



Probing the Activity Modification Space of the Cysteine Peptidase Cathepsin K with Novel Allosteric Modifiers

Marko Novinec^{1,2*}, Brigita Lenarčič^{1,3}, Antonio Baici^{2*}

1 Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia, **2** Department of Biochemistry, University of Zürich, Zürich, Switzerland, **3** Department of Biochemistry and Molecular and Structural Biology, Jožef Stefan Institute, Ljubljana, Slovenia

Abstract

Targeting allosteric sites is gaining increasing recognition as a strategy for modulating the activity of enzymes, especially in drug design. Here we investigate the mechanisms of allosteric regulation of cathepsin K as a representative of cysteine cathepsins and a promising drug target for the treatment of osteoporosis. Eight novel modifiers are identified by computational targeting of predicted allosteric sites on the surface of the enzyme. All act via hyperbolic kinetic mechanisms in presence of low molecular mass substrates, as expected for allosteric effectors. Two compounds have sizable effects on enzyme activity using interstitial collagen as a natural substrate of cathepsin K and four compounds show a significantly stabilizing effect on cathepsin K. The concept of activity modification space is introduced to obtain a global perspective of the effects elicited by the modifiers. Analysis of the activity modification space reveals that the activity of cathepsin K is regulated via multiple, different allosteric mechanisms.

Citation: Novinec M, Lenarčič B, Baici A (2014) Probing the Activity Modification Space of the Cysteine Peptidase Cathepsin K with Novel Allosteric Modifiers. PLoS ONE 9(9): e106642. doi:10.1371/journal.pone.0106642

Editor: Matthew Bogyo, Stanford University, United States of America

Received: May 27, 2014; **Accepted:** August 7, 2014; **Published:** September 3, 2014

Copyright: © 2014 Novinec et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Slovenian Research agency (grant Z1-4077 to MN), the Slovenian Research Agency - program group P1-0140 (MN and BL), the Olga Mayenfisch Foundation (grant to AB), and the Novartis Foundation (grant to AB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: marko.novinec@fkkt.uni-lj.si (MN); abaici@bioc.uzh.ch (AB)

Introduction

The classical two-state conformational model proposed by Monod and co-workers to explain allosteric interactions in hemoglobin [1,2] has been expanded in recent years as a more generalized mechanism of protein regulation than proposed originally. Major extensions to the theory include allosteric regulation of monomeric proteins and the not necessary presence of distinctive conformational changes as a prerequisite for this type of regulation [3,4].

Cysteine cathepsins belong to the papain-like peptidases and constitute one of the major groups of lysosomal enzymes. They are involved in diverse physiological and pathological processes and have long been thought to be regulated primarily by macromolecular active-site directed inhibitors (recently reviewed in ref. [2]). Alternative mechanisms of regulation operate without direct involvement of the active site. Glycosaminoglycans emerged as a group of regulators with diverse effects on various cysteine cathepsins, including the archetypal papain and lysosomal cathepsin B [5,6]. The best studied example of this mode of regulation is cathepsin K, the principal peptidase in bone turnover and a promising target for the treatment of osteoporosis (reviewed in ref. [7]). Its unique collagenolytic activity was shown to be potentiated by cartilage-resident glycosaminoglycans, foremost chondroitin sulfate [8,9]. Together with the crystal structure of chondroitin sulfate bound to cathepsin K outside of the active site [10], chondroitin sulfate and other sulfated glycosaminoglycans

have been characterized as the first allosteric regulators of a papain-like peptidase [11].

Recently, we described the first low-molecular mass allosteric modifier of cathepsin K identified by *in silico* screening of compound libraries [12] targeted at predicted allosteric sites that are connected to the active site via an evolutionarily conserved network of residues, called a protein sector [13]. Here, we expand the characterization of cathepsin K control by describing eight further compounds that affect enzyme activity by mechanisms that, kinetically, correspond to allosteric interactions. We extensively characterize the described compounds and compare their relative effects on cathepsin K by representing their activity by a novel approach termed enzyme activity modification space.

Results

Identification of modifiers

The prediction of seven putative allosteric sites on cathepsin K and targeting of these sites have been described previously [12]. Here, the same procedure was used to identify further allosteric modifiers. Each predicted site was targeted *in silico*, as described in more detail in the Materials and Methods section. The approach included a first round of high-throughput docking of compound libraries using UCSF Dock 6, followed by a second round in which the top 10% of hits were re-docked to the same site with AutoDock. Altogether, over 200 compounds were selected for experimental evaluation. Eight novel modifiers of cathepsin K are described herein. Their structures are shown in Figure 1A and

