, the ensemble may provide the cell with the means to reversibly fine-tune caspase activity by utilizing combinations of common and/or unique sites of PTMs.

Here, we examine the role of conserved water molecules in the allosteric regulation of conformational states in the native ensemble. Starting with a database of 15 high resolution structures of caspase-3, we used the DRoP (Detection of Related Solvent Positions) analysis (30) to identify 145 highly conserved water molecules. The positioning of the conserved waters correlates with regions of the protein that are also modified, suggesting that PTMs may fine-tune activity by disrupting the conserved water networks. We used saturation mutagenesis in an allosteric site to test this hypothesis, and we show that the activities of the mutants vary over four orders of magnitude. Our database of 37 high-resolution structures of caspase-3 variants shows that de-solvating the enzyme decreases activity by repartitioning states within the native ensemble. Overall, the database establishes a means to fine-tune caspase-3 activity over several

orders of magnitude through conformational selection that is facilitated by globally dehydrating the native state ensemble.

RESULTS AND DISCUSSION

Global Analysis of Caspase-3 Solvation. Protein solvation mediates folding, stabilizes native structures, and is critical for enzyme function (31). Conversely, disruption of protein-water interactions may result in unfavorable changes in stability or dynamics (31, 32). Aside from the well-known changes in solvation for the conformational transitions of hemoglobin (33), relatively little is known about the role of water in allosteric transitions (34–37). We performed a global analysis of solvation using the recently described DRoP program (30), which analyzes positions of water molecules in related X-ray crystal structures in order to cluster the waters in crystallographically-related positions. One also determines the level of conservation for each water molecule as well as conserved contact positions on the protein. Our database consisted of

15 high-resolution (<2 Å) structures of wild-type caspase-3 and contained 4,995 water molecules

(Supplemental Table S1, Fig 2A). Following clustering and conservation analysis, we identified 450 unique clusters (Fig 2B), which we grouped based on degree of conservation in the database: >93%, 80-93%, 66-79%, and <66% (Fig 2C, blue, cyan, orange, red, respectively). The conserved waters, defined as the group present in 14 of the 15 structures (>93%), comprise 145 water molecules and are the focus of the subsequent analyses. The results show that the conserved waters interact with active site loops and

helices on or near the protein surface (Fig 2C, blue spheres; Fig 3A), particularly in the dimer interface, while the least conserved waters

(<66%) are located in the substrate-binding groove and chain termini (Fig 2C, red spheres).

Formation of the intact active site involves movement of active site loop 3 (L3) from a solvent-exposed position in which the loop is mostly disordered, to the active position, where the loop forms the substrate-binding groove along the protein surface (Fig 1A) (8, 9). Although the substrate-binding pocket is "dry," that is, it does not contain conserved waters, several conserved waters are observed to connect the C-terminal side of L3 with helices 1 and 4 on the opposite side of the central cavity (supplemental Fig 1B and described below). In addition, several conserved waters interact with the N-terminal side of L3, the so-called "elbow-loop" region (20) near the central cavity of the dimer interface (supplemental Fig 1A), and the conserved waters connect L3 with the "loop-bundle" of L2, L2' and L4 (see Fig 1A), which form a network of hydrogen bonds that stabilize the active conformation (38, 39). Steric clashes between the "elbow-loop" and L2' in the "closed-loop" conformation is a primary source of allosteric regulation in maintaining the disordered-loop inactive conformation (Fig 1B, state 3). The conserved waters at the N- and C- termini of L3 likely stabilize the active conformation of the loop through extensive hydrogen bonding. Identification of the conserved waters near L3 of caspase-3 is also consistent with an allosteric network in caspase-1, where side-chains in this region are known to change H-bonding patterns in the transition from inactive to active states

(40). The DRoP analysis also shows that many conserved waters bind to helices 1, 4, and 5 (Fig 3A, blue regions). At present, no protein effectors have been observed to bind in this region of caspase-3. Interestingly, however, these helices are known to contain several sites of modification in caspases, including phosphorylation (11) and glutathionylation (41), as well as a zinc-binding site (29). The potential linkage between modifications at these sites and allosteric regulation of enzyme activity is currently not known.

In addition to the active site loops, the C-terminal region of helix 3 (Fig 3A, H3) also interacts with several conserved waters (supplemental Fig 1C). Flexibility in this helix affects the population distribution in the native ensemble, where rotation of the helix toward the dimer interface narrows the S1' binding pocket and disrupts the catalytic C163 and H121, thus abolishing activity (22, 23). Notably, the N-terminus of helix 3, as well as three short β-strands near the active site (β^1-β^3), which bind to conserved waters (supplemental Fig 1D), undergoes extensive rearrangements in the transition from the procaspase-3 zymogen to the mature caspase3 (see Fig 7 below) (24). In addition, the same region in caspase-6 undergoes a coil-to-helix transition, which increases the distance between the catalytic histidine, located on β^1, and the catalytic cysteine, located at the base of active site loop 2 (L2). The extended helix form of caspase-6 is stabilized by zinc binding to an allosteric site near helix 5 (29). The resolution of caspase-6 structures is not sufficient to include in the DRoP analysis with caspase-3, so it is not clear whether similar conserved waters are present; however, if caspase-6 has conserved waters near β^1-β^3 and Turn 6, as observed in caspase-3 (supplemental Fig 1D), then the sites would be abolished upon transition to the extended helix conformation.

We further parsed the set of 145 conserved waters into three groups based on hydrogen bond contacts, namely surface, channel, or buried, and the three groups contain 104, 14, and 27 waters, respectively (Fig 3B). The conserved surface water molecules (Fig 3D) are distributed throughout the protein, although notably the substrate-binding groove appears "dry" because the active site waters are less conserved. The conserved channel water molecules are found primarily near helix 3 and connecting helices 1 and 4 (Fig 3E), while buried water molecules (Fig 3F) connect several active site loops as well as several loops near helix 3 and helix 4.

Interestingly, the buried water molecules are distributed similarly to the PTMs described previously (42). Aside from studies of zinc binding to caspase-6 (29) and phosphorylation of helix 5 in caspase-6 (43, 44), little is known of the mechanisms by which PTMs or metal binding affect enzyme activity. The positioning of the conserved buried water molecules in caspase-3 suggests that the PTMs may affect partitioning of the native ensemble by disrupting conserved water networks.

Saturation Mutagenesis in an Allosteric Site Provides a Library of Caspase Variants with a Broad Range of Activities. In order to test the hypothesis that the inactive conformations in the native ensemble (Fig 1B) have altered solvation, we examined changes in the conserved waters resulting from mutations in the allosteric site within the central cavity of the dimer interface. We chose this site because, as described previously (26, 27), the binding of allosteric inhibitors to this site prevents active site loop formation via steric clashes with L3. The allosteric site of the central cavity is considered a common allosteric site in caspases (42), so it may report on a common mechanism for regulation of caspase activity, as opposed to a unique site that may report on a mechanism specific to

caspase-3. In addition, two mutations of V266 at the center of the allosteric site suggested context-dependent changes in activity. For example, mutation of V266 to histidine abolished enzyme activity, primarily due to destabilizing helix 3 (22, 23, 45), while mutation of V266 to glutamate decreased activity only 10-fold in the mature caspase but resulted in a substantial increase in activity of the zymogen, most likely due to expulsion of a linker from the dimer interface, allowing the active conformation to form in the zymogen (45, 46). Finally, other amino acids in or near the allosteric site have been mutated, and several highresolution structures are available for the variants (supplemental Table S4).

Using a saturation mutagenesis approach, we changed V266 to all other amino acids, resulting in seventeen new mutants at this site, since V266H and V266E were described previously. Activity measurements for the 20 proteins show that the substitution of V266 with glutamate is unique in increasing activity of the zymogen because all other mutations at this site had little effect on zymogen activity (Fig 4A, supplemental Table S2). With the exception of V266H (previously described (45)), V266Y, and V266N, all of which had no activity, the substitutions at V266 in the zymogen resulted in changes between 0.1- and 3-fold that of wildtype procaspase-3, with an average activity of $6.7x10^2$ $M^{-1}s^{-1}$ (excluding the three inactive variants and V266E).

In contrast to results for the zymogen, we observed large effects on the activity of mature caspase-3 when V266 was substituted with other amino acids (Fig 4B). With the exceptions of V266H (described previously (45)) and V266P, both of which exhibited no activity, the activities of the mutants varied nearly four orders of magnitude, with V266F exhibiting the lowest measurable activity. The mutations resulted in increased K_M values

of up to ~10-fold, where the average K_M was 23.2 μM for this data set, and large decreases in k_{cat} of up to ~100-fold (supplemental Table S2), with the combined changes resulting in a large distribution of activities. As described previously for the V266E and V266H variants (45), there were no changes in dimer stability, so the changes in enzyme activity were not due to lower dimer formation. In addition, there was no correlation in enzyme specificity (Fig 4C), k_{cat} (Fig 4D), or K_M (Fig 4E) with changes in side-chain volume upon substitution of V266. Likewise, there was no apparent correlation with side-chain chemistry, since changes in activity did not correlate with hydrophobicity, aromatic, small polar, or charged groups of side-chains.

Structural and Dynamic Studies Show Flexibility in the Allosteric Site Disrupts Conserved Waters and Propagates to the Active Site. In order to further examine changes in enzyme activity for the V266 mutant dataset and potential changes in the conserved waters, we determined the structures of 13 of the 17 new variants by X-ray crystallography. Noting that structures of the V226H and V266E variants were described previously (22, 46), we were therefore successful in obtaining a nearly complete structural description of the saturation mutagenesis dataset at V266, where only V266G, V266P, V266T, and V266R did not crystallize (Supplemental Table S3). Of the 13 new V266X structures, 12 structures had resolution >2Å and were used in the subsequent DRoP analysis (described below).

During formation of the active conformation of caspase-3, R164, from active site loop 2 (L2), intercalates between Y197 (β-strand 6) and P201 (“elbow-loop” of L3) in the dimer interface (20). These are the movements that are prevented through binding of

allosteric effectors in the central cavity (26, 27). In wild-type caspase-3, six water molecules, three from each protomer, form a hydrogen-bonding network across the dimer interface and stabilize R164 and R164' (from the second protomer) (Fig 5A). The water-mediated hydrogen-bonding network across the dimer interface is important for positioning active site loop 2 and C163 for catalysis (22). Using the DRoP analysis, we identified two of the three water molecules as conserved (>93%) (Fig 5A: Wat40, Wat50), while the third water is present in >80% of the structures in our database of wild-type caspase-3 (Fig 5A: Wat177). Our structures of the V266 variants show that substitution of valine with small side chains (A, S, C) generally retains the H-bonding network in the interface (Supplemental Fig S2). In contrast to the small side-chains, substitution of V266 with larger side chains generally disrupted the H-bonding network by displacing the conserved waters (Wat40, Wat50, and Wat177) in the interface (Fig 5).

Five of the V266 variants showed two conformations for the side-chain – V266S (Supplemental Fig S2B), V266D (Supplemental Fig S3C), V266F (Fig 5B), V266Y (Fig 5C), and V266W (Fig 5D). The two conformations can be characterized generally based on their positions relative to R164. In the first conformation, the bulk of the side-chain is positioned close to R164; in the second conformation, the bulk of the side-chain is positioned away from R164 and closer to Y195 on β-strand 6. In the V266E variant (Supplemental Fig S3A), the larger sidechain (compared to valine) displaces the conserved waters but maintains the hydrogen-bonding network in the interface, while the side-chain of aspartate, which is shorter than that of glutamate, is observed in two conformations, with longer hydrogen bonds between the waters and the side-chain (Supplemental Fig S3C). The asparagine side-chain is locked between these two

conformations (Supplemental Fig S3D), as is the histidine side-chain (Fig 5E), and disrupts the conserved waters. Overall, either the changes in hydrogen-bonding in the interface or increased side-chain dynamics, or both, may explain the lower activity in V266S, V266D and V266N compared to V266E (Fig 4B).

In the case of V266I, the isoleucine is observed in conformation 1, where the bulk of the side-chain is near R164, while the leucine side-chain in V266L is observed in conformation 2, near Y195 (Supplemental Fig S4A-D). The positioning of I266 results in a shift of Y197 and of Y195, on β-strand 5, away from β-strand 6, suggesting steric clashes in the interface. In contrast, the leucine side-chain forms close van der Waals contacts with L266' of the second protomer, with no observed shifts in Y195 or Y197 (Supplemental Fig S4C-D). Although the V266I variant retains the water network in the interface, the less optimal hydrophobic contacts across the interface as well as potential steric clashes with the neighboring β-strand may explain the lower activity ($1.6x10^3$ $M^{-1}sec^{-1}$) compared to the V266L variant ($9.0x10^3$ $M^{-1}sec^{-1}$). Finally, while the side-chains for F266 (Fig 5B), Y266 (Fig 5C), and W266 (Fig 5D) are observed in two conformations, the V266Y variant exhibits significantly higher activity than the two other mutants (Fig 4B). This is most likely due to formation of a hydrogen bond between the hydroxyl group of Y266 and the side-chain of R164. Overall, the 15 V266X mutants show no large structural changes that affect the active sites, so substitution of V266 in the dimer interface results in localized structural changes that affect conserved waters in the interface, side-chain fluctuations at 266, and residues on the neighboring β-strand 5. Because all of the structures contain an active site inhibitor, the data show that the

mutations have little global effects on the ground-state active conformation observed in the crystals (Fig 1A, state 1).

We showed previously that the lower activity of mutants coupled to V266H correlated to increased dynamics of several active site loops (22, 23). In order to determine whether other V266 variants also affected active site loop dynamics, we performed molecular dynamics simulations for 50 ns on six mutants – V266Y, V266F, V266D, V266E, V266Q, and V266N – because of the differences in enzyme activity and structural features described above. Data for wild-type caspase-3 (22) provide the baseline for comparison with the V266 mutants and show that the largest fluctuations are observed in two active site loops (L1 and L4) and Turn 6, which contains E123 and E124 (Supplemental Fig S5A-B). The positioning of Turn 6 is important for activity because it contains residues that form part of the oxyanion hole during catalysis as well as residues that contribute to the hydrogen-bonding network of the dimer interface.

The MD simulations for the V266 variants show that the fluctuations are generally within 12Å of those observed for wild-type caspase-3 (Supplemental Fig S5C). Here, the data are reported as ΔRMSF ($RMSF_{Mut} – RMSF_{WT}$), so values greater than zero show larger fluctuations in the mutant compared to those of wild-type caspase-3, while values less than zero reflect larger fluctuations in wild-type caspase-3 compared to those of the mutant. Overall, the data show that active site loops L1, L3, and part of L4 are more flexible in the mutants. Comparisons of fluctuations in V266D (Fig 6A) to those in V266E (Fig 6B) and in V266F (Fig 6C) to those in V266Y (Fig 6D) show increased fluctuations in Turn 6 that result in new, albeit transient, interactions between E123, in

Turn 6, and the catalytic H121. In wild-type caspase-3, H121 fluctuates between two conformations during the simulation, which likely reflects its movements during proton transfer of the catalytic reaction. In these conformations, the H121 side-chain fluctuates between ~7Å and ~12Å from E123. In the case of the V266D and V266F variants, movements in Turn 6, along with the fluctuations in H121, result in shorter distances between H121 and E123, between 3.2Å and 4.1Å, so the electrostatic interactions between H121 and E123 may affect the catalytic activity by decreasing the rate of proton transfer. In addition, the position of the surface β-sheet (β^1-β^3) and Turn 6 is more similar to that of the inactive zymogen than to the active conformation (Fig 7). In the active enzyme, Turn 6 is in the "down" conformation (Fig 7, blue), where E124 is positioned to H-bond with R164 in the central cavity of the dimer interface (see also Fig 5A). In the "up" conformation observed in the zymogen (Fig 7, orange), Turn 6 is rotated away from the central cavity such that E124 is solvent-exposed and the surface β-sheet, β^1-β^3, is not well-formed. In the V266 variants, fluctuations in Turn 6 displace the catalytic H121 by 2.7Å (Fig 7A), while the catalytic cysteine is moved 1.7Å from the active position (Fig 7B). In Figure 7, representative data for V266D are shown in yellow. In addition, all variants showed increased fluctuations in helix 3, which, as described previously for V266H, also affect the positions of the catalytic groups through movements propagated through the surface β-sheet, β^1-β^3 (22, 23). Together, the molecular dynamics data show that the lower activity of some V266 variants may result from fluctuations in Turn 6, helix 3, and active site loop 2, culminating in displacement of the catalytic groups.

Lower Activity Correlates with Changes in Solvation. We performed the DRoP analysis on the V266 variants for which we obtained X-ray crystal structures of >2Å

resolution (Supplemental Table S3). In order to expand the database, we also included 24 additional caspase-3 variants for which we previously determined enzyme activity and high-resolution structures, providing a database of 37 caspase-3 variants containing a total of 10,736 water molecules (Supplemental Table S4). Following DRoP analysis, we determined which water molecules within the set of 145 conserved waters were absent in each mutant, and the data are presented as total number of conserved waters displaced for each mutant. We note that all three classes of the conserved waters (surface, channel, buried) were affected by the mutations, with conserved surface waters demonstrating the largest changes (Fig 3C). The database of 37 variants provides a range of enzyme activities, nearly four orders of magnitude (Fig 4F), and the DRoP analysis shows that the change in enzyme activity correlates with a loss of conserved waters, where the lower catalytic efficiency (Fig 4F) was due to lower k_{cat} (Fig 4G) and higher K_M (Fig 4H) values. Each point in Figure 4 (panels F-H) represents one mutant in the caspase-3 database, where an average of 27.9 waters are displaced per mutant, with a broad range between zero (wild-type) and 70 waters. Fits of the data show that when the 145 conserved waters are present (no change in the waters), values for k_{cat}, 0.65 sec^{-1} (Fig 4G, range of 0.42–1.04 sec^{-1}), and K_M, 9.7 µM (Fig 4H, range of 9.3–10.2 µM), yield an enzyme specificity of $6.7x10^4$ $M^{-1}sec^1$, which is similar to that obtained from a fit of the data in Figure 4F, $1.65x10^5$ (range of $1.36x10^5$ – $2.01x10^5$) $M^{-1}sec^{-1}$, and agrees well with values previously determined for wild-type caspase-3 (Supplemental Table S2). Overall, the data for the mutant library show that each 10fold decrease in activity (Fig 4F) is accompanied by the loss of approximately 23 conserved water molecules. We note that the broad range of values in Figures 4F-4H suggest that the data may reflect the heterogeneity of the ensemble and report on multiple inactive conformations rather than a

single discrete state as represented in Figure 1B (state 2). For example, an improved understanding of inactive conformations characterized by increased dynamics of Turn 6 *versus* those with increased dynamics of helix 3, as well as the dehydration associated with each state, may allow one to parse data such as those shown in Figure 4 to resolve multiple inactive conformations.

Changes in the conserved water molecules also report on potential "hot-spots" on the protein where water networks may be sensitive to conformational changes or to ligand binding. This is analogous to the solvent mapping of Ras, where binding of multiple small organic compounds provides "maps" of potential ligand binding sites (30). Our data show that, overall, there is not a single defined pathway between the allosteric site of the dimer interface and the active site of caspase-3 because conserved waters in multiple regions of the protein were affected. Of the 145 conserved waters, 68 waters (47%) were not affected by the mutations, as they changed in fewer than 10% of the variants. In addition, only 11 waters changed in >60% of the variants, and they are generally found at chain termini. Of the remaining waters, three regions of the protein show significant changes in the conserved waters, where 20-60% of the structures report disrupted water networks: the helix 3-dimer interface cavity and two areas of the face containing helices 1, 4, and 5 (Fig 8A). First, of 21 conserved water molecules that interact with the helix 3-β^1-β^3Turn 6 region, 13 of those waters are disrupted in the mutant database. For example, several water molecules near Turn 6 form part of an extensive hydrogen-bonding network between the dimer interface, Turn 6, and the N-terminus of helix 3. Five water molecules in this network are disrupted in the caspase-3 variants (Fig 8B), which may result in the increased flexibility in this loop observed in MD simulations (Fig 6) and movements toward the "up" position observed in the inactive zymogen (24) (Fig 7). Second, 15 of 24

conserved waters that bind to helices 1 and 4 are disrupted in the mutant database. In one region of the helix-1,-4,-5 face, near the active site (Fig 8C, upper panel), the waters form extensive hydrogen bonds between the two helices and with residues in the active site. The second region includes the C-terminal ends of helices 1 and 4 (Fig 8C, lower panel). Several amino acids in this region are known to be modified on caspases, including glutathionylation of caspase-3 (41), but it is not clear how modifications in the Cterminal ends of helices 1 and 4 affect activity. We note, however, that the region is close to other modification sites (42), so the conserved waters may be important for bridging communication networks between sites of PTMs. Finally, helix 5 interacts with eight conserved waters, and five of the waters are affected by mutations. While the waters do not make interprotomer contacts, helix 5 contains a binding site for zinc (29), ionic interactions (47), and phosphorylation sites (11) that are important in allosteric regulation of several caspases. Overall, the data show that the entire face of caspase-3 containing helices-1, -4, and -5, which are on the opposite side of the central β-sheet from the allosteric site, is perturbed by mutations in the dimer interface.

CONCLUSIONS

Our study provides insight into the intrinsic nature of allosteric mechanisms and the role of water molecules in conformational selection. The analysis of water molecules in the caspase-3 mutational database, combined with structural and dynamic studies, shows that disrupting conserved waters on the protomer is correlated with a shift in the population of states toward an inactive conformation (state 2 in Fig 1B). We note that the analysis reports on changes in the conserved waters within the protomer, as there are few

changes in inter-protomer water interactions, aside from the six conserved waters in the dimer interface noted above. In addition, the data show that the allosteric networks are global, in that there is not a single pathway to describe the allosteric connections. Rather, the regions with higher changes in conserved waters may report on communication between allosteric sites facilitated by the bridging water molecules. This fact suggests that, in addition to targeting the allosteric site of the dimer interface to stabilize the disordered-loop conformation (state 3 in Fig 1B), ligands that remove conserved water molecules from the protomer could effectively fine-tune caspase-3 activity by shifting the population toward the high-energy inactive conformation (state 2 in Fig 1B).

The disordered-loop conformation (Fig 1B, state 3) appears to be a low-energy state since it is readily trapped by small drug compounds. It is also observed in the so-called "closed-loop" conformations of several caspases, where binding of active site loop L2' in the dimer interface prevents insertion of the "elbow loop" of L3 and subsequent formation of the substrate binding groove (20, 21, 43, 44, 48, 49). Interestingly, the allosteric site of the dimer interface is not modified by PTMs in the cell, and a global analysis of caspase modifications (42), along with trapping high-energy states of caspases-2 (28) and -3 (22, 23), shows that the disordered-loop conformation is only one of potentially several inactive states in the ensemble. Studies of the high-energy inactive conformations of caspases-2 and -3 show that disruption of the catalytic groups can be achieved without large scale disordering of the active site loops, as small shifts in the catalytic cysteine and histidine are sufficient to inactivate the enzyme. The high-energy inactive state may provide an advantage for regulation of caspase activity in the cell, particularly if the state is stabilized by post-translational modifications, because it

provides the cell a means to reversibly control activity by coupling one or more modifications to increased population of the inactive state. Removal of the modification would result in an unstable high-energy state that would rapidly convert to the more stable active conformation. The database described here provides a tunable allosteric library of caspase-3 variants with nearly four orders of magnitude change in activity, which we suggest will be useful for examining caspase-3 signaling in cells. In current transfection technologies, caspase activity is manipulated in cells through knock-down or knock-out strategies coupled with expression of caspase variants. Levels of endogenous caspase activity are difficult to control, and interpretation of results from overexpressed caspase variants can be problematic. New genome editing techniques, however, combined with our caspase-3 database, should provide tools to fine-tune caspase-3 activity while simultaneously maintaining endogenous protein levels.

METHODS

Cloning, expression and protein purification. Site-directed mutagenesis was performed as described previously with plasmids pHC332 (wild-type caspase-3) (50) and pHC33209 (caspase3(D9A,D28A,D175A)) (uncleavable procaspase-3) (50, 51) and the primers shown in

Supplemental Table S5. *Escherichia coli* BL21(DE3)pLysS cells were transformed with each of the plasmids, and proteins were expressed and purified as previously described (22, 45, 52, 38, 53).

Enzyme activity assay. The initial velocity of substrate cleavage was measured at 25 ^{o}C in the presence of varying concentrations of substrate (Ac-DEVD-AFC), as described

previously (52). The final protein concentration for the active mutants was 10 nM, whereas a protein concentration of 100 nM was used for the largely inactive mutants. The total reaction volume was 200 μL. Briefly, substrate was added to the sample which contained protein in activity assay buffer (150 mM Tris–HCl, 100 mM DTT, 0.1% CHAPS, 50 mM NaCl, 1% sucrose pH 7.5), and samples were immediately excited at 400 nm while the fluorescence emission was measured at 505 nm for 60 seconds.

Analysis of water clusters. Water molecules were analyzed using DRoP, as described previously (30), and the web server http://dropinthemattoslab.org/. Briefly, all structures were first aligned to that of wild-type caspase-3, 2J30, and then loaded onto the web server. The DRoP program returns a PDB file with clustered and renumbered waters as well as the level of conservation of each water molecule in the database. The conserved waters were further characterized based on the number of hydrogen bonds to the protein and to other water molecules: buried, form at least 3 H-bonds to protein side-chains or backbone atoms; channel, generally form 2 H-bonds to the protein and at least 1 H-bond to other channel or buried water molecules; surface, form 1-2 H-bonds with the protein. For analysis of the caspase-3 mutants in the I222 spacegroup, DRoP analysis was carried out after alignment with wild-type caspase-3 (2J30). For caspase-3 mutants in the C121 space group, which consists of a dimer of protomers in the asymmetric unit, one protomer was removed from the file prior to alignment with wildtype caspase-3. Following DRoP analysis, the conserved water molecules in each mutant were inspected for displacement.

Crystallization and data collection. Caspase-3 variants were crystallized in the presence of Ac-DEVD-CMK as previously described (22, 38), and most crystals appeared within about 3 days, although some took as long as three weeks to grow. Cryo-protectants included 20% PEG 400/80% reservoir solution or 20% MPD/80% reservoir solution. Data sets were collected and structures were solved by molecular replacement using the wild-type caspase-3 structure for initial phasing (PDB code 2J30); structural models were refined using Phenix (54), as previously described (23, 46, 38). Structures were determined for all V266X variants except for V266G, V266P, V266R, and V266T. The 13 structures determined were within 0.095 Å RMSD from wild-type caspase-3. A summary of the data collection and refinement statistics is shown in Supplementary Table S3.

Molecular dynamics simulations. Molecular dynamics simulations were performed for 50 ns with GROMACS 4.5 (55) using the Amber99 force field (56) and the TIP3P water model (57), as previously described (22, 23).

Accession numbers. Protein data bank accession numbers: caspase 3(V266A) (PDB ID: 5I9B), caspase 3(V266C) (PDB ID: 5I9T), caspase 3(V266D) (PDB ID: 5IAB), caspase 3(V266F) (PDB ID: 5IAE), caspase 3(V266I) (PDB ID: 5IBC), caspase 3(V266K) (PDB ID: 5IBR), caspase 3(V266L) (PDB ID: 5IAJ), caspase 3(V266M) (PDB ID: 5IBP), caspase 3(V266N) (PDB ID: 5IAN), caspase 3(V266Q) (PDB ID: 5IAG), caspase 3(V266S) (PDB ID: 5IAK), caspase 3(V266W) (PDB ID: 5IAR), caspase 3(V266Y) (PDB ID: 5IAS).

Acknowledgments. The authors thank Denise Appel and Matthew Willson for technical expertise. This work was supported by a grant from the NIH (National Institutes of Health) [grant number GM065970 (to A.C.C)]. The authors thank the research agencies of North

Carolina State University and the North Carolina Agricultural Research Service. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract number W-31-109-ENG-3

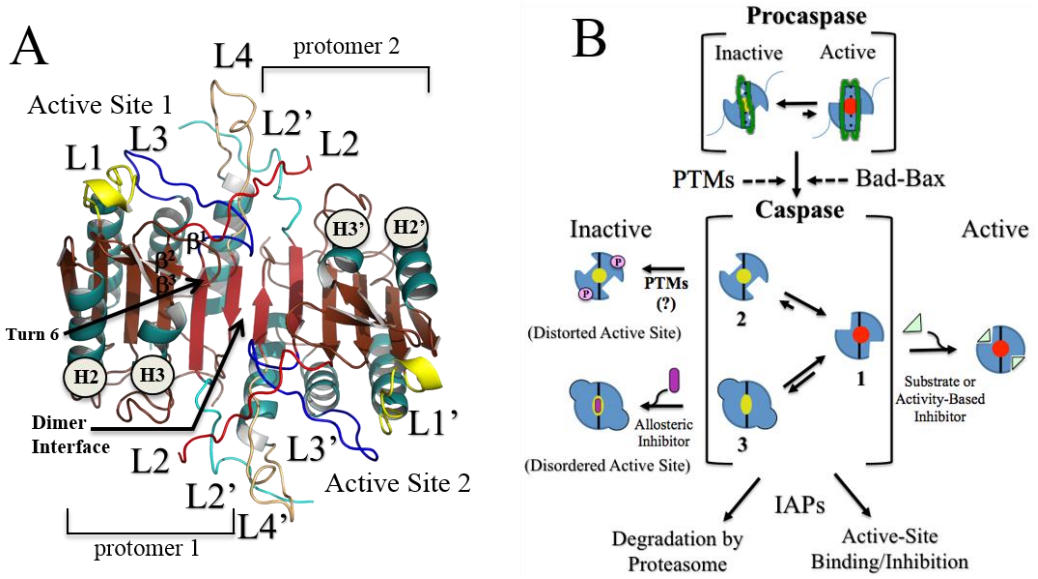


Figure 1. Caspase Allosteric Regulation.

