Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 14 (2004) 275–278
Bioorganic & Medicinal Chemistry Letters

Exploration of the P^1 SAR of aldehyde cathepsin K inhibitors

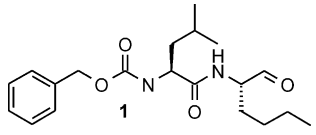
John G. Catalano,[a] David N. Deaton,[a] Eric S. Furfine,[b,†] Annie M. Hassell,[c] Robert B. McFadyen,[a,*] Aaron B. Miller,[c] Larry R. Miller,[d] Lisa M. Shewchuk,[c] Derril H. Willard, Jr.[e,*] and Lois L. Wright[b]

[a]*Department of Medicinal Chemistry, GlaxoSmithKline, Five Moore Drive, Research Triangle Park, NC 27709-3398, USA*
[b]*Department of Molecular Biology, GlaxoSmithKline, Five Moore Drive, Research Triangle Park, NC 27709-3398, USA*
[c]*Department of Structural Chemistry, GlaxoSmithKline, Five Moore Drive, Research Triangle Park, NC 27709-3398, USA*
[d]*Department of Molecular Pharmacology, GlaxoSmithKline, Five Moore Drive, Research Triangle Park, NC 27709-3398, USA*
[e]*Department of Gene Expression and Protein Biochemistry, GlaxoSmithKline, Five Moore Drive, Research Triangle Park, NC 27709-3398, USA*

Received 21 August 2003; accepted 3 September 2003

Abstract—The synthesis and biological activity of a series of aldehyde inhibitors of cathepsin K are reported. Exploration of the properties of the S^1 subsite with a series of α-amino aldehyde derivatives substituted at the P^1 position afforded compounds with cathepsin K IC_{50}s between 52 μM and 15 nM.

Cathepsin K is a cysteine protease located primarily in osteoclasts that plays a crucial role in bone resorption.[1] Cathepsin K cleaves type I collagen and other components of the bone matrix. Inhibition of this process should decrease bone resorption and help correct the imbalance in bone remodelling which favors bone resorption over bone formation.



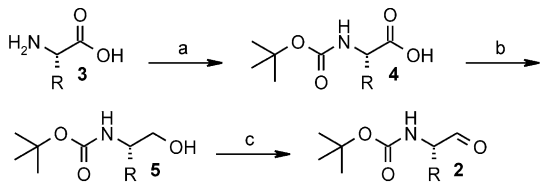
1

As part of a more extensive effort to develop new treatments for bone disorders, this group sought to discover cathepsin K inhibitors. A preliminary screen of known aldehyde inhibitors of cysteine and serine proteases in our compound collection identified calpeptin **1** (Cbz-Leu-Nle-H) as a potent inhibitor of cathepsin K ($IC_{50} = 0.11$ nM).[2,3] Further focussed screening of close analogues to calpeptin pinpointed the truncated analogue **2l** (Boc-Nle-H) as a cathepsin K inhibitor ($IC_{50} =$ 51 nM). Although the norleucine-based aldehyde **2l** was nearly 500-fold less potent than calpeptin **1**, its small size and absence of peptide bonds made it an excellent starting lead to develop cathepsin K inhibitors. Initial efforts focussed on exploring the S^1 subsite via alterations of the P^1 moiety and are detailed in this report.

The primary synthetic route to P^1 aldehyde analogues **2** is depicted in Scheme 1. Whenever possible the aldehydes **2** were synthesized from commercially available amino acids **3**, *tert*-butyloxycarbonyl amino acids **4**, or *tert*-butyloxycarbonyl amino alcohols **5**. Other amino acids were prepared according to literature precedent. Amino acids **3** were protected using di-*tert*-butyl dicarbonate and aqueous sodium hydroxide.[4] The protected amino acids **4** were then reduced to the amino alcohols **5** via the mixed anhydride method using isopropyl chloroformate and sodium borohydride.[5] Moffatt oxidation of the alcohols **5** afforded the aldehydes **2** after an extractive work up without further purification.[6] The yields of the reduction and oxidation steps as well the cathepsin K activity of P^1 aldehyde analogues **2** are shown in Table 1.

