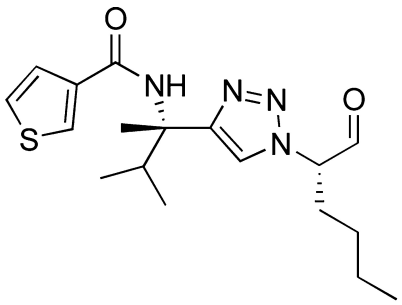
Article

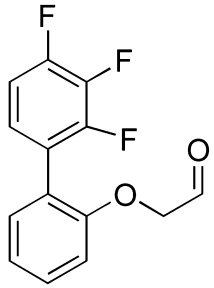
Substrate Activity Screening: A Fragment-Based Method for the Rapid Identification of Nonpeptidic Protease Inhibitors

Warren J. L. Wood, Andrew W. Patterson, Hiroyuki Tsuruoka, Rishi K. Jain, and Jonathan A. Ellman

J. Am. Chem. Soc., **2005**, 127 (44), 15521-15527• DOI: 10.1021/ja0547230 • Publication Date (Web): 13 October 2005



K_i = 9 ± 3 nM
MW = 362 Da



K_i = 490 ± 50 nM
MW = 266 Da

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 13 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 10/13/2005

Substrate Activity Screening: A Fragment-Based Method for the Rapid Identification of Nonpeptidic Protease Inhibitors

Warren J. L. Wood, Andrew W. Patterson, Hiroyuki Tsuruoka, Rishi K. Jain,† and Jonathan A. Ellman*

Contribution from the Department of Chemistry, University of California−Berkeley, Berkeley, California 94720

Received July 14, 2005; E-mail: jellman@uclink.berkeley.edu

Abstract: A new fragment-based method for the rapid development of novel and distinct classes of nonpeptidic protease inhibitors, Substrate Activity Screening (SAS), is described. This method consists of three steps: (1) a library of *N*-acyl aminocoumarins with diverse, low molecular weight *N*-acyl groups is screened to identify protease substrates using a simple fluorescence-based assay, (2) the identified *N*-acyl aminocoumarin substrates are optimized by rapid analogue synthesis and evaluation, and (3) the optimized substrates are converted to inhibitors by direct replacement of the aminocoumarin with known mechanism-based pharmacophores. The SAS method was successfully applied to the cysteine protease cathepsin S, which is implicated in autoimmune diseases. Multiple distinct classes of nonpeptidic substrates were identified upon screening an *N*-acyl aminocoumarin library. Two of the nonpeptidic substrate classes were optimized to substrates with >8000-fold improvements in cleavage efficiency for each class. Select nonpeptidic substrates were then directly converted to low molecular weight, novel aldehyde inhibitors with nanomolar affinity to cathepsin S. This study demonstrates the unique characteristics and merits of this first substrate-based method for the rapid identification and optimization of weak fragments and provides the framework for the development of completely nonpeptidic inhibitors to many different proteases.

Introduction

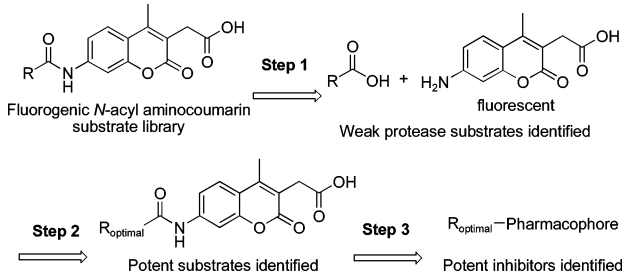
Proteases play a vital role in many physiological processes and are essential to the life cycles of many pathogens. As a result, protease inhibitors are actively being pursued to treat a large number of different diseases, including cancer, cardiovascular, autoimmune and neurodegenerative diseases, and viral and parasitic infections.[1] Typically, protease inhibitors have been developed on the basis of the preferred peptide substrates of the protease being targeted. While this approach often rapidly provides potent peptidic inhibitors, it has proven to be incredibly challenging to convert these peptidic inhibitors into nonpeptidic, drug-like structures necessary to achieve good oral bioavailability, relevant tissue penetration, and long circulating half-lives. High throughput screening (HTS) of libraries of drug-like compounds is currently the standard drug discovery method for identifying small molecule ligands for receptors and enzymes. Unfortunately, for proteases, there are few examples for which HTS has provided useful, nonpeptidic inhibitors.[2−6] For challenging protein targets, such as proteases, fragment-based approaches have been developed wherein low molecular weight fragments are identified that bind to the desired target with modest affinity and are subsequently optimized to yield more potent compounds.[7,8]

The two key challenges in fragment-based screening methods are (1) the accurate and efficient identification of weak binding fragments, and (2) an effective means for the rapid optimization of the initial weak binding fragments into higher affinity compounds. Inhibition assays common to HTS often are not effective for identifying weak-binding fragments because a high incidence of false positives can result from aggregation and nonspecific protein binding at the high fragment concentrations necessary to detect weak inhibition.[9−11] False positives can be eliminated by direct detection of binding using nuclear magnetic resonance and X-ray crystallography. These approaches have the added advantage that structural information can be obtained to provide insight for fragment optimization. However, these methods require large quantities of protein, have relatively modest throughput with dedicated use of expensive instrumentation, and are only effective when both the fragments and the proteins are soluble and well behaved at concentrations neces-

† Present address: Novartis Institute for Biomedical Research, Cambridge, MA.