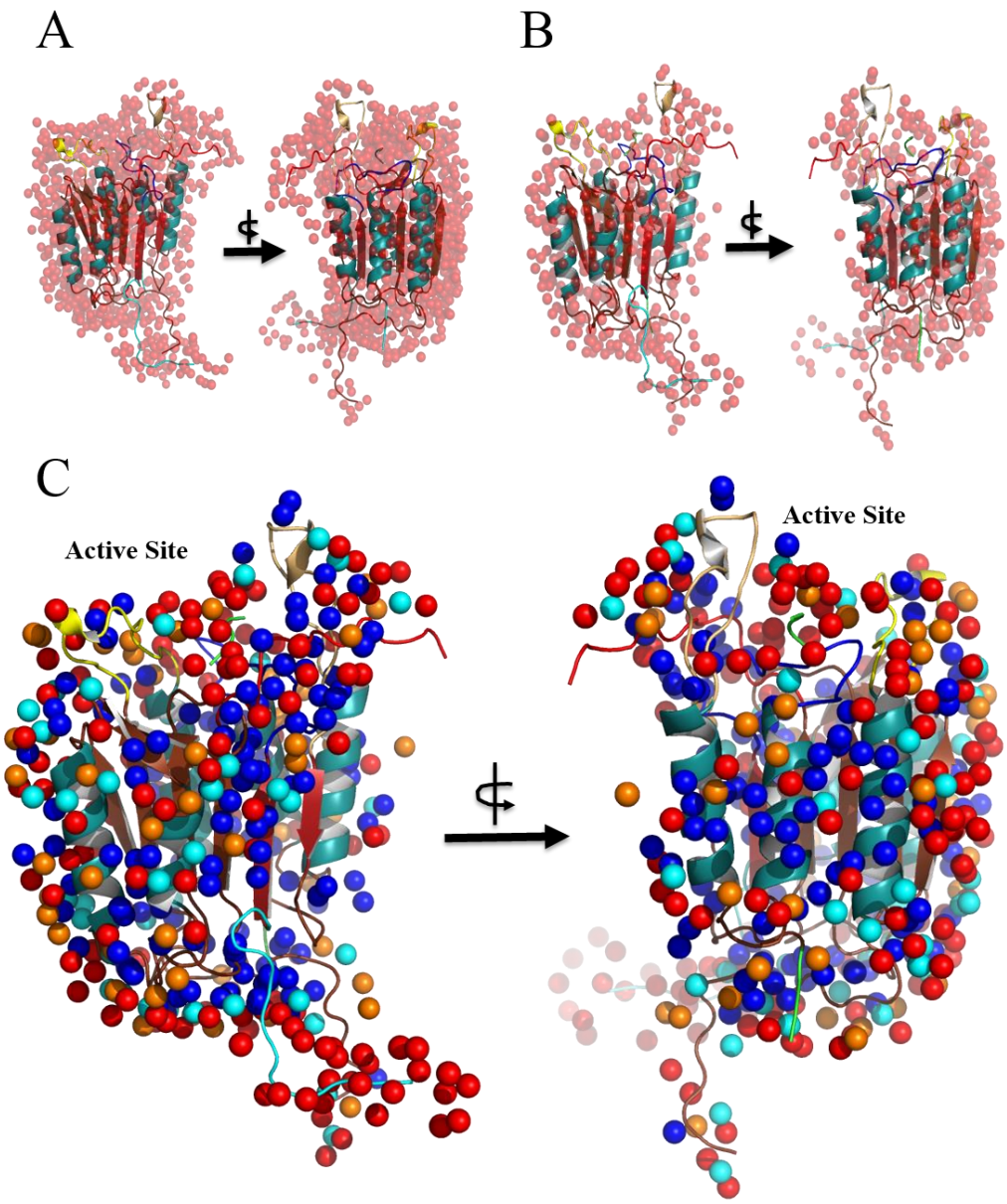


Figure 2. DRoP Analysis of Wild-type Caspase-3.

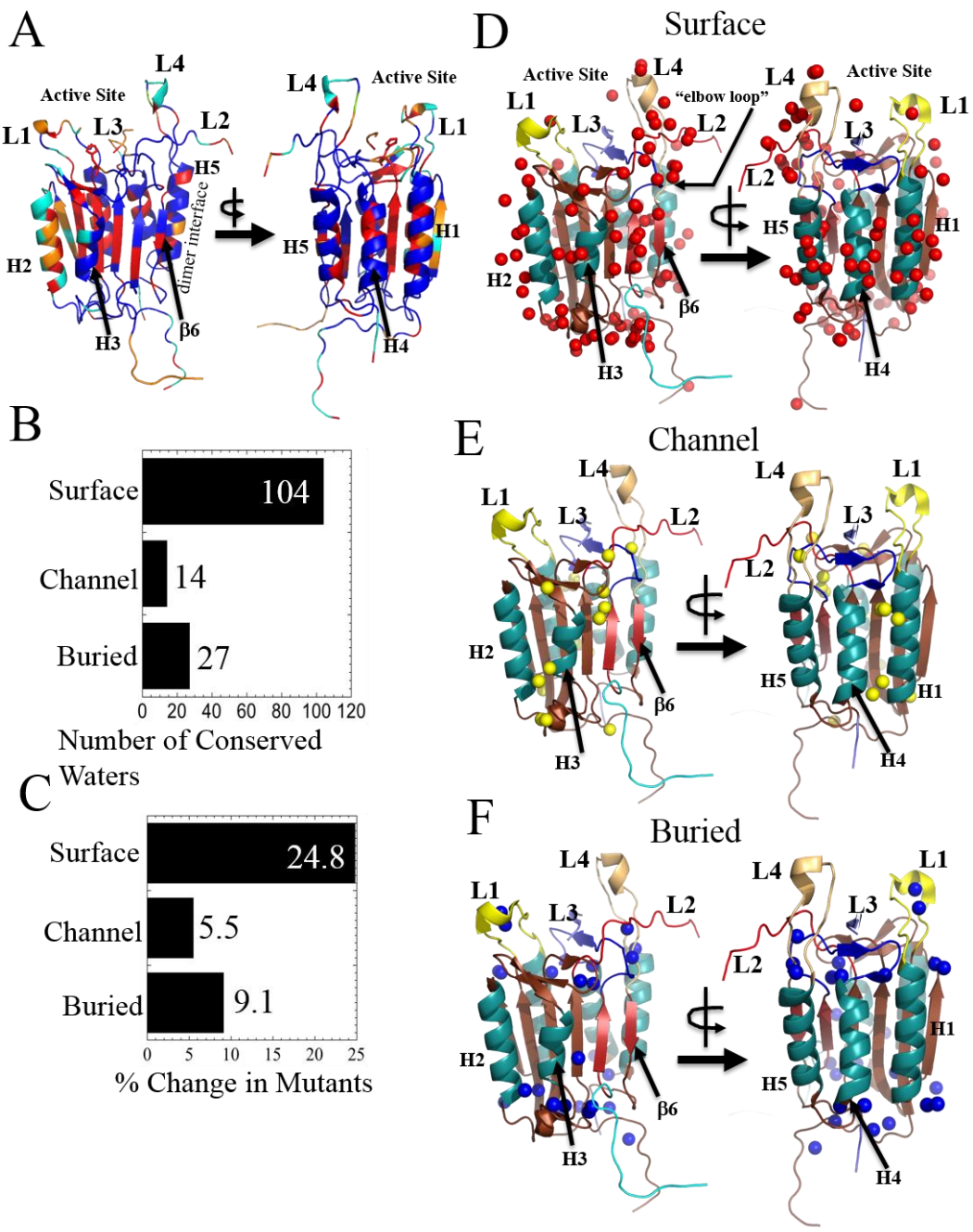


Figure 3. Conserved water molecules identified in DRoP analysis.

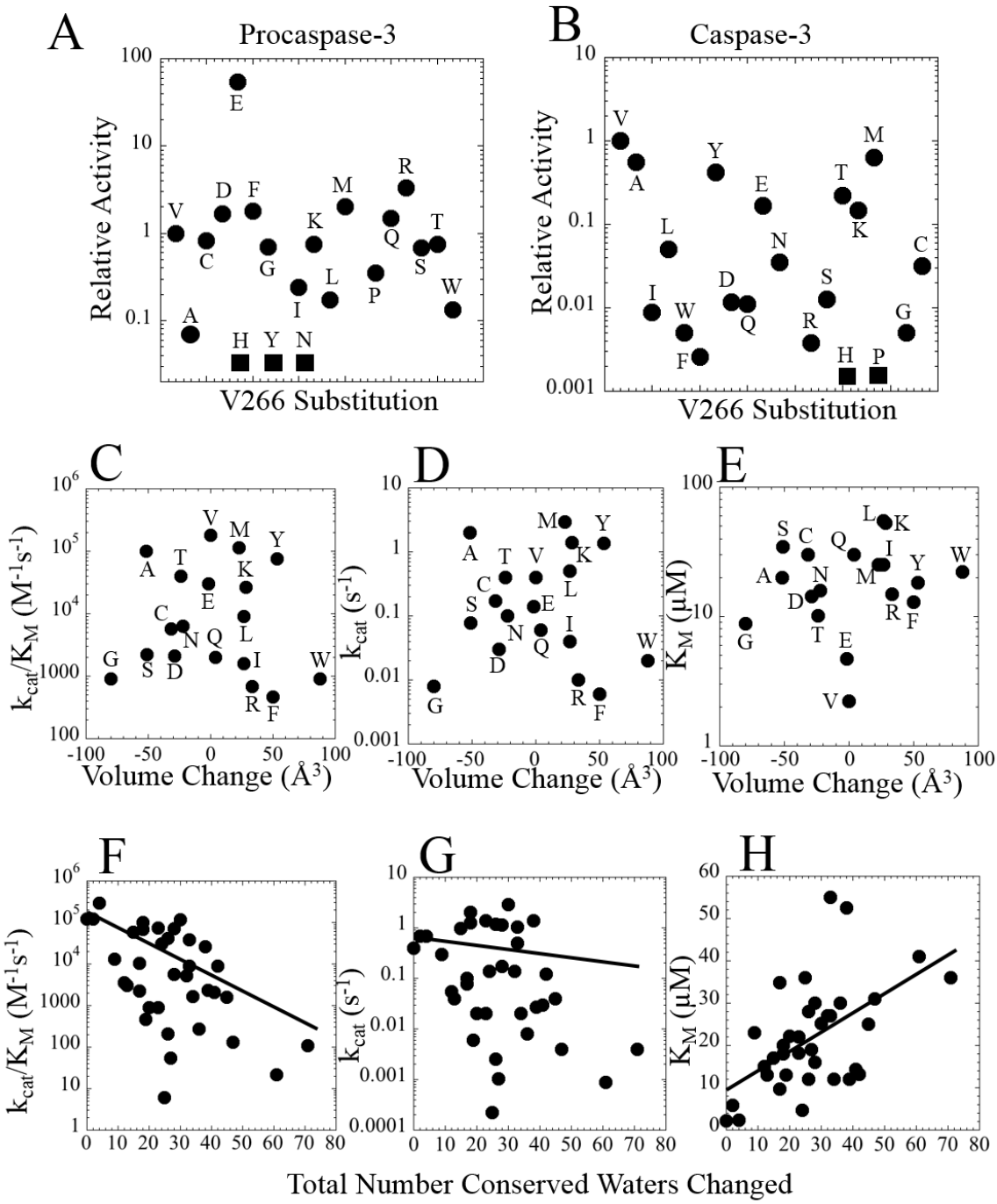


Figure 4. Saturation Mutagenesis of Dimer Interface Allosteric Site.

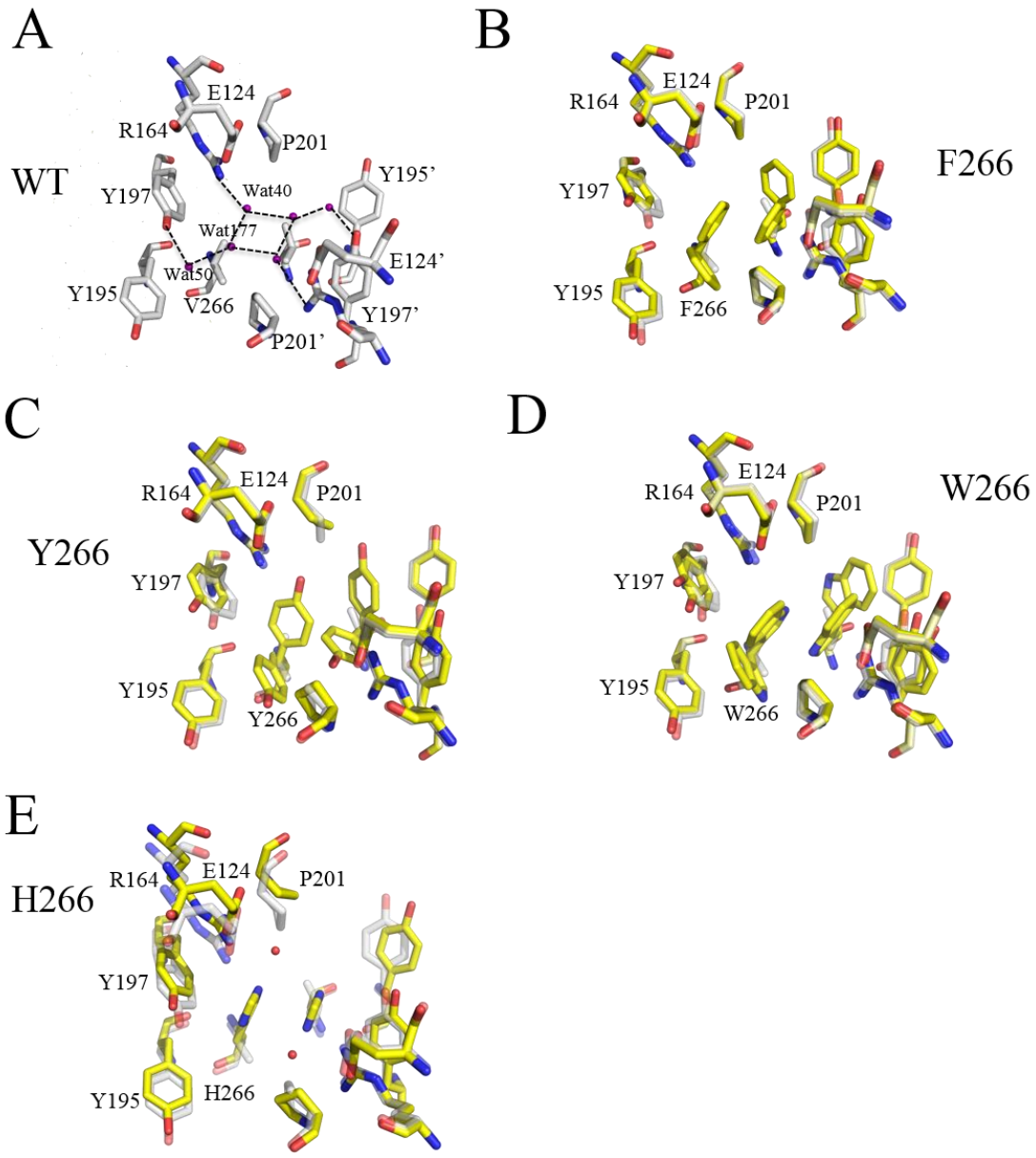


Figure 5. Comparison of Wild-type and V266X Caspase-3 Dimer Interfaces.

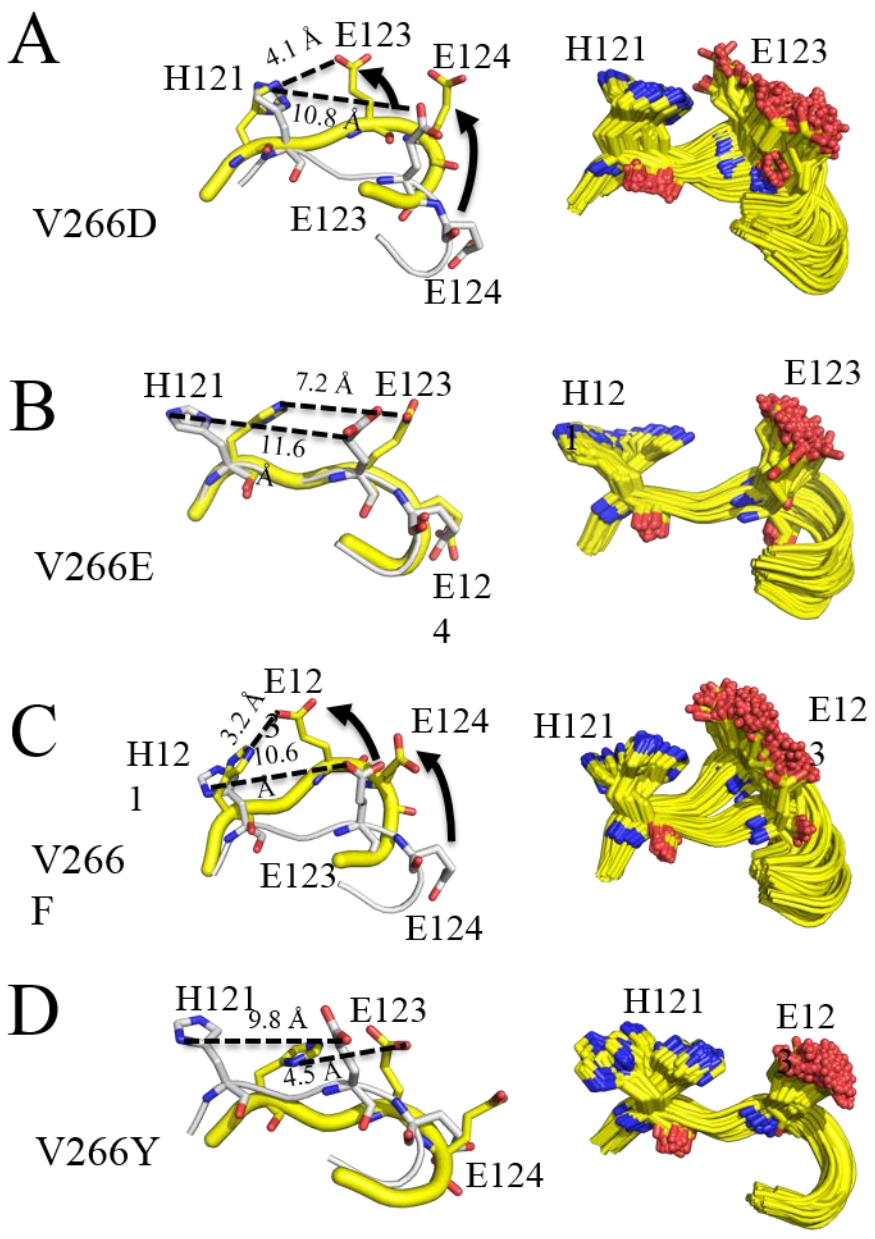


Figure 6. Molecular Dynamic Simulations of Various Caspase-3 Mutants

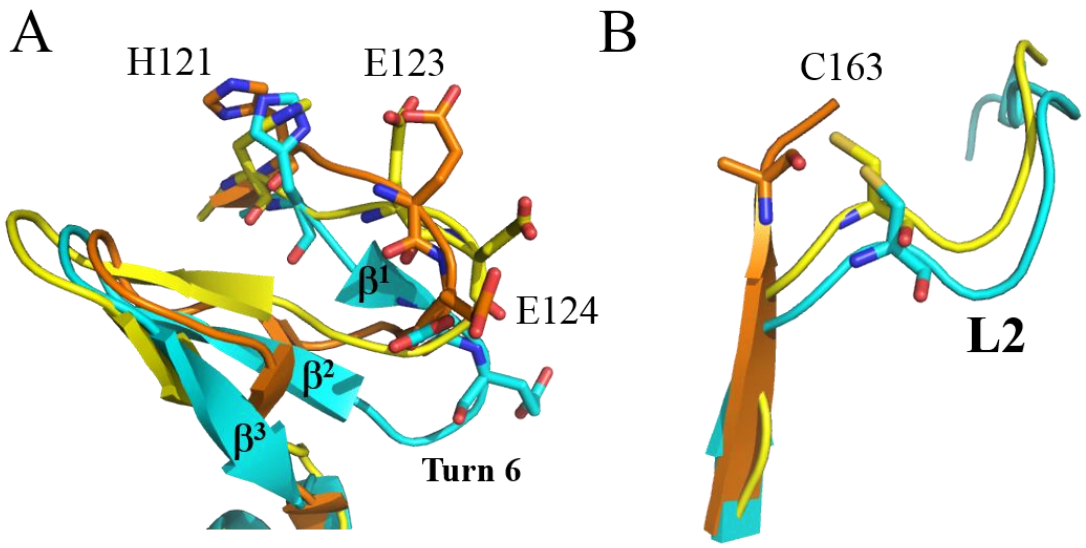


Figure 7. Comparison of Turn 6 and catalytic groups in the caspase-3 zymogen, active caspase-3 and V266D variant.

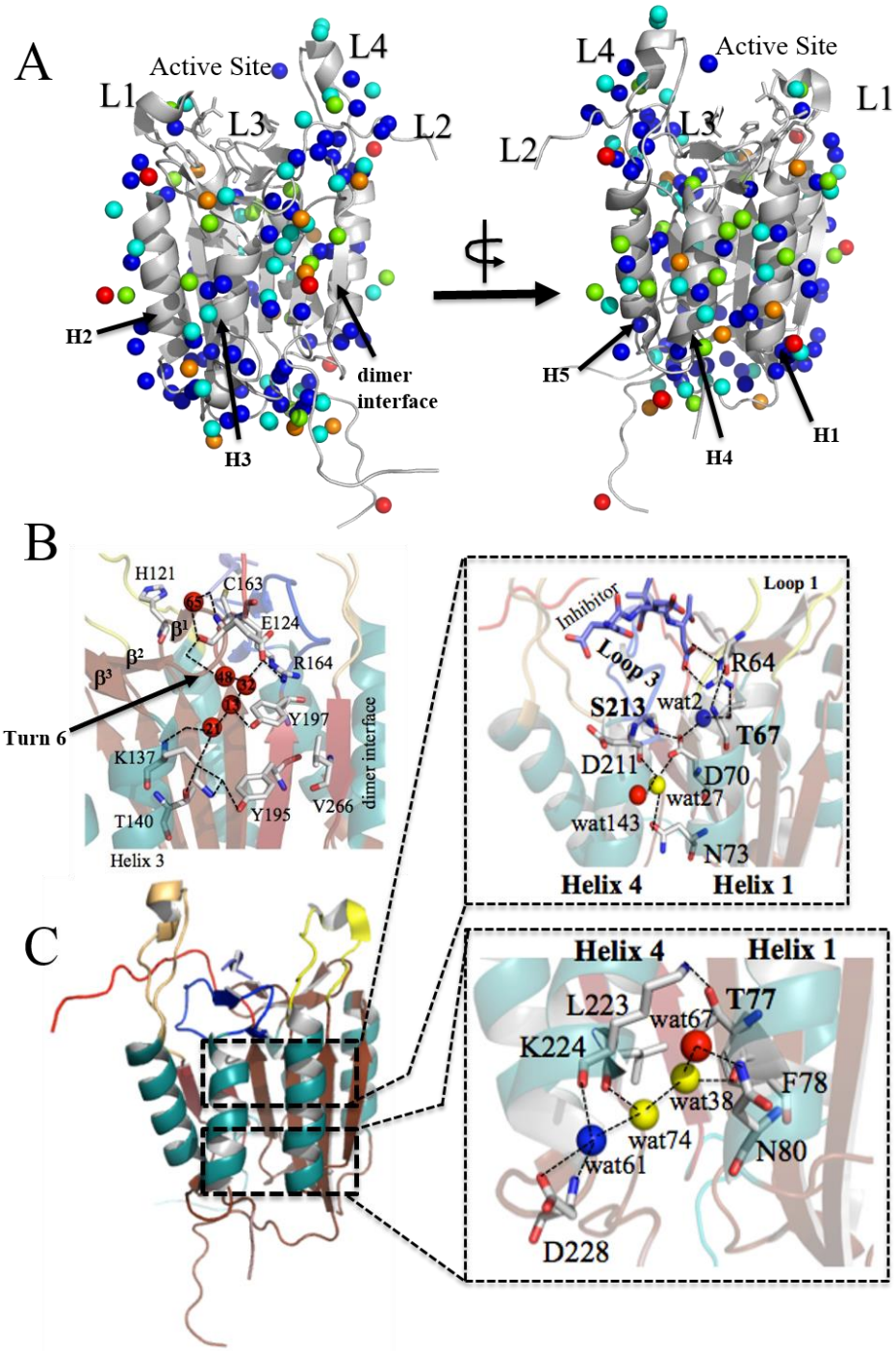


Figure 8. "Hot-spots" of conserved water molecules affected by mutations in allosteric site of the dimer interface.

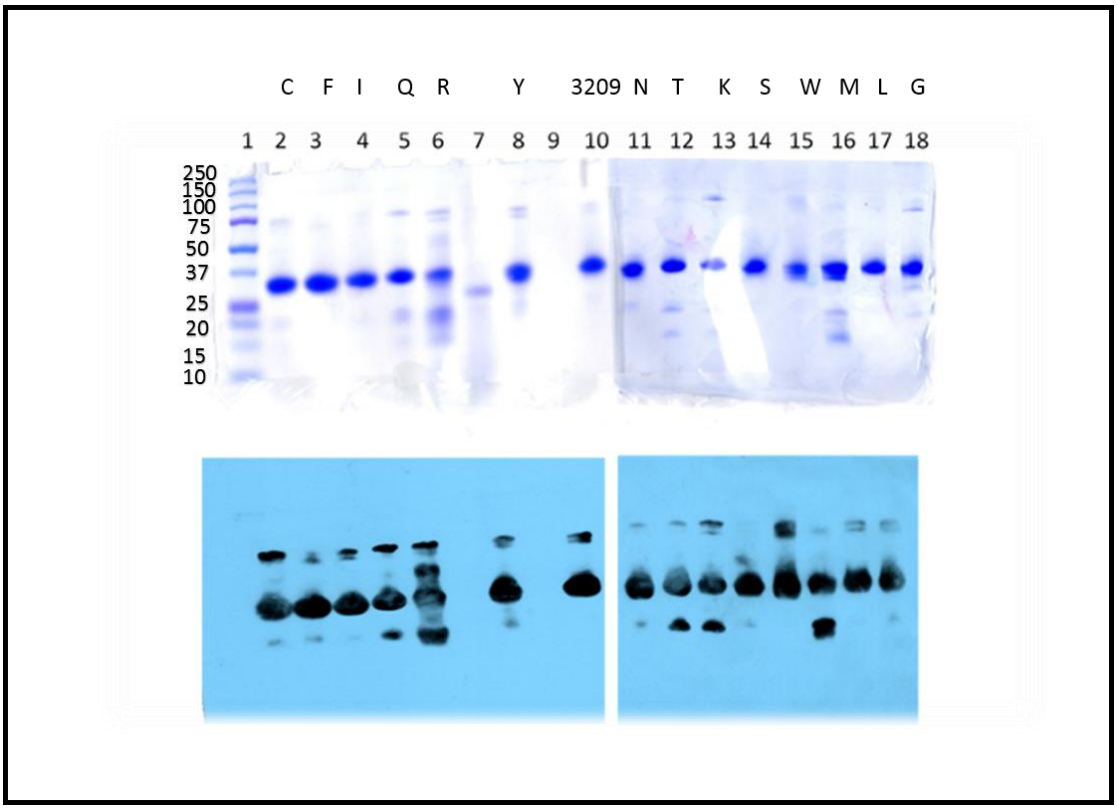
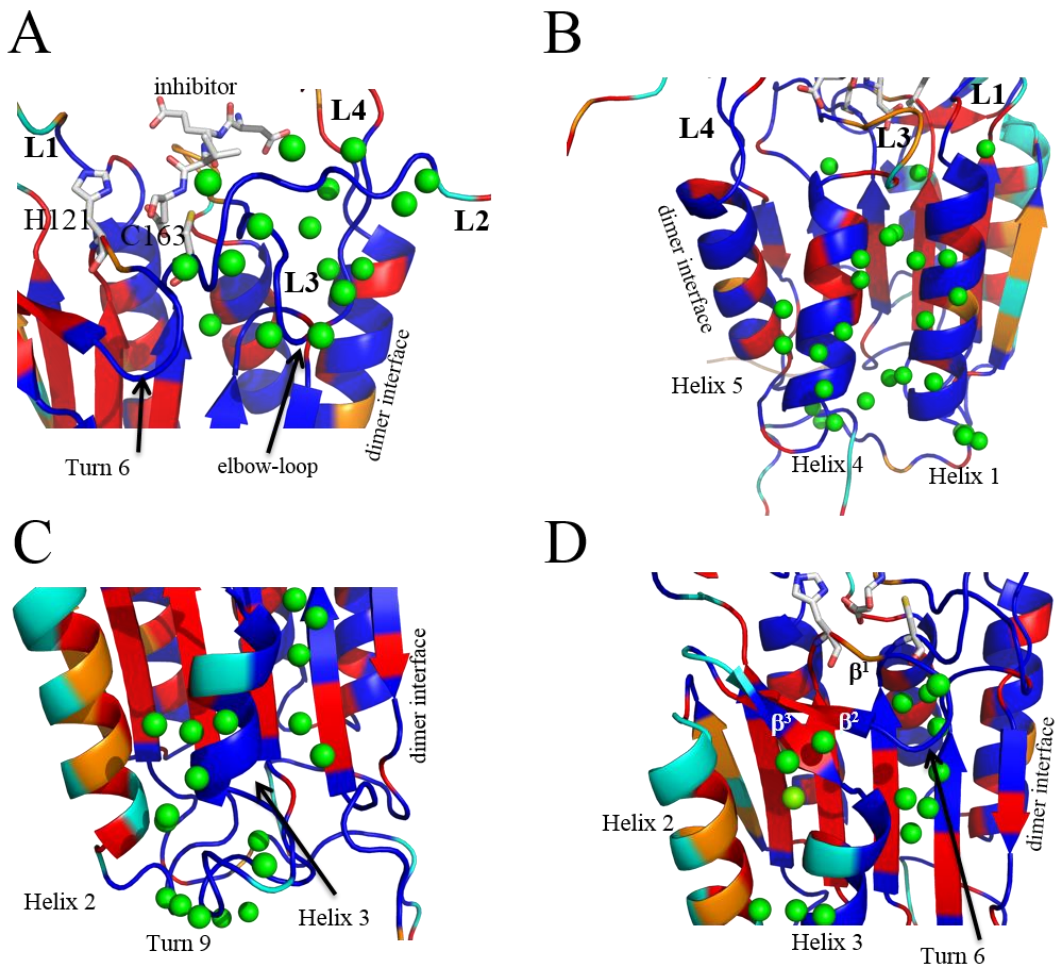
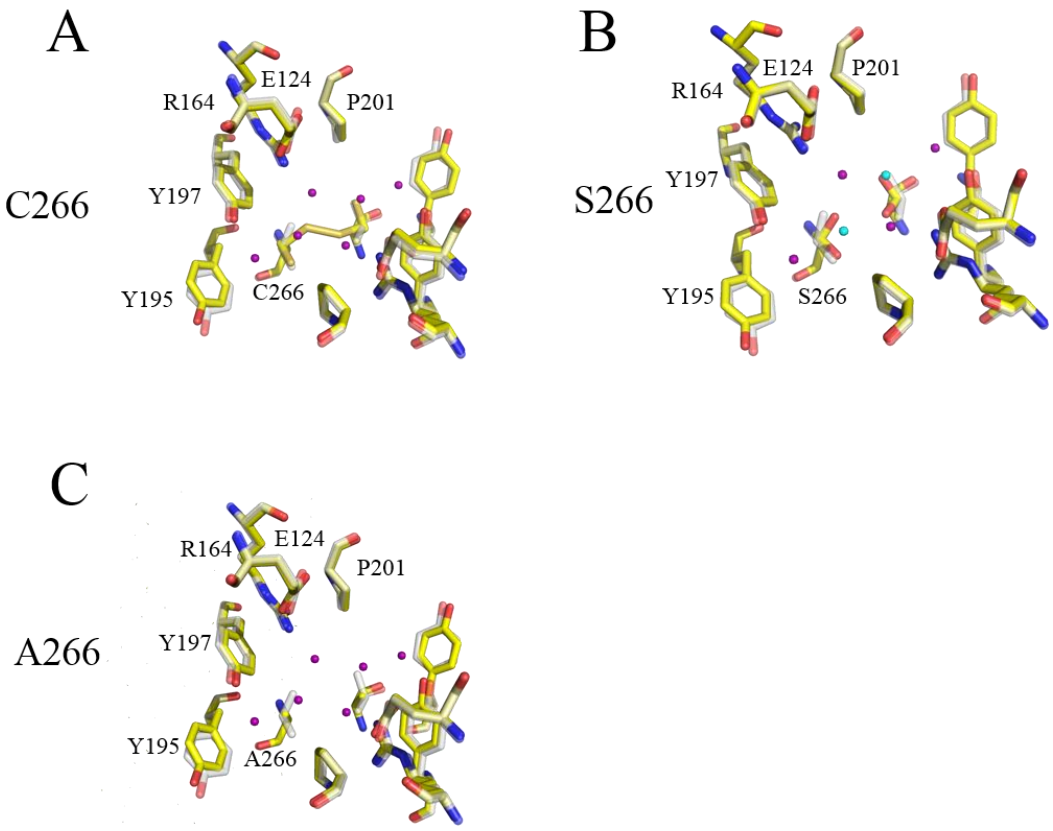


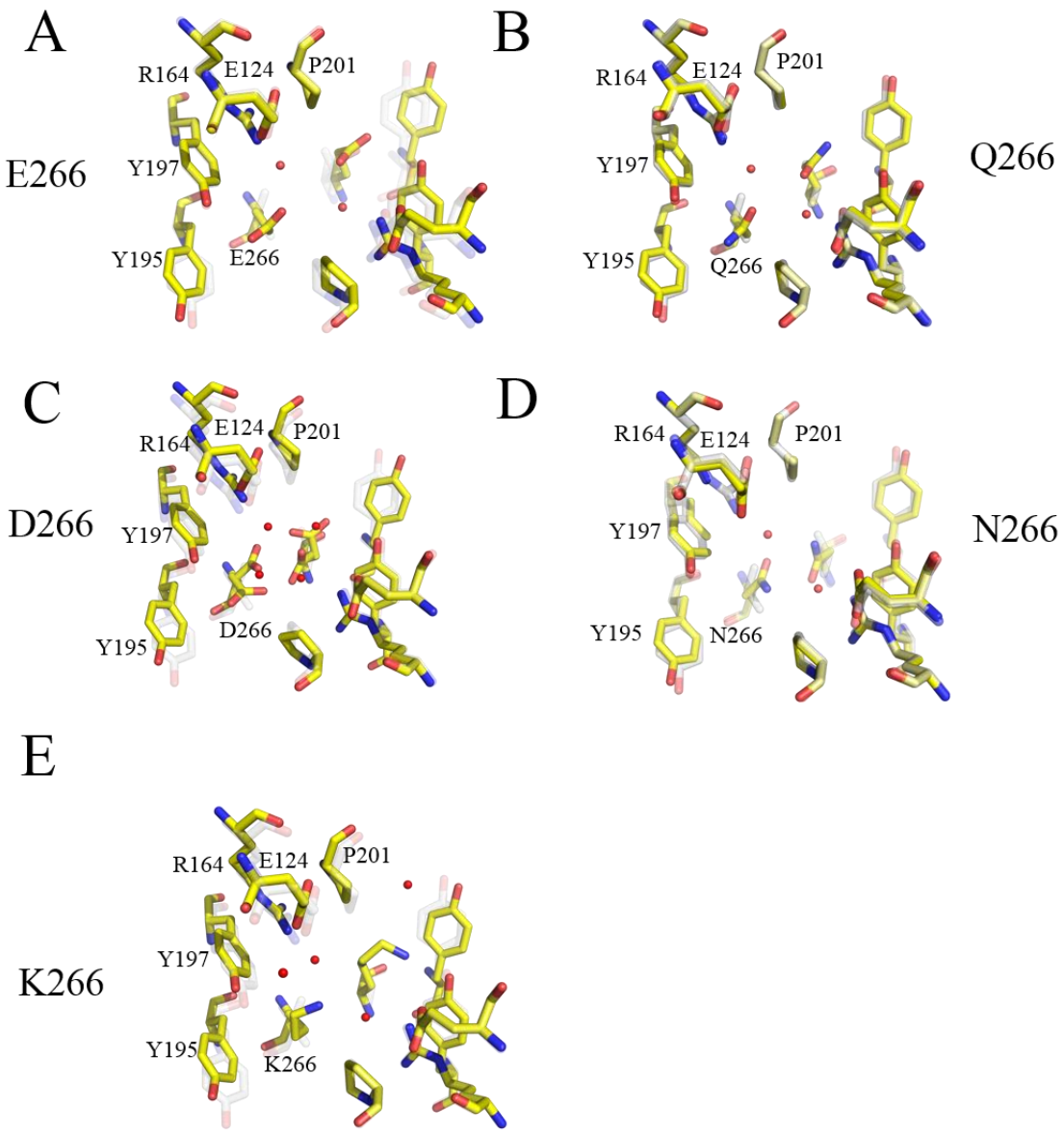
Figure 9. SDS-PAGE (top) and Western Analyses (bottom) of Caspase-3 D3A Mutants.



