SUPPLEMENTAL INFORMATION

Paired Carboxylic Acids in Enzymes and Their Role in Selective Substrate Binding, Catalysis, and Unusually Shifted pK_a Values

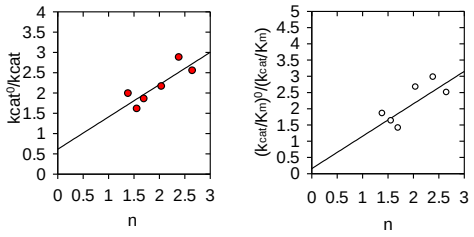
Ilja V. Khavruttski[†], *Jaimee R. Compton*[‡], *Kayla M. Jurkouich*[°], *Patricia M. Legler*[‡*]

[†]Uniformed Services University, Armed Forces Radiobiology Research Institute, Bethesda, MD 20889-5648; [°]Case Western Reserve University, Dept. of Biomedical Engineering, Cleveland, OH 44106; [‡]U.S. Naval Research Laboratories, 4555 Overlook Ave., Washington, DC 20375.

Substrate Stickiness. Solvent viscosity was varied using glycerol. The relative viscosity ηrel was calculated from measured values (1). The k_{cat}^{o} and $(k_{cat}/K_m)^{o}$ are the values in the absence of viscogen. The k_{cat}^{o}/k_{cat} and $(k_{cat}/K_m)^{o}/(k_{cat}/K_m)$ were plotted against ηrel and produced slopes near 1.0 (2).

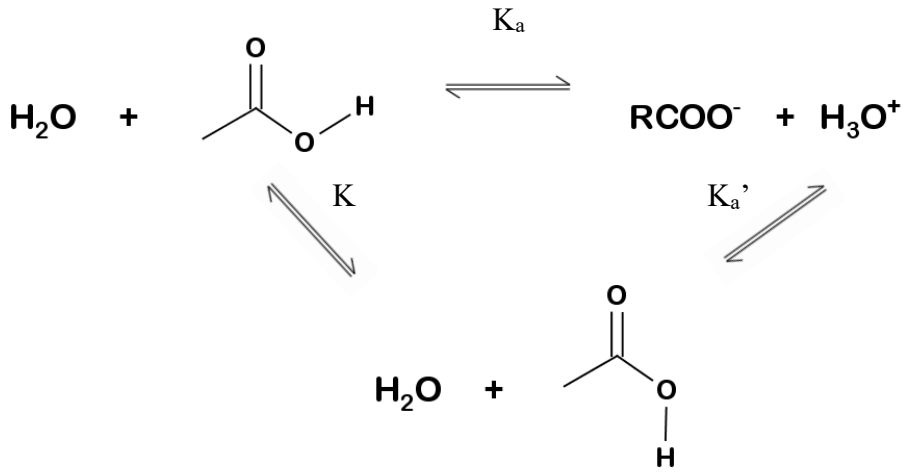
$$1/(k_{cat}/K_m) = \eta_{rel}/k_1 + (k_2/k_3)/k_1$$

S_r was calculated by fitting the viscosity data to the above equation (3). $S_r = k_3/k_2$ and is the stickiness ratio. From our data we calculated a value for $S_r = 4.3$. The relationship between K_m and K_d for a sticky substrate is in Werner, et al. (4) and Gadda, et al (5). The correction to the apparent pK_a values can be calculated from S_r (Cook and Cleland, page 338 (6)).

