Docking-Based Virtual Screening of Covalently Binding Ligands: An Orthogonal Lead Discovery Approach

Jörg Schröder,[†,∥,⊥] Anette Klinger,[†,‡,∥,#] Frank Oellien,[†] Richard J. Marhöfer,[†] Michael Duszenko,[‡] and Paul M. Selzer*[,†,‡,§]

[†]MSD Animal Health Innovation GmbH, Zur Propstei, D-55270 Schwabenheim, Germany

[‡]Interfakultäres Institut für Biochemie, Universität Tübingen, Hoppe-Seyler-Straße 4, D-72076 Tübingen, Germany

[§]Wellcome Trust Centre for Molecular Parasitology, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, 120 University Place, Glasgow G12 8TA, Scotland, United Kingdom

S *Supporting Information*

ABSTRACT: In pharmaceutical industry, lead discovery strategies and screening collections have been predominantly tailored to discover compounds that modulate target proteins through noncovalent interactions. Conversely, covalent linkage formation is an important mechanism for a quantity of successful drugs in the market, which are discovered in most cases by hindsight instead of systematical design. In this article, the implementation of a docking-based virtual screening workflow for the retrieval of covalent binders is presented considering human cathepsin K as a test case. By use of the docking conditions that led to the best enrichment of known actives, 44 candidate compounds with unknown activity on cathepsin K were finally selected for experimental evaluation. The most potent inhibitor, 4-(N-phenylanilino)-6-pyrrolidin-1-yl-1,3,5-triazine-2-carbonitrile (CP243522), showed a K_i of 21 nM and was confirmed to have a covalent reversible mechanism of inhibition. The presented approach will have great potential in cases where covalent inhibition is the desired drug discovery strategy.
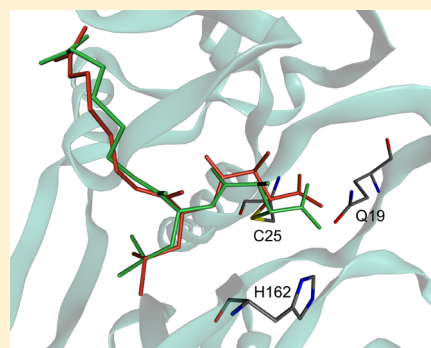
INTRODUCTION

A common concept in pharmaceutical drug design is the discovery of small molecules that predominantly undergo noncovalent interactions, such as van der Waals and electrostatic interactions in protein–ligand binding.[1,2] Besides noncovalent interactions, it is also possible that a covalent linkage is formed between protein and ligand upon their binding. Ligand molecules with reactive functional groups are able to covalently bind to proteins, DNA, or glutathion, which may result in unfavorable toxicological outcomes.[3,4] In addition, such ligands often exhibit an irreversible or mechanism-based inhibition toward a given protein target which requires more complex assay procedures.[5] Thus, it is not surprising that compounds or screening hits showing this binding profile were frequently filtered out in common screening campaigns or even precluded during the setup of compound stocks. However, successful examples of covalently binding drugs are available in the market, and promising candidates are in clinical research, such as the prostaglandin endoperoxidase synthase (cyclooxygenase) inhibitor aspirin,[6] the proton pump inhibitors (PPIs) such as omeprazole, esomeprazole, or lansoprazole,[7,8] the monoamine oxidase B (MAO-B) inhibitors selegiline and rasagiline,[9,10] the proteasome inhibitor bortezomib,[11] the hepatitis C virus (HCV) inhibitors boceprevir and telaprevir,[12,13] or the antiparasitic drug eflornithine (DFMO) which inhibits the enzyme ornithine decarboxylase (ODC) in trypanosomes.[14,15]

These examples provide clear evidence that covalent binders can be safe and effective therapeutics. Moreover, covalently binding drugs that form persistent, nonlabile covalent bonds can exhibit unique therapeutic advantages including rapid onset of inhibition, greater potency, longer duration of drug action, and potent and persistent activity against mutations that would otherwise lead to drug resistance.[16] An opportunity for covalent inhibitor design may be bacterial, viral, and parasitic diseases because the different protein targets have different turnover rates and the corresponding drug is only administered for a short period of time, thereby reducing any off-target toxicity.[17] Thus, covalent binding is an interesting drug discovery strategy to the regulation of certain protein targets. While in common noncovalent assays the covalent binding of a compound was often realized by hindsight, the strategy presented here was the implementation of a docking-based virtual screening (DBVS) workflow for the retrieval of covalently binding ligands right from the start of the lead discovery phase. This in silico approach involves the covalent docking of a screening compound database into the active site of the 3D structure of the target, followed by the ranking and selection of putative binders using suitable scoring functions. DBVS has gained attention in pharmaceutical industry over the past decades and

successfully identified novel scaffolds for different targets.[18] An issue in the past was the limited availability of software applications that are capable of covalently docking molecules. Today, a number of software packages for protein−ligand docking (e.g., GOLD,[19,20] FlexX,[21] ICM,[22,23] or QXP/FLO[24]) also offer implementations for covalent docking, which successfully have been applied to oligopeptidase[25] and proteasome inhibitors[26] as well as to the cathepsin B-like enzyme of *Eimeria tenella*[27] in low-throughput range. In contrast, the herein presented DBVS strategy identifies covalent binders in the high-throughput range in large compound libraries, e.g., different commercial libraries. Thus, only a small set of putative active compounds was selected and tested in in vitro assays as opposed to the high-throughput screening of an entire compound library.

In order to verify the covalent docking approach, human cathepsin K (EC 3.4.22.38) was selected as the target enzyme and test case. Cathepsins are lysosomal proteases, and many of them belong to the papain-like cysteine protease family.[28] This protease family originates from a wide variety of organisms like plants (e.g., papain, caricain), mammalians (e.g., cathepsins B, L, K), and parasitic protozoans (e.g., cruzain, falcipains, or cysteine proteases from *Leishmania* known as CPA, CPB, and CPC, for example).[29−31] Human cathepsin K, a 27 kDa cysteine protease, is highly and selectively expressed in osteoclast cells important in bone degradation and formation.[32] It cleaves type I collagen, the predominant component of the extracellular matrix, at multiple sites by nucleophile attack of the thiolate from C25 and subsequent proton donation from H162.[33,34] Thus, inhibition of cathepsin K represents a potential therapeutic approach for diseases characterized by excessive bone resorption such as osteoporosis.[35] Many inhibitors active against cathepsin K in vitro have been designed, which include derivatives of aldehydes,[36] vinyl sulfones,[37] ketobenzoxazoles,[38] hydroxymethyl ketones,[39] aminomethyl ketones,[40,41] and most recently nitriles.[42,43] However, the only specific cathepsin K inhibitor currently undergoing phase III studies is odanacatib (Merck & Co.). Odanacatib is a nonbasic selective inhibitor with good pharmacokinetic parameters such as minimal in vitro metabolism, long half-life, and oral bioavailability.[44] Examples of discontinued cathepsin K inhibitor developments include relacatib (GlaxoSmithKline) and balicatib (Novartis). Relacatib was discontinued following a phase I study that showed possible drug−drug interactions with the commonly prescribed medications paracetamol, ibuprofen, and atorvastatin.[45] Balicatib phase II trials were discontinued because of dermatologic adverse effects, including a morphea-like syndrome.[46]

Like odanacatib, relacatib, and balicatib, most cathepsin K inhibitors utilize a reactive electrophilic warhead that interacts with the active site thiol moiety.[47] In all three cases electrophilic nitrile or ketone groups act as warheads that mediate covalent reversible mechanism of inhibition.

Despite of this fact, most experimental and virtual screening campaigns performed on cathepsin K in the past have been set up to identify novel inhibitors that bind to the protease through noncovalent interactions.[48,49] In order to implement the DBVS approach for the retrieval of covalent binders in high-throughput range, one of the most popular molecular docking packages GOLD was used. In contrast to other covalent docking applications, GOLD allows a generic and ligand-independent way to link electrophilic chemical groups to nucleophile terminals within the binding site of a protein as

part of the covalent docking process. The basic chemical reactions of the warheads used in this study are shown in Figure 1.
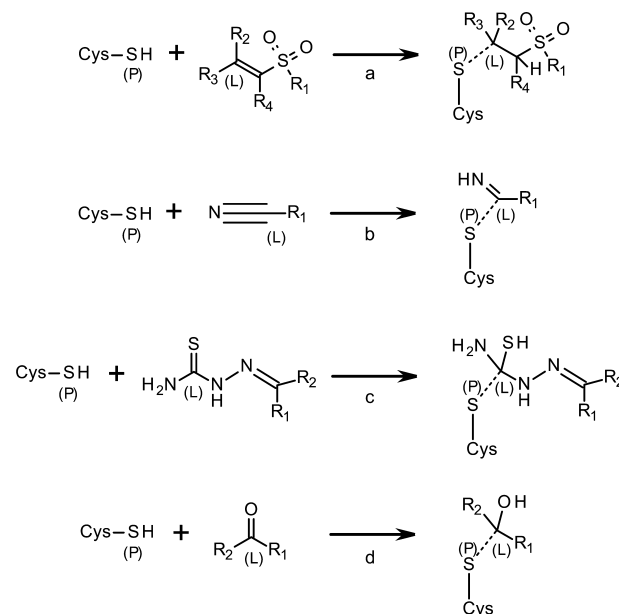

Figure 1. Proposed covalent binding of warheads. Atoms that form the covalent bond are labeled P and L for the atom in the protein and in the ligand. The covalent bond between P and L is indicated with a dash line. (a−d) Reaction between the vinyl sulfone, nitrile,[43] thiosemicarbazone,[50] and carbonyl warhead of the ligand and the catalytic cysteine residue of the protein.
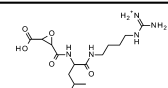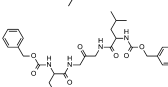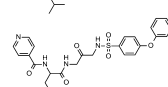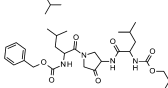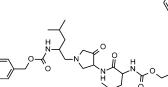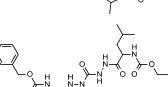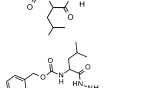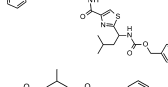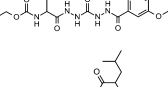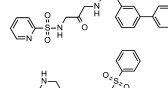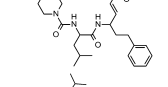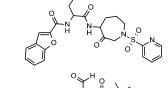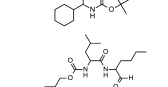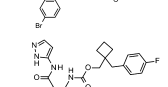
Thus, a focused library was built by application of an automated substructure searching and warhead preparation routine for compounds bearing electrophilic chemical groups of interest from a pool of commercial databases. Covalent docking of the focused library into the crystal structure of human cathepsin K finally identified a set of 44 compounds as putative inhibitors. Experimental evaluation confirmed 21 compounds as covalent reversible binding inhibitors. Out of these, three inhibitors exhibited nanomolar potency on cathepsin K and were qualified as lead structures. The results support the successful implementation of a high-throughput covalent docking strategy in lead discovery.