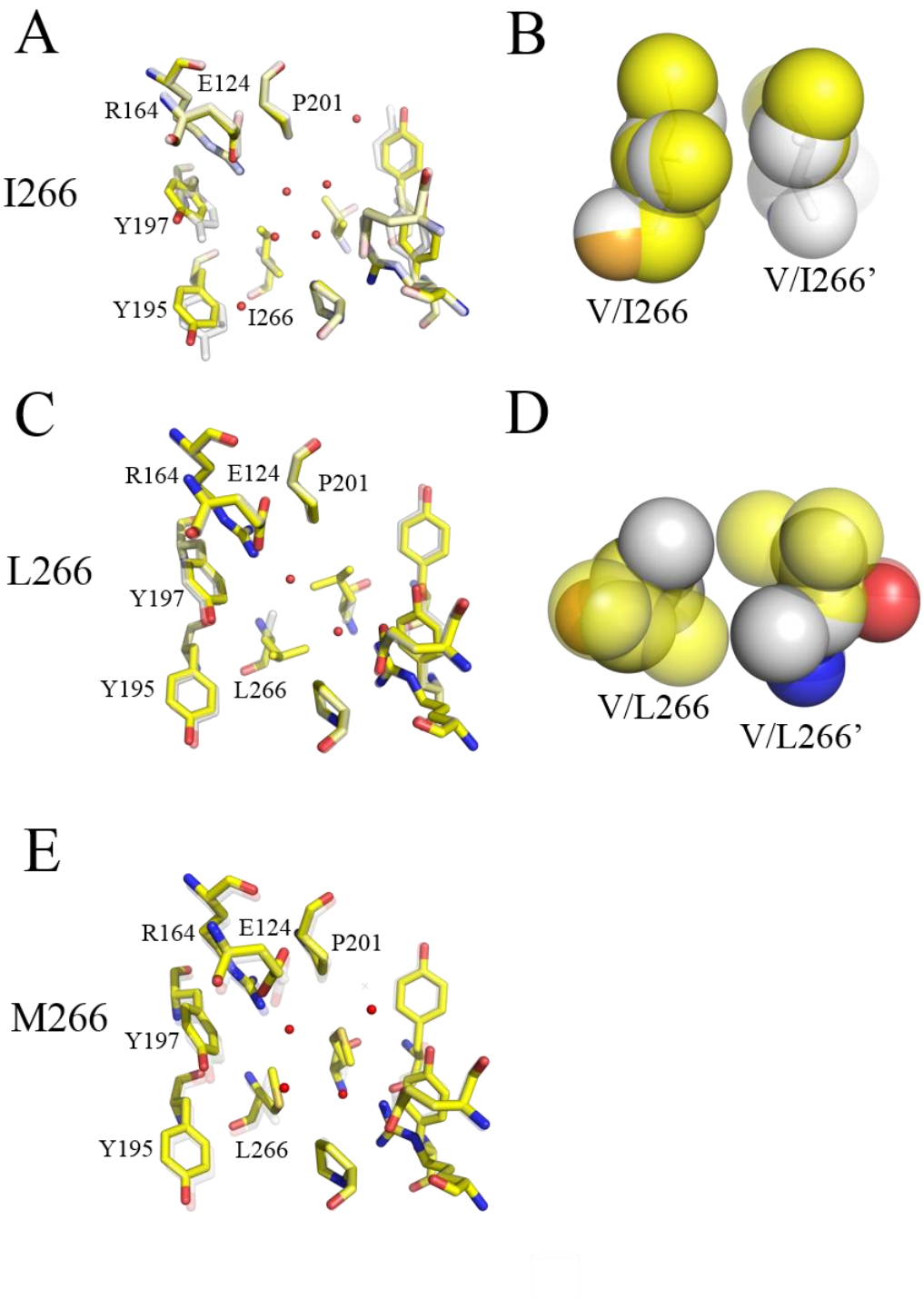
Supplemental Figure S1. Conserved Water Molecules



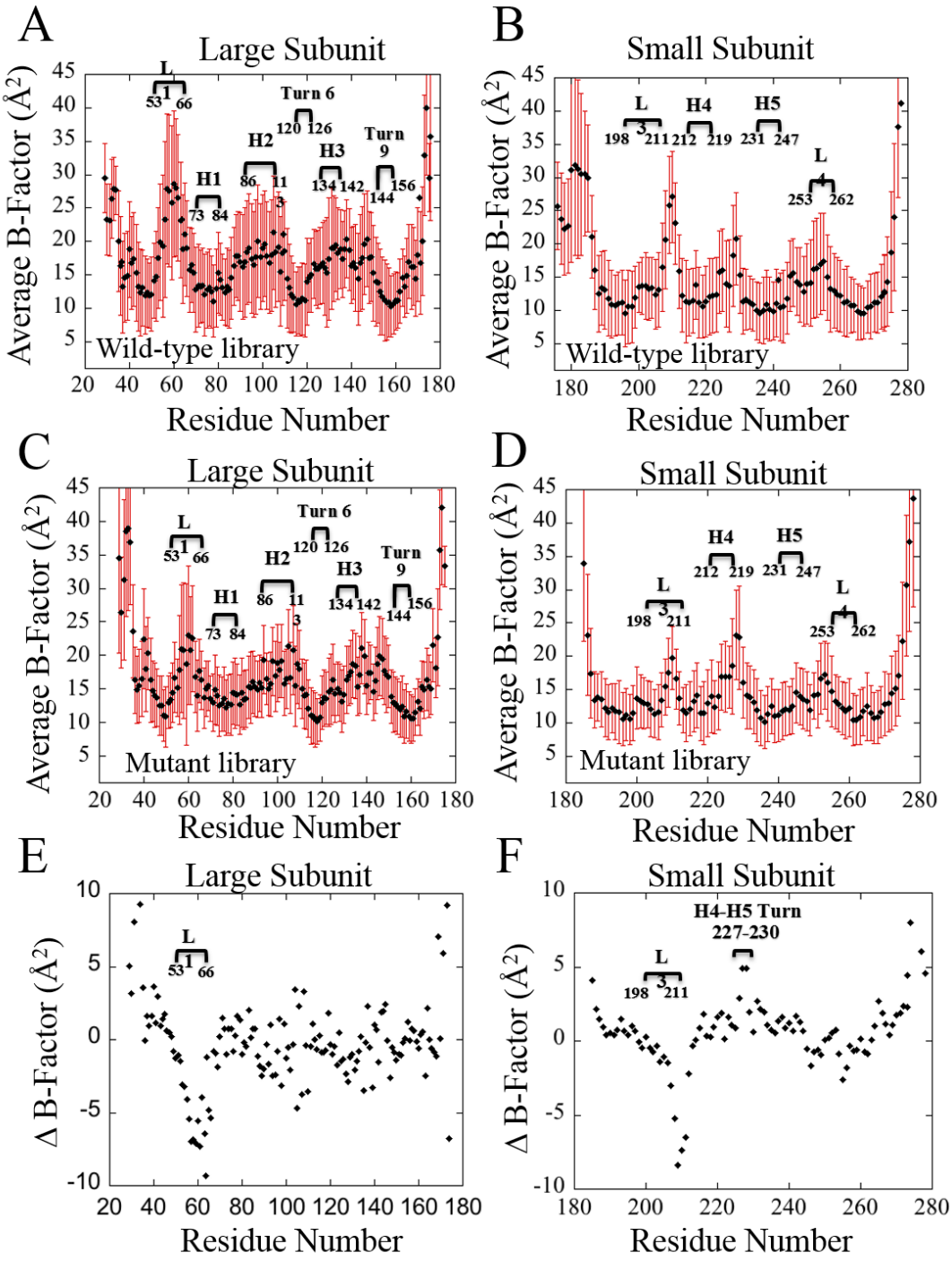
Supplemental Figure S2. Comparison of Wild-type and V266X Caspase-3 Dimer Interfaces: Small/Polar Amino Acids.



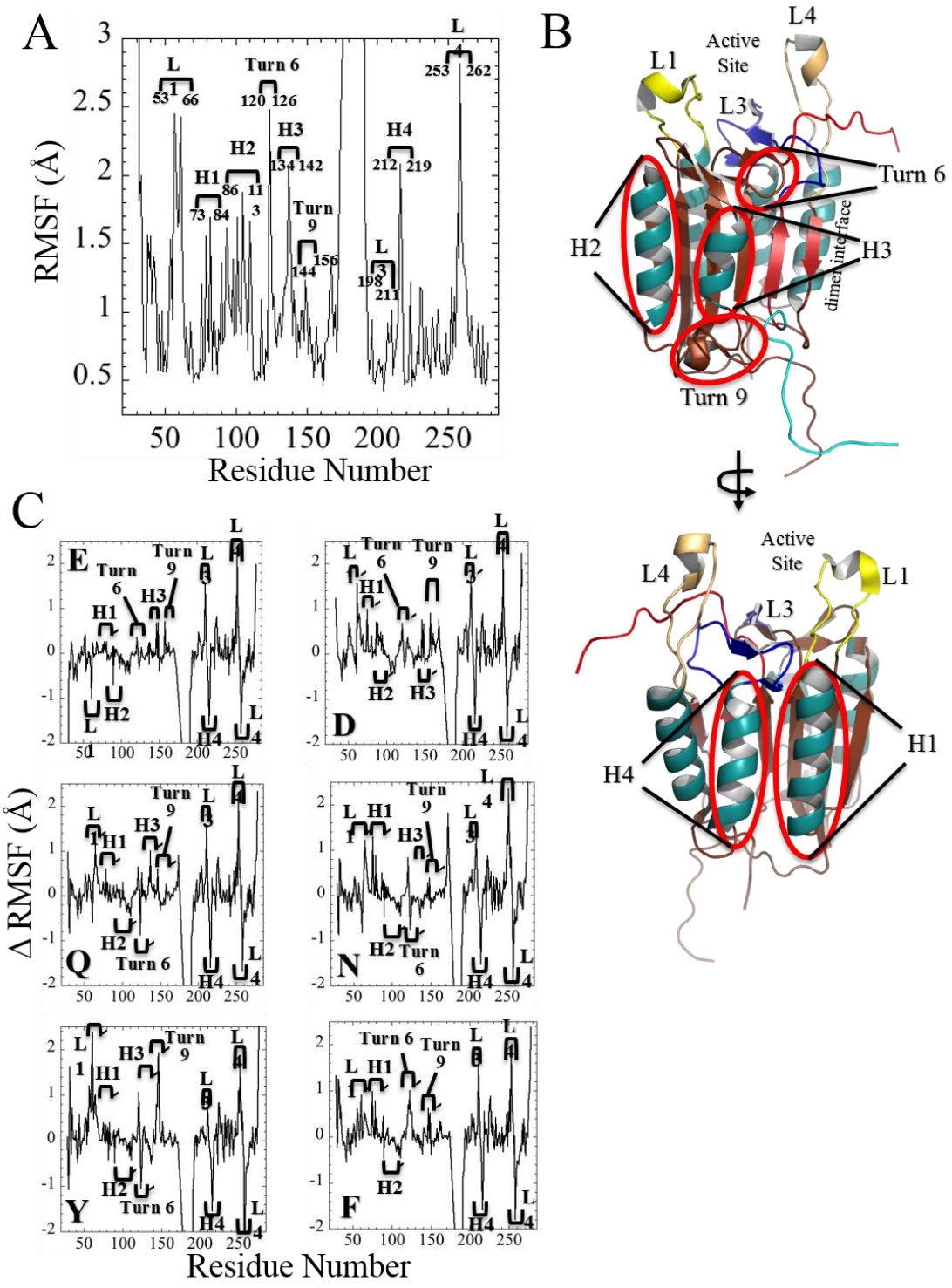
Supplemental Figure S3. Comparison of Wild-type and V266X Caspase-3 Dimer Interfaces: Large Polar and Charged Amino Acids.



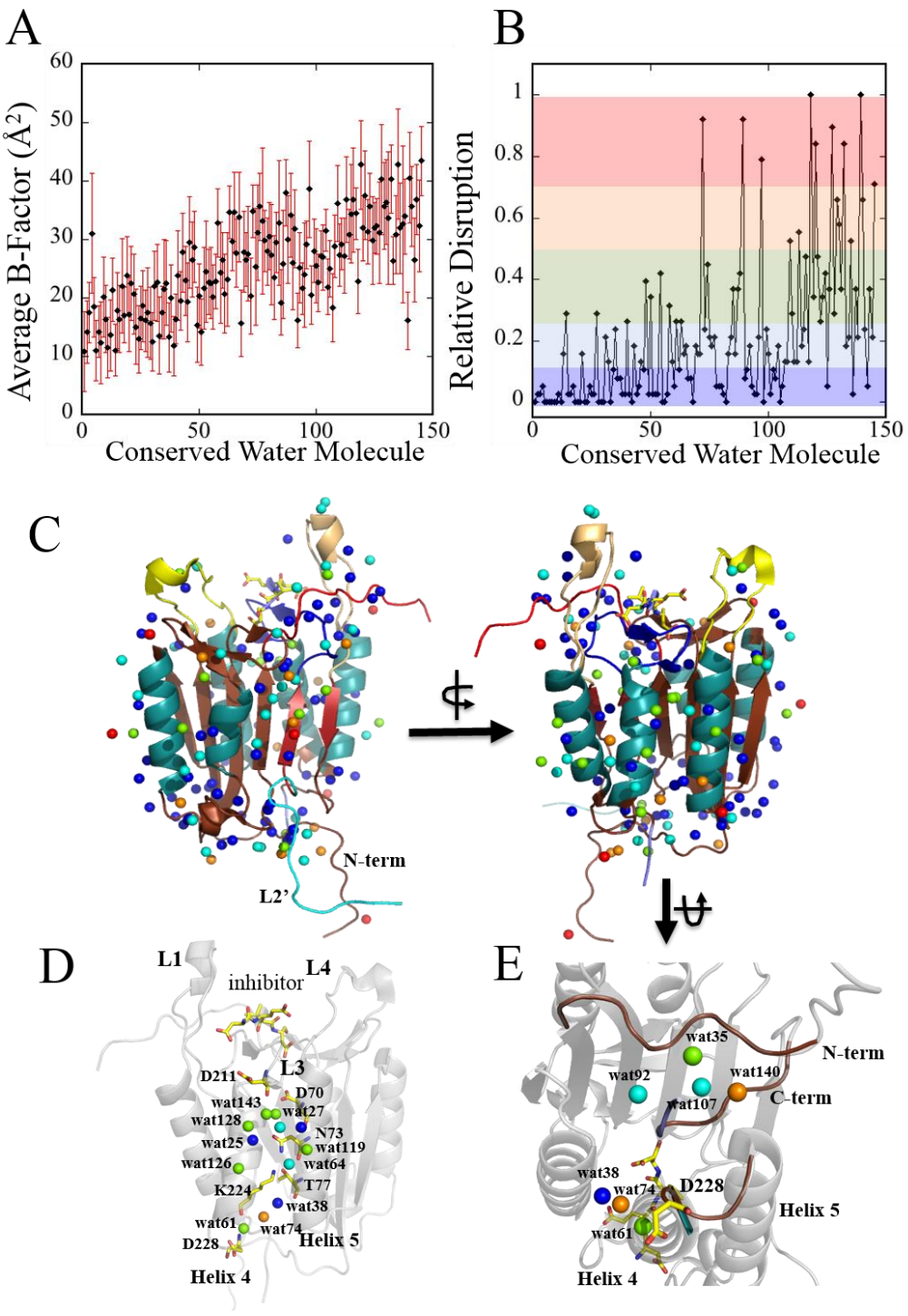
Supplemental Figure S4. Comparison of Wild-type and V266X Caspase-3 Dimer Interfaces: Hydrophobic Amino Acids.



Supplemental Figure S5. B-factor Analysis of Caspase-3 Databases.



Supplemental Figure S6. Molecular Dynamics Simulations of Caspase-3 and Selected Mutants.



Supplemental Figure S7. B-factor Analysis of Conserved Water Molecules.

Supplemental Table S1. Wild-type caspase-3 structures used in DRoP[1] analysis

PDB Code	Resolution (Å)	Space Group	Number water molecules	Reference
1NME	1.6	I222	303	(3)
1QX3	1.9	$P2_12_12$	147	(4)
2C1E	1.77	I222	414	(5)
2C2K	1.87	I222	367	(5)
2C2M	1.94	I222	359	(5)
2CDR	1.7	I222	369	(6)
2CJX	1.7	I222	363	(7)
2CJY	1.67	I222	337	(7)
2CNK	1.75	I222	403	(6)
2CNL	1.67	I222	361	(6)
2CNN	1.7	I222	348	(6)
2J30	1.4	I222	287	(8)
2XYG	1.54	I222	426	(9)
2XYH	1.89	I222	297	(9)
4JJE	1.48	I222	214	(10)
Total			**4,995**	

[1]DRoP described in Reference (11)

Supplemental Table S2. Enzymatic activity of (pro)caspase-3 V266 variants

Mutant	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
Wild-type (V266)[1]	0.4[(3)]	2.2	$1.8x10^5$
Procaspase-3 (D_3A)[1,2]	.003	3.5	$8.6x10^2$
Caspase-3 (V266A)	2.0	20	$1.0x10^5$
Procaspase-3 (D_3A,V266A)	0.0025	40	$6.25x10^1$
Caspase-3 (V266C)	0.17	30	$5.66x10^3$
Procaspase-3 (D_3A,V266C)	.025	45	$5.55x10^2$
Caspase-3 (V266D)	.03	14.4	$2.08x10^3$
Procaspase-3 (D_3A,V266D)	.03	27	$1.11x10^3$
Caspase-3 (V266E)[1]	0.14	4.7	$3.0x10^4$
Procaspase-3 (V266E)[1]	0.2	5.6	$3.7x10^4$
Caspase-3 (V266F)	.006	13	$4.60x10^2$
Procaspase-3 (D_3A,V266F)	.045	38	$1.18x10^3$
Caspase-3 (V266G)	.008	8.8	$9.09x10^2$
Procaspase-3 (D_3A,V266G)	.014	30	$4.70x10^2$
Caspase-3 (V266H)[1]	N.D.[4]	N.D.	N.D.
Procaspase-3 (D_3A,V266H)[1]	N.D.	N.D.	N.D.
Caspase-3 (V266I)	0.04	25	$1.60x10^3$
Procaspase-3 (D_3A,V266I)	.004	25	$1.60x10^2$
Caspase-3 (V266K)	1.40	52.5	$2.66x10^4$
Procaspase-3 (D_3A,V266K)	.015	30	$5.00x10^2$
Caspase-3 (V266L)	.50	55	$9.00x10^3$
Procaspase-3 (D_3A,V266L)	.004	35	$1.14x10^2$
Caspase-3 (V266M)	2.90	25.2	$1.15x10^5$
Procaspase-3 (D_3A,V266M)	0.04	30	$1.33x10^3$
Caspase-3 (V266N)	0.10	15.9	$6.29x10^3$
Procaspase-3 (D_3A,V266N)	N.D.	N.D.	N.D.
Caspase-3 (V266P)	N.D.	N.D.	N.D.
Procaspase-3 (D_3A,V266P)	.007	30	$2.33x10^2$

Supplemental Table S2 Continued

Caspase-3 (V266Q)	0.06	30	$2.00x10^3$
Procaspase-3 (D_3A,V266Q)	.033	33	$1.00x10^3$
Caspase-3 (V266R)	0.01	14.8	$6.75x10^2$
Procaspase-3 (D_3A,V266R)	.067	30	$2.23x10^3$
Caspase-3 (V266S)	.078	34.8	$2.24x10^3$
Procaspase-3 (D_3A,V266S)	.016	35	$4.57x10^2$
Caspase-3 (V266T)	0.40	10.1	$3.96x10^4$
Procaspase-3 (D_3A,V266T)	.018	35	$5.14x10^2$
Caspase-3 (V266W)	0.02	22.1	$9.04x10^2$
Procaspase-3 (D_3A,V266W)	.004	45	$8.80x10^1$
Caspase-3 (V266Y)	1.36	18.2	$7.47x10^4$
Procaspase-3 (D_3A,V266Y)	N.D.	N.D.	N.D.

[1]Data for wild-type (pro)caspase-3 are described in reference (12)

[2]D_3A refers to uncleavable procaspase-3(D9A,D28A,D175A) as described in reference (13)

[3]Errors for k_{cat} and K_M were <5% of the reported values

[4]N.D., no detectable activity

Supplemental Table S3. Crystallographic parameters for V266X variants.

Parameters	V266V	V266A	V266C	V266D	V266F	V266Q	V266L
PDB Code	2J30	5I9B	5I9T	5IAB	5IAE	5IAG	5IAJ
Space Group	I222	I222	C121	C121	C121	I222	I222
Unit Cell Param.							
a (Å)	68.73	68.556	108.315	108.590	128.223	69.040	69.375
b (Å)	84.40	84.545	96.299	96.709	68.765	85.104	84.871
c (Å)	96.35	96.146	68.019	68.929	84.706	96.273	96.564
α (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00
β (°)	90.00	90.00	128.90	126.30	131.34	90.00	90.00
γ (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00
Temperature (K)	100	100	100	100	100	100	100
Resolution (Å)	1.40	1.799	1.949	1.787	1.545	1.977	1.581
# Reflections	54279	26138	39517	54012	80456	16493	36041
Completeness (%)	97.9	99.24	99.61	99.52	99.47	81.59	91.71
R_{work}(%)	19.6	16.71	16.39	17.34	15.46	17.41	17.73
R_{free} (%)	20.7	20.35	19.76	21.34	18.22	23.14	20.22
Minimum $F_{obs}/\sigma F_{obs}$	--------	1.33	1.34	1.34	1.34	1.34	1.34
RMSD Bonds (Å)	0.005	0.011	0.007	0.007	0.006	0.006	0.007
RMSD Angles (Å)	1.30	0.835	0.845	0.996	0.898	1.093	1.171
# Protein Atoms	1974	2046	4028	3991	4279	2045	2013
# Water Atoms	290	218	347	351	590	89	195

Supplemental Table S3 Continued

Parameters	V266S	V266N	V266W	V266Y	V266I	V266M	V266K
PDB Code	5IAK	5IAN	5IAR	5IAS	5IBC	5IBP	5IBR
Space Group	I222	I222	I222	I222	I222	I222	C121
Unit Cell Param							
a (Å)	69.090	69.021	68.871	69.027	68.591	68.568	108.968
b (Å)	84.856	84.910	84.628	84.941	84.325	84.464	96.338
c (Å)	96.286	96.386	96.237	96.459	96.617	96.201	68.700
α (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00
β (°)	90.00	90.00	90.00	90.00	90.00	90.00	129.13
γ (°)	90.00	90.00	90.00	90.00	90.00	90.00	90.00
Temperature (K)	100	100	100	100	100	100	100
Resolution (Å)	1.58	2.703	1.761	1.542	1.661	1.381	1.738
# Reflections	39019	7938	28121	41942	31355	57375	56480
Completeness (%)	99.75	98.72	99.66	99.67	93.87	99.76	99.86
R_{work}(%)	15.50	16.26	15.92	16.11	17.24	15.51	14.59
R_{free} (%)	17.70	24.70	18.53	19.37	20.76	17.36	17.89
Minimum $F_{obs}/\sigma F_{obs}$	1.33	1.34	1.35	1.34	1.33	1.36	1.39
RMSD Bonds (Å)	0.006	0.007	0.007	0.007	0.005	0.005	0.007
RMSD Angles (Å)	1.107	0.937	0.962	1.166	0.871	0.872	1.066
# Protein Atoms	2066	2055	2111	2069	2010	2063	4042
# Water Atoms	270	24	218	232	256	301	547

Supplemental Table S4. Caspase-3 mutants used in DRoP analysis

Mutant	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)	PDB code	Space group	Resolution (Å)	# Waters	Reference
E124A	0.14	27	5190	4EHH	I222	1.78	239	(14)
E124A,V266H	ND	ND	ND	4EHL	C2	1.86	369	(14)
Y197C	0.1	9.7	$1.05x10^4$	4EHD	I222	1.64	263	(14)
Y197C,V266H	0.02	12	$1.67x10^3$	4EHF	I222	1.72	207	(14)
E124A,Y197C	0.008	30	$2.67x10^2$	4EHK	C2	1.68	401	(14)
E124A,Y197C,V266H	0.02	22	$9.1x10^2$	4EHN	I222	1.69	261	(14)
E190A	1.2	18	$6.9x10^4$	4QTY	I222	1.60	237	(15)
K137A	0.98	17	$5.8x10^4$	4QUB	I222	1.69	279	(15)
F128A	0.054	15	$3.6x10^3$	4QU9	I222	1.56	303	(15)
F128A,V266H	0.003	12	$2.1x10^2$	4QUI	C121	1.76	461	(15)
F55W	0.004	36	$1.1x10^2$	4QU2	C2	1.90	207	(15)
M61A	1.04	27	$3.8x10^4$	4QUG	C121	1.92	423	(15)
M61A,V266H	0.001	19	$5.4x10^1$	4QU8	I222	1.72	210	(15)
T140G	1.15	16	$7.2x10^4$	4QUH	C121	1.76	421	(15)
T140G,V266H	0.0002	36	6.1	4QUJ	I222	1.50	249	(15)
T140V	0.027	12	$2.3x10^3$	4QU5	I222	1.91	154	(15)
Y195F	1.18	28	$4.2x10^4$	4QUA	I222	1.89	212	(15)
Y195F,V266H	0.001	41	$2.2x10^1$	4QUE	C121	1.84	304	(15)
E167A	0.12	13	$9.1x10^3$	2J31	I222	1.50	204	(8)
E173A	0.69	2.4	$2.9x10^5$	2J32	I222	1.30	306	(8)
Y203F	0.68	5.9	$1.2x10^5$	2J33	I222	2.00	252	(8)
K242A	0.04	13	$3.1x10^3$	3PD1	I222	1.62	268	(16, 17)
E246A	0.3	23	$1.3x10^4$	3PD0	I222	2.00	215	(16, 17)
K242A,E246A	0.004	31	$1.3x10^2$	3PCX	I222	1.50	317	(16, 17)
V266A	2.0	20	$1.0x10^5$	5I9B	I222	1.80	218	This manuscript

Supplemental Table S4 Continued

V266S	0.078	34.8	$2.24x10^3$	5IAK	I222	1.58	270	This manuscript
V266C	0.17	30	$5.66x10^3$	5I9T	C121	1.95	347	This manuscript
V266D	0.03	14.4	$2.08x10^3$	5IAB	C121	1.79	351	This manuscript
V266E	0.14	4.7	$3.0x10^4$	3ITN	I222	1.63	89	(12, 18)
V266Y	1.36	18.2	$7.47x10^4$	5IAS	I222	1.54	232	This manuscript
V266F	0.006	13	$4.60x10^2$	5IAE	C121	1.55	564	This manuscript
V266W	0.02	22	$9.04x10^2$	5IAR	I222	1.76	218	This manuscript
V266H	ND	ND	ND	4EHA	C2	1.70	386	(14)
V266L	0.50	55	$9.0x10^3$	5IAJ	I222	1.58	195	This manuscript
V266I	0.04	25	$1.6x10^3$	5IBC	I222	1.66	256	This manuscript
V266M	2.90	25.2	$1.15X10^5$	5IBP	I222	1.38	301	This manuscript
V266K	1.40	52.5	$2.66X10^4$	5IBR	C121	1.74	547	This manuscript
Total Waters								**10,736**