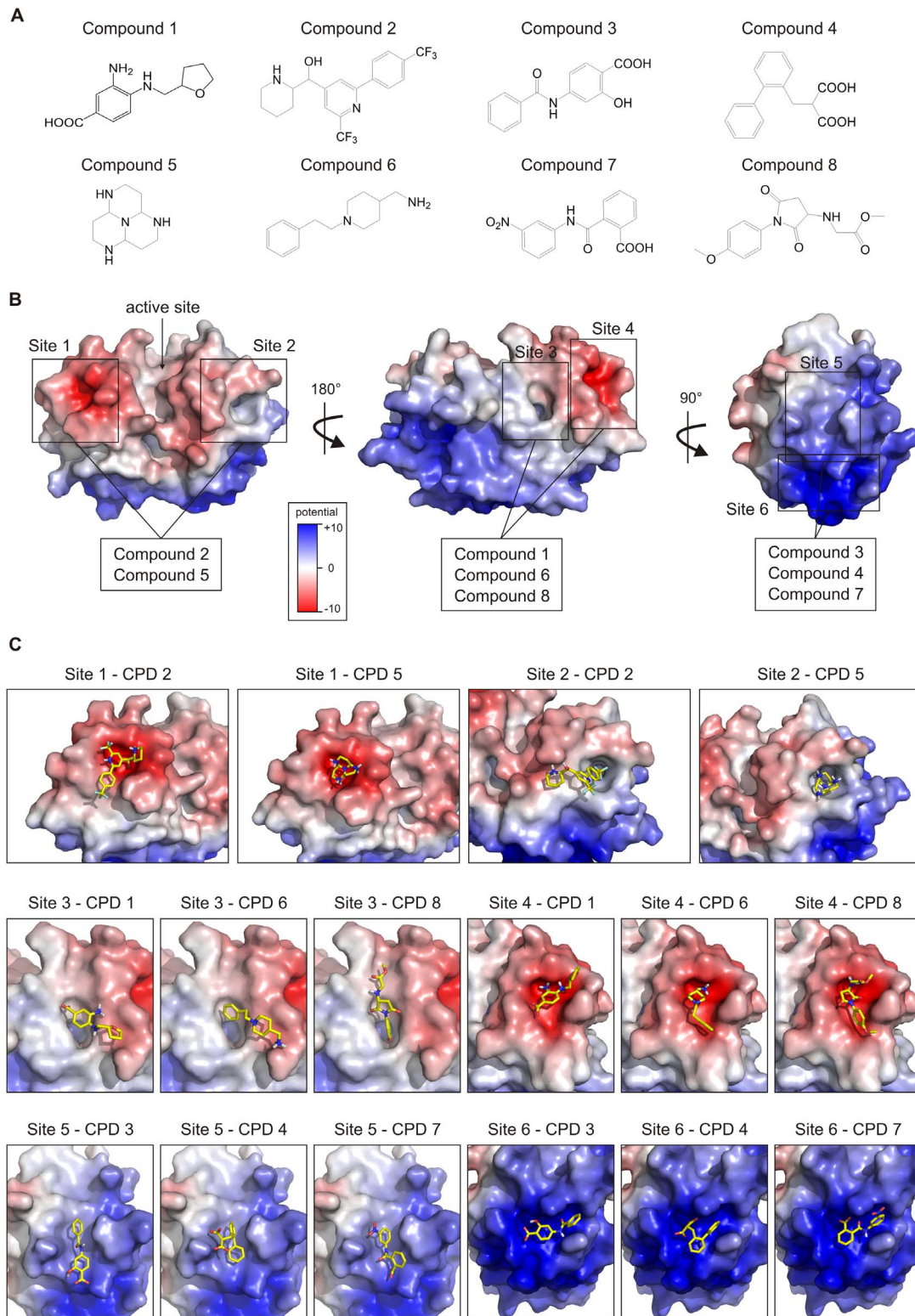


Figure 1. Structures, predicted binding sites and docked poses of the identified modifiers. (A) Structural formulas of the identified modifiers. (B) The locations of six potential allosteric sites on cathepsin K targeted in this study. Binding sites for each modifier marked. For reference, the location of the active site is marked by an arrow. (C) Binding poses of identified binders at their target sites on cathepsin K. In both panels, the receptor is shown in surface representation and colored according to the electrostatic potential (the colorbar is shown in the inset in panel A and is given in dimensionless units calculated by APBS) and the compounds are shown as sticks. The shown binding poses were calculated with AutoDock. Images was created with PyMOL (Schrodinger, Inc., Portland, OR, USA). Abbreviation: CPD – compound.
doi:10.1371/journal.pone.0106642.g001

further information on their identity and origin is given in Table 1. Each compound was predicted to bind to two of six different sites on cathepsin K (Figure 1B). The lowest energy binding modes calculated with AutoDock are shown in Figure 1C and the corresponding binding energies and equilibrium dissociation constants are collected in Table 2. The seventh putative allosteric site on the bottom of the molecule is not shown in Figure 1B since none of the compounds were predicted to bind there.

Cathepsin K is known for its high positive charge at about half of its molecular surface (Figure 1B) and electrostatic interactions are key to its regulation by glycosaminoglycans [10]. Charge-charge interactions are apparently also important in small molecule binding, as each of the identified modifiers contains at least one ionizable group (Figure 1A). The predicted binding sites are diverse both in shape and electrostatic potential. Whereas sites 5 and 6 are located in the positively charged area, sites 1 and 4 represent negatively charged patches on the protein's surface and the potential of sites 2 and 3 is near-neutral. Sites 1 and 2 were each bound by compounds 2 and 5 which both contain amino groups that can interact favorably with the slightly negative or near-neutral electrostatic potential of these sites. Sites 3 and 4 were bound by compounds 1, 6 and 8. At site 4, interactions were again driven by electrostatic interactions between the negatively charged enzyme surface and amino groups of the ligands. Interestingly, site 3 has the shape of a deep narrow groove and thus appears as a typical small molecule-binding site. However, none of the predicted compounds bound within the narrow groove. Instead, interactions are limited to the region surrounding the cleft. Finally, sites 5 and 6 were bound by compounds 3, 4 and 7 which contain carboxyl groups that interact with positively charged surface residues on the enzyme. These two sites were also predicted previously as binding sites for the allosteric modifier NSC13345 and site 6 was revealed as the true binding site [12].

Determination of kinetic mechanisms

All modifiers were tested for their effect on the hydrolysis of the low molecular mass synthetic substrate Z-FR ↓ AMC. Since all progress curves were linear from the beginning of the reaction, initial velocities were calculated from the slopes by linear regression. No significant deviations from linearity were observed for the first 120 s of the reaction and no substrate inhibition was apparent at the substrate concentrations used in this study. The basic mechanism considered for the analysis of the interactions was the general modifier mechanism shown in Figure 2, in which the effect of modifier is determined by the values of the coefficients α and β that multiply the kinetic parameters K_S , K_A (α) and k_{cat} (β) [14]. In the thermodynamic box embodied in Figure 2, the

reciprocal of the coefficient α represents the allosteric coupling constant that characterizes energy linkage between the modifier and the substrate [15], while a non-zero value of β represents the effect exerted by the modifier on catalysis. The specific velocity plot [16] was used for the primary screening of the kinetic behavior of the modifiers (Equation 2), as it allows an immediate judgment of the type of interaction and to gather preliminary estimates of the coefficients α and β , as well as of the dissociation constant of the modifier. Refinement was then performed by fitting the rate Equation 1 to data.

Specific velocity plots for all modifiers are shown in Figure 3 along with relative velocity versus modifier concentration curves recorded at fixed substrate concentrations. The kinetic parameters calculated from the combinations of both methods are collected in Table 3, together with activity parameters determined from other assays described in the next section. The calculated equilibrium dissociation constants of the enzyme/modifier complex K_A were in the 10^{-5} to 10^{-3} molar range, which was expected considering the size of the compounds and the fact that they were selected without optimization of the binding geometry. All eight modifiers acted by a hyperbolic mechanism, i.e. they caused a reduction of the reaction rate without driving it to zero at saturation. They can be subdivided into four groups according to their mechanisms of modification as determined from combinations of the coefficients α and β in Figure 2. Compounds 1 through 3 were assigned to a first group with characteristic values of $\alpha > 1$ and $\beta < 1$, which can thus be classified as hyperbolic mixed, predominantly competitive inhibitors. Compounds 1 and 2 had large values of α , indicating a strong competitive component, whereas that of compound 3 was only slightly above 1, a property that accounts for the different behavior of compound 3 with respect to compounds 1 and 2 in assays described in the following subsections. The second group comprised the pair of compounds 4 and 5, which belong to a class of hyperbolic mixed modifiers with a dual behavior as inhibitors or activators depending on substrate concentration and characterized by values of α and β both larger than 1 and $\alpha > \beta$. This rare mechanism, observed previously in the case of compound NSC13345 [12], is characterized by a critical substrate concentration, which in terms of the $[S]/K_m$ ratio σ , can be defined as $\sigma^0 = (\beta - \alpha)/(1 - \beta)$. Actually, this calculation is based on the $[S]/K_s$ ratio that, in virtue of the quasi-equilibrium assumption made with the use of the specific velocity plot, is approximated to $[S]/K_m$. At substrate concentrations that give $[S]/K_m$ smaller than σ^0 these compounds behave as inhibitors, while at substrate concentrations giving $[S]/K_m$ larger than σ^0 they behave as activators. If the substrate concentration is such that $[S]/K_m = \sigma^0$, no effect is observed. Compound 6, which is the only representative of the

Table 1. Modifiers used in this work.

Identifier	IUPAC name	Supplier (Cat. No.)
Compound 1	3-Amino-4-[(tetrahydro-2-furanyl)methyl]amino] benzoic acid	Chembridge Corp. (7607879)
Compound 2	Piperidin-2-yl-[2-(trifluoromethyl)-6-[4-(trifluoromethyl)phenyl]pyridin-4-yl] methanol	The NCI/DTP Open Chemical Repository (NSC305798)
Compound 3	N-Benzoyl-p-aminosalicylic acid	The NCI/DTP Open Chemical Repository (NSC159686)
Compound 4	(2-Biphenyl)methyl malonic acid	The NCI/DTP Open Chemical Repository (NSC94914)
Compound 5	1, 4, 7, 9B-Tetraazaphenylene	The NCI/DTP Open Chemical Repository (NSC81462)
Compound 6	1-[1-(2-Phenylethyl)-4-piperidinyl]-methanamine	Chembridge Corp. (4009651)
Compound 7	2-[(3-Nitrophenyl)carbamoyl]benzoic acid	The NCI/DTP Open Chemical Repository (NSC408860)
Compound 8	Methyl N-[1-(4-methoxyphenyl)-2,5-dioxopyrrolidin-3-yl]glycinate	Chembridge Corp. (5855661)

doi:10.1371/journal.pone.0106642.t001

Table 2. Binding energies and equilibrium dissociation constants calculated by AutoDock.