■ RESULTS AND DISCUSSION

Cognate Covalent Ligand Docking Accuracy. To validate the applicability of GOLD for covalent docking, the capabilities of the implemented algorithm were explored in a cognate covalent docking scenario. A data set of 14 cathepsin K X-ray complexes with covalently bound ligands along with their corresponding binding affinities were selected from the PDB (Table 1).

A major bottleneck in DBVS is finding a suitable scoring function for the target of interest to enable the identification of potential active compounds. In the present study, the two scoring functions GoldScore and ChemScore available in the covalent docking mode of GOLD were evaluated for their ability to accurately dock each of the cognate ligands back into its respective protein structure. To remove any conformational and positional information from the ligand before docking, a low energy conformation was generated by application of the

Table 1. Results of Cognate Covalent Ligand Docking for a Set of 14 Cathepsin K Complexes[a]

Compound	Entry	Mean GoldScore	Mean RMSD [Å]	Mean ChemScore	Mean RMSD [Å]	Values from literature	
						Potency measure	Potency human cathepsin K [nM]
(structure)	1ATK	47.5	0.9	14.4	1.3	K_i[51,52]	1.4
(structure)	1AU0	48.1	4.1	17.3	3.1	$K_{i,app}$[53]	22
(structure)	1AU2	46.7	4.6	16.5	11.5	$K_{i,app}$[53]	13
(structure)	1AU3	53.5	1.6	17.0	1.5	$K_{i,app}$[40]	2.3
(structure)	1AU4	53.1	1.8	17.8	3.4	$K_{i,app}$[40]	3.5
(structure)	1AYU	45.2	2.9	11.5	3.4	$K_{i,app}$[54]	0.7
(structure)	1AYV	50.3	4.8	14.3	3.6	$K_{i,app}$[54]	10
(structure)	1AYW	55.5	4.1	20.5	6.5	ND[b,54]	ND[b]
(structure)	1BGO	62.8	1.1	20.2	1.6	$K_{i,app}$[55]	3.5
(structure)	1MEM	46.8	1.3	21.3	5.9	K_i[56]	11.5
(structure)	1NLJ	6.0	4.2	4.0	2.8	K_i[57]	0.16
(structure)	1Q6K	29.7	1.1	16.0	1.1	IC_{50}[58]	26
(structure)	1SNK	36.1	2.0	14.8	9.5	IC_{50}[59]	0.25
(structure)	1TU6	41.6	1.9	14.6	2.3	IC_{50}[60]	0.83

[a]Calculations were done with the top scoring poses of ten docking runs. [b]ND, not determined.

3D structure generator CORINA[61] on the 2D structure of the ligand. Only chirality information from the native ligand structures was carefully considered during conformation generation. Subsequently, a transition state and stereoisomer generator based on the software package CACTVS[62] in combination with an energy minimizing protocol was applied (see Experimental Section for details). This generator converts an electrophilic warhead into the corresponding tetrahedral transition state. An example is shown in Figure 2 for the ligand E64 of the X-ray structure 1ATK, where (a) and (c) are structures of E64 before and after the covalent linkage is formed with the protein, respectively.

Accurate modeling of the tetrahedral bound state of covalent inhibitors is an important prerequisite for the proper prediction of the bound structure. Many inhibitors undergo extensive conformational changes upon formation of the covalent interaction, as the atom being attacked switches from planar to tetrahedral geometry. Additionally, formation of the tetrahedral state usually gives rise to two stereoisomers, one of which is likely to be more active than the other. As it is impossible to predict a priori which will be the most active, both stereoisomers have to be generated. The structure generator protocol also performs a renumbering of the ligand atoms in the way that the attached sulfur dummy receives the atom number 1 (Figure 3B). This allows a generic and ligand-
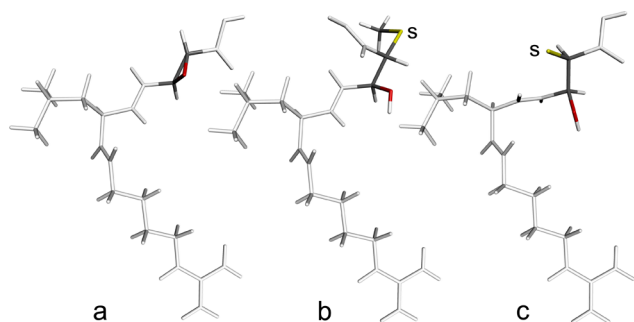
Figure 2. Illustration of the E64 structure preparation for covalent docking with GOLD. (a) An intact E64 molecule with the epoxide ring identified by the transition state generator. (b) Breaking the epoxide ring with —SCH₃ dummy atoms. The structure is not minimized so that the newly added atoms do not satisfy molecular geometry. (c) After minimization, molecular geometry is satisfied and the temporary —CH₃ group is removed, resulting in a free valence on the S-atom. E64 is depicted in capped-stick representation and colored light gray. Atoms modified during transition state generation are colored by atom type.

independent way to fit the sulfur dummy of the ligand onto the active site sulfur atom of C25 as part of the covalent docking process.

For each ligand, 10 top scoring poses resulting from 10 independent covalent docking runs were retained for further analysis. The docking accuracy of the top scored ligand pose relative to its crystallographic orientation was measured by the heavy atom root-mean-squared deviation (rmsd).[63] If a docked pose had a heavy-atom rmsd lower than or equal to 2.0 Å, it was regarded as a reasonable pose. Mean rmsd values and mean scoring values of the top scoring poses are listed in Table 1. The examination of the top scoring poses using initial generated ligand conformations provided the most difficult test for reproducing the X-ray complexes because the crystallographically determined coordinates were not supplied to GOLD as input. This procedure mimics the "real world" situation in which the optimal docking pose is not known a priori.

The mean rmsd values of the top scoring poses ranged from 0.9 to 4.8 Å for GoldScore and from 1.1 to 11.5 Å for ChemScore (Table 1). These results imply that mean rmsd values can be skewed because of a relatively small number of covalent docking poses having very large rmsd values, while the rmsd value itself has a lower bound of 0.0 Å. Using 2.0 Å as a cutoff, GoldScore correctly identified top scoring poses with a mean rmsd below 2.0 Å for seven complexes, while ChemScore was able to reach this threshold for four complexes. By comparison of the GoldScore and ChemScore scoring values to the mean rmsd values of the top scoring poses, it appeared that low-rmsd poses were not consistently scored better than high-rmsd poses (Table 1). This suggests that the scoring functions, in general, are prone to errors in which the poses closest to the X-ray conformation do not consistently receive the best score. One reason for this observation might be the fact that the current scoring functions do not reward the covalent linkage formation over noncovalent docking. Instead, an angle-bending energy term for the link atom is included in the calculation of the covalent link energy.[64] The contribution of this energy term (S_{cov} for GoldScore, E_{cov} for ChemScore) depends on the geometry of the ligand and on the position of the link atom because the cathepsin K structure is held rigid apart from protein and ligand OH and NH_3^+ hydrogen atoms, whose positions were optimized during docking runs. To compensate clashes and poor link geometries, a clash term and a torsion term are included in the calculation of the ChemScore scoring function, which may additionally influence its accuracy. From noncovalent docking validations using GOLD it is known that ChemScore gives equally good docking results as GoldScore for "drug-like" and "fragment-like" ligands, but for larger ligands with many rotatable bonds, GoldScore gives superior results. This, together with several other observations, indicates that the search algorithm of GOLD currently has a problem finding a global minimum when ChemScore is used for docking larger compounds.[65]

The best cognate covalent docking pose was obtained for the cathepsin K X-ray structure 1ATK and the inhibitor E64 using GoldScore (Figure 3A). GoldScore was able to generate top scoring poses that closely resemble the X-ray conformation of
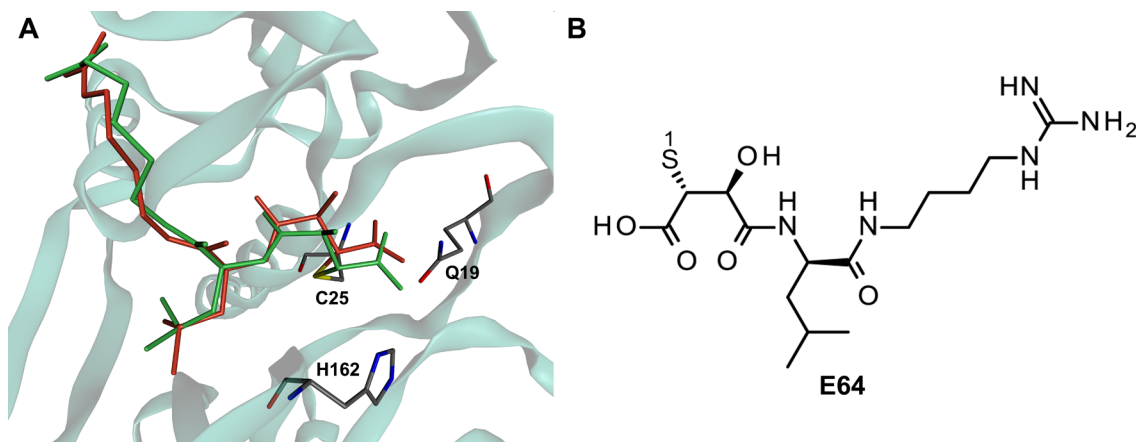
Figure 3. Cognate covalent docking of E64 into the X-ray structure of human cathepsin K (PDB code 1ATK). (A) The X-ray pose of E64 is depicted in capped-stick representation and colored green. The protein backbone of cathepsin K is depicted in ribbon-type rendering. C25 is indicated as sticks colored by atom type. The best cognate covalent docking pose of E64 obtained using GoldScore is shown in capped-stick representation and colored red. (B) Chemical structure of E64 prepared for covalent docking in the bound state. The epoxide ring was opened, and the attached sulfur dummy received the atom number 1. As part of the covalent docking process, the sulfur dummy of the ligand is fitted onto the active site sulfur atom of C25.

