

MOL #49486

Thermostable Variants of Cocaine Esterase for Long-Time Protection against Cocaine Toxicity

Daquan Gao,¹ Diwahar L. Narasimhan,¹ Joanne Macdonald, Remy Brim, Mei-Chuan Ko, Donald W. Landry, James H. Woods, Roger K. Sunahara, and Chang-Guo Zhan

Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 725 Rose Street, Lexington, KY 40536 (D.G., C.-G.Z.); Department of Pharmacology, University of Michigan Medical School, 1301 Medical Sciences Research Building III, Ann Arbor, MI 48109 (D.L.N., R.B., M.-C.K., J.H.W., R.K.S.); Division of Clinical Pharmacology and Experimental Therapeutics, College of Physicians & Surgeons, Columbia University, 630 West 168th Street, Box 84, New York, NY 10032 (J.M., D.W.L)

MOL #49486

Running Title: Thermostable Variants of Cocaine Esterase

Correspondence:

Chang-Guo Zhan, Ph.D.

Professor

Department of Pharmaceutical Sciences

College of Pharmacy

University of Kentucky

725 Rose Street

Lexington, KY 40536

Voice: 859-323-3943

FAX: 859-323-3575

E-mail: zhan@uky.edu

Number of text pages: 11

Number of tables: 0

Number of figures: 4

Number of references cited: 39

Number of words in the Abstract: 125

Number of words in the Introduction: 750

Number of words in the Discussion: 169

ABBREVIATIONS: CocE, cocaine esterase; MD, molecular dynamics; RMSD, root-mean-square-deviation; PME, particle mesh Ewald; NPT, constant temperature and constant pressure; NVT, constant temperature and constant volume.

MOL #49486

ABSTRACT

Enhancing cocaine metabolism by administration of cocaine esterase (CocE) has been recognized as a promising treatment strategy for cocaine overdose and addiction, as CocE is the most efficient native enzyme for metabolizing the naturally occurring cocaine yet identified. A major obstacle to the clinical application of CocE is the thermoinstability of native CocE with a half-life of only a few minutes at physiological temperature (37°C). Here we report thermostable variants of CocE developed through rational design using a novel computational approach followed by *in vitro* and *in vivo* studies. This integrated computational-experimental effort has yielded a CocE variant with a ~30-fold increase in plasma half-life both *in vitro* and *in vivo*. The novel design strategy can be used to develop thermostable mutants of any protein.

MOL #49486

Cocaine strongly reinforces self administration (Mendelson et al., 2000; Singh, 2000; Paula et al., 2004) and the disastrous medical and social consequences of cocaine addiction have made the development of an effective pharmacological treatment a high priority (Gorelick, 1997; Redish, 2004). The classic CNS receptor-antagonist approach has failed to yield an anti-cocaine therapeutic but we developed a proof of principle for a peripheral blocker to accelerate cocaine metabolism in the circulation (Landry et al., 1997; Mets et al., 1998) producing biologically inactive metabolites *via* hydrolysis of cocaine benzoyl ester (Gorelick, 1997; Zhan et al., 2003; Rogers et al., 2005; Meijler et al., 2005; Pan et al., 2005; Pan et al., 2007; Zheng et al., 2008). The bacterial cocaine esterase (CocE) (Bresler et al., 2000), the most efficient native cocaine hydrolase yet identified (Larsen et al., 2002), is particularly promising. In rodent models, CocE can both prevent and reverse extreme cocaine toxicity (Cooper et al., 2006; Ko et al., 2007). However, a major obstacle to the clinical application of CocE is the thermostability of native CocE with a half-life of only a few minutes at physiological temperature (37°C) (Cooper et al., 2006). It is highly desirable to develop thermostable mutants of CocE for therapeutic treatment of cocaine toxicity.

Protein engineering for thermostability is a particularly exciting and challenging field (Lehmann et al., 2001). In addition to directed evolution, various rational methods have been developed for thermostable mutant design, and a key factor considered in these methods is the interactions of amino acids within a protein's core. Enzyme engineering for thermostability poses additional challenges, because the active site structure of an enzyme and its dynamic behavior during an enzymatic reaction often appear fine-tuned for the optimum catalytic efficiency (Korkegian et al., 2005). To stabilize an enzyme without losing catalytic efficiency, a computational design method must be capable of predicting thermostable mutations within a given fold while minimizing any shift in the backbone that might structurally disrupt the active site structure or quench its flexibility. Computational design has been used successfully to thermostabilize noncatalytic proteins (Dahiyat, 1999; Luo et al., 2002; Malakauskas et al., 1998; Dantas et al., 2003), redesign binding pockets (Benson et al., 2002; Reina et al., 2002; Shifman et al., 2002; Looger et al., 2003), create a protein fold (Kuhlman et al., 2003), and design catalytic activity into a bacterial receptor (Dwyer et al., 2004). Stoddard and associates recently demonstrated that computational design could also help to thermostabilize an enzyme, as they successfully designed thermostable mutants of a small enzyme (homodimer), *i.e.* yeast cytosine

MOL #49486

deaminase (yCD) (Korkegian et al., 2005). The 153 amino acid protein displays irreversible unfolding behavior at high temperature. RosettaDesign (Kuhlman et al., 2000) was used to evaluate the fitness of a particular sequence for a given fold, and a Monte Carlo search algorithm was used for sampling sequence space. The program requires a backbone structure as input, uses an empirical energy function, and generates sequences predicted to have the lowest energy for that fold. Excluding the residues located within 4 Å of the active site or involved in the dimer interface, or examined in other published studies, their computational design of the yCD mutants only needed to consider possible mutations on 33 amino acid residues.

CocE represents a considerably more challenging enzyme to computationally derive a thermostable enzyme due to its relatively large size (574 amino acids). Firstly, the computational design and subsequent experimental tests need to take into account many more amino acid residues in order to identify the best possible thermostable mutations. In addition, it is unnecessary for the enzyme to unfold before it becomes inactive. Enzyme inactivation could be merely associated with minor structural changes on the least-stable region of the protein, without completely unfolding the enzyme. In this case, a mutation lowering the total folding energy of an enzyme does not necessarily lead to a longer half-life of the active enzyme structure. Hence, it is crucial to uncover the inactivation mechanism while modeling thermostabilization of an enzyme.

In the present study, molecular modeling and simulations have led to important mechanistic insights into the inactivation process of CocE. Herein we report thermostable variants of CocE developed through structure-and-mechanism-based design followed by *in vitro* and *in vivo* studies. This integrated computational-experimental effort has yielded CocE variants with melting temperatures up to 7 degrees higher than the native enzyme and a ~30-fold increase in plasma half-life both *in vitro* and *in vivo*. This robust method can be used to develop thermostable mutants of any protein.