Modifier	Target site ¹	Estimated free energy of binding ² (kcal/mol)	Estimated equilibrium dissociation constant ³ (μ M)
Compound 1	Site 3	-5.98	42
	Site 4	-8.24	0.9
Compound 2	Site 1	-6.44	19
	Site 2	-5.52	90
Compound 3	Site 5	-6.73	12
	Site 6	-5.49	95
Compound 4	Site 5	-6.42	20
	Site 6	-6.33	23
Compound 5	Site 1	-8.59	0.5
	Site 2	-5.95	43
Compound 6	Site 3	-6.78	11
	Site 4	-7.18	5.0
Compound 7	Site 5	-7.52	3.0
	Site 6	-8.05	1.3
Compound 8	Site 3	-6.28	25
	Site 4	-4.33	670

¹All sites are number according to Figure 1B.

²Corresponding to the lowest energy binding modes shown in Figure 1C.

³Output by AutoDock as estimated "inhibition" constant calculated from the estimated free energy of binding at 298.15 K.
doi:10.1371/journal.pone.0106642.t002

third group, behaved as a hyperbolic competitive inhibitor in which the ESA complex is turned over with the same catalytic constant as the ES complex ($\beta = 1$). The fourth group contains a pair of hyperbolic mixed inhibitors with uncompetitive character, compounds 7 and 8, which differ from the other groups for their values of α and β being less than 1. Kinetic experiments with these two compounds were the most erratic in the series investigated. With $\alpha > \beta$ the mechanism is classified as hyperbolic mixed, predominantly uncompetitive inhibition, while with $\alpha = \beta$ the mechanism is called hyperbolic uncompetitive inhibition. In the former case, the sheaf of straight lines in the specific velocity plot should intersect the ordinate at 1 and the abscissa at a negative value. In the second case, the lines should intersect at ordinate 1 and abscissa 0. The best-fit results from the relative velocity versus modifier concentration plots shown in Figure 3 point to $\alpha = \beta$. It remains established that, for this type of modifiers, inhibitory efficiency increases with increasing substrate concentration.

Inhibition of macromolecular substrate hydrolysis

The use of synthetic substrates is practical for the determination of the mechanisms of modification. However, these results do not necessarily represent the effect of modifiers on the hydrolysis of macromolecular substrates. Therefore, all modifiers were also tested with two macromolecular substrates, calf skin collagen and azocasein. The collagenolytic activity was tested quantitatively by following the fragmentation of the characteristic type I collagen pattern into smaller products on SDS-PAGE after incubation of the samples with cathepsin K in the presence and absence of modifiers. Of the eight modifiers, only compounds 1 and 2 showed a concentration-dependent inhibitory effect (Figure 4A), whereas no inhibitory, or activating, activity was observed for the remaining six compounds. IC_{50} values were estimated at around 1 mM and 0.5 mM for compounds 1 and 2, respectively. Similarly, all compounds showed little or no inhibitory activity in azocasein degradation assays (Figure 4B). Compounds 1

through 6 reduced the activity of cathepsin K by less than 25 percent with compounds 1 and 6 being the most effective, while compounds 7 and 8 were ineffective. Due to the generally small inhibitory effects, IC_{50} values could be reliably determined only for the two most effective compounds. The calculated IC_{50} value for compound 1 was $730 \pm 190 \mu$ M and was thus comparable to the value estimated for the degradation of collagen. The calculated IC_{50} value for compound 6 was $1100 \pm 240 \mu$ M. Thus, IC_{50} was about 4-fold higher than K_A for both compounds.

For the needs of the activity space analysis presented in one of the following subsections, azocasein degradation assays were also performed in the presence of glycosaminoglycans and clusterin (Figure 4C). Two different concentrations of glycosaminoglycans were used (10 μ g/ml and 200 μ g/ml, respectively) to take into account the varying concentration-dependent effects of glycosaminoglycans in collagen degradation assays [17]. In all cases, glycosaminoglycans behaved as inhibitors, except the lowest concentration of chondroitin sulfate was ineffective. A marginal inhibitory effect was also observed for clusterin.

Stabilization of cathepsin K at physiological plasma pH

Apart from directly affecting the activity of their target, allosteric modifiers have the potential to affect other aspects of enzymatic activity, such as its stability under unfavorable conditions. In the case of cathepsin K, heparin and the extracellular chaperone clusterin have been shown to increase the stability of the otherwise unstable peptidase at the physiological plasma pH value of 7.4 [11,18]. The instability of cathepsin K has been shown to be due to at least partly to autoproteolytic degradation [11]. Therefore, allosteric modifiers can contribute to the stabilization of the enzyme by at least two mechanisms, i.e. by inhibiting its proteolytic activity or by shielding the susceptible regions of the molecule from proteolytic attack by directly binding to these regions.

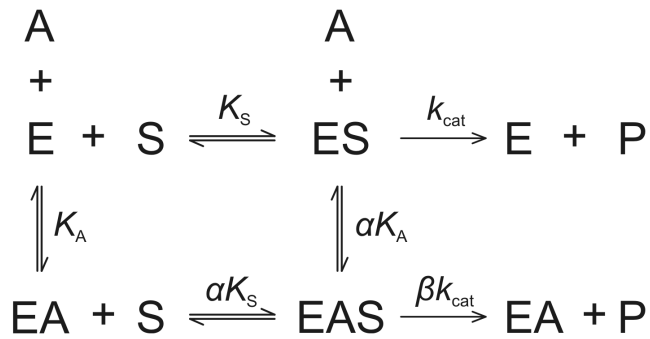


Figure 2. Reaction scheme for the general modifier mechanism. The scheme describes the interaction of enzyme (E) with substrate (S) and modifier (A). K_S and K_A are equilibrium dissociation constants of the ES and EA complexes under rapid equilibrium conditions, k_{cat} is the catalytic constant and α and β are dimensionless coefficients that determine the effect of modifier on the substrate-binding affinity of the enzyme and the catalytic rate, respectively. doi:10.1371/journal.pone.0106642.g002

Here we have examined the newly identified modifiers for their effect on the stability of cathepsin K. As shown in Table 4, compounds 1, 2 and 5 increased the stability of cathepsin K by about 2-fold, which is similar to the activation effect produced by clusterin [18]. The remaining compounds were less efficient. For the needs of the activity modification space analysis, the experiment was also performed with the previously described allosteric modifier NSC1335. This compound showed a weak stabilization effect, increasing the half-life of cathepsin K by 1.2-fold. Altogether, no clear correlation between their stabilizing effect and other kinetic or activity parameters can be determined for the modifiers. It is hence likely that their overall effects on enzyme stability are due to a combination of multiple factors, as discussed above.

The activity modification space of cathepsin K

The enzyme activity modification space is a straightforward method to combine results from different experiments and globally compare the effects of different modifiers on an enzyme. We define it as an n -dimensional space, where n is equal to the number of different experimentally determined parameters presented in the particular representation. Parameters can be plotted either in terms of thermodynamic constants, values of coefficients such as α and β in the general modifier mechanism, or relative activities. Depending on the number of dimensions, data can be plotted directly or the dimensions reduced by calculating the principal components of the activity space matrix and plotting the first two or three components. In general, the activity modification space can be viewed as a part of the complete enzyme activity space which represents all possible activity profiles of a particular enzyme, caused not only by modifiers which bind the enzyme in a defined manner, but also by other environmental parameters, such as ionic strength, pH, temperature, etc.