Database Generation | **Virtual Screening**
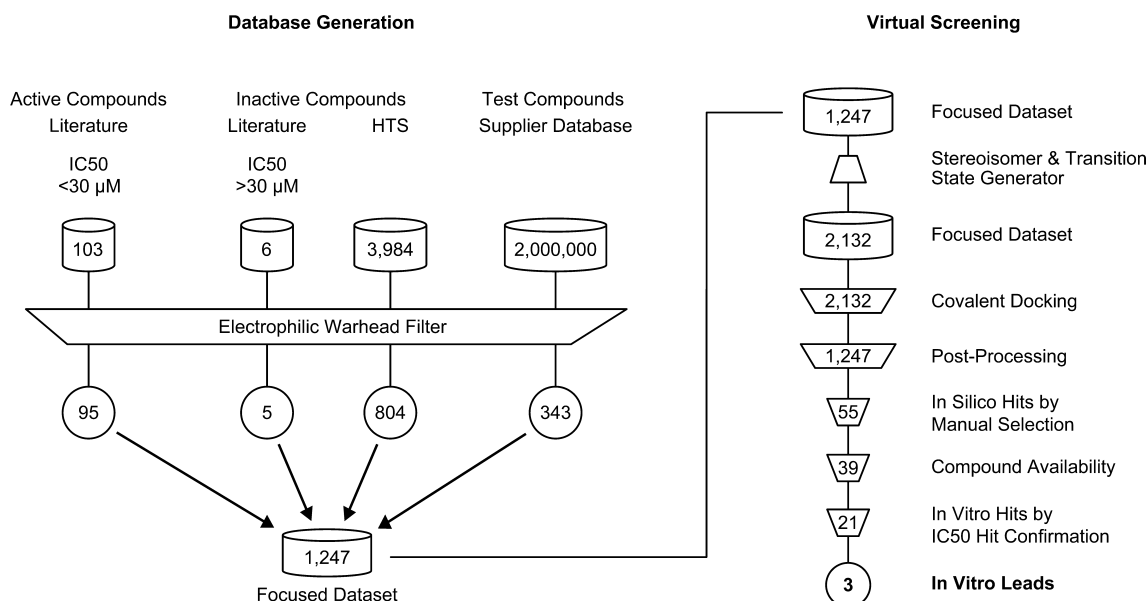
Figure 4. Filtering steps performed during database generation and virtual screening. The focused data set was built from (i) a set of known cathepsin K inhibitors from literature (103 with K_i or IC$_{50}$ of <30 μM), (ii) inactive compounds manually selected from literature (6 with K_i or IC$_{50}$ of ≥50 μM), (iii) 3984 inactive compounds identified in a cathepsin K HTS campaign, and (iiii) a set of 2 million screening compounds from different suppliers. Application of the electrophilic warhead filter resulted in 1247 structures for the focused data set. After generation of stereoisomers and transition states 2132 structures were covalently docked. During the postprocessing only the best scored pose and stereoisomer of each compound were selected for the final overall ranking. After manual selection of 55 in silico hits, 39 hits could be purchased for IC$_{50(t)}$ determination leading to 21 in vitro hits. Finally 3 new cathepsin K leads with nanomolar activity could be selected.

E64 with a mean rmsd of 0.9 Å. ChemScore reproduced the X-ray conformation of E64 with a mean rmsd value of 1.3 Å. This successful validation provides sufficient confidence in drawing meaningful conclusions from the covalent docking studies. Thus, 1ATK was selected as the X-ray structure for the following DBVS campaign.

Docking-Based Virtual Screening. As an initial step, a focused data set of 1247 compounds was prepared as depicted in Figure 4. This data set contained compounds from different suppliers, known actives and a set of known inactives as decoys. Notable effort was invested in the development of an automated procedure that allows the preselection of appropriate electrophilic warheads. Therefore, warhead classes with reasonable reactivity like nitriles, thiosemicarbazones, vinylsulfones, and cyclic ketones that allow the formation of a thiolate intermediate with C25 were selected in a first step. In a second step, a warhead filter applied further conditions that have to be fulfilled by all virtual screening compounds: (i) only compounds with one warhead are allowed; (ii) exclude compounds with a warhead that will lead to a leaving group upon reaction with C25; (iii) only allow compounds with a warhead that will react with C25 via a known or desired mechanism; (iiii) exclude compounds with killing fragments like undesired reactive functional groups. By application of these steps, a database with 2 million structurally diverse screening compounds from different suppliers was reduced to a data set of 343 test compounds. This data set was augmented with 95 known actives (K_i or IC$_{50}$ of <30 μM on human cathepsin K) selected from literature and the PDB, which also passed the warhead filter.[66–71] A data set of 3984 compounds proven to be inactive in an in-house cathepsin K HTS campaign was reduced to a subset of 804 structures by application of the warhead filter. This subset and an additional 5 inactive compounds from the literature (K_i or IC$_{50}$ of ≥50

μM) were added to the data set as decoys. Subsequently, the focused data set was used in a virtual screening workflow as shown in Figure 4. After application of the transition state generator, the data set increased to 2132 chemical entities containing all possible transition state stereoisomers (see Experimental Section for details). For each ligand and each of the two scoring functions, GoldScore and ChemScore, 10 docking poses were generated. Only the best scoring pose and the best scored stereoisomer were retained for each compound for the overall ranking, and separate ranking lists were built for both scoring functions. Each ranking list was analyzed by inspecting the occurrence of known actives and known inactives within the top scoring compound fraction to evaluate the capability of the scoring functions to separate active ligands from nonactive decoys.

The overall virtual screening performance of GOLD's covalent docking method was evaluated by application of cumulative recall curves (Figure 5).

Cumulative recall curves are a conventional way to visualize the performance of a DBVS approach if active compounds are known beforehand.[72] The curve shows the fraction of active compounds found as a function of the database fraction screened when the database compounds are ordered by decreasing score. At one extreme, if a scoring function discriminates perfectly so that all actives get higher scores than any nonactive compound, then one would find all of the actives at the front of the list. The slope of the curve would be 1 until all actives are tested; after that the slope would be 0 ("ideal recall"). At the other extreme, if the scoring function discriminates not at all, the curve would be a straight line with the slope equal to the ratio of the number of actives to the total number of compounds in the database ("random recall").

The cumulative recall curves shown in Figure 5 reveal that the two scoring functions show significant better enrichments
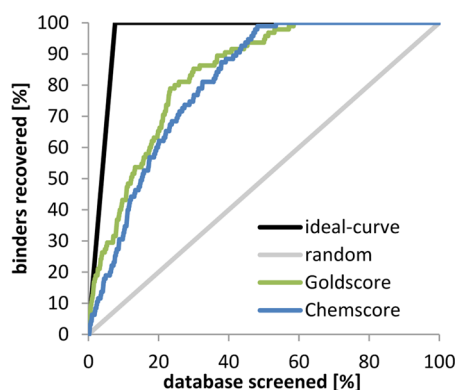
Figure 5. Cumulative recall curves representing the GoldScore- and ChemScore-driven enrichment of virtual screening results. Percentage of known binders recovered versus percentage of the ranked database screened for GoldScore (green line) and ChemScore (blue line). The enrichment level for "ideal recall" is shown as black line. The enrichment level expected for "random recall" is shown as light-gray diagonal line.

compared to the "random recall", proving their ability to discriminate between actives and nonactives. An important measure of the DBVS performance is the prediction of active ligands at the highest ranks of the database. In order to capture this "early" performance, the enrichment factors (EF) at 8% of the ligand/decoy database were calculated.[73] For the top 8% fraction of the GoldScore list, an enrichment factor of $EF_{8\%} = 4$ could be obtained that corresponds to a 4-fold enrichment compared with random screening. For ChemScore an enrichment factor of $EF_{8\%} = 3$ was calculated. Because GoldScore performed slightly better than ChemScore, the top scoring poses of the GoldScore ranking list were taken into account for the manual selection of the virtual hits. A more detailed inspection of the ranking list taking into account the warhead class of each hit revealed that some structure classes (e.g., thiosemicarbazones) show overall higher scoring values than other classes (e.g., ketones). Moreover, the enrichment performance between each warhead class was varying in a certain range (data not shown). In order not to exclude interesting structure classes from the manual selection step, a simple selection of the top ranked structures was abandoned. Instead, the manual selection was performed independently for each warhead class, taking into account high GoldScore ranking scores and activity hot spots containing known cathepsin K inhibitors of the corresponding class. All preselected docking poses were visually inspected for an overall reasonable active site fit, leading to a final data set of 55 in silico hits with unknown in vitro activity on cathepsin K. Only 39 out of the 55 manually selected candidates could be ordered in sufficient amounts from the suppliers. In addition, five low scoring compounds from the nitrile and ketone warhead classes were randomly selected from the lower part of the ranking lists as negative controls for which no or only weak activity was expected. Finally, a data set of 44 compounds including the five low scoring compounds were selected for experimental evaluation (Table 2). More detailed information for each compound is available in the Supporting Information.

In Vitro Lead Confirmation. The 44 candidate compounds were evaluated by an in vitro lead confirmation procedure, which included IC_{50} determination, kinetic experiments, and compound quality controls. All entities that passed the quality control criteria and showed nanomolar activity were

defined as confirmed in vitro leads. Hence, all compounds were tested in a fluorogenic microplate assay on human cathepsin K using the substrate Cbz-Phe-Arg-AMC (see Experimental Section for details). Four cysteine protease inhibitors, namely, leupeptin,[74] E64, K11002,[30] and K11777,[75] were used in the assay as controls.

In general, the IC_{50} value of a reversible inhibitor does not change significantly with incubation time if the inhibitor strictly follows simple Michaelis−Menten kinetics. In contrast, the IC_{50} value determined in an end-point assay of an irreversible, mechanism-based, or slow-binding covalent reversible inhibitor is time-dependent and decreases with incubation time.[76] An exact examination of these inhibitor types involve the determination of parameters like k_{inact} (maximal inactivation rate constant) or k_{obs} (observed inactivation rate), which are related to the time-dependent decrease of IC_{50} values. However, a disadvantage is that these studies can be quite laborious, especially if several binders have to be characterized. Thus, the IC_{50} at 30 min ($IC_{50(t)}$) was determined after 30 min of preincubation time.[77] The $IC_{50(t)}$ represents a compromise between time-dependent decrease of IC_{50} values, cathepsin K stability, and experimental complexity.