Supplemental Table S5. Plasmids and primers used to construct V266 library

Plasmid name	Mutant	Parent plasmid	Forward (5'-3')	Reverse(5'-3')
pHC33281	V266G	pHC332 [(1)]	CCATGTATTGGATCCATGCTCACAAAAG	CTTTTGTGAGCATGGATCCAATACATGG
pHC33282	V266G D_3A [(3)]	pHC33209 [(2)]	CCATGTATTGGATCCATGCTCACAAAAG	CTTTTGTGAGCATGGATCCAATACATGG
pHC33283	V266A	pHC332	CCATGTATTGCATCGATGCTCACAAAAG	CTTTTGTGAGCATCGATGCAATACATGG
pHC33284	V266A D_3A	pHC33209	CCATGTATTGCATCGATGCTCACAAAAG	CTTTTGTGAGCATCGATGCAATACATGG
pHC33285	V266S	pHC332	CCATGTATTTCATCGATGCTCACAAAA	CTTTTGTGAGCATCGATGAAATACATG
pHC33286	V266S D_3A	pHC33209	CCATGTATTTCATCGATGCTCACAAAA	CTTTTGTGAGCATCGATGAAATACATG
pHC33287	V266T	pHC332	CCATGTATTACTAGTATGCTCACAAAAG	CTTTTGTGAGCATACTAGTAATACATG
pHC33288	V266T D_3A	pHC33209	CCATGTATTACTAGTATGCTCACAAAAG	CTTTTGTGAGCATACTAGTAATACATG
pHC33289	V266I	pHC332	CCATGTATTATATCGATGCTCACAAAAG	CTTTTGTGAGCATCGATATAATACATG
pHC33290	V266I D_3A	pHC33209	CCATGTATTATATCGATGCTCACAAAAG	CTTTTGTGAGCATCGATATAATACATG
pHC33291	V266L	pHC332	CCATGTATTTTATCGATGCTCACAAAAG	CTTTTGTGAGCATCGATAAAATACATGG
pHC33292	V266L D_3A	pHC33209	CCATGTATTTTATCGATGCTCACAAAAG	CTTTTGTGAGCATCGATAAAATACATGG
pHC33293	V266C	pHC332	CAGATTCCATGTATCTGCAGCATGCTCACAAAAGAACTC	GAGTTCTTTTGTGAGCATGCTGCAGATACATGGAATCTG

Supplemental Table S5 Continued

pHC33294	V266C D_3A	pHC33209	CAGATTCCATGTATC TGCAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTGCAGATA CATGGAATCTG
pHC33295	V266D	pHC332	CAGATTCCATGTATC GATTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGAATCGATA CATGGAATCTG
pHC33296	V266D D_3A	pHC33209	CAGATTCCATGTATC GATTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGAATCGATA CATGGAATCTG
pHC33297	V266Q	pHC332	CAGATTCCATGTATT CAATCGATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATCGATTGAATA CATGGAATCTG
pHC33298	V266Q D_3A	pHC33209	CAGATTCCATGTATT CAATCGATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATCGATTGAATA CATGGAATCTG
pHC33299	V266K	pHC332	CAGATTCCATGTATT AAATCGATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATCGATTTAATA CATGGAATCTG
pHC33100	V266K D_3A	pHC33209	CAGATTCCATGTATT AAATCGATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATCGATTTAATA CATGGAATCTG
pHC33101	V266Y	pHC332	CAGATTCCATGTATC TACAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTGTAGATA CATGGAATCTG
pHC33102	V266Y D_3A	pHC33209	CAGATTCCATGTATC TACAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTGTAGATA CATGGAATCTG
pHC33103	V266M	pHC332	CAGATTCCATGTATC ATGAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTCATGATA CATGGAATCTG
pHC33104	V266M D_3A	pHC33209	CAGATTCCATGTATC ATGAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTCATGATA CATGGAATCTG
pHC33105	V266N	pHC332	CAGATTCCATGTATT AATTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGAATTAATA CATGGAATCTG

Supplemental Table S5 Continued

pHC332 106	V266 N D_3A	pHC332 09	CAGATTCCATGTATT AATTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGAATTAATA CATGGAATCTG
pHC332 107	V266 W	pHC332	CAGATTCCATGTATT TGGTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGAG(TGGACCAAATACATG AATCTG
pHC332 108	V266 W D_3A	pHC332 09	CAGATTCCATGTATT TGGTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGACCAAATA CATGGAATCTG
pHC332 109	V266 F	pHC332	CAGATTCCATGTATT TTTTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGAAAAAATA CATGGAATCTG
pHC332 110	V266 F D_3A	pHC332 09	CAGATTCCATGTATT TTTTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGGAAAAAATA CATGGAATCTG
pHC332 111	V266 P	pHC332	CAGATTCCATGTATT CCATCGATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATCGATGGAATA CATGGAATCTG
pHC332 112	V266 P D_3A	pHC332 09	CAGATTCCATGTATT CCATCGATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATCGATGGAATA CATGGAATCTG
pHC332 20	V266 R	pHC332	CAGATTCCATGTATT CGTAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTACGAATA CATGGAATCTG
pHC332 22	V266 R D_3A	pHC332 09	CAGATTCCATGTATT CGTAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTACGAATA CATGGAATCTG
pHC332 03	V266 H	pHC332	CAGATTCCATGTATT CATAGCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGA GCATGCTATGAATA CATGGAATCTG
pHC332 04	V266 E	pHC332 09	CAGATTCCATGTATT GAGTCCATGCTCACA AAAGAACTC	GAGTTCTTTTGTGG ACTCAATACATGGA ATCTG

(1) Plasmid pHC332 is described in reference (19)

(2) Plasmid pHC33209 is described in reference (13)

(3) D_3A refers to uncleavable procaspase-3(D9A,D28A,D175A) as described in reference (13)

REFERENCES

1. Crawford ED, Wells JA (2011) Caspase substrates and cellular remodeling. *Annu Rev Biochem* 80:1055–1087.
2. Schwerk C, Schulze-Osthoff K (2003) Non-apoptotic functions of caspases in cellular proliferation and differentiation. *Biochem Pharmacol* 66:1453–1458.
3. D'amelio M, Cavallucci V, Cecconi F (2010) Neuronal caspase-3 signaling: not only cell death. *Cell Death Differ* 17:1104–1114.
4. Han M-H, et al. (2013) The novel caspase-3 substrate Gap43 is involved in AMPA receptor endocytosis and long-term depression. *Mol Cell Proteomics* 12:3719–3731.
5. Edelmann B, et al. (2011) Caspase-8 and caspase-7 sequentially mediate proteolytic activation of acid sphingomyelinase in TNF-R1 receptosomes. *EMBO J* 30:379–394.
6. Zandy A, Lakhani S, Zheng T, Flavell RA, Bassnett S (2005) Role of the executioner caspases during lens development. *J Biol Chem* 280:30263–30272.
7. Makishima T, et al. (2011) Inner ear dysfunction in caspase-3 deficient mice. *BMC Neurosci* 12:1–9.
8. Boatright KM, Salvesen GS (2003) Mechanisms of caspase activation. *Curr Opin Cell Biol* 15:725–731.
9. MacKenzie SH, Clark AC (2012) Death by caspase dimerization. *Adv Exp Med Biol* 747:55–73.

10. Jiao S, Li Z (2011) Nonapoptotic Function of BAD and BAX in long-term depression of synaptic transmission. *Neuron* 70:758–772.

11. Parrish A, Freel CD, Kornbluth S (2013) Cellular mechanisms controlling caspase activation and function. *Cold Spring Harb Perspect Biol* 2013:1–24.

12. Chai J, et al. (2001) Structural basis of caspase-7 inhibition by XIAP. *Cell* 104:769–780.

13. Riedl SJ, et al. (2001) Structural basis for the inhibition of caspase-3 by XIAP. *Cell* 104:791–800.

14. Eckelman BP, Salvesen GS (2006) The human anti-apoptotic proteins cIAP1 and cIAP2 bind but do not inhibit caspases. *J Biol Chem* 281:3254–3260.

15. Choi YE, et al. (2009) The E3 ubiquitin ligase cIAP1 binds and ubiquitinates caspase-3 and -7 via unique mechanisms at distinct steps in their processing. *J Biol Chem* 284:12772–12782.

16. Kurokawa M, Kornbluth S (2009) Caspases and kinases in a death grip. *Cell* 138:838– 854.

17. Voss OH, Kim S, Wewers MD, Doseff AI (2005) Regulation of monocyte apoptosis by the protein kinase Cdelta-dependent phosphorylation of caspase-3. *J Biol Chem* 280:17371–17379.

18. Dix M, et al. (2012) Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. *Cell* 150:426–440.

19. Alvarado-Kristensson M, et al. (2004) p38-MAPK signals survival by phosphorylation of caspase-8 and caspase-3 in human neutrophils. *J Expimental Med* 199:449–458.

20. Fuentes-Prior P, Salvesen GS (2004) The protein structures that shape caspase activity, specificity activation and inhibition. *Biochem J* 384:201–232.

21. Romanowski MJ, Scheer JM, O'Brien T, McDowell RS (2004) Crystal structures of a ligand-free and malonate-bound human caspase-1: implications for the mechanism of substrate binding. *Structure* 12:1361–1371.

22. Walters J, Schipper JL, Swartz P, Mattos C, Clark AC (2012) Allosteric modulation of caspase-3 through mutagenesis. *Biosci Rep* 32:401–411.

23. Cade C, Swartz P, MacKenzie SH, Clark AC (2014) Modifying caspase-3 activity by altering allosteric networks. *Biochemistry* 53:7582–95.

24. Thomsen ND, Koerber JT, Wells JA (2013) Structural snapshots reveal distinct mechanisms of procaspase-3 and -7 activation. *Proc Natl Acad Sci* 110:8477–8482.

25. Clark AC (2016) Caspase allostery and conformational selection. *Chem Rev* in press.

26. Hardy JA, Lam J, Nguyen JT, O'Brien T, Wells JA (2004) Discovery of an allosteric site in the caspases. *Proc Natl Acad Sci USA* 101:12461–12466.

27. Scheer JM, Romanowski MJ, Wells JA (2006) A common allosteric site and mechanism in caspases. *Proc Natl Acad Sci USA* 103:7595–7600.

28. Schweizer A, et al. (2007) Inhibition of caspase-2 by a designed ankyrin repeat protein: specificity, structure, and inhibition mechanism. *Structure* 15:625–636.

29. Velazquez-Delgado EM, Hardy JA (2012) Zinc-mediated allosteric inhibition of caspase-
6. *J Biol Chem* 287:36000–36011.

30. Kearney BM, Johnson CW, Roberts DM, Swartz P, Mattos C (2014) DRoP: a water analysis program identifies Ras-GTP-specific pathway of communication between membrane-interacting regions and the active site. *J Mol Biol* 426:611–629.

31. Levy Y, Onuchic JN (2006) Water mediation in protein folding and molecular recognition. *Annu Rev Biophys Biomol Struct* 35:389–415.

32. Frauenfelder H, et al. (2009) A unified model of protein dynamics. *Proc Natl Acad Sci USA* 106:5129–34.

33. Colombo MF, Rau DC, Parsegian VA (1992) Protein solvation in allosteric regulaton: A water effect on hemoglobin. *Science* 256:655–659.

34. Motlagh H, Wrabl J, Li J, Hilser V (2014) The ensemble nature of allostery. *Nature* 508:331–339.

35. Davydov DR, Baas BJ, Sligar SG, Halpert JR (2007) Allosteric mechanisms in cytochrome P450 3A4 studied by high-pressure spectroscopy: Pivotal role of substrateinduced changes in the accessibility and degree of hydration of the heme pocket. *Biochemistry* 46:7852–7864.

36. Zhuravlev PI, Papoian G a (2010) Protein functional landscapes, dynamics, allostery: a tortuous path towards a universal theoretical framework. *Q Rev Biophys* 43:295–332.

37. Salvay AG, Grigera JR, Colombo MF (2003) The role of hydration on the mechanism of allosteric regulation: in situ measurements of the oxygen-linked kinetics of water binding to hemoglobin. *Biophys J* 84:564–570.

38. Feeney B, Pop C, Swartz P, Mattos C, Clark AC (2006) Role of loop bundle hydrogen bonds in the maturation and activity of (pro)caspase-3. *Biochemistry* 45:13249–13263.

39. Witkowski WA, Hardy JA (2009) L2' loop is critical for caspase-7 active site formation. *Protein Sci* 18:1459–1468.

40. Datta D, Scheer JM, Romanowski MJ, Wells JA (2008) An allosteric circuit in caspase-1. *J Mol Biol* 381:1157–1167.

41. Huang Z, Pinto JT, Deng H, Richie JP (2008) Inhibition of caspase-3 activity and activation by protein glutathionylation. *Biochem Pharmacol* 75:2234–2244.

42. Dagbay K, et al. (2014) A multipronged approach for compiling a global map of allosteric regulation in the apoptotic caspases. *Methods Enzymol* 544:215–49.

43. Hardy JA (2012) Phosphorylation Regulates Assembly of the Caspase-6 SubstrateBinding Groove. *Structure* 7:742–751.

44. Cao Q, et al. (2012) Inhibitory mechanism of caspase-6 phosphorylation revealed by crystal structures, molecular dynamics simulations, and biochemical assays. *J Biol Chem* 287:15371–15379.

45. Pop C, Feeney B, Tripathy A, Clark AC (2003) Mutations in the procaspase-3 dimer interface affect the activity of the zymogen. *Biochemistry* 42:12311–12320.

46. Walters J, et al. (2009) A consitutively active and uninhibitable caspase-3 zymogen efficiently induces apoptosis. *Biochem J* 424:335–345.

47. Feeney B, Pop C, Tripathy A, Clark AC (2004) Ionic interactions near loop L4 are important for maintaining the active site environment and the dimer stability of (pro)caspase-3. *Biochem J* 384:515–525.

48. Riedl SJ, et al. (2001) Structural basis for the activation of human procaspase-7. *Proc Natl Acad Sci USA* 98:14790–14795.

49. Chai J, et al. (2001) Crystal structure of a procaspase-7 zymogen: Mechanisms of activation and substrate binding. *Cell* 107:399–407.

50. Pop C, et al. (2001) Removal of the pro-domain does not affect the conformation of the procaspase-3 dimer. *Biochemistry* 40:14224–14235.

51. Bose K, Clark AC (2001) Dimeric procaspase-3 unfolds via a four-state equilibrium process. *Biochemistry* 40:14236–14242.

52. Bose K, Pop C, Feeney B, Clark AC (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.

53. Walters J, Swartz P, Mattos C, Clark AC (2011) Thermodynamic, enzymatic and structural effects of removing a salt bridge at the base of loop 4 in (pro)caspase-3. *Arch Biochem Biophys* 508:31–38.

54. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr* 66:213–221.

55. Pronk S, et al. (2013) GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29:845–854.

56. Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074.

57. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.

CHAPTER 3:

UTILIZING CRISPR FOR THE INCORPORATION OF CASPASE-3 MUTANTS INTO HEK293 CELLS

ABSTRACT

Recent innovations in genomic editing provides a rapid and reliable technique for the ablation of genes and incorporation of mutations. CRISPR/Cas technology allows for the incorporation of caspase-3 genes into cellular and animal models. Genomic editing will further our understanding of how cells modulate threshold levels of caspase activity. Regulation of caspase-3 provides the ability to discern specific caspase involvement in disease and is vital when striving for limited off target effects during disease treatment. Discussed herein is a brief summation of CRISPR/Cas technology and an accompanying strategy for the generation of a homogenous population of HEK293 cells containing mutated caspase-3 genes.

INTRODUCTION TO CRISPR

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs) are short repeating motifs found in the genome of prokaryotes ranging in size from 25-48 nucleotides (1,2,3,7). Initial studies focused on the function and origin of these elements elucidated that CRISPRs are well represented throughout prokaryotes, citing that 40% of eubacteria and 90% of archaea possess CRISPRs (1,2,7,11,15) . The short repeats are separated by regions of DNA coined as “spacers”. Many of the spacers, as noticed by Mujica and Bolotin, derived from a foreign origin, particularly bacteriophage (1,2). Upon examining the level of sequence conservation exhibited among spacer loci came the

elegant position that the DNA, integrated into spacers, are being utilized in a functional role and in a conserved manner among varying organisms (1). Furthermore it was observed that CRISPR loci were in close proximity to CRISPR-associated genes (*cas* encoding genes) and that the repeats and enzymes were functionally connected. The Cas enzymes were predicted to have helicase and endonuclease activity based on their sequence (2). These early studies led to the recognition of the CRISPR-Cas system as a mechanism of prokaryotic adaptive immune response. Broadly, prokaryotes incorporate invasive DNA into their genome in order to recognize and digest invading genetic elements (1,3,4,7,9,12). The capability of a bacteria to become resistant towards an invasive phage directly correlates to the organism's ability to facilitate the integration of bacteriophage DNA into its genome (4). The region of an invasive molecule destined for incorporation as a spacer is designated by its proximity to a proto-spacer-adjacent motif (PAM). PAMs are short nucleotide sequences, which differ among CRISPR-Cas systems (12). PAM sequences aid to discern self and non-self-recognition and are essential to host recognition of invasive genetic elements.

Horizontal/Lateral and vertical gene transfer also provides the ability to convey resistance, mediated by genomic material, among organisms (2,7). It should be noted that the CRISPR-Cas demonstrates both anti-DNA and anti-RNA survival strategies (6,7).

The initial discovery and classification of CRISPR loci occurred more than twenty years ago yet only recently have the mechanisms governing DNA incorporation and subsequent viral clearance been described (1,6). CRISPR investigators have classified the adaptive immune response by defining three major pathways required for host response. The first process is described as the acquisition of a novel spacer from an

invasive DNA molecule, secondly is the expression of CRISPR-Cas system and lastly is the interference of invasive molecules (13). Acquisition of novel spacer regions is poorly understood but it is evident that the enzymes Cas1 and Cas2, which are conserved throughout type I, II and III CRISPR-Cas systems, are the sole enzymes required for the incorporation of foreign DNA into the host genome. Cas1 and Cas2 both possess nuclease activity but it is reported that only Cas1 nuclease activity is required for spacer incorporation. Cas2 is also essential but it is currently recognized as a recruiter of Cas1 (15). The activation of the acquired CRISPR motifs occurs through their transcription into CRISPR-RNA (crRNA) which is used in complex with Cas enzymes to facilitate the recognition of viral DNA via base pairing (hybridization). Initially, pre-crRNA (molecule consisting of multiple crRNA units) is transcribed from the CRISPR motifs and is cleaved by Cas enzymes or RNaseIII, depending on the system, to generate mature crRNAs. Mature crRNAs are composed of a repeating unit and a spacer region transcribed from their corresponding CRISPR DNA precursors (5, 8). The repeating units of the CRISPRs function as binding sites for Cas proteins while the spacer regions confer specificity to the complex in order for the recognition and subsequent clearance of various foreign DNA and RNA molecules (5,6,7). Therefore, within the complex crRNAs can be thought of as "guides" of the functional effector Cas enzymes (5,7). CRISPR-Cas systems differ among organisms and therefore utilize varying yet related enzymes for the culmination of a cellular immune response. The systems have been organized into three primary groups; Type I, II, and III. Types I and II target foreign or invasive DNA for degradation while Type III target RNA for cleavage. Additionally, types I, II and III are delineated by their crRNA maturation methods and corresponding effector enzyme which

is essential to DNA or RNA cleavage, Cas3, Cas9/Csn1/Cas5 and Cas10 respectively (11,12). A detailed summation of type I and III systems will not be further addressed, additionally it should be noted that the classification of CRISPR-Cas systems contains further subgroupings as well as unclassified systems (12).

The Type II system demonstrates a unique set of characteristics which delineate it from its counterpart types I and III. Firstly, type II employs a dissimilar method for the maturation of crRNA through the transcription of a trans-activating crRNA (tracrRNA). tracrRNA is a small RNA molecule which complimentarily base pairs to the pre-crRNA repeat and directs an RNaseIII dependent cleavage event of dsRNA thus facilitating maturation of the crRNA (Type I and III utilize Cas6 endonucleases for crRNA maturation) (10, 13). Cas9 is also required for the initial annealing of tracrRNA to crRNA (13). The resulting hybridized complex between tracrRNA and a mature crRNA (tracrRNA:crRNA) forms an RNA structure which is bound to Cas9 (10). The absence of tracrRNA, RNaseIII or Cas9 results in the attenuation of pre-crRNA processing and compromises host immune capabilities (8). Furthermore, Type II mediated DNA cleavage is dependent on the activity of Cas9, a protein which mediates the maturation of crRNA and generates double stranded breaks upon complex recognition of DNA targets.

Cas9 is comprised of two lobes, coined by Nishimasu et al as the recognition (REC) and the nuclease (NUC) lobes. The Rec lobe assists in the recognition of the tracRNA:crRNA duplex and forms one side of the positively charged groove to which the DNA:RNA complex is bound. The NUC domain is comprised of two nuclease domains: RuvC and HNH as well as a PAM interacting domain (PI) and forms the complimentary side of the positively charged DNA binding groove. Cas9 utilizes two functional nuclease

domains, RuvC and HNH to induce a double stranded break at a site three base pairs up stream of the PAM, where each active site is responsible for the cleavage of one strand of a complimentary pair of DNA molecules (9,10,12, 25). HNH cleaves DNA bound complimentary to the tracrRNA:cRNA complex while RuvC cleaves the DNA molecule which is not hybridized to RNA (non-complimentary) (10,24). DNA recognition by the Cas9:tracrRNA:crRNA complex is reliant on the inclusion of a spacer adjacent to a PAM sequence on the non-complimentary strand of targeted DNA. PAMs interact with the PI domain in the NUC lobe and allow for the specific recognition of target strand as well as determination of self vs. non-self. (9,10,24). Cas9, tracrRNA and crRNA co-exist as an RNA-guided endonuclease (9). The transcription of both pre-crRNA, tracrRNA and the subsequent RNaseIII mediated maturation of the complex limits researcher's abilities to efficiently employ CRISP-Cas system for genomic editing. Recently the realization that this maturation event could be circumvented has led to the widespread use of CRISPR-CAS systems (10,16)

The initial innovation of a tracrRNA:crRNA chimera or single guide RNA (sgRNA) has facilitated the ease and the versatility of CRISPR-Cas usage in genomic editing(10,16). Controlling sgRNA transcription through the U6 polymerase III promoter enables transcription of a single RNA which does not require cleavage yet maintains canonical structural motifs, necessary for binding Cas9(10,16). The requirement of pre-crRNA maturation, RNaseIII mediated cleavage and multiple constructs for genetic editing is no longer a concern (10). sgRNA chimera complexes can readily be cloned to incorporate any desired crRNA targeting transcript (10,16). Additionally, a Cas9 enzyme which has been human optimized is available and facilitates translation in human cells as

organisms exhibit codon bias (16,22). The combination of easily manipulated sgRNA chimeras and the human optimized Cas9 further simplifies genetic editing techniques in the context of CRISPR-Cas type II system.

Genomic editing allows researchers to ascertain the functional contributions of a gene in cellular mechanisms. Until recently investigators have utilized non-specific and tedious mutation methods to facilitate genome alterations. Through the use of meganucleases, Zinc finger nucleases (ZFNs), transcription-activator like effector nucleases (TALENs) and CRISPR-Cas systems specific gene modulation can be achieved (17,18,21). CRISPR-Cas systems have ushered in the wide spread use of genomic editing (21). Only the CRISPR-Cas system utilizes an RNA mediated DNA recognition method, while the other nucleases implore a protein-DNA strategy. Therefore, retargeting can be rapidly achieved through the modulation of sgRNA, a luxury not shared by meganucleases, TALENs nor ZFNs (16,19,21). A recent study sought to compare the relative efficiencies of TALEN vs. CRISPR-Cas mediated genomic editing and found that CRISPR-Cas performed more favorably most likely due to the helicase activity attributed to Cas9 and the ability of cells to tolerate Cas9 (19.21). Additionally exploiting CRISPR systems as genetic editing tools provides a more precise cleavage pattern than that of TALENS and presents the ability to rapidly change target sites or to target multiple sites simultaneously (25). However CRISPR-Cas systems are not without imperfection, targeted sequences for editing must be adjacent to a PAM sequence and the potential for off target effects, while controversial, still remains a possibility (23,25).

The ability to mutate aberrant genes, whether to silence or repair, provides a new means of disease control yet more research is needed to enhance the effectiveness of

novel therapeutics (18,20). The sequence specific nucleases introduce double stranded breaks in targeted genes which causes the induction of DNA repair pathway activation. Both homologous directed repair (HDR) and non-homologous end joining (NHEJ) pathways are necessary in the functional alteration of target genes (17,18,19). By providing nucleotide templates to cells challenged with site specific nucleases allows for the incorporation of desired sequences by the HDR pathway. Conversely, the NHEJ pathway, which combines nucleic acid strands non-specifically results in the insertion or deletion (indels) of genomic material leading to frame shift alterations and the production of non-functional gene products (17,18,21). During DNA repair NHEJ occurs more habitually than that of HDR and research focused on shifting this propensity towards HDR may provide a more efficient approach during genomic editing (23).

STRATEGY FOR THE GENERATION OF A HOMOGENEOUS POPULATION OF MUTATED HEK293 CELLS

The generalized approach for the isolation of a homogenous population of mutated caspase-3 at position 266 is described in Fig.1. Briefly, a single plasmid construct containing a Cas9, green fluorescence protein (GFP) and a sgRNA is transfected simultaneously with single stranded oligonucleotides (ssODNs) into HEK293 cells(25). Cas9 containing plasmids are designed to introduce double stranded breaks near the target site while the ssODN enables specific editing of the caspase-3 gene. Cells are assayed by their expression of GFP which occurs at a stoichiometric 1:1 ratio with that of Cas9. Positively transfected cell cultures are diluted in 96 well plate in order to isolate single colonies. Single colonies are allowed to grow to confluence before being

clonally expanded into 12 well plates (Fig. 2). Expanded cells are then assayed for their incorporation of the ssODN into chromosome 4. Positively identified cultures are then expanded in 75 cm^2 flasks and aliquoted for storage in liquid nitrogen (-196 C).

Selecting Targeting Region

The caspase-3 gene is located on chromosome 4 (4q34). Positions 184,629,328-184,629,331 encompass the center of the dimer interface (V266) which is the targeted site for mutation. As discussed previously genetic material designated for cleavage must be adjacent to a PAM sequence. To assist in the determination of the locations of PAM sequences specific to spCas9, motif NGG, a sequence of up to 500 nucleotides flanking V266 on either side is submitted to a CRISPR optimization tool (http://crispr.mit.edu). This tool aids in the determination of possible sgRNA designs and ranks them based on their potential for off target effects. After submission using the Zhang CRISPR design tool three sgRNA sequences were selected based on proximity to V266 and limited off target potentiality (Fig. 3)

CRISPR CLONING STATEGY

Cloning pSpCas9(BB)-2A-GFP (PX458)

(Plasmid #48138) The PX458 plasmid purchased from Addgene codes for the Cas9 protein of S. pyogenes followed by 2A-EGFP. Cas9/EGFP expression is controlled via a CBh promoter. Upon translation of the mRNA Cas9-2A-EGFP, two independent polypeptides are produced at stoichiometric equivalents (Cas9 and EGFP). Additionally, contained within the plasmid pX458, downstream of U6 promoter, is the sgRNA cloning site which is flanked by Bbs1 restriction sites (Fig. 4). Through the ligation of a

sequence of interest into the Bbsl cloning site a unique sgRNA chimera can be generated. There is a Bbs1 restriction endonuclease site downstream of the U6 Promoter (Green) and upstream of the sgRNA scaffold (red). Bbs1 cleaves sites are indicated by blue arrows(25).

Three distinct oligo pairs were designed based on three candidates chosen from the Zhang analysis. The first oligo of each pair contains a 5' sequence CACC which binds complimentary to a Bbs1 overhang resulting from digestion of the pX458 plasmid. The oligo also contains twenty nucleotides upstream of the PAM sequence at each locus. The second oligo of each pair is designed to base pair complimentary to the twenty nucleotides of the first oligo (excluding the Bbs1 overhang binding sequence). The second oligo also contains a 5' AAAC overhang which allows for subsequent ligation of both oligo pairs into the pX458 plasmid (Fig, 4). The oligos listed in table 1 were purchased from IDT. Each pair is designed to target a specific nucleotide sequence within the human genome (Fig. 3) (Table 1). Oligo pairs must be modified by 5' phosphorylation prior to being annealed within the PX458 plasmid. Upon receipt of shipment, each single stranded oligo is reconstituted to a concentration of 100 µM. 10µM "working stocks" are then made by diluting the original stock ($1/10^{th}$ dilution). 25 µL of each (10µM) oligo is mixed with its complimentary oligo pair. Utilizing an Eppendorf Thermocycler it is possible to anneal each oligo pair by first providing an environment which promotes the denaturation of the oligo mixture, heating to 95°C, and allowing the mixture to slowly cool, annealing program termination at a temperature of 25°C. It is important to note that during the cooling process a gradual decrease in temperature is advised. Annealing procedure is listed below.

Annealing Procedure

Program “Annealing of Oligos – V266”, programmed into the Eppendorf Thermocycler in A. Clay Clark’s Laboratory at The University of Texas at Arlington. Briefly, the annealing program begins at 95°C for five minutes subsequently decreasing temperature by 1°C over the course of 15 seconds. The samples are then incubated at the new temperature ((95°C-(n)1°C), where n represents a given step number, for 1 minute before decreasing by another 1°C for 15 seconds. Oligo mixture samples are incubated at each 1°C interval for 1 min, gradually decreasing until reaching the final temperature of 25°C.

95°C: 5:00 min

95°C →94°C: 00:15 min

94°C: 1:00 min

94°C→93°C: 00:15 min

93°C: 1:00 min

93°C→92°C: 00:15

……*

…….*

….*

..*

26°C→25°C: 00:15 min

25°C: ∞

*Steps not shown (92°C through 26°C)

Annealed Oligos are labeled accordingly to their Oligo Abbreviation listed below (Table 1) and stored at -20°C in Box labeled "Joe Maciag Crispr Cloning" in A. Clay Clark's Laboratory at The University of Texas at Arlington.

Digestion of PX458 and Subsequent Ligation of Oligos into the PX458 Plasmid

PX458 was shipped via bacterial stab form Addgene, upon receipt the bacterial stab was pierced with a flame serialized loop and bacteria therein were incubated on a solid, Ampicillin containing, Agar/Lysogeny Broth (LB) plate (1% NaCl, 1% Tryptone, .5% Bacto-Yeast Extract, 1.5% Agar w/v and 50µg/mL Ampicillin) for 16 hours at 37°C. Upon completion of incubation a single colony was isolated and inoculated into a 100 mL culture of Lysogeny Broth (1% NaCl, 1% Tryptone, .5% Bacto-Yeast Extract w/v and 50µg/mL Ampicillin) for 16hrs at 37°C while gyrating. 1.6 mL aliquots were then pipetted from the 100 mL culture and transferred to 2 mL long-term storage vials and mixed with 400 mL of glycerol. The 80% Bacteria Culture (containing PX458) /20% Glycerol solution was then flash frozen in liquid nitrogen and stored in a -80°C freezer in A. Clay Clark's Laboratory at The University of Texas at Arlington.

Isolation of the pX458 plasmid from bacterial cells was initiated by T-streaking of the same Agar/LB plates described above using the same incubation procedures. 100 mL liquid growth medium (described above) was inoculated with an isolated colony produced from T-streaked plates. Liquid cultures were incubated for 16 hours at 37°C, bacterial cells were isolated, lysed and discarded in order to isolate pX458 plasmid. Briefly, isolation of DNA was achieved by utilization of the QIAprep Spin Miniprep Kit.

QIAprep kits take advantage of alkaline lysis, acidic neutralization and silica dependent binding of DNA. Isolation of pX458 was attained to a concentration of ≈ 100 ng/uL.