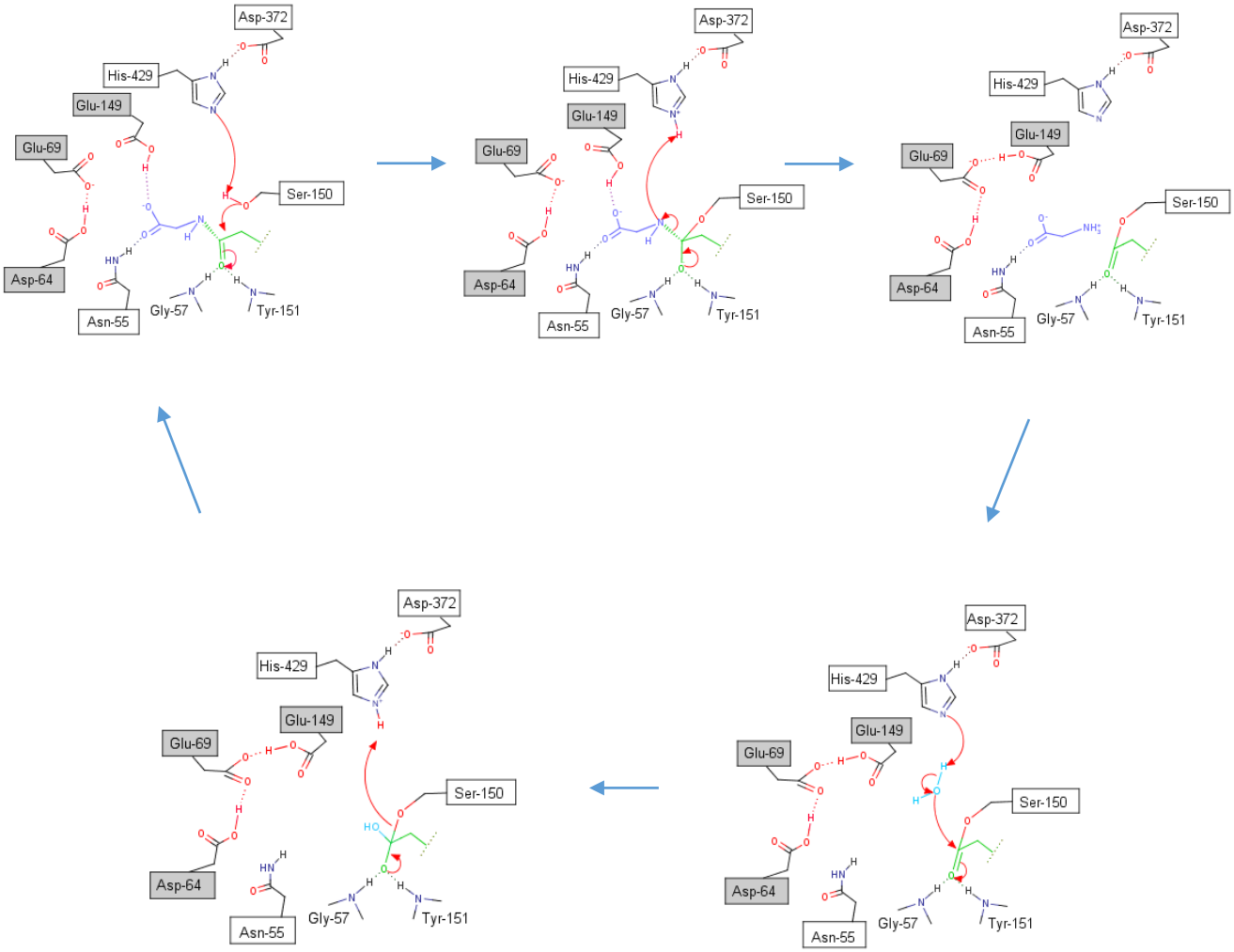


(1) Segur, J.B., Oberstar, H.E. Viscosity of Glycerol and Its Aqueous Solutions. *Industrial and engineering Chemistry* (1949) vol 43(9): 2117-2120.
(2) Adams, J.A., Taylor, S.S. Phosphorylation of peptide substrates for the catalytic subunit of cAMP-dependent protein kinase. *JBC* (1993) vol 268(11):7747-7752.
(3) Jacques, S.L., Kuliopulos, A. Protease-activated receptor-4 usees dual prolines and an anionic retention motif for thrombin recognition and cleavage. *Biochem. J.* (2003) 376: 733-740.
(4) Werner, D.S., Lee, T.R., Lawrence, D.S. Is protein kinase substrate efficacy a reliable barometer for successful inhibitor design? *JBC* (1996) 271(1):180-185.
(5) Gadda, G., Sobrado, P. Kinetic solvent viscosity effects as probes to studying mechanism of enzyme action. *Biochemistry* (2018) 57:3445-3453.
(6) Cook, P. F., and Cleland, W. W. (2007) *Enzyme kinetics and mechanism*, Garland Science, London ; New York.