The hit limit of the cathepsin K assay was chosen to be IC_{50} < 30 μM. The $[S]/K_{m}$ ratio for the assay was approximately 0.7. Significant inhibitory activity was detected for 21 of these compounds with $IC_{50(t)}$ values ranging from 0.10 to 21.4 μM (Table 2). In summary, the activity of 21 out of the selected 39 in silico hits could be experimentally confirmed and 4 out of the 5 structures predicted as inactive showed no activity on cathepsin K.

In order to confirm the actual molecular structure of the compound in the test tube with the molecular structure stored in the database, all 21 in vitro hits were analyzed by NMR spectroscopy and LCMS experiments. Subsequently, the kinetic aqueous solubility of these hits was determined using nephelometry. All in vitro hits passed the solubility criteria (solubility of up to 30 μM).

To address the question of whether the tested cathepsin K inhibitors are irreversible or reversible binders, rapid dilution experiments[79] in comparison with the potent irreversible cysteine protease inhibitor E64 were performed. These experiments represent an important prerequisite for the correct determination of K_{i} values. With this approach, 333-fold normal reaction amount of cathepsin K was preincubated for 30 min with compound at a concentration that was 333-fold greater than its IC_{50}. After dilution of the enzyme/inhibitor mixture to 333-fold into substrate containing assay buffer at the end of the incubation, the enzyme activity was determined continuously. Generally, a reversible inhibitor would dissociate quickly allowing recovery of enzyme activity, whereas an irreversible inhibitor, in contrast, should prevent recovery of enzyme activity. For an equal final inhibitor concentration, the same steady-state velocity was observed when cathepsin K and the 21 active compounds were preincubated at 333-fold higher concentration and then diluted to 333-fold into substrate containing assay buffer (data not shown). No recovery of cathepsin K activity was observed when enzyme inactivated with E64 was diluted into substrate containing buffer, indicating an irreversible inhibition. Thus, it can be suggested that the 21 active compounds are covalent reversible cathepsin K inhibitors.

The three most potent cathepsin K inhibitors CP243522, CP133433, and CP196392 showed nanomolar activities and

Table 2. Scoring and $IC_{50(t)}$ Values for Compounds Tested in the in Vitro Study[a]

compd		scoring			prediction[c]	
ID	warhead class	GoldScore	ChemScore	in vitro inhibiton, $IC_{50(t)}$ $[\mu M]$[b]	active	inactive
CP79487	nitrile	23	17	>30[c]	X	
CP84191	nitrile	27	14	19.6	X	
CP97899	nitrile	21	9	>30	X	
CP99499	nitrile	2	0	>30		X
CP133433	nitrile	28	9	0.2 ± 0.08[d]	X	
CP148830	vinylsulfone	40	19	19.3	X	
CP150531	nitrile	23	14	>30	X	
CP160888	thiosemicarbazone	34	15	3.4	X	
CP173578	ketone	16	15	21.3	X	
CP196392	thiosemicarbazone	32	14	0.1 ± 0.06[d]	X	
CP201167	vinylsulfone	33	16	>30	X	
CP228062	vinylsulfone	47	19	>30	X	
CP232977	nitrile	3	1	>30		X
CP243522	nitrile	16	8	0.4 ± 0.25[d]	X	
CP247445	nitrile	22	8	>30	X	
CP247447	thiosemicarbazone	37	18	>30	X	
CP247448	thiosemicarbazone	43	18	4.6	X	
CP247449	thiosemicarbazone	31	18	>30	X	
CP247450	vinylsulfone	22	9	>30	X	
CP247451	nitrile	10	1	>30		X
CP247453	nitrile	20	−4	4.3	X	
CP247454	thiosemicarbazone	32	18	>30	X	
CP247455	ketone	−4	−8	>30		X
CP247456	thiosemicarbazone	16	6	5.8	X	
CP247457	thiosemicarbazone	38	16	>30	X	
CP247458	thiosemicarbazone	36	18	>30	X	
CP247459	nitrile	24	9	>30	X	
CP247460	ketone	17	20	3.4	X	
CP247461	thiosemicarbazone	30	21	2.7	X	
CP247462	thiosemicarbazone	25	8	>30	X	
CP247464	thiosemicarbazone	14	7	7.6	X	
CP247465	thiosemicarbazone	31	18	10.5	X	
CP247466	ketone	28	12	>30	X	
CP247468	ketone	23	12	13.3	X	
CP247469	thiosemicarbazone	34	15	1.6	X	
CP247470	nitrile	12	9	>30	X	
CP247471	thiosemicarbazone	36	17	5.3	X	
CP247472	ketone	21	12	>30	X	
CP247473	thiosemicarbazone	36	17	9.6	X	
CP247475	nitrile	2	1	1.3		X
CP247477	ketone	23	12	>30	X	
CP247551	vinylsulfone	29	21	18.1	X	
CP247552	vinylsulfone	49	20	>30	X	
CP247553	vinylsulfone	38	21	21.4	X	

[a]The data set contains 39 compounds predicted as active and 5 structures predicted as inactive. [b]The values are the mean from three independent measurements after 30 min of incubation time. A compound is defined to be inactive with $IC_{50(t)}$ > 30 μM if this was determined for at least three measurements. [c]Based on GoldScore and/or ChemScore scoring values. [d]The ±SD is given for the three most active compounds.

therefore were selected as confirmed leads. Structures, warhead class, and K_i values are presented in Table 3. Original docking ranks prior to visual inspection and candidate selection and quality control-related experimental data are provided in the Supporting Information.

The most potent compound was the nitrile CP243522 with a K_i value of 21 nM toward human cathepsin K. From the thiosemicarbazone class, CP196392 was the most potent compound displaying a K_i value of 134 nM. In the ^{1}H NMR spectrum of CP196392 two molecular species A and B (ratio 1:1.48) were detected (Figure 6). Similar observations for

thiosemicarbazones were already discussed in literature.[80] It is likely that one species is more active than the other and can be attributed to E/Z-isomerism. On the basis of the results from the covalent docking, the E-isomer represented the higher scored pose compared to the Z-isomer and thus may be the more potent inhibitor.

Compounds CP243522 and CP133433 contain the 1,3,5-triazine-2-carbonitrile scaffold as shown in Table 3. Triazine-carbonitriles are already known inhibitors of various cysteine proteases and have been described as potential drugs for the treatment of different diseases, including HIV, cancer, arthritis,

Table 3. Structures and K_i Values[a] of the Three Identified in Vitro Leads

Compound ID	Chemical structure	Warhead class	K_i (nM)
CP243522		nitrile	21 ± 2
CP133433		nitrile	79 ± 3
CP196392		thiosemicarbazone	134 ± 20

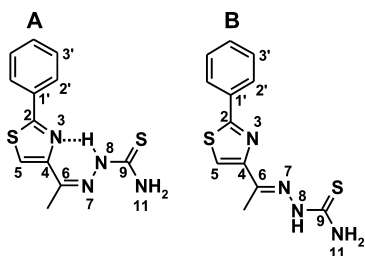[a]K_i values are shown as the mean $\pm$ SD (nM) from three independent experiments.

Figure 6. *E,Z* isomers of the thiosemicarbazone CP196392 characterized by NMR experiments. Species A (*Z*-isomer) forms a hydrogen bond between N^3 and $H(N^8)$ depicted as dotted line. Species B represents the *E*-isomer. NMR data of the isomers are provided in the Supporting Information.

atherosclerosis, osteoporosis, Chagas disease, and African sleeping sickness.[81] In addition, Bailey et al.[82] and Rankovic et al.[43] reported 1,3,5-triazine-2-carbonitriles as covalent reversible inhibitors of human cathepsin K. To the best of our knowledge, compounds CP243522 and CP133433 hitherto have not been referred as cathepsin K inhibitors. The thiosemicarbazone CP196392 contains the 2-phenyl-1,3-thiazol-4-ylethylidene scaffold. While thiosemicarbazones, in general, are known inhibitors of cysteine proteases, analogues of this specific scaffold have not been reported so far as cathepsin K inhibitors in patents or in literature. The K_i values of the confirmed lead structures are in the nanomolar range, which is a promising starting point for further chemical exploration and optimization.

■ CONCLUSIONS

In this study, it could be shown that the presented DBVS approach for the retrieval of covalently binding inhibitors on human cathepsin K, in general, is able to identify active compounds from a pool of decoys. As predicting experimental binding free energies with high accuracy is still out of reach, it is of crucial importance to select a scoring function that is capable of discriminating active from inactive compounds. Therefore, a special compound data set to evaluate the DBVS workflow using GOLD in covalent mode with the modified scoring functions GoldScore and ChemScore was prepared. In addition, an automatic ligand preparation tool for covalent docking, which is particularly useful for the setup of virtual screening data sets, where the amount of molecules is too high for manual preparation of the ligands, was established. Application of this tool allowed the automated selection of covalently binding compounds from different supplier libraries. This selection was augmented with compounds known to be active and with compounds proven to be inactive on cathepsin K as marker for the GoldScore and ChemScore ranking lists. By use of GoldScore, a 4-fold enrichment of active binders was achieved within the top 8% of the screened data set. A total of 21 compounds were confirmed by in vitro testing to be inhibitors of human cathepsin K. The three most active compounds passed the in vitro lead criteria. These novel inhibitors were from the 1,3,5-triazine-2-carbonitrile and thiosemicarbazone warhead class and showed K_i values in the nanomolar range and a covalent reversible mechanism of inhibition.

Because the presented results were obtained using "out of the box" parameters, these conclusions should be taken as more of a rough guide when selecting a covalent docking tool for a specific situation rather than a firm statement of expected performance. Updated versions of GOLD became available during the course of this study, and the results of any improvements are therefore not reflected. However, the enrichment factors achieved in this study together with in vitro activities of the tested compounds clearly point out that the presented DBVS approach is a suitable method to find new covalent leads and to prioritize ligands for synthesis and experimental testing even at the beginning of the lead discovery phase.