Here the activity modification space concept is used in two independent ways to investigate the effects of all known modifiers of cathepsin K other than active-site directed inhibitors. Data for previously known modifiers was taken either from published sources or from experiments performed specifically for this purpose, as described in the previous subsections. First, a simple two-dimensional representation was constructed only from the values of the kinetic coefficients α and β determined for the modifiers when using the synthetic substrate Z-FR \downarrow AMC (Figure 5A), the most commonly used substrate for measuring

the activity of cathepsin K and other cysteine cathepsin endopeptidases *in vitro*. This representation shows that cathepsin K can be affected by all types of hyperbolic mechanisms determined by the general modifier mechanism, including inhibition, activation and mixed inhibition/activation mechanisms, as discussed in more detail above. The distinctively colored parts of the α/β diagram represent the combinations of both coefficients that produce different types of allosteric modification, i.e. mixed inhibition, mixed inhibition/activation or non-essential activation. Hyperbolic competitive inhibitors are defined by a combination of $\alpha > 1$ and $\beta = 1$, whereas uncompetitive inhibitors are defined by the combination $0 < \alpha = \beta < 1$. By definition, no allosteric modification is possible at $\alpha = 1$ due to the absence of thermodynamic coupling between modifier and substrate [15]. A special case is represented by the point (1,1) where both coefficients α and β are equal to 1. A modifier with this combination of parameters binds to the enzyme without any measurable effect on its activity under the particular experimental conditions. One example is the recently described compound NSC13345. While the compound had no significant effect on the hydrolysis of Z-FR \downarrow AMC, other experiments have clearly shown that the compound binds and inhibits the enzyme [12]. Another class of compounds with this particular combination of coefficients are so-called liberators, represented by the recently described clusterin [18]. Again, no effect on enzymatic activity is observed in the ternary ESA complex. However, in the presence of an inhibitor, the liberator negates the effect of the inhibitor and thus “liberates” the enzyme [19]. The representation in Figure 5A also takes into account all linear mechanisms of inhibition which lie on the abscissa ($\beta = 0$). They are discriminated among each other by their respective values of the coefficient α . For linear competitive inhibitors the value of α approaches ∞ . Despite their absence from the representation in Figure 5A a number of such inhibitors have been described for cathepsin K, including endogenous macromolecular inhibitors as well as a plethora of active site-directed reversible inhibitors designed for cathepsin K targeting in medical applications (reviewed in ref. [7]). For linear uncompetitive inhibitors, $\alpha \rightarrow 0$, while linear mixed inhibitors are characterized by values of $\alpha > 1$. Finally, the mostly theoretical non-competitive mechanism would be characterized by the value of $\alpha = 1$ and $\beta = 0$.

The approach in Figure 5A summarizes the kinetic parameters and thereby aids in identifying and categorizing the kinetic mechanisms by which cathepsin K activity can be modified. It is, however, suitable only for data obtained with synthetic substrates, which allow for the determination of the individual coefficients. To evaluate comprehensively the effect of the modifiers on cathepsin K we have utilized four different assays described above, where each assay investigates a different aspect of modifier activity. The overall potency of the modifiers, and their categorization, rely on the combined effects of each modifier across all performed assays. For a global analysis, data from all four experiments was collected in the form of an activity modification matrix (Table S1) and analyzed by principal component analysis [20]. A plot of the first three principal components which account for more than 90% of the variance are shown in Figure 5B. In general, the distance of each modifier from the position of enzyme alone can be interpreted as a measure of the strength of modification. Analogously, characteristic groups formed in the activity modification space can be interpreted as distinctive modes of enzyme modification. The modifiers investigated in the present study are divided into two groups. The first group is formed by the pair of compounds 1 and 2, which act as inhibitors in all performed assays and stabilize the enzyme at neutral pH. The remaining

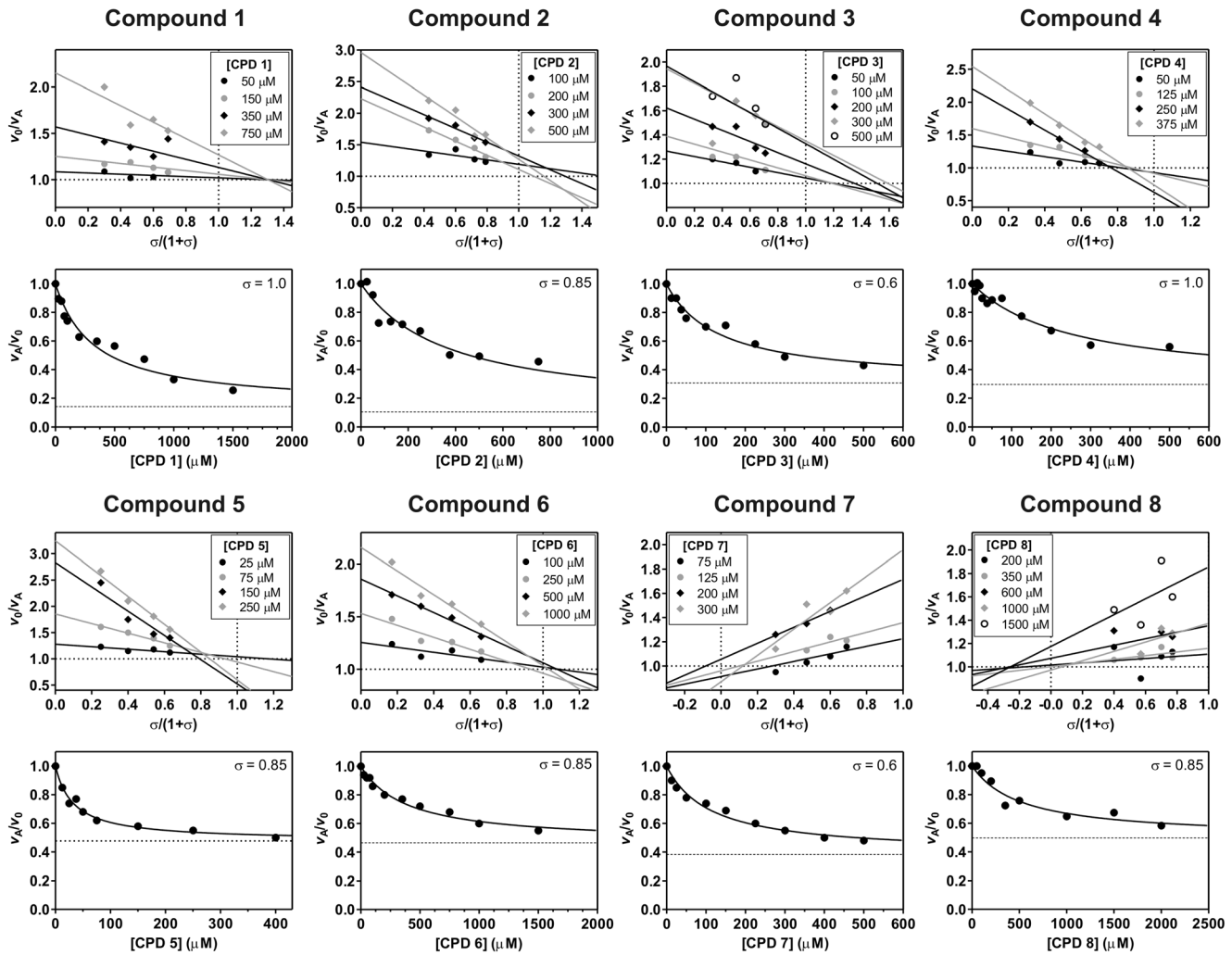


Figure 3. Kinetic mechanisms of the identified modifiers. All compounds were characterized by a combination of specific velocity plots (upper panels) and titration curves recorded at fixed substrate concentrations (lower panels). The kinetic parameters calculated from the combination of both methods are given in Table 3. All experiments were performed at $25 \pm 1^\circ\text{C}$ in 100 mM sodium acetate buffer pH 5.50 containing 1 mM EDTA and 2.5 mM DTT using the fluorogenic substrate Z-FR ↓ AMC and at a final enzyme concentration of 0.5 nM. doi:10.1371/journal.pone.0106642.g003

Table 3. Kinetic and other activity parameters of the identified modifiers.