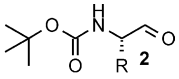
* Corresponding author. Tel.: +1-919-483-2520; e-mail: robert.b.mcfadyen@gsk.com
† Current address: Regeneron Pharmaceuticals Inc., 777 Old Saw Mill River Road, Tarrytown, NY 10591, USA.
* Deceased.

0960-894X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2003.09.088



Scheme 1. (a) $(Boc)_2O$, NaOH, dioxane, H_2O, rt; (b) *i*PrOCOCl, NEt_3, THF, −10 °C; $NaBH_4$, H_2O, 0 °C; (c) pyridine·SO_3, NEt_3, DMSO, CH_2Cl_2, −10 °C.

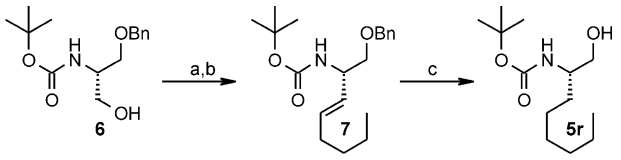
Table 1. Cathepsin K inhibition and synthesis yields



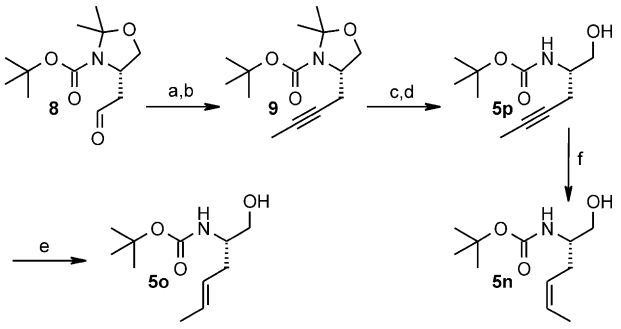
#	R	5% yd	2% yd	Cat K[a] IC_{50} nM
2a	H	na	62	5800
2b	Me	na	87	200
2c	Et	na	90	58
2d	*i*Pr	na	93	160
2e	*t*Bu	na	99	52,000
2f	Pr	77	92	38
2g	CH(Me)Et (S)	na	97	250
2h	CH(Me)Et (R)	75	87	130
2i	$CH_2C(CH_3)=CH_2$	45	92	79
2j	CH_2*i*Pr	na	98	31
2k	CH_2*t*Bu	97	87	250
2l	Bu	65	99	51
2m	$(CH_2)_3CF_3$	63	35	87
2n	Z $CH_2CH=CHCH_3$	na	83	41
2o	E $CH_2CH=CHCH_3$	na	71	150
2p	$CH_2C\equiv CCH_3$	na	79	410
2q	*n*-Pentyl	69	93	110
2r	*n*-Hexyl	na	69	23
2s	CH_2OEt	40	70	660
2t	CH_2SEt	66	72	83
2u	$(CH_2)_2OMe$	22	46	680
2v	$(CH_2)_2SMe$	74	99	78
2w	Ph	na	88	51*
2x	Cyclohexyl	66	99	26*
2y	Benzyl	na	99	110
2z	CH_2 cyclohexyl	na	96	36*
2aa	$(CH_2)_2Ph$	54	92	30*
2ab	$(CH_2)_2$cyclohexyl	na	99	26
2ac	$(CH_2)_3Ph$	62	82	15
2ad	$(CH_2)_3$cyclohexyl	na	87	16
2ae	CH_2OCH_2Ph	na	88	38*
2af	CH_2SCH_2Ph	67	99	40
2ag	CH_2NHCO_2Me	na	23	370*
2ah	$(CH_2)_4NHCO_2Me$	43	60	15*
2ai	$(CH_2)_2N(Me)COCF_3$	na	79	98
2aj	$(CH_2)_3N(Me)COCF_3$	na	92	29
2ak	$(CH_2)_4N(H)COCF_3$	na	99	41
2al	$(CH_2)_4N(Me)COCF_3$	na	41	50

[a] Inhibition of recombinant human cathepsin K activity in a fluorescence assay using 10 μM Cbz-Phe-Arg-AMC as substrate in 100 mM NaOAc, 10 mM DTT, 120 mM NaCl, pH = 5.5. The IC_{50} values are the mean of two or three inhibition assays, individual data points in each experiment were within a 3-fold range of each other. Entries marked with * represent a *n* of one.