(1) Leung, D.; Abbenante, G.; Fairlie, D. P. *J. Med. Chem.* **2000**, *43*, 305−341.
(2) Dudley, D. A.; et al. *J. Med. Chem.* **2000**, *43*, 4063−4070.
(3) Shrader, W. D.; Young, W. B.; Sprengeler, P. A.; Sangalang, J. C.; Elrod, K.; Carr, G. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1801−1804.
(4) Gopalsamy, A.; et al. *J. Med. Chem.* **2004**, *47*, 1893−1899.
(5) Altmann, E.; Cowan-Jacob, S. W.; Missbach, M. *J. Med. Chem.* **2004**, *47*, 5833−5836.
(6) Turner, S. R.; et al. *J. Med. Chem.* **1998**, *41*, 3467−3476.
(7) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. *J. Med. Chem.* **2004**, *47*, 3463−3482.
(8) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. *Nat. Rev. Drug Discovery* **2004**, *3*, 660−672.
(9) McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. *J. Med. Chem.* **2002**, *45*, 1712−1722.
(10) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 4265−4272.
(11) McGovern, S. L.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 1478−1483.

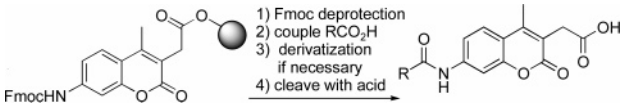
Scheme 1. Outline of the Substrate Activity Screening Method



sary to achieve significant protein binding. Tethering[12] is a mass spectrometric method for the direct detection of binding that results from disulfide interchange of a thiol-derivatized fragment and a cysteine thiol located proximal to the protein binding site. This procedure is high throughput and requires only small quantities of protein. However, a cysteine residue must be present or introduced into the protein at a position proximal to the binding site, and a library of fragments derivatized with thiols modified as mixed disulfides must be prepared. Despite these issues, fragment-based methods have been successful in identifying ligands to a number of challenging protein targets,[7,8] including the identification of inhibitors of proteases such as gelatinase B,[13] stromelysin,[14] urokinase,[15] and caspase-3.[16]

Herein, a new fragment-based method for the simple, rapid, and efficient development of novel, nonpeptidic inhibitors of proteases, called Substrate Activity Screening (SAS), is reported (Scheme 1). A library of *N*-acyl aminocoumarins with diverse, low molecular weight *N*-acyl fragments is prepared, and then in step 1, the library is screened to identify protease substrates using a one-step, high throughput fluorescence-based assay. Although nonpeptidic *N*-acyl aminocoumarins have not previously been explored, *N*-peptidyl aminocoumarins have been used extensively as fluorogenic substrates of serine and cysteine proteases for HTS, biochemical characterization, and substrate specificity determination due to the simplicity, efficiency, and sensitivity of the fluorescence-based assay.[17] Indeed, methods have been developed that enable the straightforward solid-phase synthesis of libraries of peptidyl aminocoumarins, and similarly should enable the efficient solid-phase synthesis of fragment-based *N*-acyl aminocoumarin libraries. The substrate-based screening method in step 1 (Scheme 1) has important attributes for detecting weak binding fragments in addition to being high throughput and straightforward to perform. False positives due to aggregation, protein precipitation, or nonspecific binding are not observed because active enzyme and productive active site binding are required for protease-catalyzed amide bond hydrolysis that releases the fluorescent coumarin group. In addition, in contrast to direct binding assays and traditional inhibitor screens, catalytic substrate turnover results in signal amplification, and therefore even very weak substrates can be identified at concentrations where only minimal binding to the enzyme occurs.

Scheme 2. General Method for Synthesizing Potential Substrates



Steps 2 and 3 (Scheme 1) provide a strategy for systematically and efficiently progressing substrates incorporating weak binding fragments into high affinity inhibitors, thereby addressing the second major challenge in fragment-based screening methods, the rapid optimization of fragments into inhibitors. In step 2, the activity of the substrates is rapidly optimized by the straightforward solid-phase synthesis and subsequent assay of focused libraries of substrate analogues. Step 3 then builds upon a key attribute of this mechanism-based substrate screen, that the *N*-acyl aminocoumarin must be precisely oriented in the active site to enable productive substrate cleavage, and therefore the aminocoumarin can be replaced with mechanism-based pharmacophores to directly provide protease inhibitors. The choice and versatility of pharmacophores allows reversible or irreversible inhibitors to be rapidly obtained once efficient substrates are identified.

The SAS method was successfully applied to the cysteine protease, cathepsin S, which is implicated in autoimmune diseases, such as rheumatoid arthritis and multiple sclerosis. Cathepsin S is involved in the final step of the degradation of the Invariant Chain, which is necessary for antigen presentation and subsequent immune response.[18] Multiple distinct classes of nonpeptidic substrates were identified upon screening an *N*-acyl aminocoumarin library. Two of the distinct substrate classes were optimized, leading to substrates with >8000-fold improvements in cleavage efficiency for each class. Select substrates were then directly converted to highly novel, low molecular weight, nonpeptidic aldehyde inhibitors with nanomolar affinity to cathepsin S.[19−21]

Results

Synthesis of *N*-Acyl Aminocoumarin Substrate Library. A library of *N*-acyl aminocoumarin substrates was prepared according to the procedure shown in Scheme 2. In the first step, the Fmoc group was removed from support-bound Fmoc-protected 7-amino-4-methyl coumarin acetic acid (Fmoc−AMCA) to liberate a free amine, which was then acylated with 47 commercially available carboxylic acids, which serve to introduce the diverse fragments. Carboxylic acids were selected to have a low molecular weight (av. MW ± SD is 210 ± 50), drug-like structure, and to incorporate both hydrophobic and hydrogen-bonding groups. To achieve further substrate diversity, 58 low molecular weight fragments (av. MW ± SD is 240 ± 40) were also introduced by performing transformations upon *N*-acyl aminocoumarin derivatives, including amine acylations, and 1,3-dipolar cycloadditions to afford a variety of five-membered heterocycles such as oxadiazoles, 1,2,4-triazoles, 1,2,3-triazoles, isoxazoles, isoxazolines, pyrazoles, and pyrazolines. In the final step, the *N*-acyl aminocoumarins were

