Bioorganic & Medicinal Chemistry Letters 21 (2011) 932–935

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

1*H*-Imidazo[4,5-*c*]pyridine-4-carbonitrile as cathepsin S inhibitors: Separation of desired cellular activity from undesired tissue accumulation through optimization of basic nitrogen pk_a

Wullie Arbuckle, Mark Baugh, Simone Belshaw, D. Jonathan Bennett, John Bruin, Jiaqiang Cai *, Kenneth S. Cameron, Chris Claxton, Maureen Dempster, Kathryn Everett, Xavier Fradera, William Hamilton, Philip S. Jones, Emma Kinghorn, Clive Long, Iain Martin, John Robinson, Paul Westwood

Merck Research laboratories, MSD, Lanarkshire ML1 5SH, United Kingdom

ARTICLE INFO

Article history:
Received 14 October 2010
Revised 14 December 2010
Accepted 15 December 2010
Available online 19 December 2010

Keywords:
Cathepsin s inhibitor
Lysosomotropism
Tissue sequestration
pk_a

ABSTRACT

Based on the theoretical understanding of the in vivo lysosomotropism, by adjusting the pk_a of basic nitrogen containing cathepsin S inhibitors, a set of compounds with pk_a 6–8 were identified to have excellent cell based Lip10 activity, yet avoiding undesired sequestration in spleen.

There are eleven members of the cathepsin family cysteine proteases identified in the human genome (cathepsins B, C, H, F, K, L, O, S, V, W and X).[1] In contrast to most of the cathepsins which are only function in an acidic pH environment, cathepsin S is active in both acidic and neutral conditions, and it has a dual functionality, that is, extracellular matrix degradation and intracellular invariant chain processing.[2] As we are interested in cathepsin S as a target to treat autoimmune disorders by modulating cathepsin S activity inside the low pH lysosomes, good intracellular activity measured by the accumulation of the natural substrate of cathepsin S, Lip10 in human JY cell[3] is critically important. In our previous publications[3–5] we have demonstrated the importance of basic nitrogen for the cell based Lip10 activity through so-called lysosomotropism[6,7] as shown by our previous examples in Figure 1. Although these basic nitrogen containing compounds, such as compounds **2**–**4**, have good activity in the cell based Lip10 assays, they were found to have very high plasma clearance and high distribution volume. These compounds are highly sequestered in spleen, liver and lung with spleen level as much as 50 times higher than in plasma for compound **3**. In response to concerns over safety relating to tissue accumulation, we made a strategic decision to rule out these highly sequestered compounds for further development.

The lysosomotropism can be quantitatively described[6–8] appendix by Eq. 1 for monobasic compounds where *F* is the ratio of total compound concentration (including all protein, lipid and tissue bound compound treated as homogeneous and other unbound compound in solution) between inside the lysosome and its surrounding medium; pH_M is the pH of the surrounding medium (taken as 7.4) and pH_L is pH inside the lysosome (taken as 4.8);[6,7] ff_M is the free fraction in the medium and ff_L is the free

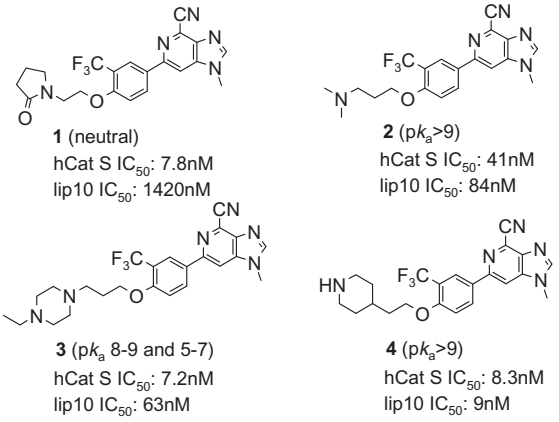


Figure 1.

* Corresponding author. Tel.: +44 01419467424; fax: +44 01698736187.
E-mail address: jiaqiang.cai@yahoo.co.uk (J. Cai).

