MedChemComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/medchemcomm

Fragment growing exploiting dynamic combinatorial chemistry of inhibitors of the aspartic protease endothiapepsin

Milon Mondal, Daphne E. Groothuis and Anna K. H. Hirsch

Abstract

Fragment-based drug design (FBDD) has emerged as an efficient hit-identification and/or -optimization strategy with a higher hit rate than high-throughput screening (HTS). Whereas fragment linking is more challenging, fragment growing has become the preferred fragment-optimization strategy, requiring synthesis of derivatives and validation of their binding mode at each step of the optimization cycle. Dynamic combinatorial chemistry (DCC) constitutes a powerful and efficient strategy to identify ligands for biological targets. Here, we have demonstrated that the novel combination of fragment-growing and DCC is a highly powerful strategy to grow a fragment into a more potent, non-covalent inhibitor of the aspartic protease endothiapepsin. We have designed a library of acylhydrazones using fragment growing starting from a known fragment in complex with endothiapepsin. We have used DCC and a fluorescence-based enzymatic assay to identify the best hit(s) from the dynamic combinatorial libraries, displaying double-digit micromolar inhibition of endothiapepsin. In addition, each DCC experiment requires only very small amounts of protein compared with established methods of analysis and the protein needs to be in the assay mixture only for a short period of time, making this protocol ideal for precious and unstable proteins. These results constitute a proof of concept that the combination of fragment-growing and DCC constitutes a powerful and efficient strategy to convert a fragment into a hit.

Introduction

Over the past decades, fragment-based design (FBDD) has become a well-established strategy for the identification of inhibitors of numerous biological targets.[1–3] FBDD has higher hit rates than high-throughput screening (HTS), with a better coverage of the chemical space.[1] Once a fragment has been identified, it has to be optimized to a lead compound by, fragment growing or linking. On the one hand, fragment growing has become the preferred fragment-optimization strategy.[2,4–8] Preserving the binding modes of the two fragments bound to adjacent pockets, as required in fragment linking, is considered

attractive given the potential for super-additivity but at the same time very challenging.[9] On the other hand, fragment growing is time-consuming as it requires synthesis of derivatives and validation of their binding mode at each step of the fragment-optimization cycle. To overcome this hurdle, we developed a strategy in which we combine fragment growing with dynamic combinatorial chemistry (DCC) for the optimization of a fragment to a hit. Over the past decade, DCC has emerged as an efficient and innovative approach in drug discovery as it allows for the formation of a dynamic combinatorial library (DCL), in which the chemical bonds that connect the multiple building blocks are continuously being made and broken, allowing for the *in situ* formation of a great variety of compounds.[10,11] There are numerous reports of DCC to facilitate hit/lead identification or optimization.[11] Proof of concept for fragment linking facilitated by DCC has been provided retrospectively when letting disconnected known hydrazine inhibitors reassemble in presence of crystals of the protein target.[12] Linking of an intermediate form biosynthesis to fragments designed to occupy adjacent pockets by disulfide-based DCC[13] as well as covalent tethering have been reported over the past decade.[14–16] FBDD in combination with DCC has, however, never been explicitly reported for conventional fragment growing in a strict FBDD context. We introduce here a fluorescence-based enzymatic assay to conveniently screen the DCL for active inhibitors, requiring little amount and short presence of the protein target, making it ideally suited for precious and unstable proteins.[17]

In the present study, we have applied this strategy to optimize a fragment into a hit for endothiapepsin. Endothiapepsin is a pepsin-like aspartic protease. Members of this class of enzymes play a causative role in numerous diseases such as malaria (plasmepsins), Alzheimer's disease (β-secretase), fungal infections (secreted aspartic proteases), and hypertension (renin).[18] Endothiapepsin has been used as the model enzyme for the development of fragments of renin[19] and β-secretase[20] as well as to better understand the mechanism of action of aspartic proteases.[21–23] Aspartic proteases consist of two similar subunits, each of which contributes an aspartic acid residue to the catalytic dyad (D35 and D219 for endothiapepsin) that hydrolyzes the peptide bond of the substrate using a catalytic water molecule.