(12) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9367−9372.
(13) Wang, X.; Choe, Y.; Craik, C. S.; Ellman, J. A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2201−2204.
(14) Hajduk, P. J.; et al. *J. Am. Chem. Soc.* **1997**, *119*, 5818−5827.
(15) Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J. *Nat. Biotechnol.* **2000**, *18*, 1105−1108.
(16) Erlanson, D. A.; et al. *Nat. Biotechnol.* **2003**, *21*, 308−314.
(17) Maly, D. J.; Huang, L.; Ellman, J. A. *ChemBioChem* **2002**, *3*, 17−37.
(18) Ward, Y. D.; et al. *J. Med. Chem.* **2002**, *45*, 5471−5482.
(19) Thurmond, R. L.; Beavers, M. P.; Cai, H.; Meduna, S. P.; Gustin, D. J.; Sun, S.; Almond, H. J.; Karlsson, L.; Edwards, J. P. *J. Med. Chem.* **2004**, *47*, 4799−4801.
(20) Jeffrey, S. C.; et al. *J. Med. Chem.* **2005**, *48*, 1344−1358.
(21) Leroy, V.; Thurairatnam, S. *Expert Opin. Ther. Pat.* **2004**, *14*, 301−311.

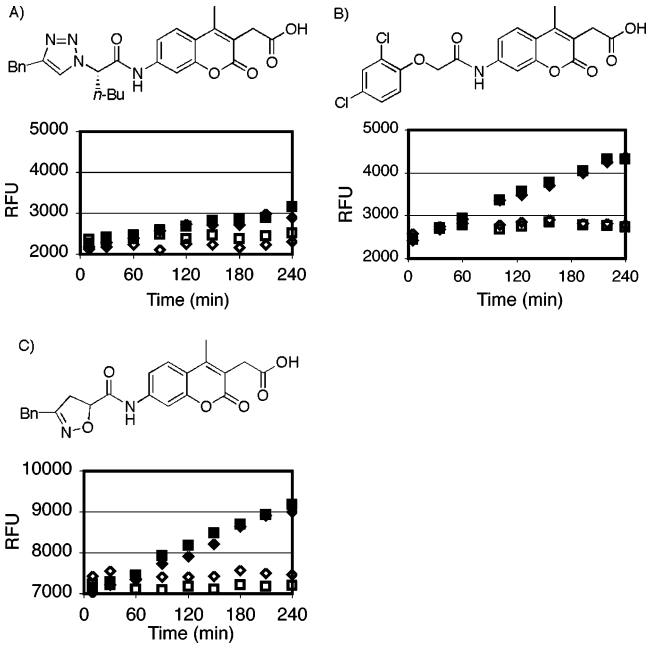


Figure 1. Relative fluorescent units (RFU) versus time for (A) 1,4-disubstituted-1,2,3-triazole, (B) phenoxyacetyl, and (C) isoxazoline substrates with cathepsin S (◆ and ■) and without cathepsin S (◇ and □).

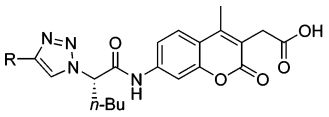
cleaved from support by treatment with trifluoroacetic acid followed by HPLC purification.

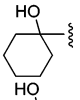
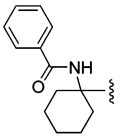
Step 1: Substrate Identification. The library of 105 *N*-acyl aminocoumarins incorporating diverse acyl fragments was initially screened at 1.0 mM in substrate and 750 nM in cathepsin S (see Supporting Information for a full list of compounds screened). From this assay, a clear increase in fluorescence over time was observed for multiple distinct substrate classes. Assay data for three of the distinct library members are shown in Figure 1. These weak initial substrate hits, with an average fragment molecular weight of 216 Da, show clear and reproducible signals over background even though they are 10^5 less efficient substrates than a standard *N*-peptidyl aminocoumarin substrate of cathepsin S, CbzLeuArg−AMCA. Based upon the facile, regioselective synthesis of 1,4-disubstituted-1,2,3-triazoles (Figure 1A), and the novelty of the phenoxyacetyl moiety for protease inhibition (Figure 1B), optimization of these two scaffolds was pursued.

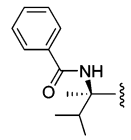
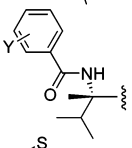
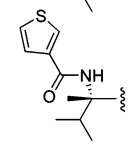
Step 2: Substrate Optimization. Analogues of the substrates were synthesized by solid-phase methods and were assayed at substrate concentrations below the K_m of the substrates to obtain substrate activities that directly correspond to the catalytic efficiencies (k_{cat}/K_m) of the substrates (see Supporting Information for a full list of compounds and activities). This was accomplished by determining the K_m of the most active substrates (~top 10%) under a set substrate concentration. If a substrate was then identified that had a K_m below the assay substrate concentration, all of the substrates assayed under that condition were then reassayed at a substrate concentration below the lowest K_m. Conducting the assay under conditions that provide relative catalytic efficiencies is necessary because transition-state theory on enzyme-catalyzed reactions predicts that the inhibitory activity (K_i) of transition-state analogue inhibitors can be correlated with the inverse of catalytic efficiency of the corresponding substrates (K_m/k_{cat}) according to eq 1, with *d* providing a measure of the effectiveness of the mechanism-based inhibitor as a transition state analogue and k_{un} equaling the rate of hydrolysis for the uncatalyzed reaction. For related series of substrates, *d* and k_{un} should remain constant, and therefore the $\log(K_i)$ should directly correspond to the $\log(K_m/k_{cat})$. This correlation has been demonstrated for multiple series of peptidic substrates and inhibitors targeting different proteases.[22]

$$\log(K_i) = \log(K_m/k_{cat}) + \log(dk_{un}) \qquad (1)$$

Table 1. Representative Structure and Activity of 1,4-Disubstituted-1,2,3-triazole Substrate Derivatives