Materials and Methods

Molecular modeling and simulation. The mechanism of enzyme inactivation is very difficult to unveil through computational simulation. For example, CocE has a half-life ($\tau_{1/2}$) of a few minutes (see below), which is too short for use as a drug, but too long for performing a molecular dynamics (MD) simulation to simulate the inactivation process. This is because a practical, fully relaxed MD simulation (with a required time step of 1 or 2 fs) on a protein system

MOL #49486

like CocE can be performed for as long as nanoseconds using currently available supercomputers. Instead, we based our computational design strategy on an analysis of the kinetic relationship between the inactivation rate constant (k_{ina}) and the temperature (T) for a given inactivation free energy barrier (ΔG_{ina}) of the inactivation process, *i.e.*

$$k_{\text{ina}} = (k_{\text{B}}T/h)\exp(-\Delta G_{\text{ina}}/RT), \quad (1)$$

which is analogous to the rate constant equation used in the well-known variational transition state theory (Corchado et al., 1998). In Eq.(1), k_{B} is Boltzmann's constant and h is Planck's constant. According to Eq.(1), the rate constant k_{ina} of the enzyme inactivation is dependent on the temperature for a given inactivation free energy barrier. The higher the temperature, the larger the rate constant and, therefore, the half-life of the active enzyme is shorter. This kinetic understanding enables us to reveal the inactivation pathway of CocE through performing MD simulation at an appropriately high temperature. The high temperature used in the MD simulations is physiologically irrelevant, but can considerably shorten the time of the protein inactivation process and, thus, allows us to observe the protein inactivation process within nanoseconds.

The initial structures constructed from the X-ray crystal structure (Larsen et al., 2002) were neutralized by adding sodium counterions and were solvated in a rectangular box of TIP3P water molecules with a minimum solute-wall distance of 10 Å. All of the MD simulations were performed by using the Sander module of Amber 8 package (Case et al., 2004). The solvated systems were carefully equilibrated and fully energy-minimized. These systems were gradually heated from $T = 10$ K to $T = 298$ K in 30 ps before production MD simulations at 298 K (in the NPT ensemble with $P = 1$ atm) and then at 575 K (in the NVT ensemble). The time step used for the MD simulations was 2 fs. Periodic boundary conditions with Berendsen temperature coupling (Berendsen et al., 1984) with isotropic molecule-based scaling (Berendsen et al., 1984) were applied. The SHAKE algorithm (Ryckaert et al., 1977) was used to fix all covalent bonds containing hydrogen atoms. The non-bonded pair list was updated every 10 steps. The particle mesh Ewald (PME) method (Darden et al., 1993) was used to treat long-range electrostatic interactions. A residue-based cutoff of 10 Å was utilized to the non-covalent interactions. Further, the protein structures simulated at 298 K were subjected to the energy minimization prior to the interaction energy calculations in the gas phase. The solvent shifts of the interaction energies were calculated by using the Delphi program (Gilson et al., 1988). In addition, the

empirical scoring function implemented in RosettaDesign program (Kuhlman et al., 2000) was also used to estimate the interaction energies.

Site-directed mutagenesis. Point mutations were generated using QuickChange (Stratagene) and CocE cDNA cloned in the bacterial expression vector, pET-22b (+). Double mutants were generated using single-point mutations as templates. Oligonucleotide sequences are available on request. All mutants were sequenced in both directions over the entire coding region. All enzymes were expressed as 6xHis-tagged proteins in *E. coli* BL-21 (DE3) cells grown at 37°C. Protein expression was induced with 1 mM isopropyl- β -thiogalactopyranoside (IPTG, Fisher) for 12 hours at 18 °C.

Purification of cocaine esterase and mutants. Cells were pelleted, resuspended in 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM dithiothreitol (DTT), and a protease inhibitor cocktail (34 μ g/ml each of L-tosylamido-2-phenylethyl chloromethyl ketone, 1-chloro-3-tosylamido-7-amino-2-heptanone and phenylmethylsulfonyl fluoride, and 3 μ g/ml each of leupeptin and lima bean trypsin inhibitor) and lysed using a French press (Thermo Fisher Scientific Corp, USA). 6xHis-tagged enzymes were enriched using Talon metal affinity chromatography (Clontech Laboratories, Inc, Mountain View CA) and purified using anion-exchange (Q-Sepharose, GE Healthcare, Piscataway NJ) chromatography. CocE was eluted from the Q-Sepharose column with 150-450 mM NaCl linear gradient buffers containing 20 mM Hepes, pH 8.0, 2 mM MgCl₂, 1 mM EDTA and 1 mM DTT. The peak fractions were pooled and concentrated by Centricon-30 (Millipore), snap frozen in liquid nitrogen and stored at -80°C.

Michaelis-Menten kinetics of cocaine hydrolysis. Cocaine hydrolysis was monitored as previously described (Turner et al., 2002). The initial rates of decay were estimated by following the change in the intrinsic absorbance of cocaine at 240 nm with time using a SpectraMax Plus 384 UV plate reader (Molecular Devices, Sunnyvale, CA) and SOFTmax Pro software (Version 3.1.2). The initial rates were estimated from the linear portion of the progress curves and spanned no longer than 5 minutes. The reaction was initiated by adding 100 μ L of a 2X enzyme solution (100 mM Phosphate Buffer, pH 7.4 and 300 mM NaCl) to 100 μ L of a 2X cocaine solution (50 ng/mL enzyme, 100 mM Phosphate Buffer, pH 7.4 and 300 mM NaCl). Final cocaine concentrations were as follows: 125, 62.5, 31.25, 15.63, 7.81, 3.91, 1.95, and 0.977 μ M. V_{\max} and K_M values were calculated using Prism (GraphPad Software, San Diego). Stability measurements were performed using the native and mutant enzymes diluted to 2X concentration

and incubated at 37°C for various times. At the end of each time point, an aliquot was removed and kinetic behavior was observed as mentioned above. Percent remaining activity was plotted against time to assess the half-life ($\tau_{1/2}$) of each enzyme.

Circular dichroism measurement and thermal denaturation experiments. CD data were collected on a J-810 spectropolarimeter (JASCO, Easton, MD). Near-UV wavelength scans were run from 250 – 320 nm with a 0.2 nm pitch, varying protein concentrations (5 – 30 μ M), temperatures increasing from 5 – 70°C in 2.5°C increments, and a 1 cm pathlength water-jacketed cylindrical quartz cell. Four elliptical curves were generated per temperature measurement, and denaturation was recorded as a change in average ellipticity over temperature. Apparent T_m 's were determined by deconvolution to component spectra using the convex constraint algorithm (Percezel et al., 1992) into 2 curves using 30 iterations, and subsequent non-linear regression curve-fitting using Prism 5 (GraphPad Software, San Diego, CA).

***In vivo* Protection against cocaine lethality.** Male NIH-Swiss mice (25 – 32 g, Harlan Inc., Indianapolis, IN) were housed in groups of 6 mice per cage, allowed *ad libitum* access to food and water, and were maintained on a 12-h light-dark cycle with lights on at 06:30 AM in a room kept at a temperature of 21 – 22°C. Mouse body temperatures are assumed to be in the well-established range of 36.5-38°C (Harkness et al., 1989). All protocols and experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals as adopted and promulgated by the National Institutes of Health and were approved by the University Committee on the Use and Care of Animals at the University of Michigan.