Suppl. Scheme S1. Energetic calculations suggest that the *syn* and *anti* conformers of a carboxylate –OH (relative to the C=O) will differ in their basicity; the K_a' is estimated to be ~10^4-fold larger than the K_a. Thus, a carboxylate is a weak base when constrained to accept a proton in the *anti* direction, and protonation is less favorable in the *anti* direction than in the *syn*. The most stable conformation for a carboxyl is *syn.*



Scheme S2. Proposed mechanism at neutral pH and pH > 5.8



Scheme S3. Proposed mechanism at pH < 5.8

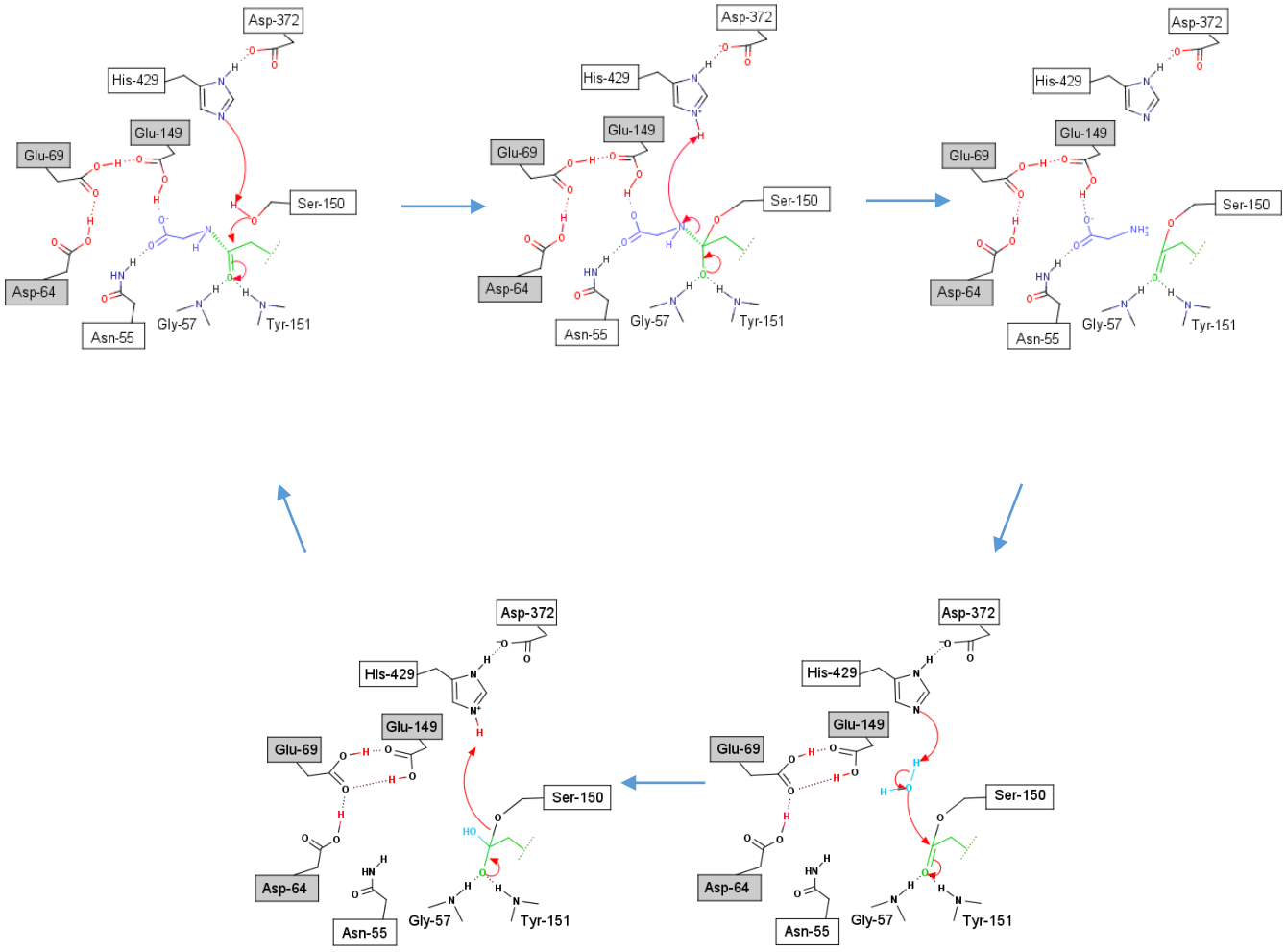


