BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

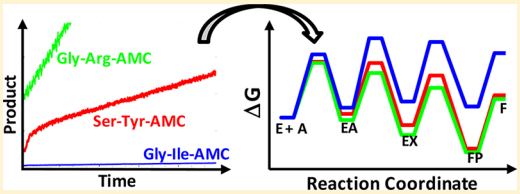
The Amino-Acid Substituents of Dipeptide Substrates of Cathepsin C Can Determine the Rate-Limiting Steps of Catalysis

Jon K. Rubach,[‡,#] Guanglei Cui,[§] Jessica L. Schneck,[‡] Amy N. Taylor,[‡] Baoguang Zhao,[§] Angela Smallwood,[§] Neysa Nevins,[§] David Wisnoski,[∥] Sara H. Thrall,[‡] and Thomas D. Meek*[,‡]

[‡]Department of Biological Reagents and Assay Development, [§]Department of Computational and Structural Chemistry, and [∥]Department of Screening and Compound Profiling, GlaxoSmithKline Pharmaceuticals, 1250 South Collegeville Road, Collegeville, Pennsylvania 19426, United States

Supporting Information

ABSTRACT: We examined the cathepsin C-catalyzed hydrolysis of dipeptide substrates of the form Yaa-Xaa-AMC, using steady-state and pre-steady-state kinetic methods. The substrates group into three kinetic profiles based upon the broad range observed for k_{cat}/K_a and k_{cat} values, pre-steady-state time courses, and solvent kinetic isotope effects (sKIEs). The dipeptide substrate Gly-Arg-AMC displayed large values for k_{cat}/K_a ($1.6 \pm 0.09\ \mu M^{-1}\ s^{-1}$) and k_{cat} ($255 \pm 6\ s^{-1}$), an inverse sKIE on k_{cat}/K_a ($^{D}(k_{cat}/K_a) = 0.6 \pm 0.15$), a modest, normal sKIE on k_{cat} ($^{D}k_{cat} = 1.6 \pm 0.2$), and immeasurable pre-steady-state kinetics, indicating an extremely fast pre-steady-state rate (>400 s^{-1}). (Errors on fitted values are omitted in the text for clarity but may be found in Table 2.) These results conformed to a kinetic model where the acylation (k_{ac}) and deacylation (k_{dac}) half-reactions are very fast and similar in value. The second substrate type, Gly-Tyr-AMC and Ser-Tyr-AMC, the latter the subject of a comprehensive kinetic study (Schneck et al. (2008) *Biochemistry 47*, 8697−8710), were found to be less active substrates compared to Gly-Arg-AMC, with respective k_{cat}/K_a values of $0.49 \pm 0.07\ \mu M^{-1}\ s^{-1}$ and $5.3 \pm 0.5\ \mu M^{-1}\ s^{-1}$, and k_{cat} values of $28 \pm 1\ s^{-1}$ and $25 \pm 0.5\ s^{-1}$. Solvent kinetic isotope effects for Ser-Tyr-AMC were found to be inverse for k_{cat}/K_a ($^{D}(k_{cat}/K_a) = 0.74 \pm 0.05$) and normal for k_{cat} ($^{D}k_{cat} = 2.3 \pm 0.1$) but unlike Gly-Arg-AMC, pre-steady-state kinetics of Gly-Tyr-AMC and Ser-Tyr-AMC were measurable and characterized by a single-exponential burst, with fast transient rates ($490\ s^{-1}$ and $390\ s^{-1}$, respectively), from which it was determined that $k_{ac} \gg k_{dac} \sim k_{cat}$. The third substrate type, Gly-Ile-AMC, gave very low values of k_{cat}/K_a ($0.0015 \pm 0.0001\ \mu M^{-1}\ s^{-1}$) and k_{cat} ($0.33 \pm 0.02\ s^{-1}$), no sKIEs, ($^{D}(k_{cat}/K_a) = 1.05 \pm 0.5$ and $^{D}k_{cat} = 1.06 \pm 0.4$), and pre-steady-state kinetics exhibited a discernible, but negligible, transient phase. For this third class of substrate, kinetic modeling was consistent with a mechanism in which $k_{dac} > k_{ac} \sim k_{cat}$, and for which an isotope-insensitive step in the acylation half-reaction is the slowest. The combined results of these studies suggested that the identity of the amino acid at the P_1 position of the substrate is the main determinant of catalysis. On the basis of these kinetic data, together with crystallographic studies of substrate analogues and molecular dynamics analysis with models of acyl-enzyme intermediates, we present a catalytic model derived from the relative rates of the acylation vs deacylation half-reactions of cathepsin C. The chemical steps of catalysis are proposed to be dependent upon the conformational freedom of the amino acid substituents for optimal alignment for thiolation (acylation) or hydrolysis (deacylation). These studies suggest ideas for inhibitor design for papain-family cysteine proteases and strategies to progress drug discovery for other classes of disease-relevant cysteine proteases.



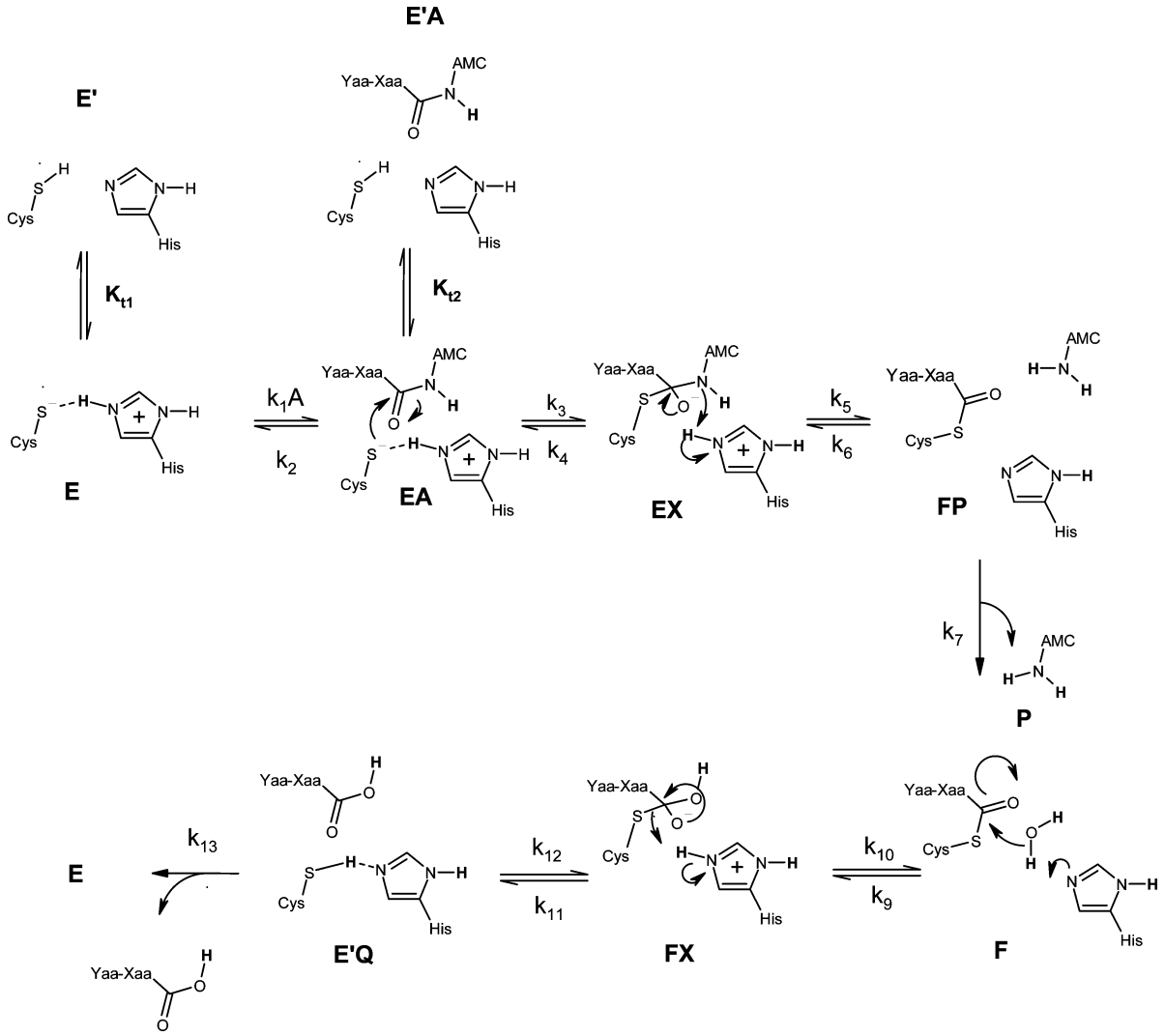
Cathepsin C (dipeptidyl peptidase I (DPP-I); E.C. 3.4.14.1), a member of the papain-family, is a lysosomal cysteine protease that sequentially removes dipeptides from the N-termini of protein and peptide substrates.[1,2] Cathepsin C activates a cascade of serine proteases, including neutrophil elastase, cathepsin G, proteinase 3, chymase, tryptase, and granzyme A and B, which elicit inflammatory responses.[3−5] Cleavage of the dipeptidyl pro-sequences of these proteases results in their activation, triggering degradation of the extracellular matrix which can result in tissue damage and chronic inflammation. Upstream inhibition of this protease cascade, by blockage of cathepsin C, could potentially provide useful therapeutics for the treatment of neutrophil-elicited inflammatory diseases such as chronic obstructive pulmonary disease (COPD).[6] Two recent studies have indicated that a high degree of inhibition or inactivation of cathepsin C activity in cellular[7] and in vivo[8] assays is required to silence its ability to process its pro-enzyme substrates.

Cathepsin C operates via a double-displacement mechanism comprised of acylation and deacylation half-reactions (Scheme 1). Previously, we completed a detailed kinetic analysis of the cathepsin C-catalyzed peptidolysis of Ser-Tyr-AMC,[1] which indicated that catalysis proceeds through rapid acylation, followed by slower, rate-limiting deacylation.[1] This study, however, was limited to a single substrate which might not fully report on the kinetics of other physiologically-relevant

Received: May 31, 2012
Revised: August 17, 2012
Published: August 28, 2012

   dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

Scheme 1



substrates. The N-terminal amino acids of known physiological substrates of cathepsin C are Gly-Glu (Cathepsin G, mast cell chymase and Granzyme B), Ser-Glu (neutrophil elastase), and Glu-Lys (Granzyme A).[2] The enzyme is generally considered to be nonspecific and will progressively cleave off many different dipeptide groups from protein and peptidyl substrates until either the N-terminus is no longer available or an uncleavable ("stop") sequence has been reached. Cathepsin C stop sequences are positively charged residues (arginine or lysine) at the N-terminus (P_2)[a] or proline at either side of the scissile bond of a substrate (P_1 and $P_{1'}$). Although no detailed kinetic analysis with multiple substrates has been reported, studies of cathepsin C-catalyzed hydrolysis using a variety of dipeptide-AMC substrates indicated that the enzyme has a wide substrate specificity, although it requires substrates to have a free N-terminal amine. Substrates bearing lysine or arginine residues at the P_1 position are preferred, while substrates containing a proline or isoleucine residue at the P_1 position are either inactive or have very low activity.[2]

The active site of cathepsin C, as with other papain-like proteases, is found in an accessible, shallow cleft which harbors the eponymous residue Cys-234 and its catalytic partner, His-381.[10] However, it differs from other cysteine proteases in that it is homotetrameric[11] and possesses an additional exclusion domain which occludes the substrate-binding sites beyond the S_2 subsite, ablating endopeptidase activity.[11,12] This exclusion domain donates the side chain of Asp-1 to form a putative ion-pair interaction between the β-carboxylate of Asp-1 and the N-terminal amine of the substrate,[12] hence the specificity for substrates with a free N-terminal amine. Structural data have shown that the S_2 subsite is a deep-binding pocket which putatively contributes most of the substrate-binding affinity.[10] The S_1 subsite of cathepsin C is a wide pocket thought to tolerate a wide range of amino acids.

The identification and characterization of substrate-specific parameters dictating the overall kinetics of cathepsin C catalysis could reveal routes toward designing small molecule inhibitors. Accordingly, we investigated the kinetic effects of varying the amino acid at the P_1 position of dipeptide substrates, in order to explore the contributions of the S_1 subsite to the kinetics of the acylation vs deacylation steps. We also characterized the substrate interactions of the residues at the P_1 position with the enzyme through crystallographic and molecular modeling studies. The crystallography studies employed nitrile-containing substrate-analogues to examine interactions at the active site, and molecular dynamics examined the effects of different amino acids binding at the S_1 subsite on the orientation of the water molecule involved in the deacylation half-reaction.

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

EXPERIMENTAL PROCEDURES

Materials. Deuterium oxide (D_2O; ≥99% ^2H), deuterium chloride (DCl), sodium deutroxide (NaOD) (>99% ^2H), sodium hydroxide, hydrochloric acid, CHAPS, DTT, triethanolamine, DMSO (97.97% pure), and EDTA were purchased from Sigma-Aldrich. The dipeptide-AMC substrates were purchased from MP Biomedicals (Solon, OH) or custom synthesized by SM Biochemicals (Anaheim, CA). All substrates had purity >95% as determined by LCMS. Ser-Tyr-nitrile was obtained externally.

Cloning, Expression, and Purification of Human Cathepsin C. Human cathepsin C was expressed and purified as previously described.[1] Briefly, prepro-cathepsin C was captured by affinity chromatography on immobilized Protein A and buffer-exchanged into sodium citrate buffer (pH 4.5) using chromatography on Sephadex G-25. Activation of cathepsin C was performed by proteolysis with recombinant cathepsin L. The resulting mature, active cathepsin C was purified by size exclusion chromatography using a Superdex 200 column equilibrated with sodium phosphate, EDTA buffer (pH 6). Purity was determined to be >95% by SDS-PAGE. Edman N-terminal sequencing confirmed three expected N-termini — DTPAN, ILHLP, and DPFNP, which correspond to the exclusion domain, heavy chain, and light chain, respectively.

Steady-State Kinetic Studies. Steady-state kinetic studies were performed in 10-μL reaction mixtures composed of 50 mM sodium acetate (pH 5.5), 30 mM NaCl, 1 mM EDTA, 1 mM CHAPS, 1 mM DTT (Buffer A), and 10% (v/v) DMSO. Peptidolysis was measured at 21 °C in 384-well Greiner low-volume polypropylene plates, and fluorescence of the amino-methyl coumarin (AMC) product was measured using an Acquest fluorimeter (Molecular Devices, Sunnyvale, CA; 360 nm excitation, 460 nm emission, 400 nm cutoff filter). Solutions of dipeptide substrates were prepared in 100% DMSO, and 0.8-μL aliquots were added to assay mixtures using a HummingBird liquid handler (Genomic Solutions, Ann Arbor, MI). Reactions were initiated by addition of 10-μL aliquots of enzyme (0.01−100 nM).

