

Mechanism of Inhibition of Cathepsin K by Potent, Selective 1,5-Diacylcarbohydrazides: A New Class of Mechanism-Based Inhibitors of Thiol Proteases

Mary J. Bossard,[‡] Thaddeus A. Tomaszek,[‡] Mark A. Levy,^{‡,§} Carl F. Ijames,^{||,⊥} Michael J. Huddleston,^{||} Jacques Briand,^{||} Scott Thompson,[⊗] Stacie Halpert,[⊗] Daniel F. Veber,[⊗] Steven A. Carr,^{||} Thomas D. Meek,[‡] and David G. Tew^{*,‡}

Departments of Molecular Recognition, Physical and Structural Chemistry, and Medicinal Chemistry, SmithKline Beecham Pharmaceuticals, P.O. Box 1539, King of Prussia, Pennsylvania 19406

Received May 24, 1999; Revised Manuscript Received September 23, 1999

ABSTRACT: The nature of the inhibition of thiol proteases by a new class of mechanism-based inhibitors, 1,5-diacylcarbohydrazides, is described. These potent, time-dependent, active-site spanning inhibitors include compounds that are selective for cathepsin K, a cysteine protease unique to osteoclasts. The 1,5-diacylcarbohydrazides are slow substrates for members of the papain superfamily with inhibition resulting from slow enzyme decarbamylation. Enzyme-catalyzed hydrolysis of 2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucinyllcarbohydrazide is accompanied by formation of a hydrazide-containing product and a carbamyl-enzyme intermediate that is sufficiently stable to be observed by mass spectrometry and NMR. Stopped-flow studies yield a saturation limited value of 43 s⁻¹ for the rate of cathepsin K acylation by 2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucinyllcarbohydrazide. Inhibition potency varies among proteases tested as reflected by 2–3 orders of magnitude differences in *K_i* and *k_{obs}/I*, but all eventually form the same stable covalent intermediate. Reactivation rates are equivalent for all enzymes tested (1 × 10⁻⁴ s⁻¹), indicating hydrolysis of a common carbamyl-enzyme form. NMR spectroscopic studies with cathepsin K and 2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucinyllcarbohydrazide provide evidence of inhibitor cleavage to generate a covalent carbamyl-enzyme intermediate rather than a tetrahedral complex. The product Cbz-Leu-hydrazide does not appear enzyme-bound after cleavage in the NMR spectra, suggesting that the stable inhibited form of the enzyme is the thioester complex. 1,5-Diacylcarbohydrazides represent a new class of unreactive cysteine protease inhibitors that share a common mechanism of action across members of the papain superfamily. Both S and S' subsite interactions are exploited in achieving high selectivity and potency.

Bone remodeling is a dynamic process involving continual deposition and resorption of bone matrix. The key cell type involved in the resorptive process is the osteoclast. Osteoclasts attach to the bone surface, resulting in the formation of a vacuole at the interface. This vacuole is the site of active bone resorption being made both acidic and rich in protease(s) by the osteoclast (1, 2). It is assumed that this protease-rich environment is responsible for matrix degradation during the resorption process. Studies with the generic cysteine protease inhibitors such as E-64, leupeptin, and cystatin have implicated a cysteine protease activity as being important in bone resorption (3–7). However, the identity of this protease has, until recently, remained elusive.

The identification of a cysteine protease of the papain superfamily, which is both highly and selectively expressed in osteoclasts, has suggested that this protease, cathepsin K,

is the enzyme responsible for matrix degradation during bone resorption by the osteoclast (8). Additional support for the role of cathepsin K in bone resorption has come from the linkage of a genetic defect of bone resorption, pycnodysostosis, being linked to mutations in the cathepsin K gene (9–11) that render cathepsin K inactive as a protease.

Several classes of mechanism-based inhibitors of cysteine proteases have been reported including aldehydes, chloromethyl ketones, (12) epoxides (13), and vinyl sulfones (14). Generally, these inhibitors share two common properties. First, they possess an inherently reactive electrophilic functional group, and second, they generally utilize interactions only toward the S side of the active site, thereby failing to take advantage of potential S' subsite interactions for enhancing potency and selectivity (15). Recently, potent inhibitors of cathepsin K based on a 1,3-diamino-2-propanone scaffold have been reported that span both S and S' sides of the active site (16, 17). However, these inhibitors also rely on the electrophilic nature of a ketone to engage the active site cysteine.

We report here detailed mechanistic studies of inactivation of cathepsin K by potent selective 1,5-diacylcarbohydrazides

* To whom correspondence should be addressed. Tel: (610)270-6198. Fax: (610)270-4094. E-mail: David_Tew@sbphrd.com.

[‡] Department of Molecular Recognition.

[§] Current address: Department of DMPK.

^{||} Department of Physical and Structural Chemistry.

[⊥] Current address: Polychip, Inc. Bethesda, MD.

[⊗] Department of Medicinal Chemistry.

(18). These inhibitors utilize interactions at both the S and S' substrate binding subsites to achieve potency and selectivity while the carbonylhydrazone functional group lacks intrinsic reactivity toward nucleophiles. The mechanism of inactivation is demonstrated to involve interaction with the active site cysteine (Cys-25) (19, 20) along with slow enzyme reactivation upon hydrolysis of the covalent intermediate. In addition, we show that the mechanism of action of these compounds is common to several members of the papain superfamily and therefore represents a new class of intrinsically unreactive mechanism based inhibitors of thiol proteases.

EXPERIMENTAL PROCEDURES

Materials. Recombinant human cathepsin K and other cathepsins were prepared as reported previously (21, 22). Peptide substrates were obtained from Bachem or Nova Biochem. Compound synthesis including $^3\text{H}_2$ -labeled 2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucylcarbohydrazide was reported previously (18).

Methods. Standard assay conditions and inhibition studies were performed in 100 mM sodium acetate, pH 5.5, 5 mM EDTA, and 5 mM cysteine as per Votta et al. (22). Cbz-Leu-Arg-AMC¹ was used as the substrate unless otherwise specified. Product fluorescence was monitored with a PerSeptive Biosystems (Framingham, MA) Cytofluor II plate reader (excitation/emission 360/460 nm). All curve fitting was achieved using Grafit 3.0 (23).

Inactivation of cathepsin K and other enzymes was performed in a 96-well plate format as described previously (22). Briefly, substrate and inhibitor were premixed in each well. Substrate was at K_m (final) for each cathepsin, and inhibitor concentrations were varied as appropriate depending on potency. Reactions were initiated by the addition of enzyme. Final enzyme concentrations were around 1 nM except for cathepsin K, which was 100 pM. Progress curves were fitted to

$$[\text{product}] = v_{ss}t + (v_o - v_{ss})[1 - \exp(-k_{obs}t)]/k_{obs} + A_o \quad (1)$$

where v_o is the initial reaction velocity, v_{ss} is the final steady-state rate, A_o is the background fluorescence, and k_{obs} is the rate of inactivation. The parameter v_o was fitted as a function of $[I]$ to eq 2b with the assumption that the inhibitor is purely competitive to obtain a K_i value for formation of the initial reversible EI complex prior to inactivation (24):

$$v_o = V_{max}[S]/(K_m(1 + [I]/K_i) + [S]) \quad (2a)$$

Since the inactivation was done under conditions of $[S] = K_m$, the equation was reduced to

$$v_o = V_{max}/(2 + [I]/K_i) \quad (2b)$$

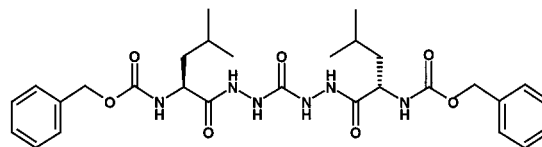
The initial slope of a simple linear fit of a secondary plot of k_{obs} vs $[I]$ yielded an apparent second-order rate constant k_{obs}/I . Uncorrected values are reported simply as k_{obs}/I . When this value was then corrected for substrate concentration and

the Michaelis constant according to eq 3, a true second-order rate constant (k_{inact}/K_i) was determined:

$$k_{obs}/[I] = (k_{inact}/K_i)(1 + [S]/K_m) \quad (3)$$

A complete discussion of this kinetic treatment has been presented (24–26), and its application to elastase inhibition has been described by Knight et al. (27).

