Identification of Plasmepsin Inhibitors as Selective Antimalarial Agents using Ligand Based Drug Design

Paul B. McKay, Martin B. Peters, Giorgio Carta, Christopher T. Flood, Enda Dempsey, Angus Bell, Colin Berry, David G. Lloyd, Darren Fayne

PII: S0960-894X(11)00462-8
DOI: 10.1016/j.bmcl.2011.04.015
Reference: BMCL 17367

To appear in: *Bioorganic & Medicinal Chemistry Letters*

Received Date: 2 March 2011
Revised Date: 1 April 2011
Accepted Date: 3 April 2011

Please cite this article as: McKay, P.B., Peters, M.B., Carta, G., Flood, C.T., Dempsey, E., Bell, A., Berry, C., Lloyd, D.G., Fayne, D., Identification of Plasmepsin Inhibitors as Selective Antimalarial Agents using Ligand Based Drug Design, *Bioorganic & Medicinal Chemistry Letters* (2011), doi: 10.1016/j.bmcl.2011.04.015

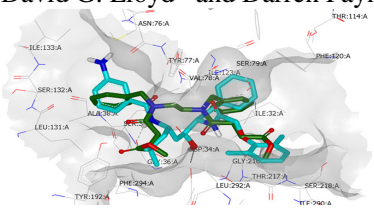
Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.

Identification of Plasmepsin Inhibitors as Selective Antimalarial Agents using Ligand Based Drug Design

Leave this area blank for abstract info.

Paul B McKay, Martin B. Peters, Giorgio Carta, Christopher T. Flood, Enda Dempsey, Angus Bell, Colin Berry, David G. Lloyd* and Darren Fayne



Bioorganic & Medicinal Chemistry Letters
journal homepage: www.elsevier.com

Identification of Plasmepsin Inhibitors as Selective Antimalarial Agents using Ligand Based Drug Design

Paul B McKay[a], Martin B. Peters[a], Giorgio Carta[a], Christopher T. Flood[a], Enda Dempsey[b], Angus Bell[b], Colin Berry[c], David G. Lloyd*[a] and Darren Fayne[a]

[a]*Molecular Design Group, School of Biochemistry and Immunology, Trinity College Dublin, Dublin 2, Ireland*
[b]*Department of Microbiology, School of Genetics and Microbiology, Moyne Institute of Preventive Medicine, Trinity College Dublin, Dublin 2, Ireland*
[c]*Cardiff School of Biosciences, Cardiff University, Park Place, Cardiff, CF10 3AT, UK*

ARTICLE INFO

Article history:
Received
Revised
Accepted
Available online

Keywords:
Ligand based drug design
Catalyst
Plasmepsin Inhibitors
Malaria
Aspartyl proteases inhibitor

ABSTRACT

We describe the application of Ligand Based Virtual Screening technologies towards the discovery of novel Plasmepsin (PM) inhibitors, a family of malarial parasitic aspartyl proteases. Pharmacophore queries were used to screen vendor libraries in search of active and selective compounds. The virtual hits were biologically assessed for activity and selectivity using whole cell Plasmodium falciparum parasites and on target in PM II, PM IV and the closely related human homologue, Cathepsin D assays. Here we report the virtual screening highlights, structures of the hits and their demonstrated biological activity.

There still remains a significant unmet need for drugs which combat the effect of malaria on approx. 500 million people each year.[1] The disease is caused by parasites from the genus *Plasmodium* and the Plasmepsins (PMs) are a family of aspartyl proteases (AP) found in each of the five species of *Plasmodium* that affect humans (*P. falciparum, P. vivax, P. malariae, P. ovale* and *P. knowlesi*). This study focuses on the identification of PM inhibitors of the most lethal species, *P. falciparum*.[1] Selectivity over the human equivalent is an important consideration when developing inhibitors of parasite enzymes. The PMs have a varied degree of sequence identity with the human aspartyl proteases, the most similar being the lysosomal enzyme Cathepsin D (CatD).[2] For this reason, CatD is commonly used as a marker for cross-inhibition. Comparison studies between PM II and CatD show a 35% sequence identity, and even higher structural identity has been observed at the active site when comparing crystal structures of PM II and CatD in complex with Pepstatin A.[3] In this regard it is therefore essential to develop PM inhibitors exhibiting specificity for the *Plasmodium* APs over the human homologue.