Substrate	Z-FR ↓ AMC			Azocasein	Calf skin collagen
	α^1	β^1	K_A (μM)	Residual activity (in %)	IC_{50} (μM) ²
Compound 1	6.7 ± 3.0	0.55 ± 0.25	190 ± 20	74 ± 9	1000
Compound 2	7.2 ± 0.5	0.74 ± 0.13	150 ± 50	93 ± 7	500
Compound 3	1.7 ± 0.5	0.44 ± 0.22	110 ± 20	86 ± 4	n.i.
Compound 4	5.6 ± 1.2	1.5 ± 0.3	100 ± 40	90 ± 3	n.i.
Compound 5	4.8 ± 1.5	1.5 ± 0.2	20 ± 5	86 ± 4	n.i.
Compound 6	3.0 ± 0.2	1	270 ± 50	77 ± 8	n.i.
Compound 7	0.19 ± 0.03	0.19 ± 0.03	290 ± 20	104 ± 5	n.i.
Compound 8	0.31 ± 0.06	0.31 ± 0.06	1100 ± 200	101 ± 3	n.i.

¹Coefficients α and β multiply the kinetic parameters K_S , K_A (α) and k_{cat} (β) and thus quantify the effect of modifier on the affinity for the substrate (α) and the catalytic rate (β), respectively.

²n.i. indicates no inhibition.

doi:10.1371/journal.pone.0106642.t003

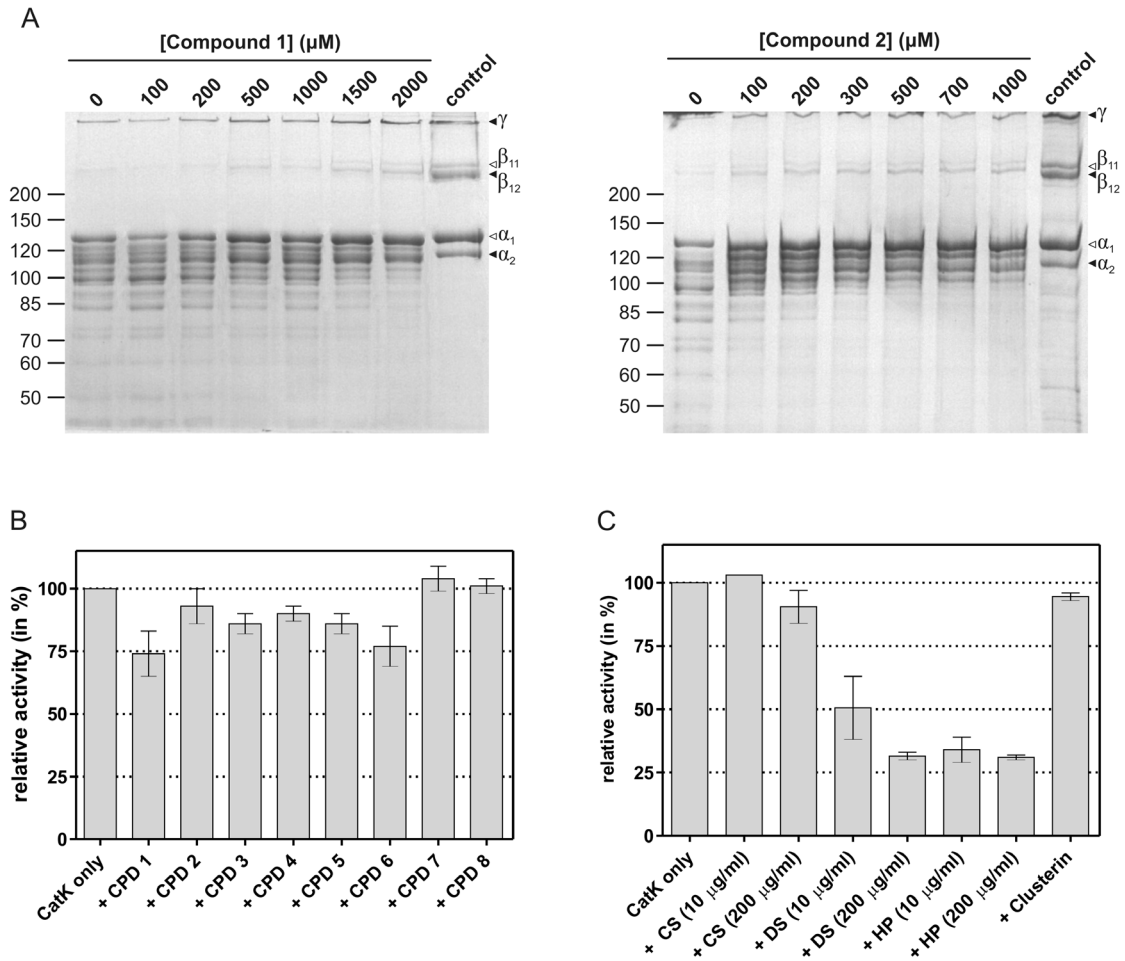


Figure 4. Effect of modifiers on the degradation of macromolecular substrates. (A) Effect of increasing concentrations of compounds 1 and 2 on the degradation of calf skin collagen. Molecular mass standards are given in kDa and positions of α , β and γ chains of intact collagen are marked by arrows. (B) Effect of modifiers on azocasein degradation. (C) Effect of glycosaminoglycans and clusterin on azocasein degradation. Two different concentrations of each glycosaminoglycan were used in the experiments (10 $\mu\text{g/ml}$ and 200 $\mu\text{g/ml}$). Error bars in panels B and C represent the SEM of three experiments. All experiments were performed in 100 mM sodium acetate buffer pH 5.50 containing 1 mM EDTA and 5 mM DTT, as described in detail in the Materials and Methods section. Abbreviations: CPD – compound, CS – chondroitin sulfate, DS – dermatan sulfate, HP – heparin. doi:10.1371/journal.pone.0106642.g004

Table 4. Effect of modifiers on the stability of cathepsin K at pH 7.4.

Modifier	Half-life (min)
enzyme only	6.4±1.2
Compound 1	13.5±2.6
Compound 2	14.3±0.7
Compound 3	8.6±0.9
Compound 4	9.5±0.8
Compound 5	13.6±0.7
Compound 6	7.5±1.1
Compound 7	10.1±1.8
Compound 8	6.5±1.0
NSC13345	7.7±0.7

The enzyme was incubated at 37°C in 50 mM HEPES buffer pH 7.40 containing 1 mM EDTA and 5 mM DTT in the presence or absence of each modifier. Aliquots were withdrawn from the reaction mixtures at regular time intervals and their residual activity measured using the substrate Z-FR ↓ AMC. doi:10.1371/journal.pone.0106642.t004