To obtain a newly constructed pX458 plasmid able to direct cleavage towards the human caspase-3 genome it is imperative to ligate the oligo pairs described previously into the isolated pX458 plasmid. In order to achieve this 5'cohesive ends of our oligo pairs must have a complimentary base pairing partner represented in the PX458 plasmid. Through endonuclease digestion of pX458 with the enzyme Bbsl, a complementary site will be available to ligate oligo pairs. In a reaction volume of 30 µL, 20 µL of isolated plasmid (≈2 µg), 3 µL of 10X NE Buffer 1 (provided with Bbsl upon purchase), 6 µL molecular grade H20 and 1 µL Bbsl enzyme (≈10 units) were incubated at 37°C for 2 hours. 1 unit of Bbsl enzyme is estimated to be sufficient to cleave 2 µg of DNA/hr. Complete digestion of pX458 was visually observed through gel electrophoresis (1% agarose w/v) and subsequent EtBr (ethidium bromide) staining.

Purification of the newly digested vector was attempted many times using the QIAquick gel extraction kit. QIAquick purification averaged yields of 5-8 ng/µL (efficiency = 7.5%-12% respectively). Although successful ligation was achieved with one group of oligo pairs (HDC3O-A/B) we were unsuccessful in ligating the additional pairs. In order to achieve a high efficiency of recovery of the digested PX458, phenol:chloroform extraction was performed in the following manner.

1. Equal amounts of phenol:chloroform were premixed. An equal volume of phenol:chloroform solution was added to the digestion sample (30 µL) within a microcentrifuge tube (1.5 mL maximum volume). Contents were mixed via pipetting and were subsequently centrifuged for 15 seconds at 12,000 g.

Additional centrifugation was required on some samples as the aqueous and organic phases were not easily distinguished, such is the case when working with small volumes.

2. Aqueous phase was aspirated and transferred to a clean 1.5 mL microcentrifuge tube.
3. Steps 1-2 are repeated until there is no visible protein between states. (Often unnecessary).
4. An equal volume of chloroform is added to the aqueous phase and mixed via pipette. Sample is then centrifuged for an additional 15 seconds at 12,000 g. Aqueous phase is aspirated and transferred to a clean 1.5 mL microcentrifuge tube.
5. Estimate the total volume of sample and add $1/10^{th}$ total volume of 3 M sodium acetate and add 2 volumes of ice cold 100 % ethanol (prechilled), store the solution at O°C for 15-30 minutes.
6. Remove the supernatant.
7. Fill the microcentrifuge tube with 750 mL of 70% ethanol, centrifuge at 12,000 g for 2 minutes, remove the supernatant and allow the tube to air dry (completely).
8. Dissolve the DNA pellet in molecular grade H20 for subsequent ligation.

Using the above protocol, the digested PX458 concentration was estimated by A260 absorption to be purified with 30% efficiency. For every 2 µg of PX458 digested, 600 ng were recovered.

Ligation of the isolated (digested) vector PX458 with the oligo pairs was attained by mixing .02 pmol of vector(PX458)/.06 pmol of insert (Oligo pairs) or 1/3 ratio of

vector to insert based on suggestions from New England Bioscicence. The ligation reaction was as follows; 1 µL T4 DNA Ligase, 0.6 µL of 0.1 µM oligo stock (final concentration = .06 pmol/20 µL), 2.87 µL of .00696 µM stock (final concentration = .02pmol/20 µL), 2 µL 10x Ligase buffer (provided with purchase of T4 DNA Ligase) and 13.53 µL of nuclease free water. The reaction was incubated at 25°C for 3 hours. Due to the low yield and tedious nature of the aforementioned protocol the method below was optimized to yield plasmids with incorporated oligos without the need for purification of digested plasmid.

Step 1

Digest 1-2 ug Plasmid DNA with BBSI 37° C for 2 hr

Cas9 Plasmid DNA ----------XuL (1-2ug final)

BBSI----------1uL

BBSI Buffer----------2 uL

ddH20----------(17-X) uL

Ligation

Add the following reagents directly to the tube in step 1

Incubate for 1 hr 37° C

Ligase (T4)----------1 uL

Ligase Buffer----------2.5 uL (10x)

Annealed Oligos----------2 uL

Post incubation, the reaction was transformed into BL21 PlysS ultra competent cells. 2 µL of reaction was mixed with 20 µL of competent cells and incubated on ice for 20 minutes, the reaction was then "heat shocked" at 42°C for 1 minute, the reaction was returned to ice for 1 minute and incubated for 1 hour with 200 µL of LB at 37°C. Upon completion of incubation, 100 µL of LB/ BL21 PlysS cells was transferred to a solid media AGAR/LB containing 50 µg/mL Ampicillin plate. Plates were incubated for 16 hours at 37°C. Individual colonies are cored from the AGAR plates and cultured in 5 mL LB medium containing 50 µg/mL Ampicillin for 16 hours at 37°C. Plasmids were isolated using the same technique described above. In order to test for the incorporation of double stranded oligo nucleotides into the pX458 plasmid, incubation of plasmids with BbsI and subsequent agarose gel analysis allows for the visualization of digests vs. undigested plasmids(undigested plasmids suggest incorporation of oilgo pairs as the BbsI site no longer exists). Additionally plasmids that remained undigested post incubation with BbsI require sequencing using a U6 primer (5'-GGGCAGGAAGAGGGCCTAT-3')(25).

Plasmids that are positive for oligo incorporation are transformed into competent cells as described above and isolated using an EndoFree Plasmid Kit (Endotoxin Free) available through Qiagen. This technique provides a large quantity of plasmid which is endotoxin free, endotoxins may decrease transfection efficiencies in later procedures. Isolated plasmids should be aliquoted and stored at -20°C.

ssODN Design

The single stranded oligo nucleotide serves as a repair template for DNA which has been cleaved by Cas9. The ssODN is designed to incorporate any mutation desired. Fig. 5 describes the ssODN design for the mutation of V266H into the human genome. ssODN design is based on the work of Corn et al. whose work has led to the increased efficiency of ssODN repair during HDR (23). Once designed the ssODNs are purchased from IDT and upon receipt are stored in water at a concentration of 10 µM.

CELL CULTURE TECHNIQUES

It is important to adhere to the proper safety guidelines and a strict aseptic technique when working with mammalian cells.

Starting Passage from Stock

Cells are stored in liquid nitrogen in complete Dulbecco's Modification of Eagle's Medium (DMEM + 10% Fetal Bovine Serum (FBS)) + 5% DMSO (v/v) at a concentration of ≈ 1,000,000 cells/mL. To begin passaging a stock of stored cells thaw an aliquot at 37°C for minute, after thawing carefully add 1mL DMEM-complete (DMEM +10% FBS) to the thawed sample drop by drop. Add the entire volume of DMEM-complete and cells to a 15 mL conical tube and add an additional 3 mL of DMEM-Complete. Centrifuge the 15 mL conical tube at 1000g and aspirate the solution assuring that the pellet is undisturbed. Pipette 2 mL directly onto the pellet and re-suspend through pipetting up and down. Prepare two 75 cm^2 flasks by adding 4 mL of DMEM-complete to each. Using the newly re-suspended cells, pipette 1 mL from the conical tube and add it to each of the flasks. Gently rock the flask side to side to allow the cells to thoroughly

mix. Using a microscope, visualize the cells and make sure that they are homogeneously spread across the flask. Allow the flask to incubate in the incubation chamere (37°C, 5% CO_2) Check for confluency in a 24-36 hours.

Splitting a Population of Adherent Cells (HEK293)

Start by setting a water bath to 37°C and place DMEM-complete, trypsin and PBS into the water bath. Allow the reagents to incubate for 30 minutes in the water bath. While the reagents incubate prepare the rest of the materials, obtain two 75 cm^2 flasks, two 15 mL conical tubes, and place them in the bio-safety cabinet. Once the reagents in the water bath have reached temperature place them in the bio-safety cabinet as well. Remove the cell culture flask (cells intended for splitting) from the incubator and place it in the bio-safety cabinet. As these cells are adherent to the surface of the flask the excess media may be poured off into a waste funnel, add 3 mL of PBS buffer to the flask and rock back and forth. This step is the wash step and will assist in the removal of excess old media and dead cells which are non-adherent. Pour off the PBS into the waste funnel. Add .5 mL of trypsin to the flask and confirm the trypsin is spread across the surface of cells evenly, return the flask to the incubation chamber for no longer than 5 minutes. Once the cells are no longer adherent the liquid trypsin should be cloudy. Remove the flask and homogenize the cells with 3 mL of DMEM-complete, deposit the solution into one of the 15 mL conical tubes and centrifuge the sample at 1000g for 30 seconds. Aspirate off the media ensuring the pellet is remained undisturbed. Re-suspend the pellet in 4 mL of DMEM-complete and pipette 50 µL of the homogenous sample into a microcentrifuge tube. Count the number of cells in your sample using K2 Cellometer (Nexcelom). A flask containing 80,000-100,000 cells will reach confluency after 3 days,

a flask containing 200,000-250,000 will reach confluency after 2 days. Determine the amount of cells/mL and add the appropriate amount of cells to each flask. The flasks should contain a total of 5 mL total volume, i.e the amount of DMEM-complete in each flask would be calculated as 5 mL –X mL of cells. Gently rock the flasks back and forth ensuring a homogeneous mixture of cells and visualize under a microscope before placing the flasks back in the incubator for the amount of time desired (2-3 days).

Plating Cells and Transfection

Cells which have been grown to confluency and exhibit no signs of contamination (i.e bacterial growth) can be seeded in 12 or 24 well plates the day before transfection. Using the above mentioned protocol, determine the concentration of cells/mL and seed 50,000 or 100,000 cells per well in a 12 or 24 well plate respectively. The newly seeded plates should be placed for 24 hours in the incubator. Cells should be at 90% confluency after 24 hours. Following the Lipofectamine 3000 protocol suggested by the manufacturer, transfect 2.5 µg of plasmid (endotoxin free) and 1 µL of 10 µM ssODN stock into each well. Allow cells and transfection reagents to incubate for 2 days. After 2 days the cells can be re-suspended and assayed for the fluorescence of GFP using the K2 cellometer of flow cytometry (Figs. 7/8).

Isolation of a Homogeneous Mutant Population of HEK293

Cells which are positive for GFP can be diluted as described in Fig. 2. Alternatively, cell sorting GFP positive cells into single wells of a 96 well plate allows for the isolation of single colonies. Cells will grow to confluency in 2-3 weeks from single colonies. Once cells are grown to confluency in 96 well plates they can be re-suspended using the cell

splitting protocol described above, albeit scaled to smaller volumes, and seeded into larger wells i.e 12 well plates. Once the expanded colonies are grown to confluency they can be assayed for ssODN incorporation into the genome. Testing for the incorporation of the ssODN can be achieved by PCR amplifying the area of interest from isolated DNA of the HEK293 cells (Fwd: 5'-CATCTTACCTTCTGCGTGTTTGCTC-3', Rev: 5'-ATCTGTGGGCATGGTCAAAGGCTC-3'). Incorporation of the ssODN is designed to remove an endogenous restriction site, conversely introducing silent mutations which insert a new restriction site would serve the same purpose.

Isolation of the genomic material contained within the HEK293 cells is achieved through the use of the Thermo Scientific Genomic DNA Purification Kit. This kit utilizes protein degradation through the use of proteinase k and RNA degradation by use of RNase A. Following the protocol included therein allows for the rapid and high yield of genomic DNA.

Amplification of the 1kb fragment encompassing the region of interest is achieved using touchdown PCR in order to limit the amplification of nonspecific regions. Touchdown PCR maximizes the probability of amplifying the region of interest by selecting for the highest base pairing specificity. The PCR protocol begins with an elevated annealing temperature which decreases with each subsequent step. The touchdown PCR protocol is as follows and is programmed into the Eppendorf Thermocycler in A. Clay Clark's Laboratory at the University of Texas at Arlington.

Step1

95°C: 2:30 min

70°C: 00:30 min

72°C: 4:00 min

Step 2

95°C→ 2:30 min

69°C: 00:30 min

72°C: 04:00 min

……*

…….*

….*

..*

95°C 2:30 min

53°C: 00:30

72°C: 14:00

15°C: ∞

*Steps encompassing annealing temperatures of 68°C through 54°C are not shown

Once amplified, the region of interest is incubated with the restriction enzyme corresponding to the restriction site which has been incorporated or removed. The design for the mutation of the dimer interface of caspase-3 utilizes the removal of an endogenous HindII site, therefore amplified regions with incorporation shall remain undigested while

wt amplified regions will digest. Amplified regions which are not positive for incorporation will digest into two .5 kb fragments while a positive fragment will remain at a molecular weight of 1kb. Additionally the ssODN removes the PAM site which is targeted by the Cas9:gRNA ribonucleoprotein complex (Fig. 6). Colonies which exhibit undigested amplified regions should be further tested through DNA sequencing. Once a clonal expansion has been identified as containing a homogeneous population of mutated HEK293 cells the cells should be further expanded into 75 mm^2. Expanded cells should be pelleted though centrifugation and transferred to a cryosolution. Cells can be frozen at a concentration of 1,000,000 cells/mL in a cryosolution composed of 95% DMEM-complete/5% DMSO.

Conclusions

Using a Type II CRISPR-Cas system, optimized for mammalian genomic editing, caspase-3 will be mutated in human embryonic kidney (HEK293) cells to determine the contribution of caspase-3 in cellular pathways and disease models. Studies which aim to assess functional roles of caspase-3 rely on knockdowns and knockouts of endogenous capase-3 while simultaneously expressing exogenous caspase plasmid donors. The levels of exogenously introduced caspases are not governed by epigenetic markers and therefore are expressed without constraint. CRISPR however offers the ability to mutate endogenous genes and therefore maintain endogenous expression levels through endogenous promoter control. The design strategy discussed in this section can be used for the generation of any mutation desired in the human genome. The specific reagents described and used in this chapter can be used to genetically edit any dimer interface mutation at position 266 (chapter 2). The only reagent required to change is that of the

repair template as the endogenous DNA targeting and cleavage machinery can be conserved throughout experiments. The methodology here demonstrates how to produce a homogeneous population of HEK293 cells expressing mutated caspase-3, under the control of endogenous promoters.

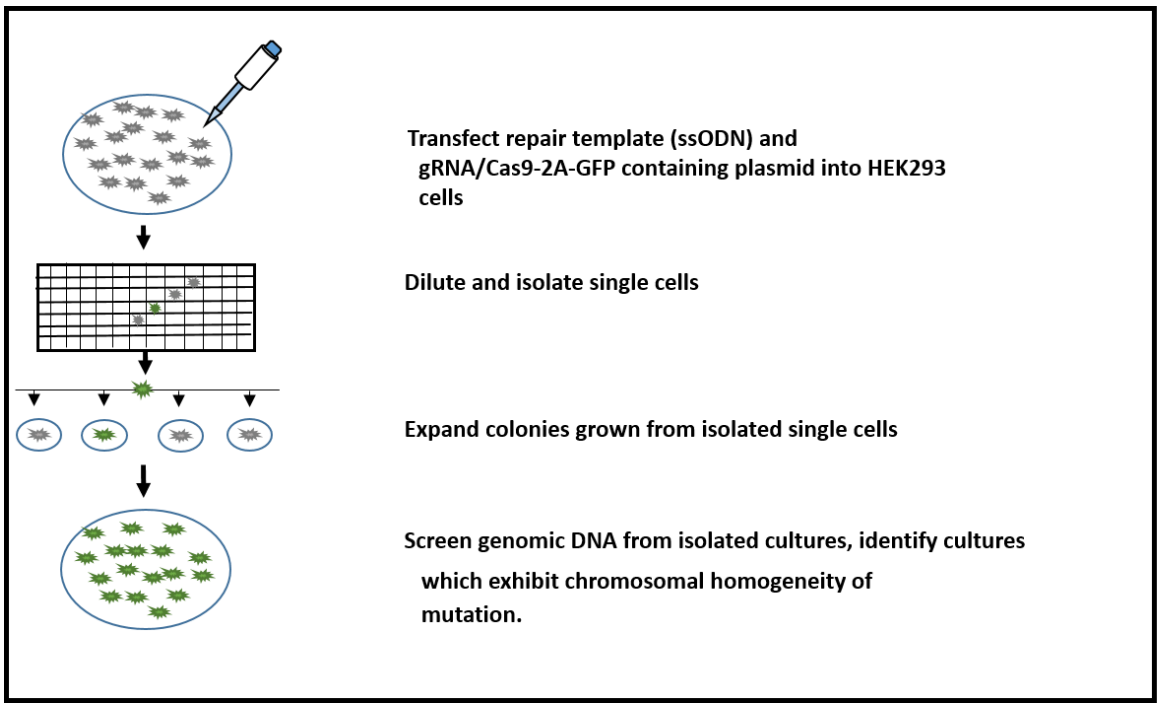


Fig 1. Overview of Cell Culture Strategy for the Generation of Homogenous Populations of V266H Caspase-3 Mutants.

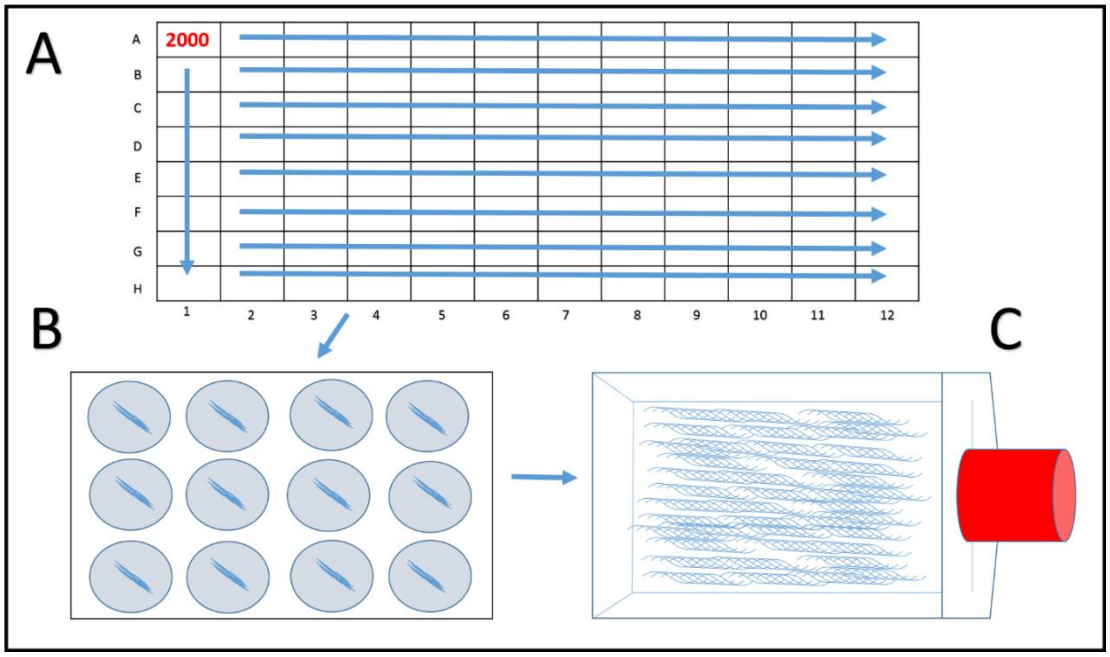


Fig 2. Representation of the Dilution and Cell Expansion of HEK293 Post Transfection.

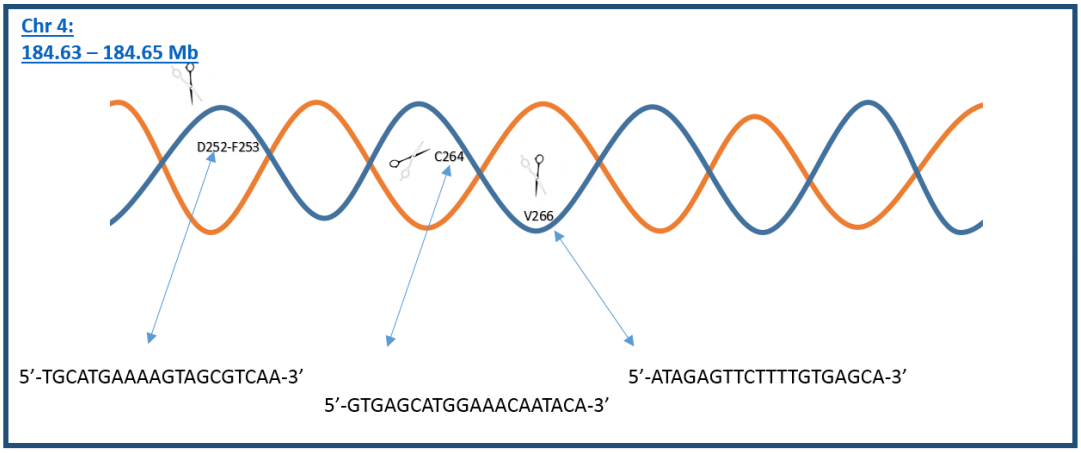


Fig. 3. Three Distinct Regions of the Caspase-3 Gene on Chromosome 4 are Targeted for Double Stranded Breaks by Cas9.

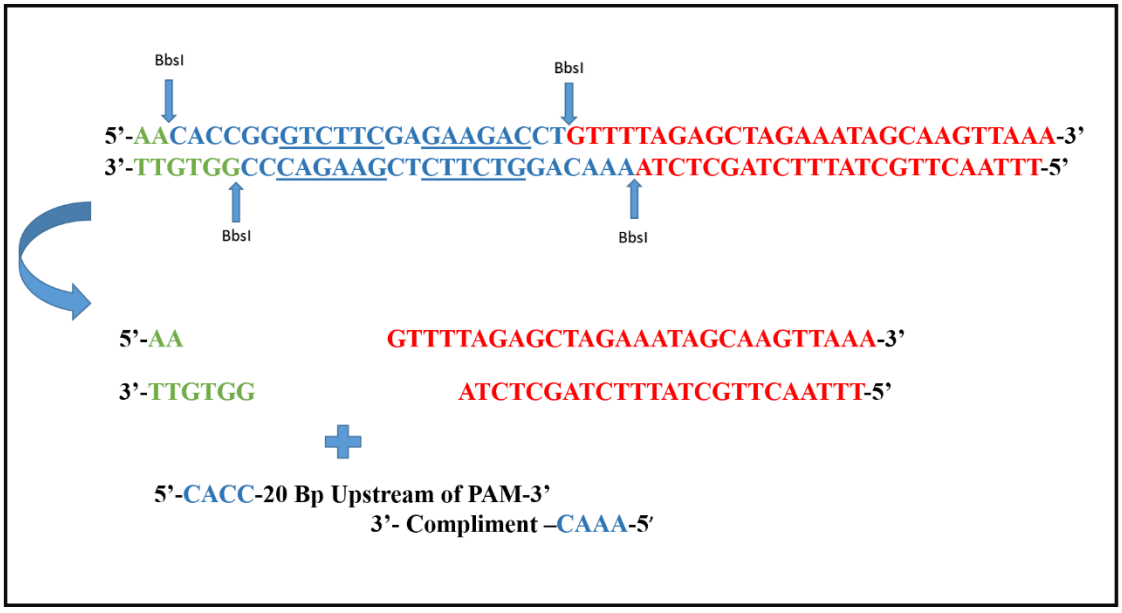


Fig. 4. Cloning Strategy of the pX458 Plasmid.

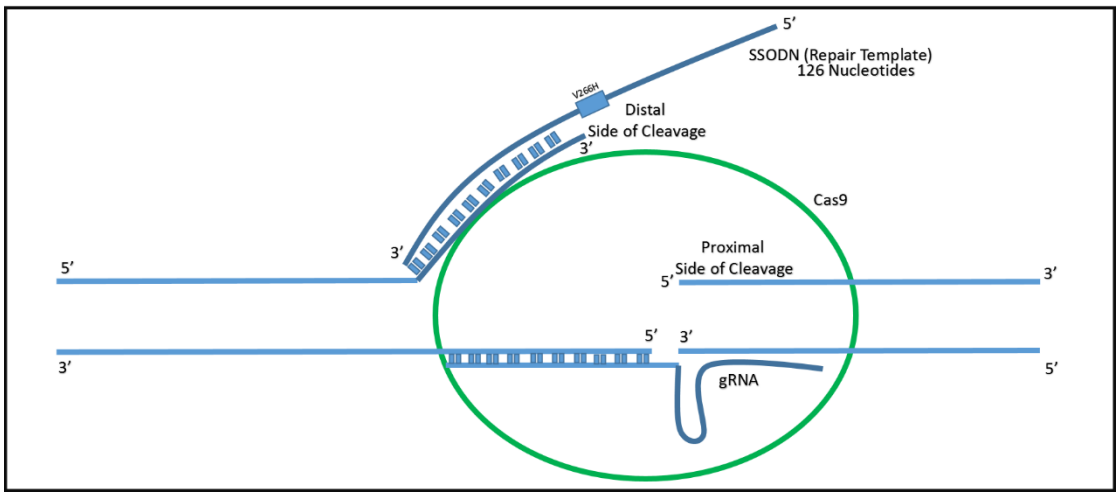


Fig. 5 SSODN Design Strategy

.

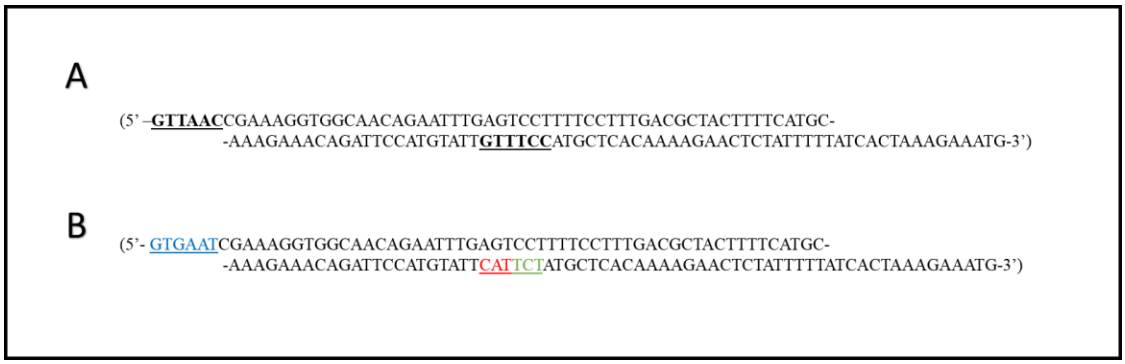


Fig. 6. Human Genomic Sequence of Caspase-3, Region Corresponding to the ssODN.

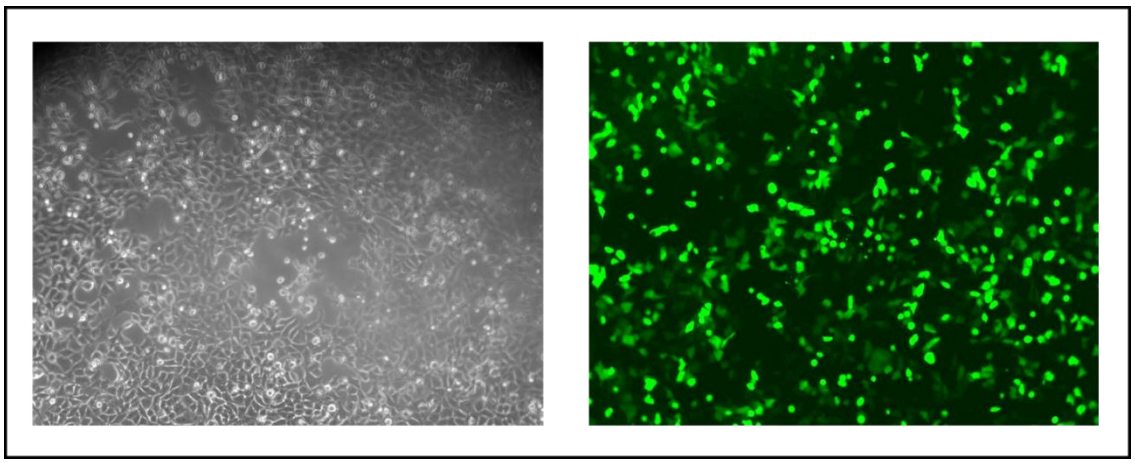


Fig. 7 White and Blue Light Exposure to Cells Transfected with Cas9-EGFP Containing plasmid.

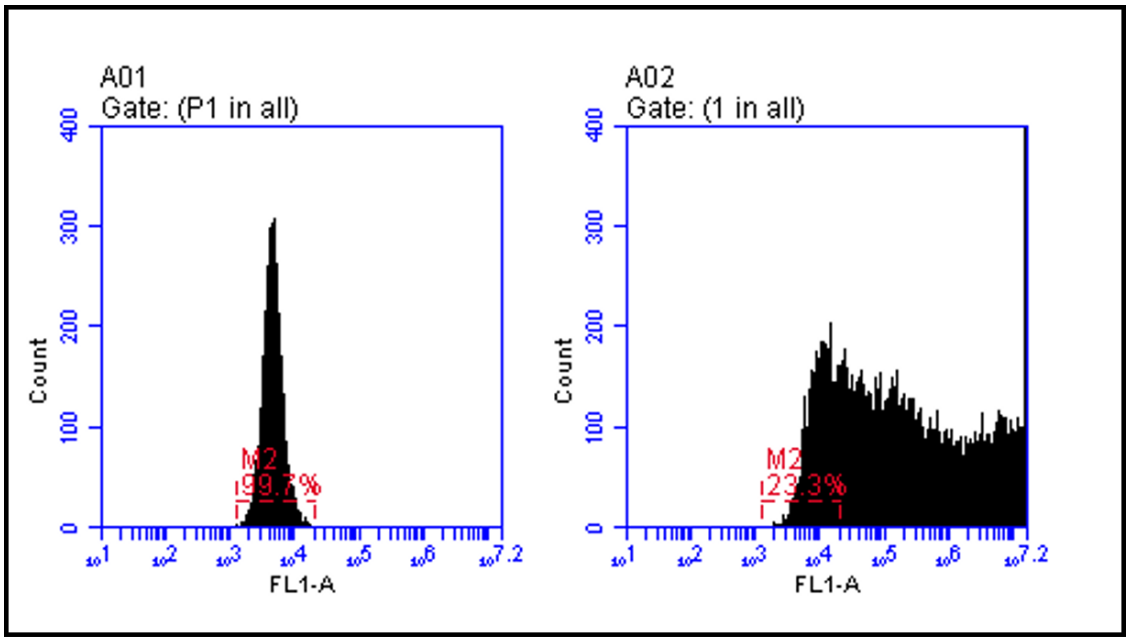


Fig. 8 Flow Cytometry of Control and Treated Cells

.

Table 1. Oligonucletide sequences which direct Cas9 cleavage within the caspase-3 genome.

Oligo Abbreviation	Oligo Sequence	Complimentary Oligo Pair	Directed Toward Caspase-3 Codon
HDC3O-A	5'-CACCGATAGAGTTCTTTTGTGAGCA -3'	HDC3O-B	V266
HDC3O-B	5'-AAACTGCTCACAAAAGAACTCTATC-3'	HDC3O-A	V266
HDC3O-1	5'-CACCGTGCATGAAAAGTAGCGTCAA -3'	HDC3O-2	D252-F253
HDC3O-2	5'AAACTTGACGCTACTTTTCATGCAC-3'	HDC3O-1	D252-F253
HDC3O-3	5'-CACCGTGAGCATGGAAACAATACA -3	HDC3O-4	C264
HDC3O-4	5'-AAACTGTATTGTTTCCATGCTCAC -3	HDC3O-3	C264

REFERENCES

1. Mojica, F. J., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution*, *60*(2), 174-182.
2. Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*(8), 2551-2561.
3. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., & Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology direct*, *1*(1), 1.
4. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., ... & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, *315*(5819), 1709-1712.
5. Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., ... & Van Der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, *321*(5891), 960-964.
6. Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., ... & Terns, M. P. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*, *139*(5), 945-956.

7. Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *science*, *322*(5909), 1843-1845.
8. Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., ... & Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602-607.
9. Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, *109*(39), E2579-E2586.
10. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821.
11. Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., ... & She, Q. (2015). Harnessing Type I and Type III CRISPR-Cas systems for genome editing. *Nucleic acids research*, gkv1044.
12. Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., ... & van der Oost, J. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, *9*(6), 467-477.
13. Karvelis, T., Gasiunas, G., Miksys, A., Barrangou, R., Horvath, P., & Siksnys, V. (2013). crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. *RNA biology*, *10*(5), 841-851.

14. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., ... & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science*, *339*(6121), 819-823.

15. Nuñez, J. K., Kranzusch, P. J., Noeske, J., Wright, A. V., Davies, C. W., & Doudna, J. A. (2014). Cas1–Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. *Nature structural & molecular biology*, *21*(6), 528-534.

16. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., ... & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science*, *339*(6121), 823-826.

17. Li, M., Suzuki, K., Kim, N. Y., Liu, G. H., & Belmonte, J. C. I. (2014). A cut above the rest: targeted genome editing technologies in human pluripotent stem cells. *Journal of Biological Chemistry*, *289*(8), 4594-4599.

18. Cox, D. B. T., Platt, R. J., & Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. *Nature medicine*, *21*(2), 121-131.

19. Ding, Q., Regan, S. N., Xia, Y., Oostrom, L. A., Cowan, C. A., & Musunuru, K. (2013). Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell stem cell*, *12*(4), 393.

20. Gori, J. L., Hsu, P. D., Maeder, M. L., Shen, S., Welstead, G. G., & Bumcrot, D. (2015). Delivery and specificity of CRISPR/Cas9 genome editing technologies for human gene therapy. *Human gene therapy*, *26*(7), 443-451.
21. Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology*, *32*(4), 347-355.
22. Gustafsson, C., Govindarajan, S., & Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in biotechnology*, *22*(7), 346-353.
23. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature biotechnology*.
24. Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., ... & Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, *156*(5), 935-949.
25. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, *8*(11), 2281-2308.

CONCLUSIONS

Caspases are highly modified enzymes which mediate numerous signaling cascades. They are profoundly involved in numerous disease states and cellular developmental pathways. Understanding the procession of caspase activation as well as the regulatory mechanisms utilized by the cell to alter their function is imperative in the ablation of disease and the re-establishment of homeostasis. It is vital to determine where caspases are allosterically modulated, how modulation adjusts caspase function and which modification sites are unique to individual caspase members. These criteria could allow for increased therapeutic specificity, decreased off target effects and the expansion of targeting sites.

Our study provides insight into the intrinsic nature of allosteric mechanisms and the role of water molecules in conformational selection. The analysis of water molecules in the caspase-3 mutational database, combined with structural and dynamic studies, shows that disrupting conserved waters on the protomer is correlated with a shift in the population of states toward an inactive conformation. The data show that the allosteric networks are global, in that there is not a single pathway to describe the allosteric connections. Rather, the regions with higher changes in conserved waters may report on communication between allosteric sites facilitated by the bridging water molecules. Ligands that remove conserved water molecules from the protomer could effectively fine-tune caspase-3 activity by shifting the population toward the high-energy inactive conformation.

Studies of the high-energy inactive conformations of caspases-2 and -3 show that disruption of the catalytic groups can be achieved without large scale disordering of the

active site loops, as small shifts in the catalytic cysteine and histidine are sufficient to inactivate the enzyme. The high-energy inactive state may provide an advantage for regulation of caspase activity in the cell, particularly if the state is stabilized by post-translational modifications, because it provides the cell a means to reversibly control activity by coupling one or more modifications to increased population of the inactive state. Removal of the modification would result in an unstable high-energy state that would rapidly convert to the more stable active conformation.

The database described here provides a tunable allosteric library of caspase-3 variants with nearly four orders of magnitude change in activity, which we suggest will be useful for examining caspase-3 signaling in cells. In current transfection technologies, caspase activity is manipulated in cells through knock-down or knock-out strategies coupled with expression of caspase variants. Levels of endogenous caspase activity are difficult to control, and interpretation of results from overexpressed caspase variants can be problematic. New genome editing techniques, however, combined with our caspase-3 database, should provide tools to fine-tune caspase-3 activity while simultaneously maintaining endogenous protein levels. Caspases are allosteric enzymes and we do not understand how cells set their threshold level of activity.

APPENDIX

APPENDIX: CRYSTALLOGRAPHY

PROTEIN PREPARATION

DAY 1

All caspase-3 variants are stored in an 80% ecoli-Lysogeny Broth/20% glycerol stock solution which is flash frozen using liquid nitrogen, all stocks are stored at -80°C. Upon initiation of protein preparation a metal loop is heat sterilized and used to acquire an aliquot of the frozen "glycerol stock" of a given caspase-3 variant. The aliquot is then T-streaked onto a solid, Agar/Lysogeny Broth plate (1% NaCl, 1% Tryptone, .5% Bacto-Yeast Extract, 1.5% Agar w/v and 50µg/mL Ampicillin). The plate is incubated at 37°C for 16 hours. During the growth period 100 mL stocks of Lysogeny Broth (1% NaCl, 1% Tryptone, .5% Bacto-Yeast Extract w/v) are prepared and sterilized via autoclaving for a minimum sterilization time of 20 minute.

Day 2

Using the plate that was incubated on Day 1, isolate a single colony using a heat sterilized metal loop and inoculate a 100 mL LB culture. Additionally, 50µg/mL of ampicillin is added to each 100 mL culture LB. The 100 mL cultures are then incubated 16-20 hours at 37°C while gyrating. 6- 1 Liter flasks of LB should be prepared and sterilized during this incubation period. The 6- 1 liters can be stored subsequent sterilization for a maximum of 24-30 hours. Lengthy incubation of Lysogeny Broth leads to an unknown contamination.

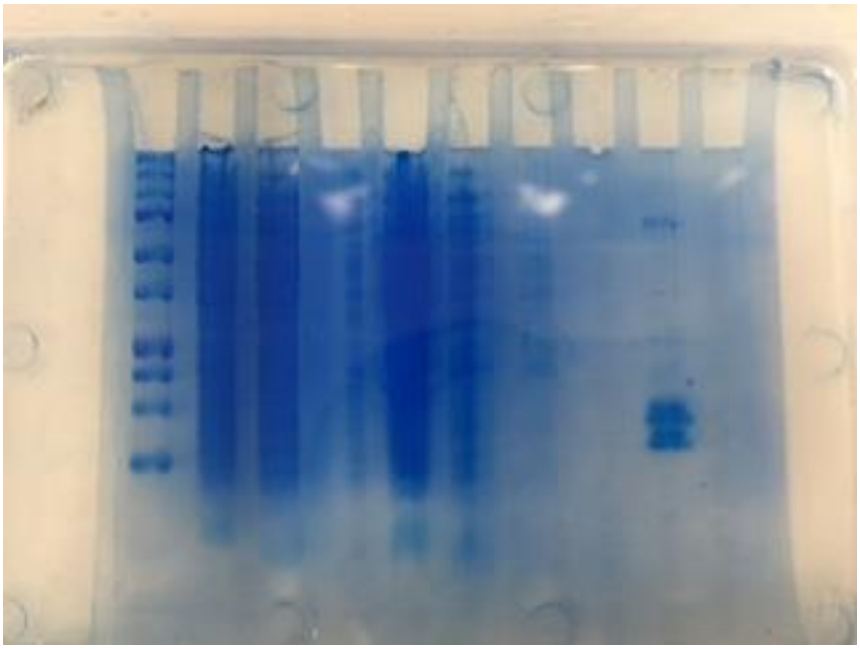
Day 3

Using a spectrophotometer, assay the optical density of the 100 mL culture, using a wavelength of 600 nm. Using the following calculation, you can calculate the volume of the 100 mL culture which will be used to inoculate each of the 6- 1 liter flasks: (OD600)*(XmL)=(1000mL)(.03) or 30/OD600. Note that .03 and OD600 do not have units as they are measurements of absorbance. Inoculate the 6 – 1 liter flasks with X mL's to a starting OD600 of .03 and 50μg/mL Ampicillin. Grow the 6 -1 Liter flasks in a 37°C shaker while gyrating. Allow the e.coli to grow until they reach an OD_{600} of 1.2, around 5.5 doubling times or 2 hours (bacterial doubling time is ≈ 20-30 min). Induce each of the 6 liters with a final concentration of IPTG (Isopropyl β-D-1-thiogalactopyranoside) = .8 mM, reduce the temperature to 25°C and continue growth for 3.5-4 hours. Harvest cells by centrifugation at 5, 000 rpm for 12 minutes. Reconstitute pellet in Lysis Buffer (50 mM NaCl, 50 mM Tris-HCl, 50 mM Imidazole pH 7.9): use 10 mL of lysis buffer per pellet. Allow pellet to stir on ice for at least 4 hours to ensure proper homogenization.

*Prior to day 4, The His chromatography column must be charged and equilibrated. To do so add 15 mL of His affinity resin (conservative estimation of ≈ 120 mg binding capacity). Chelate the resin with 1 column volume of 100 mM EDTA (Ethylenediaminetetraacetic acid) pH 8.0, remove any non-specific contaminants by washing the resin with 1 Liter of water, charge the column by incubating 100 mM Nickel Sulfate (3/4 column volume) for 15 minutes and equilibrate the column with 1 column volume of lysis buffer.

Day 4

French press the cell re-suspension at 1000 psi, note do not repeat multiple rounds of French press on the same re-suspension. The cells contained in the suspension are now lysed, to separate the cell debris from the aqueous fraction, centrifuge the samples at 14,000 rpm for 30 minutes. Retain the supernatant and store on ice. Pour the supernatant onto the His-column and re-suspend, incubate for 15 minutes and collect the flow through. Wash the column with 300 mL of lysis buffer and collect into a separate collection bottle than that of the flow through, label each accordingly. Add 50 mL of Elution buffer (500 mM Imidazole, 50 mM NaCl, 50 mM Tris-HCl pH 7.9) and collect. The majority of your His-tagged protein should be isolated in the 50 mL of elution. Dialyze the 50 mL of elution in 1 liter of 50/50 buffer (50 mM Tris-HCl, 50 mM NaCl, 1mM DTT pH 7.9) for a minimum of 4 hours. Save flow through and wash collection bottle for subsequent gel electrophoresis. The following 12% SDS-PAGE gel is an example of a post his column chromatography gel (Lanes: Ladder, Post Induction, Pre Induction, Empty, Ft, FT bleed over, Wash, Empty, Elution)

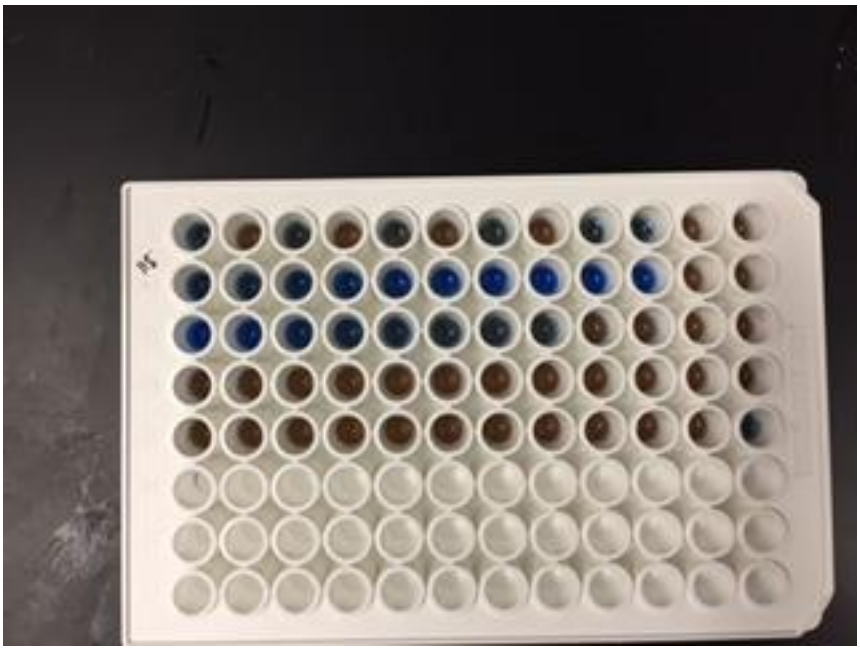


*Prior to day 5, equilibrate the Q-sepharose column using the following: using 30 mL of Q-sepharose resin strip the column of any contaminants bound to the resin by addition of 1 M NaOH. Equilibrate with 1 liter of 50/50 buffer.

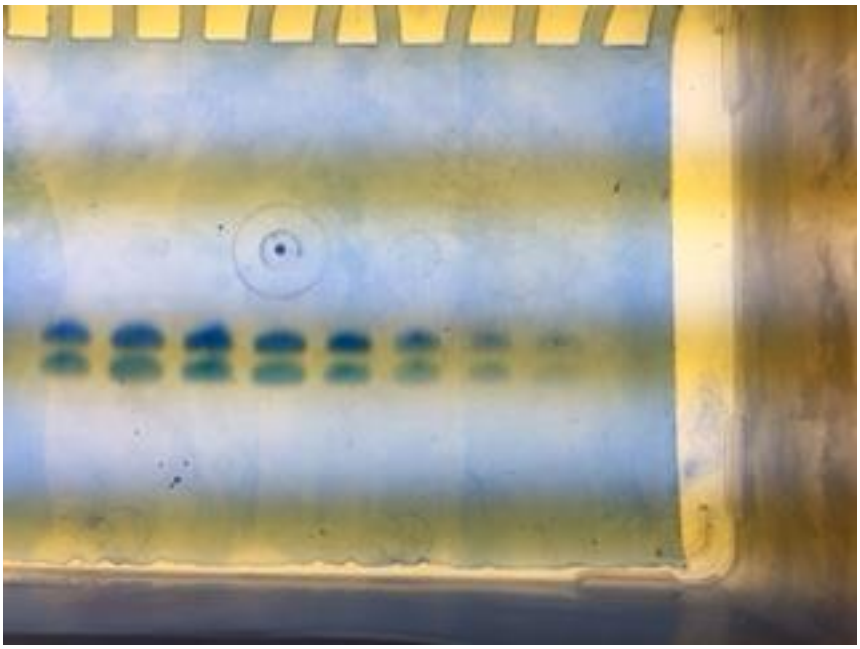
Day 5

Create a salt gradient that will be used to elute the protein from the Q-Sepharose column. The protein will elute as the concentration of salt increases. Begin by filling two 500 mL beakers with 500 mL of 50/50 buffer and 1 mM DTT. In one of the beakers add 300 mM of NaCl to the solution making the final concentration 350 mM NaCl, label this beaker "High NaCl". Connect both beakers via a hollow glass cylindrical tube. Load the dialysis solution from Day 4 onto the Q-Sepharose column and collect the flow through. After the flow through is collected, insert a hollow plastic tube into the "Low NaCl" beaker and

connect it to the top of the Q-sepharose column. Allow the entirety of the 1 liter total volume to flow from the beakers through the column whilst collecting 20 mL fractions. Save these fractions on ice and assay them for the presence of protein by Bradford assay (75 µL of Bradford reagent and 25 µL of sample).



Test positive 30 mL fractions for macromolecules corresponding to both caspase-3 subunits (≈12 and 20 kDa or 32kDa) by comparing them to known standards on an SDS-PAGE gel. The following 12% SDS-PAGE gel indicates cleaved caspasase-3.



Post gel electrophoresis, pool the fractions that are pure and positive for caspase-3. Concentrate these fractions in a 350 mL Amicon Concentrator with a 10 kDa ultrafiltration disc to a volume of ≈ 25 µL. In order to rapidly exchange the 50/50 buffer with 10 mM Tris and 100 mM NaCl pH 8.0, add 20x the volume of the initial concentration and re-concentrate to a volume of ≈ 25 µL again. Transfer the 25 µL to a Millipore desktop concentrator (Maximum Molecular weight cutoff of 10,000 Da) and concentrate to 1 mL. Estimate the concentration using Beers law, A=εlc, where ε=26500 cm^-1m^-1 and the molecular weight is 29500 Da. Aliquot the remaining solution in microcentrifuge tubes and freeze at -20°C.

CRYSTALLIZATION CONDITIONS AND STORAGE

Purified proteins were dialyzed in 10 mM Tris-HCL 100mM NaCl pH 8.0 along with 1mM DTT. The proteins are concentrated to a range of 7-10 mg/mL. Concentrated proteins were incubated with 5:1 stoichiometric equivalents of inhibitor (Ac-DEVD-

CMK), final concentrations of 10 mM DTT and 3 mM Sodium Azide (NaN_3) for 1 hour. Utilization of the hanging drop diffusion method generated crystals in a range of 2-7 days. Briefly, the solution of protein/inhibitor and reservoir solution is mixed in equivalents (2µL/2µL) and pipetted onto the underside a glass cover slip. The cover slip is sealed over a reservoir solution through the use of petroleum jelly. The reservoir solution contains varying volumes of PEG (Poly ethylene glycol) 6000 (w/v) and dH_20 while maintaining fixed volumes of 1M DTT, 100 mM NaN_3 and 100 mM sodium citrate (pH 5.0) at a total volume of 500 µL (Table 1).

Protein/Inhibitor and Plate Preparation

Initially, all reagents should be kept on ice and allowed to thaw for a period of around 30 minutes. It is important to keep the protein on ice to limit denaturation of the concentrated stock of protein. It is not necessary to store NaN_3, PEG, Na Citrate nor H_20 at -20° C and therefore unnecessary to store these on ice during the thawing process. Once the protein, inhibitor and DTT have thawed prepare the protein and inhibitor mixture. Each 24 well crystallization plate requires 2µL/well or a total volume of 48µL. It's advisable to prepare additional sample, the following example will have a total reaction volume of 60uL.

To determine the required volume of a stock inhibitor solution (100 mM) needed to inhibit mature caspase-3 use the following calculation:

Protein/Inhibitor reaction: 5X concentration of protein

((A280)/26500)(1,000) = [Protein mM]

[Inhibitor mM] = (5)([Protein mM])

(100mM Stock of Inhibitor)(X µL) = ([Inhbitor mM])(60 µL)

Upon calculation of the volume of inhibitor required use the following reaction volumes to prepare an inhibition reaction. Adjusting the final volume in the above and below reactions will alter the volumes of both DTT and NaN_3 to reach a final concentration of 10 mM and 3mM respectively.

1M DTT	0.6 µL
100 mM NaN_3	1.8 µL
Ac-DEVD-CMK	X µL
Protein	57.6-X µL

Total	60 µL

The reaction should incubate for 1 hour stored in the dark. While the reaction incubates prepare the crystallography plates as directed below:

Petroleum jelly should be applied to each lip of the crystallography plates. Application of the jelly should be done prior to preparing the reservoir solution in order to ensure the wells are not contaminated. Petroleum jelly may be heated to a consistency of 75-100% aqueous jelly (10 min). A pipe cleaner or loop with a circumference equal to that of the well lip allows for the efficient delivery of aqueous petroleum jelly to the surface of the well. Immerse the loop with jelly and gently rest it on the well for 1-2 seconds. The petroleum jelly should solidify within 1 minute. The plates may then be prepared by adding reservoir solution to each well (Fig 1 and2). Once the jelly has solidified on the plate add the appropriate volumes of reagents listed in Table 1to the desired wells. Thoroughly mix the wells by gently shaking the plates and tilting each at a 45° angle. Store the plates in the dark until the hour incubation of protein/inhibitor has concluded.

Pipette 2µL of protein/inhibitor onto the center of a glass cover slip, pipette 2µL of reservoir solution directly into the droplet of protein/inhibitor and place the cover slip onto the reservoir. Gently press the cover slip to insure a seal. Once all of the wells are sealed, place the plate into 18°C growth chamber and assay for crystal growth in 2-3 days using a microscope.

Upon crystal formation, crystals were transferred into a cryogenic solution consisting of 20% 2-Methyl-2,4-Pentanediol (MPD) and 80% reservoir solution. Mix 400 µL of the reservoir solution (from the well in which you wish to pick crystals) with 100µL of MPD in a 1.5mL microcentrifuge tube and vortex for a few seconds. Pipette the solution onto a clean plastic petri dish and place the cover slide next to the solution. Remove a .1-.2 mm crystallography loop connected to a goniometer base from the cryovial. Place the cryovial into liquid nitrogen and mount the goniometer base to a magnetic wand. To capture a crystal of interest, adjust the loop so that it is directly underneath a crystal and gently lift up slowly, the crystal should slide into the loop. As soon as the crystal is snared within the loop move the loop to the solution of cryo-protectant and submerge the loop, quickly raise it out of the solution and immediately attach the goniometer base to the cryovial already submerged in liquid nitrogen. The vial should be removed from the liquid nitrogen and placed on a labelled cane. Canes are stored in dewars containing liquid nitrogen (-196°C). Crystals are sent to SER-CAT synchrotron beamline (Advance Photon Source, Argonne National Laboratory, Argonne, IL).

SCALING DATA

The following section may be used to scale the "raw" data generated through x-ray diffraction of protein crystal structures. Using HKL2000 one can scale or process data into .log and .scal files.

Open HKL2000 and select Mar300, click on the tab labelled "DATA" at the top of the user interface. Within the DATA tab and underneath "Load/Create New Sets" select an input as well as an output pathway. The input pathway should contain the data received from SERCAT while the output directory represents the destination of the files you will generate using HKL2000. Once these two directories are set click on the box labeled "Load Data Sets". HKL2000 will now load the raw data from your input directory. You will be prompted to choose which data you would like to process, choose "1-180", this is indicative of 180 frames.

Click on "Site Configuration" (above the "DATA" tab and input the coordinates of the beam center. If you do not know the beam center coordinates, open the file labeled "den.in" or "beamcenter_22BM.dat" for the x and y beam coordinates.

Next, select the "INDEX" tab located at the top of the user interface and select "Peak Search" on the right side of the window. This will open up an additional window in which a frame of diffractions appears and diffraction spots that are circled. Click on "Peak Sear" on the left side of the screen. Add around six frames to the indexing by middle clicking the "Frame" button six times, note that the left click can be used to revert back to the previous frame. Make sure that the amount of peaks in each frame is no less than 60. HKL2000 uses the indexing of a small amount of frames to determine which point group will be assigned and the Bravais lattice. Once the correct amount of frames

and peaks are satisfactory, click “OK”. Go back to the previous user interface and click “Index” (just below the “Peak Search” clicked earlier). A new window will now open and which allows you to select the Bravais lattice for your data. The list of point groups that are displayed are accompanied by a list of numbers that are colored Red, Blue or Green. The colors indicate how well of a fit there is between the data and the point group, red being indicative of a poor fit while green represents a good fit. The first number listed to the right of the point group is a percentage, this percentage specifies how much the point group will have to be distorted in order to fit the data (the lower the better). Unless the data is a 0% distortion two rows of numbers will be listed to the right of the distortion percentage. The first row are the numbers associated with the data while the second are the numbers associated with the point group model. The first three numbers signify a,b and c unit cell lengths while the last three signify alpha, beta and gamma angles between the axis. Point groups are listed in order of symmetry (top to bottom) a good model should have a relatively low distortion percentage and high symmetry. Select the point group which matches this criteria and click “Apply and Close”

Once you have selected your point group the original user interface will display the space group you have chosen along with Chi square values, Chi square values below 2.0 are desirable. Click the “Refine” button (just below “Index”). When the Refinement displays READY click on “Fit All” on the left side of the user interface and deselect the box labeled “Mosaicity” then click the “Refine” button once more. The average Mosaicity profile may be checked using the “Check Mosaicity” button. Select the “Mosaicity” checkbox which was deselected previously and click “Refine” once more. You have now successfully indexed the peaks of the data and they can now be integrated.

In order to integrate the data you must first set the parameters of the peaks by designating which portions of the diffraction spots are background and which is the diffraction itself. To do so click on “Zoom Wind” in the pop up box which displays the frames. A new window will open which is blank, in order to fill the window middle click on the window which contains the frame. A magnified image of a portion of the frame can now be visualized in the newly created zoom window. In the zoom window click “Int. Box”. You will now see boxes with circles inside of them, it is preferable to have the circles encompass the diffraction spot and the boxes to be small enough that they do not encroach on the neighboring peak. In the original user interface “Box Size” and “Spot Size” can be used to change the respective size of each. It should be noted that the sizes will not change until you click “Refinement”. They will automatically update in the zoom window upon completion of refinement. When you are satisfied with both box size and spot size you may integrate the data by clicking “Integrate” in the original user interface. Integration may take around thirty minutes to an hour.

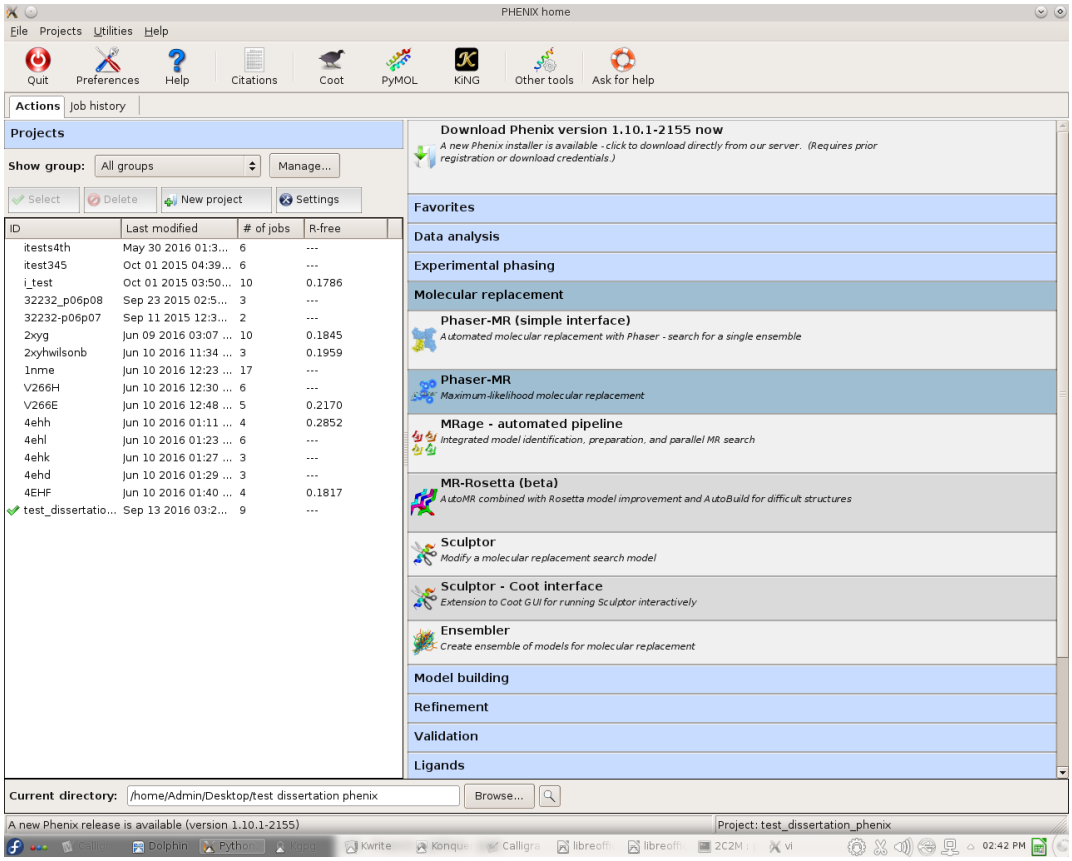
Once the integration is complete click on the “SCALE” tab. Within this tab click on “Scale Sets” at the bottom of the window. The scaling will produce a scale.log file associated with your data set in the folder that you designated as output initially. Open the scale.log file and scroll down to the final table presented in this file. This table indicates the intensity and error associated with different resolution shells. In order to determine which resolution shell will give us the highest quality data we will use the I/sigma or I/error rule of 3. The highest resolution shell with an I/sigma value of 3 along with non-zero R-values will be used to further process the data. Write down the shell which satisfies the rule of three.

Next, scroll to the table above the previous table (second from bottom) in order to determine the completeness of the high resolution shell previously identified. In this table the total completeness of your shell should be higher than 90%. The next table up indicates the number of reflections associated with each shell as well as the total number of reflections. Above the reflections table is the redundancy table which will indicate what level of redundancy is associated with each shell or the average amount of times a reflection is detected in each shell. A high redundancy gives you more confidence in the average intensity. Now that you have reviewed each of the tables of interest in the scale.log file and are confident in the quality of your data you can go back to the HKL2000 user interface and input "50" as your low resolution limit and "high resolution shell limit as determined through table analysis" as you high resolution limit. There is also a box which allows for the inclusion of previously rejected reflections, click this box, followed by clicking "Scale". A newly generated .sca file has been generated in your output directory, view this file and recheck the tables discussed previously.

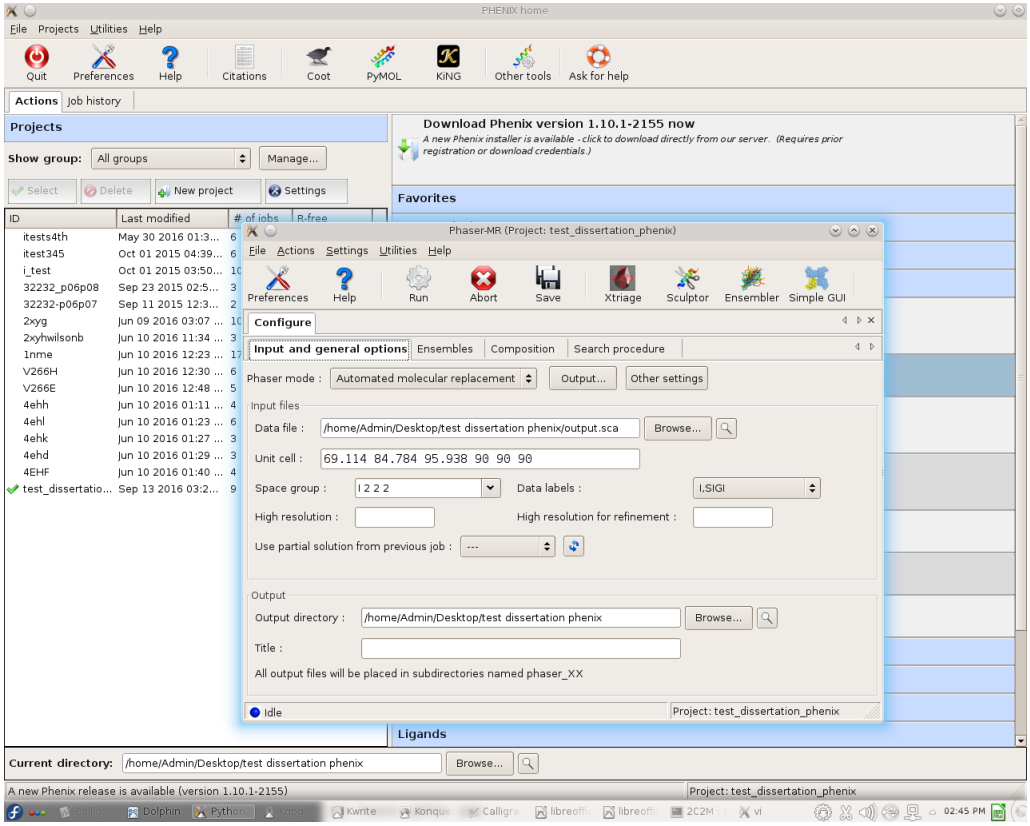
PHASING AND REFINEMENT: PHENIX

Caspase-3 structures crystallize in one of two space groups: orthorhombic I222 and monoclinic C2, where phasing was accomplished by either the 2J30 or 1NMS structures respectively. Open the program "Phenix" and select the appropriate output folder. This output folder should be the same folder that contains your .sca and .log files. On the right hand side there is a drop menu, click on the tab labeled "Molecular Replacement" to expand the menu contents. Select the tab labeled "Phenix-MR", a new

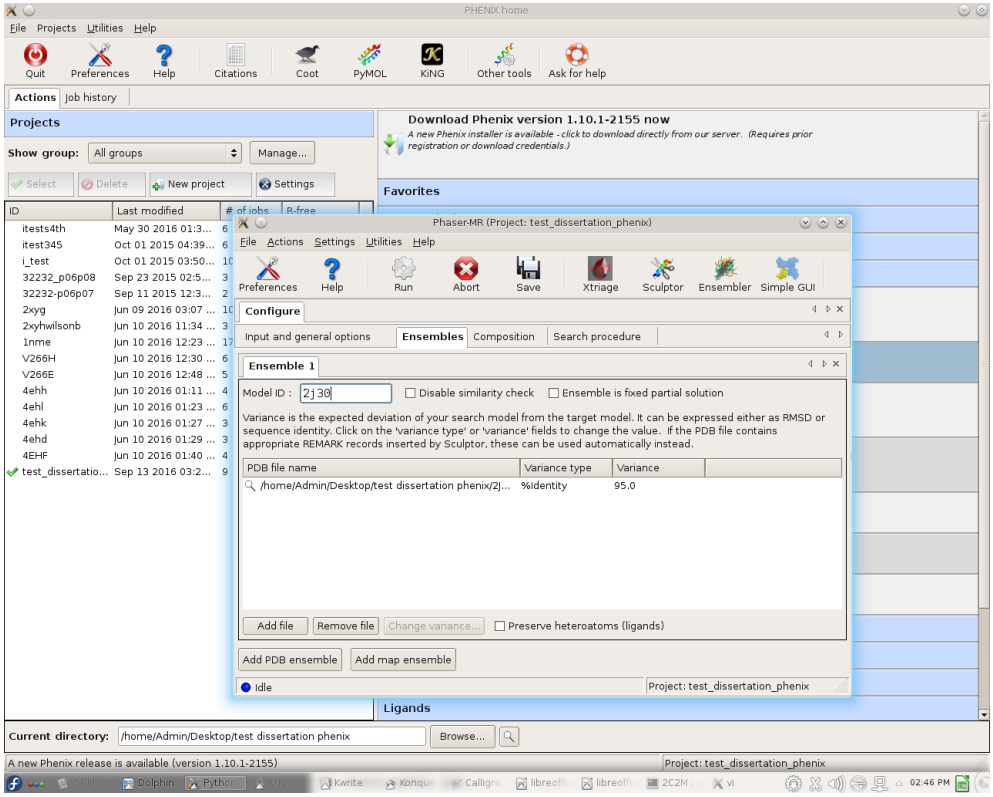
window will appear which asks for the input of a data file.