Four members of the ten member PM enzyme family are found within the acidic digestive food vacuole (DV) of *P. falciparum*, while the remaining six are located outside the DV. The intra-DV PMs are PM I, II, HAP (also known in earlier literature as PM III, where the Asp32 of the active site is replaced by a Histidine residue[4]) and IV. These four DV PMs have been shown to be involved in hemoglobin degradation which plays an essential role in malaria parasite development within infected red blood cells.[5-6] This degradative process provides a source of amino acids and is thought to help maintain intracellular osmolarity during rapid parasite growth.[7] High levels of sequence identity (~60% – 70%) are observed between PM I, II, IV, and HAP.[4] An analysis of the binding site region of PM II reveals that the binding site regions of PM I, IV, and HAP show 84%, 68% and 39% identity, respectively, while there is a lower degree of identity between the other six PMs.[8]

PM V, IX and X have been shown to be produced in intra-erythrocytic parasites[1,10-13] but to date, function has been elucidated only for PMV.[14] While these non-DV PMs may yet prove to be viable drug targets, the DV PMs have long been considered *possible* targets for the development of candidate drugs[9] due to their ability to initiate degradation of native hemoglobin. They also represent the greatest amount of available structural information, from high resolution co-crystals to large amounts of both whole cell parasite activity data and on-target DV PM inhibition data. Although these DV enzymes have shown some functional redundancy,[13] the discovery and description of novel inhibitors or chemical probes for these proteins is desirable and the design of inhibitors targeting them continues as the familial active site similarities make the prospect of a cross family inhibitor, targeting multiple PMs, possible.[16-19]

The goal of this research programme was to apply computational tools to guide the discovery of novel small

molecules targeting DV PMs, displaying selectivity over CatD and assess their ability to exert an anti-malarial effect.

Using the HypoGen module in the program Catalyst[20] (version 2004.11), a ligand based pharmacophore (Ph4) building and development method was applied in an attempt to derive a model with the ability to distinguish PM actives from inactives. The models generated were firstly validated by way of their ability to selectively retrieve actives from a database of known actives and decoys. Following this, the best Ph4 models were used as three-dimensional queries to perform a search of vendor compound databases in an effort to identify new PM inhibitors.

In order to create a Ph4 using HypoGen a training set of diverse known actives is normally required. The nature of the target under study required a slightly different approach with respect to selecting molecules to include in the training set. The PMs consist of a family of ten enzymes, of which all four DV PMs have been purified and a body of on-target inhibitor structure-activity data have been made available. This, coupled with the fact that selectivity over CatD is of crucial importance, meant that a refined method for active compound selection, as described below, was implemented. Using Pipeline Pilot,[21] the Wombat ligand database[22] was filtered for all known PM and CatD inhibitors and subsequently supplemented with additional inhibitors to include more recent data.[17,23-24]

This furnished a database of 194 active compounds with activity data (IC_{50} and/or K_i for the PMs and with additional information, where available for CatD) which was further filtered in order to match the HypoGen requirement for a diverse training set (i. dataset of greater than 16 compounds – this ensures statistical significance of the Ph4 model; ii. activity range of at least 4 orders of magnitude (OM) and each OM be represented by at the least three compounds; iii. Include most active and inactive ligands from the dataset; iv. Similar structures must differ in activity by 1 OM or only use the most active; very similar actives must be structurally distinct, otherwise use the most active). Concurrently the selectivity for the PMs over CatD was considered. In order to account for the selectivity issue, a metric was applied to rank the molecules based on their selectivity for the PMs over human CatD. The selectivity metric (SM) used is shown in Equation 1.