Entry	R	Rel. k_{cat}/K_m
1	PhH_2C–	1.0
2		6.3
3		31
4		0.0
5		700

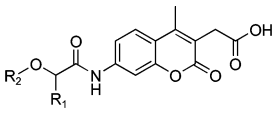
Entry	R	Rel. k_{cat}/K_m
6		35
7	 Y = H	4400
8	Y = *o*-OCH_3	110
9	Y = *m*-OCH_3	1200
10	Y = *p*-OCH_3	500
11	Y = *m*-F	3300
12		8200

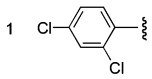
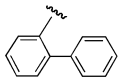
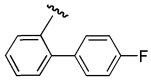
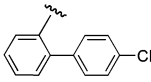
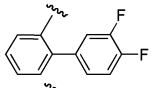
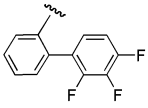
1,4-Disubstituted-1,2,3-triazole Substrates. Seventy-six derivatives of the 1,4-disubstituted-1,2,3-triazoles were synthesized by the general method outlined in Scheme 2. The support-bound aminocoumarin was acylated with an azido acid, and then a Cu(I)-catalyzed 1,3-dipolar cycloaddition was conducted with terminal alkynes. For the preparation of some of the analogues, further modification of the support-bound cycloaddition product was also performed (see Material and Methods). The substrates were then cleaved from support, purified, and assayed against cathepsin S. Table 1 shows the relative cleavage efficiencies of select substrates to highlight the progression from the initial hit (entry 1) to a substrate that is cleaved 8200-fold more efficiently (entry 12).

Thirty-four analogues of the initial hit (entry 1, Table 1) displaying diverse hydrocarbon and polar functionality were prepared and screened, resulting in the identification of hydroxy-substituted analogues (entries 2 and 3) with 6.3- and 31-fold greater cleavage efficiencies than the initial hit (entry 1). Replacement of the alcohol with an amine (entry 4) provided an inactive compound. However, evaluation of several acylated derivatives resulted in the identification of a benzamide analogue (entry 5) with 700-fold greater cleavage efficiency than the initial hit. The epimeric compounds in entries 6 and 7 were designed to combine the benzamide functionality with the aliphatic functionality that had provided the optimal cleavage efficiency for the hydroxy substituted analogues (entry 3). The enzyme showed strong chiral recognition, with one epimer (entry 7) showing 4400-fold and the other epimer (entry 6) only 35-

(22) Mader, M. M.; Bartlett, P. A. *Chem. Rev.* **1997**, *97*, 1281−1302.

Table 2. Representative Structure and Activity of Phenoxyacetyl Substrate Derivatives[a]



Entry	R_2	R_1	Rel. k_{cat}/K_m
1		H	2.6
2		H	17
3		*n*-Bu	0.0
4		H	170
5		H	2200
6		H	2800
7		H	3400
8		H	27000

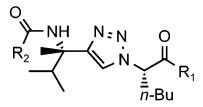
[a] The relative k_{cat}/K_m values are correlated to the relative k_{cat}/K_m reported for the initial triazole hit listed in entry 1 of Table 1.

fold greater cleavage efficiency than the initial hit (entry 1). Thirty-five substrates were then prepared and assayed to probe the binding interactions provided by the benzamide group. All of these analogues served as substrates for cathepsin S to various degrees. The analogues shown in entries 9, 10, and 11 illustrate that *meta*- and *para*-substitution is tolerated better than *ortho*-substitution (entry 8). The substrate with the highest cleavage efficiency (entry 12) incorporates a thiophene in place of the phenyl ring of the benzamide group and is cleaved 8200-fold more rapidly than the initial hit.

Phenoxyacetyl Substrates. A total of 55 derivatives of the phenoxyacetyl fragments were synthesized by the general procedure outlined in Scheme 2. The support-bound aminocoumarin was acylated with chloroacetic anhydride, and then the chloride was displaced with commercially available phenols under basic conditions. If necessary, further derivatization of the support-bound compound was carried out after chloride displacement (see Material and Methods). The substrates were then cleaved from support, purified, and assayed against cathepsin S. Table 2 shows the relative cleavage efficiencies of select substrates to highlight the progression from the initial hit (entry 1) to a substrate that is cleaved 10 000-fold more efficiently (entry 8).