The synthesis of alcohol **5r** is shown in Scheme 2. The alcohol **6** was oxidized to an intermediate aldehyde which was homologated via Wittig reaction to produce the alkene **7**.[7] Hydrogenation with Pearlman's catalyst reduced the olefin and cleaved the benzyl ether to produce the penultimate alcohol **5r**.[8]



Scheme 2. (a) Pyridine·SO_3, NEt_3, DMSO, CH_2Cl_2, −10 °C, (82%); (b) *n*BuLi, $CH_3(CH_2)_4PPh_3Br$, THF, −78 °C; aldehyde, −78 °C to rt, (19%); (c) H_2, $Pd(OH)_2$, MeOH, (43%).

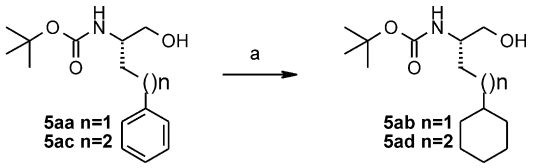


Scheme 3. (a) CBr_4, PPh_3, NEt_3, CH_2Cl_2, (80%); (b) *n*BuLi, THF, −78 °C; MeI, −78 °C to rt, (89%); (c) TFA, MeOH, 0 °C; (d) Boc_2O, Na_2CO_3, dioxane, water, (37%); (e) Li, NH_3, EtOH, (81%); (f) H_2, Lindlar's catalyst, EtOH, (83%).

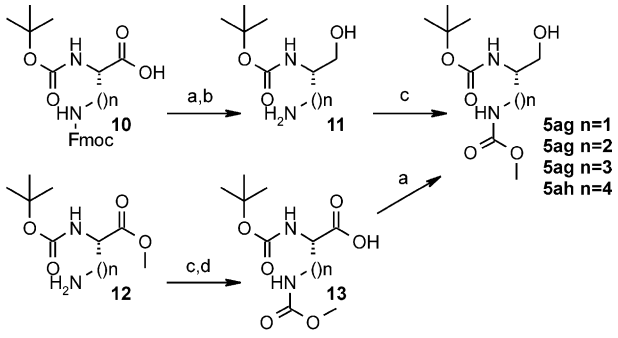
The synthesis of alcohols **5n–p** is shown in Scheme 3. Employing the Corey–Fuchs protocol, aldehyde **8** was converted to a dibromoolefin then treated with butyl lithium. The resulting acetylene anion was trapped with methyl iodine to yield the alkyne **9**.[9] Acidic hydrolysis of acetylene **9** cleaved the acetonide with the concomitant loss of the *tert*-butyl carbamate. The resulting amine was reprotected to provide the alcohol **5p**.[10] Metal/ammonia reduction of alkyne **5p** yielded the *trans* alkene **5o**.[11] Lindlar hydrogenation of the same alkyne **5p** gave the *cis* olefin **5n**.[12]

A series of compounds that contain P^1 side chains that feature a tethered phenyl or cyclohexyl ring was also made. As depicted in Scheme 4, alcohols **5ab** and **5ad** were prepared by reduction of the corresponding aryl analogues **5aa** and **5ac**, respectively.[13]