Figure S1. Similarity of human CatA (slate blue, PDB 1IVY) to *S. cerevisiae* Kex1DP (dark blue, PDB 1AC5), *T. californica* AChE (cyan, PDB 2ACE), *S. cerevisiae* SFGH (white, PDB 3C6B), human EsD (pink, PDB 3FCX), and *B. subtilis* RsbQ (purple, PDB 1WOM). The catalytic triad residues (lime green) have been superposed and all of the structures are shown from the same angle. SFGH, hEsD, and RsbQ also belong to the alpha-beta hydrolase superfamily. SFGH, hEsD, and RsbQ contain a pair of His residues. Whereas CatA and Kex1P contain a pair of Glu residues.

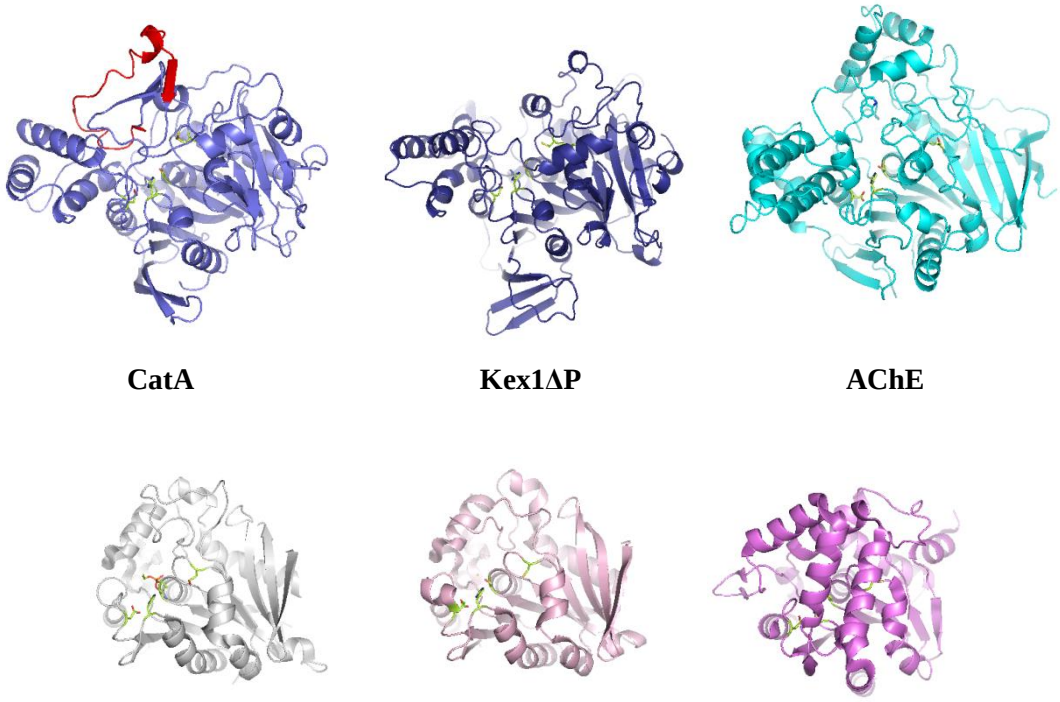
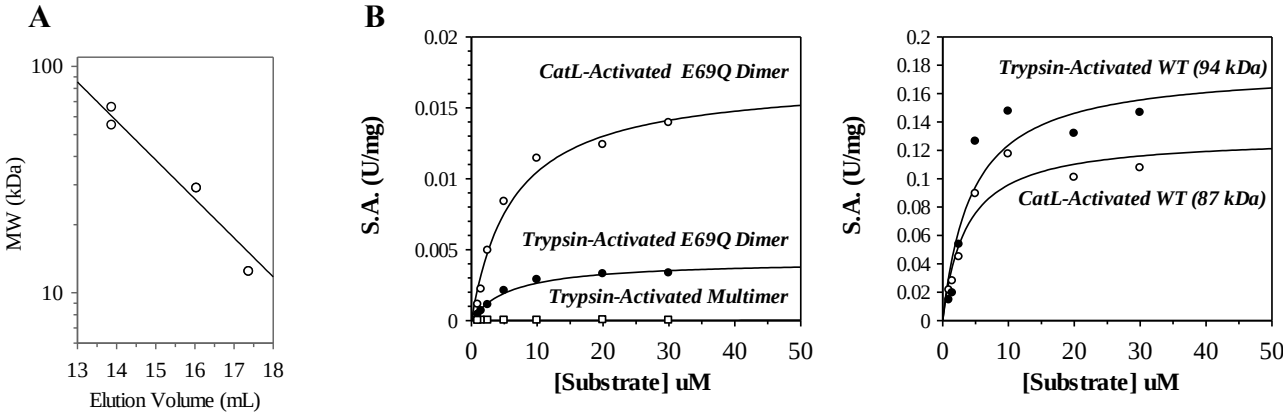