Imine-based chemistry has been extensively used for protein-templated DCC in drug design projects.[24] Owing to the inherent instability of imines in aqueous media, acylhydrazones have become the working horse for imine-type biomedical DCC projects.[25–27] Acylhydrazone-based DCC is particularly attractive given that the resulting products provide both H-bond donor and -acceptor sites and are stable enough

in acidic and physiological conditions to enable direct analysis. In our previous work, we showed that acylhydrazone-based DCC is compatible with the target endothiapepsin.[27]

Here, we will discuss how we combine FBDD and DCC for the efficient fragment-to-hit optimization for aspartic protease endothiapepsin.

Results and discussion

For our studies, we used the crystal structure of endothiapepsin in complex with fragment **1** (Protein Data Bank (PDB) code: 3PCW).[28] Fragment **1** was found to inhibit endothiapepsin during a fragment-based screening campaign.[28] As shown in Figure 1, **1** is engaged in three charged H bonds with the catalytic dyad consisting of residues D35 and D219, by using its amidine group. Except for the clogP value (3.3), **1** obeys the "rule of three",[29] having a molecular weight (M_w) of 261 Dalton, three H-bond donors, two H-bond acceptors, one freely rotatable bond and a total polar surface area (TPSA) of 49.9 Å^2. The fact that **1** shows 45% inhibition of endothiapepsin at a concentration of 1 mM, combined with its promising physicochemical properties and the fact that it only has 13 heavy atoms, encouraged us to choose it as a starting point for optimization into an inhibitor of endothiapepsin.

Fragment **1** occupies the S2 and part of the S1 and S1' pockets of endothiapepsin and addresses the catalytic dyad through charged H-bonding interactions (Figure 1). By using the molecular-modeling software Moloc[30] and the FlexX docking module in the LeadIT suite,[31] we designed a derivative of **1**, in which the amidine group is replaced by an α-amino group to address both D35 and D219 of the catalytic dyad through H-bonding interactions. As we have already shown that the acylhydrazone moiety is a suitable central scaffold to address the catalytic dyad of endothiapepsin,[27] we decided to use an acylhydrazone moiety to grow **1** towards the S1 and S3 pockets. We introduced an acylhydrazone linker pointing towards the S1 & S3 pockets. The NH group of the newly designed acylhydrazone moiety is engaged in an H-bonding interaction with D35, while the carbonyl group of the acylhydrazone linker accepts an H bond from the backbone of G80. Careful investigation of known co-crystal structures of endothiapepsin[25] and hotspot analysis[32] of the active site of endothiapepsin suggested that both aliphatic and aromatic substituents can be introduced through the aldehyde of the acylhydrazone linker, which can be involved in hydrophobic interactions with D125, I122, D116 and Y79 in the S1 & S3 pockets (Figure 2). Based on molecular modeling and docking studies, we selected a series of nine acylhydrazone-based inhibitors (Scheme 1a). All the derivatives can be conveniently

synthesized starting from hydrazide **2**, by reacting it with nine aldehydes (**3a**–**11a**) (Scheme 1b). While all aldehydes are commercially available, we synthesized **2** via an asymmetric Strecker reaction starting from commercially available *para*-trifluorobenzaldehyde (**11a**) (Scheme S1 in the supporting information).[33]

We set up nine DCLs, each containing hydrazide **2** and one of the nine aldehydes **3a–11a** to form the corresponding acylhydrazones **3h–11h** (Scheme 1b). We used a fluorescence-based enzymatic activity assay to conveniently screen the DCL for active inhibitors as this method requires very little amount of protein.[18] In the assay, we used the fluorogenic substrate of HIV protease (2-aminobenzoyl-Thr-Ile-Nle-Phe(*p*-NO_2)-Gln-Arg-NH_2, **12**), which is hydrolyzed by endothiapepsin and generates two fluorescent fragments, namely **13** and **14** enabling convenient spectrophotometric monitoring (Scheme S2 in the supporting information). Having investigated the DCLs using the fluorescence-based assay, we identified a total of two DCLs out of nine displaying higher inhibition of endothiapepsin than **1**, showing 75% (IC_{50} = 407 μM) and 79% (IC_{50} = 252 μM) inhibition at an inhibitor concentration of 1 mM (based on the acylhydrazones formed *in situ*) for the DCLs **3a** + **2** and **9a** + **2**, respectively (Figure 3). In the DCLs, three components (aldehyde, hydrazide and acylhydrazone) are in equilibrium. As a result, full conversion to the corresponding acylhydrazones is not ensured. In addition, the excess hydrazide building blocks might compete for binding. As a result, the IC_{50} values determined this way are expected to be higher than for isolated and purified compounds.

