Use of X-ray Co-crystal Structures and Molecular Modeling To Design Potent and Selective Non-peptide Inhibitors of Cathepsin K

Renee L. DesJarlais,*,†,○ Dennis S. Yamashita,*,‡ Hye-Ja Oh,‡ Irene N. Uzinskas,‡ Karl F. Erhard,‡ Andrew C. Allen,§ R. Curtis Haltiwanger,† Baoguang Zhao,∥ Ward W. Smith,∥ Sherin S. Abdel-Meguid,∥ Karla D'Alessio,⊥ Cheryl A. Janson,⊥ Michael S. McQueney,⊥ Thaddeus A. Tomaszek,⊗ Mark A. Levy,⊗ and Daniel F. Veber‡

Departments of Physical and Structural Chemistry, Medicinal Chemistry, Analytical Chemistry, Structural Biology, Protein Biochemistry and Molecular Recognition, SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

Received April 8, 1998

Cathepsin K is a cysteine protease that is selectively expressed by osteoclasts.[1] Inhibitors of this enzyme have been shown to inhibit bone resorption and provide a promising new approach for treating osteoporosis.[1] Efforts aimed at discovering selective inhibitors of cathepsin K have led to several novel classes of cysteine protease inhibitors: 1,3-bis(acylamino)-2-propanones,[2a] diacylcarbohydrazides,[2b] and diacylhydrazines.[2b] All of the previously disclosed inhibitors contain a central carbonyl adjacent to an amino-blocked leucinyl moiety, a potential source of instability in biological systems. By making use of X-ray crystal structures of cathepsin K/inhibitor complexes and molecular modeling, cathepsin K inhibitors of the 1,3-bis(acylamino)-2-propanone series that lack a leucinyl group have now been designed. These inhibitors are equipotent with their closest leucine-derived analogues and are selective for human cathepsin K over human cathepsins B, L, and S.

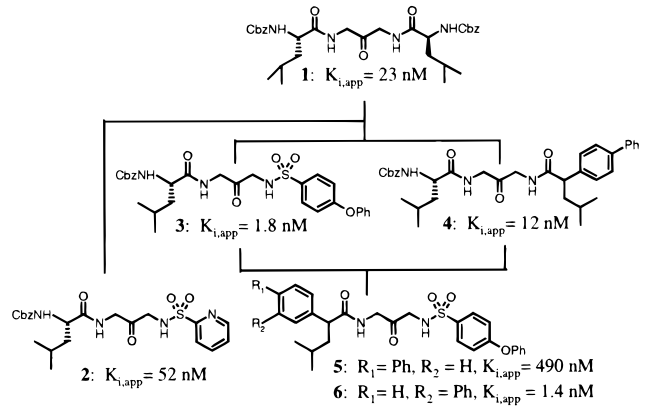


Figure 1. Evolution of non-peptide cathepsin K inhibitors.

To decrease the peptidic nature of our lead compound, 1,3-bis(Cbz-Leu-NH)-2-propanone (**1,** $K_{i,app}$ = 23 nM, Figure 1),[2a] analogues were synthesized in which one of the Cbz-Leu groups was replaced with peptidomimetics. For instance, the 2-pyridylsulfonyl analogue (**2**, Figure 1) was notable since it had increased water solubility with only a 2-fold loss in potency relative to **1**. Examination of the 3-dimensional structures of cathepsin K/inhibitor complexes, obtained by X-ray crystallography,[1i,2] indicated several important recognition elements in our cathepsin K inhibitors including the isobutyl side chain of the leucine, which binds in the hydrophobic S2[3] pocket of the enzyme and the two Cbz phenyl rings, which each form aromatic−aromatic interactions, one with Tyr 67 on the unprime side and the other with Trp 184 on the prime[3] side of the active site. This structural data facilitated the design of Cbz-Leu mimetics such as the previously described (4-phenoxyphenyl)sulfonyl group present in **3** (Figure 1),[2a] a high-potency inhibitor ($K_{i,app}$ = 1.8 nM) where the 4-phenoxy group was proposed to interact with Trp 184.

Another Cbz-Leu mimetic, the 2-(4-biphenyl)-4-methylvaleryl group, was designed to provide both an aromatic ring intended to mimic the Cbz phenyl and an isobutyl group intended to mimic the side chain of leucine. In addition, the biphenyl is bulkier and more rigid than the benzyl carbamate of the Cbz-Leu. The 4-substituted biphenyl was selected because it best matched conformation of the prime side Cbz-Leu[4] as seen in cathepsin K/inhibitor structures.[2] To test this design, **4** (Figure 1) was synthesized and found to have potency comparable to that of **1**.