Active-Site Titration. The concentration of active sites of cathepsin C was determined by measuring enzyme activity with Ser-Tyr-AMC or Gly-Arg-AMC following incubation of the enzyme with the potent, irreversible inhibitor Ala-Hph-VS-Ph,[13] using the method of Schneck et al.[1]

pH-Rate Profiles and Solvent Kinetic Isotope Effects. The pH-rate profiles of k_{cat}/K_a (V/KE_t) and k_{cat} (V/E_t) for Ser-Tyr-AMC[b], Gly-Arg-AMC, and Gly-Ile-AMC were determined by measurement of initial rates over the pH range of 3.5−8.0, in increments of 0.3 pH units, in a mixed buffer consisting of 50 mM sodium acetate, 50 mM MES, 100 mM triethanolamine, 1 mM DTT, 1 mM EDTA, 1 mM CHAPS, at constant ionic strength I = 0.1.[14] pH values were adjusted with HCl or NaOH. Solvent kinetic isotope effects were likewise determined from pH- and pD-rate profiles using an identical buffer prepared in deuterium oxide (final deuterium content: 93%). Buffers with pD values in increments of 0.3 pD units were made by adjusting the pD with DCl or NaOD, from which pD values were determined as pD = pH + 0.4.[15]

Pre-Steady-State Kinetics. Transient kinetics experiments were performed using an Applied Photophysics SX20 stopped-flow spectrophotometer (Surrey, UK) with the reaction mixtures in Buffer A and 10% (v/v) DMSO at 21 °C. Time courses of AMC production were measured by the fluorescence generated using a 348-nm wavelength for excitation and a 400-nm cutoff filter for emission. Background fluorescence was measured in the absence of enzyme and was subtracted at each substrate concentration. The formation of AMC, as measured in millivolts, was converted to nanomolar concentrations of product formed following calibration of fluorescence with standards of AMC at known concentrations.[1]

Data Analysis. Steady-state kinetic data were fitted to appropriate rate equations using the nonlinear regression function of Grafit (Erithacus Software Ltd.) or SigmaPlot (Systat Software, Inc.). Initial rate data were fitted to eq 1 to determine the kinetic parameters k_{cat} and k_{cat}/K_a.

$$v = k_{cat}AE_t/(K_a + A) \quad (1)$$

For eq 1, v is the initial rate (nM/s), k_{cat} is the turnover number (s^{-1}), E_t is the concentration of active sites (nM), and K_a is the Michaelis constant for the substrate (A).

Data from pH-rate profiles for log k_{cat} or log k_{cat}/K_a which decreased at low and high pH or pD with respective apparent slopes of 1 and −1 were fitted to eq 2, and plots with apparent slopes of 1 and −2 were fitted to eq 3. For profiles of log k_{cat} or log k_{cat}/K_a which decreased at low and high pH or pD, with apparent slopes of less than 1 and −1 or −2, respectively, were fitted to eq 5 or 4, respectively. Eq 6 was used to fit plots of log k_{cat} vs pH and pD for Gly-Arg-AMC.

$$\log y = \log\{c/[1 + H/K_1 + K_2/H]\} \quad (2)$$

$$\log y = \log\{c/[1 + H/K_1 + (K_2/H)(1 + K_3/H)]\} \quad (3)$$

$$\log y = \log\{c(1 + H/K_a)/[(1 + H/K_b)(1 + H/K_1 + (K_2/H)(1 + K_3/H))]\} \quad (4)$$

$$\log y = \log\{c(1 + H/K_a)/[(1 + H/K_b)(1 + H/K_4 + K_5/H)]\} \quad (5)$$

$$\log y = \log\{c(1 + H/K_a)/[1 + H/K_b + K_c/H + H^2/K_d]\} \quad (6)$$

For eqs 2−6, y is the observed kinetic parameter, k_{cat} or k_{cat}/K_a, c is the pH- or pD-independent value of y, H is the measured value of $[H^+]$ or $[D^+]$, K_1-K_5 are apparent acid or basic dissociation constants for acidic and basic groups on the substrate or enzyme, and K_a-K_d are mechanism-dependent equilibrium constants which comprise the multiplicative products of acid or basic dissociation constants and kinetic rate constants (K_1-K_5 and K_a-K_d are nonidentical in eqs 2−6).

SKIEs were calculated from the pH(D)-independent values of k_{cat}/K_a or k_{cat}, obtained from plots of log k_{cat} or log k_{cat}/K_a vs pH or pD from data-fitting to eqs 2−6. Error propagation was determined as described in Schneck et al.[1]

Pre-steady-state kinetic data were fitted to eq 7 for each fixed concentration of dipeptide substrate, which describes time courses conforming to a single-exponential function for which [P] is the nanomolar concentration of AMC formed, k_{ss} is the apparent steady-state rate, t is time (s), β is the apparent burst amplitude, λ is the apparent rate of the exponential phase, and C is a constant of calibration.

$$[P] = k_{ss}E_t(t) + \beta E_t(1 - \exp(-\lambda t)) + C \quad (7)$$

Additionally, pre-steady-state data conforming to a single-exponential function were fitted to eq 8, in which time courses

at all fixed concentrations of selected dipeptide substrates were simultaneously fitted as previously described,[1] where A is the changing-fixed concentration of substrate, E_t is the enzyme concentration, k_{cat}, k_{ac}, and k_{dac}, are turnover numbers for the full reaction and the acylation and deacylation half-reactions, respectively, in units of s^{-1}, and K_a and K_{ia} are, respectively, the Michaelis constant and an apparent dissociation constant for each substrate (μM).

$$[P] = k_{cat}AE_t(t)/(K_a + A) + [k_{cat}A/(k_{dac}(K_a + A))]^2 \times E_t\{1 - \exp(-t)[((k_{ac} + k_{dac})A + k_{dac}K_{ia}) /(A + K_{ia})]\} \quad (8)$$

Nomenclature. Isotope effects are expressed using the nomenclature of Cook, Cleland and Northrop.[16,17] Solvent kinetic and equilibrium kinetic isotope effects are notated by the leading superscript D.

Synthesis of Dipeptide Nitrile Compounds. See Supporting Information.

Inhibition of Cathepsin C by Dipeptide Nitriles. Inhibition of cathepsin C activity was measured by preincubation of the nitrile inhibitors (dissolved in DMSO) with cathepsin C, followed by addition of 10 μM Ser-Tyr-AMC in Buffer A. Activity was measured by fluorescence, in the same manner as the steady-state kinetics assay.

Crystallography. Inhibitor complexes were prepared by adding 2 μL of 52.2 mM inhibitor in DMSO to 100 μL of 7.25 mg/mL protein (final concentration of 1 mM inhibitor and 2% (v/v) DMSO). The crystallization samples were placed in Hampton Cryschem sitting-drop plates (Hampton Research) which were prepared by transferring 0.5-mL of crystallization solution from a 24-well block containing solutions of 16−26% polyethylene glycol-3350 at pH 6.0−6.9 in 200 mM $MgCl_2$, 100 mM Bis-Tris buffer. The sitting drops were prepared by adding 2 μL of the protein-inhibitor complex to 2 μL of the crystallization solution. The plates were sealed with tape and incubated at room temperature. A cryoprotectant solution was prepared from 20% ethylene glycol and 80% crystallization solution. Twenty microliters of this mixture was added back to the crystallization drop. The crystals were mounted on loops and cryo-frozen with liquid nitrogen.

X-ray diffraction data were collected at the APS 21-ID-D(LS) beamline at Argonne National Laboratory. The data were processed and scaled using the program HKL2000 (HKL Research Inc.). The structures were solved by molecular replacement using the program REFMAC (CCP4), and the coordinates of the pdb file 1myln as a starting model (unpublished data). The X-ray data statistics are given in Table 4.

Molecular Dynamic Simulations of Acylated Cathepsin C. The initial coordinates for the acylated enzymes were taken from the crystal structure of the Ser-Tyr-CN-Cathepsin C complex, from which the crystallographically determined water molecules were removed. To construct a model of the acyl-enzyme intermediate, the observed imidothioate linkage established between the nitrile group and the catalytic Cys234 was replaced by a thioester. The resulting structure corresponded to a form of the enzyme that is acylated by the substrate Ser-Tyr-AMC. The initial coordinates for the enzyme acylated by Gly-Arg-AMC were made by substitution of the Ser-Tyr peptide sequence with Gly-Arg. Glycosyl substituents were removed from the three glycosylated residues (Asn-5, Asn-95, and Asn-252) on the surface of the protein, such that the residues were treated as unglycosylated amino acids.

The structural changes introduced to the original coordinates were minimal and were readily accommodated during protein preparation. The preparation was carried out in Molecular Operating Environment or MOE (ver. 2009.10, Chemical Computing Group Inc.), in which the protonation states of ionizable residues were assigned in the Protonate 3D module, using the default settings. However, the catalytic His-381 was manually treated. The neutral form was assigned with the polar hydrogen added to the NE2 nitrogen, to be consistent with its putative role in the deacylation step. In addition to protonation, the side chains capable of hydrogen bonding (except that of His-381) were optimized to maximize any hydrogen bond interactions with their immediate environments. Once prepared, the resulting structure was subject to a short energy optimization in MOE with all non-hydrogen atoms tethered with a strong positional restraint using the MMFF94x force field.[18]

All of the MD simulations conducted in this study were carried out with NAMD 2.7b1. The AMBER force field *ff99SB*[19] was chosen to represent all of the standard amino acids. For the acylated active-site residue Cys234, the *gaff* force field was used,[20] and a new set of RESP partial charges[21] were calculated for each substrate studied here (Tyr-S-Cys234 and Arg-S-Cys234) with capping groups (acetyl and dimethylamino). The electrostatic potentials were computed following standard procedures, using Gaussian 03,[22] with geometries taken from the prepared structures. During the charge-fitting step, the existing atomic charges in *ff99SB* were assigned to the capping groups and were not allowed to relax. The atom types of the acylated Cys234 were assigned using antechamber.[20] Any missing force field parameters due to the mixed use of *ff99SB* and *gaff* force fields were amended with existing ones that best match the chemical environment. No new force field term was necessary for the purpose of this study. The force field parameters assembled for the acylated cysteine are provided in the Supporting Information.

The distribution of water molecules was calculated from the simulations with explicit water molecules incorporated around the acylated enzyme. Roughly 15 000 SPCE[23] water molecules were added to form a box of 85 Å × 75 Å × 80 Å in dimension, and a periodic boundary condition was employed. For van der Waals interactions, the evaluation was cut off at 10.0 Å. To avoid discontinuity, a switching function was turned on at 8.0 Å. For electrostatic interactions, a shifting function was switched on at the cutoff distance. Long-range electrostatic interactions were handled by particle-mesh-ewald (PME)[24] with a uniform grid spacing of 1.0 Å. The simulations were calculated under constant temperature (300 K) and pressure (1 atm). Temperature was regulated by a Langevin thermostat with a damping frequency of 5 per picosecond (ps). Constant pressure was achieved by a Nose-Hoover Langevin barostat with a 200 fs (fs) oscillation time and 100 fs damping time.[25] A typical equilibration procedure was used to bring the system to the desired temperature and pressure. First, the entire system underwent a series of energy minimizations with diminishing positional restraints. It was then followed by a series of short, constant-temperature, MD simulations (300 K) at a total of 30 ps with diminishing positional restraints. Once the equilibrium temperature was achieved, the Nose-Hoover barostat was switched on, and the system was allowed to relax for about 200 ps after the desired pressure was reached. A one-femtosecond

Table 1. Pre-Steady-State and Steady-State Kinetic Parameters for Dipeptide-AMC Substrates and Dipeptide-Nitrile Inhibitors of Cathepsin C[a]

substate or inhibitor type	substrate	k_{cat} (s^{-1})	k_{cat}/K_a (μM^{-1} s^{-1})	K_a (μM)	rel k_{cat}[b]	rel k_{cat}/K_a[b]	app k_{ac} (s^{-1})[c]	app k_{ac}/k_{cat}[c]
1	Gly-Tyr-AMC	28 ± 1	0.49 ± 0.067	57 ± 8	1.0	1.0	495 ± 7	17
1	Gly-Phe-AMC	54 ± 2	0.83 ± 0.073	64 ± 6	1.9	1.7	≥500	≥50
1	Ser-Tyr-AMC	25 ± 0.5	5.3 ± 0.6	5.0 ± 0.8	0.9	10.8	400 ± 5	16
1	Ala-Trp-AMC	26 ± 0.5	3.7 ± 0.4	7.0 ± 0.6	0.9	7.6	≥400	≥15
1 or 2	Gly-His-AMC	64 ± 3	0.12 ± 0.013	525 ± 40	2.3	0.2	N/A	N/A
2	Gly-Arg-AMC	255 ± 6	1.6 ± 0.09	160 ± 9	9.1	3.3	≥500	2
2	Gly-Lys-AMC	290 ± 6	1.0 ± 0.06	250 ± 14	10.4	2.0	N/A	N/A
2	Gly-Glu-AMC	126 ± 4	0.046 ± 0.003	2800 ± 170	4.5	0.1	≥500	4
2	Gly-Ala-AMC	110 ± 2	0.081 ± 0.0042	1390 ± 70	3.9	0.2	N/A	N/A
2	Gly-Leu-AMC	109 ± 2	0.86 ± 0.045	126 ± 6	3.9	1.8	≥500	5
2	Ser-Arg-AMC	142 ± 5	4.5 ± 0.6	32 ± 4	5.1	9.2	N/A	N/A
3	Gly-Ile-AMC	0.33 ± 0.2	0.0015 ± 0.00011	250 ± 26	0.012	0.003	0.5	1
3	Gly-Val-AMC	0.69 ± 0.01	0.0017 ± 0.0001	395 ± 10	0.025	0.035	<1	1
3	Gly-Pro-AMC		0.3×10^{-6}[d]					
3	Ala-Pro-AMC		0.5×10^{-6}[d]				≥500	2
3	Lys-Pro-AMC		9.0×10^{-9}[d]					
	Arg-AMC	0.005 ± 0.0005	$6.2 \pm 1.9 \times 10^{-6}$	874 ± 260				

	inhibitors	K_i (μM)[e]
1	Ser-Tyr-CN	0.1 ± 0.1 (1.0)
2	Gly-Lys-CN	5.0 ± 0.5 (50)
3	Gly-Ile-CN	110 ± 15 (1100)

[a]The kinetic constants were determined at 21 °C in 50 mM sodium acetate, 30 mM NaCl, 1 mM EDTA, 1 mM CHAPS, 1 mM DTT buffer (pH 5.5) from fitting of initial-rate data to eq 1. [b]Values are relative to data of Gly-Tyr-AMC. [c]Values of transient rates (λ) are at saturating concentrations of the substrate except for Gly-Tyr-AMC (at $5K_a$), and provide apparent values of k_{ac} when k_{ac}/k_{cat} >10. [d]Rate constant from linear fit of the data R^2 value >0.9. [e]Values in parentheses are relative to data of Ser-Tyr-CN.

time step was used throughout the equilibration stage. Various energetic terms and the system volume were checked before the production run was committed. All the production dynamics in this study were carried out uninterrupted for 10 ns (ns) with a 2-fs time step. System configurations computed were collected every 4 ps, resulting in 2500 samples per simulation for subsequent analysis.

To estimate the conformational preference of the tyrosine side chain in Ser-Tyr-S-Cys234, the adaptive biasing force (ABF) method[26,27] implemented in NAMD[28,29] 2.7b1 was applied to the $\chi 1$ dihedral angle over the entire angular range (−180° to 180°) with a grid spacing of 5°. The ABF simulation was carried out for 10 ns, during which the computed potential of mean force $W(\chi 1)$ was monitored for convergence. Additional conformational states identified in the ABF simulation were subject to further examinations.