■ EXPERIMENTAL SECTION

General Computational Tools. Two-dimensional (2D) molecular structures were either taken from supplier libraries or sketched using ISIS Draw.[83] All structures were converted to three-dimensional coordinates (3D) using the structure generator CORINA. 3D ligand and protein structure files were converted to MOL2 format using the modeling software SYBYL.[84] The Biopolymer module of the SYBYL package was used for protein preparation. Scripts for automated library filtering, ligand preparation, stereoisomer enumeration, and transition state generation were developed in-house using the CACTVS toolkit. 2D and 3D illustrations of molecular structures depicted in this paper were generated with ISIS Draw and MOLCAD,[85] respectively. All calculations were carried out on SGI Fuel workstations (MIPS R12k/ R14k CPUs).

Setup of the Covalent Docking Mode in GOLD. Version 2.2 of the GOLD docking program was used in this study. The GOLD program utilizes a GA to explore the full range of ligand conformational flexibility and the rotational flexibility of selected receptor hydrogens. The mechanism for ligand placement is based on fitting points. The program adds fitting points to hydrogen-bonding groups on the protein and ligand and maps acceptor points in the ligand on donor points in the protein and vice versa. Additionally, GOLD generates hydrophobic fitting points in the protein cavity onto which ligand CH-groups are mapped. The genetic algorithm optimizes flexible ligand dihedrals, ligand ring geometries, dihedrals of protein and ligand OH and NH_3^+ groups, and the mappings of the fitting points. The docking poses are ranked based on molecular-mechanics-like scoring functions. There are two different built-in scoring functions in the used version of GOLD, namely, GoldScore and ChemScore. The ChemScore function is an optimized version of the original ChemScore function developed by Eldridge et al.[86] In covalent mode the program assumes that there is just one atom linking the ligand to the protein. Both protein and ligand files are set up with the link atom included. Inside the GOLD least-squares fitting routine, the link atom in the ligand will be forced to fit onto the link atom in the protein. In order to make sure that the geometry of the bound ligand is correct, the angle-bending potential from the Tripos force field has been incorporated into the fitness function. Thus, an angle-bending energy term for the link atom is included in the calculation of the fitness score. The aforementioned covalent docking mode of GOLD was applied for all docking runs using standard default settings. The scoring functions GoldScore and ChemScore were used in their modified version for covalent docking. Protein and ligand structures were prepared according to the GOLD user manual. The sulfur atom of the cathepsin K amino acid C25 (mature enzyme numbering) was defined as the linking atom for the covalent bond in the protein input file.

Preparation of the Protein Structures for Covalent Docking. All water molecules were removed from the complexes because no significant role in ligand binding could be detected by 3D structural alignment of the considered PDB structures. Ligands were removed, and hydrogens were added to the X-ray structures of cathepsin K. The hydrogen on the sulfur atom of C25 was subsequently removed, and thus, it had a free valence as required by GOLD. Side chain amine

groups were protonated. Carboxylate groups were negatively charged, and hydroxyl groups were considered to be neutral. The imidazole ring of H162 (mature enzyme numbering) was deprotonated. This represents the state immediately after the reaction of the ligand with the protein.

Preparation of the Focused Data Set. A focused data set containing known actives, known inactives, and commercial screening compounds with unknown activity on cathepsin K was prepared. Only molecular structures providing one electrophilic warhead for which a reasonable reaction mechanism can be assumed were considered in the data set. Starting from a database containing 2 million structural diverse screening compounds, 6164 nonredundant compounds belonging to the chemical classes of cyclic ketones, nitriles, vinyl sulfones, and thiosemicarbazones were extracted using substructure filters. This data set was further processed by application of a script-based electrophilic warhead filter leading to a library with 343 structures. Warheads that react with cathepsin K by an undesired or unknown mechanism were excluded from the data set. Compounds containing "killing fragments" like halogen methyl ketones or carboxylic acid halides were excluded. In total, a set of 95 compounds known to be active and a set of 809 compounds known to be inactive on human cathepsin K were added to the 343 screening compounds, resulting in a data set of 1247 structures. The subset of 95 active compounds was built from literature and PDB entries and consisted of aldehydes, acyclic ketones, epoxylsuccinates, isatines, semicarbazones, and compounds from the aforementioned warhead classes.[66−71] Compounds exhibiting a K_i or IC_{50} of <30 μM were considered to be active on cathepsin K, while compounds exhibiting activity values of >30 μM were considered as inactives. Structural diversity of the compound set selected from literature was gained by considering compounds prepared via individual synthesis pathways. For the decoy subset, 804 compounds (<40% inhibition) were selected by application of the electrophilic warhead filter on a set of 3984 inactive compounds found during an in-house HTS campaign (data not shown). The decoy subset was completed by addition of five compounds from the above-cited literature exhibiting K_i or IC_{50} of ≥50 μM.

Ligand Preparation for Covalent Docking. Before application of the covalent docking procedure, all compounds in the focused data set had to be transformed to the three-dimensional molecular structure resembling the electrophilic covalently bound state. Depending on the electrophilic warhead of the ligands, different reaction mechanisms (Figure 2) had to be taken into account, which also could lead to different stereoisomers. For that purpose, a CACTVS-based script that generated all transition states considering also the corresponding stereoisomers was implemented. In detail, a sulfur dummy was added to the electrophile carbon atom of the respective warhead resembling the conformation of the bound state. If the electrophile carbon represented a prochiral center, all possible stereoisomers of the corresponding compound were generated and treated as distinct ligands for further processing. Finally, the atom IDs of each structure were changed in that way, and the attached sulfur dummy received the atom number 1. This step is crucial to allow covalent docking of multiple ligands in GOLD as required for a DBVS setup. By standardization of all ligands regarding the atom number 1, only one generic GOLD configuration file is required, which forces the docking tool always to use the first atom of each ligand (the added sulfur dummy) to be fit onto the sulfur atom of C25 in the cathepsin K active site. For energy minimization, a methyl group was attached to the sulfur dummy. Hydrogen atoms were added to all unassigned valences, and 3D coordinates were generated using CORINA. Partial charges were calculated using the SPL script "dbcharges.spl" on the basis of the MMFF94 force field. Energy minimization was done using the conjugate gradient method with a maximum of 5000 iteration steps as implemented in the SPL script "mindb.spl". Subsequently, the temporary methyl group on the sulfur dummy was removed resulting in a free valence as required by GOLD. Because stereoisomers were treated as individual compounds, the number of entries in the virtual screening data set increased from 1247 to 2132 structures.

GA Parameter Setup. All dockings were performed using GOLD's standard default settings with highest accuracy. For each independent GA run, a maximum number of 100 000 GA operations were performed on a single population of 100 individuals. Operator weights for crossover, mutation, and migration in the entry box were set to 95, 95, and 10, respectively. Fitness flags were set for allowing the GA to vary the bonds to nitrogens and to vary ligand ring conformations.

Cognate Covalent Ligand Docking. For cognate covalent ligand docking, the native ligands along with chirality information from the corresponding PDB complexes were resketched in 2D using ISIS Draw. Ligands and proteins were prepared for covalent docking as described above, and each ligand was docked back into its cognate binding site (Table 1). Binding sites are defined as all protein atoms within 6 Å of a non-hydrogen ligand atom. Serine and threonine hydroxyl groups and lysine NH_3^+ groups could rotate, flexible ligand torsions could rotate, and ring corners in ligands were allowed to flip to explore ring conformers. For each ligand, 10 docking runs were performed using either GoldScore or ChemScore, which resulted in 10 ranked poses from each run. The calculation of the heavy atom root-mean-squared deviation was carried out using the standard rmsd algorithm of the SYBYL-based script "RMS_Table_mol.spl".

Docking-Based Virtual Screening. The cathepsin K crystal structure (PDB code 1ATK) was selected as the protein target for the DBVS. The protein was prepared as described above. To maintain simplicity in the selection of docking solutions and result analysis, only the top scoring pose for each compound was kept in the final ranking list for GoldScore and ChemScore. In the case of stereoisomers only the higher scored isomer was kept. Hence, the final ranking list for each scoring function comprised 1247 docking poses.

Test Compounds. Compounds selected for in vitro testing were purchased from commercial compound suppliers. The cysteine protease inhibitors K11002 (morpholine-urea-Phe-homoPhe-vinyl-sulfone-phenyl) and K11777 (N-methyl-piperazine-Phe-homoPhe-vinylsulfone-phenyl) were gratefully obtained from J. H. McKerrow, University of San Francisco (UCSF), CA. CP243522 was ordered from Vitas-M Laboratory, Ltd., Moscow, Russia; the IUPAC name is 4-(N-phenylanilino)-6-pyrrolidin-1-yl-1,3,5-triazine-2-carbonitrile (purity of 100%). CP133433 was ordered from ABI Chem, Munich, Germany; the IUPAC name is 4-anilino-6-(propan-2-ylamino)-1,3,5-triazine-2-carbonitrile (purity of 93.7%). CP196392 was ordered from Interchim, Montluçon Cedex, France; the IUPAC name is [1-(2-phenyl-1,3-thiazol-4-yl)ethylideneamino]thiourea (purity of 97.6%).

Compound purity and molecular mass were confirmed by LCMS experiments which were performed on an Agilent 1100 LCMS (Chromtech, Idstein, Germany) device using a gradient acetonitrile–water with 0.1% formic acid on an Agilent Zorbax SB-C18 column (4.6 mm × 30 mm) at 313 K. Molecular structures of the tested compounds were additionally confirmed by NMR experiments. The NMR spectra were recorded on a Bruker Avance 400 (Bruker Biospin, Rheinstetten, Germany) spectrometer equipped with a 5 mm BBO probe head with z-gradient and Great 1/10 gradient amplifier. The temperature was kept constant at 298 K. Spectral processing was performed using the software Topspin 1.2 (Bruker Biospin, Rheinstetten, Germany). Proton (^{1}H), carbon (^{13}C), and nitrogen (^{15}N) chemical shifts were assigned using ^{1}H, H,H-NOESY, H,C-HSQC, H,C-HMBC, and H,N-HMQC experiments. Chemical shifts were determined with respect to TMS as standard. In each case the spectra were consistent with the structure representations provided by the suppliers.

