FOCUS ARTICLE

ChemComm
www.rsc.org/chemcomm

Medicinal chemistry in academia: molecular recognition with biological receptors

Fraser Hof and François Diederich*

Laboratorium für Organische Chemie, ETH Hönggerberg, HCI, CH-8093 Zürich, Switzerland. E-mail: diederich@org.chem.ethz.ch

Why carry out medicinal chemistry at a university, when it means competing with the billion-dollar research efforts of the pharmaceutical industry? In academic research, the race to get a drug to market is not the prime motivation. Instead, university-based medicinal chemistry is driven by the search for new knowledge and the opportunity to educate a new generation of chemists. Furthermore, academia can complement commercial efforts by addressing diseases neglected by private industry.

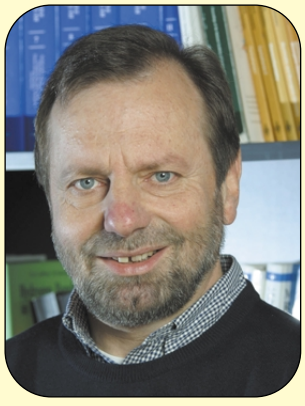
Pharmaceutical research has traditionally been considered the realm of industry, but academic laboratories have always played a role in developing new ideas and methods. No matter what the research setting, medicinal chemistry is a fertile field for learning about molecular behaviour. Medicinal chemistry is a significant activity in numerous laboratories around the world, wherever there is expertise in chemical synthesis and an interest in biological phenomena. Here we present three case histories from our own laboratory that demonstrate why and how medicinal chemistry is carried out in the academic setting.

Thrombin inhibitors from structure-based design

In the active site of a rigid enzyme (thrombin), we were able to carry out some fundamental molecular recognition investigations. This approach turns the "receptor design" paradigm of traditional molecular recognition chemistry inside out—here, one must design appropriate guests to engage the receptor from within its concave binding pockets. The detailed examination of ligand interactions within protein environments teaches valuable lessons in molecular recognition.

In the blood coagulation cascade, thrombin catalyses clot formation *via* the cleavage of the soluble blood protein fibrinogen into insoluble fibrin. Thrombin has long been a prime pharmaceutical target because clotting disorders are among the leading causes of death in the developed world.[1] Moreover, the X-ray structure of thrombin shows that its active site has limited flexibility.[2,3] These factors have driven various structure-based inhibitor design programs.[4,5]

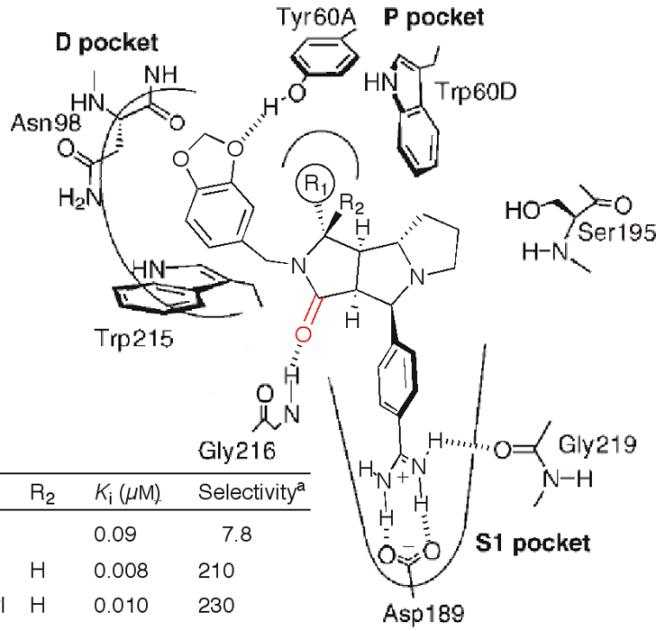
Owing to the rigidity of the active site, related ligands with functional group mutations form complexes of similar geometry, allowing direct correlation of measured changes in the complexation free enthalpy with the contributions of