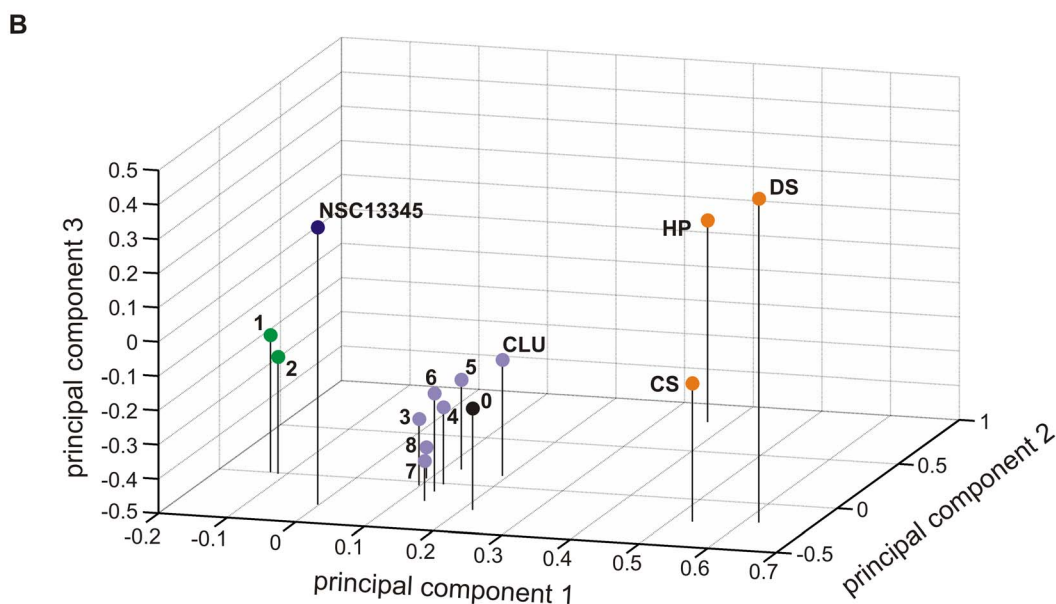
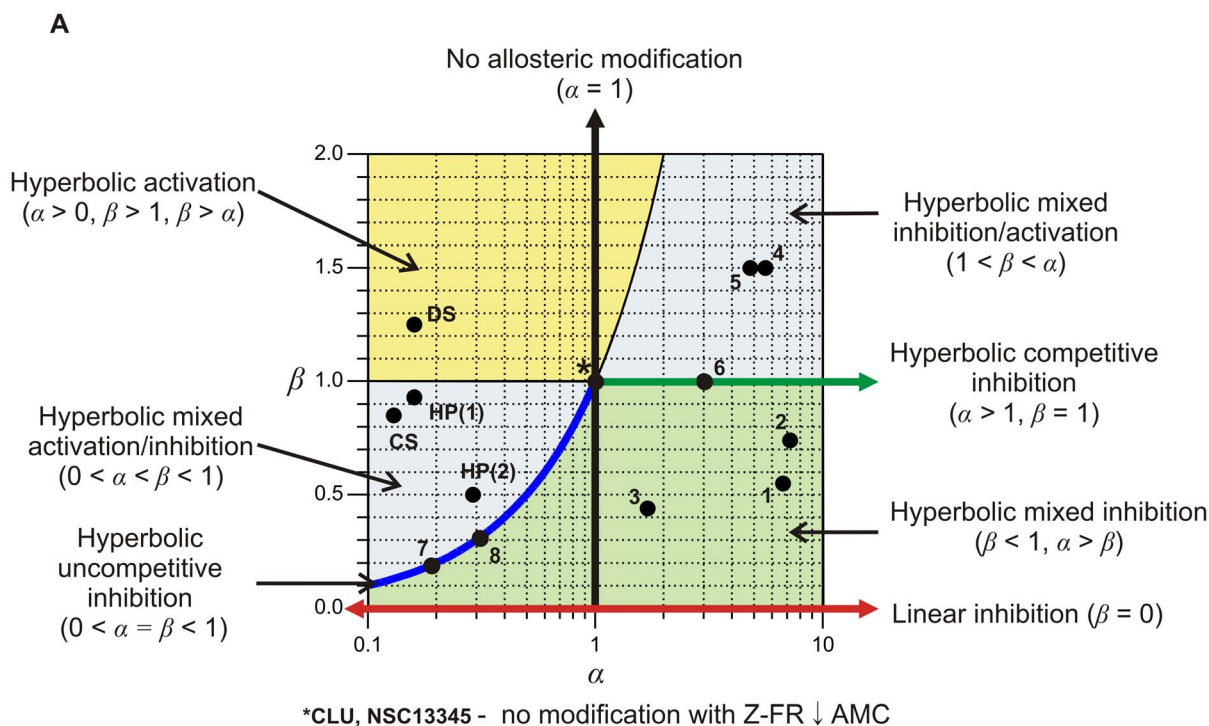


Figure 5. Activity modification space of cathepsin K. (A) The repertoire of binding affinities (parameter α) and catalytic rates (parameter β) with Z-FR \downarrow AMC as substrate. Areas of the modification space that represent distinctive modes of modification are marked accordingly. Since heparin (HP) binds to two sites on cathepsin K, data for each site are presented separately in this representation. The image was created with GraphPad Prism 5.0. (B) The global effect of all known modifiers of cathepsin K other than active site-directed competitive inhibitors was analyzed with principal component analysis. Data is presented as a three-dimensional diagram of the first three principal components, which account for more than 90% of the variance. The analysis was performed on an activity modification matrix containing data collected from all four types of assays described herein (Table S1). Vertical lines in the diagram indicate the positions of each modifier in the plane defined by principal components 1 and 2. Color-coding corresponds to the division of the modifiers into different groups, based on the results of this analysis. Modifiers identified in this study are marked with consecutively increasing Arabic numerals according to Table 1. The position of cathepsin K alone is marked by "0". The image was created with Matlab. Abbreviations: CS – chondroitin sulfate, DS – dermatan sulfate, HP – heparin, CLU – clusterin. doi:10.1371/journal.pone.0106642.g005

compounds form a loose group located in the vicinity of unmodified cathepsin K, reflecting their overall weak effect on cathepsin K. Of the previously known modifiers, NSC13345 and glycosaminoglycans, each form distinctive groups in the activity modification space distant from the position of unmodified enzyme. Not surprisingly, a significant degree of dispersion is observed in the glycosaminoglycan group (chondroitin sulfate, dermatan sulfate and heparin), resulting from their diverse effects on cathepsin K described herein and previously [11]. Last, clusterin is positioned close to unmodified cathepsin K due to its lack of effect on enzyme activity and can be considered part of the cluster formed by compounds 3 through 8.

Discussion

The primary goal of this study was to investigate the range of effects that can be produced by allosteric modifiers of cathepsin K. The theory of enzyme kinetics does not set precise limitations in terms of the effects on substrate binding affinity or catalytic rate of any enzyme. However, the activity of enzymes is restricted by their structural limitations meaning that each enzyme can exhibit only a limited range of kinetic profiles, which can result either from its intrinsic flexibility or from the binding of an allosteric modifier. The concept of activity modification space was introduced to globally analyze the effect of the identified modifiers on cathepsin K. This method is similar to the concepts of ligand-activity space (or substrate-activity space) which uses a multidimensional representation to analyze the directed evolution of enzymes using multiple different substrates or ligands (see ref. [21] and refs. therein) and chemical space of inhibitors which was recently used to investigate the selectivity of PARP inhibitors [22].