$$SM. = \sum_{i} \frac{(Act(i) - Act(CatD))}{N} \qquad \text{Equation 1}$$

Act(i) = pKi or pIC_{50} on PM I, II or IV

Act(CatD) = pKi or pIC_{50} on CatD

N = number of PMs that the compound was tested on (Max = 3)

Note: The equation is set to zero if no CatD activity value is available. The biological data needs to be consistent for each calculation, i.e. all pKi or all pIC_{50} values. If data for >1 PM are available, SM is the average difference between the PM(s) and CatD.

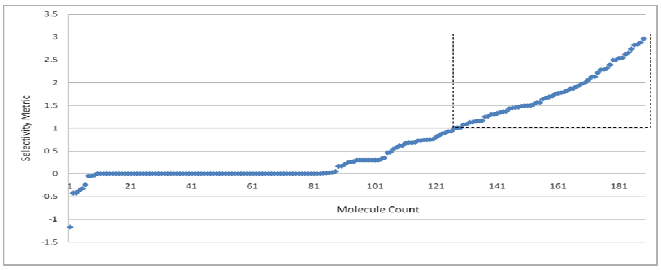


Figure 1: The selectivity score of each molecule targeting PM within the Wombat Database for the PMs vs CatD. The most selective molecules are shown in the hatched box

The resulting graph (Figure 1) allowed for our inhibitor database to be separated by selectivity for the PMs over CatD and revealed that many of the ligands described in the literature had no activity data for CatD. As this is a major consideration in this study, only those molecules with positive selectivity data for a PM over CatD were included in the dataset. The final training set consisted of 16 molecules that were chosen based on their selectivity, activity and relative structural diversity within a particular cluster.

A test set was generated so as to validate the activity predictive ability of the Ph4s generated. It was assembled using a bespoke Bayesian model implemented within Pipeline Pilot. The model was built using the PM filtered Wombat database and the following descriptors: ALogP, molecular weight, number of hydrogen bond donors (HBDs), number of hydrogen bond acceptors (HBAs), number of rotatable bonds, number of aromatic rings, number of rings and polar surface area. This model was then used to extract 960 decoy compounds with similar properties from the Maybridge compound catalogue[25]. The final test set consisted of 40 known actives and 960 decoy molecules.

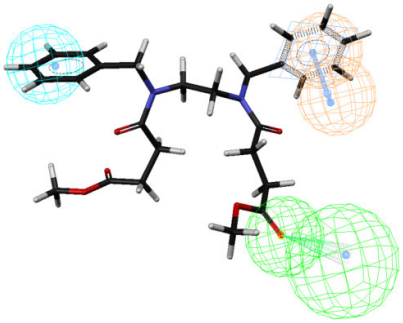
Multiple Ph4s were evaluated according to both Catalyst guidelines and by screening against the 1000 compound test set. ROC curves were generated for each output to evaluate the ability of the models to discriminate between active and decoy sets. The Ph4 yielding the optimal validation hit rate was then used to screen the Specs[26] and Asinex Platinum[27] commercial compound databases. Prior to screening, each database was pre-processed to remove those molecules with a molecular weight of less than 250 or greater than 750, followed by filtering using the same Bayesian model employed to generate the test set.

To assess the anti-malarial activity and selectivity profile of the compounds suggested from the Catalyst Ph4 screen, four biological assays were conducted. The first was an inhibition assay using *P. falciparum* cultures and the remaining three assays were on-target enzyme inhibition assays directed at PM II, IV and CatD. Details of these assays are provided in the Supporting Information.