François Diederich, born in 1952 in the Grand Duchy of Luxembourg, received his Ph.D. degree in 1979 from the University of Heidelberg. After postdoctoral studies at the University of California at Los Angeles (UCLA) and Habilitation at the Max-Planck-Institut für medizinische Forschung in Heidelberg, he joined in 1985 the Faculty of the Department of Chemistry and Biochemistry at UCLA where he became full Professor in 1989. In 1992, he moved to the ETH Zürich, where he currently serves as the head of the new Department of Chemistry and Applied Biosciences. His research interests span from medicinal chemistry, with a focus on molecular recognition studies, to advanced fullerene and acetylene-based materials with novel optoelectronic properties.

Fraser Hof was born in Medicine Hat, Canada. He obtained a B.Sc. in chemistry at the University of Alberta in 1998, and a Ph.D. at the Scripps Research Institute in 2003 under the direction of Professor Julius Rebek, Jr. He is currently a Novartis Stiftung post-doctoral fellow in the laboratory of Professor Diederich at ETH Zürich. His research centres on inhibiting the malarial protease plasmepsin II.

DOI: 10.1039/b314081b



	R_1	R_2	K_i (μM)	Selectivity[a]
(±)-**1**	=O		0.09	7.8
(±)-**2**	ethyl	H	0.008	210
(±)-**3**	cyclopropyl	H	0.010	230
(+)-**4**	isopropyl	H	0.007	740
(±)-**5**	cyclohexyl	H	1.7	2.6
(±)-**6**	phenyl	H	1.4	3.6

Fig. 1 *De novo* designed thrombin inhibitors that explore the binding affinity and selectivity available within the small, hydrophobic P pocket. The carbonyl group that hydrogen bonds to Gly216 is highlighted in red. [a] K_i (trypsin)/K_i (thrombin).

individual ligand substituents. Our own efforts at thrombin inhibition began with a rigid tricyclic core[6] which can direct vectors into the hydrophobic binding pockets P and D while introducing a phenylamidinium group to occupy the selectivity pocket S1 (Fig. 1). The crystal structure of the best first-generation inhibitor ((±)-**1**) complexed with thrombin confirmed the predicted binding mode.[7] Stepwise mutation of the inhibitor then allowed exploration of the molecular recognition properties of discrete regions within the thrombin active site.

One of the inhibitor's imide carbonyl oxygens accepts a hydrogen bond from the backbone NH group of Gly216 (Fig. 1).[8] When the hydrogen bond is removed, by replacing the carbonyl with a methylene group, the enzyme-inhibitor complex is destabilised by 3.4 kJ mol^{-1}. This provides a good estimate of the incremental free enthalpy arising from a single hydrogen bond in a biological setting.

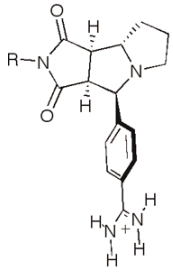
The P pocket of thrombin is a small hydrophobic cavity lined with aromatic residues (Tyr60A and Trp60D). In the first-generation inhibitors, modelling and crystallography showed that an imide carbonyl occupies this sub-site. These inhibitors demonstrate only modest (less than tenfold) selectivity for thrombin over trypsin. Since the P pocket is not present in related proteases, such as trypsin, we proposed that replacing this carbonyl with a more hydrophobic residue would increase both the affinity for thrombin and selectivity over trypsin.