Stopped-Flow Kinetic Studies. Cathepsin K (100 nM) was rapidly mixed with various concentrations of **1** in an Applied



2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucylcarbohydrazide

Photophysics SX.18MV stopped-flow spectrophotometer. The reaction was performed using standard assay buffer at pH 5.5 containing 10% DMSO in both drive syringes. All reactions were done at room temperature under pseudo-first-order conditions with **1** in excess of cathepsin K. Total tryptophan fluorescence was monitored by excitation at 295 nm with a 320 nm cutoff emission filter. Individual rate constants (k_{obs}) were determined using the instrument software to fit a single-order exponential with a floating endpoint. Values of k_{obs} at low $[I]$ were fit to the linear eq 4a to obtain a value for k_{on} (25, 26):

$$k_{obs} = k_{off} + k_{on}[I] \quad (4a)$$

Values of k_{obs} were then fit as a function of $[I]$ using eq 4b to obtain values of both K_i and k_{inact} (28):

$$k_{obs} = k_{inact}[I]/(K_i + [I]) \quad (4b)$$

Reactivation Studies. Release of inhibitor from inhibited enzyme was followed by monitoring the return of enzyme activity as a function of time as previously described for cathepsin K (22). The dilution buffer contained 50 μM Cbz-Leu-Arg-AMC as a substrate for cathepsin K and papain. The substrate Cbz-Val-Val-Arg-AMC was used for cathepsin S. Preincubation was carried out for 2 h using 1 μM enzyme and 2 μM inhibitor followed by final dilution to 50 pM enzyme and 100 pM inhibitor into substrate. A SPEX Fluorolog 3 fluorimeter (Instruments SA, Edison, NJ) was used to monitor the return of enzyme activity at 30 °C. Data were fitted using eq 4 (29, 30):

$$F = v_{ss}t + (v_o - v_{ss})[1 - \exp(-k_{react}t)]/k_{react} + F_o \quad (5)$$

where F is the observed fluorescence, F_o is the initial background fluorescence at the start of the assay, v_o is the initial velocity following dilution of the preincubated enzyme–inhibitor complex, v_{ss} is the steady-state rate of product formation after reactivation, and k_{react} is the rate of reactivation.

Mass Spectrometric Analysis of Inactivated Cathepsin K. Protein samples were analyzed by loading 100 μL of solution containing 1000 pmol of protein in assay buffer onto a 1 mm \times 5 mm reverse-phase peptide trap (Michrom BioResources, Inc., Auburn, CA) and washing with 100 μL of solvent A before back-elution into a T connector that directed

¹ Abbreviations: AMC, aminomethylcoumarin; Cbz, benzyloxycarbonyl; DMSO, dimethyl sulfoxide; HPLC, high-performance liquid chromatography; DSS, sodium 2,2-dimethyl-2-silapentane-5-sulfonate.

55 $\mu\text{L}/\text{min}$ to a UV detector/fraction collector and 5 $\mu\text{L}/\text{min}$ to the IonSpray source of a triple-quadrupole mass spectrometer (Sciex API III, Sciex, Thornhill, ON, Canada). Elution was with a gradient beginning with a 2-min hold at 20% B, followed by a ramp to 80% B in 2 min, and then a final 2-min hold. Solvent A was 0.1% TFA, and B was 90% MeCN, 10% H_2O , and 0.1% TFA. The mass spectrometer was scanned from m/z 1000 to 1800 every 4 s.

NMR Analysis of Inactivated Cathepsin K. The samples for NMR analysis were prepared in two ways. For the first sample preparation, cathepsin K (50 μM) was complexed with ^{13}C , ^{15}N -labeled **1** by adding 10 vol % of a 0.5 mM **1** solution in DMSO. The protein solution was then concentrated by ultrafiltration and dialyzed into the desired buffer to yield a final solution of 0.45 mM cathepsin K complexed with **1** in 90% H_2O –10% D_2O , 50 mM acetate- d_3 , 250 mM NaCl, and 2 mM cysteine at pH 4.0 and 5 $^\circ\text{C}$. For the second NMR sample, a 400- μL sample of the 0.5 mM mature form of cathepsin K in 50 mM acetate, 500 mM NaCl, and 2 mM cysteine, pH 4.0, was mixed with 50 μL of a 25 mM SB-240314 solution in DMSO- d_6 . Fifty microliters of buffer (50 mM acetate, 500 mM NaCl, and 2 mM cysteine, pH 4.0) was then added to yield a final solution of 0.4 mM cathepsin K in 90% H_2O –10% DMSO- d_6 , 45 mM acetate- d_3 , 450 mM NaCl, and 1.8 mM cysteine at pH 4.0 and 5 $^\circ\text{C}$. It should be noted that although the total amount of **1** in the NMR tube is about 5:1 in excess of that of cathepsin K, the final concentration of free **1** in solution was negligible due to its very poor solubility, and the only signal observed from the labeled ligand came from that bound to cathepsin K.

Turnover of **1 by Cathepsin K.** In a preliminary paper, we reported that [^3H]-2,2'- N,N' -bis(benzyloxycarbonyl)-L-leucinyldi-carbohydrazide is a substrate for cathepsin K as monitored by the production of 0.92–0.96 equiv of [^3H]-labeled Cbz-Leu-hydrazide by HPLC (18). Formation of one product, [^3H]-Cbz-Leu-hydrazide, as detected by HPLC was complete within 5 min as prolonged incubations and different ratios of inhibitor–enzyme did not increase the stoichiometry (data not shown).

An active site titration of 1 nM cathepsin K was done with varying concentrations of **1** using the standard fluorescent kinetic assay. The dilute protein was found to be completely inhibited by 0.97 equiv of **1**. Cathepsin K (12 μM) was preincubated with 41 μM nonlabeled **1** for 20 min to ensure complete inhibition (as monitored by dilution of an aliquot into the standard fluorescent kinetic assay). ^3H -Labeled **1** was added to the reaction. At various time points, an aliquot of the enzyme reaction was quenched with 1.5 vol of DMSO. To ensure recovery of small amounts of product from the HPLC, nonlabeled carrier was added upon quenching to both the enzyme-catalyzed reaction and the control samples. Aliquots were injected onto a reverse-phase Beckman ODS standard analytical 4.6 mm \times 25 cm HPLC column. Fractions were monitored by UV absorbance, collected, and counted in a Beckman LS3801 scintillation counter.