Results obtained with the eight modifiers described in this paper and previous findings show that cathepsin K has a broad activity space and that it is very sensitive to modification by specific mechanisms, such as binding of modifier, as well as non-specific mechanisms, such as variations in the ionic strength of the medium [11]. At the structural level, this is probably related to a high structural flexibility of the molecule, which can be easily explained by its broad specificity for macromolecular proteinaceous substrates. The analysis of the activity modification space (Figure 5) identified several functionally distinct groups of compounds, indicating the presence of multiple allosteric mechanisms in the molecule. Glycosaminoglycans and NSC13345 each formed clearly distinguishable groups within the activity space related to their distinctive allosteric mechanisms of action, which have been characterized both structurally and functionally [10,11,12]. The newly identified modifiers are separated from these two groups indicating that they are functionally distinct. These compounds either bind to allosteric sites elsewhere on the molecule or they elicit different allosteric responses by binding to the same site on the enzyme. The latter was observed in several systems, e.g. the conformational changes associated with GTP and GDP binding to G proteins such as Ras and cAMP-dependent protein kinase (discussed in ref. [23]). The compounds form two groups in the activity modification space, one is formed by compounds 1 and 2 which exhibit the overall strongest performance in activity and stability assays, whereas the remaining compounds form a group that is closely located near unmodified cathepsin K, reflecting their overall weak activity. Given the diverse kinetic profiles and structural diversity of the latter group, these compounds are likely to bind to several different binding sites that do not efficiently transmit allosteric signals through the molecule. A comparison of computationally predicted and experimentally determined equilibrium constants (Tables 2 and 3, respectively) shows that for all

compounds except 6 and 7 at least one of the predicted sites remains a viable candidate for the true binding site. Minor differences between both values can be easily accounted for by receptor flexibility, as observed in the case of NSC13345 where the experimentally determined binding constant was 5-fold lower than the computationally predicted value [12]. While the compounds 1 and 2 are commendable of further development for drug design, the remaining six are interesting not only for their *in vitro* kinetic effects, but also for the development of specific probes to follow cathepsin K activity *in vivo*. Currently available probes bind into the active site of their targets and thereby abolish the target's activity, binding at a distant site would, however, allow to follow the enzyme in action.

In general, the distinguishable groups in the activity space identify several trends of activity modification. The inhibitory effect of NSC13345 scales proportionally with substrate size, ranging from no interaction with Z-FR ↓ AMC to full inhibition with collagen [12], whereas the inverse trend is observed for compounds 3 through 6. Glycosaminoglycans and compounds 1 and 2 show no clear trends in this respect, however, modifier activity is again dependent on the substrate. Finally, clusterin provides additional diversity by acting as a liberator without affecting enzyme activity *per se*. While the plethora of different effects may seem erratic, it is a clear indication that allosteric mechanisms serve to fine-tune the activity of papain-like peptidases in nature, while emergency inhibition is reserved for macromolecular inhibitors.

Materials and Methods

Materials

Recombinant human cathepsin K was prepared as described previously [24]. Chemical probes were obtained from the U.S. NCI/DTP Open Chemical Repository or from ChemBridge (USA). Information on their identity and origin is given in Table 1. All compounds were soluble in aqueous buffers at concentrations of at least 2 mM, except for compound 2, which was soluble up to a maximal concentration of 1 mM. Stock solutions of all compounds were prepared in DMSO. The fluorogenic substrate Z-FR ↓ AMC (benzyloxycarbonyl-Phe-Arg-7-amino-4-methylcoumarin) was from Bachem (Switzerland).

Docking of compound libraries

Potential allosteric modifiers were predicted as described previously [12]. In short, the NCI Diversity Set III and ChemBridge Building Blocks compound libraries were obtained from the ZINC database [25] and filtered to retain only molecules with molecular masses between 150 and 500 Da and less than seven rotatable bonds. Both libraries were docked to all seven predicted allosteric sites on cathepsin K [12]. Cathepsin K coordinates were retrieved from the Protein Data Bank (PDB ID 1ATK). Docking was first performed with the UCSF Dock 6.4 program suite (available online at <http://dock.compbio.ucsf.edu/>) using default values of all parameters. The best hits (about top 10%) were selected based on the binding energies calculated by the internal scoring function and re-docked to the receptor using AutoDock 4.2 [26]. The best binders were again selected on the basis of the binding affinities calculated by the AutoDock internal scoring function. Each ligand was parameterized to allow the maximum number of rotatable bonds, but excluding amide and ring bonds. Docking was performed with the Lamarckian genetic algorithm. For each ligand, 20 runs were performed using a population of 300 individuals over 3000 generations, starting from random orientations and conformations. For presentation of the results, the lowest

energy poses calculated by AutoDock were selected and visualized with PyMOL. The receptor surface was colored according to its electrostatic potential calculated with APBS [27].

Kinetic measurements

All kinetic measurements were performed in 100 mM sodium acetate buffer pH 5.50 containing 1 mM EDTA and 2.5 mM dithiothreitol (DTT) in single-use acrylic cuvettes (1×1 cm) thermostatted at 25±1°C with magnetic stirring. The hydrolysis of Z-FR ↓ AMC was followed fluorometrically at λ_{ex} 383 nm and λ_{em} 455 nm. All tested modifiers were analyzed spectroscopically to determine their absorption properties under the conditions used in the assays. When necessary, appropriate corrections were made to the calculated reaction rates to take into account the inner filter effect [28].

Kinetic model

The effects of the modifiers on enzymatic activity were analyzed with the general modifier mechanism [14] which describes the interaction between enzyme (E), substrate (S) and modifier (A), as shown in Figure 2. Assuming that binding of A to E and ES is at quasi-equilibrium and that steady-state conditions are established in the fluxes around ES and ESA (i.e. the catalytic steps), the reaction rate for the mechanism in Figure 2 is defined by Equation 1:

$$v_A = \frac{v_0(1+\sigma) \left(1 + \beta \frac{[A]}{\alpha K_A}\right)}{1 + \frac{[A]}{K_A} + \sigma \left(1 + \frac{[A]}{\alpha K_A}\right)} \quad (1)$$

where v_0 is the reaction rate in the absence of modifier, K_A is the equilibrium dissociation constant of the EA complex, α and β are dimensionless coefficients and $\sigma = [S]/K_m$, approximated to $[S]/K_s$. To determine the mechanism of interaction, plot the kinetic data and calculate preliminary values of the interaction parameters, the specific velocity plot was used [16]. For this purpose, Equation 1 is rewritten as:

$$\frac{v_0}{v_A} = \frac{[A] \left(\frac{1}{\alpha K_A} - \frac{1}{K_A}\right)}{1 + \beta \frac{[A]}{\alpha K_A}} \frac{\sigma}{1 + \sigma} + \frac{1 + \frac{[A]}{K_A}}{1 + \beta \frac{[A]}{\alpha K_A}} \quad (2)$$

The plot of v_0/v_A versus $\sigma/(1+\sigma)$ produces straight lines with intersection points at $v_0/v_A = 1$, regardless of the interaction mechanism. The parameters α , β and K_A are determined by reploting the extrapolated values of the straight lines at $\sigma/(1+\sigma) = 0$ (a) and $\sigma/(1+\sigma) = 1$ (b) versus $1/[A]$ in the form:

$$\frac{a}{a-1} = \frac{\alpha K_A}{\alpha - \beta} \frac{1}{[A]} + \frac{\alpha}{\alpha - \beta} \quad (3)$$

$$\frac{b}{b-1} = \frac{\alpha K_A}{1 - \beta} \frac{1}{[A]} + \frac{1}{1 - \beta} \quad (4)$$

Collagenolytic assays

Soluble calf skin collagen (Worthington Biochemical Corporation, USA) was diluted in 100 mM sodium acetate buffer pH 5.50 containing 1 mM EDTA and 5 mM DTT to a final concentration of 0.5 mg/ml. Experiments were started by addition of cathepsin K (final enzyme concentration 0.6 μM) to the reaction mixture containing collagen and increasing concentrations of modifier. All assays were performed at 28°C with shaking on an Eppendorf Thermomixer Comfort. Samples were incubated for 2 hours and the reactions then blocked by addition of E-64 to a final concentration of 1 μM. 10 μg of digested collagen were analyzed by SDS-PAGE on 8% polyacrylamide gels and protein bands visualized by Coomassie Brilliant Blue R-250 staining.