The imide oxygen in question was replaced with a variety of alkyl groups to yield inhibitors (±)-**2**–(±)-**6** (Fig. 1).[8] Substitution with an isopropyl group ((+)-**4**) gives an impressive combination of increased affinity (7 nM) and selectivity over trypsin (740-fold). The crystal structure of (+)-**4** complexed with thrombin shows that, as predicted, the isopropyl group is nestled snugly against the aromatic residues that line the P pocket.[8]

Unlike the cosy P pocket, the D pocket is spacious enough to accommodate cyclic binding elements such as cyclohexylmethyl ((±)-**7**) and benzyl ((±)-**8**) groups (Fig. 2). According to modelling, both substituents adopt an edge-to-face orientation with respect to Trp215. Their similar affinity for thrombin suggests that specific electrostatic CH–π interactions need not be invoked for the benzyl complex, supporting earlier proposals that edge-to-face aromatic interactions are mainly dispersive in nature.[9]

Fluorinated organic compounds are well represented among current therapeutic agents, yet the nature of noncovalent interactions experienced by organic fluorine is still an issue for discussion.[10] We therefore treated the D pocket of thrombin to a fluorine scan,[11,12] an approach that might not have been favoured in an industrial setting, but one which yielded results of interest to researchers in academia and industry alike.

Compounds (±)-**9**–(±)-**14** (Fig. 2) each inhibited thrombin with similar efficiency, with one notable exception. The 4-fluoro derivative (±)-**11** shows fivefold enhanced activity ($\Delta\Delta G$ = 4.4 kJ mol^{-1}) relative to the parent compound (±)-**8**. The crystal structure of the enzyme-inhibitor complex reveals that the carbon-bound fluorine atom is in close contact with the enzyme. The fluorine atom is only 3.1 Å away from the α-carbon of Asn98, forming a dipolar C–H···F–C contact. It is also only 3.5 Å away from the carbonyl carbon of the same residue. Further exploration of the latter unusual interaction led to the discovery within the Cambridge Structural Database[13] of 43 examples of sub-van der Waals (d(F···C) < 3.09 Å) contacts between organic fluorine atoms and carbonyl carbons. The fluorine atoms clearly prefer a position close to the pseudo-trigonal axis of the carbonyl unit. At short distances, the F···C=O angle tends toward 90°; at longer separation the angle dependence is weaker. This attraction may arise from the interaction of the partially negative fluorine with the partially positive carbonyl carbon. Protein



	R	K_i (μM)
(±)-**7**	cyclohexylmethyl	0.35
(±)-**8**	benzyl	0.27
(±)-**9**	2-fluorobenzyl	0.50
(±)-**10**	3-fluorobenzyl	0.36
(±)-**11**	4-fluorobenzyl	0.057
(±)-**12**	2,6-difluorobenzyl	0.61
(±)-**13**	3,5-difluorobenzyl	0.59
(±)-**14**	pentafluorobenzyl	0.27

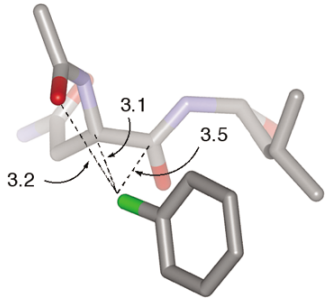


Fig. 2 Exploration of the thrombin D pocket, including a series of compounds that constitute a fluorine scan. On the right are depicted the unusually short fluorine contacts (Å) observed in the complex of thrombin with the active enantiomer of (±)-**11**.