RESULTS AND DISCUSSION

Steady-State Kinetics of Cathepsin C Di-Peptide Substrates. The steady-state kinetic parameters for a panel of dipeptide-AMC substrates of cathepsin C are summarized in Table 1. Values of k_{cat}, k_{cat}/K_a, and K_a for the dipeptide substrates comprise a range of values spanning several orders of magnitude. In most cases, subject to the availability of substrates, we examined the effects of substituting the scissile amino acids of the substrate (P_1 position), while maintaining the N-terminal amino acid (P_2) as glycine.

In previous studies with a single substrate Ser-Tyr-AMC,[1] we showed that the rate of the acylation half-reaction (k_{ac}) was 20-fold greater than that of the deacylation half-reaction (k_{dac}). Accordingly, the kinetic parameter that best characterizes substrate recognition for binding to cathepsin C, and its subsequent acylation, is k_{cat}/K_a, while values of the Michaelis constant, K_a, for this double-displacement chemical mechanism are complicated by the fact that they are dependent on whether or not k_{cat} is rate-limited by k_{ac} or k_{dac}.[c] As a result, for most of the following discussion, we emphasize k_{cat}/K_a, rather than K_a, as the dominant determinant of substrate recognition. To this point, variation of the amino acid at P_1 leads to a range of k_{cat}/K_m values that span 7 orders of magnitude, from 0.34×10^{-6} μM^{-1} s^{-1} to 5.3 μM^{-1} s^{-1} (Table 1), while it contributes only a ∼1000-fold range to the values of substrate K_a, from about 5 μM to 2800 μM, for which an uncharged, aromatic residue at P_1 appears to contribute an enhancement to this value (lower values of K_a; Table 1). For substrates bearing an invariant glycyl residue at the P_2 position, k_{cat}/K_a values ranging from 0.08 to 1.6 μM^{-1} s^{-1} were found for straight-chain residues at P_1, for which the larger values corresponded to longer and cationic side-chains (Glu < Ala < Leu < Lys < Arg). This indicates a slight preference for the identity of the amino acid found at the P_1 position. With regard to the effects of the P_2 residues, the data for Ser-Tyr-AMC vs Gly-Tyr-AMC ($(k_{catSer\text{-}Tyr\text{-}AMC})/(k_{catGly\text{-}Tyr\text{-}AMC})$ = 0.9, $(k_{cat}/K_{aSer\text{-}Tyr\text{-}AMC})/(k_{cat}/K_{aGly\text{-}Tyr\text{-}AMC})$ = 11) and Ser-Arg-AMC vs Gly-Arg-AMC ($(k_{catSer\text{-}Arg\text{-}AMC})/(k_{catGly\text{-}Arg\text{-}AMC})$ = 0.6, $(k_{cat}/K_{aSer\text{-}Arg\text{-}AMC})/(k_{cat}/K_{aGly\text{-}Arg\text{-}AMC})$ = 3.3) demonstrated that a serinyl, rather than a glycyl residue, in the P_2 position promotes a 3−11-fold increase in catalytic efficiency, expressed as k_{cat}/K_a, with little effect on k_{cat}. Aromatic groups at the P_1 position (His, Phe, Tyr) for Gly-Xaa-AMC substrates exhibited comparable values of k_{cat}/K_a = 0.12−0.83 μM^{-1} s^{-1}, which were lower than substrates containing charged and straight-chained residues at P_1. However, Gly-His-AMC is likely to be cationic under our assay conditions, and so also exhibits kinetic characteristics of substrates bearing a cationic group at P_1 (k_{cat}/K_a < 1 μM^{-1} s^{-1}). Dipeptides bearing a β-branched amino acid at P_1 (Ile, Val)

dx.doi.org/10.1021/bi300719b | Biochemistry 2012, 51, 7551−7568

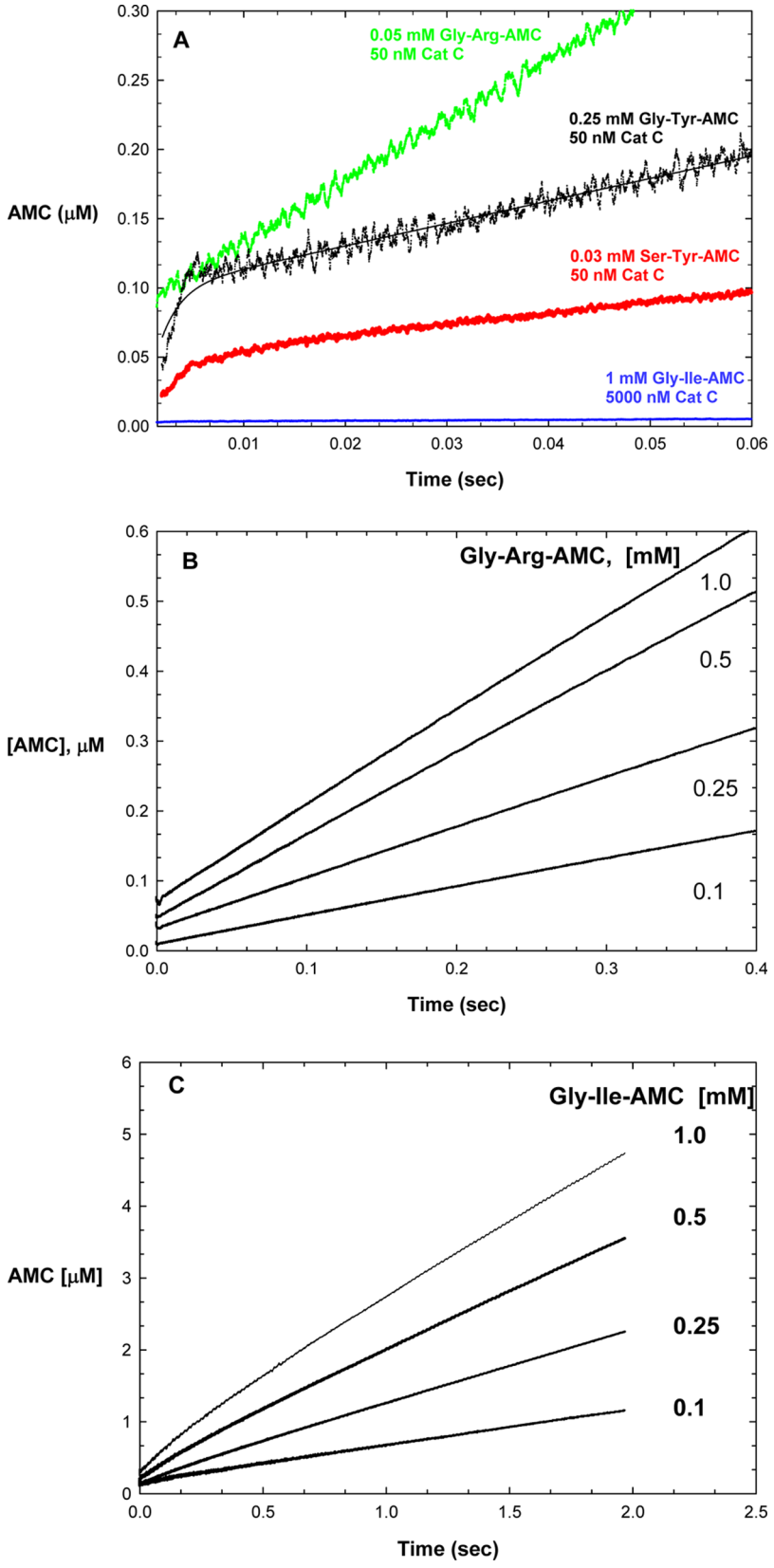


Figure 1. Pre-steady-state kinetics of cathepsin C substrates. (A) Pre-steady-state kinetics of Ser-Tyr-AMC (red), Gly-Tyr-AMC (black) Gly-Arg-AMC (green), and Gly-Ile-AMC (blue) in a stopped-flow fluorimeter detecting the AMC product between 0 and 60 ms at the indicated substrate concentrations and using 50 nM cathepsin C (for Gly-Ile-AMC, enzyme was at 5000 nm enzyme and the product formed was normalized for the plot to a value of 50 nM enzyme). Experimental data points for all four time courses were fitted to eq 7 (lines). (B) Time courses of Gly-Arg-AMC at 50 nM enzyme and the indicated concentrations of Gly-Arg-AMC where the experimental data were fitted to eq 9. (C) Time courses of Gly-Ile-AMC at 5000 nM enzyme and the indicated concentrations of Gly-Ile-AMC where the experimental data were fitted to eqs 7 and 9.

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

were poorly accommodated by cathepsin C (k_{cat}/K_a = 0.0015−0.0017 $\mu M^{-1}s^{-1}$), suggesting that the branched nature of the residues at P_1 may hinder facile attack by Cys-234 on the scissile amide in the acylation half-reaction, and/or water during the deacylation half-reaction for these substrates. This trend is also seen with substrates containing a proline at P_1, as previously reported,[30] since in our studies both Ala-Pro-AMC and Gly-Pro-AMC were almost completely inactive.

The similar catalytic efficiencies of Ala-Trp-AMC and Ser-Tyr-AMC (k_{cat}/K_a = 3.7 and 5.3 μM^{-1} s^{-1}, respectively) suggested that a substrate bearing at least a methyl group occupying the S_2 pocket can aid with substrate recognition and acylation. This is further supported by the finding that the value of k_{cat}/K_a = 5.3 μM^{-1} s^{-1} for Ser-Tyr-AMC is an order of magnitude greater than that of Gly-Tyr-AMC (k_{cat}/K_a = 0.49 μM^{-1} s^{-1}), and that of Ser-Arg-AMC is 3-fold greater than Gly-Arg-AMC, although both sets of substrates had comparable values of k_{cat}. This suggests that the dissociation constants (K_{ia}) of substrates of the type Ser-Xaa-AMC will be lower in value by as much as an order of magnitude than those of the form Gly-Xaa-AMC. The similarity in k_{cat} values for these substrates likely indicates that the rate-limiting steps for Ser-Tyr-AMC and Gly-Tyr-AMC occur in the deacylation half-reaction, as previously shown for the Ser-Tyr-AMC,[1] since the nature of the P_2 residue would probably not affect the rate of hydrolysis of the acyl-enzyme thioester intermediate as much as the initial recognition by the protease during acylation. The very low k_{cat} observed for the monopeptide Arg-AMC supports the idea that a minimum of two amino acids is required, suggesting that the carboxylic acid side chain of Asp-1 of cathepsin C is not likely to be sufficiently proximal to form an ionic bond with the N-terminal amino group of Arg as it does with the amino group of the P_2 residues of dipeptide substrates. This is also consistent with previously published studies which showed substitutions of the N-terminal amino acid or modification of the N-terminal amine can have significant detrimental effects on k_{cat}/K_a.[2]

Similar to the wide range of k_{cat}/K_a values measured for the dipeptide substrates of cathepsin C, the substitutions at P_1 observed here also resulted in k_{cat} values that spanned over 3 orders of magnitude. The k_{cat} values for these substrates indicated some general trends. Substrates with aromatic amino acids at P_1 (Tyr, Phe, Trp, His) displayed k_{cat} values less than 100 s^{-1} (25−64 s^{-1}); substrates with β-branched aliphatic amino acids at P_1 were poorly hydrolyzed with k_{cat} < 1 s^{-1}, while substrates bearing straight-chain residues at P_1 such as Lys, Arg, Glu, Ala, and Leu resulted in k_{cat} > 100 s^{-1} (109−290 s^{-1}). Gly-Pro-, Ala-Pro-, and Lys-Pro-AMC were such poor substrates that Michaelis−Menten kinetics were not obtained.

In summary, steady-state data indicate here three general types of dipeptide substrates for cathepsin C of the form Gly-Xaa-AMC: Type 1; where Xaa = Phe, Tyr, Trp, His; k_{cat} values = 20−60 s^{-1} and k_{cat}/K_a values = 0.1−0.8 μM^{-1} s^{-1}; Type 2; where Xaa = His, Lys, Arg, Leu, Glu, Ala; k_{cat} values = 100−300 s^{-1} and k_{cat}/K_a values = 0.05−1.6 μM^{-1} s^{-1}; and Type 3; where Xaa = Ile or Val; k_{cat} values = 0.3−0.7 s^{-1} and k_{cat}/K_a values = 0.0015−0.0017 μM^{-1} s^{-1}. Note that Gly-His-AMC has properties of both Type 1 and Type 2 substrates, which may arise from the fact that it is both aromatic and likely cationic at the P_1 position.

Transient Kinetics Studies. Expanding on earlier studies with Ser-Tyr-AMC, we next investigated the role of the amino acid at the P_1 position on the transient kinetics of selected cathepsin C dipeptide substrates using a stopped-flow fluorimeter.[1] As previously reported, Ser-Tyr-AMC (in Figure 1: 30 μM; 6K_a), a Type 1 substrate, demonstrated a single-exponential burst, which upon fitting to eq 7 rendered apparent values of k_{ss} = 12.23 ± 0.01 s^{-1}, λ = 150 ± 1 s^{-1} and β/E_t = 0.604 ± 0.001. As described before,[1] the single-turnover transient phase is commensurate with the acylation half-reaction (k_{ac}) of Scheme 1, while the steady-state phase described the deacylation half reaction (k_{dac}), which contains the rate-limiting step(s) of the peptidolysis of Ser-Tyr-AMC. Evaluation of the transient kinetics of Ser-Tyr-AMC at changing-fixed concentrations produced a family of single-exponential time courses (data not shown) which upon fitting to eq 8 led to the limiting values of k_{ac} = 400 ± 5 s^{-1} and k_{cat} = 16.7 ± 0.01 s^{-1} and k_{dac} = 15.3 ± 0.03 s^{-1}. These results are nearly identical to previously reported results,[1] with differences likely attributable to variation in the fraction of active cathepsin C enzyme in the preparations used.

Similar pre-steady-state kinetics were observed for the other Type 1 substrates, Ala-Trp-AMC and Gly-Tyr-AMC (Figure 1; at 250 μM; 5K_a). Time courses for these substrates were described by single-exponential bursts, followed by linear kinetics, similar to Ser-Tyr-AMC. In contrast, the pre-steady-state time course of Gly-Phe-AMC (data not shown) displayed only a nonzero intercept with no apparent transient phase, indicating that the transient "burst" occurred within the 2-ms dead-time of the stopped-flow spectrophotometer. For Gly-Tyr-AMC, the transient observed in Figure 1 led, upon fitting to eq 7, to values of k_{ss} = 28 ± 0.01 s^{-1}, λ = 490 ± 7 s^{-1}, β/E_t = 1.3 ± 0.7. The steady-state kinetic rate constant obtained here at [Gly-Tyr-AMC] = 5K_a (k_{ss} = 28 ± 0.01 s^{-1}) is, as expected, identical to its value of k_{cat} = 28 ± 1 s^{-1}, while the apparent transient rate of λ = 490 ± 7 s^{-1} is 20% faster than that of Ser-Tyr-AMC. The apparent transient rate is nearly equal to the rate of desorption of the AMC product (460 s^{-1}) as determined for Ser-Tyr-AMC,[1] suggesting that this product-release step for Gly-Tyr-AMC is the slowest in the acylation half-reaction. Also, as with Gly-Tyr-AMC, the apparent burst amplitude, β/E_t = 1.3 ± 0.7, is within experimental error effectively a value of unity, thereby suggesting that $k_{cat} \sim k_{dac}$. The pre-steady-state data for Gly-Tyr-AMC are therefore similar to that of Ser-Tyr-AMC, indicating that for both of these substrates, the rate of acylation of cathepsin C (k_{ac}) greatly exceeds that of deacylation (k_{dac}), such that $k_{cat} \sim k_{dac}$, which is likely a common property of the Type 1 substrates of cathepsin C which display single-exponential bursts.