RESULTS AND DISCUSSION

We have been unable to detect any intrinsic reactivity of these compounds toward thiols in solution. No detectable thiol addition products were observed by proton NMR analysis on exposure of **1** (50 mM) to *p*-methoxybenzyl

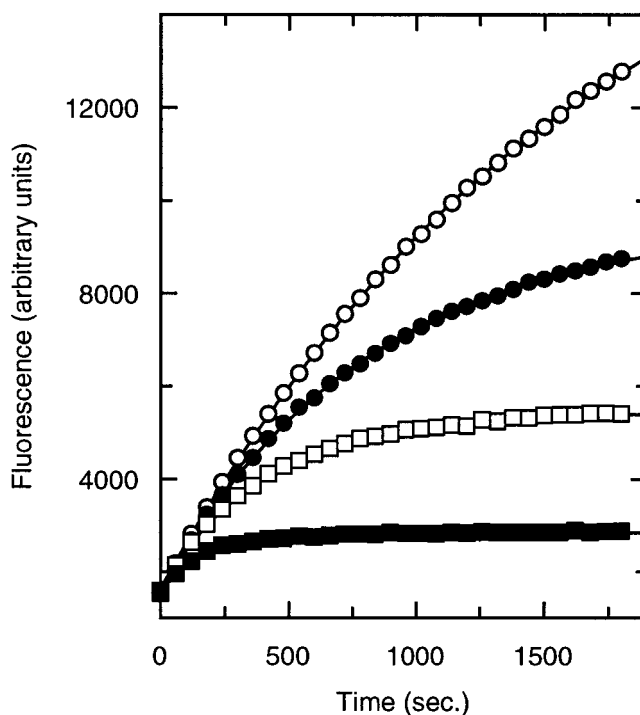


FIGURE 1: Inhibition of cathepsin K by **1**. Individual rate constants were determined by fitting the progress curves: 0.25 (○), 0.5 (●), 1 (□), and 2.5 nM (■).

mercaptan (250 mM) in methanol- d_4 under neutral, acidic (250 mM acetic acid) or basic (250 mM triethylamine) conditions (data not shown). That **1** is a time-dependent inhibitor of cathepsin K has already been reported (18). A crystal structure of cathepsin K complexed with **1** was also reported at this time that was best interpreted as having the bis-hydrazide urea group bound as a tetrahedral adduct at the active site cysteine. The X-ray resolution was, however, not at a high enough level to make an unequivocal interpretation. However, at the time of that paper, the mechanism of inhibition of cathepsin K by **1** was unclear. We have sought, therefore, to better understand the mechanism of inhibition of cathepsin K by **1** and also to extend this to related molecules and other members of the papain superfamily through kinetic and other solution-based physical studies.

Figure 1 shows the effect of various concentrations of **1** on the rate of hydrolysis of the fluorescent cathepsin K substrate Cbz-Phe-Arg-AMC. The nonlinear nature of these progress curves indicate that **1** is a time-dependent inhibitor of cathepsin K. The progress curves at each substrate concentration were fit to eq 1 to yield values for the first-order rate constant, k_{obs} , and initial velocity.

Figure 2 shows both a plot of k_{obs} against **1** concentration and initial rate of substrate hydrolysis against **1** concentration. Analysis of the initial rates of substrate hydrolysis as a function of **1** concentration yields a $K_{\text{i(app)}}$ value of 2.7 nM for the reversible association of **1** with cathepsin K prior to the time-dependent step. A second-order rate constant for inactivation, $k_{\text{inact}}/K_{\text{i}}$, of $5.2 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ can be obtained from the initial linear portion of the replot. That **1** shows both potent inhibition of the initial rate of substrate hydrolysis by cathepsin K and a time-dependent inactivation of cathepsin K suggests that a minimum of two steps is involved in the inhibition of cathepsin K. These two steps would likely

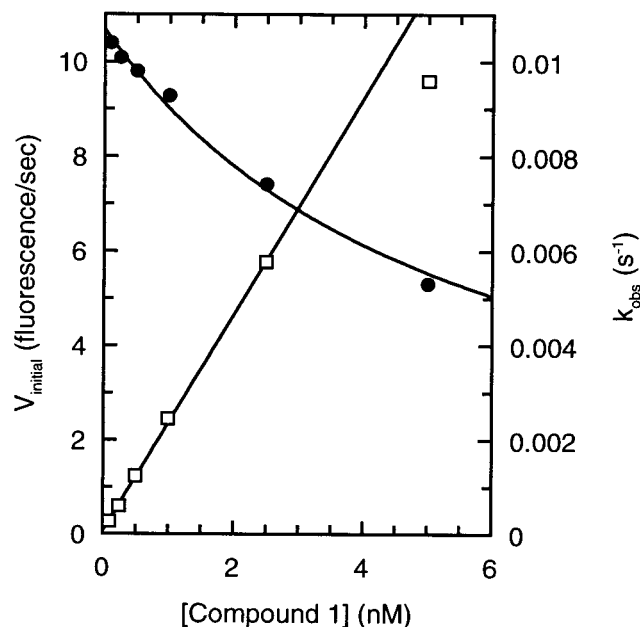
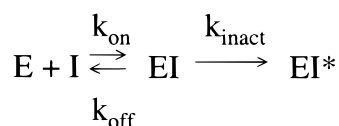


FIGURE 2: Replot of both k_{obs} vs I (□) and V_{initial} vs I (●) determined from the progress curves in Figure 1.

Scheme 1



be a rapid reversible binding of **1** to cathepsin K followed by a slow chemical (inactivation) step (Scheme 1).

This mechanism would predict that $k_{\text{obs}}/[\text{I}]$ should show a saturating maximal value. No indication of saturation of the k_{obs} vs $[\text{I}]$ plot is observed at the relatively low concentrations of **1** used in this study. At **1** concentrations of 10 nM and greater, fitting of progress curves becomes unreliable due to the low initial rates of substrate hydrolysis and rapid enzyme inactivation.

Determination of the Kinetics of Inactivation of Cathepsin K by 1. Cathepsin K exhibits an intrinsic fluorescence increase, upon binding of **1** as monitored in the stopped-flow spectrophotometer. This method does not follow enzyme activity and so is not limited to low inhibitor concentrations used for activity measurements. Under pseudo-first-order conditions where $[\text{I}] \gg [\text{E}]$, the increase in cathepsin K tryptophan fluorescence over time seen upon binding **1** is best described by a single-exponential Figure 3A. A value of $5.1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ for k_{inact}/K_i for the inactivation process can be obtained from the initial slope of this secondary plot. The initial slope of the secondary plot [k_{inact}/K_i is a lower limit for k_{on} (eq 4b)]. The values of k_{inact}/K_i derived from the stopped-flow experiments ($5.1 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$) and substrate hydrolysis ($5.2 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$) experiments are in excellent agreement, suggesting the observation of common enzyme complexes with both techniques.

Fitting k_{obs} as a hyperbolic function of **1** (eq 4) yields an upper limit for k_{inact} of 43 s^{-1} at saturating $[\text{I}]$ (Figure 3B). This hyperbolic dependence of k_{obs} is consistent with a minimal two-step mechanism for the inactivation of cathepsin K by **1** (Scheme 1). Setting $k_{\text{inact}}/K_i = k_{\text{on}}$ allows k_{off} to be

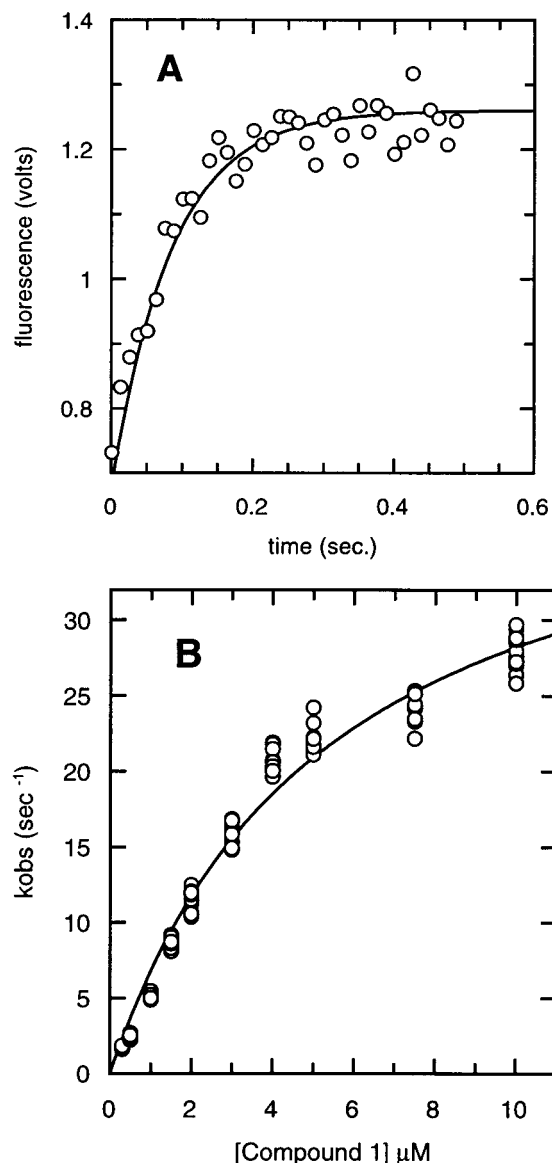


FIGURE 3: Binding of **1** to cathepsin K by stopped-flow analysis. (A) Stopped-flow fluorescence trace upon **1** binding to cathepsin K (○). The solid line is the curve fit to a single-exponential equation as described in the Experimental Procedures. For display purposes only, every tenth data point is shown. (B) Replot of k_{obs} vs $[\text{I}]$ for cathepsin K (○). Cathepsin K was held constant at 100 nM. The solid line is the fit to eq 4b.