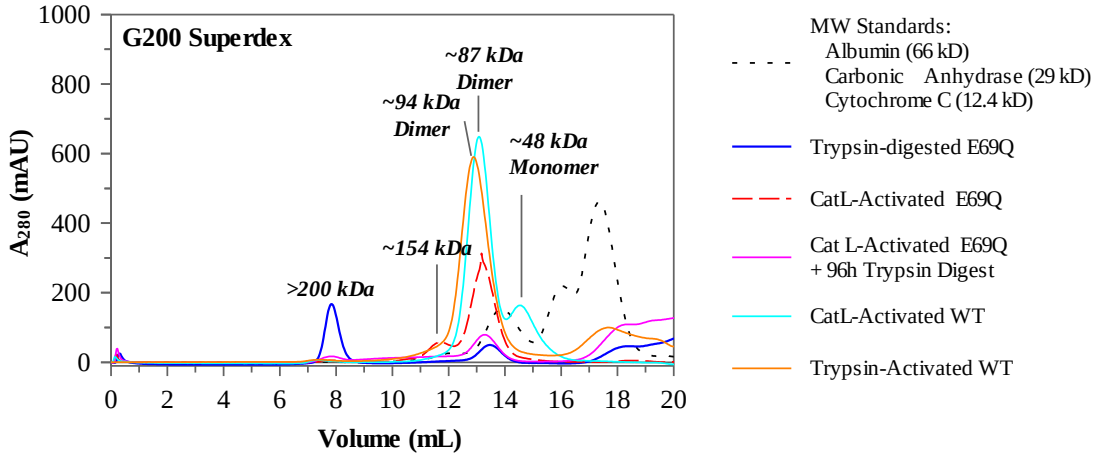
Figure S2. ***Cathepsin A Expression Construct used in this study.*** The natural signal peptide sequence is shown in red, the signal peptide is cut and the protein is secreted; the cleavage site of the signal peptide was previously confirmed by N-terminal sequencing (*1*). The blocking peptide (residues 272–277, yellow) and the excision peptide (residues 285–293, green) are highlighted. The protein is 54 kDa after glycosylation (homodimeric). The 1.6 kDa excision peptide was predicted by Bonten (green) its removal is thought to produce the 30 and 19 kDa subunits (*2*). After cleavage a conformational change is thought to occur in the blocking peptide. Larger fragments may be removed by trypsin-like enzymes (residues 266-292 and 263-292, (*3*)). Ser150-His429-Asp372 forms the catalytic triad (residues underlined in red). The N-terminal sequences, APDQDEVQRL and LDPPCTNTTA (underlined), were reported for the related bovine spleen cathepsin A (*4*).