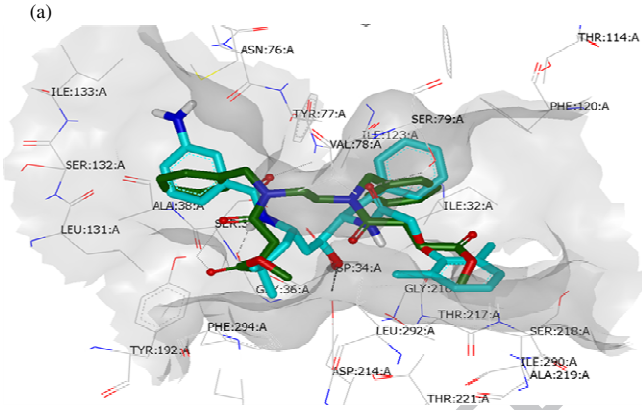
Two vendor databases, Asinex Platinum[27] and Specs[26], totalling 236,054 compounds were filtered as described above. Following pre-processing, the database was parsed to provide for ionisation states of between pH 4 – 6 and then imported into Catalyst for conformer generation (100 conformers for each compound were generated) using the catConf FAST algorithm. The output molecules were subsequently sorted according to their calculated fit values. From this collection of molecules, a visual inspection was undertaken to select a subset of 15 putative hits with tractable drug-like characteristics. The anti-parasitic potential of all 15 molecules was initially investigated in the *P. falciparum* 3D7 cell line, followed by progression of active compounds to an enzyme inhibition assay in which their ability to bind to PM II, PM IV and CatD was assessed. Details of the assays are given within the Supporting Information, with biological results shown in Table 1.

From the active subset, MDG422 is shown below as it aligns to the Ph4, in Figure 2A. This compound exhibited an IC_{50} of 5.24 µM in the parasite assay (after 72 h) and a K_i of 7 µM for both PM II and PM IV in the enzyme inhibition assay. MDG422 was docked using FRED[28] into the active site of PMII (PDB ID: 1LEE) with results shown in Figure 2B. The docking predicts

that key subsite-pockets of the active site are occupied by the ligand. Of particular note is the fact that this compound does not contain a hydrogen bond donor (HBD), analogous to the traditional hydroxyl-like moiety which typically coordinates with the Asp catalytic dyad and is contained within the vast majority of known AP inhibitors. This suggests lead optimisation potential for this scaffold and the possibility to evolve molecular SAR distant from a central hydroxyl-group to enable the design of enhanced specificity and potency which is independent of this common interaction inherent within most AP inhibitors.

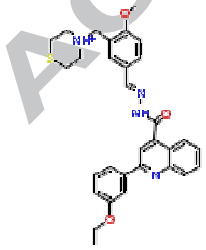
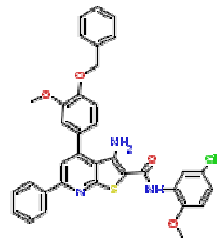
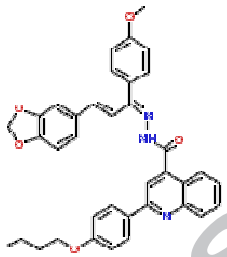
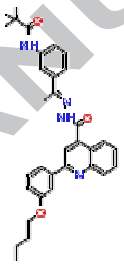
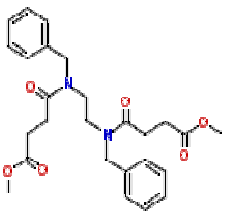
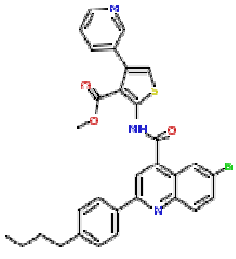


(a)



(b)

Figure 2. (A) MDG422 aligned to the Ph4 generated by Catalyst. Feature keys: Green = HBA, Orange: HYD|ARO and Cyan = HYD. (B) MDG422 docked (FRED) with PM II (PDB ID:1LEE). MDG422 shown in green and the reference ligand (R36) is shown in cyan.

