Lead Structure Optimization

DOI: 10.1002/anie.201411206

Tracing Binding Modes in Hit-to-Lead Optimization: Chameleon-Like Poses of Aspartic Protease Inhibitors**

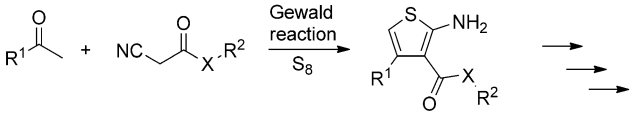
*Maren Kuhnert, Helene Köster, Ruben Bartholomäus, Ah Young Park, Amir Shahim, Andreas Heine, Holger Steuber, Gerhard Klebe, and Wibke E. Diederich**

Abstract: *Successful lead optimization in structure-based drug discovery depends on the correct deduction and interpretation of the underlying structure–activity relationships (SAR) to facilitate efficient decision-making on the next candidates to be synthesized. Consequently, the question arises, how frequently a binding mode (re)-validation is required, to ensure not to be misled by invalid assumptions on the binding geometry. We present an example in which minor chemical modifications within one inhibitor series lead to surprisingly different binding modes. X-ray structure determination of eight inhibitors derived from one core scaffold resulted in four different binding modes in the aspartic protease endothiapepsin, a well-established surrogate for e.g. renin and β-secretase. In addition, we suggest an empirical metrics that might serve as an indicator during lead optimization to qualify compounds as candidates for structural revalidation.*

In the field of drug discovery and medicinal chemistry, the concept of target-structure-based design considerations in the optimization process from hit-to-lead and (pre-)clinical candidate to an eventually approved drug has been well established over the last three decades.[1] Success stories of this structure-guided design approach range from the discovery of various HIV-1 protease inhibitors via the design of the neuraminidase inhibitor oseltamivir, the BCR-Abl kinase inhibitor imatinib, and various HCV protease inhibitors to a virtually unlimited number of current design campaigns that make use of structure-educated decision-making in prioritizing medicinal chemistry activities.[2]

This optimization typically starts with one or more identified hits sharing a common basic scaffold. The binding mode of one representative with this scaffold and the target protein is determined by crystallography, and subsequently a series of derivatives is synthesized and their activity is monitored by appropriate assays. Typically, structural modifications are realized assuming that an identified binding mode of a given parent scaffold is maintained during structure–activity relationship (SAR) exploration and optimization. However, despite the meanwhile long history of structure-guided lead discovery there is an ongoing debate on how frequently structural support is required to establish a reliable SAR by monitoring putative changes of ligand binding geometry. Moreover, what criteria can be applied to indicate the requirement to assess putative conservation or changes in the binding mode throughout the optimization campaign? Herein, we follow these questions by utilizing a well-suited model system and suggest advice that may support the decision as to whether the (re)-validation of a binding mode is required.

In lead discovery, the aspartic protease endothiapepsin (EP) commonly serves as a model enzyme and has successfully been exploited as a surrogate for renin and β-secretase in structure-based inhibitor design.[3] Based on the Gewald reaction (Scheme 1),[4] we have synthesized a series of substituted 2-aminothiophene-type EP inhibitors that cover a broad range of affinities from two-digit μM up to submicromolar potency. Starting from the two-digit micromolar inhibitors **1** and **2** (Figure 1) we embarked on a structure-guided SAR exploration of this ligand series.



Scheme 1. General synthetic route to the 2-aminothiophene core through the Gewald reaction.