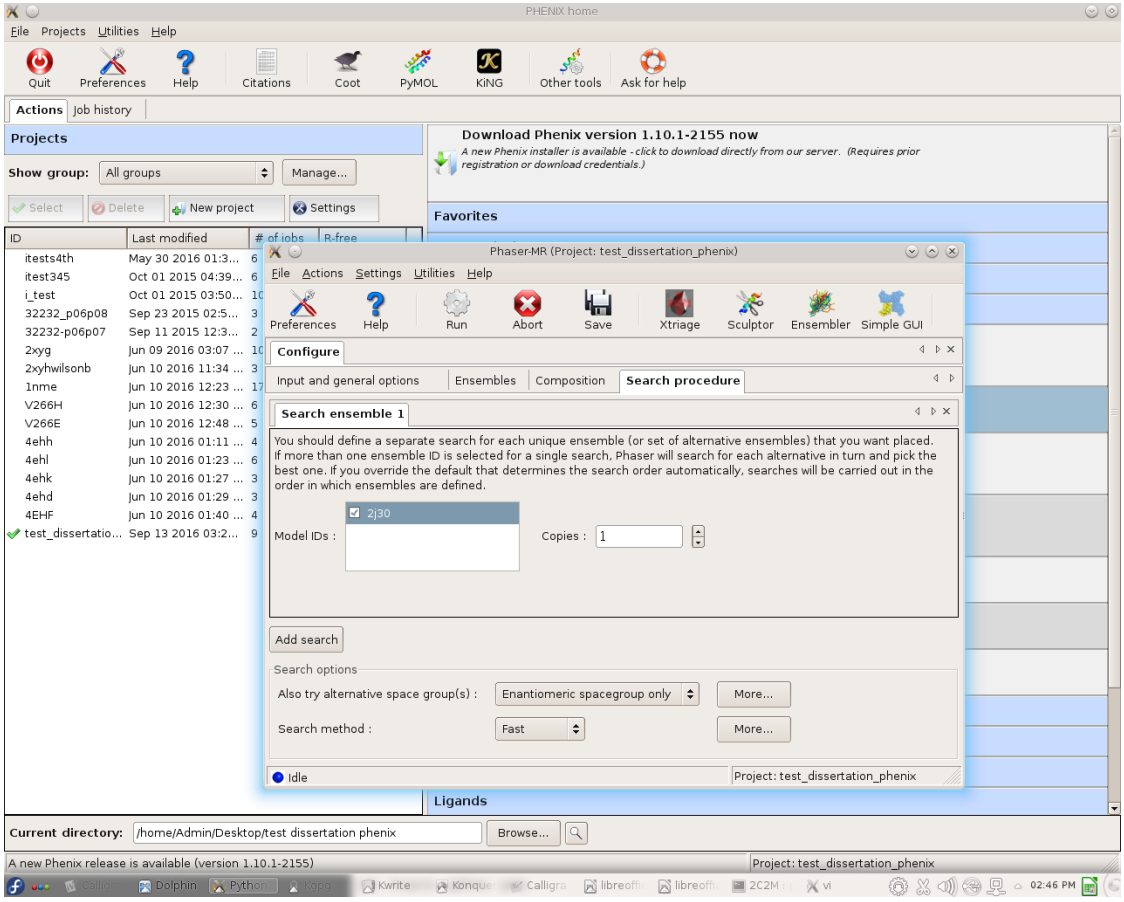
The data file is the .sca which was generated during the scaling procedure above. Once you have uploaded the ,sca file select “Ensembles” in the newly opened window.



Upload the replacement model, either 2J30 or 1NMS, and title the model ID appropriately. At the bottom of the same window change the variance to 95%.

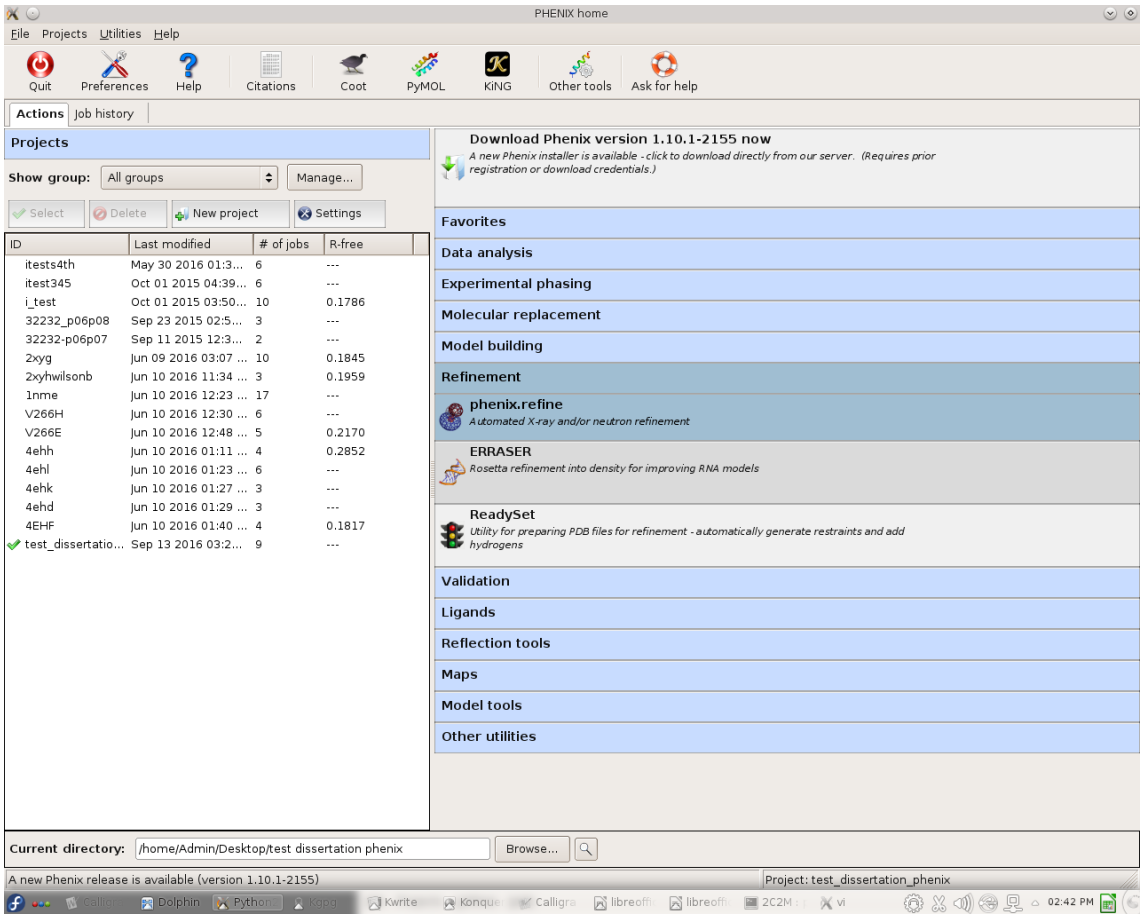


At the top of the window is a tab labeled “Search Procedure” click this tab to open a new window and check the box next to the model ID entered previously.

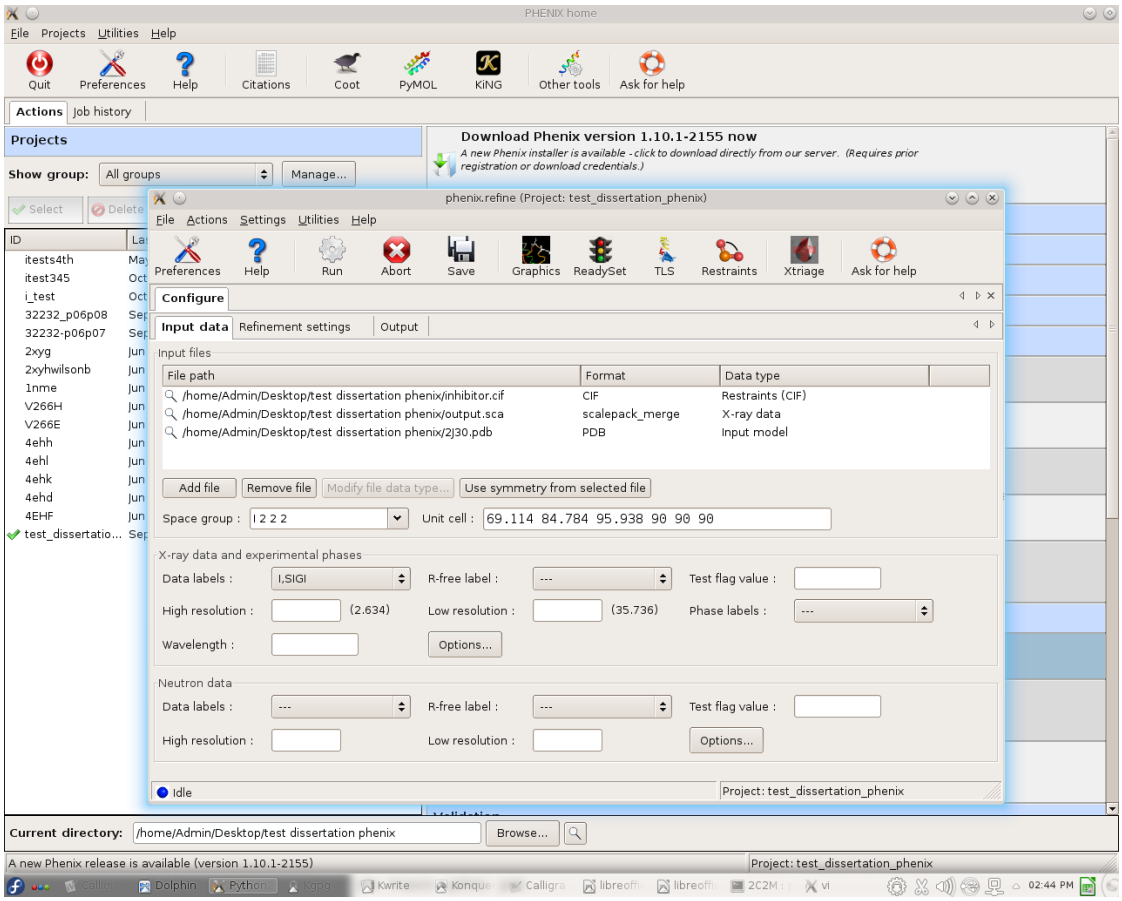


Then select “Run” followed by “Run Locally” at the top of the window to begin the phaser-mr. Three new files will be added to the output directory: a phase map (.mtz), an output model (.pdb) and a processed input file (data.mtz). Once the run is complete you may open the files in Coot automatically by selecting view in Coot underneath the summary of the phaser-mr. In Coot it is important to make sure the model fits the map, additionally if the data is derived from a mutant form of caspase the model can be changed by selecting the residue of interest and mutating it to the correct amino acid, save any changes made to the .pdb file.

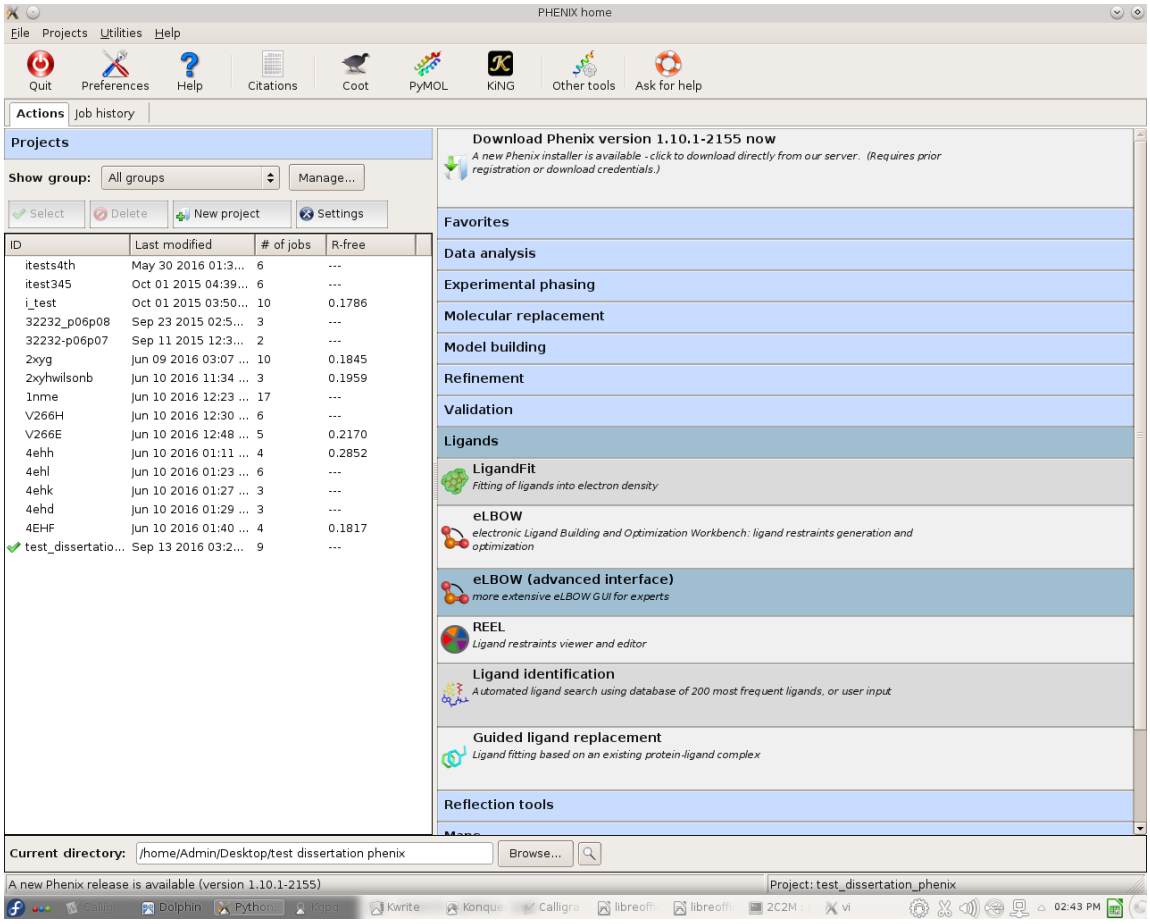
After checking the fit of the model begin the first refinement. Open phenix and select the “Refinement’ drop down menu on the right. In the expanded menu is a tab labeled “pheinx.refine” select this tab.



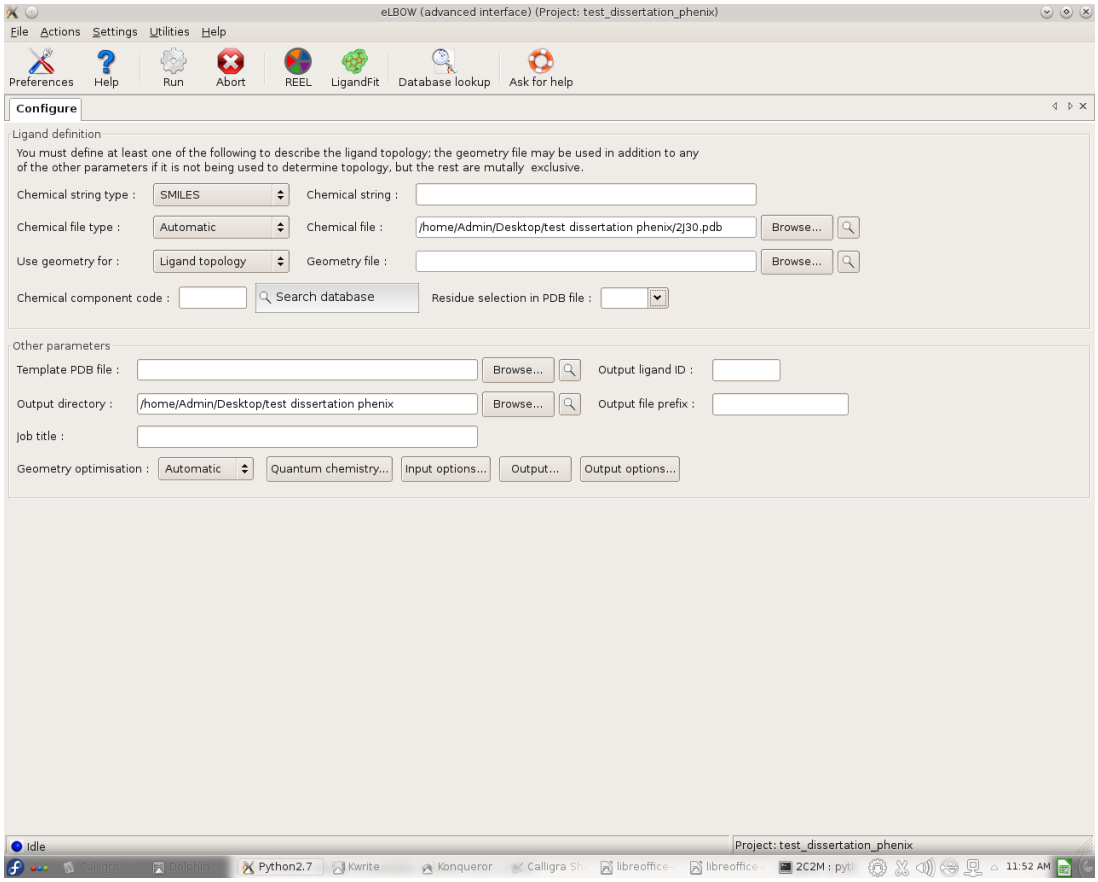
A new window will open, upload the input model (most recent .pdb), x-ray data (data.mtz) and the restraints file (CIF). The CIF file is used to recognize the inhibitor. Ensure that the check box “update waters” remains unchecked.



At the top of the window select “Run” and “Run Locally”. If an error message indicates that the model possesses unknown residue names the problem can be corrected by generating a new restraints file. To generate a new restraints file return to the phenix home screen and select the drop down menu on the right labeled “Ligands”, in the expanded menu select “eLBOW (advanced interface)”.



Upload the .pdb file into the chemical string input box and select the “residue selection in PDB file” menu.



Indicate the three letter code of interest. At the top of the window select “Run” and “Run Locally”. The newly made CIF file can be used to refine the structure using the above procedure. All screen shots and material used in this subsection are adapted from the PHENIX program (1).

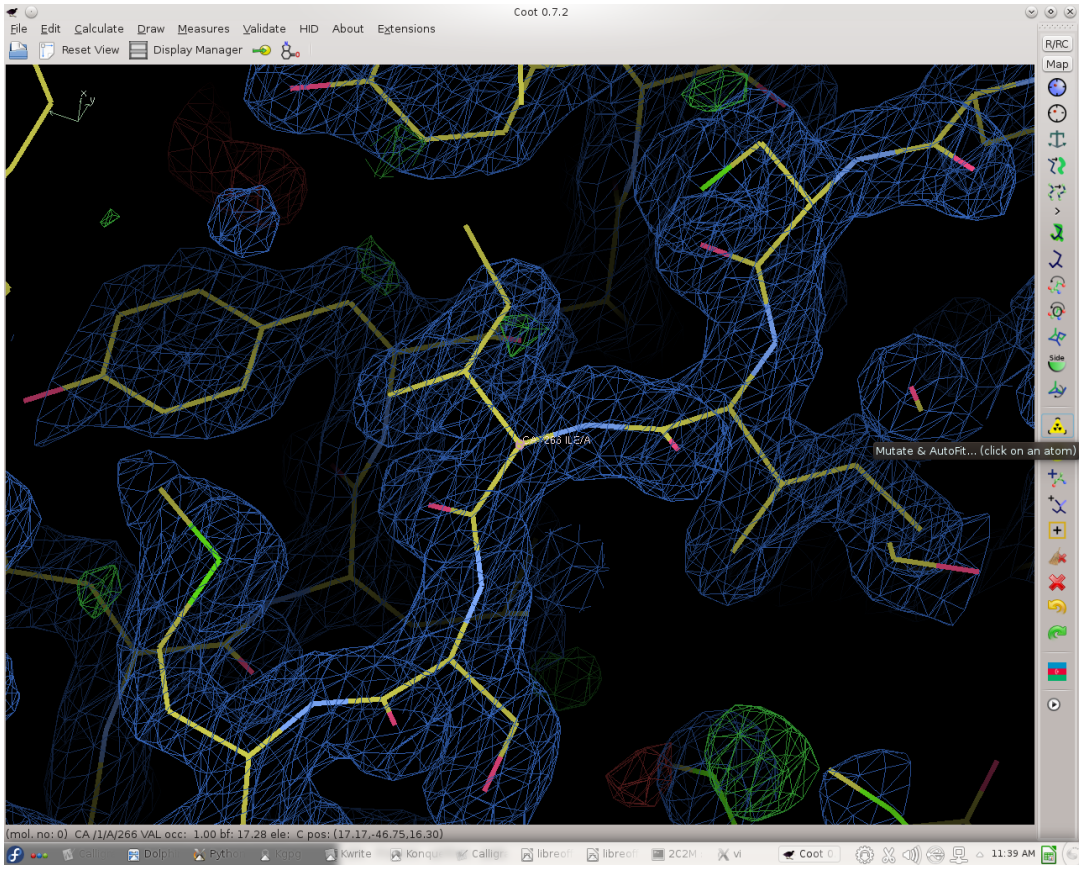
COOT

The following section describes common techniques used in the creation of a protein model using the program Coot. It is important to thoroughly investigate the entirety of the structure after each refinement. Using the following tools one is able to make changes throughout the model.

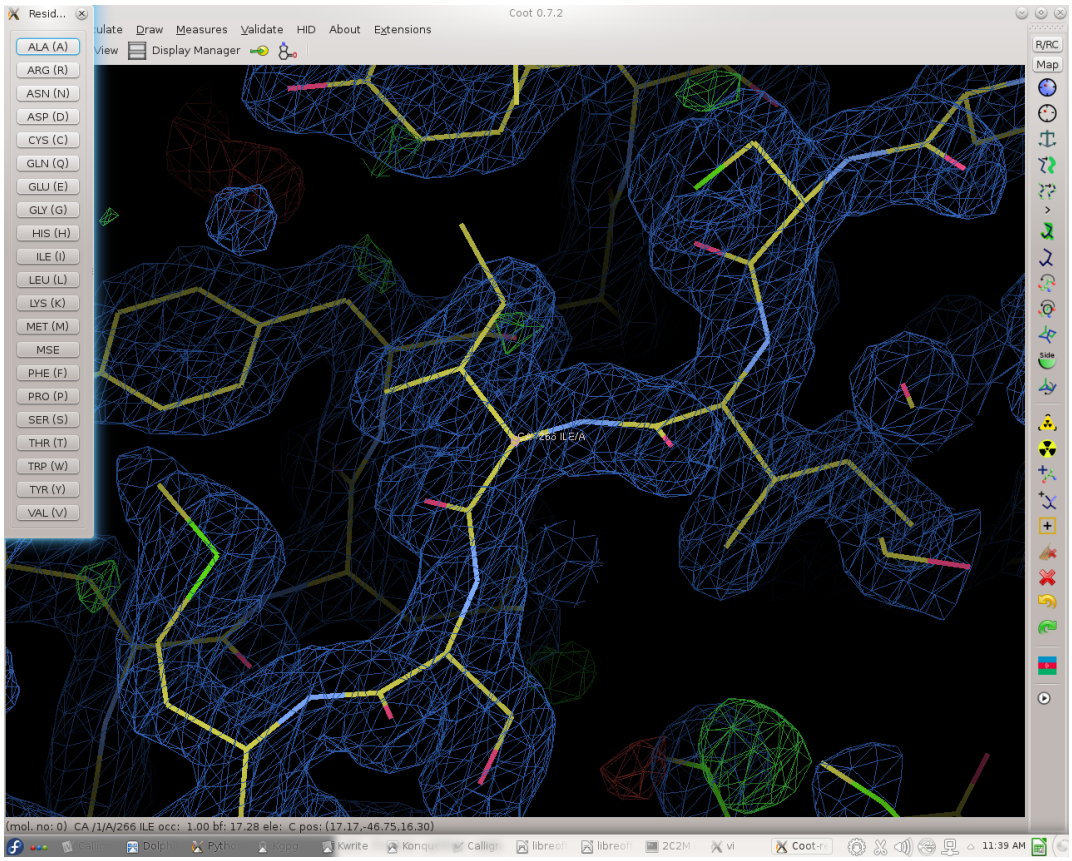
Open the program “Coot” and go to “File”, then click “Open Coordinates” and select the .pdb file form the previous section. Then go to “File”, “Auto Open MTZ” and select the .mtz file associated with your molecular replacement or most recent refinement. The model (.pdb) and the electron density map(.mtz) are now open in coot. Using the wheel, center of left and right mouse buttons, on the mouse set the map RMSD to 3.0 Å. When fitting new molecules use an RMSD of 1.5 Å. The following commands will aid in the generation of the most complete model that fits the electron density observed.

Mutating Residues

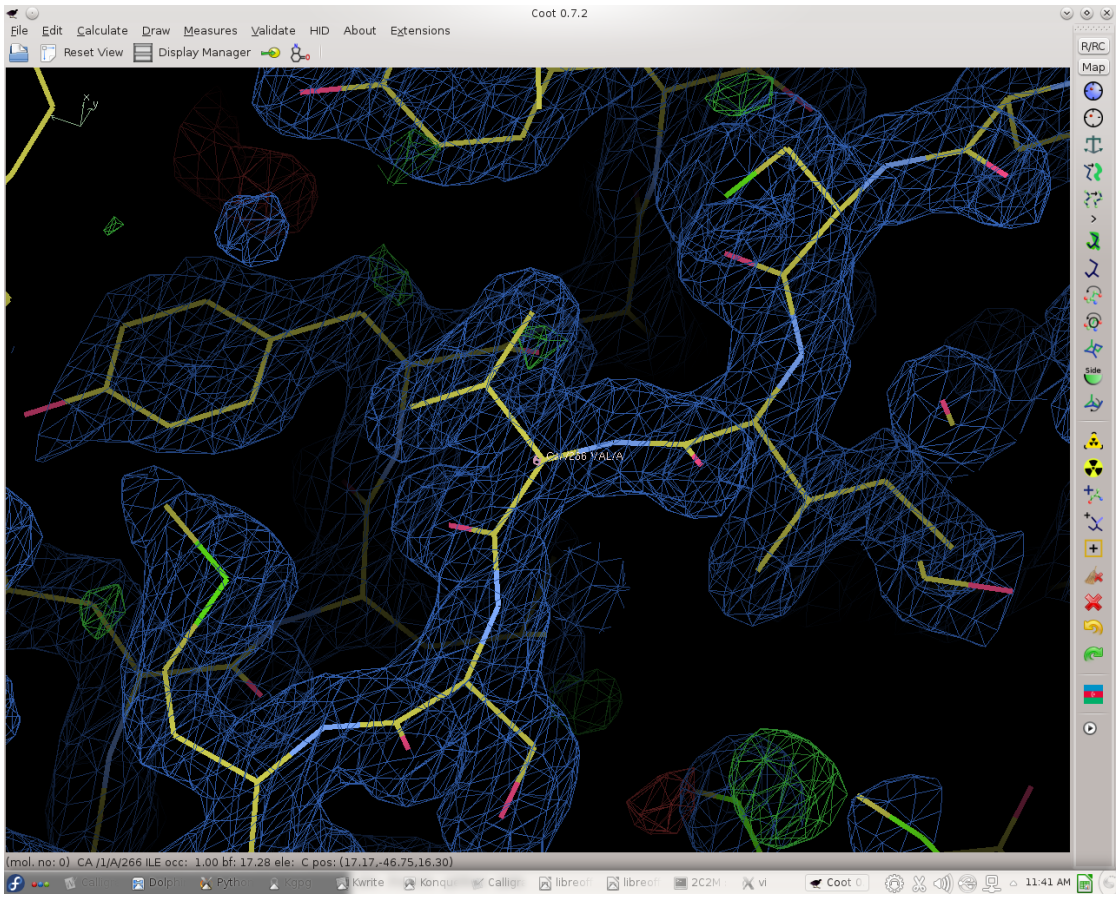
To mutate a residue in a structure select “Mutate and Auto Fit” on the right side user interface.