"MTSSPRAPPGEQGRGGAEMIRAAPPPLFLLLLLLLLLVSWASRG
EAAPDQDEIQRLPGLAKQPSFRQYSGYLKGSGSKHLHYWFVESQKDPENSPVVLWLNG
GPGCSSLDGLLTEHGPFLVQPDGVTLEYNPYSWNLIANVLYLESPAGVGFSYSDDKFY
ATNDTEVAQSNFEALQDFFRLFPEYKNNKLFLTGESYAGIYIPTLAVLVMQDPSMNLQ
GLAVGNGLSSYEQNDNSLVYFAYYHGLLGNRLWSSLQTHCCSQNKCNFYDNKDLECVT
NLQEVARIVGNSGLNIYNLYAPCAGGVPSHFRYEKDTVVVQDLGNIFTRLPLKRMWHQ
ALLRSGDKVRMDPPCTNTTAASTYLNNPYVRKALNIPEQLPQWDMCNFLVNLQYRRLY
RSMNSQYLKLLSSQKYQILLYNGDVDMACNFMGDEWFVDSLNQKMEVQRRPWLVKYGD
SGEQIAGFVKEFSHIAFLTIKGAGHMVPTDKPLAAFTMFSRFLNKQPYHHHHHH-"

(1) Jackman, H. L., Tan, F. L., Tamei, H., Beurling-Harbury, C., Li, X. Y., Skidgel, R. A., and Erdos, E. G. (1990) A peptidase in human platelets that deamidates tachykinins. Probable identity with the lysosomal "protective protein", *J Biol Chem 265*, 11265-11272.
(2) Rudenko, G., Bonten, E., d'Azzo, A., and Hol, W. G. (1995) Three-dimensional structure of the human 'protective protein': structure of the precursor form suggests a complex activation mechanism, *Structure 3*, 1249-1259.
(3) Kolli, N., and Garman, S. C. (2014) Proteolytic activation of human cathepsin A, *J Biol Chem 289*, 11592-11600.
(4) Matsuzaki, H., Ueno, H., Hayashi, R., and Liao, T. H. (1998) Bovine spleen cathepsin A: characterization and comparison with the protective protein, *J Biochem 123*, 701-706.

Figure S3. ***Quaternary structure determination using gel-filtration***. (A) The G200 Superdex column was calibrated using protein standards (66 kDa bovine serum albumin, 29 kDa bovine carbonic anhydrase, and 12.3 kDa horse heart cytochrome c) from Sigma-Aldrich. The albumin dimer elutes first (12.05 mL), followed by the monomer (13.9 mL), carbonic anhydrase (16.05 mL) and cytochrome c (17.4 mL). (B) The specific activity was measured for the WT & E69Q variant for the corresponding fractions. (C) Summary of the gel-filtration sizing experiment. (D) Kinetic data of column fractions.





D

Enzyme	pH	V_{max} (U/mg)	K_m (μM)	k_{cat} (1/sec)
WT CatL-Activated	4.5	0.20 ± 0.02	7 ± 2	0.17 ± 0.02
WT CatL-Activated Dimer	4.5	0.13 ± 0.02	5 ± 1	0.11 ± 0.02
WT-Trypsin-Activated Dimer	4.5	0.18 ± 0.03	4 ± 2	0.16 ± 0.03
E69Q CatL-Activated	4.5	0.034 ± 0.002	2.5 ± 0.6	0.03 ± 0.002
E69Q CatL-Activated Dimer	4.5	0.017 ± 0.001	6 ± 1	0.015 ± 0.001
E69Q Trypsin-Activated Dimer	4.5	0.0042 ± 0.0003	6 ± 1	0.0037 ± 0.0003

Figure S4. ***In silico*** **mutagenesis and the calculated** ***pKa values.*** (A) Using the structures of CatA the pK_a were calculated by PROPKA (http://propka.org). Individual residues were mutated in each structure using Coot and pK_a's were calculated from the models. (B) pH Dependencies of $-\log(K_m)$.