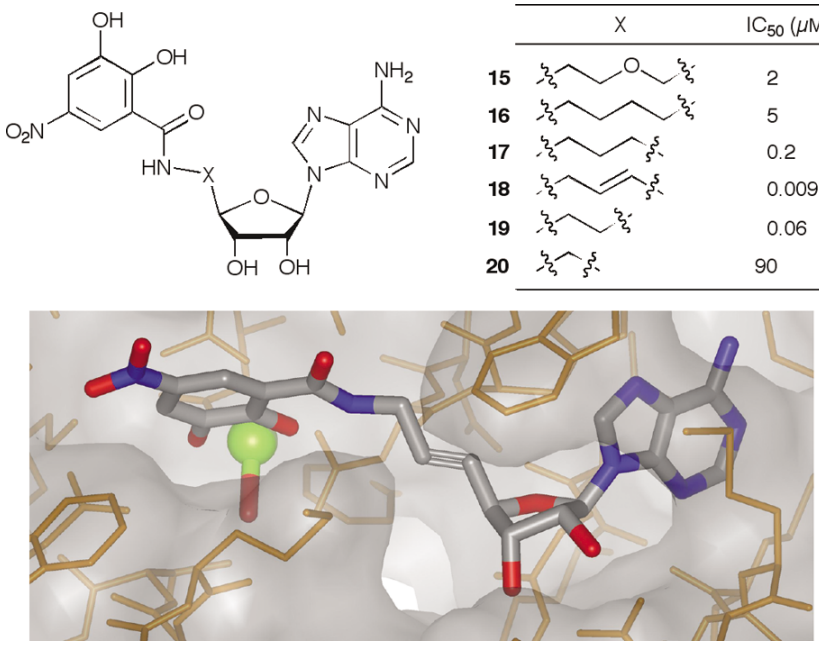
data base mining subsequently showed that the fluorine atoms of ligand CF_3 groups may also be able to form short dipolar C–F···C(O) contacts. Using this new information, structure-based designs should be able to use rational fluorine substitution to improve ligand affinity.

Novel design of bisubstrate inhibitors of catechol *O*-methyltransferase

Much can be learned about molecular behaviour through the detailed scrutiny of issues such as linker effects and enzyme kinetics. The enzyme catechol ***O***-methyltransferase (COMT) uses the cofactor *S*-adenosylmethionine (SAM) and a Mg^{2+} ion to methylate biological catechols, including L-dopa and dopamine, which can hinder treatment of Parkinson's disease by L-dopa. Concomitant treatment with COMT inhibitors decreases the rate of L-dopa metabolism, increasing the amount of the drug delivered to the brain.[14] Nitrocatechol derivatives that mimic the natural substrates have been developed as COMT inhibitors, including two currently marketed drugs (Tasmar® and Comptan®).[15,16] In a novel approach to COMT inhibition, we created inhibitors occupying both the SAM and catechol binding sites of the enzyme.[17] These involve an inhibitor design where a catechol derivative is attached *via* a variable linker to the 5′ position of adenosine. The methionine element of SAM is omitted on the understanding that its complementary polar binding pocket could be hydrated without great cost in free enthalpy. The success of the resulting inhibitors justifies this simplification, providing an important lesson in inhibitor design.

A small family of inhibitors was synthesised in which only the structure of the linker between adenosine and catechol was varied (Fig. 3).[17–20] The effect of the linker on inhibitory strength is striking; within this family, the inhibitory concentrations (IC_{50} = inhibitor concentration at which 50% V_{max} is observed) span five orders of magnitude! Evidence for the bisubstrate action of the best inhibitor **18** was provided by enzyme kinetics and X-ray crystallography.[18] Structural analysis of the ternary complex of COMT, Mg^{2+} and **18** shows the adenine and ribose moieties contacting the expected residues within the SAM pocket, while the nitrocatechol moiety occupies the catechol binding pocket as predicted (Fig. 3).

What gives rise to the variable performance of inhibitors **15–20**? Linkers that are too long (**15** and **16**) or too short (**20**) provide only micromolar-range inhibition of COMT. Comparing inhibitors **17** and **18** reveals a more subtle effect.



	X	IC_{50} (μM)
15		2
16		5
17		0.2
18		0.009
19		0.06
20		90

Fig. 3 Linker effects within a family of bisubstrate inhibitors of COMT, and the crystal structure showing the elongated binding conformation of the best inhibitor, **18**. The water-coordinated magnesium ion is shown in green.

Inhibitor **18** has a fully flexible linker, allowing hydrophobic collapse of the two aromatic binding elements: catechol and adenine. Introduction of one double bond in the linker of **18** removes one degree of freedom, and the effect of this structural rigidification is significant: **18** is 20 times more potent than **17**.