Cocaine-induced toxicity was characterized by the occurrence of lethality, as defined by the cessation of observed movement and respiration. Lethality was recorded following intraperitoneal (i.p.) cocaine administration (LD_{100} , 180 mg/kg) on mice subjected to pretreatment of purified native CocE or mutants (i.v. injection of 0.2 mL) for various times (1, 5, 10, or 30 min, 1, 2, 3, 4, 5 hours after enzyme administration). Each treatment used 8 mice to assess the percent of lethality (*i.e.* protection) in mice pretreated with a single dose of an esterase at a single time point.

Blood Collection and Western Blotting: Blood was collected 1 hr and 2 hr post-treatment from the submandibular area. This area was punctured using a mouse bleeding lancet (GoldenRod 4.0 mm animal lancet; MEDpoint Inc., Mineola, NY). Blood was collected into BD Microtainer® collection tubes (BD Diagnostics Serums Separator Tubes, 365956)) and placed on

MOL #49486

ice. Samples were spun at 4000 rpm for 5 minutes to separate the plasma. 10 μ g of plasma proteins were loaded onto a 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane. Membranes were then probed with anti-CocE rabbit antibody followed by HRP-linked anti-rabbit secondary antibody. As a loading control, blots were stripped and then reprobed with mouse anti-apolipoprotein-A1 antibody.

Results

We used Amber (Case et al., 2004) to carry out MD simulations on native CocE in water. The simulations, starting from the X-ray crystal structure (Larsen et al., 2002) of CocE, were performed first at room temperature (298 K) for 250 ps followed by an increase in temperature to 575 K for 1050 ps (*i.e.* 250 ps to 1300 ps in Figure 1(A)). Depicted in Figure 1(A) is the time-dependence of the root-mean-square-deviation (RMSD) of the atomic positions in the simulated CocE structure from the corresponding positions in the X-ray crystal structure. The RMSD value reflects the magnitude of the overall change of the protein structure during the MD simulation. A stable protein structure is usually expected to have a reasonable RMSD value of ~ 2 Å for the overall protein structure due to dynamics of protein in water. A significantly larger RMSD value may be associated with some major structural change that could cause the protein to lose its normal function. Figure 1(A) shows that the CocE structure was quickly stabilized during the MD simulation at 298 K. The RMSD value became increasingly larger during the further MD simulation at 575 K and reached a value for the overall protein structure of 4 Å at ~ 1000 ps. In order to identify the pathway for the major structural change of CocE during the MD simulation at 575 K, we further visualized the detailed structural changes during the MD simulation and found that amino acid residues G165 to I195 in domain II (consisting of amino acid residues D145 to L240) of CocE first started to have the major structural changes. Figure 1(B) depicts the three domains (I to III) of CocE. As seen in Figure 1(B), part of the substrate-binding pocket belongs to domain II. Based on the MD simulation, the region from residues G165 to I195 in domain II is the least stable region of the CocE structure. As seen in Figure 1(A), the RMSD value for residues G165 to I195 in domain II is much larger than the corresponding RMSD value for the overall CocE structure. Elevation to 575 K imposes distortions in Domain II that would likely lead to disruption of the substrate-binding pocket and loss of catalytic activity due to their relative close proximity. Eventually, the distortion may lead to global alterations in protein

MOL #49486

structure. The enzyme is expected to be inactive following disruption of domain II structure, regardless of how stable domains I and III are. Without altering the side chains on residues G165 to I195 in domain II, decreasing the total folding energy by stabilizing domain I or III should not increase the half-life of the active enzyme structure.

The above mechanistic insights obtained from the MD simulations suggest that rational design of possible thermostable mutants of CocE should therefore focus on domain II, particularly residues G165 to I195. This may be accomplished by increasing either hydrogen bonding between domains I and II or by increasing some favorable interactions between amino acid residues within domain II. Hence, our current computational design only targets residues of domain II that are at least 5 Å away from (-)-cocaine atoms according to our modeled CocE-(-)-cocaine complex structure. The energetics of possible mutations on these amino acid residues were estimated by using two different approaches. The first approach calculates the shift in the interaction energy between the mutated residue and the remaining part of the enzyme, through a combined application of Amber and Delphi (Gilson et al., 1988) programs. The structures of both native CocE and mutants were optimized (through energy minimization) prior to the interaction energy calculations.

The other approach estimates the change in folding energy of the mutant by using the RosettaDesign program for folding energy calculations on the optimized structures of native CocE and its mutants. These two computational approaches can complement each other. Whereas the former uses a more sophisticated theoretical energetic approach, the later uses the empirical scoring function developed specifically for folding energy calculations. We can reasonably expect that when both approaches consistently predict that a mutation on a residue in domain II can stabilize domain II, the prediction should be more reliable. Indeed, both computational approaches consistently predicted that several mutations, including T172R and G173Q, should stabilize domain II. Mutation G173Q was predicted to build a bridge between domains I and II through a hydrogen bond between residues P43 of domain I and Q173 of domain II (see Figure 1(C)), whereas mutation T172R was predicted to have an improved interaction between residue #172 and the other residues in domain II (with a lower interaction/folding energy). The CocE-wt crystal structure (Larsen et al., 2002) reveals that the alkyl chain of Thr172 on helix 5 points toward helix 6 within domain II but away from domain I. Substitution of arginine at this location introduces a longer alkyl moiety and a charge that we

MOL #49486

rationalize may act in concert to contact and stabilize helix 6. The stabilization of domain II itself may influence the overall stability of the protein. Although the computational evaluations suggest that the effects of mutations T172R and G173Q on the energetics should be additive, improving the integrity of domain II itself may help G173Q to facilitate interactions with domain I in a more cooperative manner.

To validate our computational predictions we assessed the effect of T172R, G173Q, and T172R/G173Q in comparison to native CocE (CocE-wt) using *in vitro* and *in vivo* experimentation. Using purified enzyme preparations we directly compared their catalytic activities and half-lives at a physiological temperature (37°C). The purified preparations are greater than 95% pure by SDS-PAGE and coomassie blue staining (Figure 2(A)) and display appropriate Michaelis-Menten kinetics with the following catalytic constants: CocE-wt ($K_M \sim 21 \mu\text{M}$, and $k_{\text{cat}} \sim 2323 \text{ min}^{-1}$), T172R ($K_M \sim 24 \mu\text{M}$ and $k_{\text{cat}} \sim 2502 \text{ min}^{-1}$), G173Q ($K_M \sim 15 \mu\text{M}$ and $k_{\text{cat}} \sim 2384 \text{ min}^{-1}$), and T172R/G173Q ($K_M \sim 24 \mu\text{M}$ and $k_{\text{cat}} \sim 2247 \text{ min}^{-1}$). The effects of temperature on thermostability are illustrated in Figure 2 (B) and summarized in Figure 4.