For the Type 2 substrates with k_{cat} values exceeding 100 s^{-1}, Gly-Arg-AMC (Figure 1B), Gly-Glu-AMC, and Gly-Leu-AMC, while no transient time courses were observed, a pre-steady-state burst (nonzero intercept) was evident, indicating that the transient phase (the rates of the acylation) for these substrates occurred during the 2-ms dead-time of the instrument. Accordingly, upon fitting the data for the time courses of each concentration of Gly-Arg-AMC shown in Figure 1A,B to eq 7, we obtained discrete values of the steady-state rate (k_{ss}) and the burst amplitude size (β), while the transient rate (λ) could not be determined. The apparent steady-state rates and burst amplitudes were determined at each fixed concentration of Gly-Arg-AMC by fitting to eq 9 to obtain limiting values of k_{ss} = 300 ± 30 s^{-1} and $K_{Gly\text{-}AMC}$ = 10 ± 4 μM, which, as expected, are in good agreement with the steady-state kinetic parameters shown in Table 1. The resulting superstoichiometric value of β/E_t = 21 ± 3 underscores the difficulty of obtaining a

dx.doi.org/10.1021/bi300719b | Biochemistry 2012, 51, 7551−7568

Table 2. Steady-State Kinetic Parameters, pH-Rate Profile Data, and Solvent Kinetic Isotope Effects of Cathepsin C Dipeptide Substrates[a]

substrate	solvent	k_{cat}/K_a (mM^{-1} s^{-1})	pK value (k_{cat}/K_a)	eq fitted	k_{cat} (s^{-1})	pK value (k_{cat})	eq fitted	solvent kinetic isotope effects
Ser-TyrAMC	H_2O	6.2 ± 0.34	pK_1 = 4.17 ± 0.06	3	24.6 ± 0.8	pK_4 = 3.45 ± 0.07	2	
			pK_2 = 6.5 ± 0.3			pK_5 = 7.2 ± 0.05		$^D(k_{cat}/K_a)$ = 0.74 ± 0.05
			pK_3 = 7.8 ± 0.1					$^Dk_{cat}$ = 2.3 ± 0.1
	D_2O	8.2 ± 0.4	pK_1 = 5.0 ± 0.04	2	11.2 ± 0.4	pK_4 = 4.32 ± 0.02	2	
			pK_2 = 7.25 ± 0.05			pK_5 = 8.9 ± 0.05		
			pK_1 = 4.5 ± 0.08			pK_4 = 4.6 ± 0.09		
Gly-IleAMC	H_2O	0.0015 ± 0.0001	pK_2 = 7.0 ± 0.4	3	0.40 ± 0.012	pK_5 = 7.24 ± 0.09	2	$^D(k_{cat}/K_a)$ = 1.0 ± 0.5
			pK_3 = 7.8 ± 0.4					$^Dk_{cat}$ = 1.06 ± 0.04
	D_2O	0.0017 ± 0.0008	pK_1 = 4.74 ± 0.02	2	0.38 ± 0.01	pK_4 = 4.39 ± 0.06	2	
			pK_2 = 8.0 ± 0.09			pK_5 = 7.83 ± 0.05		
			pK_1 = 5.8 ± 0. 1			pK_a = 4.8 ± 0.3		
			pK_2 = 6.5 ± 0.7			pK_b = 5.4 ± 0.3		$^D(k_{cat}/K_a)$ = 0.58 ± 0.15
Gly-ArgAMC	H_2O	4.4 ± 1.2	pK_3 = 8.18 ± 0.04	4	320 ± 50	pK_c = 6.9 ± 0.1	6	$^Dk_{cat}$ = 1.6 ± 0.2
			K_a = 0.08 ± 0.015			pK_d = 9.2 ± 0.5		
			K_b = 0.01 ± 0.002					
			pK_1 = 6.6 ± 0. 3			pK_a = 4 ± 1		
	D_2O	7.6 ± 1.5	pK_1 = 7.8 ± 0.3	5	212 ± 17	pK_b = 5.2 ± 0.2	6	
			K_a = 0.13 ± 0.013			pK_c = 7.5 ± 0.08		
			K_b = 0.016 ± 0.002			pK_d = 9 ± 2		

[a]Results were obtained from fitting of experimental data shown in Figure 1 to eqs 2−6 as indicated. Solvent kinetic isotope effects were calculated from the appropriate c values, and error was determined, as described in ref 1.

reliable burst amplitude for such an active substrate at relatively higher concentration of fluorescent substrates.

$$X = x_{max}[\text{Gly-Arg-AMC}]/([\text{Gly-Arg-AMC}] + K_{\text{Gly-Arg-AMC}}) \quad (9)$$

The value of k_{cat} = 300 ± 30 s^{-1} for Gly-Arg-AMC is only slightly less than the rate constant of acylation for Ser-Tyr-AMC (λ = 397 ± 5 s^{-1}), where the slowest step in the acylation half-reaction was determined to be the release of the product AMC (460 s^{-1}),[1] which is a common step for both substrates. Accordingly for Gly-Arg-AMC, if assuming that k_{ac} ∼ 460 s^{-1}, then a value of k_{cat} = 260 ± 30 s^{-1} is obtained when the rate of deacylation is ∼ 640 s^{-1}, which is ∼50-fold faster than the rate of deacylation for Ser-Tyr-AMC (∼13 s^{-1}). As a result for Gly-Arg-AMC, the observed value of k_{cat} will represent nearly equal contributions from both the acylation (k_{ac}) and deacylation (k_{dac}) half-reactions.

With the Type 3 substrates, Gly-Ile-AMC (Figure 1C) and Gly-Val-AMC (data not shown), slow kinetics were observed with no pronounced pre-steady state burst or nonzero intercept. These results suggested that both the acylation and deacylation half-reactions are slow. For each concentration of Gly-Ile-AMC (Figure 1C), a small transient is slightly discernible, but the two phases expected in a single-exponential time course are not distinct and appear to merge. Moreover, these experiments were performed using 5 μM of cathepsin C, rather than 50 nM enzyme, which was used for the other substrates. Attempts to fit the family of time courses to eq 8 were unsuccessful, but fitting the curves at each concentration of Gly-Ile-AMC to eq 7 (Figure 1C) yielded values of k_{ss} and β/E_t at all concentrations of substrate, while values of the transient rate (λ) were poorly determined. Limiting values of k_{ss} = 0.47 ± 0.02 s^{-1} (compare to k_{cat} = 0.33 s^{-1} from steady-state data) and β/E_t = 0.1 ± 0.04 were obtained by use of eq 7. Since at saturating substrate concentrations $\beta/E_t = k_{cat}/k_{dac}$ = 0.1 ± 0.04, we can solve for estimated values of k_{dac} = 5 s^{-1} and k_{ac} = 0.5 s^{-1} for Gly-Ile-AMC, indicating that for this substrate, the rate of the acylation half-reaction is ∼10 times slower than that of deacylation. As the same is likely true for Gly-Val-AMC, which likewise displayed no pre-steady-state burst, then the occurrence of a β-branched residue at the P_1 position of these two substrates results in an apparent change of the rate-limiting step(s) for cathepsin C from deacylation, seen for type 1 substrates, to the acylation half-reaction. It is possible that steric hindrance from the P_1 side chain of Gly-Ile-AMC hinders access of Cys-234 to the scissile amide group of the substrate more than it hinders access of the catalytic water to the thioester of the acyl-enzyme formed during deacylation.

pH-Rate Profiles. We next investigated the pH-rate profiles for the hydrolysis of Gly-Arg-AMC and Gly-Ile-AMC by cathepsin C (Table 2). The pH(D)-rate profiles obtained in H_2O and D_2O for Gly-Arg-AMC and Gly-Ile-AMC are shown in Figure 2. As described before for Ser-Tyr-AMC,[1] a type 1 substrate, and repeated for this paper, the plot of log k_{cat}/K_a vs pH was a bell-shaped curve with apparent slopes of +1 and −2 on the respective acidic and basic sides, respectively, of the profile (data not shown). Fitting of the data to both eqs 2 and 3 led to a better fit for eq 3, for which $c = k_{cat}/K_a$ = 6.2 μM^{-1} s^{-1}, pK_1 = 4.17, pK_2 = 6.5, and pK_3 = 7.8.[2] These results are nearly identical to those reported previously, in which we attributed pK_1 = 4.17 and pK_2 = 6.5, respectively, to the active-site Cys-234 in thiolate form and the active-site His-381 in imidazolium form. Consistently, the third pK value, pK_3 = 7.8, appears only in the plot of log k_{cat}/K_a vs pH. The absence of a third pK value in this and other plots of log k_{cat}/K_a vs pD may reflect the inability to obtain reliable data above pD = 8.0. The plot of log k_{cat} vs pH (data not shown) was also defined by a bell-shaped curve comprised of slopes of +1 and −1, which upon fitting to eq 2 rendered values of $c = k_{cat}$ = 24.6 s^{-1}, pK_4 = 3.4, and pK_5 =

dx.doi.org/10.1021/bi300719b | Biochemistry 2012, 51, 7551−7568

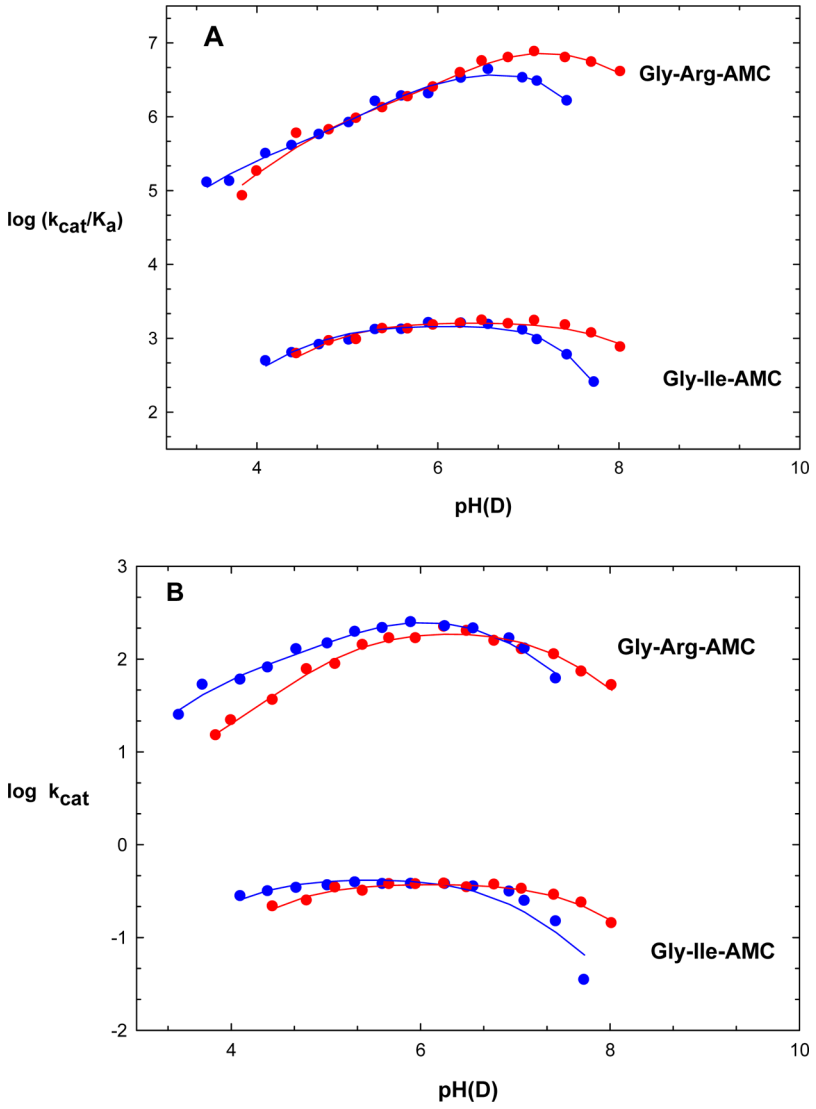


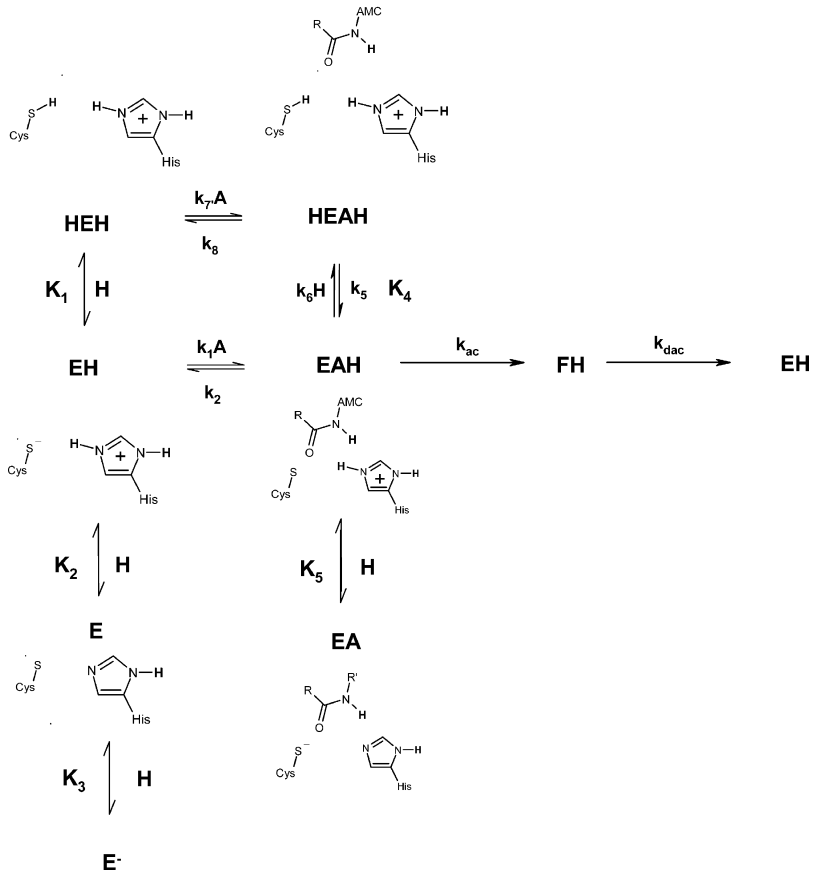
Figure 2. pH and pD-rate profiles for k_{cat}/K_a (2A) and k_{cat} (2B) for Gly-Arg-AMC and Gly-Ile-AMC determined in H_2O (blue) and D_2O (red). For Gly-Ile-AMC, the lines drawn through experimental data for log k_{cat}/K_a vs pH and pD were from fitting to eqs 3 and 2, respectively, and for log k_{cat} vs pH and pD, to eq 2. For Gly-Arg-AMC, the lines drawn through experimental data for log k_{cat}/K_a vs pH and pD were from fitting to eqs 5 and 4, respectively, and for log k_{cat} vs pH and pD, to eq 6.