To determine the accurate IC_{50} values of the hit compounds identified from screening the DCLs, we synthesized the corresponding acylhydrazones **3h** and **9h** individually (Scheme S3 in the supporting information). By adopting the same fluorescence-based enzyme activity assay, we confirmed the results from the screening of the DCLs and established that acylhydrazones **3h** and **9h** display IC_{50} values of 210 μM and 85 μM, respectively. We calculated the experimental Gibbs free energies of binding (ΔG_{EXPT} (**3h**) = –23 kJ mol,$^{-1}$ ΔG_{EXPT} (**9h**) = –25 kJ mol^{-1}) and ligand efficiencies (LE (**3h**) = 0.24, LE (**9h**) = 0.27) from the IC_{50} values using the Cheng-Prusoff equation.[34]

As shown in Figure 4, we designed **3h** and **9h** so as to bind in a similar way, addressing the catalytic dyad using their α-amino group to form direct H bonds with D35 and D219. The amide groups of both compounds donate an H bond to D35. The carbonyl group of the acylhydrazone moieties also accept an H bond from the backbone amide of G80. The *para*-trifluorophenyl group of both acylhydrazones occupies the S2 and part of the S1 and S1' pockets of endothiapepsin and is involved in several

hydrophobic interactions with I300, I302, I304, I217 and F19, which maintains the binding mode of the fragment. The phenyl and cyclopentyl groups of the acylhydrazones fill the S3 pocket and are engaged hydrophobic contacts with D125, I122, D116 and Y79.

Conclusions

In conclusion, we have demonstrated for the first time, that the combination of fragment growing and DCC is a powerful technique for the rapid and efficient identification of novel hits for the aspartic protease endothiapepsin. Moreover, by using a fluorescence-based assay, we could directly screen the DCLs for active inhibitors. Advantages of this approach are that only very small amounts of protein are required compared with established methods of analysis and that the protein needs to be in the assay mixture only for a short period of time, making this protocol ideal for precious and unstable proteins. Among the acylhydrazones identified, the most potent inhibitor **9h** displays an IC_{50} value of 85 μM. This synergistic combination of computational and analytical methodologies proved to be successful for the development of inhibitors of the aspartic protease endothiapepsin and could be applied to a wide range of targets.

Contacts:

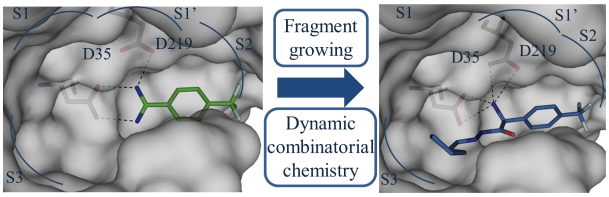
M. Mondal, D. E. Groothuis, Dr. A. K. H. Hirsch
Stratingh Institute for Chemistry, University of Groningen
Nijenborgh 7, 9747 AG Groningen (The Netherlands)
E-mail: a.k.h.hirsch@rug.nl
Homepage: http://www.rug.nl/research/bio-organic-chemistry/hirsch/

Acknowledgments:

Funding was granted by the Netherlands Organisation for Scientific Research (NWO-CW, ChemThem grant to A.K.H.H.) and by the Dutch Ministry of Education, Culture, Science (gravitation program 024.001.035).

Graphics

Table of contents



The novel combination of fragment growing and DCC constitutes a powerful and efficient strategy to convert a fragment into a hit.