calculated from the relationship $K_i = k_{\text{off}}/k_{\text{on}}$. This gives k_{off} a value of $1.4 \times 10^{-2} \text{ s}^{-1}$.

Reactivation Studies. It has previously been reported (18) that, after inactivation of cathepsin K with **1**, inhibition was essentially irreversible upon dilution or dialysis. However, these experiments had been of a qualitative nature rather than quantitative. In addition, inactivation of papain by azapeptides (30) has been shown to give rise to slowly reactivating acyl enzyme species of a type that may be similar to those formed by the inactivation of cathepsin K by **1**. Thus, it was of interest to investigate the reactivation of inactivated cathepsin K further. Attempts to demonstrate reversibility by **1** were complicated by the extreme potency of the inhibitor along with instability of both the free and the inactivated enzyme complex over long periods of time. (The rate of turnover of **1** will be addressed below.) Use of eq 5 for direct determination of k_{react} requires reduction of the

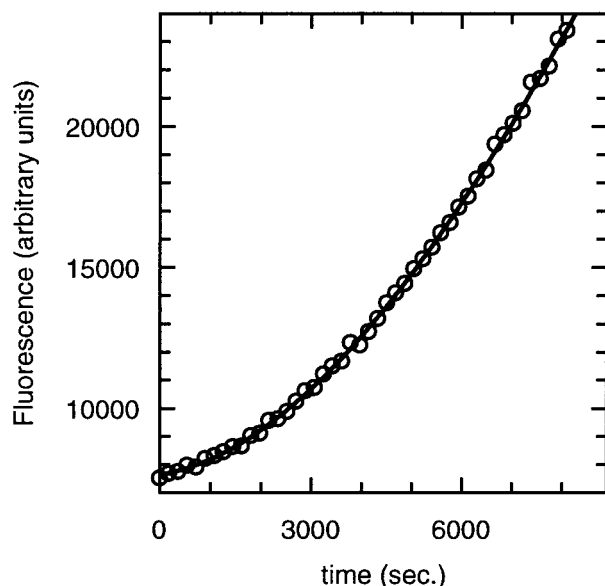


FIGURE 4: Progress curve for the recovery of cathepsin K activity after preincubation with **1** (O). Activity was monitored by hydrolysis of Z-Leu-Arg-AMC. The solid line represents the fit of the data to eq 5. From this fit, a value of k_{off} was determined directly.

Table 1: Inhibition of Papain Superfamily Members by **1**^a

enzyme	$K_{\text{i(app)}}$ (nM)	k_{obs}/I ($\text{M}^{-1} \text{s}^{-1}$)	k_{react} (s^{-1})
cathepsin K	2.7 (0.2)	2.6×10^6 (3.1×10^4)	1.1×10^{-4}
cathepsin B	560 (200)	2.0×10^3 (1.5×10^2)	nd
cathepsin L	<i>b</i>	8.3×10^4 (5.8×10^3)	nd
cathepsin S	<i>b</i>	4.7×10^3 (3.8×10^2)	2.5×10^{-4}
papain	<i>b</i>	4.6×10^3 (0.9×10^2)	0.7×10^{-4}

^a Errors are shown in parentheses. nd, not done. ^b Initial velocities did not change sufficiently to determine a K_{i} .

inhibitor substantially below $K_{\text{i(app)}}$ to avoid reactivation of the enzyme. To satisfy the requirement that $[1] \ll K_{\text{i(app)}}$, reactivation experiments were carried out by diluting enzyme and inhibitor concentration 20 000-fold to final concentrations of 50 pM cathepsin K and 100 pM **1**. Figure 4 shows the recovery of cathepsin K activity after inactivation with compound and dilution into excess substrate. Using this approach, significant recovery of cathepsin K activity was observed over approximately 2 h. The reactivation rate of cathepsin K after inhibition by **1** was found to be $1.1 \times 10^{-4} \text{ s}^{-1}$. Although this rate of reactivation is reasonable, the apparent extent of reactivation is significantly below 100% due to enzyme stability problems.

Inhibition of Other Members of the Papain Superfamily by 1. The time-dependent inhibition of **1** is not unique to cathepsin K among members of the papain superfamily. However, **1** shows a very high degree of selectivity for the inhibition of cathepsin K over other members of the papain superfamily. Comparative data with other family members shown in Table 1 indicate that cathepsin L is inactivated 30-fold less efficiently than cathepsin K. Slower rates of inactivation are reflected in the lower values for k_{obs}/I . Inhibition parameters of the other members of the papain superfamily studied (cathepsin B, cathepsin S, and papain) were approximately 3 orders of magnitude lower than those determined for cathepsin K. In keeping with the changes in $k_{\text{obs}}/[I]$ seen with other members of the papain superfamily, a K_{i} value for the reversible association of **1** was only

measurable for cathepsin B, and this was more than 200-fold greater than that found for cathepsin K; 560 nM as compared with 2.7 nM.

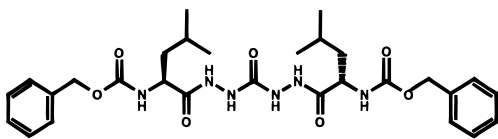
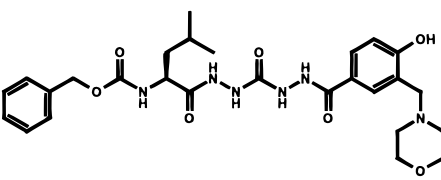
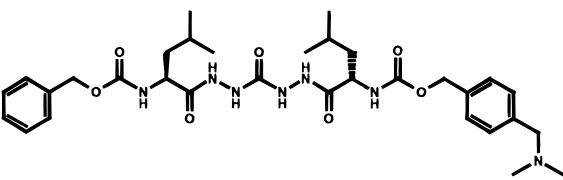
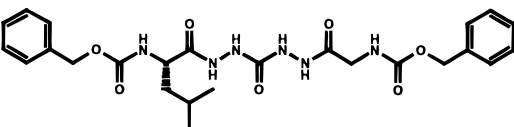
Cathepsin S and papain are inhibited by **1**, albeit slowly as shown in Table 1. However, the reactivation rates are found to be approximately the same (10^{-4} s^{-1}) for all the enzymes examined. The similarity in reactivation rates for papain, cathepsin S, and cathepsin K after inactivation with **1** suggests that the reactivation process is governed by the chemistry of reactivation rather than any specificity determinants in the inactivated complex.

Inhibition of Cathepsin K by Analogues of 1. Table 2 lists the second-order rate constants for the inactivation of cathepsin K by a number of compounds structurally related to **1**. All are good inhibitors of cathepsin K with $k_{\text{obs}}/[I]$ values in excess of $10^5 \text{ M}^{-1} \text{ s}^{-1}$. Although there is an order of magnitude difference in the second-order rate constants for inactivation, the values of k_{react} are essentially identical. It is likely that two of the three analogues give rise to the same inactivated complex as that derived from cathepsin K and **1** (vide infra). It is therefore not surprising that the rates of inactivation are all similar.