Compound Solubility Measurements. A 2-fold serial dilution of the compounds was performed (0.5–100 μM) in DMSO (Acros Organics, Morris Plains, NJ, U.S.) and added in duplicate to a 96-well microplate. Phosphate buffered saline (PBS, pH 7.4) (Medicago AB, Uppsala, Sweden) was added to give a total volume of 200 μL and a final DMSO concentration of 5% (v/v) in all wells. The plate was incubated at room temperature for 22 min, and the relative solubility of the compounds was determined by measuring forward-scattered light using a NEPHELOstar laser-based microplate nephelometer (BMG LABTECH GmbH, Offenburg, Germany). Wells containing

only buffer and 5% (v/v) DMSO were used as controls. Data analysis was carried out using Excel (Microsoft Corporation, Redmond, WA, U.S.).

In Vitro Inhibition Assays. Human cathepsin K was purchased from CalBiochem (San Diego, CA, U.S.) in recombinant form with His-Tag. Enzymatic activity was measured using a resonance energy transfer fluorogenic assay modified from literature.[43] Progress curves were monitored by following the increase in fluorescence induced by cleavage of the 7-amino-4-methylcoumarin moiety (excitation wavelength of 355 nm and emission wavelength of 460 nm) from the synthetic peptide substrate Cbz-Phe-Arg-AMC (Bachem, Weil am Rhein, Germany). For the inhibition assays, black 384-well microplates were used and read on a SPECTRAFluor Plus (Tecan Inc., Durham, NC, U.S.) plate reader. Assays were performed at 310 K in a 100 mM sodium acetate buffer (pH 5.5) containing 5 mM DTT (Carl Roth GmbH, Karlsruhe, Germany), 5 mM Titriplex II (Merck, Darmstadt, Germany), 0.1 mg/mL BSA (PAA Laboratories GmbH, Pasching, Austria), and 3.5% DMSO using 25 μM Cbz-Phe-Arg-AMC as substrate. Prior to the addition of substrate, different concentrations of the inhibitor ranging from 30 μM to 3 nM were preincubated for 30 min with the enzyme (2 nM) to allow the establishment of the enzyme–inhibitor complex. Substrate was then added and the enzyme activity measured from the increase of OD at 460 nm. The final volume of the reaction was 300 μL. The cysteine protease inhibitors leupeptin (Ac-Leu-Leu-Arg-CHO, Sigma-Aldrich Inc., St. Louis, MO, U.S.), E-64 (trans-epoxysuccinyl-L-leucylamido-(4-guanidino)butane, Sigma-Aldrich Inc., St. Louis, MO, U.D.), K11002, and K11777 were used as controls. The percent inhibition of the reaction was calculated from a control reaction containing only the compound vehicle. $IC_{50(t)}$ values were calculated using the four-parameter equation "205" and the option "unlock" from the XLfit add-in (IDBS, Guildford, U.K.) in Excel (Microsoft Corporation, Redmond, CA, U.S.). K_i values for time-independent inhibition were obtained by plotting $v_0/v_i - 1$ versus [I] using the equation $v_0/v_i = 1 + [I]/K_{i,app}$ and correction to zero substrate concentration from $K_i = K_{i,app}/(1 + [S]/K_M)$.[87] The K_M value determined for cathepsin K to correct $K_{i,app}$ values was 36.6 μM. The $[S]/K_m$ ratio for the assay was approximately 0.7. All values are the mean from at least three independent assays to ensure statistically significant results. The experiment was controlled further for background fluorescence of the substrate, for fluorescence of fully cleaved substrate, and for fluorescence quenching or augmentation from solutions.

Kinetic Analysis. To verify whether the selected compounds are reversible or irreversible inhibitors, rapid dilution experiments were performed. A 333-fold normal amount of enzyme (666 nM) was mixed with compound at a final concentration of 333-fold IC_{50} for each control. After incubation at 310 K for 30 min the mixtures were subsequently diluted to 333-fold in substrate-containing assay buffer and the residual cathepsin K activity was repeatedly sampled for 60 min.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Dose–response curves, analytical data, and docking ranks of the most potent compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +49-6130-948-206. Fax: +49-6130-948-517. E-mail: paul.selzer@msd.de.

Present Addresses

⊥Ottonenstrasse 32, D-55218 Ingelheim, Germany.

#MicroCombiChem e.K., Rheingaustrasse 190-196, Building E512, D-65203 Wiesbaden, Germany.

Author Contributions

‖These authors contributed equally.

Notes

◼ ACKNOWLEDGMENTS

◼ ABBREVIATIONS USED

AMC, 7-amino-4-methylcoumarin; DBVS, docking-based virtual screening; DFMO, α-difluoromethylornithine; GA, genetic algorithm; $IC_{50(t)}$, concentration of inhibitor required to decrease activity by 50% after a preincubation time (t); MMFF94, Merck molecular force field; SPL, SYBYL programming language; TCL, tool command language

◼ REFERENCES

(1) Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644–2676.

(2) Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. *Angew. Chem., Int. Ed.* **2004**, *43*, 6596–6616.

(3) Kalgutkar, A. S.; Gardner, I.; Obach, R. S.; Shaffer, C. L.; Callegari, E.; Henne, K. R.; Mutlib, A. E.; Dalvie, D. K.; Lee, J. S.; Nakai, Y.; O'Donnel, J. P.; Boer, J.; Harriman, S. P. A comprehensive listing of bioactivation pathways of organic functional groups. *Curr. Drug Metab.* **2005**, *6*, 161–225.

(4) Wiliams, D. P. Toxicophores: investigations in drug safety. *Toxicology* **2006**, *226*, 1–11.

(5) Maurer, T. S.; Tabrizi-Fard, M. A.; Fung, H. L. Impact of mechanism-based enzyme inactivation on inhibitor potency: implications for rational drug discovery. *J. Pharm. Sci.* **2000**, *89*, 1404–1414.

(6) Van der Ouderaa, F. J.; Buytenhek, M.; Nugteren, D. H.; Van Dorp, D. A. Acetylation of prostaglandin endoperoxide synthetase with acetylsalicyclic acid. *Eur. J. Biochem.* **1980**, *109*, 1–8.

(7) Shin, J. M.; Cho, Y. M.; Sachs, G. Chemistry of covalent inhibition of the gastric (H⁺, K⁺)-ATPase by proton pump inhibitors. *J. Am. Chem. Soc.* **2004**, *126*, 7800–7811.

(8) Besancon, M.; Simon, A.; Sachs, G.; Shin, J. M. Sites of reaction of the gastric H, K-ATPase with extracytoplasmic thiol reagents. *J. Biol. Chem.* **1997**, *272*, 22438–22446.

(9) Salach, J. I.; Demeterm, K.; Youdim, M. B. H. The reaction of bovine and rat liver monoamine oxidase with [¹⁴C]-clorgyline and [¹⁴C]-deprenyl. *Mol. Pharmacol.* **1979**, *16*, 234–241.

(10) Van Houten, K. A.; Kim, J. M.; Bogdan, M. A.; Ferri, D. C.; Mariano, P. S. A new strategy for the design of monoamine oxidase inactivators. Exploratory studies with tertiary allylic and propargylic amino alcohols. *J. Am. Chem. Soc.* **1998**, *120*, 5864–5872.

(11) Groll, M.; Berkers, C. R.; Ploegh, H. L.; Ovaa, H. Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. *Structure* **2006**, *14*, 451–456.

(12) Prongay, A. J.; Guo, Z.; Yao, N.; Pichardo, J.; Fischmann, T.; Strickland, C.; Myers, J., Jr.; Weber, P. C.; Beyer, B. M.; Ingram, R.; Hong, Z.; Prosise, W. W.; Ramanathan, L.; Taremi, S. S.; Yarosh-Tomaine, T.; Zhang, R.; Senior, M.; Yang, R. S.; Malcolm, B.; Arasappan, A.; Bennett, F.; Bogen, S. L.; Chen, K.; Jao, E.; Liu, Y. T.; Lovey, R. G.; Saksena, A. K.; Venkatraman, S.; Girijavallabhan, V.; Njoroge, F. G.; Madison, V. Discovery of the HCV NS3/4A protease inhibitor (1R,5S)-N-[3-amino-1-(cyclobutylmethyl)-2,3-dioxopropyl]-3-[2(S)-[[[(1,1-dimethylethyl)amino]carbonyl]amino]-3,3-dimethyl-1-oxobutyl]-6,6-dimethyl-3-azabicyclo[3.1.0]hexan-2(S)-carboxamide (Sch 503034) II. Key steps in structure-based optimization. *J. Med. Chem.* **2007**, *50*, 2310–2318.

(13) Matthews, S. J.; Lancaster, J. W. Telaprevir: a hepatitis C NS3/4A protease inhibitor. *Clin. Ther.* **2012**, *34*, 1857–1882.

(14) Byers, T. L.; Bush, T. L.; McCann, P. P.; Bitonti, A. J. Antitrypanosomal effects of polyamine biosynthesis inhibitors correlate with increases in *Trypanosoma brucei* S-adenosyl-L-methionine. *Biochem. J.* **1991**, *274*, 527–533.

(15) Vincent, I. M.; Creek, D.; Watson, D. G.; Kamleh, M. A.; Woods, D. J.; Wong, P. E.; Burchmore, R. J. S.; Barrett, M. P. A molecular mechanism for eflornithine resistance in african trypanosomes. *PLoS Pathog.* **2010**, *6*, 6–9.

(16) Smith, A. J.; Zhang, X.; Leach, A. G.; Houk, K. N. Beyond picomolar affinities: Quantitative aspects of noncovalent and covalent binding of drugs to proteins. *J. Med. Chem.* **2009**, *52*, 225–233.

(17) Copeland, R. A. Irreversible Enzyme Inactivators in Evaluation of Enzyme Inhibitors. *Evaluation of Enzyme Inhibitors in Drug Discovery. A Guide for Medicinal Chemists and Pharmacologists*; John Wiley & Sons: New York, 2005; Chapter 8, pp 214–248.

(18) Tuccinardi, T. Docking-based virtual screening: recent developments. *Comb. Chem. High Throughput Screening* **2009**, *12*, 303–314.

(19) Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **1995**, *245*, 43–53.

(20) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727–748.

(21) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.

(22) Abagyan, R. A.; Totrov, M. M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **1994**, *235*, 983–1002.

(23) Abagyan, R. A.; Totrov, M. M.; Kuznetsov, D. N. ICM—a new method for protein modeling and design. Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **1994**, *15*, 488–506.