Mutant enzymes were generated by site-directed mutagenesis and expressed and purified as described previously (Cooper et al., 2006). The *in vitro* enzyme activity depicted in Figure 2 (B) reveal that preincubation of native CocE at 37°C decreased the enzymatic activity exponentially with time ($\tau_{1/2} = \sim 11 \pm 0.9 \text{ min}$). In contrast, the CocE mutants T172R, G173Q, and T172R/G173Q displayed significantly longer half-lives without changing the enzymatic activity prior to preincubation at 37°C: $\tau_{1/2}(\text{T172R}) = \sim 78 \pm 6.5 \text{ min}$ (~ 7 -fold increase), $\tau_{1/2}(\text{G173Q}) = \sim 75 \pm 9.9 \text{ min}$ (~ 7 -fold increase), and $\tau_{1/2}(\text{T172R/G173Q}) = \sim 305 \pm 38 \text{ min}$ (~ 30 -fold increase). Careful analysis of the activity of T172R/G173Q following incubation at 37°C suggests a more complicated mechanism of inactivation at 37°C compared to CocE-wt or the single mutants. While CocE-wt, T172R and G173Q fit very well to a single-phase exponential decay, T172R/G173Q displays multi-phase inactivation kinetics with $\tau_{1/2\text{short}} \sim 180 \text{ min}$ and $\tau_{1/2\text{long}} \sim 4900 \text{ min}$. Although we have little data to pinpoint a precise mechanism for the biphasic appearance the data do reflect the cooperative, rather than additive, effects of combining T172R and G173Q mutations. These data correlated well with subsequent circular dichroism measurements (Figure 3) of thermal denaturation, revealing a 3.1 – 3.5°C increase in apparent denaturation temperature (T_m) for single mutants and a 7.1°C increase for the double mutant

MOL #49486

T172R/G1734Q. The data suggest an overall loss in helicity of the enzyme preparations with respect to temperature and indicates the thermal denaturation of the enzyme structure. CocE is approximately 28% helical in nature however domain II accounts for approximately one-fifth of the overall helicity. One may therefore rationalize that mutants which stabilize domain II or the interaction with domain I and domain II should preserve the observed helicity at higher temperatures in comparison to the native enzyme.

Assessment of the thermostability using *in vivo* models based on protection of native CocE (CocE-wt) and its T172R and T172R/G173Q mutants against cocaine-induced lethality correlated well with the *in vitro* data. Depicted in Figure 4(A) are the determined *in vivo* half-lives of the enzymes in comparison with the corresponding *in vitro* half-lives at 37°C. As seen in Figure 4(A), the *in vivo* half-life of CocE-wt is $\sim 11 \pm 1.4$ min, whereas both the T172R and T172R/G173Q mutants have significantly longer half-lives, *i.e.* $\sim 95 \pm 13$ min (~ 9 -fold increase) and $\sim 262 \pm 46$ min (~ 24 -fold increase), respectively, which are consistent with the *in vitro* data. Western Blotting analysis of CocE-wt and the T172R/G173Q mutant in serum following i.v. injection in mice suggests that the intact, immunoreactive species is present for at least 2 hours (Figure 4(B)) with no detectable proteolysis. Careful analysis of the data in Figure 4(B) suggests that CocE-wt appears to disappear faster than the T172R/G173Q mutant (after 1 hour) but at a slower rate than suggested by the protection from lethality data (~ 11 min). In contrast, serum levels of the T172R/G173Q mutant appear to be unchanged two hours post-injection. These data support the notion that the improved thermostability of the T172R/G173Q mutant may account for its prolonged *in vivo* protective effect on the cocaine-induced lethality. The identification of thermostable mutants of CocE that display significantly longer circulation time in the body provide promising therapeutic potential for the treatment of cocaine overdose. The T172R/G173Q mutant displays an extended half-life beyond the plasma half-life of cocaine in humans (~ 45 to 90 min) (Landry et al., 1993). More extensive *in vivo* studies on these thermostable mutants of CocE will be described elsewhere.

Discussion

Biotechnology-based therapeutics has recently garnered considerable attention in the pharmaceutical industry. The relatively lower-toxic and metabolic potential of protein-based therapeutics are their main attraction. One of the most challenging problems complicating the use

MOL #49486

of proteins as drugs is the issue of thermostability. The success in generating more thermostable forms of CocE described here suggests that our structure-and-mechanism-based design approach is a promising and efficient general computational approach for protein redesign. In particular, to design thermostable mutants of a protein, one can first perform the MD simulation on the protein at an appropriately chosen high temperature to understand the protein inactivation pathway. Based on the detailed understanding of the protein structure and inactivation mechanism, one can carry out a virtual screening of various possible mutants through the interaction energy calculations in order to predict the most likely thermostable mutants for wet experimental tests *in vitro* and *in vivo*. This computational approach will likely represent a valuable strategy for thermostabilization of other proteins and have dramatic implications on their therapeutic potential.

Acknowledgements

The authors also acknowledge the Center for Computational Sciences (CCS) at University of Kentucky for supercomputing time on an IBM X-series Cluster with 1,360 processors.

References

- Benson DE, Haddy AE, and Hellinga HW (2002) Converting a maltose receptor into a nascent binuclear copper oxygenase by computational design. *Biochemistry* **41**: 3262-3269.
- Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, and Haak JR (1984) Molecular-dynamics with coupling to an external bath. *J Chem Phys* **81**: 3684-3690.
- Bresler MM, Rosser SJ, Basran A, and Bruce NC (2000) Gene cloning and nucleotide sequencing and properties of a cocaine esterase from *Rhodococcus* sp. strain MB1. *Applied & Environmental Microbiology* **66**: 904-908.
- Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister C, Caldwell JW, Ross WS, and Kollman PA (2004) *Amber 8*, University of California, San Francisco.
- Cooper, Z. D., Narasimhan, D., Sunahara, R. K., Mierzejewski, P., Jutkiewicz, E. M., Larsen, N A., Wilson, I. A., Landry, D. W. & Woods, J. H. (2006) Rapid and Robust Protection against Cocaine-Induced Lethality in Rats by the Bacterial Cocaine Esterase. *Mol Pharmacol* **70**: 1885-1891.
- Corchado JC, Coitino EL, Chuang YY, Fast PL, and Truhlar DG (1998) Interpolated vibrational transition-state theory by mapping. *J Phys Chem A* **102**: 2424-2438.
- Dahiyat BI (1999) *In silico* design for protein stabilization. *Curr Opin Biotechnol* **10**: 387-390.
- Dantas G, Kuhlman B, Callender D, Wong M, and Baker DA (2003) Large scale test of computational protein design: folding and stability of nine completely redesigned globular proteins. *J Mol Biol* **332**: 449-460.
- Darden TA, Lee H, and Pedersen LG (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J Chem Phys* **98**: 10089-10092.
- Dwyer MA, Looger LL, and Hellinga HW (2004) Computational design of a biologically active enzyme. *Science* **304**: 1967-1971.
- Gilson MK, Sharp KA, and Honig BH (1988) Calculating the Electrostatic Potential of Molecules in Solution: Method and Error Assessment. *J Comput Chem* **9**: 327-335.
- Gorelick DA (1997) Enhancing cocaine metabolism with butyrylcholinesterase as a treatment strategy. *Drug Alcohol Depend* **48**: 159-165.

Harkness JE and Wagner JE (1989) *The Biology and Medicine of Rabbits and Rodents*. 3rd edition, Lea and Febiger, Philadelphia.