A series of aldehydes incorporating nitrogen functionality (amide/carbamate groups) into the P^1 position was also investigated. The synthesis of alcohols **5ag–ah** is shown in Scheme 5. The orthogonally protected diaminoacids **10** were reduced to the alcohols. Subsequent Fmoc removal afforded the amines **11** which were reacted with methyl chloroformate to provide methylcarbamates **5ag** ($n = 1$–3).[14] The lysine derivative **5ah** ($n = 4$) was synthesized from the methyl ester **12**. Conversion of the free amine to the methyl carbamate and subsequent hydrolysis of the ester afforded the acid **13** ($n = 4$). Then, reduction of the acid **13** as previously described gave the alcohol **5ah** ($n = 4$). Oxidation of alcohol **5ag** ($n = 1$) as in Scheme 1 afforded the aldehyde **2ag** ($n = 1$). However, the extended amine derivatives ($n = 2$–3) cyclized to the five- and six-membered hemiaminals upon oxidation. Alcohol **5ah** ($n = 4$) gave a mixture of aldehyde **2ah** and seven-membered hemiaminals upon oxidation.



Scheme 4. (a) H_2, $RhCl_3{\cdot}H_2O$, Aliquat® 336, $ClCH_2CH_2Cl$, H_2O, (99%).

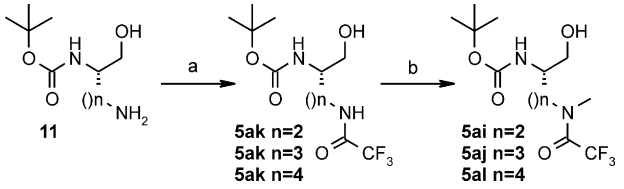


Scheme 5. (a) *i*PrOCOCl, NEt_3, THF, −10 °C; $NaBH_4$, H_2O, 0 °C, (30%); (b) NEt_2, THF; (c) MeOCOCl, iPr_2NEt, THF, (43%); (d) LiOH, THF, H_2O.

With the aim of preventing hemiaminal formation, a series of methylated trifluoroacetamides was synthesized. As shown in Scheme 6, the amines **11** ($n=2$–4) were converted into the trifluoroacetamides **5ak** ($n=2$–4).[15] The trifluoroacetamides were then deprotonated with sodium hydride, followed by alkylation with methyl iodide to give the tertiary amides **5ai**, **5aj**, and **5al**.[16] Moffat oxidation of the alcohols **5** as in Scheme 1 afforded the aldehydes **2**.

With an X-ray crystal structure of cathepsin K unavailable at the start of this work, a traditional structure–activity relationship study was employed varying steric and electronic properties of the P^1 moiety. As shown in Table 1, an *n*-alkyl side chain of 1–6 carbons leads to an increase in cathepsin K activity over the glycine derivative **2a**. The alanine derivative **2b** is roughly 30-fold more potent than the glycine analogue **2a**. The longer linear P^1 side chain aldehydes like **2c**, **2f**, and **2l** have equivalent potency and are slightly more potent than the alanine derivative **2b**. Additional substitution of the α-carbon of the side chain is detrimental to enzyme inhibitory activity (compare **2c** to **2d** or **2f** to **2g** and **2h**) with the fully substituted **2e** being roughly 900-fold less potent than **2c**. Steric branching at the β-carbon of the P^1 side chain appears to be tolerated (compare **2f** and **2j**) although potency does drop off in the fully substituted case **2k**. Incorporation of π-bonds into the P^1 side chain gave mixed results. The *Z*-alkene **2n** was experimentally indistinguishable from the norleucine analogue **2l**, while the *E*-alkene **2o** is 4-fold less active. The alkyne **2p** is 10-fold less potent than *Z*-alkene **2n**.

Hoping to pick up favorable π–π stacking interactions with potential aryl groups in the S^1 subsite, a phenyl ring was attached to the amino acid backbone with varying tether lengths. These P^1 side chains were found



Scheme 6. (a) CF_3CO_2Et, NEt_3, MeOH; (b) NaH, THF; MeI, 0 °C to rt.

to have good potency. Cathepsin K inhibitory activity of phenylglycinal **2w** is 51 nM. The phenylalaninal analogue **2y** is the least active member of this series ($IC_{50}=110$ nM). Activity then increases with the phenethyl derivative **2aa** ($IC_{50}=30$ nM) and the phenpropyl analogue **2ac** ($IC_{50}=15$ nM). The cyclohexyl derivatives are also potent inhibitors of cathepsin K. The cyclohexyl glycine derivative **2x** is roughly as potent as the propyl analogue **2f**. It is interesting to note that the branching at the beta carbon does not lead to the same decrease in activity observed in **2d**. The methyl-, ethyl- and propylcyclohexyl analogues **2z**, **2ab**, and **2ad**, respectively, are potent inhibitors with IC_{50}s of 36, 26, and 16 nM. These results appear to show that no favorable π–π interactions have been picked up by presence of an aryl group in these analogues.