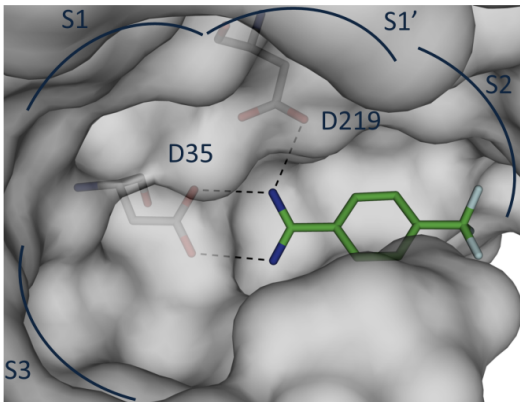


Figure 1. X-ray crystal structure of endothiapepsin in complex with fragment **1** bound in the active site (PDB code: 3PCW).[25] Color code: protein surface: C: gray, O: red and N: blue; fragment skeleton: C: green, N: blue, F: cyan. Hydrogen bonds below 3.0 Å are shown as black dashed lines.

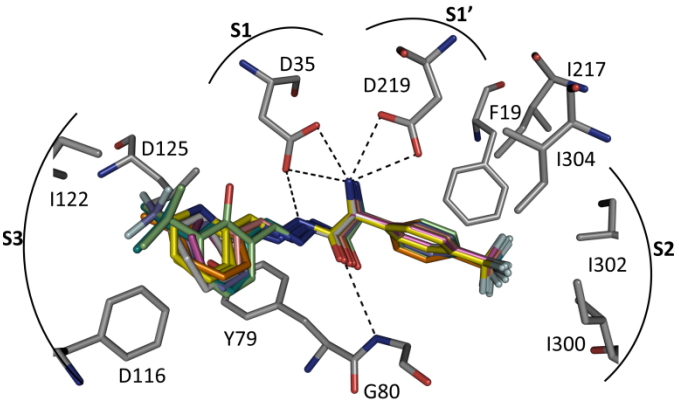
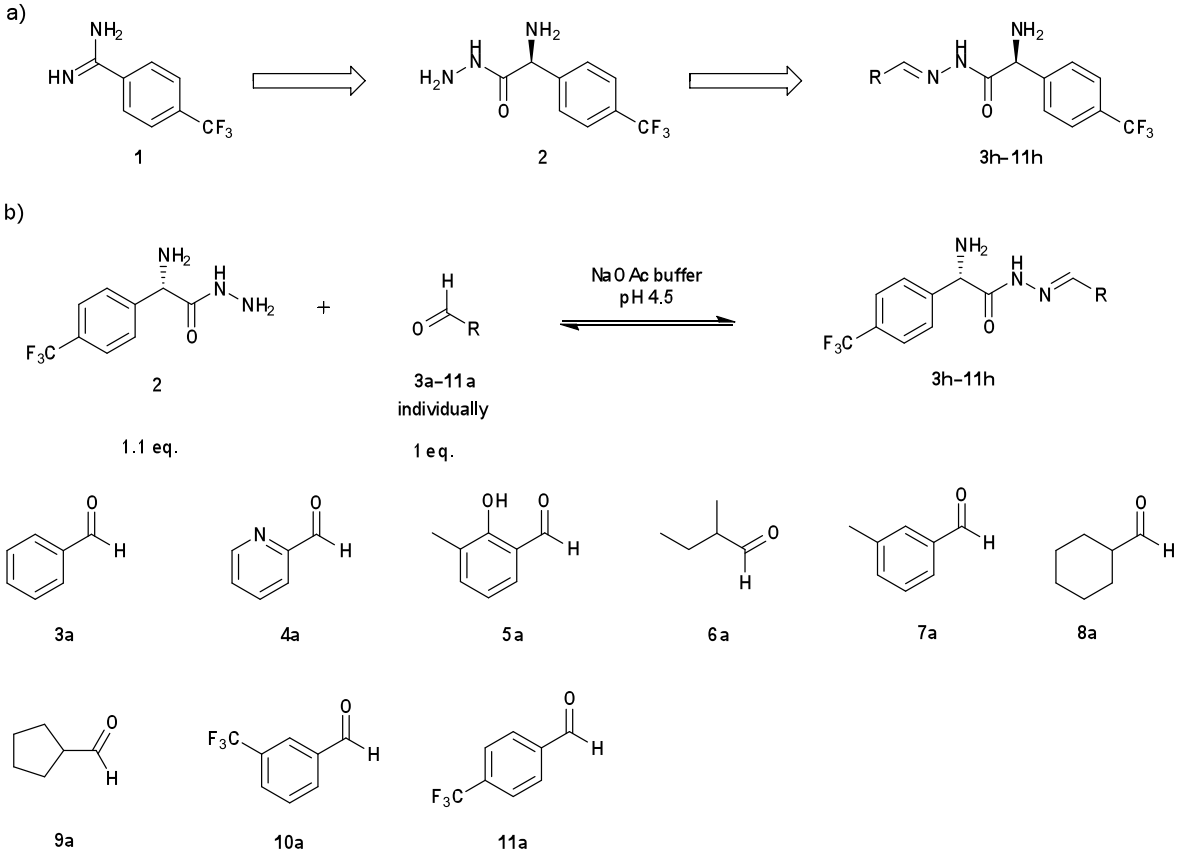