Ko MC, Bowen LD, Narasimhan D, Berlin AA, Lukacs NW, Sunahara RK, Cooper ZD, and Woods JH (2007) Cocaine Esterase: Interactions with Cocaine and Immune Responses in Mice. *J Pharmacol & Exp Therap* **320**: 926-933.

Korkegian A, Black ME, Baker D, and Stoddard BL (2005) Computational thermostabilization of an enzyme. *Science* **308**: 857-860.

Kuhlman B and Baker D (2000) Native protein sequences are close to optimal for their structures. *Proc Natl Acad Sci USA* **97**: 10383-10388.

Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, and Baker D (2003) Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**: 1364-1368.

Landry DW, Zhao K, Yang GXQ, Glickman M, and Georgiadis TM (1993) Antibody-catalyzed degradation of cocaine. *Science* **259**: 1899-1901.

Larsen NA, Turner JM, Stevens J, Rosser SJ, Basran A, Lerner RA, Bruce NC, and Wilson IA (2002) Crystal structure of a bacterial cocaine esterase. *Nature Struct Biol* **9**: 17-21.

Lehmann M and Wyss M (2001) Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Current Opinion in Biotechnology* **12**: 371-375.

Looger LL, Dwyer MA, Smith JJ, and Hellinga HW (2003) Computational design of receptor and sensor proteins with novel functions. *Nature* **423**: 185-190.

Luo P, Hayes RJ, Chan C, Stark DM, Hwang MY, Jacinto JM, Juvvadi P, Chung HS, Kundu A, Ary ML, and Dahiyat BI (2002) Development of a cytokine analog with enhanced stability using computational ultrahigh throughput screening. *Protein Sci* **11**: 1218-1226.

Malakauskas SM and Mayo SL (1998) Design, structure and stability of a hyperthermophilic protein variant. *Nature Struct Biol* **5**: 470-475.

Meijler MM, Kaufmann GF, Qi LW, Mee JM, Coyle AR, Moss JA, Wirsching P, Matsushita M, and Janda KD (2005) Fluorescent cocaine probes: a tool for the selection and engineering of therapeutic antibodies. *J Am Chem Soc* **127**: 2477-2484.

Mendelson JH and Mello NK (1996) Management of cocaine abuse and dependence. *New Engl J Med* **334**: 965-972.

MOL #49486

Mets B, Winger G, Cabrera C, Seo S, Jamdar S, Yang G, Zhao K, Briscoe RJ, Almonte R, Woods JH, and Landry DW (1998) A catalytic antibody against cocaine prevents cocaine's reinforcing and toxic effects in rats. *Proc Natl Acad Sci USA* **95**: 10176-1081.

Pan Y, Gao D, Yang W, Cho H, Yang GF, Tai HH, and Zhan CG (2005) Computational redesign of human butyrylcholinesterase for anticocaine medication. *Proc Natl Acad Sci USA* **102**: 16656-16661.

Pan Y, Gao D, and Zhan CG (2007) Free energy perturbation (FEP) simulation on the transition-states of cocaine hydrolysis catalyzed by human butyrylcholinesterase and its mutants. *J Am Chem Soc* **129**: 13537-13543.

Paula S, Tabet MR, Farr CD, Norman AB, and Ball Jr WJ (2004) Three-dimensional quantitative structure-activity relationship modeling of cocaine binding by a novel human monoclonal antibody. *J Med Chem* **47**: 133-142.

Perczel A, Park K, and Fasman GD (1992) Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide (1992) *Anal Biochem* **203**: 83-93.

Redish AD (2004) Addiction as a computational process gone awry. *Science* **306**: 1944-1947.

Reina J, Lacroix E, Hobson SD, Fernandez-Ballester G, Rybin V, Schwab MS, Serrano L, and Gonzalez, C. Computer-aided design of a PDZ domain to recognize new target sequences. *Nature Struct Biol* **9**: 621-627.

Rogers CJ, Mee JM, Kaufmann GF, Dickerson TJ, and Janda KD (2005) Toward cocaine esterase therapeutics. *J Am Chem Soc* **127**: 10016-10017.

Ryckaert JP, Ciccotti G, and Berendsen HJC (1977) Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. *J Comput Phys* **23**: 327-341.

Shifman JM, and Mayo SL (2002) Modulating calmodulin binding specificity through computational protein design. *J Mol Biol* **323**: 417-423.

Singh S (2000) Chemistry, design, and structure-activity relationship of cocaine antagonists. *Chem Rev* **100**: 925-1024.

MOL #49486

Turner JM, Larsen NA, Basran A, Barbas CF, Bruce C, Wilson IA, and Lerner RA. (2002) Biochemical characterization and structural analysis of a highly proficient cocaine esterase. *Biochemistry* **41**: 12297-12307.

Zhan CG, Zheng F, and Landry DW (2003) Fundamental reaction mechanism for cocaine metabolism in human butyrylcholinesterase. *J Am Chem Soc* **125**: 2462-2474.

Zheng F, Yang W, Ko MC, Liu J, Cho H, Gao D, Tong M, Tai HH, Woods JH, and Zhan CG (2008) Most efficient cocaine hydrolase designed by virtual screening of transition states. *J Am Chem Soc* **130**: 12148-12155.

MOL #49486

Footnotes

This research was supported by NIH grants DA021416, DA025100, DA013930, and GM007767.

Persons to receive reprint requests:

Donald W. Landry, M.D. & Ph.D., Professor, Division of Clinical Pharmacology and Experimental Therapeutics, College of Physicians & Surgeons, Columbia University, 630 West 168th Street, Box 84, New York, NY 10032. E-mail: dwl1@columbia.edu.

James H. Woods, Ph.D., Professor, Department of Pharmacology, University of Michigan Medical School, 1301 Medical Sciences Research Building III, Ann Arbor, MI 48109. E-mail: jhwoods@umich.edu.

Roger K. Sunahara, Ph.D., Associate Professor, Department of Pharmacology, University of Michigan Medical School, 1301 Medical Sciences Research Building III, Ann Arbor, MI 48109. E-mail: sunahara@umich.edu.

Chang-Guo Zhan (to whom galley proofs should be sent), Ph.D., Professor, Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 725 Rose Street, Lexington, KY 40536. E-mail: zhan@uky.edu.

¹ These authors contributed equally to this work.

MOL #49486

Legends for figures

Figure 1. (A) Time-dependence of the root-mean-square deviation (in Å) of the MD-simulated atomic positions of CocE from those in the initial structure built from the X-ray crystal structure. The temperature used in the MD simulation was 298 K for first 250 ps and 575 K for the remaining 1050 ps (including temperature increase for 50 ps). Domain II refers to the plot for the atoms of residues G165 to I195 in domain II. (B) CocE(-)-cocaine binding structure modeled by molecular docking and MD simulation at 298 K. (C) Optimized structure of T172R/G173Q mutant of CocE. The hydrogen bond between Q173 domain II and P43 of domain I is indicated by a dash line.