Drugs for a developing world—plasmepsin II, a new antimalarial target

Another major role for academic medicinal research is to tackle diseases of the developing world which are often neglected by the pharmaceutical industry. Ideally, these publicly produced lead compounds will be further developed into viable drugs in collaboration with commercial organizations. Malaria remains one of the most significant infectious diseases, affecting up to 500 million people annually.[21] The increasing prevalence of *Plasmodium* strains resistant to existing treatments demands the development of new antimalarial drugs with novel modes of action.[22]

Plasmepsin II is an aspartic protease used by *Plasmodium* parasites for the proteolysis of human haemoglobin to provide nutrients during their rapid growth and multiplication.[23] The related human aspartic protease, renin, undergoes a large conformational change that exposes a new hydrophobic *flap* pocket upon binding appropriate inhibitors.[24–26] Owing to the homology between renin and plasmepsin II, we hypothesised that a similar conformational shift may also occur in plasmepsin II. In the absence of a plasmepsin II structure in the *flap*-open conformation,[27] we initiated a structure-based design project for plasmepsin II inhibition using a homology model that superimposes the conformational shifts observed for renin onto plasmepsin II.[28]

Using computer modelling, we identified a suitable scaffold for inhibiting plasmepsin II. The structure engages the catalytic aspartate dyad with a bicyclic ammonium ion while directing a large aromatic substituent into each of the hydrophobic S1/S3 cleft and the *flap* pocket (Fig. 4).[28–30] In the modelled structure of plasmepsin II the S1/S3 cleft is shallow and solvent-exposed, suggesting that little specificity can be gained in this area. But the addition of a single chlorine atom (as in (±)-**24**) to fill space at the rear of the hypothetical *flap* pocket provides a fourfold boost in activity.[28] Omission of either of the two aromatic binding elements (structures not shown) results in tenfold decreased activity, suggesting that both the S1/S3 cleft and the hypothesised *flap* pocket must be filled for efficient ligand binding.

The biological results obtained with these first-generation inhibitors suggest that the opening of the *flap* pocket is possible for plasmepsin II. This hypothesis has now been validated with a conceptually similar ligand. This work, carried out at Roche, provided the first crystal structure of plasmepin II in the *flap*-open conformation.[31] The good agreement between the experimental structure and the homology model confirms the accuracy of the computer-based structure prediction.