Select the alpha carbon of the amino aid that is to be mutated. A new window will open on the left side of the screen which lists the amino acid residues, select the amino acid which corresponds to the correct mutation

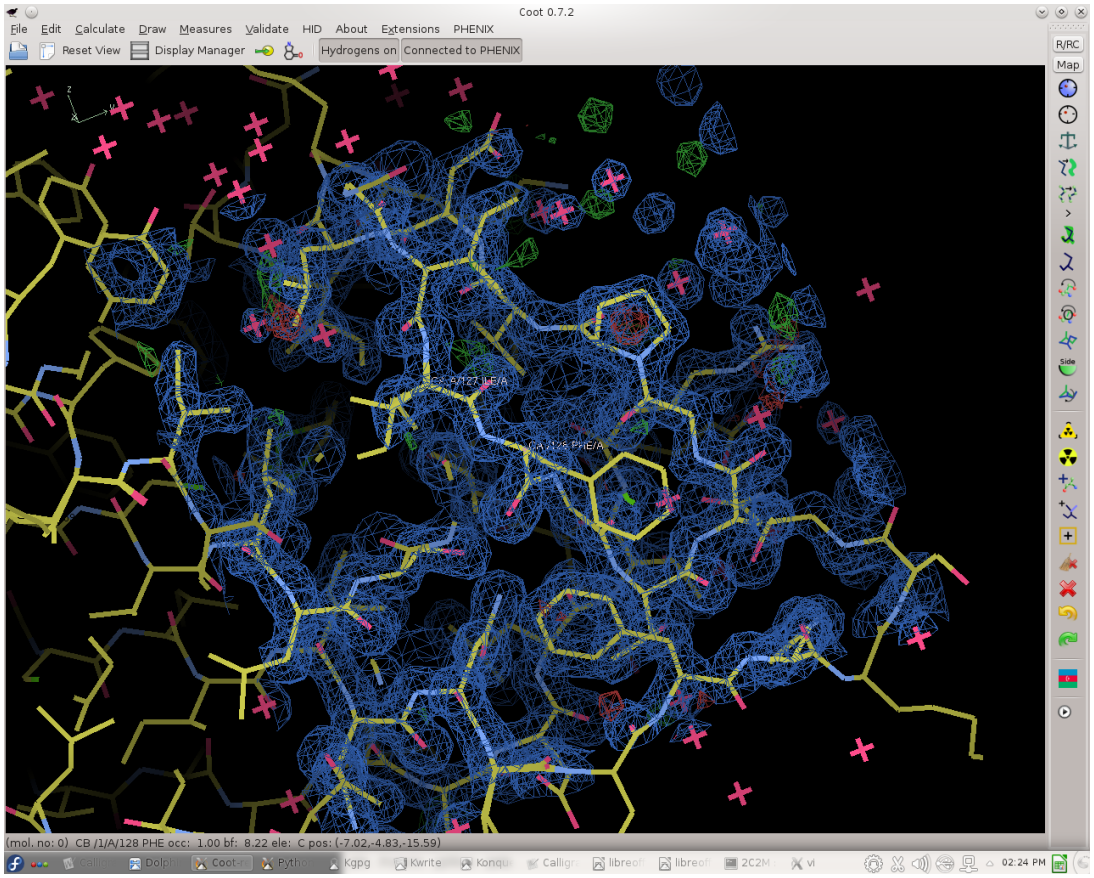


Once the amino acid is selected Coot replaces the pre-existing residue and auto fits the new residue in to the density.

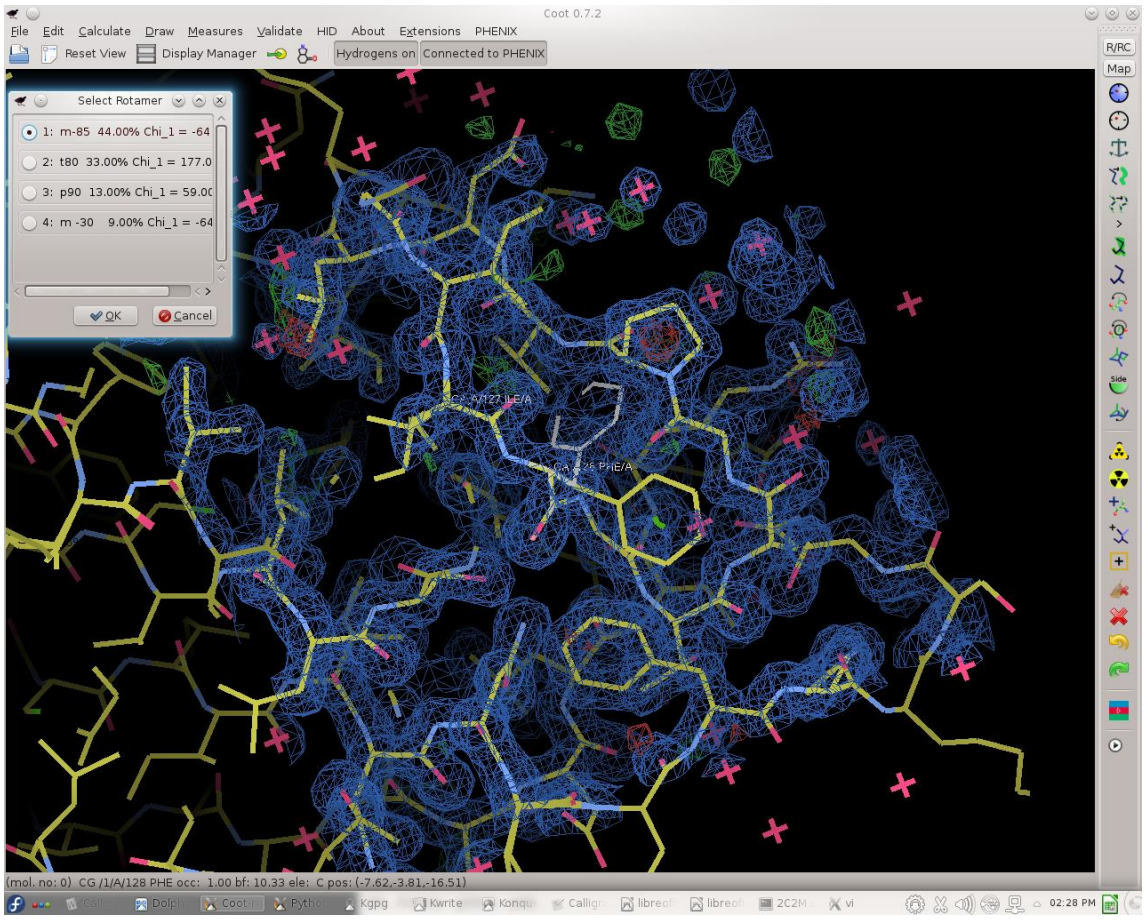


Manually Fitting a Residue into Density

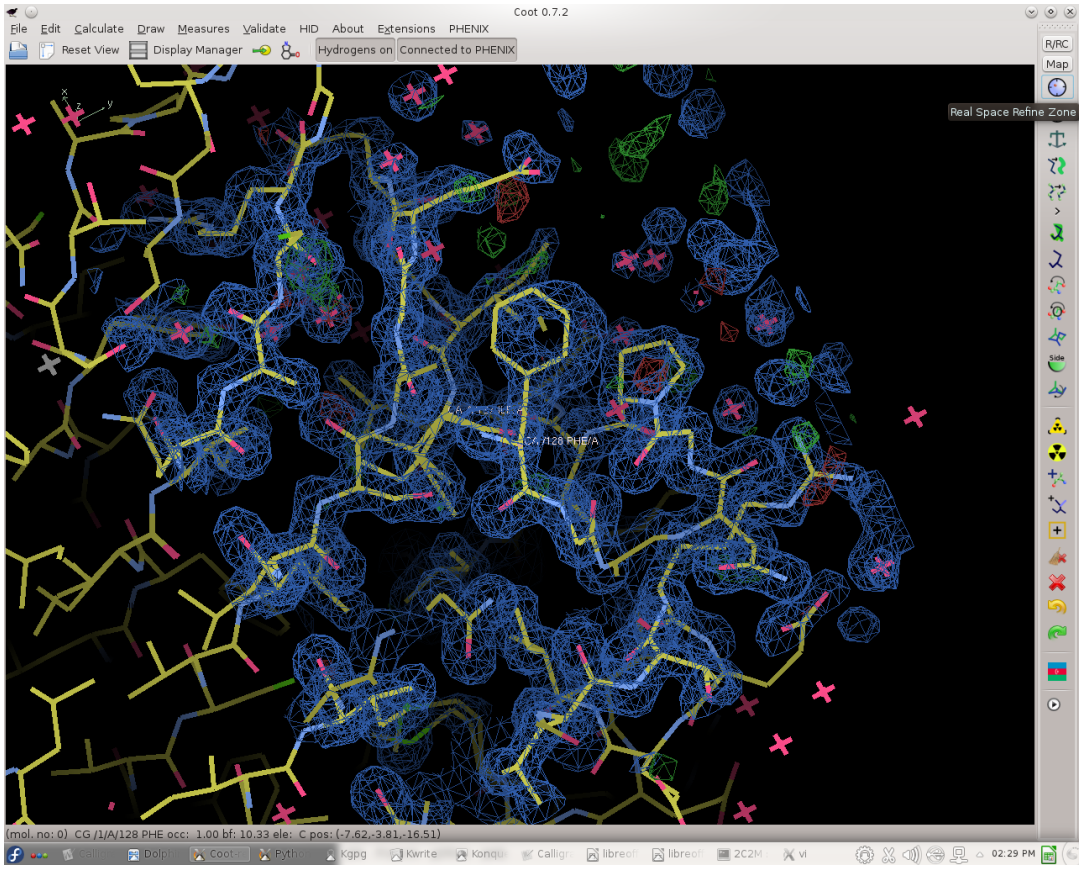
To move a side chain into the correct density select “Rotamers….(Click on Atom)” in the user interface on the right side of the screen and click on the carbon alpha of the amino acid that is to be altered.



A new window appears on the left side of the screen with varying rotamer choices. Select the rotamer which most closely fits into the electron density to which the side chain is being moved. A light-white colored side chain will appear upon selection of each rotamer choice. Click “OK” once satisfied with the rotamer.

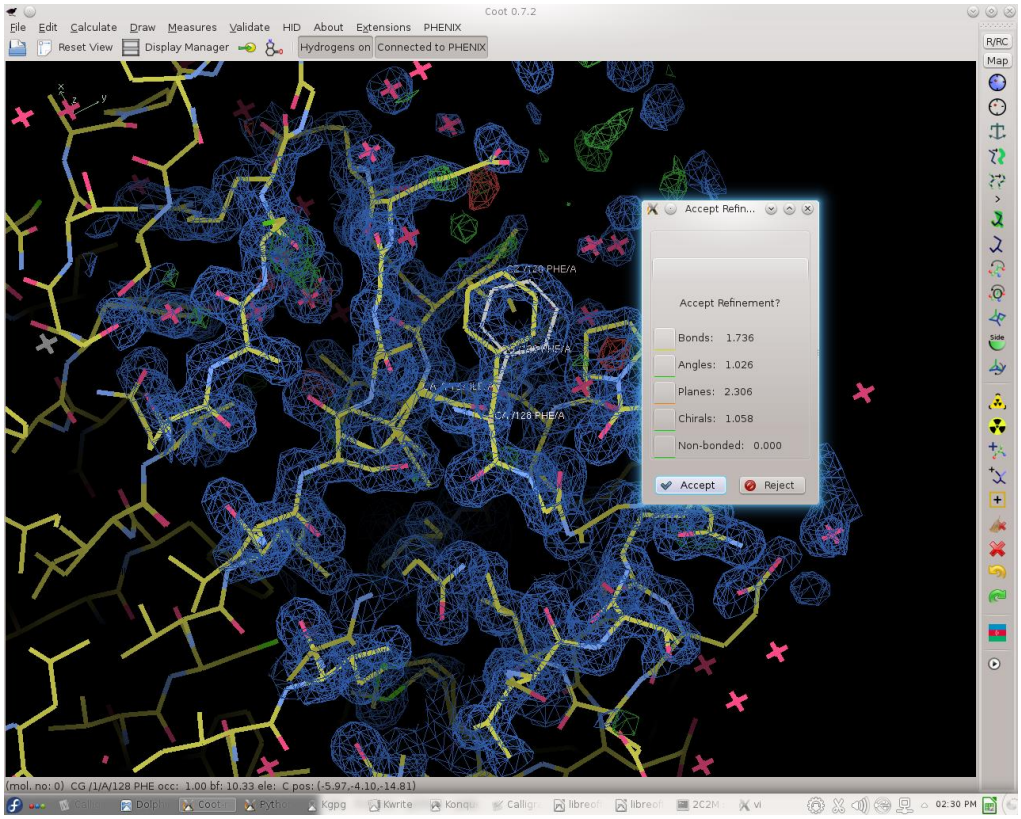


Click “Real Space Refine” on the user interface on the right side of the screen, the cselect two atoms on the side chain which is to be moved.



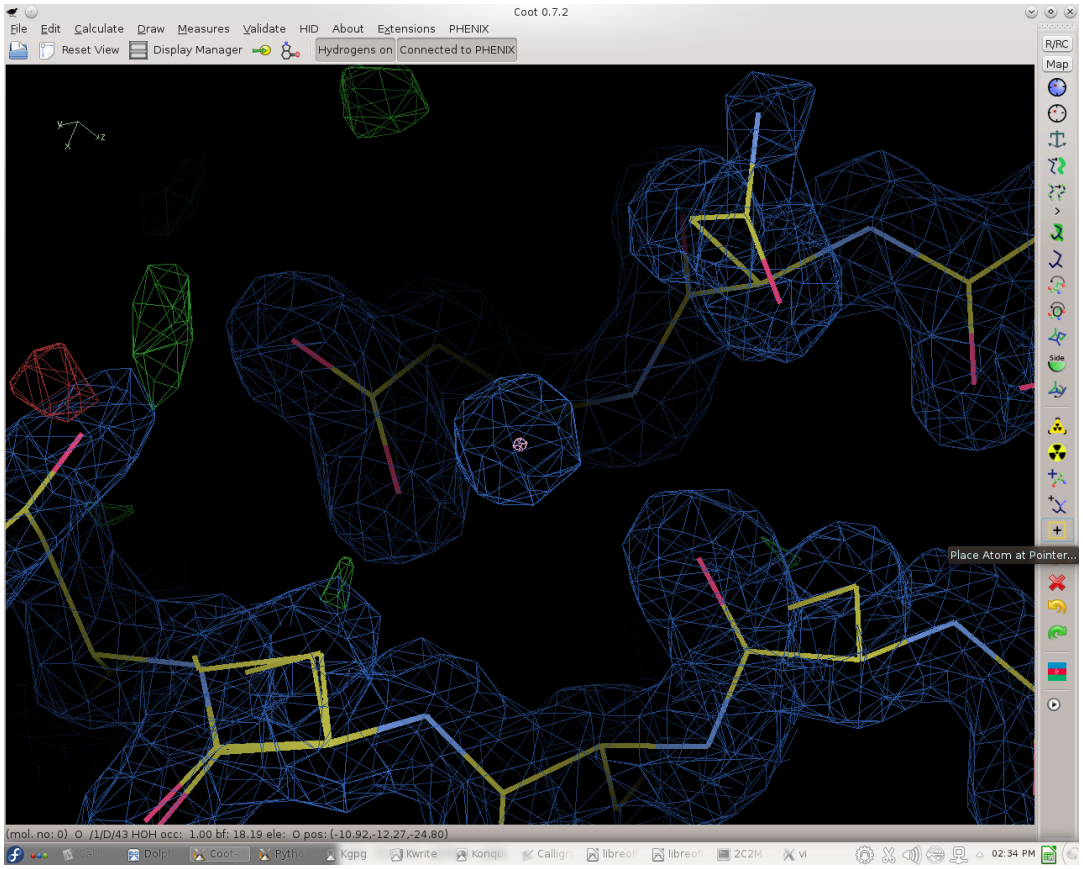
After selecting the two atom positions the side chain will move into the electron density as a light-white side chain and a new window will appear which prompts you to either

Accept or Reject the refinement.

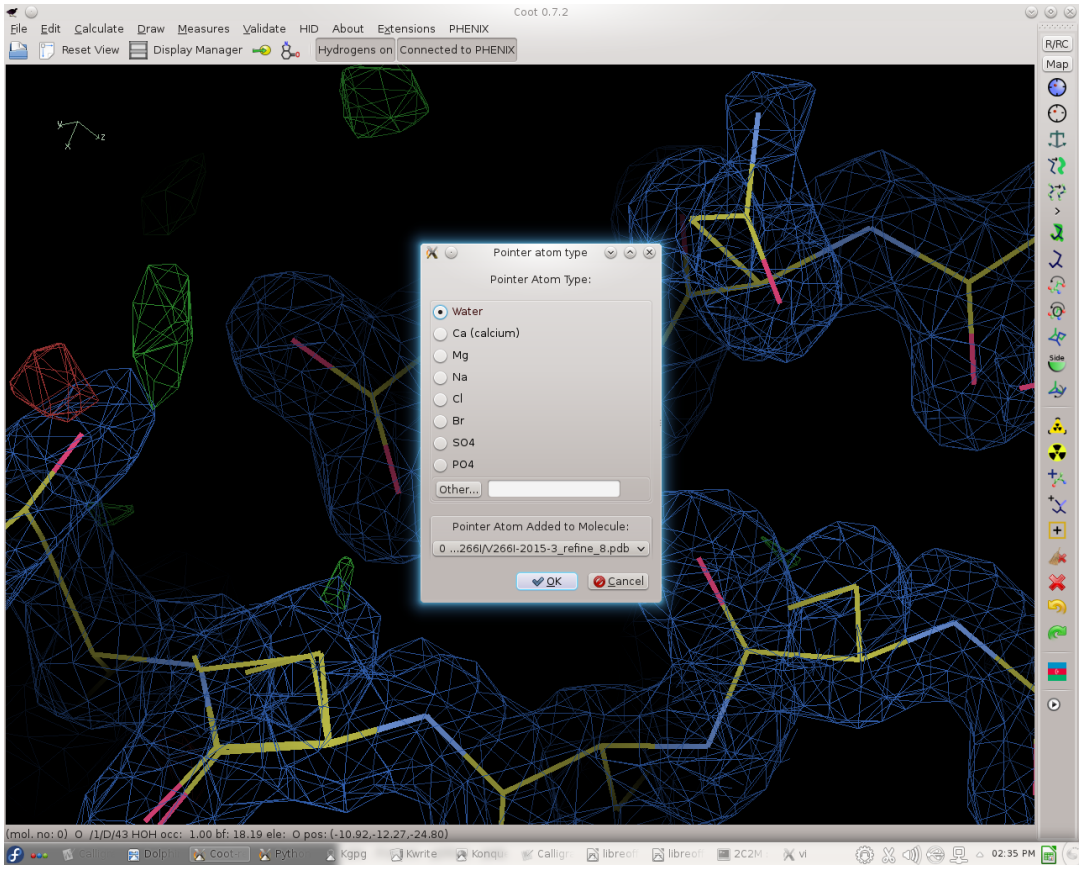


Addition of Atoms

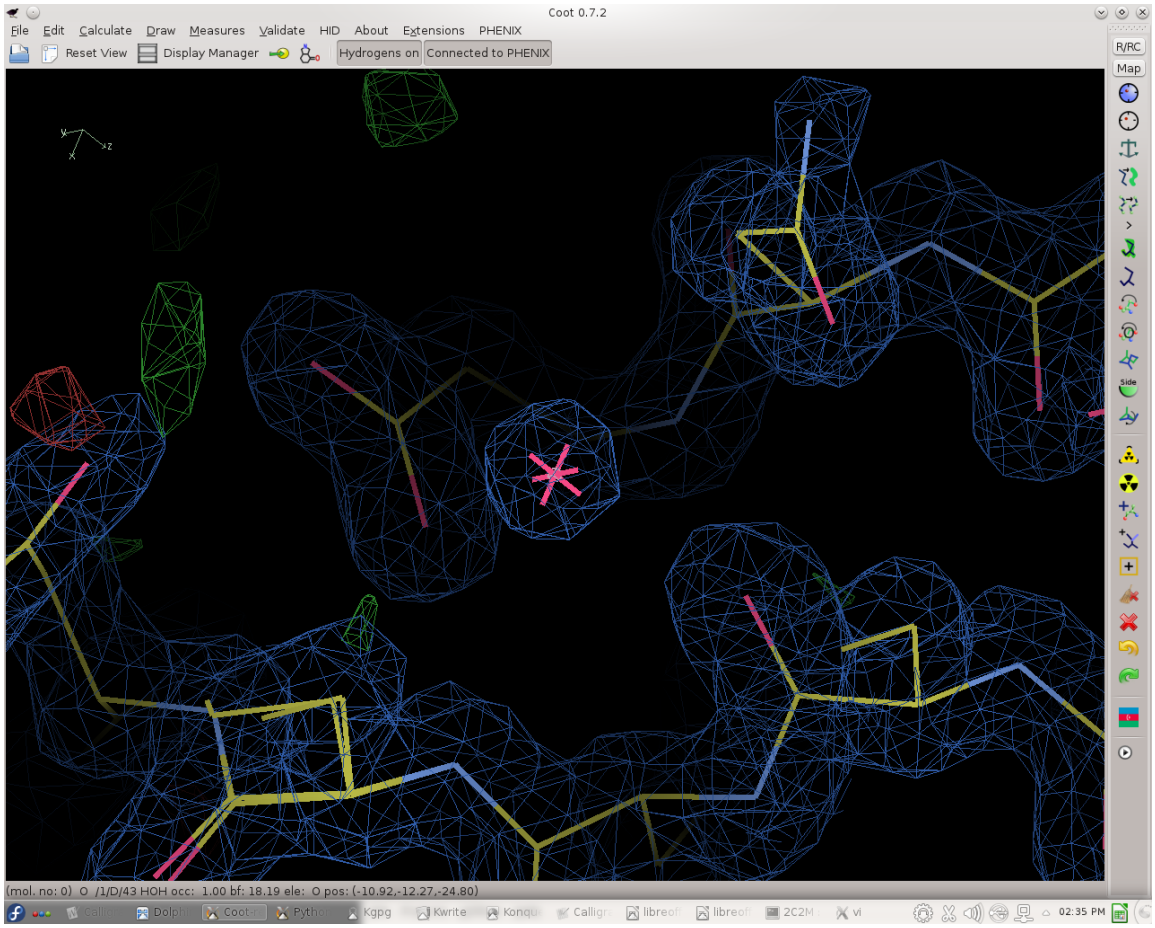
Move the cursor into the center of an area of electron density by holding “Ctrl” and dragging the mouse. It is important to check the placement of the cursor in multiple dimensions as this is a 3D model. Select “Place Atom at Pointer” in the user interface on the right side of the screen.



Select the atom of interest you wish to place in the electron density, additionally select the .pdb file you wish to save the molecule to by clicking the drop down menu labeled “Pointer Atom Added to Molecule”. If you do not select the .pdb file Coot will save the atom into a new .pdb file.



One you click “OK” a new atom will be present in the electron density, subsequent refinements using Phenix will indicate whether or not this atom fits well in the electron density data.



Useful Coot Pathways

- "Draw→Go to Atom", Allows the user to select an atom or molecule and to have the cursor placed on the area of interest.
- "Measures→Distances & Angles→Distance", Allows the user to calculate the distance between two atoms by licking each of them in succession.
- "Draw→Cell & Symmetry", Click Ok on Master Switch: Show Symmetry Atoms? in order to generate the symmetry mate of the model.
- "Validate→Distance Map Peaks", Displays negative and positive density peaks between the model and the density map.

- "File→ Save Coordiantes", Saves the .pdb file.
- Additionally, Using the "Space Bar" will move the user to the next residue in the protein sequence

All screen shots and material used in this subsection are adapted from the COOT program (2).

DROP ANALYSIS PROTOCOL

1. Begin by creating a file directory for your analysis. This file must contain the 2J30 .pdb file available on the pdb website. Additionally you will need to add a .pdb file for the mutant you wish to compare to 2J30. Both of these files must be in the same folder. After both files are present you will compress the folder as a .zip folder. This .zip folder can now be uploaded to http://dropinthemattoslab.org/ .

2. After the website says that your analysis is finished you can download the results by clicking on the HERE in the following line. Run Complete: click Here for results.

3. Once the file is finished downloading unzip the file to the directory of your choosing. Open the folder and click on the image watercluster_labeled.png. This image shows a bar graph with the number of clusters on the y-axis and cluster size on the x-axis. This graphs tells you how many waters are only seen in one protein (cluster size=1) and how many waters are present in both molecules (cluster size=2). If you wish to compare >2 .pdb files the cluster size will increase accordingly. The most important piece of data that you need from this image is the number of clusters in the AFTER category (i.e. 161 clusters in a cluster size of two means that there are 161 waters conserved between both molecules).

4. Once you have acquired the “number” go back to the folder containing your results. Open the file clusters.pdb in a text editor program (i.e. Kwrite for linux). This .pdb file contains all of the water molecules in order of cluster size. This means that the waters conserved in both structures are listed first followed by the waters only present in one structure. Using our previous example of 161 waters, we would copy the first 161 waters listed in the clusters.pdb and paste them into a new file. This file must be saved as a .pdb and will serve as the conserved water file between both of your structures. Save this file in the same directory as your results.

5. In a pymol session open the 2J30 file along with the 93-100% conserved waters .pdb which was previously created. Delete the endogenous waters in the 2J30 file. Go to save molecule and select both the 2J30 and the 93-100% conserved water file. You have now made a single file containing the 2J30 molecule and the 93-100 conserved waters. The reason for this is because you are not allowed to align a file with only waters to a structure file. This file will be used for all caspase-3 water comparisons.

6. Now open a new pymol session. In this window open one of the structure files from your analysis and the conserved water file you created by cutting and pasting. Make sure they line up. If they are not lined up you have done something wrong. Once you have checked the alignment open your 2J30-93-100%conserved file and align it to the structure file that you opened at the beginning of step 6. Once this is done go to show-->spheres for the conserved water file you made in step 4. Render this file yellow or green using the color control on the right side of the pymol window. Now we must create spheres for only the water molecules of the 2J30-93-100%conserved file. The waters are

all on chain c in the .pdb so you can type “show spheres, chain c and 2J30-93-100%conserved”. This will now show all of the waters as red.

7. Once you have a set of red and yellow/green spheres you can hide all of the structures until you only see the red spheres and the yellow/green spheres. You are now able to see which waters are conserved by the presence of two overlapping colors. Red waters which do not overlap the yellow/green spheres can be selected and colored blue by right clicking the water-->atom-->color-->blue. It is important to note that only the 2J30-93-100%conserved waters should be changed to blue as the conserved file contains more waters than are represented in the initial 93-100conserved analysis (145 waters).

8. When you are finished triple check that you have not made any mistakes. Hide the conserved water file you created in step 4

B-FACTOR ANALYSIS

The b-factor analysis was completed using the following method. Firstly open the final .pdb structure in the program “Pymol”. In the Pymol command line enter “create carbonalpha, name ca” this will create a new object named carbonalpha which displays only the alpha carbons. Next, go to save molecule and select the carbonalpha object. This object can be saved as a new .pdb file named as “MutantName CarbonAlpha”. Open the new .pdb into an excel workbook, word or text editor and copy the right most numerical column corresponding to each amino acid. The right most numerical column is the b-factor of the atom. Copy this column and paste it into a new excel sheet, label the column with the appropriate .pdb heading. For the wild-type structures an excel sheet was created with each of the fifteen structures, it is important that the residue numbers line up correctly as some structures possess more or less electron density for a given residue. The

average of the wild-type structures can then be generated along with standard deviations. The carbon alpha b-factors for the mutants are then compared to the wild-type structures using by calculating ΔB-Factor=(B-Factor$_{mut}$ – B-Factor$_{wild\text{-}type}$).

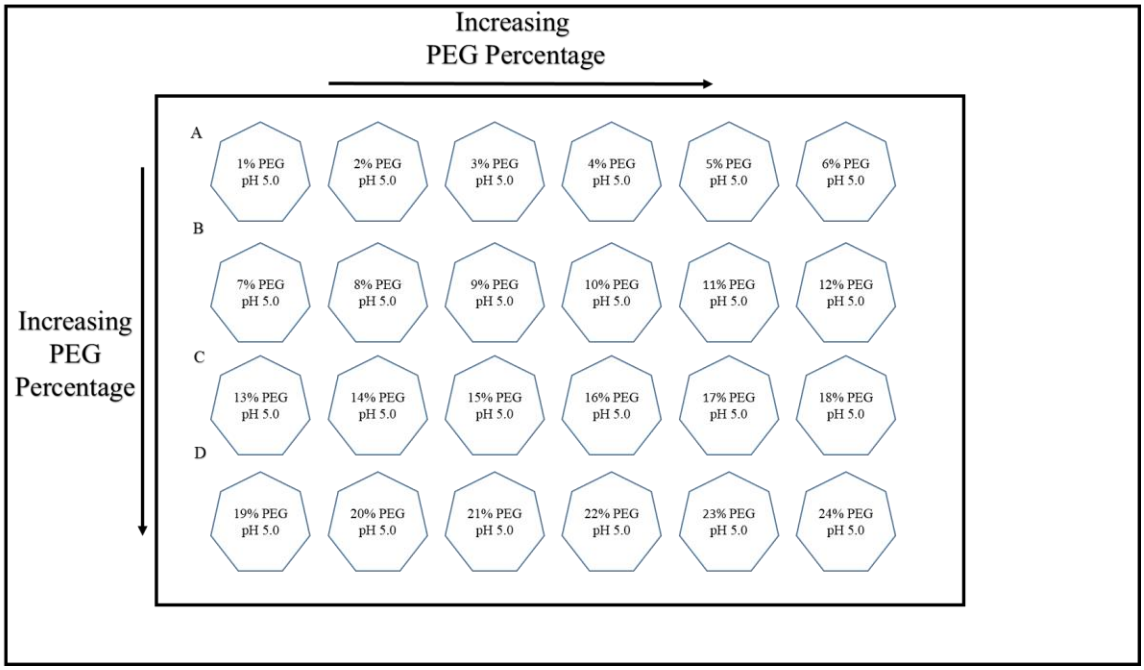


Fig. 1 Initial crystallography screening method for caspase mutants.

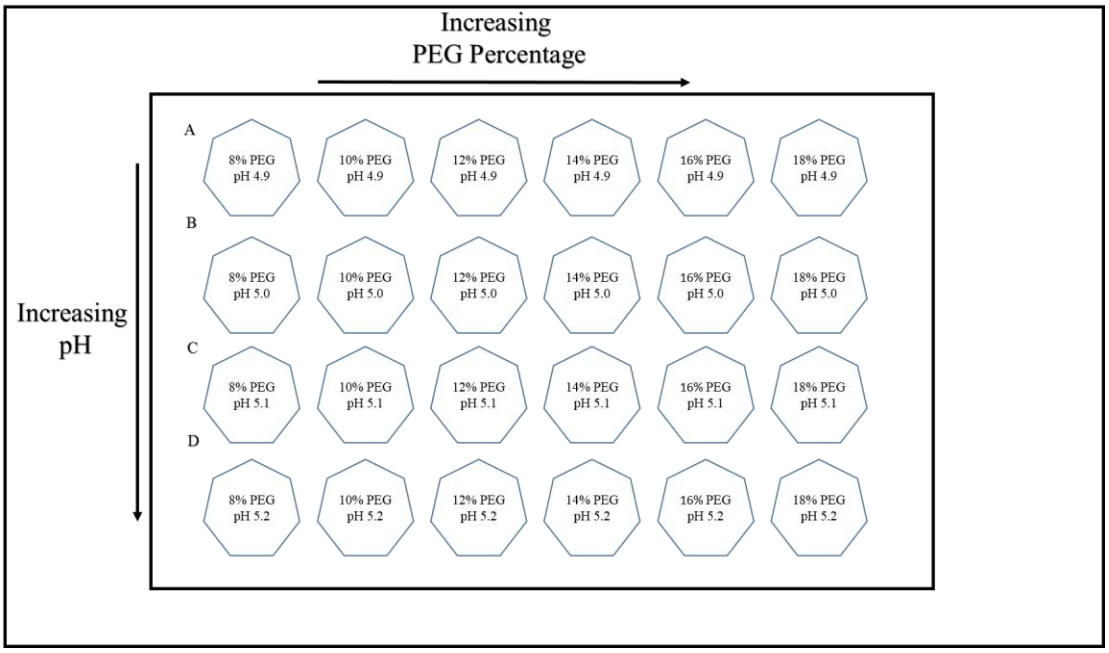


Fig. 2 pH screen for caspase mutants.

REFERENCES

1.) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., ... & McCoy, A. J. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 213-221.

2.) Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 486-501.