A

	Calculated pK_a				
Cathepsin A	**D64**	**E69**	**E149**	**D372**	**H429**
				Triad	
WT (PDB 4MWS)	7.27	5.84	13.11	4.64	6.74
WT (PDB 4CI9)	6.77	5.84*	13.46*	4.40	6.74
D64N (PDB 4MWS)	Mutated	5.80	12.54	4.64	6.74
E69Q (PDB 4MWS)	6.15	Mutated	9.26	4.64	6.48
E149Q (PDB 4MWS)	6.23	8.5	Mutated	4.63	6.57
E69Q/E149Q (PDB 4MWS)	6.14	Mutated	Mutated	4.63	6.57
WT- 8a Inhibitor Bound (PDB 4AZ0)	8.71*	5.26*	14.08*	4.74	6.02
WT-15a Inhibitor Bound (PDB 4AZ3)	9.61*	5.70*	14.39*	3.49	4.51
WT- 1 Inhibitor Bound (PDB 4CIA)	8.01*	5.65*	13.97*	2.94	4.03
WT- 2 Inhibitor Bound (PDB 4CIB)	9.16*	5.65*	14.46*	3.02	6.17
D64N Inhibitor Bound (PDB 4CIA)	Mutated	5.71	13.14	2.94	4.03
E69Q Inhibitor Bound (PDB 4CIA)	6.28	Mutated	9.68	2.94	4.03
E149Q Inhibitor Bound (PDB 4CIA)	6.17	9.01	Mutated	2.94	4.00
E69Q/E149Q Inhibitor Bound (PDB 4CIA)	6.27	Mutated	Mutated	2.94	4.00
D64N/E69Q/E149Q Inhibitor Bound (PDB 4CIA)	Mutated	Mutated	Mutated	2.94	4.00