Figure 2. (A) SDS-PAGE of purified CocE-wt, T172R, G173Q and T172R/G173Q. (B) Decay in activity at 37°C. 50 ng/ml wt-CocE and the mutants were incubated at 37°C and enzymatic activity measured over time. The half-lives for each mutant, $T_{1/2}$ were determined by analyzing the catalytic efficiency at different preincubation times. Data were fitted using a signal phase exponential decay (Prism 5, GraphPad Software, La Jolla, CA). Native CocE (black), T172R(red), G173Q (blue) and T172R/G173Q (green) showed 11.2 +/-0.9 min, 78+/-6.5 min, 75+/-9.9 and 305+/-38 min half-lives ($\tau_{1/2}$), respectively.

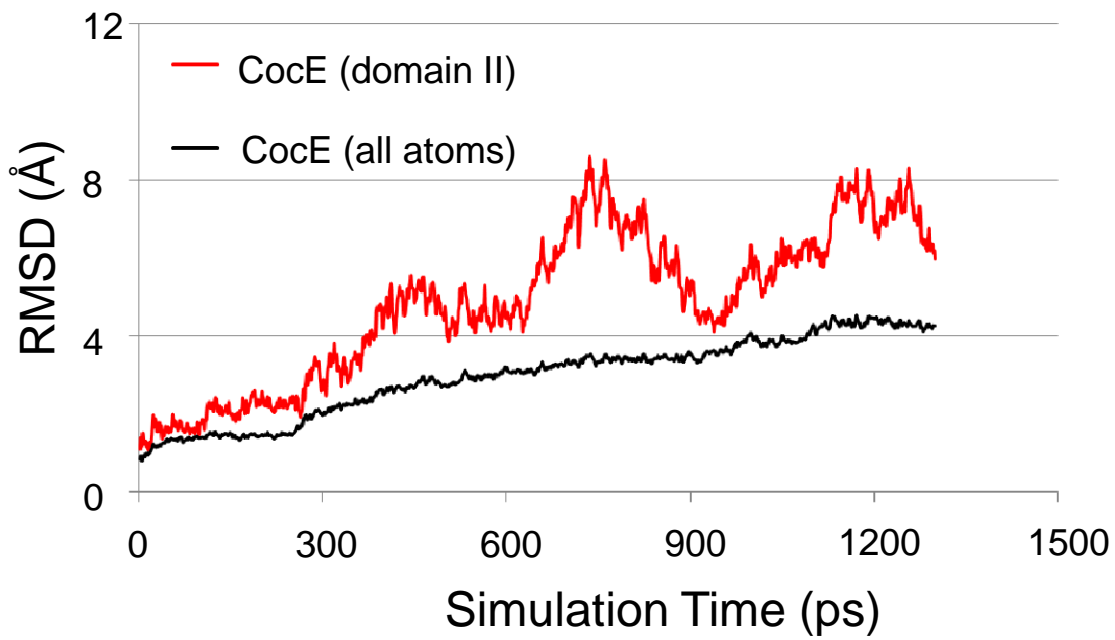
Figure 3. Data from circular dichroism measurements: (A) Native CocE; (B) T172R mutant; (C) G173Q mutant; (D) T172R/G173Q mutant. The temperatures of each scan for each mutant are listed in color-coded form in the figure legend.

Figure 4. (A) Prolonged protection against cocaine toxicity. The time following CocE administration (1 mg i.v.) pretreatment at which the enzyme protects 50% of the animals from a lethal cocaine administration (180 mg/kg, i.p., n=8 for each dosing condition). The *in vivo* half-lives ($\tau_{1/2}$) for CocE-wt (open box), T172R (hatched) and T172R/G173Q (black) are compared to the $\tau_{1/2}$ values determined by the *in vitro* measurements of cocaine hydrolysis at 37°C from Figure 2. $T_{1/2}$ for the *in vivo* data was determined by analyzing the lethality time-course data. Data were fitted using a signal phase exponential decay (Prism 5, GraphPad Software, La Jolla, CA) and the times which leads to 50% lethality were interpolated. (B) Western blot analysis of mice plasma collected 1 and 2 hour post-treatment with and CocE-wt and the T172R/G173Q mutant (1 mg i.v.). 10 µg of plasma proteins were loaded onto a 10% SDS-PAGE gel and transferred onto a nitrocellulose membrane. Membranes were then probed with anti-CocE rabbit

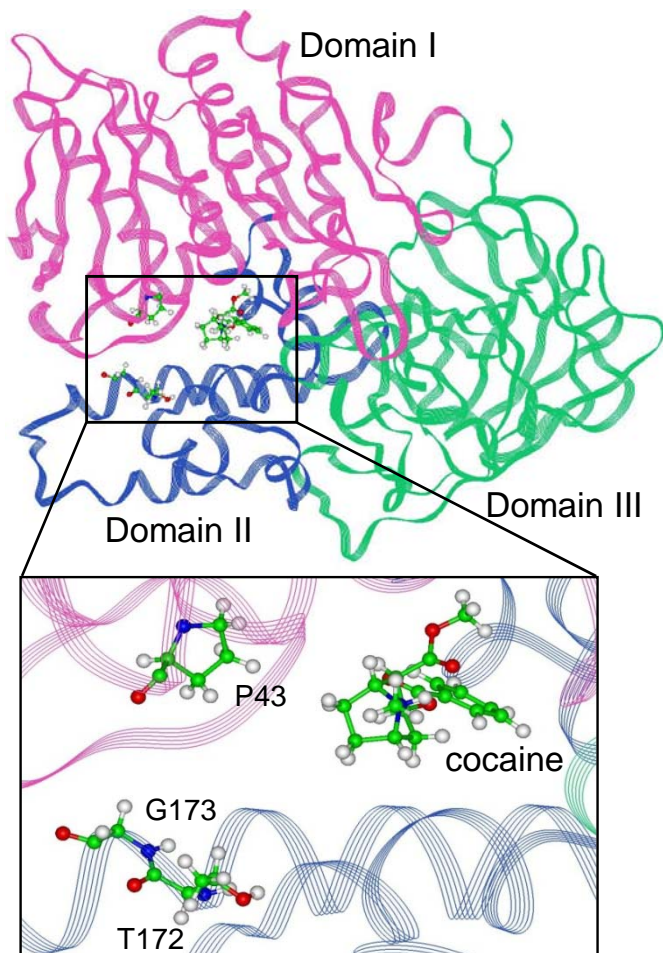
MOL #49486

antibody followed by HRP-linked anti-rabbit secondary antibody. As a loading control, blots were stripped and then reprobed with mouse anti apolipoprotein-A1 antibody.