Figure 2. Superposition of MOLOC-generated molecular models of potential acylhydrazone inhibitors featuring direct H-bonding interactions with the catalytic dyad (D35 and D219) in the active site of endothiapepsin (PDB

code: 3PCW).[25] Color code: protein skeleton: C: gray; inhibitor skeletons: C: green, violet, purple, blue, yellow, N: blue, F: cyan, O: red. Hydrogen bonds below 3.4 Å are shown as black dashed lines.



Scheme 1. (a) Generation of acylhydrazone-based inhibitors **3h–11h** from fragment **1**. (b) Generation of acylhydrazone-based DCLs from hydrazide **2** and structures of the individual aldehydes **3a–11a** for fluorescence-based assay screening.

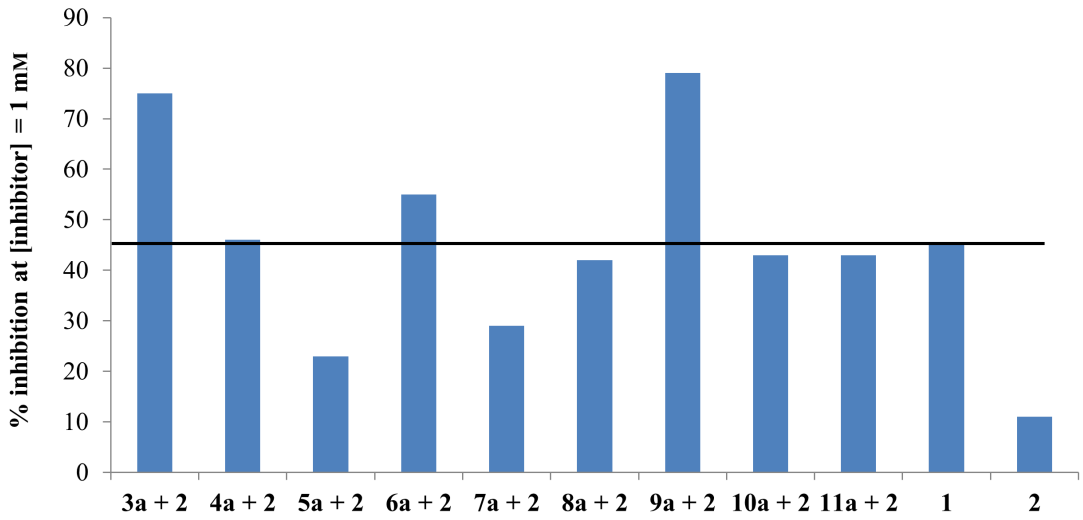


Figure 3. Bar graph of fluorescence-based assay results of all 9 DCLs as well as fragments **1** and **2**. The %inhibition (mM) is plotted for each DCL as well as fragments **1** and **2**.