7.2. In the k_{cat} profile, we speculated that the group represented by pK_5 = 7.2 was also the imidazolium form of His-381, which catalyzes the rate-limiting collapse of the tetrahedral intermediate in the deacylation half-reaction (FX in Scheme 1). Also in concert with our previous results, both plots of log k_{cat}/K_a and k_{cat} vs pD were described by bell-shaped curves, best fitted by eq 2: for which $c = k_{cat}/K_a$ = 8.2 μM^{-1} s^{-1}, pK_1 = 5.0, pK_2 = 7.25; and $c = k_{cat}$ = 11.2 s^{-1}, pK_4 = 4.32, and pK_5 = 8.9. For Ser-Tyr-AMC, analysis of the proton inventory of k_{cat} at pH(D) = 6.8 was best fitted to a model in which a single proton was transferred in the transition state, which is also consistent with the proposal that proton transfer from the imidazolium group of the FX intermediate in Scheme 1 constitutes the rate-limiting step for this substrate (data not shown).

Shown in Figure 2, for Gly-Ile-AMC, a type 3 substrate, the pH-rate profiles of log k_{cat}/K_a vs pH and log k_{cat} vs pH were of the same form as those of Ser-Tyr-AMC, and each were best fitted to eqs 3 and 2, respectively, yielding pK values shown in Table 2. It is noteworthy that while the values of pK_1, pK_2, and pK_3 from the profile of log k_{cat}/K_a vs pH are nearly identical to those found for the more active substrate, Ser-Tyr-AMC, the acidic and basic dissociation constants of the plot of log k_{cat} vs pH do not agree with those of Ser-Tyr-AMC. Indeed, the near equality of pK_1 to pK_4 and pK_2 to pK_5 for Gly-Ile-AMC suggests that the protonation states of Cys-234 (pK_1) and His-381 (pK_2) are the same for both free enzyme (as expressed by k_{cat}/K_a), and the overall rate-limiting step(s) for Gly-Ile-AMC peptidolysis (as expressed by k_{cat}). If the attack of Cys-234 on Gly-Ile-AMC is among the slowest catalytic steps, one would expect that the pK values of thiolate and imidazolium forms of Cys-234 and His-381, respectively, would be expressed in the overall k_{cat} for this substrate. This is consistent with findings from the transient kinetics of Gly-Ile-AMC which suggested that the rate of acylation was 10-fold slower than deacylation. In contrast, the pK = 4.5 of the thiolate is not observed in profiles of log k_{cat} vs pH for Ser-Tyr-AMC, since deacylation is rate-limiting for this substrate, and for the catalytic steps of deacylation, the thiol group of Cys-234 is not available as a prototropic species. For Gly-Ile-AMC in D_2O, both bell-shaped

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

Scheme 2



plots of log k_{cat}/K_a vs pD and log k_{cat} vs pD were best fitted to eq 2 yielding pK values shown in Table 2.

For the type 2 substrate Gly-Arg-AMC, both pH-and pD-rate profiles of log k_{cat}/K_a and log k_{cat} resemble bell-shaped curves in which the slopes on the acidic side of the curves are <1.0, while on the basic side, apparent slopes of either −2 (log k_{cat}/K_a vs pH) or −1 (log k_{cat}/K_a vs pD and log k_{cat} vs pH and pD) were observed. The acidic phases of the plots have been described as a "hollow", which is observed in cases where either the substrate and/or a proton is "sticky" in the Michaelis complex.[31] A "hollow" has been characterized for the log V/K vs pH profile for phosphocreatine for rabbit creatine kinase.[32]

A comprehensive model for the pH-rate data of cathepsin C is shown in Scheme 2. The EH form represents the catalytically competent form of the unbound enzyme in which Cys-234 and His-381 comprise tautomeric forms where either a thiolate-imidazolium (shown) or thiol-imidazole dyad exist (EH is E in Scheme 1). Protonation of EH produces the inactive form HEH in which both Cys-234 and His-381 are protonated, while single or double deprotonation of EH results in the inactive E and E^- forms, respectively, the former arising from the inactive thiolate-imidazole dyad. Both of these inactive protonated (HEAH) and deprotonated species (EA) may be formed from the substrate-bound Michaelis complex (EAH). The classically bell-shaped log k_{cat} vs pH(D) profiles for the substrates Ser-Tyr-AMC and Gly-Ile-AMC suggest that these substrates are not very sticky in their respective EAH complexes or that the protonation described by the k_6 step is very rapid. In Scheme 2, acylation and deacylation half-reactions are encompassed by the single rate constants, k_{ac} and k_{dac}, respectively. In contrast, for the very active type 2 substrate Gly-Arg-AMC, it is likely that the k_5 and k_6 steps and the EAH form achieve steady-state. On the basis of precedent,[31] expressions for the pH(D)-dependence of log k_{cat}/K_a and log k_{cat} for Scheme 2 may be derived:

$$\log k_{cat}/K_a = \log\{c(1 + H/K_1K_a)/[(1 + H/K_b)(1 + H/K_1 + (K_2/H)(1 + K_3/H))]\}$$

(same form as eq 4)

for which K_1, K_2, and K_3 are the acidic and basic dissociation constants described in Scheme 2, and $c = k_1k_{ac}/(k_2 + k_{ac})$, $K_a = (k_5 + k_8)/k_5k_7$, $K_b = k_1(k_2 + k_{ac})(k_5 + k_8)/k_2k_5k_7$; and

$$\log k_{cat} = \log\{c(1 + H/K_a)/[1 + H/K_b + K_c/H + H^2/K_d]\}$$ (same form as eq 6)

for which K_5 is the basic dissociation constant of Scheme 2, and $c = k_{ac}k_{dac}/(k_{ac} + k_{dac})$, $K_4 = k_5/k_6$, $K_a = K_4k_2(k_5 + k_8)/k_5k_8$, and $K_b = K_4k_2(k_{ac} + k_{dac})(k_5 + k_8)/(k_{ac}k_5k_8 + k_{ac}k_8k_{dac} + k_2k_8k_{dac} + k_5k_8k_{dac})$, $K_c = K_5k_{dac}/(k_{ac} + k_{dac})$, and $K_d = K_4^2k_2(k_{ac} + k_{dac})(k_5 + k_8)/k_5k_8k_{dac}$.

The pH-rate profiles for Gly-Arg-AMC were fitted to eqs 2−6, and the best fits were determined from linear regression and goodness-of-fit of the resulting kinetic parameters (Table 2). The profiles of log k_{cat}/K_a vs pH and pD were best fitted to eqs 4 and 5, respectively, resulting in values shown in Table 2. Again, the second basic group (pK_3) is observed in the plot of log k_{cat}/K_a vs pH but not in the plot in D_2O, as was seen with both Ser-Tyr-AMC and Gly-Ile-AMC. Apart from the value of pK_1 = 5.8 for Gly-Arg-AMC, which differs from that of the

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

Table 3. Solvent Kinetic Isotope Effects of Di-Peptide Substrates of Cathepsin C[a,b]

	Gly-Arg-AMC			Gly-Ile-AMC		
kinetic parameters	experimental values	calculated values[b]	equation used for calculated values	experimental values	calculated values[b]	equation used for calculated values
k_{cat}/K_a (μM^{-1} s^{-1})	1.6 ± 0.2	1.6	10	0.0015 ± 0.0001	0.0015	10
k_{cat} (s^{-1})	255 ± 6	250	11	0.33 ± 0.012	0.33	12
$^D(k_{cat}/K_a)$	0.58 ± 0.15	0.75	15	1.0 + 0.5	1.02	15
$^Dk_{cat}$	1.6 ± 0.09	1.6	16	1.06 ± 0.04	1.02	17
$K_{ia} = k_2/k_1 = 62$ μM, $k_3 = 8000$ s^{-1}, $k_4 = 13$ s^{-1}, $k_3/k_2 = a = 80$, $k_5/k_4 = b = 690$, $c = k_6/k_7 = 0.01$, $K_{t2} = 0.7$, $k_9 = 2000$ s^{-1}, $k_{10}/k_{11} = x = 0.05$, $k_9/k_{11} = z = 1.0$				$K_{ia} = k_2/k_1 = 2300$ μM, $k_3 = 0.7$ s^{-1}, $k_4 = 10$ s^{-1}, $k_3/k_2 = a = 0.0011$, $k_5/k_4 = b = 5$, $c = k_6/k_7 = 0.01$, $k_5/k_7 = d = 0.1$		

[a]Values for kinetic constants for Ser-Tyr-AMC taken from Table 3 of ref 1; $K_{ia} = k_2/k_1 = 36$ μM, $k_3 = 1500$ s^{-1}, $k_4 = 13$ s^{-1}, $k_3/k_2 = a = 2.3$, $k_5/k_4 = b = 220$, $c = k_6/k_7 = 0.01$, $k_9 = 115$ s^{-1}, $k_{10}/k_{11} = x = 2.8$, $k_9/k_{11} = z = 3.2$. [b]Common candidate values used in all expressions were $K_{t1} = 0.7$, $^DK_t = 2.56$. $^DK_{eq\,5} = 0.91$, $k_7 = 460$ s^{-1}, $k_{13} > 1000$ s^{-1}, $^Dk_5 = 3.2$, $^Dk_9 = 2.6$, and $^Dk_{11} = 3.4$.

other substrates, values of pK_2 and pK_3 for the three substrates are in good agreement.

The profiles of log k_{cat} vs pH and pD were best fitted to eq 6, resulting in values also shown in Table 2. Despite the complexity of the fitting, it is noteworthy that the fitted values $pK_c = \log(K_5k_{dac}/(k_{ac} + k_{dac})) = 6.9$ (H_2O) and 7.5 (D_2O) are in good agreement with those of the other two substrates ($pK_5 = 7.2$ (H_2O) and 8.9 (D_2O)) that were obtained from fitting to the simpler bell-shaped expression (eq 2), which is due in part to the comparable values of $k_{ac} \sim 460$ s^{-1} and $k_{dac} \sim 640$ s^{-1} for Gly-Arg-AMC.

Solvent Kinetic Isotope Effects. Interpretation of solvent kinetic isotope effects of enzymatic reactions can be complicated by many factors, including effects of D_2O on protein conformation that would affect kinetic commitment factors and therefore report on events that do not reflect proton-transfer chemistry (W. W. Cleland, personal communication). Accordingly, the determination of sKIEs with multiple, structurally similar substrates of widely different kinetic parameters, as in our study here, are expected to give correspondingly different sKIEs on these parameters under conditions where the rates of chemical steps change, while if the observed sKIEs arose solely from changes in protein conformation, one might expect similar values on a panel of substrates of widely different rates of catalysis.

The pD-rate profiles obtained in 93% D_2O for Gly-Arg-AMC and Gly-Ile-AMC are shown in Figure 2, and with data from the fittings of the pD data for Gly-Arg-AMC, Gly-Ile-AMC and Ser-Tyr-AMC found in Table 2. As is typical, the pD-rate profiles are shifted to the right side compared to their counterpart plots in H_2O, which arises from solvent equilibrium isotope effects on the acidic and basic dissociation constants of groups involved in binding and catalysis (Figure 2). The sKIEs for Ser-Tyr-AMC and Gly-Ile-AMC substrates were determined from the pH(D)-independent values of k_{cat}/K_a and k_{cat} parameters found in Table 2, and for Gly-Arg-AMC, from the maximal values of k_{cat}/K and k_{cat} found in the plots of Figure 2. For Ser-Tyr-AMC, we ascertained in the current study, values of $^D(k_{cat}/K_a) = 0.74 \pm 0.05$ and $^Dk_{cat} = 2.3 \pm 0.1$, which are as previously published.[1] In this previous study, we were able to effect solution of simultaneous equations to find suitable, "candidate" values of all individual rate constants ($k_1 - k_{13}$), the intrinsic kinetic isotope effects ($^Dk_5$, $^Dk_9$, and $^Dk_{11}$) and tautomerization of the active-site Cys-234 and His-381 (K_{t1}) for the chemical mechanism of cathepsin C depicted in Scheme 1, from which we were able to recalculate the steady-state kinetic parameters and sKIEs that closely matched the experimental values, and from which we were able to generate reaction time courses that closely matched the experimental data.

Expressions for the steady-state kinetic parameters and the sKIEs for all three substrates, as derived in Schneck et al.,[1] are found in eqs 10−18.

$$k_{cat}/K_a = k_5K_{eq3}/[(1 + K_{t1})K_{ia}(1 + b(1 + a) + c)] \quad (10a)$$

$$K_a = \{[(1 + K_{t1})K_{ia}(1 + b(1 + a) + c)]/k_5K_{eq3}\} \times [k_{ac}/(1 + (k_{ac}/k_{dac}))] \quad (10b)$$

$$k_{cat} = k_{ac}/[(1 + (k_{ac}/k_{dac}))] = k_5K_{eq3}/[(1 + K_{t2})(1 + b + c) + K_{eq3}(1 + c + d) + (k_5K_{eq3}/k_9)(1 + z(1 + y + w) + x(1 + y))] \quad (11)$$

$$k_{ac} = k_5K_{eq3}/[(1 + K_{t2})(1 + b + c) + K_{eq3}(1 + c + d)] \quad (12)$$

$$k_{dac} = k_9/[1 + z(1 + y + w) + x(1 + y)] \quad (13)$$

wherein

$$K_{t1} = [E']/[E],\ K_{t2} = [E'A]/[EA]),\ K_{ia} = k_2/k_1,\ K_{eq3} = k_3/k_4,\ a = k_3/k_2,\ b = k_5/k_4,\ c = k_6/k_7,\ d = k_5/k_7,\ x = k_{10}/k_{11},\ y = k_{12}/k_{13},\ z = k_9/k_{11},\ w = k_{11}/k_{13} \quad (14)$$

$$^Dk_{cat}/K_a = [(1 + K_{t1}/{}^DK_t)/(1 + K_{t1})][{}^DK_5 + b(1 + a) + c{}^DK_{eq5}]/[1 + b(1 + a) + c] \quad (15)$$

$$^Dk_{cat} = [(1 + K_{t2}/{}^DK_t)({}^Dk_5 + b + c{}^DK_{eq5}) + K_{eq3}({}^Dk_5 + c{}^DK_{eq5} + d)] + (k_5K_{eq3}/k_9) \times [{}^Dk_9 + z({}^Dk_{11} + y{}^DK_{eq11} + w) + x({}^Dk_{11}{}^DK_{eq9} + y{}^DK_{eq9}{}^DK_{eq11})] / [(1 + c + b)(1 + K_{t2}) + K_{eq3}(1 + c + d) + (k_5K_{eq3}/k_9)(1 + z(1 + y + w) + x(1 + y))] \quad (16)$$

dx.doi.org/10.1021/bi300719b | Biochemistry 2012, 51, 7551−7568

$$^{D}k_{ac} = [(1 + K_{t2}/^{D}K_{t})(^{D}k_{5} + b + c^{D}K_{eq5}) + K_{eq3}(^{D}k_{5} + c^{D}K_{eq5} + d)] /[(1 + c + b)(1 + K_{t2}) + K_{eq3}(1 + c + d)] \quad (17)$$

$$^{D}k_{dac} = [^{D}k_{9} + z(^{D}k_{11} + y^{D}K_{eq11} + w) + x(^{D}k_{11}{}^{D}K_{eq9} + y^{D}K_{eq9}{}^{D}K_{eq11})] /(1 + z(1 + y + w) + x(1 + y)) \quad (18)$$

In Table 3, we have attempted to ascertain reasonable "candidate values" for the intrinsic solvent kinetic isotope effects and the individual rate constants of the mechanism in Scheme 1 for Gly-Arg-AMC and Gly-Ile-AMC, that when inserted into eqs 10a, 11, 12, and 15−18 would generate values for the four kinetic parameters that then compared favorably to corresponding experimental values. We used some of the candidate values previously obtained for Ser-Tyr-AMC (ref 1, and repeated here in footnotes a and b in Table 3). For Gly-Arg-AMC and Gly-Ile-AMC, the ratio of the Cys-SH/neutral His tautomer (E′) to the Cys-S$^-$/HisH$^+$ form (E) of cathepsin C as given by K_{t1} and K_{t2} was fixed as 0.7.