Sixteen derivatives of the initial hit were first prepared and assayed to investigate the effect of substitution on the phenoxy group. A number of more active substrates were identified, including the compounds in entries 2 and 4, which show 6.6- and 64-fold greater cleavage efficiency than the initial hit, respectively. Comparison of the phenoxyacetic acid substrates to the triazole substrates prompted introduction of an *n*-butyl group alpha to the carbonyl, but the substrate was not active under the assay conditions, which confirms the very different binding modes of the two substrate classes (entry 3). Next, 11 analogues of the compound in entry 4 were prepared primarily to explore substitution at the *ortho*-, *meta*-, and *para*-positions of the terminal phenyl ring of the biaryl moiety. Substrates with substituents at the *para*-position of the terminal phenyl ring showed the greatest increase in cleavage efficiency. A 13-fold increase in activity was observed when a *para*-fluoro group was

Table 3. Activity of 1,4-Disubstituted-1,2,3-triazole Substrates and Inhibitors



Entry	R_2	R_1 = AMCA k_{cat}/K_m ($M^{-1}s^{-1}$)	R_1 = AMCA k_{cat} (s^{-1})	R_1 = AMCA K_m (μM)	R_1 = H K_i (nM)
1	Y = *o*-OCH_3	90 ± 10	0.040 ± 0.004	430 ± 50	1800 ± 800
2	Y = *m*-OCH_3	360 ± 70	0.069 ± 0.009	190 ± 30	93 ± 8
3	Y = *m*-F	900 ± 200	0.33 ± 0.04	350 ± 60	48 ± 5
4	Y = H	1100 ± 100	0.45 ± 0.04	410 ± 40	27 ± 5
5		2600 ± 500	0.77 ± 0.08	300 ± 40	9 ± 3

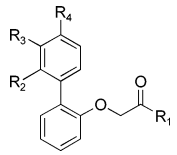
introduced on the phenyl ring (entry 5) relative to the unsubstituted system (entry 4). Based on these results, another 29 potential substrates were prepared, which primarily included various substituents at the *para*-position and various halogenated derivatives of the terminal phenyl ring. Many of these analogues were active substrates, and in general the analogues incorporating fluorine substituents were the most active, such as in entries 6, 7, and 8. The substrate with the highest cleavage efficiency (entry 8) contained three fluorine atoms on the terminal phenyl ring and was cleaved 10 000-fold more rapidly than the initial hit substrate.

Step 3: Conversion of Substrates into Inhibitors. The final step of the SAS method involves conversion of the optimal substrates to inhibitors by replacement of the aminocoumarin with a mechanism-based pharmacophore. For this study, the aminocoumarin group was replaced with a hydrogen atom to provide an aldehyde, which is a minimal mechanism-based pharmacophore for cysteine proteases. Upon binding of the inhibitor in the active site and addition of the active-site cysteine thiol to the aldehyde, a tetrahedral adduct is formed, which is an analogue of the transition-state. Based on transition-state theory (eq 1), good correlation between $\log(K_m/k_{cat})$ and $\log(K_i)$ is expected for compounds in related chemical series.[22] The substrates with greatest cleavage efficiency for each series should therefore provide the highest affinity inhibitors.

The two triazole substrates with greatest cleavage efficiency were converted to potent aldehyde inhibitors with low nanomolar K_i values (entries 4 and 5, Table 3). These nonpeptidic inhibitors are quite structurally distinct from any of the previously reported cathepsin S inhibitors. The optimal phenoxyacetyl substrate resulted in an inhibitor with nanomolar activity (490 nM) (entry 4, Table 4). This inhibitor also serves as an extremely promising starting point for drug discovery because of its low molecular weight (266 Da) and completely nonpeptidic structure, which would have unlikely been identified by traditional methods.

Substrates with a range of cleavage efficiencies for each substrate class were also converted to inhibitors to establish the correlation between the substrate cleavage efficiency and inhibitor affinity for compounds in related chemical series. Inhibitors were prepared corresponding to five triazole substrates with an 80-fold range in cleavage efficiencies (Table 3) and four phenoxyacetyl substrates with an 160-fold range in cleavage efficiencies (Table 4). As shown in Figure 2A and B, outstanding correlation between $\log(K_m/k_{cat})$ and $\log(K_i)$ was observed for both series, $R^2 = 0.98$ and 0.94, respectively, confirming that rapid substrate optimization followed by conversion of the

Table 4. Activity of Phenoxyacetyl Substrates and Inhibitors[a]



entry	R_2	R_3	R_4	R_1 = AMCA rel. k_{cat}/K_m	R_1 = H K_i (μM)
1	H	H	H	170	39 ± 9
2	H	H	CF_3	2800	1.6 ± 0.1
3	H	F	F	3400	1.3 ± 0.2
4	F	F	F	27 000	0.49 ± 0.05

[a] The relative k_{cat}/K_m values are correlated to the relative k_{cat}/K_m reported for the initial triazole hit listed in entry 1 of Table 1.

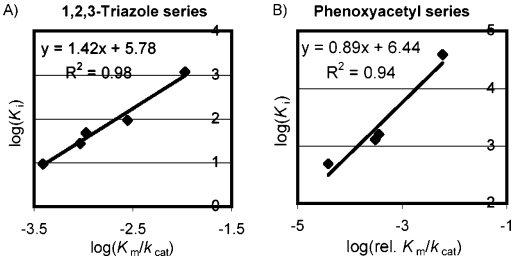


Figure 2. (A) Plot of $\log(K_m/k_{cat})$ and $\log(K_i)$ for triazole substrates and corresponding inhibitors. (B) Plot of log(relative K_m/k_{cat}) and $\log(K_i)$ for phenoxyacetyl substrates and corresponding inhibitors.

optimal substrates to inhibitors is indeed a valid approach for the progression of weak fragment-based substrates to high affinity inhibitors.