Having found several replacements for one of the Cbz-Leu groups, we sought a mimetic to replace the remaining amino acid. Both modeling and structural data[2] suggested that compounds with a single Cbz-Leu were likely to bind with the Cbz-Leu on the unprime side of the active site and that successful mimicry would require filling both the S2 and S3 pockets. The 2-(4-biphenyl)-4-methylvaleryl group contains substituents that can mimic both the Cbz phenyl and the leucine isobutyl, so we hypothesized that this group might serve as an unprime side Cbz-Leu mimetic. To test this idea, **5**, which contains both of the above Cbz-Leu mimetics (present in **3** and **4**), was prepared and was found to have a $K_{i,app}$ of 490 nM, a 270-fold loss in potency relative to **3** (Figure 1).

To understand the diminished activity of **5**, a more rigorous modeling study was conducted. First, a thiohemiketal adduct of

* Corresponding authors.
† Department of Physical and Structural Chemistry.
‡ Department of Medicinal Chemistry.
§ Department of Analytical Chemistry.
∥ Department of Structural Biology.
⊥ Department of Protein Biochemistry.
⊗ Department of Molecular Recognition.
○ Current address: 3-Dimensional Pharmaceuticals, Inc., 665 Stockton Drive, Suite 104, Exton, PA 19341.

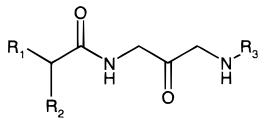
(1) (a) Drake, F. H.; Dodds, R. A.; James, I. E.; Conner, J. R.; Debouck, C.; Richardson, S.; Lee-Rykaczewski, E.; Coleman, L.; Rieman, D.; Barthlow, R.; Hastings, G.; Gowen, M. *J. Biol. Chem.* **1996**, *271*, 12511−12516. (b) Bossard, M. J.; Tomaszek, T. A.; Thompson, S. K.; Amegadzie, B. Y.; Hanning, C. R.; Jones, C.; Kurdyla, J. T.; McNulty, D. E.; Drake, F. H.; Gowen, M.; Levy, M. A. *J. Biol. Chem.* **1996**, *271*, 12517−12524. (c) Inaoka, T.; Bilbe, G.; Ishibashi, O.; Tezuka, K.; Kumegawa, M.; Kokubo, T. *Biochem. Biophys. Res. Commun.* **1995**, *206*, 89−96. (d) Bromme, D.; Okamoto, K. *Biol. Chem. Hoppe-Seyler* **1995**, *376*, 379−384. (e) Bromme, D.; Okamoto, K.; Wang, B. B.; Biroc, S. *J. Biol. Chem.* **1996**, *271*, 2126−2132. (f) Shi, G.-P.; Chapman, H. A.; Bhairi, S. M.; DeLeeuw, C.; Reddy, V. Y.; Weiss, S. J. *FEBS Lett.* **1995**, *357*, 129−134. (g) Tezuka, K., Tezuka, Y.; Maejima, A.; Sato, T.; Nemoto, K.; Kamioka, H.; Hakeda, Y.; Kumegawa, M. *J. Biol. Chem.* **1994**, *269*, 1106−1109. (h) Gelb, B. D.; Shi, G.-P.; Chapman, H. A.; Desnick, R. J. *Science* **1996**, *273*, 1236−1238. (i) 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. *Nat. Struct. Biol.* **1997**, *4*, 109−111. (j) McGrath, M. E.; Klaus, J. L.; Barnes, M. G.; Bromme, D. *Nat. Struct. Biol.* **1997**, *4*, 105−109. (k) McQueney, M. S.; Amegadzie, B. Y.; D'Alessio, K. D.; Hanning, C. R.; McLaughlin, M. M.; McNulty, D.; Carr, S. A.; Ijames, C.; Kurdyla, J.; Jones, C. S. *J. Biol. Chem.* **1997**, *272*, 13955−13960.

(2) (a) Yamashita, D. S.; Smith, W. W.; Zhao, B.; Janson, C. A.; Tomaszek, T. A.; Bossard, M. J.; Levy, M. A.; Oh, H.-J.; Y. 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.; DesJarlais, R. L.; Gleason, J. G.; Veber, D. F. *J. Am. Chem. Soc.* **1997**, *119*, 11351−11352. (b) 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.; Lee-Rykaczewski, L.; James, I. E.; Lark, M. W.; Drake, F. H.; Gowen, M.; Gleason, J. G.; Veber, D. F. *Proc. Natl. Acad. Sci. U.S.A.*, **1997**, *94*, 14249−14254.