0960-894X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2010.12.065

fraction in the lysosome. When only unbound compound concentration (both protonated and unprotonated) is considered, the compound concentration ratio (f) between lysosome and the medium can be simplified as Eq. 2. When f is plotted against pk_a, a titration curve as shown by Figure 2 can be obtained. Based on this equation, the theoretical maximum concentration ratio for a monobasic nitrogen containing compound in solution is ca. 400. As lysosomes occupy only a very small portion of the tissue (0.004–0.0068 cm^3/g estimated for liver tissue[6]), the maximum lysosomotropism should only account for less than three-fold extra compound level (i.e., $400 \times 0.0068 = 2.76$). The observed high sequestration ratio of 25–50 is probably due to other tissue bindings, most likely phospholipid binding in the described cases. As we are targeting lysosome cathepsin S, the lysosomotropism should be considered as a positive factor. This lysosomotropism also seems essential for achieving desired cellular functional activity against Lip10 degradation. It therefore becomes a formidable challenge to obtain a compound which possesses desired cellular Lip10 activity without undesired organ accumulation.

$$F = \frac{10^{pH_M}}{10^{pH_L}} \times \frac{10^{pK_a} + 10^{pH_L}}{10^{pK_a} + 10^{pH_M}} \times \frac{ff_M}{ff_L} \quad (1)$$

$$f = \frac{10^{pH_M}}{10^{pH_L}} \times \frac{10^{pK_a} + 10^{pH_L}}{10^{pK_a} + 10^{pH_M}} = \frac{10^{7.4}}{10^{4.8}} \times \frac{10^{pK_a} + 10^{4.8}}{10^{pK_a} + 10^{7.4}} \quad (2)$$

In this paper we will report our effort in designing such compounds by selecting a narrow pk_a window from the theoretical understanding of the lysosomotropism.

Expansion of the Figure 2 graph between pk_a 5.5–8.5 is shown in Table 1; compounds with pk_a 6–8 should have ~16 to ~300-fold (f) higher level in lysosome than other neutral compartments. In theory it should be possible to design compounds with pk_a in this range to have reasonably good cellular Lip10 activity. Compounds with pk_a in the lower end of 6–8 should be expected to have relatively low phospholipid binding as judged by their in vivo distribution volume (V_{ss}) and should then have low sequestration level in organs such as spleen, liver and lung. On top of what was discussed above, these low pk_a (6–8) compounds will mostly be positively charged in the low pH (4.5–5) lysosomal environment with a low log D and should have lower binding affinities towards proteins and lipids than the corresponding uncharged neutral spices, and as such a higher free fraction (ff_L) level for target engagement.

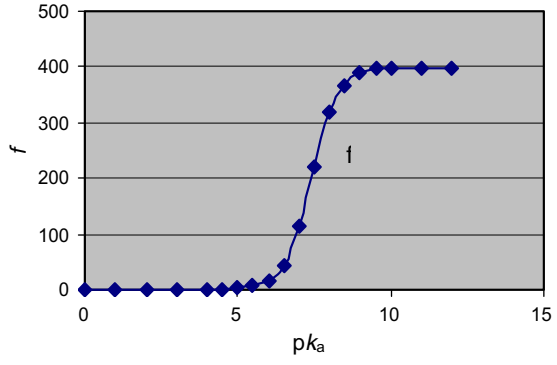


Figure 2. f plotted against pk_a.

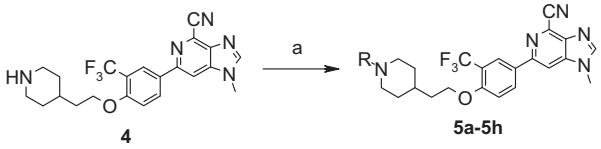
Table 1
relationship between f and pk_a

pk_a	5.5	6	6.5	7	7.5	8	8.5
f	6	16.2	45	113	221	318	368

For three reasons compound **4** was selected for exploring the pk_a and in vitro Lip10 relationship. The first is that compound **4** is the most active from the series and it has excellent microsome stability across all species examined (SD rat, C57 mouse and human). Secondly, structural biology work indicated that an attachment to the piperidine nitrogen is exposed to solvent outside the S2 pocket and should have relatively little impact on the human cathepsin S inhibitory activity, facilitating future data interpretation. Thirdly, a wide range of pk_a for the piperidine nitrogen could be easily achievable by judicious choice of substituents with different electronic properties.