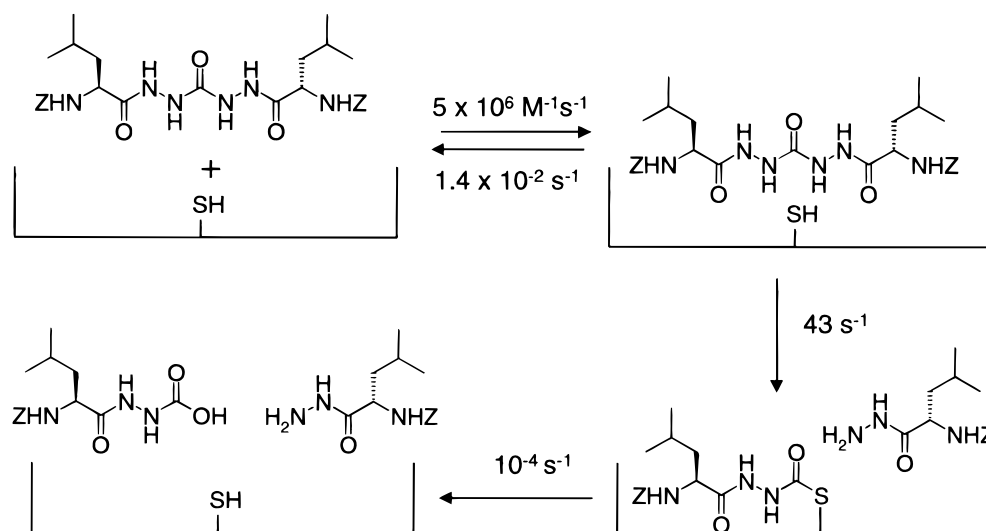
Turnover of 1 by Cathepsin K. As the studies on the reactivation of the cathepsin K–**1** complex were complicated by difficulties in demonstrating a high level of recovery of activity, an alternative approach to obtaining the cathepsin K–**1** reactivation rate was pursued. Although the rate of reactivation ($1.1 \times 10^{-4} \text{ s}^{-1}$) is reasonable, we wished to elucidate whether reactivation was complete. In a preliminary paper (18), we reported that incubation of ^3H -labeled **1** with cathepsin K rapidly generated ~ 1 mol of ^3H -labeled Cbz-Leu-hydrazide per mole of cathepsin K. Prolonged incubations did not increase the stoichiometry of ^3H -labeled Cbz-Leu-hydrazide production. This indicated that the rapid inactivation of cathepsin K by **1** was accompanied by production of Cbz-Leu-hydrazide. This suggested an alternative approach toward estimating release of enzyme could be done by pulse chase studies.

If **1** inactivated cathepsin K were to reactivate in the presence of radiolabeled **1**, then it would be rapidly inactivated again, but now radiolabeled Cbz-Leu-hydrazide would be released. Thus the rate of ^3H -labeled Cbz-Leu-hydrazide release should provide a measure of the rate of reactivation of **1** inactivated cathepsin K. Cathepsin K was completely inactivated with excess **1**, and ^3H -labeled **1** was then added. At varying times, an aliquot was quenched and analyzed for ^3H -labeled Cbz-Leu-hydrazide production. An HPLC profile for production of ^3H -labeled Cbz-Leu-hydrazide as compared to a control reaction is shown in Figure 5A. Figure 5B shows the increase in tritiated Cbz-Leu-hydrazide formation with respect to time when ^3H -labeled **1** is added to cathepsin K that has been previously inactivated by **1**. From the initial linear portion of this curve, the rate of ^3H -labeled Cbz-Leu-hydrazide production was calculated to be $1.2 \times 10^{-4} \text{ s}^{-1}$, assuming that the cathepsin K is 100% active. This is experimentally indistinguishable from the value of $1.1 \times 10^{-4} \text{ s}^{-1}$ found for the reactivation of the cathepsin K–**1** complex measured by dilution into excess substrate and supports the supposition that the reactivation process measured by the dilution experiment is real and corresponds to 100% of the inactivated cathepsin K.

Table 2: Inhibition of Cathepsin K by **2** Analogues

Inhibitor	Structure	k_{obs}/I ($\text{M}^{-1}\text{s}^{-1}$)	k_{react} (s^{-1})
Compound 1		3.1×10^6	1.1×10^{-4}
Compound 2		2.1×10^5	1.2×10^{-4}
Compound 3		5.2×10^6	1.8×10^{-4}
Compound 4		1.3×10^5	1.2×10^{-4}

Scheme 2



Nature of the Inactivated Cathepsin K–1 Complex: Mass Spectral Analysis. Upon HPLC analysis of enzyme samples that had been incubated with ^3H -labeled **1**, a peak of radioactivity was found to coelute with protein as monitored by UV. The peak of radioactivity increased in size upon increasing enzyme concentration. Protein recovery issues precluded accurate determination of the stoichiometry of the enzyme–inhibitor complex with the HPLC system employed. To further understand the nature of the inhibited cathepsin K–**1** complex, HPLC–electrospray mass spectrometric

analysis was undertaken. Cathepsin K was incubated with **1** as described in Experimental Procedures, and mass spectrometric analysis was performed to determine the chemical nature of the inhibited enzyme. Cathepsin K is a 24-kDa protein, which upon inactivation, is found to increase in mass by 305.6 Da. The expected mass increase for a carbamyl adduct of the type shown in Scheme 2 is 305.3 Da. Thus, the formation of this +305.6 Da higher mass enzyme species upon inactivation by **1** corresponds to carbamylation of cathepsin K by **1** concomitant with loss of Cbz-Leu-