(3) Nomenclature of Schechter and Berger described in Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157−162.

S0002-7863(98)01171-8 CCC: $15.00
Published on Web 08/22/1998

Table 1. Activities[9,10] vs Cathepsins K, B, L, and S[a]



compound	R_1	R_2	R_3	cath K $K_{i,app}$ (nM)	cath B $K_{i,app}$ (nM)	cath L $K_{i,app}$ (nM)	cath S $K_{i,app}$ (nM)
3 (*S*)	Cbz-NH	*i*-Bu	4-PhOPhSO$_2$	1.8	>10000	1400	80
4 (*S*)	Cbz-NH	*i*-Bu	*a*	12	*b*	*b*	*b*
5 (*R*/*S*)	4-BP[c]	*i*-Bu	4-PhOPhSO$_2$	490	4100	>10000	1100
6 (*R*)	3-BP	*i*-Bu	4-PhOPhSO$_2$	1.4	>10000	>1000[d]	910
6 (*S*)	3-BP	*i*-Bu	4-PhOPhSO$_2$	200	*b*	*b*	*b*
7	3-BP	H	4-PhOPhSO$_2$	12000	>10000	>10000	>10000
8	4-BP	H	4-PhOPhSO$_2$	9900	1600	>1000[d]	440
9 (*R*)	3-BP	*i*-Bu	2-py-SO$_2$	3.5	6400	680	>10000

[a] (*R*,*S*)-2-(4-biphenyl)-4-methylvaleryl. [b] nd = not determined. [c] BP = biphenyl. [d] nl = nonlinear kinetics.

inhibitor **5** and an isolated cysteine amino acid was constructed using MacroModel version 5.5,[5] minimizing with the MacroModel implementation of the AMBER force field.[6] Then, the molecule was fitted interactively into the active site of cathepsin K using UCSF MidasPlus.[7] The fitting involved first overlapping the tetrahedral adduct of the proposed inhibitor with that from a cathepsin K/inhibitor[1i,2] crystal structure and then adjusting torsion angles to optimize the fit. This modeling experiment revealed that the terminal phenyl of the 4-biphenyl moiety of **5** was not positioned to interact with Tyr 67 in either of the crystallographically observed subsites (between Tyr 67 and Asp 61, a site thought to represent the S3 subsite,[2b] or between Tyr 67 and Leu 160, a site that probably accommodates the extended peptide backbone of a substrate[2a]). *More importantly, the 3-biphenyl regioisomer of* **5** *(rather than the 4-biphenyl) was modeled and predicted to better engage Tyr 67, particularly by binding in the S3 subsite.*

To test this hypothesis, **6** (Figure 1), the 3-biphenyl regioisomer of **5**, was prepared and, indeed, **6** had a $K_{i,app}$ of 1.4 nM (equipotent with **3**[2a]) and also had greater than 500-fold selectivity over human cathepsins B, L, and S (Table 1). To explore the more specific structural predictions of our models, limited SAR studies were undertaken. The desisobutyl analogues of **6** (and **5**) and **7** (and **8**) both had dramatic decreases in activity. Furthermore, measurement of the activity of each of the two enantiomers of **6** revealed that the high potency resides in the R isomer,[8] which corresponds to L-leucine.

An X-ray crystal structure of the complex of cathepsin K and **9**, a more water soluble analogue of **6** in which the 4-phenoxyphenyl was replaced by a 2-pyridyl group (as present in **2**), was solved.[9] The resulting structure delineated interactions with the biphenyl and isobutyl portions of the inhibitor that correspond closely to those predicted by modeling (Figure 2).