(A)



(B)



(C)

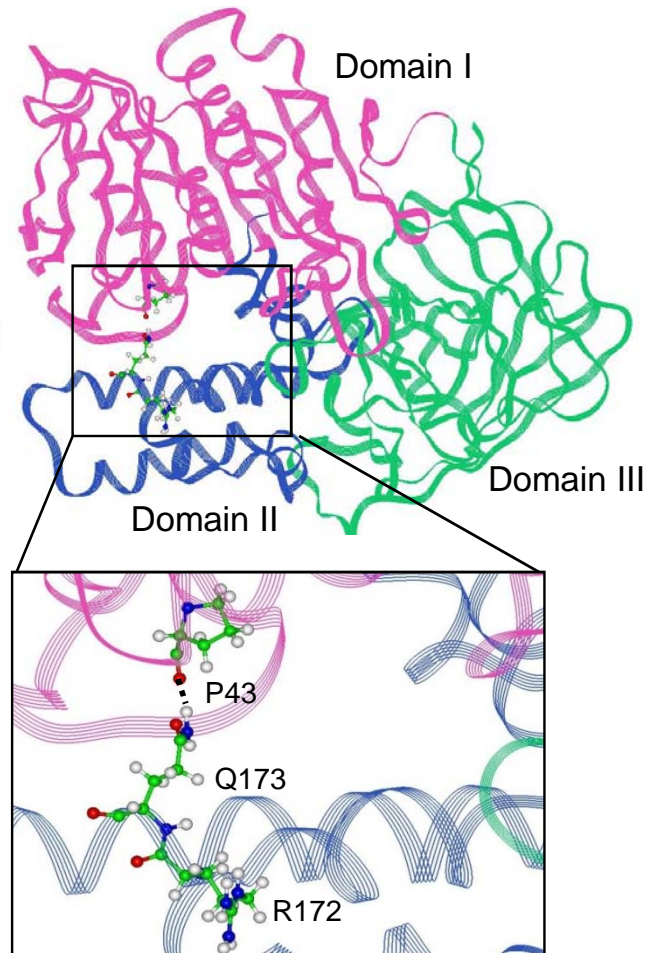


Figure 1

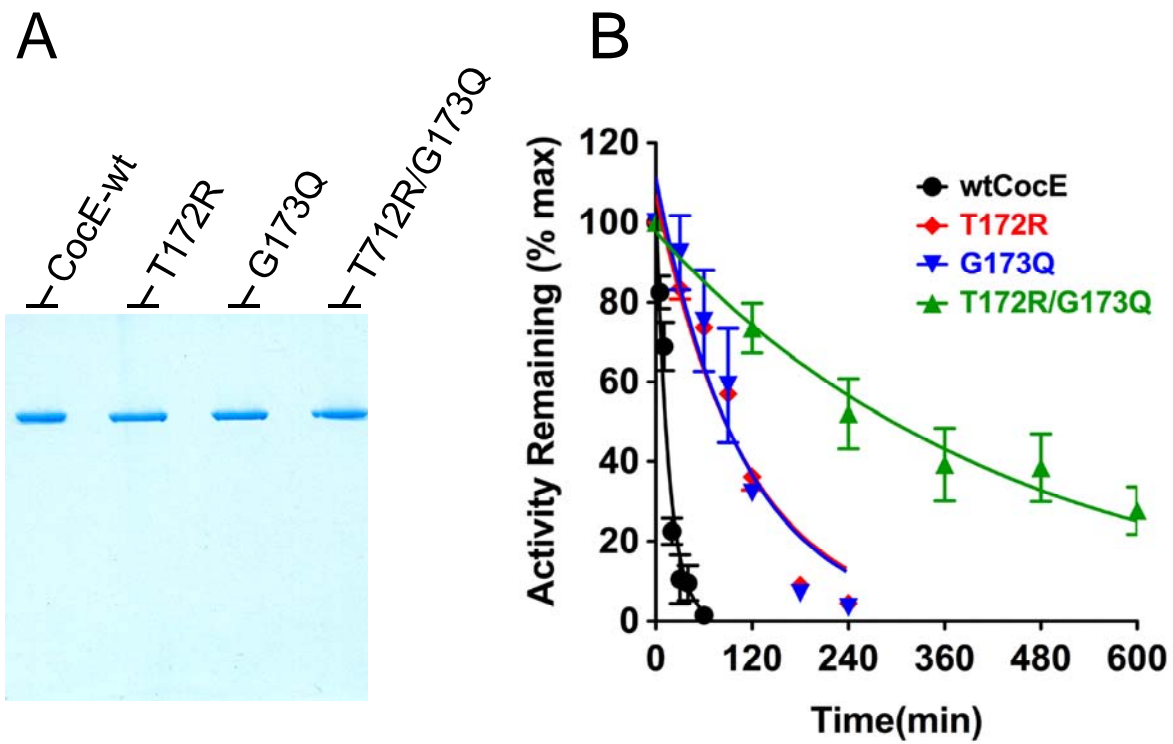


Figure 2