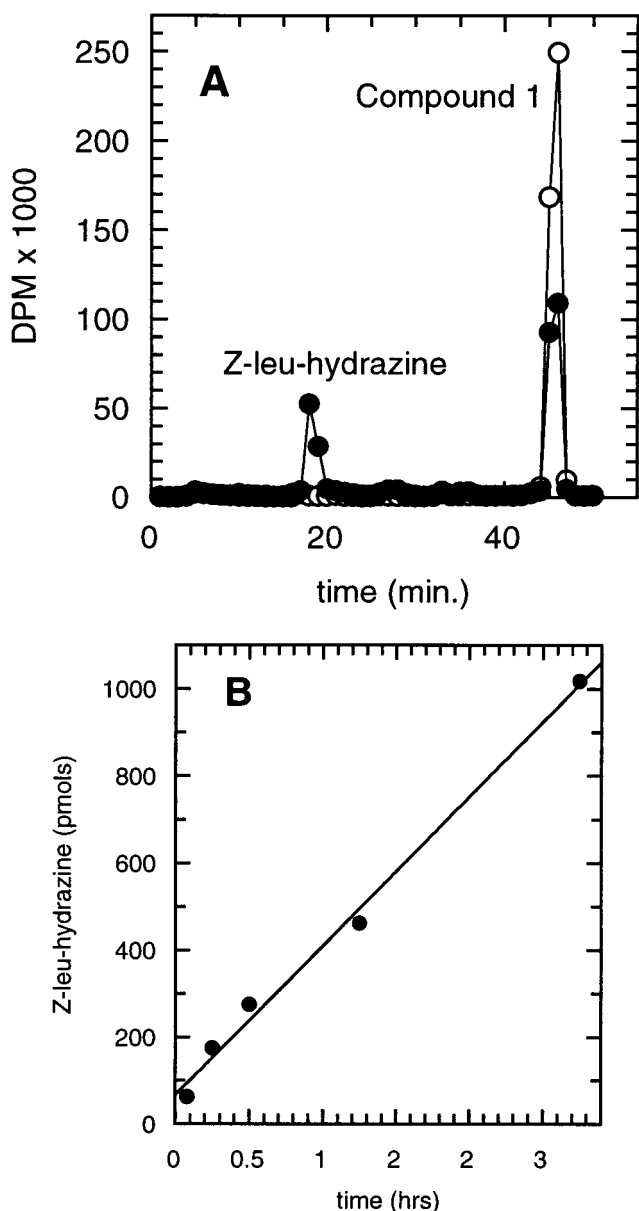


FIGURE 5: Production of ^3H -labeled Z-Leu-hydrazine by exchange of ^3H -labeled **1** into a cathepsin K–**1** complex. Carbamyl–enzyme was formed by preincubation of cathepsin K with excess **1**. A “hot” chase of ^3H -labeled **1** was added. At various time points, aliquots were quenched with DMSO and analyzed by HPLC. (A) The amount of ^3H -labeled Z-Leu-hydrazine formed (●) vs control (○) 2.5 h after the addition of ^3H -labeled **1**. (B) Production of ^3H -labeled Z-Leu-hydrazine as a function of time (●). For production of ^3H -labeled Z-Leu-hydrazine to occur, the nontritiated carbamyl–enzyme must decarbamate followed by subsequent binding of a ^3H -labeled **1**, which then cleaves to generate ^3H -labeled Z-Leu-hydrazine and a new carbamyl–enzyme.

hydrazide. All of the cathepsin K is found to be modified in this way, indicating both that inactivation is complete and that only one carbamyl–enzyme species is formed. In addition, Scheme 2 depicts the formation of 1 equiv of Cbz-Leu-hydrazide upon inactivation.

To investigate this apparent inactivation by carbamylation with concomitant production of Cbz-Leu-hydrazide, we examined both the inactivation of cathepsin K by a number of analogues of **1** and also the inactivation of cathepsin S and papain by **1**. Table 3 shows the mass adducts observed when cathepsin S (25 kDa) and papain (23 kDa) are

Table 3: Mass Spectral Analysis of Members of the Papain Superfamily Inhibited by **1**

enzyme	average adduct mass
cathepsin K	305.6
cathepsin S	305.2
papain	305.1

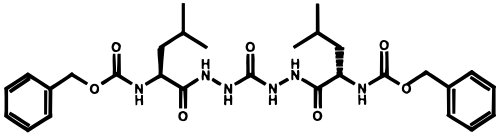
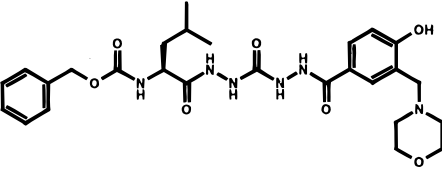
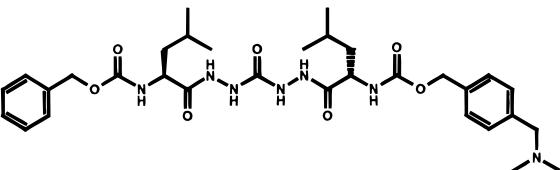
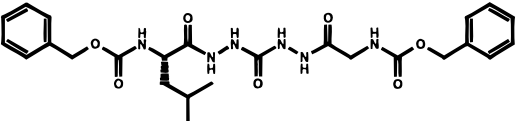
inactivated by **1**. Although **1** is a poor inactivator of cathepsin S and papain (Table 1), complete inactivation is possible. Both enzymes give rise to a mass adduct of +305 Da, indicating that these members of the papain superfamily form a similar inactivated species as that seen for cathepsin K. As seen with cathepsin K, cathepsin S and papain show complete conversion of native enzyme to the increased mass species that is consistent with the carbamyl–enzyme shown in Scheme 2. Thus, as judged by mass spectrometry of inactivated enzyme, **1** and analogues inactivate cathepsin K and other members of the papain superfamily by forming a carbamyl–enzyme species, and this mechanism is common to all of the analogues and enzymes studied in this work.

Mass spectroscopic analysis of the adducts generated by the inactivation of cathepsin K with a number of analogues of **1** is summarized in Table 4. All compounds show complete formation of adducts relative to unreacted cathepsin K consistent with the mechanism in Scheme 2. All but one compound give rise to the same adduct as seen for **1**, indicating that the preferred route of carbamylation is to retain a Cbz-leucinyldihydrazinyl carbamyl–enzyme species. The exception to this is **3**, which shows two mass adducts in a ratio of 2:1 indicating that dual binding modes exist such that either half of the compound is retained on the enzyme but the dimethylaminobenzyl group is slightly favored as compared to simple phenyl. Since the central carbonyl is retained in both adducts, the results suggest that the inhibitor may initially bind in either of two possible orientations of an unsymmetrical diacylcarbohydrazide.

If we consider this carbamylation reaction to be analogous to the first half of the peptide bond cleavage cycle catalyzed by cathepsin K, then it appears that a Cbz-leucinyld moiety is favored on the S side of the active site, in keeping with substrate preferences (20), and the half of the molecule binding to the S' side is the product released upon carbamylation. Thus, **2** and **4** demonstrate this preference by the formation of only one adduct. **1** is symmetrical; therefore, it is obliged to form only one adduct. **3** offers two binding orientations, both of which satisfy the requirement to bind a leucine moiety on the S side. The appearance of the dimethylamino group in two-thirds of the adduct reflects the fact that the orientation of **3** with respect to the active site is almost random. This near randomness of the retention of the dimethylaminomethyl-bearing half of **3** is not surprising given that this change is relatively minor in terms of its affect on subsite specificities. Also, it is noteworthy that a positively charged amino acid is tolerated at both P3 and P3' as exemplified by the cleavage of osteonectin by cathepsin K, so this substitution would not obviously disfavor either orientation of binding (21).

*Nature of the Inactivated Cathepsin K–**1** Complex: NMR Studies.* It was of concern that our inability to see an intact molecule of **1** bound to cathepsin K, as seen in the crystal structure, may be due to the conditions employed in the

Table 4: Mass Spectral Analysis of Cathepsin K Inhibited by 2,2'-N,N'-Bis(benzyloxycarbonyl)-L-leucinecarbohydrazide Analogues^a

Inhibitor	Structure	Chemical	Other
		average	adducts
Compound 1		305.6	
Compound 2		305.7	
Compound 3		306.15 (33%)	362.0 (67%)
Compound 4		305.9	

^a One-half of each inhibitor is the same as **1**, and so the expected adduct mass for that reaction is 305.34 Da. No evidence was observed for unreacted cathepsin K in any sample besides the control. Abundances are based on the observed ratio of peak heights in the MS data. All samples contained a minor amount of protein with two adducts. This may be an artifact of the mass spec conditions and was not pursued.

experiment, namely, reverse-phase HPLC and ionization of the sample in the mass spectrometer. In an attempt to influence the nature of the inactivated cathepsin K complex as little as possible, NMR studies were undertaken using ¹⁵N- and ¹³C-labeled **1** with the labels in the two hydrazine moieties and the central carbonyl group, respectively. To further investigate the nature of the cathepsin K–**1** adducts, experiments (Figure 6) were performed under conditions that would mimic the solution-phase kinetic experiments as closely as possible.

The ¹³C spectrum of labeled **1** is shown in Figure 6A. The spectrum consists of a 1:2:1 triplet resulting in the two adjacent ¹⁵N nuclei coupling to the central ¹³C nucleus. Geminal coupling ²*J*(¹³C,¹⁵N) to the more distant two ¹⁵N nuclei is not evident in this spectrum.