Discussion

The SAS method is the first substrate-based fragment identification and optimization method and offers clear advantages over the generation of protease inhibitor leads from peptides. For example, active peptidomimetics with drug-like properties are typically difficult to identify and often require synthetically demanding modifications such as ring constraints and bioisostere replacements. In contrast, nonpeptidic hit fragments reliably identified by SAS can be constructed through facile chemistry or commercial building blocks, making them highly chemically tractable. The fragments can then serve as flexible starting points for the efficient development of lead candidates by substrate optimization to select for higher affinity compounds with drug-like properties followed by direct conversion to inhibitors.

Indeed, the SAS method was successfully applied to the identification of two distinct classes of novel, nonpeptidic inhibitors of cathepsin S with nanomolar affinity. In the first step, a diverse library of fragment-based *N*-acyl aminocoumarins was screened and resulted in the identification of multiple distinct substrate classes. In the second step, substrate optimization by rapid analogue synthesis and evaluation resulted in substrates with >8000-fold greater cleavage efficiency than the initial substrates. Direct conversion of the substrates to inhibitors was then accomplished simply by replacement of the aminocoumarin with a hydrogen atom to introduce the aldehyde mechanism-based pharmacophore. The novel, nonpeptidic inhibitors that resulted from this process are distinct from previously reported inhibitors of cathepsin S.

The substrate-based screening method reported here for fragment identification and progression to inhibitors could be directly applied to many different proteases. Many of these enzymes are important therapeutic targets for the treatment of a variety of human diseases.[1] Moreover, substrate-based fragment screening and optimization approaches could be applied to many different enzyme families with the key requirements being an appropriate method for efficiently and accurately monitoring substrate turnover and the availability of mechanism-based pharmacophores for converting optimized substrates into inhibitors.

Material and Methods

General Methods. Low amine content DMF was purchased from EM Science (Cincinnati, OH), Wang resin was purchased from Novabiochem (San Diego, CA), and HATU was obtained from PerSeptive Biosystems (Foster City, CA). Reactions were conducted in polypropylene cartridges equipped with 70 mm PE frits (Applied Separations, Allentown, PA) and Teflon stopcocks. Reactions were gently rocked on an orbital shaker table during the solid-phase reactions. Heating solid-phase reactions were conducted in a glass vial (Kimble Artible No. 60700-5) with a screw cap (Kontes Article No. 410119-2015 open cap with Kimble Article No. 749110-0022 valve) and gently stirred in a heating module (PIERCE Reacti-Therm III). Fmoc-protected 7-amino-4-methyl coumarin acetic acid (Fmoc−AMCA) was synthesized according to a method analogous to the synthesis of 7-amino-4-carbamoylmethylcoumarin.[23] The HCl salt of (*S*)-3-amino-3,4-dimethyl-1-pentyne was prepared by the addition of lithium trimethylsilylacetylide to *N*-sulfinyl isopropyl methyl ketimine according to literature procedures.[24]

General Synthesis of Substrates. Loading of Fmoc−AMCA onto Wang Resin (Scheme 2). A solution of Fmoc−AMCA (2.3 equiv, 0.3 M) and pyridine (3.8 equiv, 0.5 M) in anhydrous DMF was added to a cartridge containing 5 g of Wang resin (100−200 mesh, 0.9−1.30 mmol/g) pre-swollen in 50 mL of anhydrous DMF. After slow addition of 2,6-dichlorobenzoyl chloride (2.3 equiv), the mixture was shaken for 48 h. After solvent removal, the resin was washed with three portions (50 mL) each of DMF, THF, CH_3OH, THF, and CH_2Cl_2. The remaining solvent was removed in vacuo, and the Fmoc−AMCA−Wang resin was stored at −20 °C until further use.

General Procedure for Loading Carboxylic Acids onto AMCA−Wang Resin (Scheme 2). A 20% solution of piperidine (vol/vol) in DMF (5 mL) was added to a cartridge containing 0.15 g of Fmoc−AMCA−Wang resin (0.30−0.72 mmol/g by Fmoc quantitation). The mixture was shaken for 5 min, the solution was removed, and the resin was washed with 5 mL of DMF. The process was repeated once, and the resin was washed with DMF (3 × 5 mL). A solution of HATU (0.40 M, 5 equiv), collidine (0.40 M, 5 equiv), and carboxylic acid (0.40 M, 5 equiv) in DMF was added to the resin, and the mixture was shaken for 12 h. After removal of the solution, the resin was washed with 5 mL of DMF, and the process was repeated. After removal of the solution, the resin was washed with three portions (5 mL) each of DMF, THF, CH_3OH, THF, and CH_2Cl_2, and after the solvent was removed in vacuo, the resin was further derivatized or the product was cleaved from the solid support [see General Procedure for Support Cleavage and Purification of Product].

General Procedure for Synthesis of Triazole Substrates Listed in Entries 1−4 (Table 1). N_3−Nle−OH was prepared from L-norleucine according to a literature procedure[25] and then was loaded onto AMCA−Wang resin according to the general procedure for loading carboxylic acids described above. To 0.10−0.15 mmol of the resulting

(23) Maly, D. J.; Leonetti, F.; Backes, B. J.; Dauber, D. S.; Harris, J. L.; Craik, C. S.; Ellman, J. A. *J. Org. Chem.* **2002**, *67*, 910−915.
(24) Cogan, D. A.; Liu, G. C.; Ellman, J. *Tetrahedron* **1999**, *55*, 8883−8904.
(25) Lundquist, J. T. t.; Pelletier, J. C. *Org. Lett.* **2001**, *3*, 781−783.