Compounds for these studies were synthesized according to Scheme 1 by simple alkylation or acylation of the piperidine nitrogen of compound **4**. Results of these studies are shown in Table 2.

The two most basic compounds **4** and **5a** share a similar activity against purified human cathepsin S enzyme and also in cell based Lip10 assays. Attaching a *N*,*N*-dimethylaminocarbonylmethyl group to the piperidine nitrogen (**5b**) reduces pk_a to 7.5. Compound **5b** is not only highly active against purified human cathepsin S, it also has high activity in the cellular assay with a Lip10 IC_{50} of 22 nM. In order to reduce the pk_a further to the region of 6–7, by using our in-house pk_a calculation tools,[10] it was found that a heteroarylmethyl group could affect the piperidine nitrogen pk_a significantly. By arranging the heteroatoms O, N and S inside the five-membered heteroaryl ring, it looked possible to achieve the required pk_a range based on the calculations. A selected set of four compounds **5c**–**5f** were synthesized and the results are shown in



Scheme 1. Reagents and conditions: (a) RCl, DIPEA, NMP, rt, 16 h.

Table 2
Inhibitory activity against both purified cathepsin S enzyme[a] and natural substrate in cell based assays[a] and measured pk_a

Compds	R–N	IC_{50}[b] (nM)		Lip10 (nM)	pk_a[c]
		hCatS	hCatK		
2		41	776	84	>9 (c)
3		7.2	331	63	8–9 (c)
4	HN	8.3	178	9	>9 (c)
5a	MeN	7.9	na	10	>9 (c)
5b		8.3	407	22	7.5 (m)
5c		3.3	288	12	7.9 (m)
5d		7.9	708	39	6.3 (m)
		1.6[d]			
5e		14.1	813	123	5.6 (m)
5f		18.6	1148	294	5.3 (m)
5g		13	1905	235	Neutral
5h		27	2344	353	Neutral

[a] For assay conditions see Ref. 3.
[b] All compounds have IC_{50} of >10 μM for human cathepsins B, L and V.
[c] pK_a of conjugated base, values were determined using a Sirius GLpKa potentiometric titrator[9]; m = measured; c = calculated.[10]
[d] IC_{50} for mouse cathepsin s.

Table 3
Total compound level in plasma and spleen measured at 2 h time point after 30 mg/kg po dosing (C57Bl/6 mice)

Compds	Total plasma (nM)	Total spleen (nM)	Spleen/plasma ratio
2	94 (4 h)	2433 (4 h)	26
3	90	4570	51
5b	1031	11,738	11.4
5c	3400	19,567	5.8
5d	6610	12,502	1.9
5e	6369	18,165	2.9
5f	5664	4087	0.7

Compounds were dosed as HCl salt. 18–25 g C57Bl/6 mice were dosed by gavage with 30 mg/kg compound (0.5% gelatin, 5% mannitol in water). Blood and spleen were collected at 2 h time point after dosing and compound level measured.

Table 2. The pk_a of these compounds were measured and the values are in the expected range from 5 to 8. It is interesting to see that there is a relationship between pk_a and cathepsin S inhibitory potency, the higher the pk_a, the higher the cathepsin S inhibitory potency (or lower IC_{50}). There seems to be very little difference between low pk_a compound (**5f**) and the neutral analogs **5g** and **5h**[11] in terms of both biochemical and cellular potency. Although the cellular activities (Lip10) share a parallel relationship with pk_a as IC_{50}s do, the gap between Lip10 and IC_{50} becomes bigger when pk_a is lower, for example, Lip10/IC_{50} ratio is nearly 16 for compound **5f** and it is only 3.6 for the higher pk_a compound **5c**. These differences, although relatively small, could be easily explained by the accumulation factor *f*. Compounds **5b–5f**, together with compounds **2** and **3** were assessed in an in vivo spleen sequestration assay using C57Bl/6 mice. These sequestration results are shown in Table 3. As the data shows, all the compounds with pk_a lower than 7 do not have significant sequestration in the mouse spleen. For compounds with pk_a 7–8, such as **5b** and **5c**, the ratio between spleen and plasma is also quite low (<12) as compared with the higher pk_a analogs **2** and **3**.