The crystal structure of EP in complex with **1**, determined at 1.30 Å resolution, unambiguously shows the ligand's tryptamine moiety bound into the S_1 pocket, while the thiophene core resides between the S_1' and S_2 pockets (Figure 2a). Key hydrogen bonds are formed between the ligand's sp^3-hybridized nitrogen and the two aspartates of the catalytic dyad (Asp35 and Asp219), mediated by the lytic

[*] M. Kuhnert, A. Shahim, Prof. Dr. W. E. Diederich
Institut für Pharmazeutische Chemie, Philipps-Universität Marburg
Hans-Meerwein-Strasse 3, 35032 Marburg (Germany)
E-mail: wibke.diederich@staff.uni-marburg.de
Dr. H. Köster, Dr. R. Bartholomäus, Dr. A. Y. Park, Prof. Dr. A. Heine, Prof. Dr. G. Klebe
Institut für Pharmazeutische Chemie, Philipps-Universität Marburg
Marbacher Weg 6, 35032 Marburg (Germany)
Dr. H. Steuber[+]
LOEWE-Zentrum für Synthetische Mikrobiologie, Philipps-Universität Marburg
Hans-Meerwein-Straße, 35032 Marburg (Germany)
[+] Current address: Bayer Pharma AG, Lead Discovery Berlin - Structural Biology
Müllerstraße 178, 13353 Berlin (Germany)

[**] We acknowledge the support of the beamline staff at BESSY II, Berlin (Germany), Petra III, Hamburg (Germany), and SLS, Villingen (Switzerland) and travel grants from the Helmholtz-Zentrum für Materialien und Energie, Berlin (Germany).
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201411206.

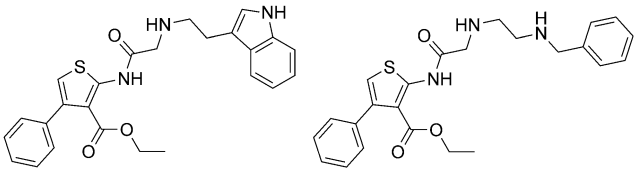


Figure 1. Chemical structures of **1** and **2** and their affinity data against EP.

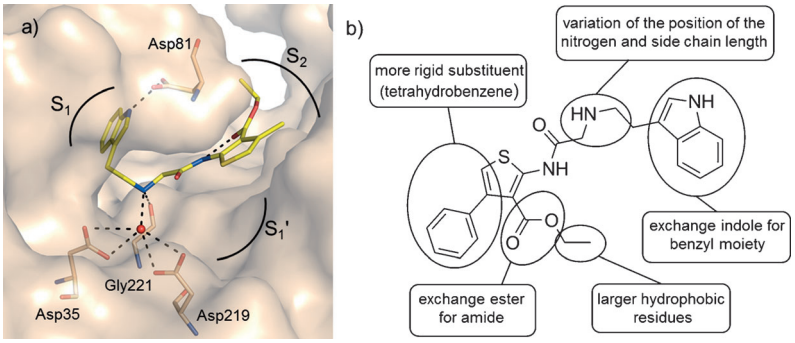


Figure 2. a) Observed binding mode of **1** in complex with EP. Amino acids involved in key hydrogen bonds (dashed lines) are shown as rods. For the detailed interaction pattern, see Figure S5 in the Supporting Information. b) Sites for structural variations derived from compound **1**.

water molecule. Likely, this amine is protonated and serves as a hydrogen-bond donor to the carbonyl group of Gly221. The indole NH contributes a charge-assisted H-bond to Asp81 located at the tip of the hairpin-type flap region. Noteworthy, for the 4-phenyl substituent of the ligand, no properly defined difference electron density could be observed, suggesting a distribution over multiple conformational states in the binding site, and therefore it was omitted from the structural model. Interestingly, the amide NH and the ester carbonyl oxygen of the ligand establish an intramolecular H-bond, keeping the ester and the thiophene plane approximately coplanar (dihedral angle ca. 18°) and the spatially demanding substituents (ethyl ester and phenyl ring) in an eclipsed-type conformation.

Based on the observed binding mode we designed and synthesized a series of compounds with different structural variations to explore the underlying SAR of this inhibitor scaffold (Figure 2b). Via extension of the ethyl ester with larger hydrophobic substituents aimed to address the hydrophobic S_2/S_4 pocket, a gain in affinity was anticipated. In addition, to modulate the electronic properties of the intramolecular H-bond, the influence of an exchange of the ester for an amide moiety was investigated. A thereby induced change in the rotameric state of the carbonyl group might additionally enable a polar contact to Tyr226 (Figure S1 in the Supporting Information). Further sites for variation consisted in changing the length of the tryptamine side chain as well as the positioning of the basic nitrogen. The replacement of the indole by morpholine was considered to enhance solubility. Finally, the observation that the 4-phenyl substituent is disordered prompted us to evaluate a more rigid scaffold such as a tetrahydrobenzothiophene.