	Parasite Assay	PM II		PM IV		Cat D
MDG ID	**72hrs IC_{50} [µM]**	**IC_{50}**	**K_i**	**IC_{50}**	**K_i**	**% Inhibition at 5 µM**
 MDG 418	40 ± 11	*	*	*	*	*
 MDG 419	2.3 ± 1	2.3 ± 0.6	1.6 ± 0.4	6.2 ± 3	5.4 ± 2.5	12
 MDG 420	24 ± 6	*	*	*	*	*
 MDG 421	3.7 ± 1.7	11.5 ± 2	8 ± 1.4	15 ± 3.5	13 ± 3	12
 MDG 422	5.24 ± 2.2	10 ± 2.5	7 ± 2	9 ± 3.5	7 ± 3	35
 MDG 424	1.24 ± 0.5	*	*	*	*	*

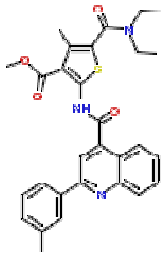
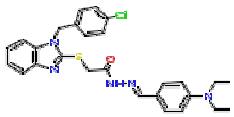
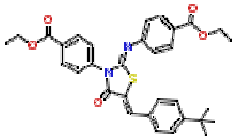
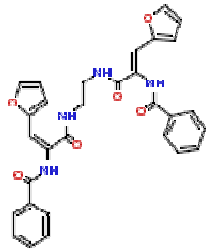
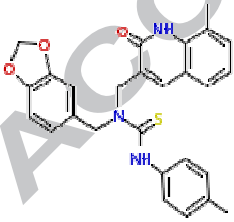
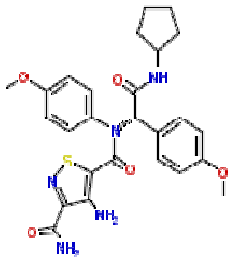
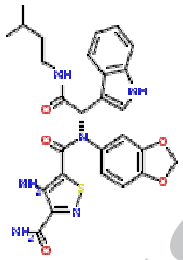
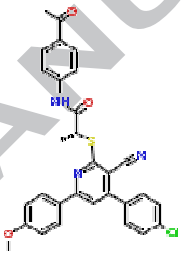
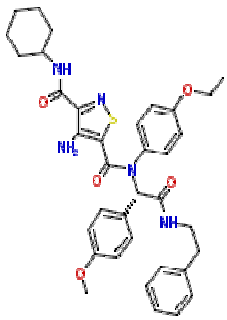
 MDG 425	4.5 ± 2.1	2.5 ± 1	2 ± 0.8	2.6 ± 1	2.5 ± 0.9	20
 MDG 426	2.18 ± 0.5	2.4 ± 0.5	1.2 ± 0.3	0.62 ± 0.2	0.47 ± 0.2	10
 MDG 427	5 ± 1.8	6.1 ± 2.2	4 ± 2	3.7 ± 1.7	3.23 ± 1.5	15
 MDG 428	12 ± 5.7	*	*	*	*	*
 MDG 430	9.6 ± 3.5	7 ± 2	4.6 ± 1.5	6.3 ± 3	5.5 ± 2.7	35
Chiral  MDG 431	25 ± 8.1	*	*	*	*	*
Chiral  MDG 432	2.21 ± 0.8	21.3 ± 7.5	10.6 ± 4	26 ± 10	20 ± 7.5	12
Chiral  MDG 433	32% [a]	*	*	*	*	*
Chiral  MDG 435	0.35 ± 0.1	30 % [b]		20 % [b]		20

Table 1. Inhibition values for the hits resulting from the PM virtual screen with Catalyst. Inhibition assays of *P. falciparum* infected RBCs and on-target PM assays were performed for all compounds. All values are quoted in micromolar. ± refers to the 95% confidence interval. (a) % Parasite Inhibition at 16 µM. (b) % Enzyme Inhibition at 5 µM. (*) Value not determined.