As discussed above, the rate of the acylation half-reaction of Gly-Arg-AMC is >400 s^{-1}, such that the rate-limiting step of acylation is likely to be the desorption of the AMC product, which was calculated for Ser-Try-AMC to be k_7 = 460 s^{-1}.[1] For an overall value of k_{cat} = 255 s^{-1} for this Type 2 substrate, a value of k_{ac} = 460 s^{-1} would mean that k_{dac} = 640 s^{-1}. Under these conditions, we would expect k_3 and k_5 to be very large for Gly-Arg-AMC, such that the commitment factors k_5/k_4 and k_3/k_2 would be large enough to ablate the effect of the intrinsic isotope effect of $^{D}k_5$ = 3.2 on the value of $^{D}(k_{cat}/K_a)$ = 0.6, a "composite" isotope effect which is dominated by the effect of D_2O on the enrichment of the more active thiolate-imidazolium tautomeric form of the free enzyme. If we select values of K_{t1} = 0.7, $K_{ia} = k_2/k_1$ = 62 μM, k_3 = 8000 s^{-1}, k_4 = 13 s^{-1}, $k_3/k_2 = a$ = 80 and $k_5/k_4 = b$ = 690, and k_7 = 460 s^{-1}, we obtained calculated values of $^{D}(k_{cat}/K_a)$ = 0.75 and k_{cat}/K_a = 1.6 μM^{-1} s^{-1}, which are in good agreement with experimental values. For Gly-Arg-AMC the experimental value of $^{D}k_{cat}$ = 1.6 ± 0.09 is significantly lower than that for Ser-Tyr-AMC ($^{D}k_{cat}$ = 2.3), while the value of k_{cat} for Gly-Arg-AMC is an order of magnitude greater. This suggests that for the more active Gly-Arg-AMC, the chemical steps k_9 and k_{11} are faster in the deacylation half-reaction than for Ser-Tyr-AMC, such that product release contributes more to rate-limitation thereby suppressing the expression of the intrinsic isotope effects $^{D}k_9$ and $^{D}k_{11}$ on $^{D}k_{cat}$. By choosing candidate values of k_9 and k_{11} = 2000 s^{-1}, which are 20-fold higher than those of Ser-Tyr-AMC and greater than the estimated off-rate for the carboxylic product (k_{13} > 1000 s^{-1}), along with $d = k_5/k_7$ = 0.1, and other candidate values above used for k_{cat}/K_a, we calculated candidate values for k_{cat} = 250 s^{-1} and $^{D}k_{ac}$= 1.6, in excellent agreement with the corresponding experimental values.

Surprisingly, for the exceedingly poor, Type 3 substrate, Gly-Ile-AMC, for which one would imagine that chemical steps were rate-limiting in either one or both of the half-reactions, negligible SKIEs were found for $^{D}(k_{cat}/K_a)$ = 1.0 ± 0.5 and $^{D}k_{cat}$ = 1.06 ± 0.04. As determined from the transient kinetics for this substrate, the rate of acylation (k_{ac}) is ≤10% the rate of deacylation (k_{dac}), such that $k_{cat} \sim k_{ac}$ and $^{D}k_{cat} \sim {}^{D}k_{ac}$. The absence of solvent isotope effects on either of these kinetic parameters could mean, among other explanations, that a prechemical step, such as substrate binding or substrate "positioning" may be the slowest step in enzymatic acylation, or that the isotope-insensitive attack of the thiolate group (k_3) on the substrate may be the slowest step. Accordingly, with candidate values of K_{t1} = 0.7, $K_{ia} = k_2/k_1$ = 2300 μM, a very low value of k_3 = 0.7 s^{-1}, k_4 = 10 s^{-1}, and k_5/k_4 = 5, k_3/k_2 = 0.0011, we calculated values of $^{D}(k_{cat}/K_a)$ = 1.02 and k_{cat}/K_a = 0.0015 μM^{-1} s^{-1}, in excellent agreement with experimental results. Similarly, we obtained calculated values of $^{D}k_{cat}$ = 1.02 and k_{cat} = 0.33 s^{-1}, which also agree with experimental results. The calculated values are in accord with low rates of both substrate binding (k_1 = 0.28 μM^{-1} s^{-1}), and of the low rate of attack of the thiolate on the substrate (k_3 = 0.7 s^{-1}), which are purported here to be far slower than the other two substrates. We propose that this is the result of steric hindrance of the β-branched isoleucyl side chain at the P_1 position of this and other similar dipeptide substrates, which impairs both productive binding and the ability of the scissile carbonyl to adopt the correct orientation for nucleophilic attack by the thiolate.

Dipeptide Nitrile−Cathepsin C Complexes. We determined crystal structures of cathepsin C complexed to the nitrile analogues of the Ser-Tyr and Gly-Lys substrates in order to characterize structural mimics of the acyl-enzyme intermediate formed in the catalytic reaction. The cathepsin C-inhibitor complexes all crystallized in the *I*222 space group (Table 4). The general folds and conformations of the peptide backbones are essentially identical to previously published structures.[12,30] Each of the structures is determined with high resolution, for which 1.55 Å was the highest for the outer shell reflections. For

Table 4. Data Collection and Refinement Statistics for Ser-Tyr-CN and Gly-Lys-CN bound to Human Cathepsin C

	(Gly-Lys-CN, **3xgrl**)	(Ser-Tyr-CN, **1pjbl**)
Data Statistics		
space group	*I*222	*I*222
unit cell (*a*, *b*, *c*) (Å)	84.9, 89.0, 115.1	85.5, 89.3, 115.3
mol/ASU	1	1
resolution (Å) (last shell)	50−1.52 (1.57−1.52)	50−1.4 (1.45−1.40)
redundancy	6.1	6.7
I/σ (last shell)	25.9(2.8)	26.6(2.2)
completeness (%) (last shell	99.5(96.5)	98.4(88.7)
R_{sym}[a] (last shell)	5.9(53.1)	6.3/59.0
Wilson B (Å^2)	20.2	19.0
Refinement Statistics		
resolution range (Å) (last shell)	40−1.52(1.56−1.52)	40−1.4(1.44−1.40)
R_{cryst}[b] (%)	19.9(26.5)	21.1(28.6)
R_{free}[c] (%)	20.8(30.3)	22.5(28.7)
rmsd from ideal geometry		
bond length (Å)	0.005	0.005
bond angle (°)	0.888	0.909
average B factor	19.8	18.9
reflection used	59748	74264
total non-hydrogen atoms	3154	3208

[a] $R_{sym} = \Sigma_{hkl} |I - \langle I \rangle| / \Sigma I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity from observations of symmetry-related reflections. [b] $R_{cryst} = \Sigma_{hkl} ||F_{obs}| - |F_{cal}|| / \Sigma |F_{obs}|$, where F_{obs} and F_{cal} are the observed and calculated structure factor amplitudes, respectively, for the *hkl* reflections. [c] R_{free} is calculated from a set of reflections (6%) that were not included in atomic refinement

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

the Ser-Tyr nitrile adduct (K_i = 100 nM), the terminal amino group interacts with the carboxylate on the side chain of Asp-1, the serine side chain of the inhibitor binds into the S_2 pocket, the tyrosine side chain binds in the S_1 pocket with well-defined electron density, and an apparent thioimidate adduct is formed between the cyano group and the active site thiol (Figure 3).

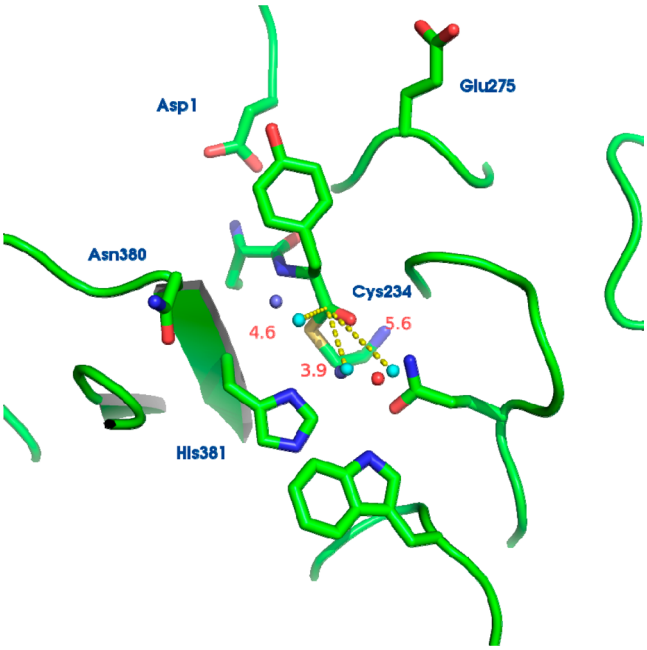


Figure 3. A close-up view of the substrate binding site of cathepsin C acylated by Ser-Tyr-AMC and the computed hydration sites near the thioester bond, shown as cyan and blue spheres for the Ser-Tyr-AMC and Gly-Arg-thioimidate simulations respectively. The crystallographically determined hydration (red sphere) site from the crystal structure of cathepsin C/Ser-Tyr-CN was superimposed to provide as a reference. The distances (in Angstroms) to the reaction center (carbonyl carbon of thioimidate) to the hydration sites from Ser-Tyr-AMC are displayed numerically (red). The water number density of the three sites (from left to right) calculated from the simulation was 0.08, 0.11, and 0.07 per cubic angstroms, respectively.

The nitrogen of the thioimidate group binds in the oxyanion hole of cathepsin C, which is formed from the backbone nitrogen of the active site cysteine (Cys-234) and the side chain of Gln-228.[30] In this structure these groups are 3.3 Å and 3.1 Å, respectively, from the nitrogen of the thioimidate group. The binding of the putative imine in the oxyanion hole suggests that the nitrile inhibitor mimics the thioester acyl-enzyme intermediate of the catalyzed reaction. No significant interactions exist between the phenolic group of the Tyr side chain and the protein. The nearest polar contact from the protein to the hydroxyl group of the side chain Ser at P_2 is the side chain of Asp-1, which are separated by more than 4 Å.

We did not have access to the nitrile analogue of the Gly-Arg substrate, but we were successful in determining the structure of a complex with the Gly-Lys-CN. Given the similar kinetic parameters of Gly-Lys-AMC and Gly-Arg-AMC, the Gly-Lys-CN-cathepsin C complex should be analogous to that of Gly-Arg-CN. The peptide backbone of Gly-Lys-CN binds in the same conformation as the other peptide nitrile ligands. The main difference with this inhibitor is the lysine side chain is best modeled in an extended conformation, protruding directly out of the active site. The electron density for this side chain clearly shows CB and CG in an extended conformation, and CD, CE, and NZ of unknown placement due to lack of electron density suggesting that these atoms are in a highly mobile, disordered state. Even in a fully extended conformation, the ε-amine group of the side chain cannot interact with the γ-carboxylate group of Glu-275, which is 3.7 Å away.

Overall, the crystal structures for the nitrile analogues of the substrates show that the peptide backbones for the Ser-Tyr-CN and Gly-Lys-CN compounds bind as expected but are uninformative with respect to how aromatic amino acids in the P_1 position of substrates contribute to the slower k_{cat} and deacylation rates as compared to those of Type 2 substrates which contain aliphatic, non-beta-branched amino acids. These structures show that the side chain bound in the S_1 pocket extends upward toward the solvent. There is significant electron density observed for the Tyr side chain of Ser-Tyr-CN; however, in the Gly-Lys complex structure the electron density of the Lys side chain is not well defined. Attempts to crystallize Gly-Ile-CN were unsuccessful. Even with cocrystallization conditions for Gly-Ile-CN at a concentration of 100K_i, we only obtained active sites that did not show bound inhibitor.

Inhibition of Cathepsin C by Dipeptide Nitrile Compounds. Analogues of all three substrate types where the C-terminal amide bond was replaced with a nitrile group were made to explore the source of differences in binding and kinetic modalities for the three kinetic types of dipeptide substrates. As shown here and as previously reported,[33] these nitrile-containing compounds react with Cys-234 of cathepsin C to form a thioimidate adduct, which is believed to structurally mimic the acyl-enzyme intermediate. Despite the apparent covalent nature of these compounds, they do not display time-dependent inhibition and as such, the formation of the covalent adducts must be readily reversible (data not shown). The apparent inhibition constants (K_i) were found to be 0.1, 5, and 100 μM for Ser-Tyr-CN, Gly-Lys-CN, and Gly-Ile-CN, respectively (Table 1). The 50-fold difference in values of K_i for Ser-Tyr-CN and Gly-Lys-CN corresponds exactly to the ratio of the Michaelis constants of these two substrates, Ser-Tyr-AMC and Gly-Lys-AMC (Table 1), while this correlation does not extend to Gly-Ile-CN and Gly-Ile-AMC. As we could obtain crystallographic structures demonstrating that the establishment of a thioimidate adduct of both Ser-Tyr-CN and Gly-Arg-CN (analogue of Gly-Lys-CN) with cathepsin C, but not for Gly-Ile-CN, it is likely that Gly-Ile-CN fails to form a thioimidate adduct of cathepsin C, and for this reason, the poor, noncovalent inhibition exerted by Gly-Ile-CN results in a lack of correlation of the values of K_i and K_a for Gly-Ile-CN and Gly-Ile-AMC, respectively, as found for Ser-Tyr-CN/AMC and Gly-Lys-CN/AMC.