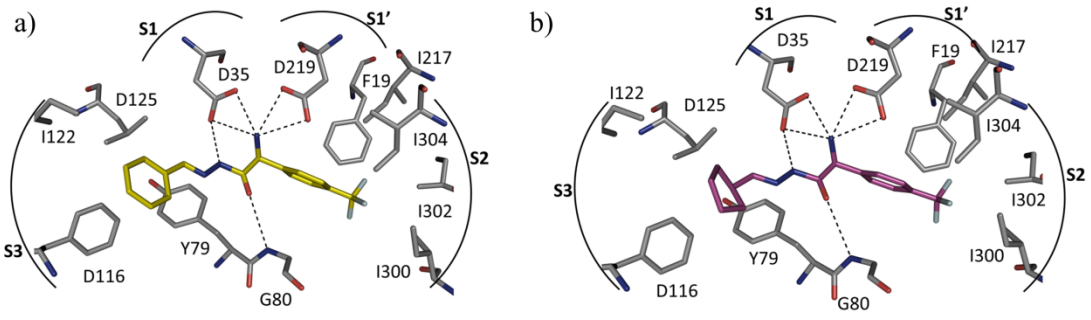


Figure 4. Predicted full binding mode of (a) **3h** and (b) **9h** in the catalytic active site of endothiapepsin (PDB code: 3PCW).[25] Color code: enzyme skeleton: C: gray; inhibitor skeleton: C: yellow and violet, N: blue, O: red and F: cyan. Hydrogen bonds below 3.4 Å are shown as black dashed lines.

References

1. P. J. Hajduk and J. Greer, *Nat. Rev. Drug Discov.*, 2007, **6**, 211–219.

2. G. E. de Kloe, D. Bailey, R. Leurs, and I. J. P. de Esch, *Drug Discov. Today*, 2009, **14**, 630–646.

3. D. A. Erlanson, *Curr. Opin. Biotechnol.*, 2006, **17**, 643–652.

4. S. J. Taylor, A. Abeywardane, S. Liang, I. Muegge, A. K. Padyana, Z. Xiong, M. Hill-Drzewi, B. Farmer, X. Li, B. Collins, J. X. Li, A. Heim-Riether, J. Proudfoot, Q. Zhang, D. Goldberg, L. Zuvela-Jelaska, H. Zaher, J. Li, and N. A. Farrow, *J. Med. Chem.*, 2011, **54**, 8174–8187.

5. Y. Cheng, T. C. Judd, M. D. Bartberger, J. Brown, K. Chen, R. T. Fremeau, D. Hickman, S. A. Hitchcock, B. Jordan, V. Li, P. Lopez, S. W. Louie, Y. Luo, K. Michelsen, T. Nixey, T. S. Powers, C. Rattan, E. A. Sickmier, D. J. St. Jean, R. C. Wahl, P. H. Wen, and S. Wood, *J. Med. Chem.*, 2011, **54**, 5836–5857.

6. M. Congreve, G. Chessari, D. Tisi, and A. J. Woodhead, *J. Med. Chem.*, 2008, **51**, 3661–3680.

7. G. Chessari and A. J. Woodhead, *Drug Discov. Today*, 2009, **14**, 668–675.

8. E. Edink, P. Rucktooa, K. Retra, A. Akdemir, T. Nahar, O. Zuiderveld, R. Van Elk, E. Janssen, P. Van Nierop, J. Van Muijlwijk-Koezen, A. B. Smit, T. K. Sixma, R. Leurs, and I. J. P. De Esch, *J. Am. Chem. Soc.*, 2011, **133**, 5363–5371.

9. S. Chung, J. B. Parker, M. Bianchet, L. M. Amzel, and J. T. Stivers, *Nat. Chem. Biol.*, 2009, **5**, 407–413.

10. A. Herrmann, *Chem. Soc. Rev.*, 2014, **43**, 1899–933.

11. M. Mondal and A. K. H. Hirsch, *Chem. Soc. Rev.*, 2015, **44**, 2455–2488.

12. M. S. Congreve, D. J. Davis, L. Devine, C. Granata, M. O'Reilly, P. G. Wyatt, and H. Jhoti, *Angew. Chemie - Int. Ed.*, 2003, **42**, 4479–4482.