The electronics of the S^1 subsite were probed by substituting heteroatoms into the P^1 group. Incorporation of polar heteroatoms into the side chain close to the amino acid backbone leads to decreased cathepsin K inhibitory activity. Ether analogues **2s** and **2u** are over 10-fold less potent than the norleucinal **2l**. Amine derivative **2ag** ($IC_{50}=370$ nM) is also less potent than **2l** ($IC_{50}=51$ nM). Sulfur is an exception to this trend. With available d orbitals and more diffuse electron lone pairs, the 'soft' sulfur is closer to carbon than it is to oxygen or nitrogen. The thioethers **2t** and **2v** are equipotent with the butyl analogue **2l**. In contrast to P^1 side chains with 'hard' heteroatoms close to the amino acid backbone like **2ag**, the ornithine **2aj** and lysine derivatives **2ah**, **2ak**, and **2al** are among the more potent aldehydes surveyed. Thus, inhibitor solubility could be enhanced by these P^1 groups.

During our SAR investigations, an X-ray co-crystal structure of inhibitor **2x** bound to the active site of cathepsin K was obtained (Fig. 1). The structure lends some support to the observed SAR of these P^1 analogues. The S^1 binding site is more aptly described as a wall rather than pocket, since one half of the subsite is solvent exposed. It is created by residues ^{23}Gly, ^{24}Ser, ^{64}Gly, and ^{65}Gly. All of the polar backbone atoms that form the wall are hydrogen bonded which results in the subsite having a very hydrophobic character, explaining the preference for linear hydrophobic substituents in this position. Furthermore, P^1 side chains with 'hard' heteroatoms close to the amino acid backbone should have unfavorable interactions with this hydrophobic enzyme surface. However, the enzyme wall is not extremely large (~5 Å in length). Thus, heteroatoms attached by longer tethers should eventually become exposed to solvent. In accord with this, lysine and orni-

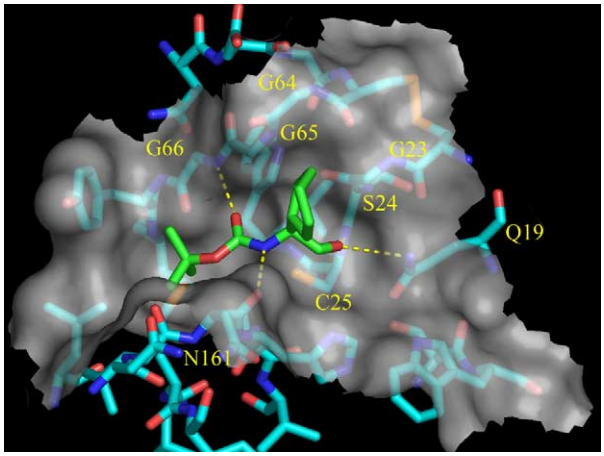


Figure 1. Active site of the X-ray co-crystal structure of compound **2x** bound to cathepsin K. The cathepsin K carbons are colored cyan and compound **2x** in green. The semi-transparent white surface represents the molecular surface while the hydrogen bonds are depicted by the dashed yellow lines. The coordinates have been deposited in the Brookhaven Protein Data Bank, accession number 1Q6K. This figure was generated with the program PyMOL.[17]

thine derivatives like **2ah** and **2aj** are well tolerated. The hydrophobic portions of their tethers are able to make favorable contacts with the S^1 wall, while their polar ends can reach above the wall and into a solvent-exposed region.