The synthesized compounds and their inhibitory activities are shown in Figure 3, with **4** and **9** being the most potent representatives (460 and 545 nM, respectively). However, the interpretation of the underlying SAR turned out to be challenging based on the initially determined binding geometry of **1**. Surprisingly, the simple ester-to-amide exchange resulted in an affinity improvement by a factor of five (**3**, 4.0 μM). Is this gain solely attributed to the changed interaction of the introduced amide function, and how does this relate to the intramolecular H-bond observed in **1**? The replacement of the ethyl moiety in **3** by a benzyl substituent leads to the most potent compound of our series, **4**, with a K_i value of 460 nM and thus an improvement in the inhibitory activity of one order of magnitude. Removal of the 4-phenyl substituent and annulation of a tetrahydrobenzene moiety at the [b] site of the thiophene ring (**6**) reduces affinity by about fourfold, although in **1** this substituent appeared disordered and was not expected to contribute significantly to the establishment of a directional interaction to EP. The replacement of the indole moiety (**4**) by a benzyl group (**5**), presumably accompanied by the loss of the charge-assisted H-bond to Asp81,

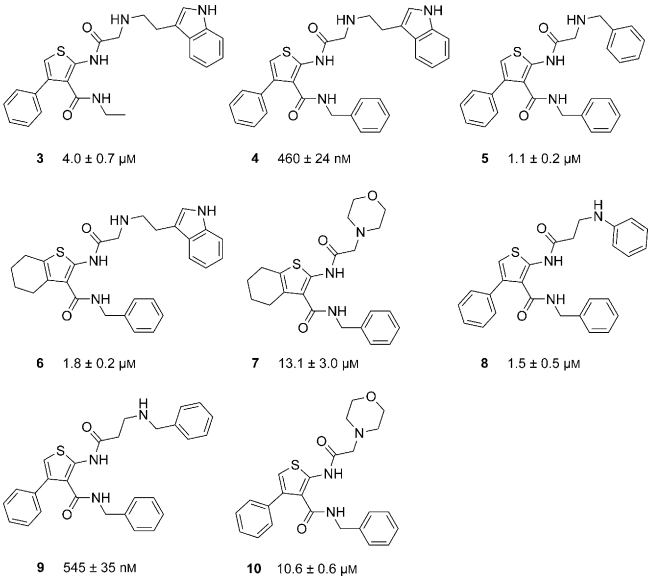


Figure 3. Chemical structures of the synthesized compounds **3**–**10** and their corresponding K_i values against EP.

reduces the affinity by approximately twofold (1.1 μM). A positional shift of the aliphatic nitrogen that was involved as a donor in an H-bond to the lytic water resulted in **8** (1.5 μM) and additionally a one-carbon insertion rendered **9**, which, with an activity of 545 nM, was unexpectedly found to be nearly as potent as **4**.

Overall, this rather incongruent SAR prompted us to perform a comprehensive structural analysis of the underlying

binding modes by crystal structure analysis. To our surprise, the binding modes of the synthesized ligands successfully determined for **1** and **3–9**, split into four clusters of binding poses, and our starting compound **1** turned out to represent a structural singleton, in other words, the sole representative of this binding mode. Figure 4a shows the crystal structures determined in this study and the affiliation for each inhibitor to the corresponding cluster, as well as the superposition of one cluster representative with the binding pose of our initial lead **1**. Among the eight crystal structures, all adopting the same crystal packing, the protein itself manifests nearly identical conformation: rmsd values for the all-atom fitting are below 0.5 Å, backbone fitting yielded rmsd values below 0.3 Å.[5] Hence, the deviating binding modes are neither caused by significant rearrangements of the protein environment, nor are they induced by different crystal forms that might create a different packing environment. Surprisingly, the two most potent compounds belong to different binding mode clusters. The largest cluster (green) comprises four derivatives (**3**, **4**, **5**, and **6**), followed by a two-membered cluster (**8** and **9**) and two structural singletons (**1** and **7**).