Compounds **5b** and **5d** were further evaluated for full pharmacokinetic properties in male C57 mice, Sprague Dawley rats and Cynomolgus monkeys. The results are shown in Table 4. Although their rat and monkey oral bioavailabilities are less than 10%, both compounds showed excellent mouse PK profiles with oral bioavailabilities at 62% and 112% for **5b** and **5d**, respectively.

Considering compound **5d** is also highly active against mouse cathepsin S (1.6 nM, Table 2), it was further assessed for its in vivo cathepsin S inhibition and its efficacy in collagen induced arthritis model (CIA) dosed in a therapeutically relevant setting. The full results of these in vivo studies will be published in a separate account.

In summary, based on the theoretical understanding of the in vivo lysosomotropism and by adjusting the pk_a of basic nitrogen containing cathepsin S inhibitors, a set of compounds with pk_a 6–8 were identified to have excellent cell based Lip10 activity, yet avoiding undesired accumulation in spleen.

Table 4
Pharmacokinetic parameters (iv 2 mg/kg; po 10 mg/kg, gelatin manitol)

Compds		Cl (mL/min/kg)	V_{ss} (L/kg)	C_{max} (ng/mL)	*F* (%)
5b	C57	79.4	6.17	295.5	62
	SD	93.5	9.8	13.37	8.93
5d	C57	13.9	1.48	3160	112
	Cyno	10	0.6	234	9
	SD	73.8	2.43	9.64	3.6

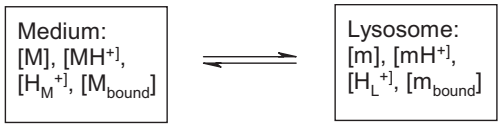


Figure 3.

Appendix

Consider lysosome and its surrounding medium as a two compartment system as shown in Figure 3, where:

[M]: free unprotonated compound concentration in the medium,
[MH^+]: unbound protonated compound concentration in the medium,
[M_{bound}]: all other bound compound concentration in the medium considered as homogeneous,
[H_M^+]: medium proton concentration,
[m]: free unprotonated compound concentration in the lysosome,
[mH^+]: unbound protonated compound concentration in the lysosome,
[m_{bound}]: all other bound compound concentration in the lysosome considered as homogeneous, and
[H_L^+]: lysosome proton concentration.

As lysosmal membrane should prevent all species moving in or out lysosomes except free uncharged molecules and at the steady-state, this uncharged molecule will have equal concentration across the two compartments, that is, $[M] = [m]$

Total medium concentration Mt:

$$[\mathrm{Mt}] = [\mathrm{M}] + [\mathrm{MH}^+] + [\mathrm{M}_{\mathrm{bound}}] = [\mathrm{M}] + [\mathrm{MH}^+] + [\mathrm{Mt}](1 - [ff_{\mathrm{M}}])$$

$$[\mathrm{Mt}] = \frac{[\mathrm{M}] + [\mathrm{MH}^+]}{[ff_{\mathrm{M}}]}$$

Total lysosome concentration [mt]:

$$[\mathrm{mt}] = \frac{[\mathrm{m}] + [\mathrm{mH}^+]}{[ff_{\mathrm{L}}]}$$

$$F = \frac{[\mathrm{mt}]}{[\mathrm{Mt}]} = \frac{[\mathrm{m}] + [\mathrm{mH}^+]}{[\mathrm{M}] + [\mathrm{MH}^+]} \times \frac{[ff_{\mathrm{M}}]}{[ff_{\mathrm{L}}]}$$

With $[\mathrm{MH}^+] = \frac{[\mathrm{M}][\mathrm{H}_{\mathrm{M}}^+]}{K_a}$, $[\mathrm{mH}^+] = \frac{[\mathrm{m}][\mathrm{H}_{\mathrm{L}}^+]}{K_a}$, and $[\mathrm{M}] = [\mathrm{m}]$ at the equilibrium, then:

$$F = \frac{[\mathrm{mt}]}{[\mathrm{Mt}]} = \frac{K_a + [\mathrm{H}_{\mathrm{L}}^+]}{K_a + [\mathrm{H}_{\mathrm{M}}^+]} \times \frac{[ff_{\mathrm{M}}]}{[ff_{\mathrm{L}}]}$$