N_3−Nle−AMCA−Wang resin were added alkyne (0.02 M, 2.0 equiv), *i*-Pr_2EtN (1 M, 100 equiv), and CuI (0.0025 M, 0.25 equiv) in THF, and the mixture was shaken for 20−48 h. The solution was removed, the resin was washed with 20 mL of THF, and the process was repeated. After removal of the solution, the resin was washed with three portions (20 mL) each of THF, CH_3OH, THF, and CH_2Cl_2, and then the product was cleaved from the solid support [see General Procedure for Support Cleavage and Purification of Product].

General Procedure for Synthesis of Triazole Substrates Listed in Entries 5−12 (Table 1). To the cartridge containing the pre-swollen N_3−Nle−AMCA−Wang resin (0.10−0.15 mmol) were added the HCl salt of (*S*)-3-amino-3,4-dimethyl-1-pentyne (0.02 M, 1.5 equiv), *i*-Pr_2EtN (1.3 M, 100 equiv), and CuI (0.04 M, 3.0 equiv) in THF, and the mixture was shaken for 20−48 h. After removal of the solution, the resin was washed with three portions (25 mL) each of THF, CH_3OH, acetonitrile, and THF. After solvent removal, *i*-Pr_2EtN (8 equiv) was added to the derivatized resin. To a THF solution of carboxylic acid (0.1 M, 3.5 equiv) and triphosgene (0.03 M, 1.1 equiv) was added 2,4,6-collidine (0.3 M, 10 equiv). The resulting slurry was stirred for 1 min and then was added to the cartridge containing the derivatized resin. The resulting mixture was shaken for 4−12 h. After removal of the solution, the resin was washed with 25 mL of THF, and the coupling was repeated two more times. After removal of the solution, the resin was washed with three portions (25 mL) each of THF, CH_3OH, THF, and CH_2Cl_2, and then the product was cleaved from the solid support [see General Procedure for Support Cleavage and Purification of Product].

General Procedure for the Synthesis of Phenoxyacetyl Substrates (Table 2). Loading of Chloroacetic Acid. A 20% solution of piperidine (vol/vol) in DMF (17 mL) was added to a 40 mL cartridge containing 2.0 g (1.1 mmol, 0.553 mmol/g) of Fmoc−AMCA−Wang resin. The mixture was shaken for 15 min, and the resin was washed with three portions (20 mL) each of DMF, CH_3OH, and CH_2Cl_2. After addition of 20 mL of CH_2Cl_2−pyridine (20:1, vol/vol) and chloroacetic acid anhydride (752 mg, 4.4 mmol), the reaction mixture was agitated for 2 h at room temperature, and then the resin was washed with two portions (20 mL) each of DMF, water, DMF, and ethyl ether. The resin was finally dried under reduced pressure to give the desired resin-bound chloroacetamide−AMCA.

Loading of Phenol Derivatives. To resin-bound chloroacetamide−AMCA (0.1 mmol) was added a solution of the appropriate phenol derivative (5 equiv) and 2-*tert*-butylimino-2-diethylamino-1,3-dimethyl-perhydro-1,3,2-diazaphosphorine (3 equiv) in THF (2 mL). After being stirred gently at 70 °C for 15−18 h, the resin was transferred into a 10 mL polypropylene cartridge and subsequently rinsed with two portions (2 mL) each of DMF, CH_3OH, and CH_2Cl_2 to give the desired resin.

Suzuki Coupling (Table 2, Entries 5−8). To the resin obtained above using 2-bromophenol (0.1 mmol) was added either DME (2 mL), a 2 M aqueous solution of Na_2CO_3 (0.25 mL, 0.5 mmol), arylboronic acid (0.5 mmol, 5 equiv), and tetrakis(triphenylphosphine)palladium (11 mg, 0.01 mmol) or potassium phosphate (106 mg, 0.5 mmol), arylboronic acid (5 equiv), tris(dibenzylideneacetone)dipalladium(0) chloroform adduct (10 mg, 0.01 mmol), 2-(dicyclohexylphosphino)-biphenyl (14 mg, 0.04 mmol), and dioxane (2 mL). After being purged with N_2, the capped reaction vessel was placed in a 80−90 °C oil bath, and the mixture was gently stirred for 15−18 h. The resin was transferred into a 10 mL polypropylene cartridge and subsequently rinsed with two portions (2 mL) each of DMF, water, CH_3OH, and CH_2Cl_2 to give resin-bound biaryloxyacetamide−AMCA. The product was then cleaved from support according to the general cleavage and purification conditions.

General Procedure for Support Cleavage and Purification of Product. The resin was swollen in CH_2Cl_2. To the swollen resin was then added a 5 mL solution of 9:1 CH_2Cl_2:(95% CF_3CO_2H, 2.5% H_2O, 2.5% triisopropylsilane). The mixture was shaken or allowed to sit for 1−2 h. Upon removal of the solution, the resin was washed with one portion of the cleavage solution and three portions of CH_2Cl_2. The combined washes were concentrated under reduced pressure. The crude product mixture was purified by HPLC [preparatory reverse phase C18 column (24.1 × 250 mm), acetonitrile/H_2O−0.1% CF_3CO_2H; 5−95% over 50 min; 8 mL/min; 254 nm detection for 60 min] and either lyophilized or extracted with ethyl acetate followed by solvent removal. The identity and purity of each compound was confirmed by LCMS analysis of the purified material.