Noteworthy, in between the four clusters, the binding modes of the inhibitors with respect to the two catalytic aspartates and the occupancy of the remaining binding pockets differ substantially. The lytic water molecule adopts a versatile role: it acts either as an H-bond donor or an H-bond acceptor or is released from the complex. Inhibitor **1** addresses the Asp35/219 with its sp^3-hybridized nitrogen via the lytic water molecule, the latter acting as an H-bond acceptor (Figure 4a). In the structures of **8** and **9** (purple cluster) the benzyl-substituted amide nitrogen forms a polar contact to the lytic water molecule, even though the interaction geometry appears rather unfavorable for a typical H-bond. Inhibitor **7** interacts through its morpholinomethylene-substituted carbonyl oxygen mediated by the lytic water molecule with the catalytic aspartates; in this case, however, the water molecule acts as an H-bond donor. All representatives of the green cluster directly address the catalytic dyad, while the lytic water molecule is released upon binding of the inhibitors. The significantly different binding modes provoke substantially deviating H-bond networks between the inhib-

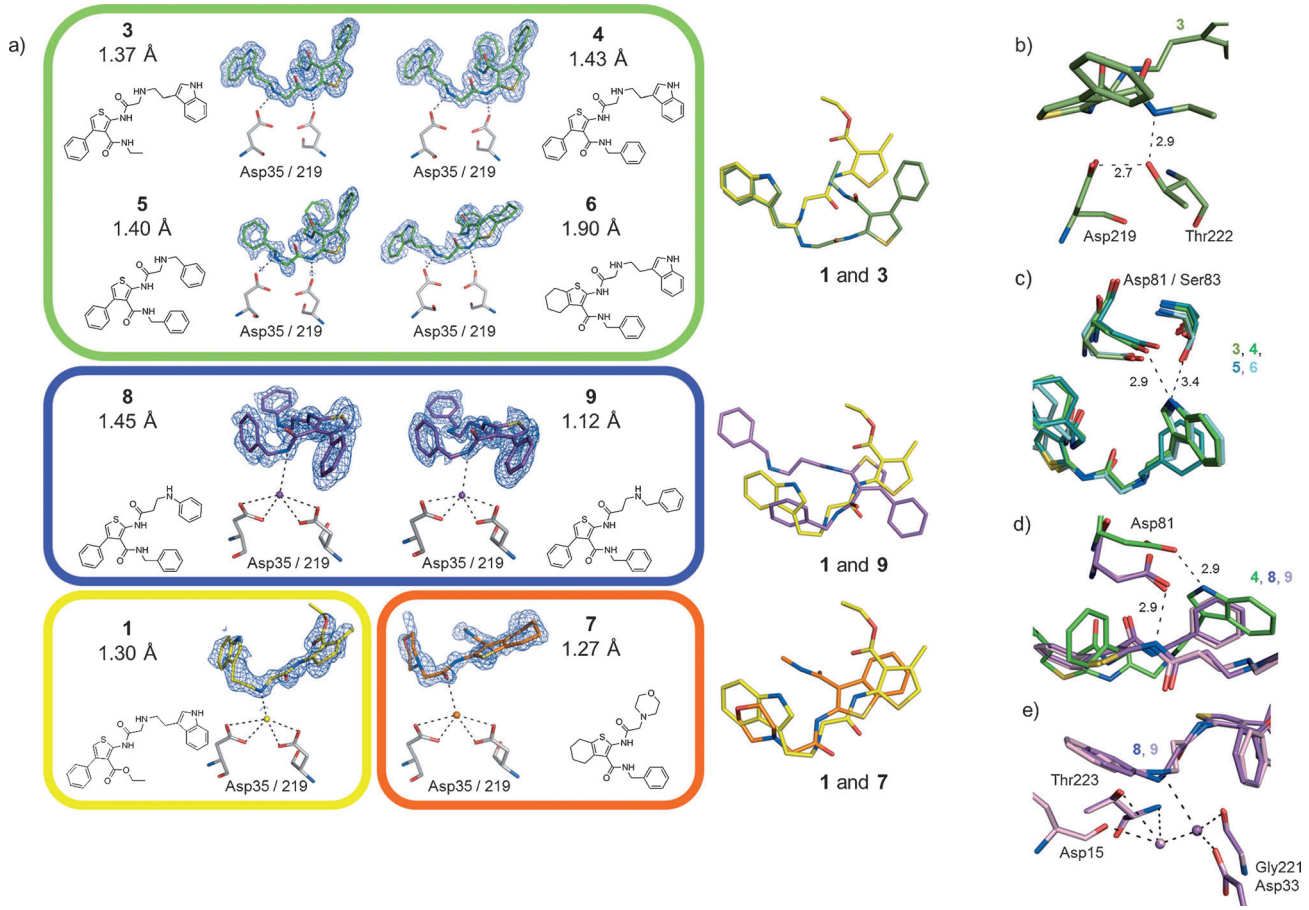


Figure 4. a) Left: Binding modes and assignment to binding mode clusters. Green, purple, yellow, and orange frames indicate the four different binding modes. Resolutions of the X-ray structures are shown in Å. The $2F_o-F_c$ electron densities are depicted for the ligands as a blue mesh at 1σ. Right: Superposition of **1** (yellow) with one representative of each binding mode cluster in the active site. b) H-bond of the amide NH of **3** to Thr222 Oγ. Likely Thr222 forms an H-bond as an H-bond donor to Asp219 and therefore can only act as an H-bond acceptor to the ligand. Distances in Å. c) H-bond of the indole NH of **3**, **4**, and **6** to Asp81 and Ser83. This interaction is not