(24) McMartin, C.; Bohacek, R. S. QXP: Powerful, rapid computer algorithms for structure-based drug design. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 333–344.

(25) Lawandi, J.; Toumieux, S.; Seyer, V.; Campbell, P.; Thielges, S.; Juillerat-Jeanneret, L.; Moitessier, N. Constrained peptidomimetics reveal detailed geometric requirements of covalent prolyl oligopeptidase inhibitors. *J. Med. Chem.* **2009**, *52*, 6672–6684.

(26) Zhang, S.; Shi, Y.; Jin, H.; Liu, Z.; Zhang, L. Covalent complexes of proteasome model with peptide aldehyde inhibitors MG132 and MG101: docking and molecular dynamics study. *J. Mol. Model.* **2009**, *15*, 1481–1490.

(27) Schaeffer, M.; Schroeder, J.; Heckeroth, A. R.; Noack, S.; Gassel, M.; Mottram, J. C.; Selzer, P. M.; Coombs, G. H. Identification of lead compounds targeting the cathepsin B-like enzyme of *Eimeria tenella*. *Antimicrob. Agents Chemother.* **2012**, *56*, 1190–1201.

(28) Turk, B.; Turk, D.; Turk, V. Lysosomal cysteine proteases: more than scavengers. *Biochim. Biophys. Acta* **2000**, *1477*, 98–111.

(29) Selzer, P. M.; Chen, X.; Chan, V. J.; Cheng, M.; Kenyon, G. L.; Kuntz, I. D.; Sakanari, J. A.; Cohen, F. E.; McKerrow, J. H. Leishmania major: molecular modeling of cysteine proteases and prediction of new nonpeptide inhibitors. *Exp. Parasitol.* **1997**, *87*, 212–221.

(30) Selzer, P. M.; Pingel, S.; Hsieh, I.; Ugele, B.; Chan, V. J.; Engel, J. C.; Bogyo, M.; Russell, D. G.; Sakanari, J. A.; McKerrow, J. H. Cysteine protease inhibitors as chemotherapy: lessons from a parasite target. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11015–11022.

(31) Mottram, J. C.; Coombs, G. H. Leishmania, C. P. A, CPB and CPC Cysteine Proteases. In *Handbook of Proteolytic Enzymes*, 2nd ed.;

Barrett, A. J., Rawlings, N. D., Woessner, J. F., Eds; Elsevier: London, 2004; pp 1166−1169.

(32) Brömme, D.; Okamoto, K.; Wang, B. B.; Biroc, S. Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts: functional expression of human cathepsin O2 in spodoptera frugiperda and characterization of the enzyme. *J. Biol. Chem.* **1996**, *271*, 2126−2132.

(33) Zeng, G. Z.; Pan, X. L.; Tan, N. H.; Xiong, J.; Zhang, Y. M. Natural biflavones as novel inhibitors of cathepsins B and K. *Eur. J. Med. Chem.* **2006**, *41*, 1247−1252.

(34) Pan, X.; Tan, N.; Zeng, G.; Han, H.; Huang, H. 3D-QSAR and docking studies of aldehyde inhibitors of human cathepsin K. *Bioorg. Med. Chem.* **2006**, *14*, 2771−2778.

(35) Drake, F. H.; Dodds, R. A.; James, I. E.; Connor, J. R.; Debouck, C.; Richardson, S.; Lee-Rykaczewski, E.; Coleman, L.; Rieman, D.; Bathlow, R.; Hastings, G.; Gowen, M. Cathepsin K, but not cathepsins B, L, or, S is abundantly expressed in human osteoclasts. *J. Biol. Chem.* **1996**, *271*, 12511−12516.

(36) Votta, B. J.; Levy, M. A.; Badger, A.; Bradbeer, J.; Dodds, R. A.; James, I. E.; Thompson, S.; Bossard, M. J.; Carr, T.; Conner, J. R.; Tomaszek, T. A.; Szewczuk, L.; Drake, F. H.; Veber, D. F.; Gowen, M. J. Peptide aldehyde inhibitors of cathepsin K inhibit bone resorption both in vitro and in vivo. *J. Bone Miner. Res.* **1997**, *12*, 1396−1406.

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

(38) McGrath, M. E.; Sprengeler, P. A.; Hill, C. M.; Martichonok, V.; Cheung, H.; Somoza, J. R.; Palmer, J. T.; Janc, J. W. Peptide ketobenzoxazole inhibitors bound to cathepsin K. *Biochemistry* **2003**, *42*, 15018−15028.

(39) Mendonca, R. V.; Venkatraman, S.; Palmer, J. T. Novel route to the synthesis of peptides containing 2-amino-1′-hydroxymethyl ketones and their application as cathepsin K inhibitors. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2887−2891.

(40) 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.; Veber, D. F. Conformationally constrained 1,3-diamino ketones: a series of potent inhibitors of the cysteine protease cathepsin. *J. Med. Chem.* **1998**, *41*, 3563−3567.

(41) Marquis, R. W.; Ru, Y.; Yamashita, D. S.; Oh, H. J.; Yen, J.; Thompson, S. K.; Carr, T. J.; Levy, M. A.; Tomaszek, T. A.; Ijames, C. F.; Smith, W. W.; Zhao, B.; Janson, C. A.; Abdel-Meguid, S. S.; D'Alessio, K. J.; McQueney, M. S.; Veber, D. F. Potent dipeptidylketone inhibitors of the cysteine protease cathepsin K. *Bioorg. Med. Chem.* **1999**, *7*, 581−588.

(42) Palmer, J. T.; Bryant, C.; Wang, D. X.; Davis, D. E.; Setti, E. L.; Rydzewski, R. M.; Venkatraman, S.; Tian, Z. Q.; Burrill, L. C.; Mendonca, R. V.; Springman, E.; McCarter, J.; Chung, T.; Cheung, H.; Janc, J. W.; McGrath, M.; Somoza, J. R.; Enriquez, P.; Yu, Z. W.; Strickley, R. M.; Liu, L.; Venuti, M. C.; Percival, M. D.; Falgueyret, J. P.; Prasit, P.; Oballa, R.; Riendeau, D.; Young, R. N.; Wesolowski, G.; Rodan, S. B.; Johnson, C.; Kimmel, D. B.; Rodan, G. Design and synthesis of tri-ring P3 benzamide-containing aminonitriles as potent, selective, orally effective inhibitors of cathepsin K. *J. Med. Chem.* **2005**, *48*, 7520−7534.

(43) Rankovic, Z.; Cai, J.; Kerr, J.; Fradera, X.; Robinson, J.; Mistry, A.; Hamilton, E.; McGarry, G.; Andrews, F.; Caulfield, W.; Cumming, I.; Dempster, M.; Waller, J.; Scullion, P.; Martin, I.; Mitchell, A.; Long, C.; Baugh, M.; Westwood, P.; Kinghorn, E.; Bruin, J.; Hamilton, W.; Uitdehaag, J.; Van Zeeland, M.; Potin, D.; Saniere, L.; Fouquet, A.; Chevallier, F.; Deronzier, H.; Dorleans, C.; Nicolai, E. Design and optimization of a series of novel 2-cyano-pyrimidines as cathepsin K inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1524−1527.

(44) Gauthier, J. Y.; Chauret, N.; Cromlish, W.; Desmarais, S.; Duong, L. E. T.; Falgueyret, J. P.; Kimmel, D. B.; Lamontagne, S.; Leger, S.; Leriche, T.; Li, C. S.; Masse, F.; McKay, D. J.; Nicoll-Griffith, D. A.; Oballa, R. M.; Palmer, J. T.; Percival, M. D.; Riendeau, D.; Robichaud, J.; Rodan, G. A.; Rodan, S. B.; Seto, C.; Therien, M.;

Truong, V. L.; Venuti, M. C.; Wesolowski, G.; Young, R. N.; Zamboni, R.; Black, W. C. The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 923−928.

(45) Podgorski, I. Future of anticathepsin K drugs: dual therapy for skeletal disease and atherosclerosis? *Future Med. Chem.* **2009**, *1*, 21−34.

(46) Peroni, A.; Zini, A.; Braga, V.; Colato, C.; Adami, S.; Girolomoni, G. Drug-induced morphea: report of a case induced by balicatib and review of the literature. *J. Am. Acad. Dermatol.* **2008**, *59*, 125−129.

(47) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discovery* **2011**, *10*, 307−317.

(48) Ravikumar, M.; Pavan, S.; Bairy, S.; Pramod, A. B.; Sumakanth, M.; Kishore, M.; Sumithra, T. Virtual screening of cathepsin K inhibitors using docking and pharmacophore models. *Chem. Biol. Drug Des.* **2008**, *72*, 79−90.

(49) Stumpfe, D.; Sisay, M. T.; Fritzler, M.; Vogt, I.; Gütschow, M.; Bajorath, J. Inhibitors of cathepsins K and S identified using the DynaMAD virtual screening algorithm. *Chem. Med. Chem.* **2010**, *5*, 61−64.

(50) Du, X.; Guo, C.; Hansell, E.; Doyle, P. S.; Caffrey, C. R.; Holler, T. P.; McKerrow, J. H.; Cohen, F. E. Synthesis and structure−activity relationship study of potent trypanocidal thiosemicarbazone inhibitors of the trypanosomal cysteine protease cruzain. *J. Med. Chem.* **2002**, *45*, 2695−2707.

(51) 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.; Abdel-Meguid, S. S. Crystal structure of human osteoclast cathepsin K complex with E-64. *Nat. Struct. Biol.* **1997**, *4*, 109−111.

(52) Li, W. A.; Barry, Z. T.; Cohen, J. D.; Wilder, C. L.; Deeds, R. J.; Keegan, P. M.; Platt, M. O. Detection of femtomole quantities of mature cathepsin K with zymography. *Anal. Biochem.* **2010**, *401*, 91−98.

(53) Yamashita, D. S.; Smith, W. W.; Zhao, B.; 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.; D'Alessio, K. J.; Amegadzie, B. Y.; Hanning, C. R.; Abdel-Meguid, S. S.; DesJarlais, R. L.; Gleason, J. G.; Veber, D. F. Structure and design of potent and selective cathepsin K inhibitors. *J. Am. Chem. Soc.* **1997**, *119*, 11351−11352.

(54) 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.; Rykaczewski, L. L.; James, I. E.; Lark, M. W.; Drake, F. H.; Gowen, M.; Gleason, J. G.; Veber, D. F. Design of potent and selective human cathepsin K inhibitors that span the active site. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14249−14254.