Other interactions also contribute to the binding of these inhibitors. The *t*-butyl group is bound in the deep S^2 pocket formed by residues ^{67}Tyr, ^{68}Met, ^{134}Ala, ^{163}Ala, and ^{209}Leu and hydrophobic groups would be preferred based on the 'greasy' nature of these residues. A covalent hemithioketal intermediate is formed by the aldehyde of the inhibitor and the active site cysteine (^{25}Cys) of the protein. The OH of the hemithioketal is stabilized by two hydrogen bonds to the backbone amide of ^{25}Cys and the sidechain of ^{19}Gln. Two other hydrogen bonds formed by the inhibitor occur in the peptide recognition site: the NH and carbonyl oxygen of the inhibitor carbamate form hydrogen bonds to the backbone carbonyl of ^{161}Asn and the backbone HN of ^{66}Gly, respectively.

A series of aldehyde inhibitors of cathepsin K with varied substituents at the P^1 position has been synthesized. Linear alkanes were some of the more potent analogues in the enzymatic assay. Tethered ring substituents provided similar potencies, but might be less attractive for further investigation because of synthetic complexity and increased size. Long alkyl chains with polar caps are well tolerated and could potentially add aqueous solubility. Information gained from these studies has proved to be useful in the design of other cathepsin K inhibitors. These inhibitors will be reported in due course.

References and notes

1. Yamashita, D. S.; Dodds, R. A. *Curr. Pharm. Des.* **2000**, *6*, 1.
2. Votta, B. J.; Levy, M. A.; Badger, A.; Bradbeer, J.; Doods, R. A.; James, I. E.; Thompson, S.; Bossard, M. J.; Carr, T.; Connor, J. R.; Tomaszek, T. A.; Szewczuk, L.; Drake, F. H.; Veber, D. F.; Gowen, M. *J. Bone Miner. Res.* **1997**, *12*, 1396.
3. Tsujinaka, T.; Kajiwara, Y.; Kambayashi, J.; Sakon, M.; Higuchi, N.; Tanaka, T.; Mori, T. *Biochem. Biophys. Res. Commun.* **1988**, *153*, 1201.
4. Tarbell, D. S.; Yamamoto, Y.; Pope, B. M. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 730.
5. Kubota, M.; Nagase, O.; Yajima, H. *Chem. Pharm. Bull.* **1981**, *29*, 1169.
6. Hamada, Y.; Shiabat, M.; Sugiura, T.; Kato, S.; Shioiri, T. *J. Org. Chem.* **1987**, *52*, 1252.
7. Hann, M. M.; Sammes, P. G.; Kennewell, P. D.; Taylor, J. B. *J. Chem. Soc., Perkin Trans. 1* **1982**, 307.
8. Oikawa, H.; Matsuda, I.; Kagawa, T.; Ichihara, A.; Kohmoto, K. *Tetrahedron* **1994**, *50*, 13347.
9. Corey, E. J.; Fuchs, P. L. *Tetrahedron Lett.* **1972**, *13*, 3769.
10. Garner, P.; Park, J. M. *J. Org. Chem.* **1987**, *52*, 2361.
11. De Medeiros, E. F.; Herbert, J. M.; Taylor, R. J. K. *J. Chem. Soc., Perkin Trans. 1* **1991**, 2725.
12. Ulery, H. E.; Richards, J. H. *J. Am. Chem. Soc.* **1964**, *86*, 3113.
13. Blum, J.; Amer, I.; Zoran, A.; Sasson, Y. *Tetrahedron Lett.* **1983**, *24*, 4139.
14. Islam, A. M.; Raphael, R. A. *J. Chem. Soc.* **1955**, 3151.
15. Joullie, M. M.; Day, A. R. *J. Am. Chem. Soc.* **1954**, *76*, 2990.
16. Norman, M. H.; Rigdon, G. C.; Hall, W. R.; Navas, F., III *J. Med. Chem.* **1996**, *39*, 1172.
17. DeLano, W. L. *The PyMOL Molecular Graphics System*. DeLano Scientific LLC: San Carlos, CA, USA (http://www.pymol.org).