P_1-Side Chain Conformational Studies. The distribution of water molecules near the acylated Cys-234, the impact of side chain conformations at the P_1 position, and their proximity to the reaction center may be important factors to characterize the deacylation half-reaction of Ser-Tyr-AMC. Molecular dynamics simulations were carried out for models of the acyl-enzyme intermediates of the Ser-Tyr and Gly-Arg substrates based on the crystal structures of the covalent inhibitors Ser-Tyr-CN and Gly-Lys-CN. Cathepsin C is a relatively compact and rigid protein with five disulfide bonds. During the simulations (about 10 ns each), no significant conformational change was observed for the backbone of the acylated-enzyme. Only 12 out of 354 residues (Glu85 to Lys89, Ile204 to His206, Asp371, Pro372, Leu439, and Glu440) demonstrated signifi-

dx.doi.org/10.1021/bi300719b | Biochemistry 2012, 51, 7551−7568

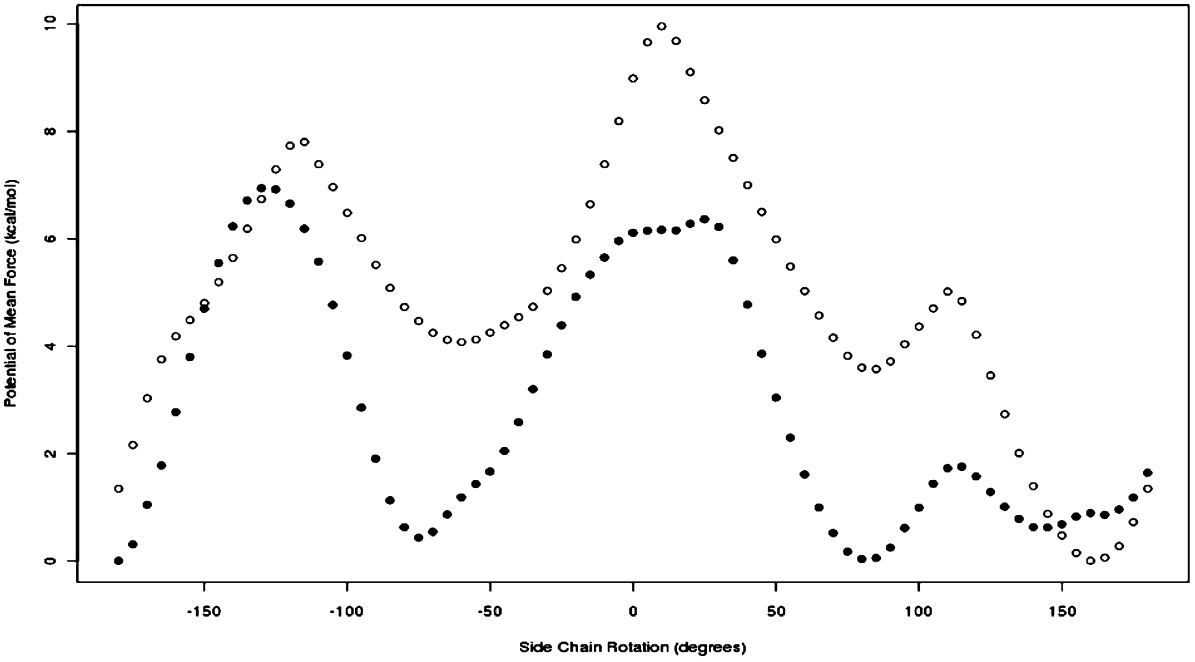


Figure 4. The computed free energy profile of P_1 side chain rotations from adaptive biasing force (ABF) simulations for Ser-Tyr-Cys234 (solid circle, χ1 dihedral angle C-CA-CB-CG) and Gly-Arg-Cys234 (open circle, pseudo dihedral angle C-CA-CB-((CG, CD, NE) center of mass).

cant backbone fluctuation (greater than 1.5 Å). The hydration pattern of cathepsin C and Ser-Tyr-CN complex was well resolved given the high resolution data collected (1.4 Å), which enabled the assignment of 347 crystallographic water molecules during the refinement. The calculated SPCE water density distribution[23] near the surface of the acylated enzyme was able to locate 489 high density sites (with the density greater than 0.05 per cubic Å). Even with the protein thermomotion and the void of crystal contacts in the simulation, there was a significant level of agreement between the MD-calculated sites and the crystallographic sites, 98 of which were reproduced with a discrepancy of 1.0 Å or less. One of the sites reproduced was near the acylated Cys-234 with a density twice that of the bulk water. Two additional water sites that are even closer to the electrophilic carbonyl carbon were also identified with a density of 0.08 and 0.11 per cubic Å, respectively (see Figure 3). This hydration pattern is relatively conserved in the simulation of the Cat C-Gly-Arg-CN complex, especially the water site that is within 3.9 Å of the carbonyl carbon (see Figures 3 and 5). The close proximity of these two water sites potentially poises them in an ideal position to gain access to the reaction center.

During the simulation of the Ser-Tyr-Cys234 complex of cathepsin C, the tyrosine side chain adopts a similar conformation as in the crystal structure with the distribution of the χ1 dihedral centered near 175 degrees. The same distribution was similarly observed for the arginine side chain of Gly-Arg-Cys-234. The conformational preference of the arginine side chain is intuitively reasonable as the result of the strong bias due to the salt-bridge interaction with Glu-275. However, the unimodal distribution of the tyrosine side chain from a brief molecular dynamics simulation needs to be further investigated as the sampling of other rotameric states can be kinetically hindered. Therefore, to fully explore the conformational space of the amino acid side chain at the P_1 position, adaptive biasing force (ABF) was applied to the χ dihedral angle of the tyrosine side chain of Ser-Tyr-Cys234, and similarly to a pseudo dihedral angle of the arginine side chain of Gly-Arg-Cys-234, defined as C-CA-CB-(CG, CD, NE), in which the mass center of atom CG, CD, and NE was used. The computed potential of mean force (PMF) along the dihedral angle provided an estimate of the free energy of different conformational states. Comparing the PMF profiles of the tyrosine and arginine side chain rotations (Figure 4), the bias from interacting with the negative charge of Glu-275 on the arginine side chain is easily visible, with the trans conformation almost 3.0 kcal/mol more favored than two other minima (near −65 and 90 deg). On the other hand, the tyrosine side chain may populate two equally favorable conformational states (near −75 and 85 deg) other than the trans. The rotational transition among these states required the crossing of substantial free energy barriers, which resulted in the kinetic hindrance and the unimodal distribution of the tyrosine side chain in the regular molecular dynamics simulation. Structurally, the difficulty in transitioning from the trans conformation to the state near −75 deg is possibly originated from the blockage of the side chain of Asn-380, while the transition to the state near 85 deg is largely free of such steric impediments, hence smaller "penalties" of free-energy. A multitude of kinetically isolated states of the P_1 tyrosine may have a negative impact on the deacylation rate if only a limited number of states are productive and the interstate transition is slow. One possibility is that water is blocked from accessing the reaction center in some of the states identified by the ABF calculations. Data to support this idea were readily accessed by calculating the hydration pattern in the region with the motion of the tyrosine side chain restricted to the individual states. As shown in Figure 5, the change to the hydration pattern near the acylated cysteine was most noticeable when the tyrosine side chain adopted the −75 deg state. The two nearest water sites to the acylated cysteine carbonyl were eliminated entirely, presumably trapping the enzyme in a nonproductive state, and thereby reducing the rate of deacylation. For Gly-Arg-Cys-234, such a difficulty was largely removed by the conformational control from Glu-275, which ensures the access of water molecules to the reaction

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

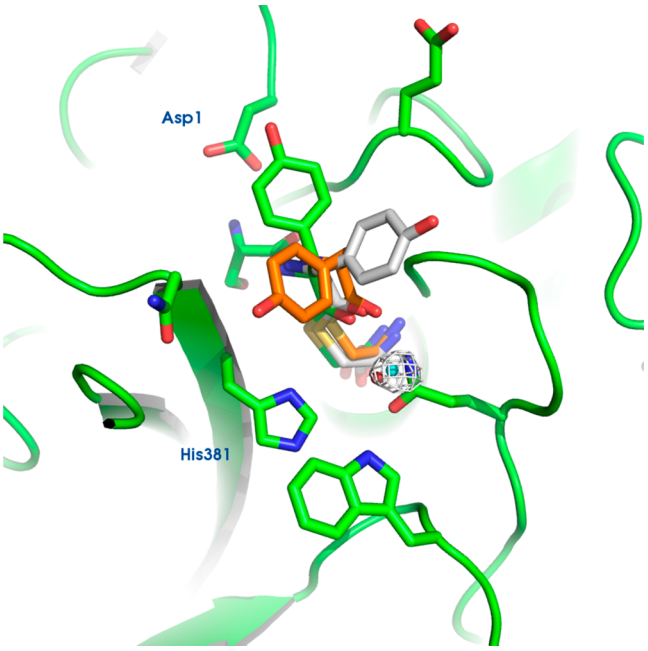


Figure 5. The alternative side chain conformations of the P_1 tyrosine computed from the ABF simulation, (1) −75 deg (orange), (2) 85 deg (white), and (3) 180 deg (green, original). Only one hydration site remained when the tyrosine side chain adopted the −75 deg state, shown as cyan sphere surrounded by an iso-density mesh contoured at 0.05 per cubic Å. This is the site evident in most cathepsin C crystal structures.

center. The different sKIE values of the two substrates may reflect this difference in catalytic water accessibility or reactivity: $^{D}k_{cat}$ = 2.3 for Ser-Tyr-AMC and 1.6 for Gly-Arg-AMC, which suggests that the attack of the lytic water molecule is more facile, and thereby less rate-limiting, for the latter substrate.

Kinetic Comparison of the Dipeptide Substrates. Figure 6 is a proposed free-energy diagram of the acylation half-reaction to summarize and compare the kinetics of the three substrate types as catalyzed by cathepsin C. The relative free-energy levels of the transitory enzyme complexes are based on the free energies of the equilibrium constants of each catalytic step, for example, $K_{eq3} = k_3/k_4$ for the EA-to-EX step, from the candidate values found for the three substrates in Table 3. The free energies of activation are graphically estimated from the solved rate constants shown in Table 3 for concentrations of substrates set at their Michaelis constants. In the case of the Type 1 substrate Ser-Tyr-AMC (red), the proclivity of the tetrahedral adduct of Cys-234 and the substrate in the EX complex to progress through the isotope-sensitive k_5 step rather than return to the Michaelis complex (EA) leads to a suppression, albeit, not ablation, of the intrinsic sKIE ($^{D}k_5$ = 3.2) on this step. The slowest step in the acylation half-reaction for Ser-Tyr-AMC was found to be release of AMC product (k_7),[1] as depicted by the highest kinetic barrier for this half-reaction. The inverse isotope effect of $^{D}(k_{cat}/K_a)$ = 0.74 for this substrate results from an increased amount of the thiolate-imidazolium form of the free enzyme (E) in D_2O. With this form of the enzyme the expression of the normal kinetic isotope effect of the k_5 step is abrogated by the high free-energy barriers on either side of this kinetic step.

For the more active, Type 2 substrate, Gly-Arg-AMC (green), the free-energy diagram is similar to that of Ser-Tyr-AMC, except that the largest kinetic barrier for both half-reactions is the k_7 step, describing desorption of the AMC product, while for Ser-Tyr-AMC, the kinetic barriers are larger for the catalytic k_9 and k_{11} deacylation steps (deacylation half-reaction not shown). More than with Ser-Tyr-AMC, there is negligible impact of the intrinsic isotope effect $^{D}k_5$ = 3.2 on $^{D}(k_{cat}/K_a)$ = 0.6, and the lower value of $^{D}k_{cat}$ = 1.6 that is observed with Gly-Arg-AMC results because the desorption of the final carboxylic product via the k_{13} step is similar in rate to that of the catalytic k_9 and k_{11} steps.

The lower stability of the E-Gly-Ile-AMC complex (blue in Figure 6) and the high energy barrier required to form a

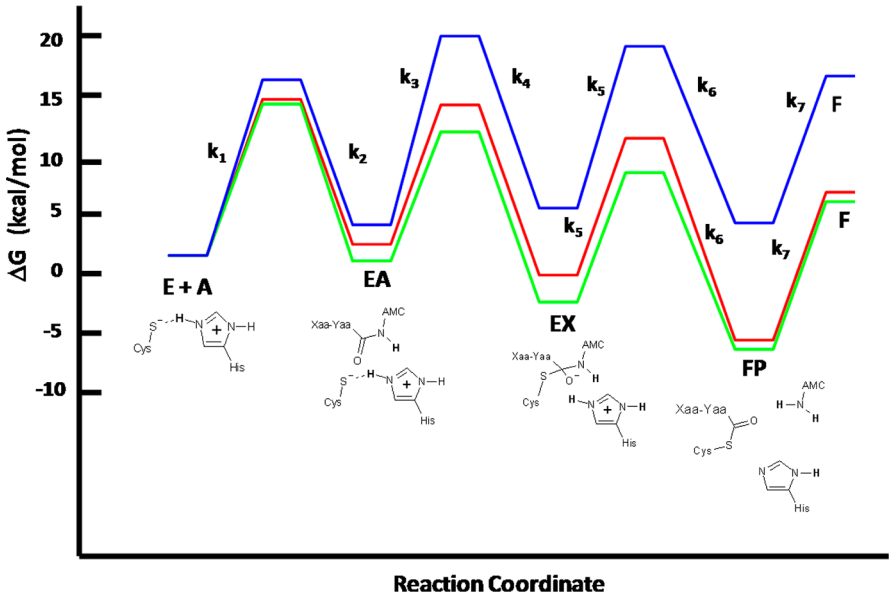


Figure 6. Free energy diagrams for the acylation half-reaction of Cathepsin C with the substrate Ser-Tyr-AMC (red), Gly-Arg-AMC (green), and Gly-Ile-AMC (blue). Free energy levels of enzyme forms were determined from the candidate values found in Table 3; for example, for EX, $\Delta G = -nRT \ln(k_3/k_4)$. Substrate concentrations used were those of the Michaelis constants for all three substrates.

dx.doi.org/10.1021/bi300719b | *Biochemistry* 2012, 51, 7551−7568

thioester with this substrate (k_3) suppress the inverse isotope effect afforded by the tautomerization of free enzyme in D_2O as expressed in the $^D(k_{cat}/K_a)$ parameter. Diminution of the rate of the k_5 step in D_2O is inadequate to raise the free energy barrier of the k_5 step to offset the commitment rendered by the preceding, isotope-insensitive k_3 step, leading to an absence of an observable sKIE on k_{cat}/K_a. The k_{cat} of this substrate is equivalent to the rate of acylation (k_{ac}), such that $^D(k_{cat}/K_a) = {}^Dk_{cat} = 1$.

Since structural studies revealed no specific interactions between the side chain of the amino acid bound at P_1 and the enzyme, it is possible that the relationship between the structure of the amino acid side chain at P_1 and its relative mobility while bound to the enzyme has a role in the rates of catalysis. Type 2 substrates with Ala, Leu, Glu, Arg, and Lys at the P_1 position have extremely fast acylation and deacylation steps demonstrated by "transient-free", pre-steady-state bursts observed for this substrate type. The P_1 side chains of these substrates likely bind in conformations that do not interfere with the catalytic events that elaborate higher energy barriers, such as attack of the active site thiol or the activation of water during the deacylation reaction. The lack of interference by these side chains is supported by the molecular dynamics studies with the Gly-Arg-Cys-234 intermediate. These calculations show the Arg side chain of the substrate is distal from the active-site thiol and does not interfere with the probability of water binding properly for deacylation to occur.

However, this may differ for the aromatic amino acids of Type 1 substrates. The planar aromatic rings could cause the side chains to sweep around a wide area as various rotamer conformations are sampled. In contrast, the Arg side chain has four adjacent freely rotatable bonds with fewer conformational restraints, so that many conformations can be sampled without occluding the attack of water.

The presence of β-branched amino acids (e.g., Ile and Val) in the S_1 subsite decreases the rate of the acylation step to less than that of deacylation. This is apparent from the presteady-state kinetics in that the rate of the transient phase appears to merge with the steady-state phase. However, from molecular modeling, the structural reason(s) that the acylation half-reaction becomes rate-limiting with these substrates is not evident. One possible reason is that for acylation to occur the carbonyl of the scissile bond needs to be in the correct orientation, and with the Val and Ile side chains the substrates bind with the carbonyl group in an unproductive conformation for catalysis.