possible for **5** due to the replacement of the indole by benzyl. d) The charge-assisted H-bond interaction to Asp81, which in the green cluster (here shown for **4**) is formed with the indole NH, is established in **8** and **9** by means of the 2-amino function of the thiophene. e) H-bond network of **8** and **9** formed with the ligands' aniline NH or benzyl NH, respectively. Involved water molecules are shown in spheres.

itors and the protein, which are shown in Figures S3–S6 in the Supporting Information.

The binding mode of **1** apparently relies on the presence of the ester moiety, whereas the amide derivative (**3**) adopts an altered binding mode. Interestingly, the intramolecular H-bond observed in **1** (Figure 2a) breaks up in favor of the establishment of an H-bond to Thr222 Oγ through the ethyl-substituted amide nitrogen. How does the mere replacement of the ester by an amide translate into such a predominant change in binding geometry? In all the EP crystal structures determined in our study, the interaction pattern between Thr222 Oγ and the Asp219 carboxylate (2.7 Å) is conserved . Likely, this Thr side chain donates an H-bond to the carboxylate of Asp219 keeping the proton of Thr222 Oγ oriented towards the aspartate (Figure 4b). This observation supports the conclusion that Thr222 Oγ itself can only be involved in favorable polar interactions if addressed by appropriate H-bond donors on the ligand site. This is impossible for the ester **1** as its oxygen can only serve as an H-bond acceptor; hence the ester prefers to saturate its polar contact inventory by an intramolecular H-bond. However, when the ester is changed to an amide (**3**), the introduced NH function is able to address Thr222 Oγ as a donor rather than maintaining the intramolecular H-bond, while the concomitantly nonsaturated glycine amide NH rotates by ≈100° to establish an H-bond to the Asp219 carboxylate (Figure 4a).