General Synthesis of 1,2,3-Triazole-aldehyde Inhibitors (Entries 1−5, Table 3). The HCl salt of (*S*)-3-amino-3,4-dimethyl-1-pentyne was acylated with the appropriate benzoic anhydride derivative. A Cu(I)-catalyzed 1,3-dipolar cycloaddition between the terminal alkyne of the benzamide product and N_3−Nle−OH was next performed. The carboxylic acid of the resulting 1,2,3-triazole derivative was then methylated with diazomethane, and the resulting methyl ester was then reduced with DIBAL to give the corresponding aldehyde. For the synthesis of the aldehyde inhibitor listed in entry 5 (Table 3), it was necessary to first protect (*S*)-3-amino-3,4-dimethyl-1-pentyne with Boc-anhydride, followed by performing the copper-catalyzed 1,3-dipolar cycloaddition with norleucine azide methyl ester (N_3−Nle−OCH_3). The Boc group on the amine of the resulting triazole product was cleaved with CF_3CO_2H, and the resulting amine was then acylated with the acid chloride of thiophene-3-carboxylic acid. The methyl ester was reduced with DIBAL to give the aldehyde inhibitor. Due to facile aldehyde to enol tautomerization, significant epimerization of the aldehyde inhibitors occurred upon reverse-phase HPLC purification and isolation [detailed procedures and full analytical characterization are provided in the Supporting Information].

General Synthesis of Phenoxyacetyl Aldehyde Inhibitors (Entries 1−4, Table 4). The appropriate 2-hydroxybiphenyl derivative, which for the inhibitors listed in entries 1−4 (Table 4) was prepared by Suzuki reaction between 2-bromophenol and the appropriate aryl boronic acid, was coupled with ethyl bromoacetate under basic conditions followed by reduction with $LiAlH_4$. Dess−Martin oxidation of the resulting alcohol provided the desired aldehyde inhibitor, which was purified either by alumina or by silica gel chromatography [detailed procedures and full analytical characterization are provided in the Supporting Information].

General Assay Procedure. The proteolytic cleavage of *N*-acyl aminocoumarins by cathepsin S was conducted in DYNATECH Microfluor fluorescence 96-well microtiter plates (black plates), and readings were taken on a Molecular Devices Spectra Max Gemini XS instrument. The excitation wavelength was 370 nm, and the emission wavelength was 455 nm with a cutoff of 435 nm. For CbzLeuArg−AMC, an excitation wavelength of 355 nm and an emission wavelength of 450 were used. The assay buffer consisted of a 100 mM solution of pH 6.1 sodium phosphate buffer with 100 mM sodium chloride, 1 mM of DTT, 1 mM of EDTA, and 0.001% Tween-20.

Assay Procedure for *N*-Acyl Aminocoumarin Substrates. Assays were conducted at 37 °C in duplicate with and without the enzyme. In each well were placed 38 μL of enzyme solution and 2 μL of a DMSO solution of an *N*-acyl aminocoumarin. For assays longer than 10 min, SealPlate (adhesive sealing films for microplates) was used to seal the plate between readings, to prevent evaporation. Relative fluorescent units (RFU) were measured at regular intervals over a period of time (maximum 4 h). A plot of RFU versus time was made for each library member with and without cathepsin S.

To determine the K_m and k_{cat} of selected substrates, assays were conducted at 37 °C in duplicate with and without the enzyme at six substrate concentrations above and below the K_m of each substrate. In each well were placed 38 μL of enzyme solution and 2 μL of a DMSO substrate solution. RFU were measured at regular intervals over a period of time (maximum 1 h). A plot of RFU/s versus substrate concentration, analyzed using KaleidaGraph, was used to determine K_m and V_{max}.

Using the RFU per μM for AMCA and the cathepsin S concentration, determined by E-64c titration,[26] k_{cat} was determined.

Assay Procedure for Aldehyde Inhibitors. Time-dependent inhibition assays were conducted with the enzyme and inhibitor preincubated for 10, 30, and 60 min. For all K_i determinations, the assays were conducted at a time when equilibrium between the enzyme and inhibitor was established, as determined by the time-dependent inhibition assays. The dissociation constants (K_i) were calculated by the method of Dixon.[27] Two concentrations of CbzLeuArg−AMC (AMC = 7-amino-4-methylcoumarin, Peptide International, Louisville, KY) were used (2 and 5 μM, K_m = 23 μM[28]), and the cathepsin S concentration in the assays was 0.6 nM. Assays were conducted in duplicate with and without inhibitor at five inhibitor concentrations to provide from 15% to 90% enzyme inhibition. In each well were placed 180 μL of enzyme solution and 10 μL of a DMSO inhibitor solution. The resulting solutions were incubated for 30 min at 37 °C, and then 10 μL of the substrate CbzLeuArg−AMC was added and cleavage of CbzLeuArg−AMC was monitored over 5 min. The reversibility of the inhibitors was confirmed by performing dialysis using Pierce Slide-A-Lyzer extra strength dialysis cassettes, with a molecular weight cutoff of 10 kDa. Enzyme and inhibitor were preincubated with an inhibitor concentration sufficient to completely inactivate the enzyme, and then dialysis was conducted followed by evaluation of regain of activity.

Acknowledgment. We acknowledge support from the NIH (Grant GM54051). H.T. also gratefully acknowledges support from Sankyo Co., Ltd.

Supporting Information Available: Tables listing all substrates and detailed procedures for the synthesis of all substrates and inhibitors, and full author lists for refs 2, 4, 6, 16, 18, and 20. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0547230

(26) Barrett, A. J.; Kembhavi, A. A.; Brown, M. A.; Kirschke, H.; Knight, C. G.; Tamai, M.; Hanada, K. *Biochem. J.* **1982**, *201*, 189−198.
(27) Dixon, M. *Biochem. J.* **1953**, *55*, 170−171.
(28) Bromme, D.; Bonneau, P. R.; Lachance, P.; Storer, A. C. *J. Biol. Chem.* **1994**, *269*, 30238−30242.