Figure 6, panels B and C show the ¹³C NMR spectra obtained after reacting ¹⁵N- and ¹³C-labeled **1** with cathepsin K. The two spectra arise from different methods of sample preparation as described above. The NMR spectra in Figure 6B,C are identical, indicating that the result is independent of method of sample preparation. Both spectra show a broad doublet, which is consistent with the ¹³C nucleus coupled to one adjacent ¹⁵N nucleus. ¹⁵N broad band decoupling causes this doublet to collapse to a singlet, confirming that the doublet is indeed due to coupling to an ¹⁵N nucleus. The transformation of the ¹³C NMR spectrum from a triplet derived from coupling to two ¹⁵N nuclei in the parent compound to a doublet derived from coupling to only one

¹⁵N nucleus cannot arise from an intact molecule of **1** bound to cathepsin K but is consistent with a carbamyl–enzyme species similar to that shown in Scheme 2. In the presence of cathepsin K, a downfield shift of the carbonyl resonance (Figure 6B,C) relative to the free ligand (Figure 5A) is seen. This is consistent with a stable form of an carbamyl–enzyme adduct and not a tetrahedral adduct. Furthermore, the doublet structure of the carbonyl resonance (Figure 6B,C) is in sharp contrast to the triplet structure observed for the free ligand in DMSO-*d*₆, indicating cleavage of the ligand upon binding at one of the two ¹³C–¹⁵N bonds. That this spectrum is a doublet suggests that the cathepsin K–**1** complex no longer has both hydrazine moieties of **1** covalently bound to the central carbonyl group. The slight difference in chemical shift of the carbonyl resonance observed in the spectra of Figure 6, panels B and C, can be attributed to the slightly different buffer conditions, as described in Experimental Procedures.

The close similarity of the results observed in Figure 6, panels B and C, indicate that the absence or presence, respectively, of the Cbz-Leu-hydrazide, which is a product of the cleavage of **1** upon binding, does not affect the stability of the carbamyl–enzyme adduct.

SUMMARY

Analysis of the kinetics and mechanism of inhibition of cathepsin K by **1** is summarized in Scheme 2. The rate of association of **1** approximates *k*_{obs}/[**1**] when the rate of dissociation is small in comparison to the forward rate

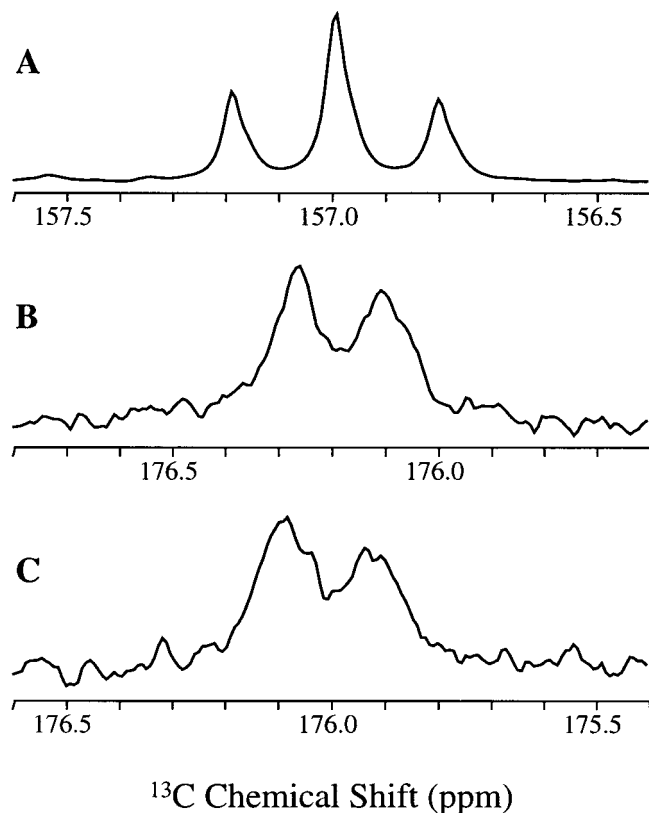


FIGURE 6: ^{13}C NMR spectra of $[\text{}^{13}\text{C}, \text{}^{15}\text{N}_4]\text{-2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucinyldihydrazide (1)}$ in the presence and absence of cathepsin K at 125.76 MHz. (A) $^{13}\text{C}, \text{}^{15}\text{N}_4$ -labeled **1** (15 mM) in $\text{DMSO-}d_6$ at 20 °C, 1.31-s acquisition time, 1-s relaxation delay, 3000 scans, ^1H broadband decoupling. The triplet structure of the ^{13}C -labeled carbazide carbonyl resonance peak is caused by the one-bond $^{13}\text{C}\text{--}^{15}\text{N}$ spin–spin coupling ($^1J_{\text{CN}} = 19.5$ Hz) with two adjacent ^{15}N -labeled carbazide amines. (B) First NMR sample preparation (as described in Experimental Procedures) in which the sample was dialyzed after formation of the enzyme inhibitor complex; 0.52-s acquisition time, 1.48-s relaxation delay, 123 283 scans, ^1H broadband decoupling (C) Second NMR sample preparation (as described in Experimental Procedures) without dialysis; 0.52-s acquisition time, 1.48-s relaxation delay, 144 000 scans, ^1H broadband decoupling. The doublet structure in panels B and C is caused by the $^{13}\text{C}\text{--}^{15}\text{N}$ spin–spin coupling with only one adjacent ^{15}N -labeled amine and coalesces into a singlet upon ^{15}N decoupling. The spectrum in panel A was referenced to the $\text{DMSO-}d_6$ ^{13}C resonance whereas the spectra in panels B and C were referenced to an external DSS standard dissolved in the buffers used in panels B and C, respectively.

constant for the second step (eq 4a). This allows the individual rate constants to be calculated as shown in Scheme 2. **1** binds rapidly and with high affinity to cathepsin K. This is followed by carbamylation of cathepsin K. We have assumed that in the minimal mechanism shown that the second step evident from the stopped-flow studies (rate constant of 43 s^{-1}) is the carbamylation step. While we have no direct proof of this, the value of 43 s^{-1} is in keeping with known k_{cat} values for good substrates of cysteine proteases (31).

Reactivation of cathepsin K occurs slowly, 10^{-4} s^{-1} . It is likely that reactivation represents the rate of carbamyl–enzyme hydrolysis rather than dissociation of either of the two products from a noncovalent complex. This is supported by the fact that cathepsin S and papain reactivate at the same rate as cathepsin K, despite being inactivated more slowly. All three enzymes form the same carbamyl–enzyme com-

plex as judged by mass spectrometry so the rate of carbamyl–enzyme hydrolysis would be expected to be very similar for all three enzymes. It is known that carbamyl–enzyme complexes are refractory toward hydrolysis. Generation of similar carbamyl–enzyme intermediates by inhibition with aza-peptide esters has been reported (32). These complexes are found to reactivate very slowly with $t_{1/2}$ values of several hours (30, 32). The stability of these carbamyl–enzyme complexes has been attributed to the decreased electrophilic nature of the carbonyl group in carbamates relative to esters/thioesters due to delocalization involving the nitrogen adjacent to the carbonyl group (32). If product release were rate-limiting, then it would perhaps be surprising that the rates of reactivation were so similar. In addition, the affinity of Cbz-Leu-hydrazide for cathepsin K is so low that we have been unable to determine a K_i .

It is clear from the work described here that in solution the only observable covalent intermediate in this inactivation process is the carbamyl–enzyme complex shown in Scheme 2. We have been unable to detect a discrete long-lived complex that would correspond to the tetrahedral species that was suggested in the crystal structure for the cathepsin K–**1** complex reported previously (18). This apparent discrepancy may be due to a number of reasons. Under the conditions used to obtain crystals it may be possible to kinetically ‘stall’ the enzyme prior to carbamylation and so obtain a snapshot of the cathepsin K–**1** complex that is not normally isolatable in solution. (The crystal structure would then be analogous to looking at a single frame from a movie.) It may also be possible that the crystal structure actually corresponds to the carbamyl–enzyme complex but with the product Cbz-Leu-hydrazide still bound to the S' side. The X-ray data was at a resolution (2.2 \AA) that does not distinguish other possible interpretations that are also consistent with the solution data reported here. This includes binding of the hydrolysis product (Cbz-Leu-hydrazide) to the carbamylated enzyme. The X-ray data also did not preclude an interpretation where the carbamylated species discussed here binds with similar populations at both the S and S' sites. We are unable to unequivocally determine the presence or absence of this in solution.