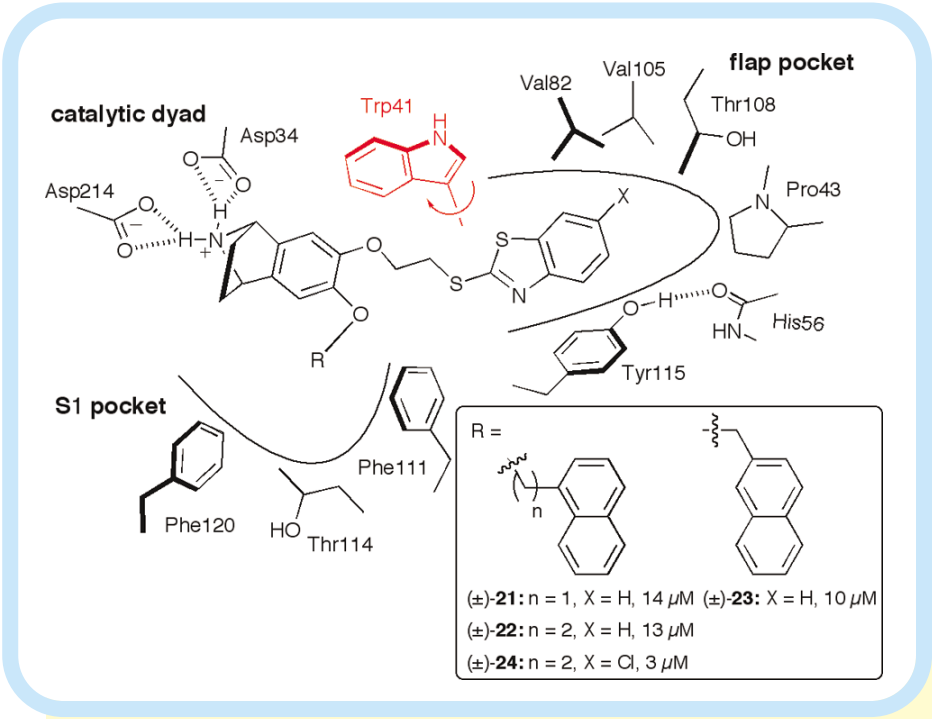


Fig. 4 First-generation inhibitors of plasmepsin II, and their proposed occupation of the active site. The Trp41 residue that swings aside to open the flap pocket is highlighted in red. The design of these inhibitors was based on an enzyme structure constructed by homology modelling.

In conclusion

We are well aware that there are other important motives for academic medicinal chemistry that are not addressed within the limited space of this article. Notable examples are the searches for practical solutions for interfering with protein–protein binding and for binding to flat protein surfaces.[32,33] Such advances would make available a plethora of new targets for disease treatment. Another important objective for academic medicinal chemistry is to contribute to the development of efficient *in silico* screening methods of virtual libraries.

Along with new lessons in molecular recognition, design strategies and leads for neglected diseases, medicinal research in academia provides another significant benefit. It is the most important, and often overlooked, reason for carrying out research in an academic environment: the introduction of students and postdoctoral chemists to the fascinating problems of biological recognition within the context of chemical synthesis. In this way, interdisciplinary research in academia acts as an 'incubator'—training new researchers, many of whom go on to make valuable contributions to the pharmaceutical industry.

Notes and references

1 M. F. Scully, in *Essays in Biochemistry*, ed. K. F. Tipton, Portland Press, London, 1996, p. 17.

2 W. Bode, I. Mayr, U. Baumann, R. Huber, S. R. Stone and J. Hofsteenge, *EMBO J.*, 1989, **8**, 3467.

3 D. W. Banner and P. Hadvary, *J. Biol. Chem.*, 1991, **266**, 20085.

4 P. E. J. Sanderson and A. M. Naylor-Olsen, *Curr. Med. Chem.*, 1998, **5**, 289.

5 K. Menear, *Curr. Med. Chem.*, 1998, **5**, 457.

6 R. Grigg and S. Thianpatanagul, *J. Chem. Soc., Chem. Commun.*, 1984, 180.

7 U. Obst, V. Gramlich, F. Diederich, L. Weber and D. W. Banner, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 1739.

8 U. Obst, D. W. Banner, L. Weber and F. Diederich, *Chem. Biol.*, 1997, **4**, 287.

9 S. Paliwal, S. Geib and C. S. Wilcox, *J. Am. Chem. Soc.*, 1994, **116**, 4497.

10 J. D. Dunitz and R. Taylor, *Chem. Eur. J.*, 1997, **3**, 89.

11 J. A. Olsen, D. W. Banner, P. Seiler, U. Obst Sander, A. D'Arcy, M. Stihle, K. Müller and F. Diederich, *Angew. Chem., Int. Ed.*, 2003, **42**, 2507.

12 C.-Y. Kim, J. S. Chang, J. B. Doyon, J. T. A. Baird, C. A. Fierke, A. Jain and D. W. Christianson, *J. Am. Chem. Soc.*, 2000, **122**, 12125.

13 CSD version 9.24, November 2001.

14 P. T. Mannistö, I. Ulmanen, K. Lundström, J. Taskinen, J. Tenhunen, C. Tilgmann and S. K. Kaakkola, *Prog. Drug Res.*, 1992, **39**, 291.