In summary, this work shows how a rationally designed virtual screening protocol can be successfully applied for the discovery of novel small molecule inhibitors of the plasmepsin aspartyl proteases, active against *P. falciparum.* The utility of Catalyst screening has been validated against a challenging flexible target ligand dataset, affording a more rigorous examination of the software's capabilities in 'real world' design challenges. The identified hits contain several different molecular scaffolds and attenuate the growth of the malaria parasite, *P.*

falciparum, in the low micromolar range. It is interesting to note that some compounds demonstrate a lower IC_{50} on the parasite than on PM II and IV in isolation and there are a number of considerations that may explain this phenomenon. Leaving aside the possibilities of target synergy, both MDG424 and 435 are weakly basic (as are 12 of the 15 hit compounds described in this work). Accordingly, and analogous to agents such as chloroquine, such physicochemical properties of the compounds could lead to their accumulation in the DV, arising from the pH in this compartment. In the whole parasite assay, under physiological conditions, protonation of the compounds would decrease their ability to traverse the DV membrane and so the concentration of compound could increase within the target organelle. In conjunction with this, it is suggested by docking studies that PM V and other members of the DV and non-DV PM family may also be targeted by these compounds (data not shown) and so an additive/synergistic anti-parasitic effect from inhibiting multiple-PMs could arise. On target IC_{50} values range from 2.3-21.3 μM for PM II and 0.62-26 μM for PM IV and importantly, as designed, demonstrate selectivity for the PMs over Cat D. The majority of compounds are equipotent on PM II and IV; however MDG419 displays a 2.7 fold selectivity for PM II over PM IV and MDG426 is almost four times more selective for PM IV over PM II. Adjusting and further quantifying the PM familial selectivity profile of these compounds will be the focus of future work. In parallel, MDG422, which does not display the typical hydrogen bonding interactions with the Asp catalytic dyad, will undergo SAR elucidation and PM binding affinity optimisation. Our intention is that these molecules will provide the basis for the development of additional selective molecular probes targeting aspartic peptidases for intervention in malaria.

Acknowledgements

We gratefully acknowledge funding support from Science Foundation Ireland (SFI) and the Irish Research Council for Science, Engineering & Technology (IRCSET). We thank the software vendors for their continuing support of such academic research efforts, in particular the contributions of Accelrys, the Chemical Computing Group and OpenEye Scientific, which significantly underpinned the work performed in this study and the Trinity Centre for High Performance Computing.