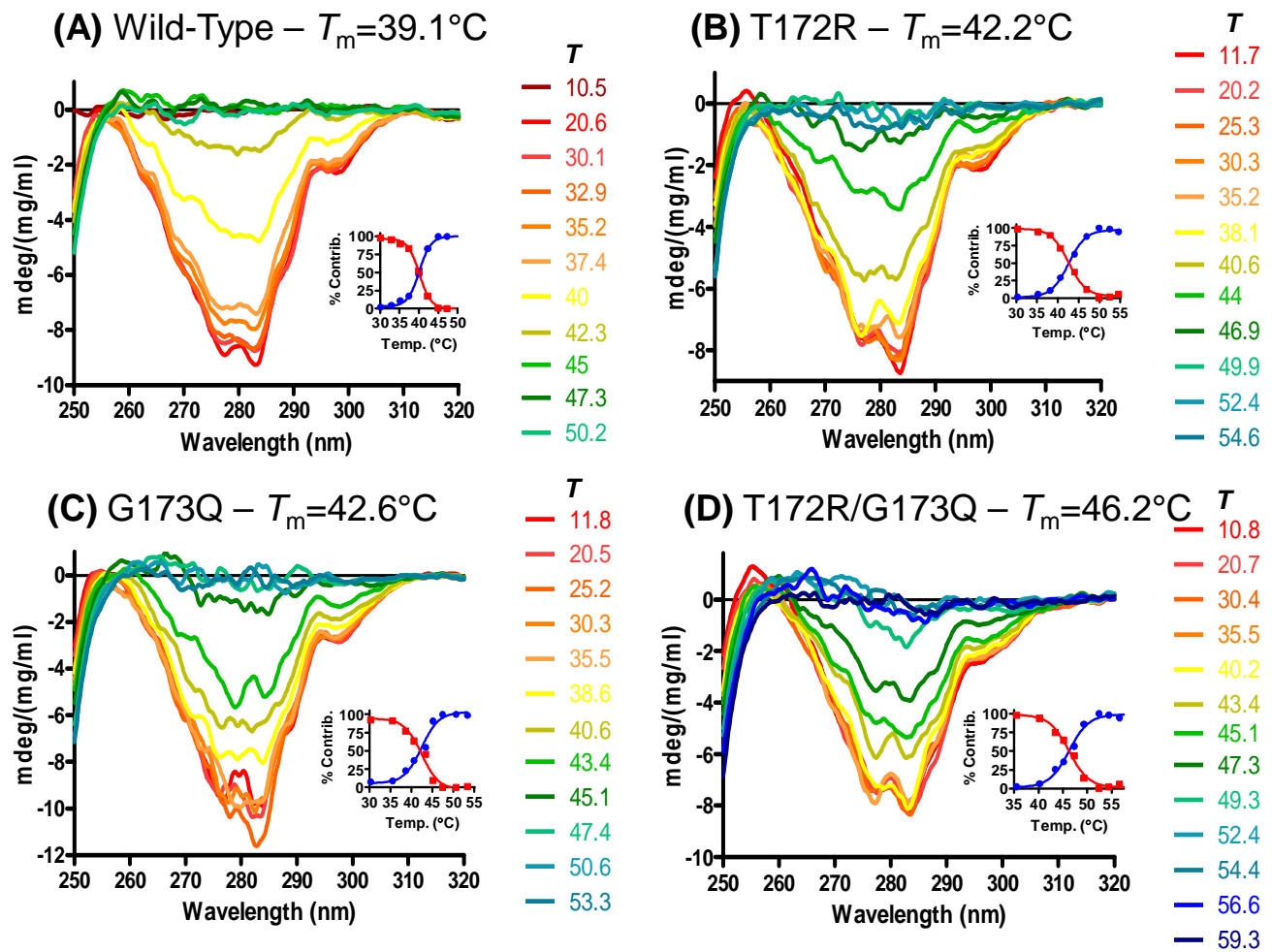


Figure 3

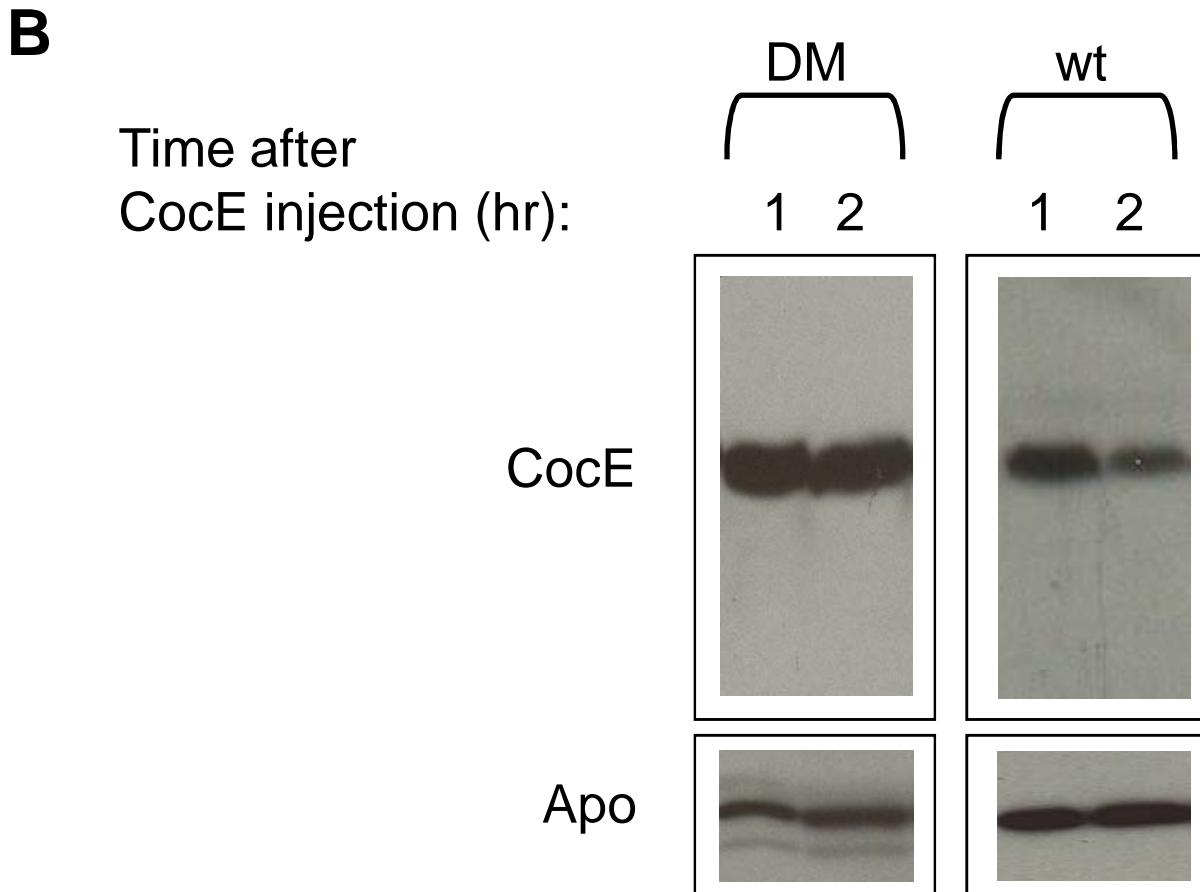
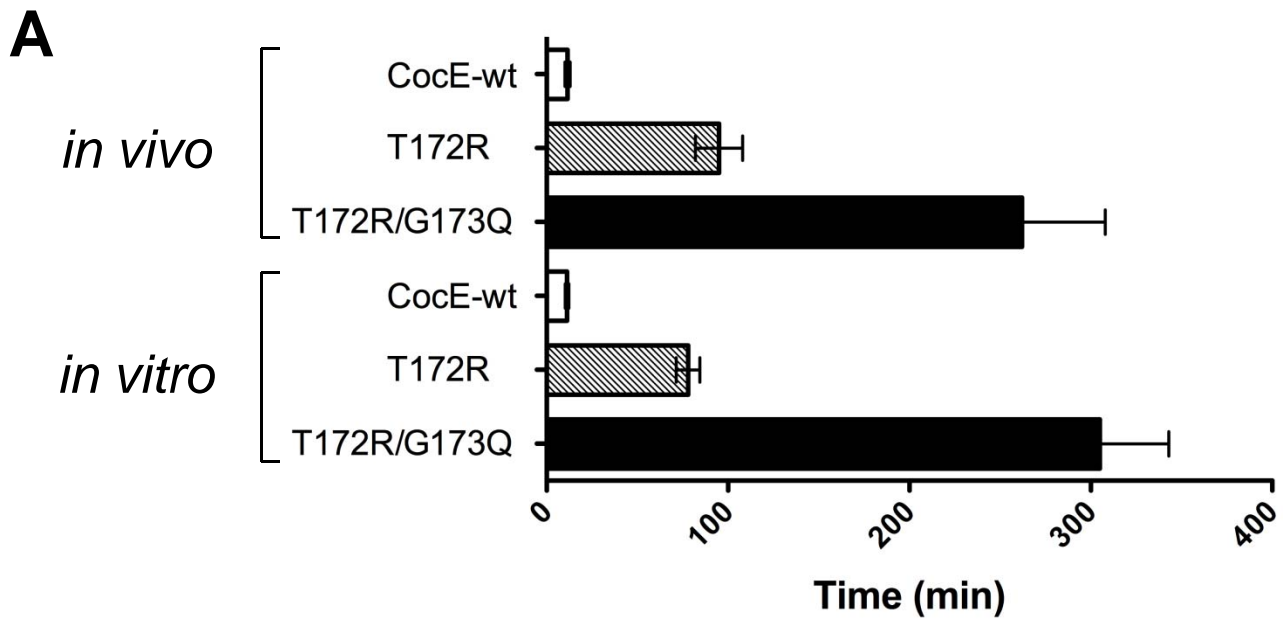


Figure 4