15 J. Borgulya, M. D. Prada, R. Dingemanse, B. Scherschlicht, B. Schläppi and G. Zuercher, *Drugs Future*, 1991, **16**, 719.

16 E. Nissinen, I. B. Linden, E. Schultz and P. Pohto, *Arch. Pharmacol.*, 1992, **346**, 262.

17 B. Masjost, P. Ballmer, E. Borroni, G. Zuercher, F. K. Winkler, R. Jakob-Roetne and F. Diederich, *Chem. Eur. J.*, 2000, **6**, 971.

18 C. Lerner, A. Ruf, V. Gramlich, B. Masjost, G. Zuercher, R. Jakob-Roetne, E. Borroni and F. Diederich, *Angew. Chem., Int. Ed.*, 2001, **40**, 4040.

19 C. Lerner, R. Siegrist, E. Schweizer, F. Diederich, V. Gramlich, R. Jakob-Roetne, G. Zuercher and E. Borroni, *Helv. Chim. Acta*, 2003, **86**, 1045.

20 C. Lerner, B. Masjost, A. Ruf, V. Gramlich, R. Jakob-Roetne, G. Zuercher, E. Borroni and F. Diederich, *Org. Biomol. Chem.*, 2003, **1**, 42.

21 M. Nacher, *Trends Parasitol.*, 2001, **17**, 563.

22 C. H. Sibley, J. E. Hyde, P. F. G. Sims, C. V. Plowe, J. G. Kublin, E. K. Mberu, A. F. Cowman, P. A. Winstanley, W. M. Watkins and A. M. Nzila, *Trends Parasitol.*, 2001, **17** , 582.

23 C. Boss, S. Richard-Bildstein, T. Weller, W. Fischli, S. Meyer and C. Binkert, *Curr. Med. Chem.*, 2003, **10**, 883.

24 C. Oefner, A. Binggeli, V. Breu, D. Bur, J. P. Clozel, A. D'Arcy, A. Dorn, W. Fischli, F. Gruninger, R. Guller, G. Hirth, H. P. Marki, S. Mathews, M. Muller, R. G. Ridley, H. Stadler, E. Vieira, M. Wilhelm, F. K. Winkler and W. Wostl, *Chem. Biol.*, 1999, **6**, 127.

25 M. G. Bursavich and D. H. Rich, *Org. Lett.*, 2001, **3**, 2625.

26 M. G. Bursavich, C. W. West and D. H. Rich, *Org. Lett.*, 2001, **3**, 2317.

27 O. A. Asojo, S. V. Gulnik, E. Afonina, B. Yu, J. A. Ellman, T. S. Haque and A. M. Silva, *J. Mol. Biol.*, 2003, **327**, 173.

28 D. A. Carcache, S. R. Hörtner, A. Bertogg, C. Binkert, D. Bur, H. P. Märki, A. Dorn and F. Diederich, *ChemBioChem*, 2002, **3**, 1137.

29 D. A. Carcache, S. R. Hörtner, A. Bertogg, F. Diederich, A. Dorn, H. P. Märki, C. Binkert and D. Bur, *Helv. Chim. Acta*, 2003, **86**, 2192.

30 D. A. Carcache, S. R. Hörtner, P. Seiler, F. Diederich, A. Dorn, H. P. Märki, C. Binkert and D. Bur, *Helv. Chim. Acta*, 2003, **86**, 2173.

31 J. Yuvaniyama, A. D'Arcy, C. Oefner, H. Loetscher, D. Bur, G. E. Dale, H. P. Märki, F. K. Winkler, R. P. Moon and R. G. Ridley, Symposium on "Parasite Proteases as Drug Targets", University of Glasgow, UK, 2000.

32 H. S. Park, Q. Lin and A. D. Hamilton, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5105.

33 L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. A. Liu, *Science*, 2004, in press.