In summary, **1** is a potent, high-affinity but intrinsically unreactive inactivator of cathepsin K that is dependent on the catalytic process for inhibition. This molecule represents a new class of cysteine protease inhibitors that share a common mechanism of action across members of the papain superfamily. Both S and S' site interactions are exploited in achieving high potency and rapid inactivation.

ACKNOWLEDGMENT

We thank Keith Garnes, Art Shu, and J. Richard Heys for production of $[\text{}^3\text{H}]\text{2,2'-N,N'-bis(benzyloxycarbonyl)-L-leucinyldihydrazide}$. We thank Michael McQueeney, Karla J. D'Alessio, Bernard Amegadzie, and Charles R. Hanning for the cathepsins used in this study.

REFERENCES

- Zaidi, M., Pazianas, M., Shankar, V. S., Bax, B. E., Bax, C. M., Bevis, P. J., Stevens, C., Huang, C. L., Blake, D. R., and Moonga, B. S. (1993) *Exp. Physiol.* 78, 721–739.
- Baron, R. (1995) *Acta Orthop. Scand. Suppl.* 266, 66–70.

3. Delaisse, J. M., Eeckhout, Y., and Vaes, G. (1984) *Biochem. Biophys. Res. Commun.* 125, 441–447.
4. Delaisse, J. M., Eeckhout, Y., and Vaes, G. (1980) *Biochem. J.* 192, 365–368.
5. Delaisse, J. M., Boyde, A., Maconnachie, E., Ali, N. N., Sear, C. H., Eeckhout, Y., Vaes, G., and Jones, S. J. (1987) *Bone* 8, 305–313.
6. Lerner, H. H., and Grubb, A. (1992) *J. Bone Miner. Res.* 7, 433–439.
7. Hill, P. A., Buttle, D. J., Jones, S. J., Boyde, A., Murata, M., Reynolds, J. J., and Meikle, M. C. (1994) *J. Cell. Biochem.* 56, 118–130.
8. Drake, F. H., Dodds, R. A., James, I. E., Connor, J. R., Richardson, S., Coleman, L., Rieman, D., Barthlow, R., Hastings, G., and Gowen, M. (1996) *J. Biol. Chem.* 271, 12511–12516.
9. Gelb, B. D., Moissoglu, K., Zhang, J., Martignetti, J. A., Bromme, D., and Desnick, R. J. (1996) *Biochem. Mol. Med.* 59, 200–206.
10. Gelb, B. D., Shi, G. P., Chapman, H. A., and Desnick, R. J. (1996) *Science* 273, 1236–1238.
11. Johnson, M. R., Polymeropoulos, M. H., Vos, H. L., Ortiz, d. L. R., and Francomano, C. A. (1996) *Genome Res.* 6, 1050–1055.
12. Rich, D. H. (1998) in *Proteinase Inhibitors* (Barrett, A. J., and Salvesen, G., Eds.) pp 153–178, Elsevier, New York.
13. Albeck, A., and Fluss, S. P. R. (1996) *J. Am. Chem. Soc.* 118, 3591–3596.
14. Bromme, D., Klaus, J. L., Okamoto, K., Rasnick, D., and Palmer, J. T. (1996) *Biochem. J.* 315, 85–89.
15. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* 27, 157–162.
16. Marquis, R. W., Yamashita, D. S., Ru, Y., LoCastro, S. M., Oh, H. J., Erhard, K. F., DesJarlais, R. L., Head, M. S., Smith, W. W., Zhao, B., Janson, C. A., Abdel-Meguid, S. S., Tomaszek, T. A., Levy, M. A., and Veber, D. F. (1998) *J. Med. Chem.* 41, 3563–3567.
17. Yamashita, D. S., Smith, W. W., Zhao, B. G., Janson, C. A., Tomaszek, T. A., Bossard, M. J., Levy, M. A., Oh, H. J., Carr, T. J., Thompson, S. K., Ijames, C. F., Carr, S. A., McQueney, M., Dalessio, K. J., Amegadzie, B. Y., Hanning, C. R., Abdelmeguid, S., DesJarlais, R. L., Gleason, J. G., and Veber, D. F. (1997) *J. Am. Chem. Soc.* 119, 11351–11352.
18. Thompson, S. K., Halbert, S. M., Bossard, M. J., Tomaszek, T. A., Levy, M. A., Zhao, B., Smith, W. W., Abdel-Meguid, S. S., Janson, C. A., D'Alessio, K. J., McQueney, M. S., Amegadzie, B. Y., Hanning, C. R., DesJarlais, R. L., Briand, J., Sarkar, S. K., Huddleston, M. J., Ijames, C. F., Carr, S. A., Garnes, K. T., Shu, A., Heys, J. R., Bradbeer, J., Zembryki, D., and Veber, D. F. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 14249–14254.
19. Zhao, B., Janson, C. A., Amegadzie, B. Y., D'Alessio, K., Griffin, C., Hanning, C. R., Jones, C., Kurdyla, J., McQueney, M., Qiu, X., Smith, W. W., and Abdel-Meguid, S. S. (1997) *Nat. Struct. Biol.* 4, 109–111.
20. McGrath, M. E., Klaus, J. L., Barnes, M. G., and Bromme, D. (1997) *Nat. Struct. Biol.* 4, 105–109.
21. Bossard, M. J., Tomaszek, T. A., Thompson, S. K., Amegadzie, B. Y., Hanning, C. R., Jones, C., Kurdyla, J. T., McNulty, D. E., Drake, F. H., Gowen, M., and Levy, M. A. (1996) *J. Biol. Chem.* 271, 12517–12524.
22. Votta, B. J., Levy, M. A., Badger, A., Bradbeer, J., Dodds, R. A., James, I. E., Thompson, S., Bossard, M. J., Carr, T., Connor, J. R., Tomaszek, T. A., Szwczuk, L., Drake, F. H., Veber, D. F., and Gowen, M. (1997) *J. Bone Miner. Res.* 12, 1396–1406.
23. Leatherbarrow, R. J. (1996) in *GrafIt 3.0*, Erithacus Software Ltd., Staines, U.K.
24. Morrison, J. F., and Walsh, C. T. (1988) *Adv. Enzymol.* 61, 201–301.
25. Williams, J. W., and Morrison, J. F. (1979) *Methods Enzymol.* 63, 437–467.
26. Cha, S. (1975) *Biochem. Pharmacol.* 24, 2177–2185.
27. Knight, W. B., Green, B. G., Chabin, R. M., Gale, P., Maycock, A. L., Weston, H., Kuo, D. W., Westler, W. M., Dorn, C. P., and Finke, P. E. (1992) *Biochemistry* 31, 8160–8170.
28. Knight, C. G. (1986) in *Proteinase Inhibitors* (Barrett, A. J., and Salvesen, G., Eds.) pp 23–51, Elsevier, New York.
29. Brady, K. A. R. H. (1990) *Biochemistry* 29, 7608–7617.
30. Baggio, R., Shi, Y. Q., Wu, Y., and Ables, R. H. (1996) *Biochemistry* 35, 3351–3353.
31. Kirschke, H., Barrett, A. J., and Rawlings, N. D. (1998) *Protein Profile* 2, 1587–1643.
32. Gupton, B. F., Carroll, D. L., Tuhy, P. M., Kam, C. M., and Poweres, J. C. (1984) *J. Biol. Chem.* 259, 4279–87.

BI991193+