B

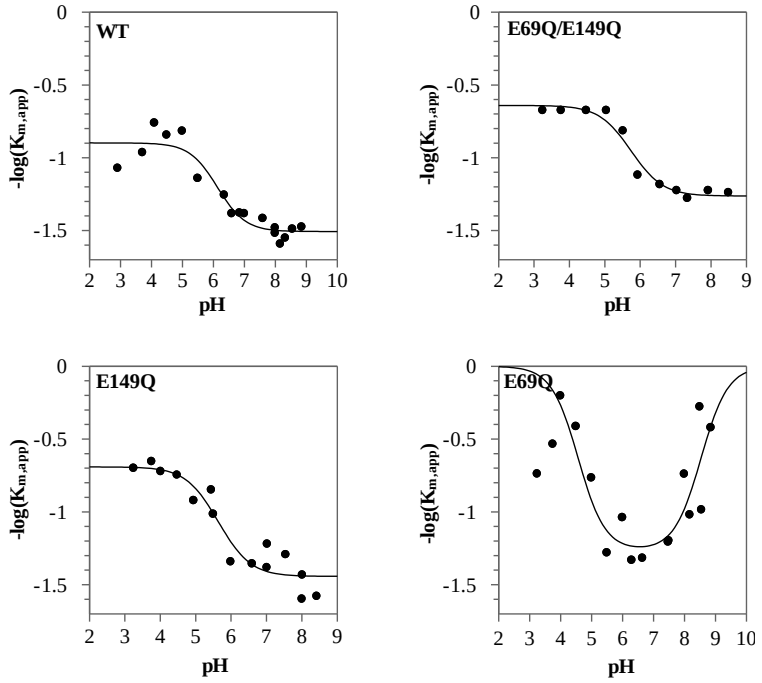
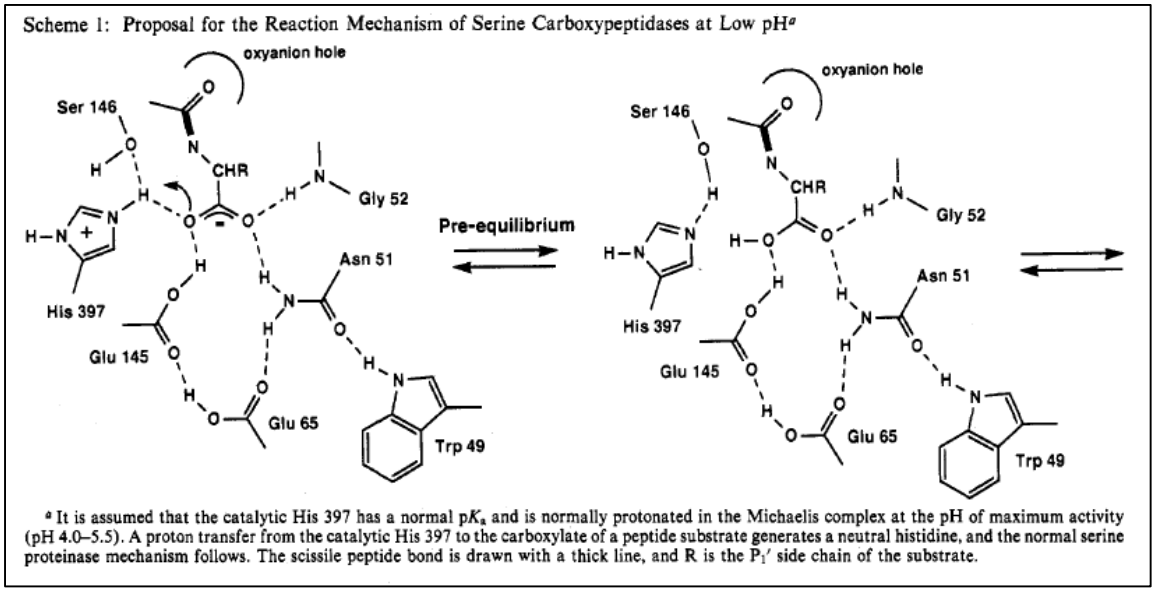
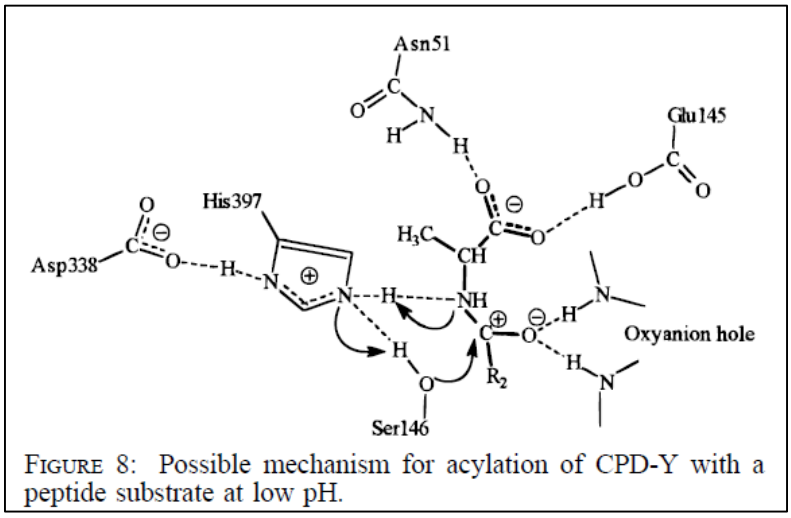


Figure S5. Mechanisms proposed for CPY that involve at least one of the conserved glutamate residues (Glu-145/Glu-65). (A) Mechanism proposed by Bullock et al. (B) Mechanism proposed by Stennicke et al. (C) Calculated pK_a from crystallized WT and E65A/E145A CPY.

A



B



C

CPY	Calculated pK_a			
	E65	**E145**	**D338**	**H397**
			Triad	
WT (PDB 1YSC)	6.15	11.76	5.89	6.39
E65A/E145A (PDB 1CPY)	Mutated	Mutated	5.62	6.44