To replace K_a with more widely used pK_a and $[H^+]$ with pH with $K_a = \frac{1}{10^{pK_a}}$, $[\mathrm{H}_{\mathrm{L}}^+] = \frac{1}{10^{\mathrm{pH_L}}}$ and $[\mathrm{H}_{\mathrm{M}}^+] = \frac{1}{10^{\mathrm{pH_M}}}$ then

$$F = \frac{10^{\mathrm{pH_M}}}{10^{\mathrm{pH_L}}} \times \frac{10^{pK_a} + 10^{\mathrm{pH_L}}}{10^{pK_a} + 10^{\mathrm{pH_M}}} \times \frac{ff_{\mathrm{M}}}{ff_{\mathrm{L}}}$$

and

$$f = \frac{[\mathrm{mt}][ff_{\mathrm{L}}]}{[\mathrm{Mt}][ff_{\mathrm{M}}]} = \frac{10^{\mathrm{pH_M}}}{10^{\mathrm{pH_L}}} \times \frac{10^{pK_a} + 10^{\mathrm{pH_L}}}{10^{pK_a} + 10^{\mathrm{pH_M}}}$$

Based on this quantitative model, high pk_a compounds with high phospholipid binding could be highly sequestered in high phospholipids containing organs, such as spleen and lung where free fraction level (*ff*) will be low.

References and notes

1. Gupta, S.; Singh, R. K.; Dastidar, S.; Ray, A. *Expert Opin. Ther. Targets* **2008**, *12*, 291.
2. Leroy, V.; Thurairatnam, S. *Expert Opin. Ther. Pat.* **2004**, *14*, 301.

3. Cai, J.; Baugh, M.; Black, D.; Long, C.; Bennett, J.; Dempster, M.; Fradera, X.; Gillespie, J.; Andrews, F.; Boucharens, S.; Bruin, J.; Cameron, K. S.; Cumming, I.; Hamilton, W.; Jones, P. S.; Kaptein, A.; Kinghorn, E.; Maidment, M.; Martin, I.; Mitchell, A.; Rankovic, Z.; Robinson, J.; Scullion, P.; Uitdehaag, J. C. M.; Vink, P.; Westwood, P.; van Zeeland, M.; van Berkom, L.; Bastiani, M.; Meulemans, T. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4350.
4. Cai, J.; Bennett, J.; Rankovic, Z.; Dempster, M.; Fradera, X.; Gillespie, J.; Cumming, I.; Finlay, W.; Baugh, M.; Boucharens, S.; Bruin, J.; Cameron, K.; Hamilton, W.; Kerr, J.; Kinghorn, E.; McGarry, G.; Robinson, J.; Scullion, P.; Uitdehaag, J. C. M.; van Zeeland, M.; Potin, D.; Saniere, L.; Fouquet, A.; Chevallier, F.; Deronzier, H.; Dorleans, C.; Nicolai, E. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4447.
5. Cai, J.; Fradera, X.; van Zeeland, M.; Dempster, M.; Cameron, K.; Bennett, J.; Robinson, J.; Popplestone, L.; Baugh, M.; Westwood, P.; Bruin, J.; Hamilton, W.; Kinghorn, E.; Long, C.; Uitdehaag, J. C. M. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4507.
6. de Duve, C.; de Barsy, T.; Poole, B.; Trouet, A.; Tulkens, P.; van Hoof, F. *Biochem. Pharmacol.* **1974**, *23*, 2495.
7. MacIntyre, A.; Cutler, D. J. *Biopharm. Drug Dispos.* **1988**, *9*, 513.
8. Black, W. C.; Percival, M. D. *ChemBioChem* **2006**, *7*, 1525.
9. Takacs-Novac, K.; Box, K. J.; Avdeef, A. *Int. J. Pharm.* **1997**, *151*, 235.
10. ACD PhysChem Batch, version 4.76; Advanced Chemistry Development, Toronto, ON, Canada.
11. 1*H*-imidazo[4,5-*c*]pyridine-4-carbonitrile core has a calculated pk_a of 1.8. Compounds with this pk_a should be unprotonated inside lysosomes, it is then regarded as neutral.