(55) DesJarlais, R. L.; Yamashita, D. S.; Oh, H.; Uzinskas, I. N.; Erhard, K. F.; Allen, A. C.; Haltiwanger, R. C.; Zhao, B.; Smith, W. W.; Abdel-Meguid, S. S.; D'Alessio, K.; Janson, C. A.; McQueney, M. S.; Tomasz, T. A. Use of X-ray co-crystal structures and molecular modeling to design potent and selective non-peptide inhibitors of cathepsin K. *J. Am. Chem. Soc.* **1998**, *120*, 9114−9115.

(56) McGrath, M. E.; Klaus, J. L.; Barnes, M. G.; Bromme, D. Crystal structure of human cathepsin K complexed with a potent inhibitor. *Nat. Struct. Biol.* **1997**, *4*, 105−109.

(57) Marquis, R. W.; Ru, Y.; LoCastro, S. M.; Zeng, J.; Yamashita, D. S.; Oh, H. J.; Erhard, K. F.; Davis, L. D.; Tomaszek, T. A.; Tew, D.; Salyers, K.; Proksch, J.; Ward, K.; Smith, B.; Levy, M.; Cummings, M. D.; Haltiwanger, R. C.; Trescher, G.; Wang, B.; Hemling, M. E.; Quinn, C. J.; Cheng, H. Y.; Lin, F.; Smith, W. W.; Janson, C. A.; Zhao, B.; McQueney, M. S.; D'Alessio, K.; Lee, C. P.; Marzulli, A.; Dodds, R. A.; Blake, S.; Hwang, S. M.; James, I. E.; Gress, C. J.; Bradley, B. R.; Lark, M. W.; Gowen, M.; Veber, D. F. Azepanone-based inhibitors of human and rat cathepsin K. *J. Med. Chem.* **2001**, *44*, 1380−1395.

(58) Catalano, J. G.; Deaton, D. N.; Furfine, E. S.; Hassell, A. M.; McFadyen, R. B.; Miller, A. B.; Miller, L. R.; Shewchuk, L. M.; Willard, D. H.; Wright, L. L. Exploration of the P1 SAR of aldehyde cathepsin K inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 275−278.

(59) Boros, E. E.; Deaton, D. N.; Hassell, A. M.; McFadyen, R. B.; Miller, A. B.; Miller, L. R.; Paulick, M. G.; Shewchuk, L. M.; Thompson, J. B.; Willard, D. H.; Wright, L. L. Exploration of the P2-P3 SAR of aldehyde cathepsin K inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3425−3429.

(60) Barrett, D. G.; Catalano, J. G.; Deaton, D. N.; Hassell, A. M.; Long, S. T.; Miller, A. B.; Miller, L. R.; Shewchuk, L. M.; Wells-Knecht, K. J.; Willard, D. H.; Wright, L. L. Potent and selective P2-P3 ketoamide inhibitors of cathepsin K with good pharmacokinetic properties via favorable P1′, P1, and/or P3 substitutions. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4897−4902.

(61) Sadowski, J.; Rudolph, C.; Gasteiger, J. The generation of 3D-models of host−guest. *Anal. Chim. Acta* **1992**, *265*, 233−241.

(62) Ihlenfeldt, W. D.; Takahashi, Y.; Abe, H.; Sasaki, S. Computation and management of chemical properties in CACTVS: an extensible networked approach toward modularity and compatibility. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 109−116.

(63) Bissantz, C.; Folkers, G.; Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* **2000**, *43*, 4759−4767.

(64) Hartshorn, M. J.; Verdonk, M. L.; Chessari, G.; Brewerton, S. C.; Mooij, W. T.; Mortenson, P. N.; Murray, C. W. Diverse, high-quality test set for the validation of protein−ligand docking performance. *J. Med. Chem.* **2007**, *50*, 726−741.

(65) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved protein−ligand docking using GOLD. *Proteins* **2003**, *52*, 609−623.

(66) Falgueyret, J. P.; Oballa, R. M.; Okamoto, O.; Wesolowski, G.; Aubin, Y.; Rydzewski, R. M.; Prasit, P.; Riendeau, D.; Rodan, S. B.; Percival, M. D. Novel, nonpeptidic cyanamides as potent and reversible inhibitors of human cathepsins K and L. *J. Med. Chem.* **2001**, *44*, 94−104.

(67) Smith, R. A.; Bhargava, A.; Browe, C.; Chen, J.; Dumas, J.; Hatoum-Mokdad, H.; Romero, R. Discovery and parallel synthesis of a new class of cathepsin K inhibitors. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2951−2954.

(68) Robichaud, J.; Oballa, R.; Prasit, P.; Falgueyret, J. P.; Percival, M. D.; Wesolowski, G.; Rodan, S. B.; Kimmel, D.; Johnson, C.; Bryant, C.; Venkatraman, S.; Setti, E.; Mendonca, R.; Palmer, J. T. A novel class of nonpeptidic biaryl inhibitors of human cathepsin K. *J. Med. Chem.* **2003**, *46*, 3709−3727.

(69) Robichaud, J.; Bayly, C.; Oballa, R.; Prasit, P.; Mellon, C.; Falgueyret, J. P.; Percival, M. D.; Wesolowski, G.; Rodan, S. B. Rational design of potent and selective NH-linked aryl/heteroaryl cathepsin K inhibitors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4291−4295.

(70) Katunuma, N.; Matsui, A.; Inubushi, T.; Murata, E.; Kakegawa, H.; Ohba, Y.; Turk, D.; Turk, V.; Tada, Y.; Asao, T. Structure-based development of pyridoxal propionate derivatives as specific inhibitors of cathepsin K in vitro and in vivo. *Biochem. Biophys. Res. Commun.* **2000**, *267*, 850−854.

(71) Duffy, K. J.; Ridgers, L. H.; DesJarlais, R. L.; Tomaszek, T. A.; Bossard, M. J.; Thompson, S. K.; Keenan, R. M.; Veber, D. F. Design and synthesis of diaminopyrrolidinone inhibitors of human osteoclast cathepsin K. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1907−1910.

(72) Maiorov, V.; Sheridan, R. P. Enhanced virtual screening by combined use of two docking methods: getting the most on a limited budget. *J. Chem. Inf. Model.* **2005**, *45*, 1017−1023.

(73) McInnes, C. Virtual screening strategies in drug discovery. *Curr. Opin. Chem. Biol.* **2007**, *11*, 494−502.

(74) Hozumi, M.; Ogawa, M.; Sugimura, T.; Takeuchi, T.; Umezawa, H. Inhibition of tumorigenesis in mouse skin by leupeptin, a protease inhibitor from actinomycetes. *Cancer Res.* **1972**, *32*, 1725−1728.

(75) Kerr, I. D.; Lee, J. H.; Farady, C. J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K. C.; Caffrey, C. R.; Legac, J.; Hansell, E.; McKerrow, J. H.; Craik, C. S.; Rosenthal, P. J.; Brinen, L. S. Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *J. Biol. Chem.* **2009**, *284*, 25697−25703.

(76) Krippendorff, B. F.; Neuhaus, R.; Lienau, P.; Reichel, A.; Huisinga, W. Mechanism-based inhibition: deriving Ki and kinact directly from time-dependent IC50 values. *J. Biomol. Screening* **2009**, *14*, 913−923.

(77) Chang, S. Y.; Chen, C.; Yang, Z.; Rodrigues, D. Further assessment of 17α-ethinyl estradiol as an inhibitor of different human cytochrome P450 forms in vitro. *Drug Metab. Dispos.* **2009**, *37*, 1667−1675.

(78) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. *J. Med. Chem.* **2004**, *47*, 45−55.

(79) Chatterjee, S.; Ator, M. A.; Bozyczko-Coyne, D.; Josef, K.; Wells, G.; Tripathy, R.; Iqbal, M.; Bihovsky, R.; Senadhi, S. E.; Mallya, S.; O'Kane, T. M.; McKenna, B. A.; Siman, R.; Mallamo, J. P. Synthesis and biological activity of a series of potent fluoromethyl ketone inhibitors of recombinant human calpain I. *J. Med. Chem.* **1997**, *40*, 3820−3828.

(80) Siles, R.; Chen, S. E.; Zhou, M.; Pinney, K. G.; Trawick, M. L. Design, synthesis, and biochemical evaluation of novel cruzain inhibitors with potential application in the treatment of Chagas' disease. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4405−4409.

(81) Mott, B. T.; Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Kean-Hooi Ang, K.; Leister, W.; Shen, M.; Silveira, J. T.; Doyle, P. S.; Arkin, M. R.; McKerrow, J. H.; Inglese, J.; Austin, C. P.; Thomas, C. J.; Shiochet, B. K.; Maloney, D. J. Identification and optimization of inhibitors of trypanosomal cysteine proteases: cruzain, rhodesain, and TbCatB. *J. Med. Chem.* **2010**, *53*, 52−60.

(82) Bailey, A.; Pairaudeau, G.; Patel, A.; Thom, S. Preparation and New Use of Pyrimidine- or Triazine-2-carbonitriles for Treating Diseases Associated with Cysteine Protease Activity. International Patent Application WO2004000819 A1, 2005.

(83) *ISIS Draw*, version 2.5; Symyx Technologies Inc. (3100 Central Expressway, Santa Clara, CA 95051, U.S.), 2002.

(84) *SYBYL*, version 6.91; Tripos International (1699 South Hanley Road, St. Louis, MO, 63144, U.S.), 2003.

(85) Brickmann, J.; Goetze, T.; Heiden, W.; Moeckel, G.; Reiling, S.; Vollhardt, H.; Zachmann, C. D. Interactive Visualization of Molecular Scenarios with MOLCAD/SYBYL. In *Data Visualization in Molecular Science*; Bowie, J. E., Ed.; Addison-Wesley: Reading, MA, 1995; pp 83−97.

(86) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions. 1. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.

(87) Vicik, R.; Busemann, M.; Gelhaus, C.; Stiefl, N.; Scheiber, J.; Schmitz, W.; Schulz, F.; Mladenovic, M.; Engels, B.; Leippe, M.; Baumann, K.; Schirmeister, T. Aziridide-based inhibitors of cathepsin L: synthesis, inhibition activity, and docking studies. *ChemMedChem* **2006**, *1*, 1126−1141.