References

1. Moura, P. A.; Dame, J. B.; Fidock, D. A. *Antimicrob. Agents Chemother.* **2009**, *53*, 4968.
2. Silva, A. M.; Lee, A. Y.; Erickson, J. W.; Goldberg, D. E. *Adv. Exp. Med. Biol.* **1998**, *436*, 363.
3. Silva, A. M.; Lee, A. Y.; Gulnik, S. V.; Majer, P.; Collins, J.; Bhat, T. N.; Collins, P. J.; Cachau, R. E.; Luker, K. E.; Gluzman, I. Y.; Francis, S. E.; Oksman, A.; Goldberg, D. E.; Erickson, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10034.
4. Coombs, G. H.; Goldberg, D. E.; Klemba, M.; Berry, C.; Kay, J.; Mottram, J. C. *Trends Parasitol.* **2001**, *17*, 532.
5. Francis, S. E.; Gluzman, I. Y.; Oksman, A.; Knickerbocker, A.; Mueller, R.; Bryant, M. L.; Sherman, D. R.; Russell, D. G.; Goldberg, D. E. *EMBO J.* **1994**, *13*, 306.
6. Salas, F.; Fichmann, J.; Lee, G. K.; Scott, M. D.; Rosenthal, P. *J. Infect. Immun.* **1995**, *63*, 2120.
7. Naughton J.A.; Nasizadeh S.; Bell A. *Mol. Biochem. Parasitol.*. **2010**, 173, 81.
8. Nezami, A.; Freire, E. *Int. J. Parasitol.* **2002**, *32*, 1669.
9. Banerjee, R.; Goldberg, D. E. The Plasmodium food vacuole; Humana Press: Totowa, NJ, **2001**.
10. Omara-Opyene, A. L.; Moura, P. A.; Sulsona, C. R.; Bonilla, J. A.; Yowell, C. A.; Fujioka, H.; Fidock, D. A.; Dame, J. B. *J. Biol. Chem.* **2004**, *279*, 54088.
11. Liu, J.; Gluzman, I. Y.; Drew, M. E.; Goldberg, D. E. *J. Biochem.* **2005**, *280*, 1432.
12. Bonilla, J. A.; Moura, P. A.; Bonilla, T. D.; Yowell, C. A.; Fidock, D. A.; Dame, J. B. *Int. J. Parasitol.* **2007**, *37*, 317.
13. Bonilla, J. A.; Bonilla, T. D.; Yowell, C. A.; Fujioka, H.; Dame, J. B. *Mol. Microbiol.* **2007**, *65*, 64.
14. Russo I.; Babbitt S.; Muralidharan V.; Butler T.; Oksman A.; Goldberg DE. *Nature.* **2010**, 463, 632.
15. Bailly, E.; Jambou, R.; Savel, J.; Jaureguiberry, G. *J. Protozool.* **1992**, *39*, 593.
16. Luksch, T.; Chan, N. S.; Brass, S.; Sotriffer, C. A.; Klebe, G.; Diederich, W. E. *Chem. Med. Chem.* **2008**, *3*, 1323.
17. Corminboeuf, O.; Dunet, G.; Hafsi, M.; Grimont, J.; Grisostomi, C.; Meyer, S.; Binkert, C.; Bur, D.; Jones, A.; Prade, L.; Brun, R.; Boss, C. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 6194.
18. Boss, C.; Corminboeuf, O.; Grisostomi, C.; Weller, T. *Expert Opin. Ther. Pat.* **2006**; *16*, 295.
19. Boss, C.; Richard-Bildstein, S.; Weller, T.; Fischli, W.; Meyer, S.; Binkert, C. *Curr. Med. Chem.* **2003**, *10*, 883.
20. Catalyst; Accelrys Software Inc.: San Diego, CA, http://accelrys.com.
21. PipelinePilot; Scitegic: San Diego, CA, http://accelrys.com/products/scitegic.
22. Meyer, K. *J. Zhejiang Univ. Sci. B.* **2007**, *8*, 815.
23. Ersmark, K.; Nervall, M.; Hamelink, E.; Janka, L. K.; Clemente, J. C.; Dunn, B. M.; Blackman, M. J.; Samuelsson, B.; Aqvist, J.; Hallberg, A. *Int. J. Med. Chem.* **2005**, *48*, 6090.
24. Hof, F.; Schutz, A.; Fah, C.; Meyer, S.; Bur, D.; Liu, J.; Goldberg, D. E.; Diederich, F. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2138.
25. Maybridge; Fisher Scientific: Cornwall, United Kindom, www.maybridge.com.
26. SPECS; Specs: Delft - The Netherlands, www.specs.net.
27. Asinex http://www.asinex.com.
28. FRED; Openeye Scientific Software: Santa Fe, New Mexico, USA, www.eyesopen.com.

Figure captions

Figure 1: The selectivity score of each molecule targeting PM within the Wombat Database for the PMs vs CatD. The most selective molecules are shown in the hatched box

Figure 2. A) MDG422 aligned to the Ph4 generated by Catalyst. Feature keys: Green = HBA, Orange: HYD|ARO and Cyan = HYD. B) MDG422 docked (FRED) with PM II (PDB ID:1LEE). MDG422 shown in green and the reference ligand (R36) is shown in cyan.