Azocasein degradation assays

Azocasein (Sigma-Aldrich, USA) was dissolved in 100 mM sodium acetate buffer pH 5.50 containing 1 mM EDTA and 5 mM DTT at a concentration of 3 mg/ml. Samples were incubated with 0.1 μM cathepsin K for 30 min at 37°C in the presence of each compound. All compounds were used at concentrations increasing up to 2 mM, except compound 2, which was used at a maximal concentration of 1 mM due to its limited solubility at higher concentrations. Reactions were stopped by addition of trichloroacetic acid (final concentration 5% v/v), centrifuged, clear supernatants mixed with 0.5 M NaOH and the concentrations of solubilized peptides determined by measuring the absorbance at 440 nm. All experiments were performed in triplicates and appropriate blanks were run to take into account the absorption of azocasein and compounds alone. IC_{50} values were determined as described previously [12].

Thermal stability of cathepsin K in the presence of modifiers

Cathepsin K was incubated at 37°C in 50 mM Hepes buffer pH 7.40 containing 1 mM EDTA and 5 mM DTT in the presence or absence of each modifier. The final enzyme concentration was 0.2 μM and final compounds concentrations were 2 mM, except compound 2, which was used at a concentration of 1 mM. All experiments were performed in Eppendorf Protein LoBind Tubes to minimize the adsorption of proteins on tube walls. Aliquots were withdrawn from the reaction mixtures at regular time intervals and their residual activity measured using the substrate Z-FR ↓ AMC (5 μM final concentration). Inactivation curves had exponential profiles that were fitted with the mathematical model for one-phase exponential decay to determine the half-life of the enzyme in each reaction mixture.

Principal component analysis of the activity modification space

For analysis, the results of each experiment were presented as one data set termed the activity modification matrix (Table S1). Data from experiments performed with the synthetic substrate Z-FR ↓ AMC was included as the geometric mean of residual activities (v_A) at a low substrate concentration ($0.1 \times K_m$) and a high substrate concentration ($10 \times K_m$), both in the presence of a saturating modifier concentration ($10 \times K_A$). For each modifier, Equation 1 was used to calculate the appropriate values of v_A using experimentally determined values of the parameters α and β . Stabilization of the enzyme was expressed as $\ln(sf)$, where the stabilization factor sf is the ratio of half-lives determined in the presence and absence of modifier. Azocasein degradation data was input as the fraction of residual enzyme activity (a value of 1 equals

the activity of unmodified enzyme). For collagenolytic assays, qualitative values of 1 and 0 were used to signify inhibition and lack thereof, respectively. For glycosaminoglycans, a value of -1 was used to signify their concentration- and environment-dependent activation/inhibition activity. Principal component calculations and graphical visualizations were performed in the Matlab environment.

Supporting Information

Table S1 The activity modification matrix used for principal component analysis.
(PDF)

References

- Monod J, Wyman J, Changeux JP (1965) On the Nature of Allosteric Transitions: A Plausible Model. *J. Mol. Biol.* 12: 88–118.
- Novinec M, Lenarčič B (2013) Papain-like peptidases: structure, function, and evolution. *Biomol. Concepts* 4: 287–308.
- Tsai CJ, del Sol A, Nussinov R (2008) Allosteric: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol.* 378: 1–11.
- Nussinov R, Tsai CJ (2013) Allostery in disease and in drug discovery. *Cell* 153: 293–305.
- Almeida PC, Nantes IL, Rizzi CC, Judice WA, Chagas JR, et al. (1999) Cysteine proteinase activity regulation. A possible role of heparin and heparin-like glycosaminoglycans. *J. Biol. Chem.* 274: 30433–30438.
- Almeida PC, Nantes IL, Chagas JR, Rizzi CC, Faljoni-Alario A, et al. (2001) Cathepsin B activity regulation. Heparin-like glycosaminoglycans protect human cathepsin B from alkaline pH-induced inactivation. *J. Biol. Chem.* 276: 944–951.
- Novinec M, Lenarčič B (2013) Cathepsin K: a unique collagenolytic cysteine peptidase. *Biol. Chem.* 394: 1163–1179.
- Li Z, Hou WS, Brömme D (2000) Collagenolytic activity of cathepsin K is specifically modulated by cartilage-resident chondroitin sulfates. *Biochemistry* 39: 529–536.
- Li Z, Hou WS, Escalante-Torres CR, Gelb BD, Brömme D (2002) Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate. *J. Biol. Chem.* 277: 28669–28676.
- Li Z, Kienetz M, Cherney MM, James MN, Brömme D (2008) The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex. *J. Mol. Biol.* 383: 78–91.
- Novinec M, Kovačič L, Lenarčič B, Baici A (2010) Conformational flexibility and allosteric regulation of cathepsin K. *Biochem. J.* 429: 379–389.
- Novinec M, Korenč M, Cafilisch A, Ranganathan R, Lenarčič B, et al. (2014) A novel allosteric mechanism in the cysteine peptidase cathepsin K discovered by computational methods. *Nat. Commun.* 5: 3287.
- Halabi N, Rivoire O, Leibler S, Ranganathan R (2009) Protein Sectors: Evolutionary Units of Three-Dimensional Structure. *Cell* 138: 774–786.
- Botts J, Morales M (1953) Analytical description of the effects of modifiers and of enzyme multivalency upon the steady state catalyzed reaction rate. *Trans. Faraday Soc.* 49: 696–707.
- Fenton AW (2008) Allostery: an illustrated definition for the ‘second secret of life’. *Trends Biochem. Sci.* 33: 420–425.
- Baici A (1981) The specific velocity plot. A graphical method for determining inhibition parameters for both linear and hyperbolic enzyme inhibitors. *Eur. J. Biochem.* 119: 9–14.
- Li Z, Yasuda Y, Li W, Bogoy M, Katz N, et al. (2004) Regulation of collagenase activities of human cathepsins by glycosaminoglycans. *J. Biol. Chem.* 279: 5470–5479.
- Novinec M, Lenarčič B, Baici A (2012) Clusterin is a specific stabilizer and liberator of extracellular cathepsin K. *FEBS Lett.* 586: 1062–1066.
- Keleti T (1967) The liberator. *J. Theor. Biol.* 16: 337–355.
- Ringner M (2008) What is principal component analysis? *Nat. Biotechnol.* 26: 303–304.
- Mannervik B, Runarsdottir A (2010) The quest for molecular quasi-species in ligand-activity space and its application to directed enzyme evolution. *FEBS Lett.* 584: 2565–2571.
- Wahlberg E, Karlberg T, Kouznetsova E, Markova N, Macchiarulo A, et al. (2012) Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors. *Nat. Biotechnol.* 30: 283–288.
- Nussinov R, Tsai CJ (2014) Unraveling structural mechanisms of allosteric drug action. *Trends Pharmacol. Sci.* 35: 256–264.
- Novinec M, Pavšič M, Lenarčič B (2012) A simple and efficient protocol for the production of recombinant cathepsin V and other cysteine cathepsins in soluble form in *Escherichia coli*. *Protein Expr. Purif.* 82: 1–5.
- Irwin JJ, Shoichet BK (2005) ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* 45: 177–182.
- Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* 30: 2785–2791.
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98: 10037–10041.
- Palmier MO, Van Doren SR (2007) Rapid determination of enzyme kinetics from fluorescence: overcoming the inner filter effect. *Anal. Biochem.* 371: 43–51.

Acknowledgments

The authors thank the U.S. NCI/DTP Open Chemical Repository (<http://nci.cancer.gov>) for providing free samples of the probes used in this study.

Author Contributions

Conceived and designed the experiments: MN BL AB. Performed the experiments: MN. Analyzed the data: MN BL AB. Contributed reagents/materials/analysis tools: MN BL AB. Contributed to the writing of the manuscript: MN AB.