For all of the kinetic, structural, and computational data of the three classes of substrates of cathepsin C investigated here, we may summarize from the data in Tables 1 and 4 that in general, (a) the presence of a Ser as opposed to a Gly residue at P_2 leads to a faster value k_1 or a slower value of k_2 (and therefore lower K_{ia}) value for the substrate, reflected in the 10-fold difference in k_{cat}/K_a for Gly-Tyr-AMC and Ser-Tyr-AMC, but with comparable values of k_{ac} and k_{dac}; (b) as substituents at the β- and γ-carbon atoms of the P_1 amino acid become larger than a methylene group, the value of k_{cat} decreases. This suggests that the highest rates of deacylation are achieved when there is mimimal steric interference to hydrolysis from these carbon atoms, as demonstrated by the kinetic data and suggested by current computational analysis using the available structures of enzyme with Ser-Tyr-CN and Gly-Lys-CN.

The relationship between the substrate(s) existing in the correct orientation for catalysis and the rate of the catalyzed reaction has been discussed in the literature and has been referred to by several ideas including, but not limited to, proximity and orientation.[34−37] The near-attack conformation model, which relates the probability of reacting groups being in a near-attack conformation to the rate of the reaction that occurs[38] fits well with our data, suggesting that the binding of the substrate in a nonproductive conformation, such as with Gly-Ile-AMC, or the occlusion of water by the substrate side chain, as seen in the molecular dynamics simulation with Ser-Tyr-Cys-234, decreases the probability of a near-attack conformation and increases in the energy barriers for the chemical steps (Figures 4 and 5).

Physiologically, cathepsin C activates serine proteases including neutrophil elastase, cathepsin G, proteinase 3, chymase, tryptase, and granzyme A and B.[2−5] The scissile dipeptides in the pro-sequences are similar for these serine proteases: Gly-Glu in Pro-Cathepsin G, Pro-granzyme B, and Pro-chymase; Ser-Glu in Pro-elastase; Ala-Glu in Pro-Proteinase 3; Glu-Lys in Pro-Granzyme A. The pro-sequence in each of these cases corresponds to the highly active type 2 substrates, where both acylation and deacylation are very fast and where desorption of the AMC product may be overall rate-limiting for the reaction. Thus, the kinetics of dipeptide substrates such as Gly-Glu-AMC, Gly-Arg-AMC, and Gly-Lys-AMC may be more representative of the in vivo kinetics of cathepsin C for the efficient activation of the pro-serine proteases into their mature, active forms. Likewise, each of these serine proteases contains a "stop-sequence", that is, a dipeptide that terminates the stepwise processing of a protein substrate by cathepsin C, preventing untoward destruction of the activated protease. This stop sequence is also highly conserved among these serine proteases, with amino acids following the Pro-sequence of Ile-Ile in granzyme A and B, cathepsin G, and chymase and Ile-Val in proteinase 3 and elastase.[2] Our present studies demonstrate that such dipeptides would comprise very poor substrates for human cathepsin C.

Cysteine proteases are challenging drug targets. Most efforts in designing new inhibitors for cathepsin C have centered on the use of electrophilic compounds with reactive groups that react with the active-site cysteine.[10] Unfortunately these types of reactive compounds are associated with poor metabolic stability and safety concerns.[10] Prior studies have demonstrated high fractional inhibition of cathepsin C (>50%, depending on the substrate) must be achieved to see the desired pharmacokinetic effect,[8] speaking to the need for fast-acting, long-duration inhibitors of cathepsin C as opposed to freely reversible inhibitors.

In this study, we have found substrates that display very slow rates of both acylation and deacylation, which may allow for the design of pseudosubstrates that effect avid binding, facile acylation, but then display a poor rate of deacylation. Such a substrate could function as a potent inhibitor in vivo. Indeed, it would be ideal to develop a dipeptide, pseudosubstrate which has the kinetic properties of Gly-Arg-AMC and Ser-Tyr-AMC for the acylation half-reaction, but those of Gly-Ile-AMC for the deacylation half-reaction. It might be possible to achieve something similar with a compound that reacts with cathepsin C or another papain-like cysteine protease and blocks the attack of water in the deacylation step. This suggests that the Ser-Tyr-CN "inhibitor" may actually be a good starting point for inhibitor design. This type of inhibition has been accomplished in nature with the inhibition of caspase-8 by the protein p35.[39] This protein forms a covalent bond with the active site cysteine

dx.doi.org/10.1021/bi300719b | Biochemistry 2012, 51, 7551−7568

of caspase-8 and the carbonyl of Asp87 in p35. It is thought that hydrolysis of this covalent caspase-8-p35 adduct is blocked by conformational exclusion.[39] One therapeutically successful example of this type of inhibitor is rivastigmine, a slowly reversible covalent inactivator of acetylcholinesterase[40] used in the treatment of cognitive disorders. Rivastigmine competes with substrate in its extremely rapid carbamoylation of the active-site serine of acetylcholinesterase. The carbamoylated form of the enzyme then slowly hydrolyzes to restore active enzyme. Adaptation of some of the structural attributes of rivastigmine to the dipeptide substrates of cathepsin C may afford slowly reversible inactivators of this enzyme.

■ ASSOCIATED CONTENT

Supporting Information

Descriptions of the synthesis and chemical characterization of Gly-Ile-nitrile and Gly-Lys-nitrile inhibitors of human cathepsin C via published procedures as described. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University New Brunswick, NJ (http://www.rcsb.org/).

■ AUTHOR INFORMATION

Corresponding Author

*Phone (610) 917-6786; fax: (610) 422-2636; e-mail: thomas.meek@gsk.com.

Present Address

[#]Kemin Industries, 2100 Maury St., Des Moines, IA 50306.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Professor Paul Cook for helpful comments and Dr. Gaochao Tian for assistance with the construction of the free-energy diagram.

■ ABBREVIATIONS USED

AMC, 7-amino-4-methyl-coumarin; MES, 2-(*N*-morpholino) ethanesulfonic acid; DTT, dithiothreitol; CHAPS, 3-[3-(cholamidopropyl)dimethylammonio]-1-propanesulfonate; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; k_{ac}, rate of acylation half-reaction; k_{dac}, rate of deacylation half-reaction; sKIE, solvent kinetic isotope effect(s); Ala-Hph-VS-Ph, 3-(*N*-alanyl)amino-5-phenyl-1-(phenylsulfonyl)-1-pentene (hydrochloride)

■ ADDITIONAL NOTES

[a]P_1, P_2, S_1, and S_2 refer to the amino-acid residues and enzyme subsites, respectively, using the convention of Schechter & Berger.[9]

[b]We repeated evaluations of the steady-state, pre-steady-state, pH-rate profiles, and solvent kinetic isotope effects of Ser-Tyr-AMC as found in ref 1 as a baseline of comparison, and for consistency using different enzyme preparations for the current panel of substrates. We have also included additional data for this substrate.

[c]From examination of eq 10b, it can be shown that $K_a = [k_{ac}/(1 + (k_{ac}/k_{dac}))]K_{a'}$ where $K_{a'} = [(1 + K_{t1})K_{ia}(1 + b(1 + a)) + c]/k_5K_{eq\ 3}]$, and that $K_a = k_{dac}K_{a'}$ when $k_{ac} \gg k_{dac}$, $K_a = (k_{ac}/2)K_{a'}$ when $k_{ac} \sim k_{dac}$, and $K_a = K_{a'}$ when $k_{dac} \gg k_{ac}$.

■ REFERENCES

(1) Schneck, J., Villa, J., McDevitt, P., McQueney, M., Thrall, S., and Meek, T. (2008) Chemical mechanism of a cysteine protease, cathepsin C, as revealed by integration of both steady-state and pre-steady-state solvent kinetic isotope effects. *Biochemistry 47*, 8697−8710.

(2) Tran, T. (2002) Dipeptidyl peptidase I: importance of progranzyme activation sequences, other dipeptide sequences, and the N-terminal amino group of synthetic substrates for enzyme activity. *Arch. Biochem. Biophys. 403*, 160−170.

(3) Adkison, A., Raptis, S., Kelley, D., and Pham, C. (2002) Dipeptidyl peptidase I activates neutrophil-derived serine proteases and regulates the development of acute experimental arthritis. *J. Clin. Invest. 109*, 363−371.

(4) Pham, C. T., and Ley, T. J. (1999) Dipeptidyl peptidase I is required for the processing and activation of granzymes A and B in vivo. *Proc. Natl. Acad. Sci. U. S. A. 96*, 8627−8632.

(5) Wolters, P. J., Pham, C. T., Muilenburg, D. J., Ley, T. J., and Caughey, G. H. (2001) Dipeptidyl peptidase I is essential for activation of mast cell chymases, but not tryptases, in mice. *J. Biol. Chem. 276*, 18551−18556.

(6) Kim, S., and Nadel, J. (2004) Role of neutrophils in mucus hypersecretion in COPD and implications for therapy. *Treat. Respir. Med. 3*, 147−159.

(7) Methot, N., Rubin, J., Guay, D., Beaulieu, C., Ethier, D., Reddy, J., Riendeau, D., and Percival, D. (2007) Inhibition of the activation of multiple serine proteases with a cathepsin C inhibitor requires sustained exposure to prevent pro-enzyme processing. *J. Biol. Chem. 282*, 20836−20846.

(8) Methot, N., Guay, D., Rubin, J., Ethier, D., Ortega, K., Wong, S., Normandin, D., Beaulieu, C., Reddy, J., Riendeau, D., and Percival, D. (2008) In vivo inhibition of serine protease processing requires a high fractional inhibition of cathepsin C. *Mol. Pharmacol. 73*, 1857−1865.

(9) Berger, A., and Schechter, I. (1970) Mapping the active site of papain with the aid of peptide substrates and inhibitors. *Philos. Trans. R. Soc. London, B 257*, 249−264.

(10) Laine, D., and Busch-Petersen, J. (2010) Inhibitors of cathepsin C (dipeptidyl peptidase I). *Expert Opin. Ther. Patents 20*, 497−506.

(11) Dolenc, I., Turk, B., Pungercic, G., Ritonja, A., and Turk, V. (1995) Oligomeric structure and substrate induced inhibition of human cathepsin C. *J. Biol. Chem. 270*, 21626−21631.

(12) Molgaard, A., Arnau, J., Lauritzen, C., Larsen, S., Petersen, G., and Pedersen, J. (2007) The crystal structure of human dipeptidyl peptidase I (cathepsin C) in complex with the inhibitor Gly-Phe-CHN2. *Biochem. J. 401*, 645−650.

(13) Palmer, J. T., Rasnick, D., Klaus, J. L., and Bromme, D. (1995) Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J. Med. Chem. 38*, 3193−3196.

(14) Ellis, K. J., and Morrison, J. F. (1982) Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol. 87*, 405−426.

(15) Schowen, K. B., and Schowen, R. L. (1982) Solvent isotope effects of enzyme systems. *Methods Enzymol. 87*, 551−606.

(16) Cook, P. F., and Cleland, W. W. (1981) Mechanistic deductions from isotope effects in multireactant enzyme mechanisms. *Biochemistry 20*, 1790−1796.

(17) Northrop, D. B. (1977) Determining the Absolute Magnitude of Hydrogen Isotope Effects, in *Isotope Effects on Enzyme-Catalyzed Reactions* (Cleland, W. W., O'Leary, M. H., and Northrop, D. B., Eds.) pp 122−152, University Park Press, Baltimore.

(18) Halgren, T. (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem. 17*, 490−519.

(19) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and

development of improved protein backbone parameters. *Proteins 65*, 712−725.
(20) Wang, J., Wang, W., Kollman, P., and Case, D. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell. 25*, 247−260.
(21) Cieplak, P., Cornell, W., Bayly, C., and Kollman, P. (1995) Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem. 16*, 1357−1377.
(22) Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., DClifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) *Gaussian 03*, Revision C.02, Gaussian, Inc., Wallingford, CT.
(23) Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The missing term in effective pair potentials. *J. Phys. Chem. 91*, 6269−6271. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(24) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
(25) Evans, D. J., and Holian, B. L. (1985) The Nose-Hoover thermostat. *J. Chem. Phys. 83*, 4069−4074.
(26) Henin, J., and Chipot, C. (2004) Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys. 121*, 2904−2914.
(27) Darve, E., Rodriguez-Gomez, D., and Pohorille, A. (2008) Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys. 128*, 144120.
(28) Phillips, J., Zheng, G., Kumar, S., and Kale, L. (2002) NAMD: Biomolecular Simulation on Thousands of Processors, in *Proceedings of SC 2002*, pp 1−18.
(29) Phillips, J., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R., Kalo, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(30) Turk, D., Janjio, V., Stern, I., Podobnik, M., Lamba, D., Dahl, S. W., Lauritzen, C., Pedersen, J., Turk, V., and Turk, B. (2001) Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J. 20*, 6570−6582.
(31) Cleland, W. W. (1977) Determining the chemical mechanisms of enzyme-catalyzed reactions by kinetic studies. *Adv. Enzymol. Relat. Areas Mol. Biol. 45*, 273−387.
(32) Cook, P. F., Kenyon, G. L., and Cleland, W. W. (1981) Use of pH studies to elucidate the catalytic mechanism of rabbit muscle creatine kinase. *Biochemistry 20*, 1204−1210.
(33) Guay, D., Beaulieu, C., Truchon, J. F., Jagadeeswar Reddy, T., Zamboni, R., Bayly, C., Methot, N., Rubin, J., Ethier, D., and vid Percival, M. (2009) Design and synthesis of dipeptidyl nitriles as potent, selective, and reversible inhibitors of cathepsin C. *Bioorg. Med. Chem. Lett. 19*, 5392−5396.
(34) Bruice, T. C., and Benkovic, S. J. (2000) Chemical basis for enzyme catalysis. *Biochemistry 39*, 6267−6274.
(35) Bruice, T., and Lightstone, F. (1998) Ground state and transition state contributions to the rates of intramolecular and enzymatic reactions. *Acc. Chem. Res. 32*, 127−136.
(36) Cannon, W. R., Singleton, S. F., and Benkovic, S. J. (1996) A perspective on biological catalysis. *Nat. Struct. Biol. 3*, 821−833.
(37) Mesecar, A. D., Stoddard, B. L., and Koshland, D. E. (1997) Orbital steering in the catalytic power of enzymes: small structural changes with large catalytic consequences. *Science (New York, N. Y.) 277*, 202−206.
(38) Bruice, T. (2002) A view at the millennium: the efficiency of enzymatic catalysis. *Acc. Chem. Res. 35*, 139−148.
(39) Xu, G., Cirilli, M., Huang, Y., Rich, R. L., Myszka, D. G., and Wu, H. (2001) Covalent inhibition revealed by the crystal structure of the caspase-8/p35 complex. *Nature 410*, 494−497.
(40) Bar-On, P., Millard, C. B., Harel, M., Dvir, H., Enz, A., Sussman, J. L., and Silman, I. (2002) Kinetic and structural studies on the interaction of cholinesterases with the anti-Alzheimer drug rivastigmine. *Biochemistry 41*, 3555−3564.