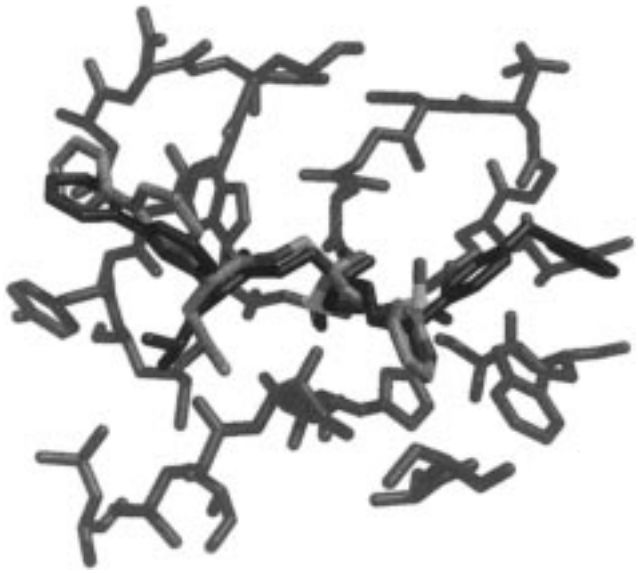


Figure 2. X-ray structure of **9** (colored by atom) and cathepsin K (cyan) and the predicted binding model of **6** (purple).

In conclusion, we designed two Cbz-Leu mimetics, the 2-(4-biphenyl)-4-methylvaleryl and the 2-(3-biphenyl)-4-methylvaleryl groups, in the context of cathepsin K binding by the 1,3-bis-(acylamino)-2-propanone class of inhibitors. The former can replace the Cbz-Leu on the prime side, but not the unprime side, and the latter can replace the Cbz-Leu on the unprime side of the active site. Detailed modeling of the peptidomimetics in the active site was critical to the design of the optimal regioisomer.

Acknowledgment. We thank Stacie M. Halbert and Dr. Scott K. Thompson for providing the 2-(4-biphenyl)-4-methylvaleric acid, Lee M. Katrincic and Priscilla Offen for NMR spectra and interpretation, and Dr. Fred H. Drake and Dr. William H. Miller for helpful discussions.

Supporting Information Available: Experimental details and structural characterization (13 pages, print/PDF). The X-ray structural information for the *p*-bromphenethylene salt of 2-(3-biphenyl)-4-methylvaleric acid is available in CIF format, through the Internet only. See any current masthead page for ordering information and Web access instructions.

JA981171V

(4) A model of the 2-(4-biphenyl)-4-methylvaleryl group was constructed in SYBYL version 6.3 and minimized with MAXIMIN using the standard settings. This model was then fit onto the unprimed side Cbz-Leu group from the crystal structure of **1** with cathepsin K to best match the 4-phenyl, the carbon bearing the isobutyl group, and the carbonyl with the Cbz phenyl, the leucine α carbon, and the leucine carbonyl. Sybyl is available commercially from Tripos Associates, Inc.

(5) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440−467.

(6) (a) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−784. (b) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252. (c) McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7743−7746.

(7) Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R. *J. Mol. Graphics* **1988**, *6*, 13−27.

(8) Both enantiomers were characterized by CD, and the active enantiomer is consistent with having the R configuration based on the observed negative Cotton effect at 240 nm ($\pi-\pi^*$). Also, 2-(3-biphenyl)-4-methylvaleric acid was resolved by recrystallization of the (*S*)-*p*-bromophenethylamine salt, and a single-crystal X-ray structure determination of this salt (clear colorless needles, monoclinic, $P2_1$, a = 14.337(7) Å, b = 6.477(3) Å, c = 14.386(6) Å, β = 113.41(7)°, U = 1226(1) Å^3, Z = 2, R = 0.064, and GOF = 1.077) confirmed the R configuration. Active cathepsin K inhibitor analogues were prepared from this enantiomerically pure material.

(9) X-ray crystallography and synthesis experimental conditions and data are described in the Supporting Information. The coordinates for the complex of cathepsin K and **9** have been deposited in the Brookhaven Protein Data Bank, accession number 1BGO.

(10) Inhibitors were evaluated for inhibition against purified recombinant cathepsin K as described in ref 1b. Inhibitors were assayed against human liver cathepsin B (Calbiochem), human liver cathepsin L (Calbiochem), and human recombinant cathepsin S (Shi, G. P.; Munger, J. S.; Meara, J. P.; Rich, D. H.; Chapman, H. A. *J. Biol. Chem.* **1992**, *267*, 7258−7262) with the following substrates: Cbz-Phe-Arg-AMC at 50 μM (K_m = 140 μM), Cbz-Phe-Arg-AMC at 5 μM (K_m = 3 μM), and Cbz-Val-Val-Arg-AMC at 50 μM (K_m = 70 μM), respectively, in 100 mM sodium acetate, 20 mM cysteine, 5 mM EDTA, pH 5.5, buffer with a final DMSO concentration of 10%.