13. D. E. Scott, G. J. Dawes, M. Ando, C. Abell, and A. Ciulli, *ChemBioChem*, 2009, **10**, 2772–2779.

14. B. M. R. Liénard, R. Hüting, P. Lassaux, M. Galleni, J.-M. Frère and C. J. Schofield, *J. Med. Chem.*, 2008, **51**, 684–688.

15. N. R. Rose, E. C. Y. Woon, G. L. Kingham, O. N. F. King, J. Mecinović, I. J. Clifton, S. S. Ng, J. Talib-Hardy, U. Oppermann, M. A. McDonough and C. J. Schofield, *J. Med. Chem.*, 2010, **53**, 1810–1818.

16. M. T. Cancilla, M. M. He, N. Viswanathan, R. L. Simmons, M. Taylor, A. D. Fung, K. Cao and D. A. Erlanson, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 3978–3981.

17. M. V .Toth and G. R. Marshall, *Int. J. Pept. Protein Res.*, 1990, **36**, 544–550.

18. J. B. Cooper, *Curr. Drug Targets*, 2002, **106**, 3652−3711.

19. J. Cooper, W. Quail, C. Frazao, S. I. Foundling, T. L. Blundell, C. Humblet, E. A. Lunney, W. T. Lowther, and B. M. Dunn, *Biochemistry*, 1992, **31**, 8142–8150.

20. S. Geschwindner, L. L. Olsson, J. S. Albert, J. Deinum, P. D. Edwards, T. De Beer, and R. H. A. Folmer, *J. Med. Chem.*, 2007, **50**, 5903–5911.

21. L. Coates, P. T. Erskine, S. Mall, R. Gill, S. P. Wood, D. A. A. Myles, and J. B. Cooper, *Eur. Biophys. J.*, 2006, **35**, 559–566.

22. L. Coates, H.-F. Tuan, S. Tomanicek, A. Kovalevsky, M. Mustyakimov, P. Erskine, and J. Cooper, *J. Am. Chem. Soc.*, 2008, **130**, 7235–7237.

23. L. Coates, P. T. Erskine, S. P. Wood, D. A. A. Myles, and J. B. Cooper, *Biochemistry*, 2001, **40**, 13149–13157.

24. I. Huc and J. M. Lehn, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 2106–2110.

25. A. J. Clipson, V. T. Bhat, I. McNae, A. M. Caniard, D. J. Campopiano, and M. F. Greaney, *Chem. – Eur. J.*, 2012, **18**, 10562–10570.

26. V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, and M. F. Greaney, *Nat. Chem.*, 2010, **2**, 490–497.

27. M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, and A. K. H. Hirsch, *Angew. Chem. Int. Ed.*, 2014, **53**, 3259–3263.

28. H. Köster, T. Craan, S. Brass, C. Herhaus, M. Zentgraf, L. Neumann, A. Heine, and G. Klebe, *J. Med. Chem.*, 2011, **54**, 7784–7796.

29. M. Congreve, R. Carr, C. Murray, and H. Jhoti, *Drug Discov. Today*, 2003, **8**, 876–877.

30. P. R. Gerber and K. Müller, *J. Comput.-Aided. Mol. Des.*, 1995, **9**, 251–268.

31. BioSolveIT GmbH, Sankt Augustin. http://www.biosolveit.de, *LeadIT*, version 2.1.3.

32. H. Gohlke, M. Hendlich, and G. Klebe, *Perspect. Drug Discovery* Des., 2000, **20**, 115–144.

33. Y. Pérez-Fuertes, J. E. Taylor, D. A. Tickell, M. F. Mahon, S. D. Bull, and T. D. James, *J. Org. Chem.*, 2011, **76**, 6038–6047.

34. H. C. Cheng, *J. Pharmacol. Toxicol. Methods*, 2001, **46**, 61–71.