Although the overall SAR interpretation is inconclusive in the absence of structural data, the SAR within the clusters reveals a consistent picture, which will be discussed only exemplarily. In the green cluster, **3** bearing an ethyl group at P_2 has the lowest affinity (4.0 μM); this most likely results from the reduced hydrophobic interactions, compared to those of the inhibitors equipped with a benzyl group at this position. Furthermore, **3** and **6** lack the intramolecular π–π interaction between the two phenyl rings observed in **4** and **5**, which leads to a loss of preorganization corresponding to the bound conformation. In **5**, due to the exchange of the indole moiety for the benzyl group, the H-bond to Asp81 and the weak polar contact to Ser83 as observed in **3**, **4**, and **6** cannot be established (Figure 4c). In **4**, the representative with the highest affinity (460 nM) of this cluster, the four varied substituents seem to be optimally chosen among all investigated variations, as the indole moiety is able to form an H-bond to Asp81 whereas the benzyl moiety interacts via π–π stacking to Tyr226. Inhibitor **8** (purple cluster) differs from **5** (green cluster) only by the position of the secondary amino nitrogen in the P_1 substituent of the inhibitor. Based on the X-ray structure of **5**, a lower affinity caused by a less efficient interaction of the relocated nitrogen to the catalytic aspartates would have been expected for **8**. Interestingly though, its affinity is nearly unchanged. The X-ray structure of **8** provides the explanation for this initially surprising observation: The switch in binding geometry between **5** and **8** seems to be mainly triggered by the loss of the H-bond to Asp35 that would be expected for an unchanged geometry of **5**. The thiophene moiety formerly located in the S_1' pocket (green cluster) flips by about 180° in order to populate now the S_2 pocket. Likewise, inhibitor **9**, in which the P_1 substituent is extended by one carbon atom, shows this altered binding geometry, which enables **8** and **9** to maintain the charge-assisted H-bond interaction to Asp81, although the indole moiety that previously accomplished this task is replaced by the phenyl/benzyl group with the amide NH in thiophene 2-position (Figure 4d). Even though the aryl- and benzyl-substituted nitrogens are located at virtually identical positions, the increase in affinity by about threefold from **8** to **9** might be explained by the different pK_a values of the aniline NH compared to the benzyl NH. Thus, the latter, which is being most likely protonated to a larger extent, is capable of forming a more efficient charge-assisted H-bond to the water molecule that mediates the H-bond to Asp33 and Gly221 (Figure 4e).

In our inhibitor series also less or inactive compounds were synthesized which are essential for SAR interpretations as well. However, in the present case the interpretation of these data is almost impossible, as the assignment to one of the binding mode clusters remains speculative.

The current example highlights the complexity of binding events and their strong dependence on seemingly minor effects of scaffold decoration and modifications. That **1**, **7**, **8**, and **9** do not adopt the same binding geometry as observed in the green cluster can be rationalized retrospectively, but without the crystal structure determinations the adopted binding modes would have been difficult to predict, and, if so, would have hardly been believed or attracted sustained attention without our experimental evidence.

Such an unexpected flipping of binding modes after modification of the scaffold substitutions has been detected sporadically, however, mostly by serendipity.[6] As the commonly accepted hypothesis that similar ligands bind in a similar fashion[7] apparently does not hold true on a comprehensive basis, the question arises, how frequently a (re-)validation of an assumed binding mode is required. Are there any easily available indicative hints to estimate whether a crystallographic (re-)validation of the assumed binding mode used as working hypothesis is required throughout the lead exploration campaign?

In addition to the affinity determination via a fluorescence-based assay (K_i values), we performed a thermal shift assay (TSA) to investigate the extent to which the studied compounds stabilize the protein, especially with respect to the different binding modes. TSAs are commonly used for hit identification and affinity ranking of inhibitors.[8,9]

The results for our inhibitor series (Table 1) clearly suggest that the observed shifts in the melting temperature do not generally correlate with the K_i values. However, within one binding mode cluster, the ΔT_m values are ranked correctly, with **5** being an exception: This might be attributed to the loss of the charge-assisted H-bond to Asp81 due to the absence of the indole moiety, while the other members of this cluster establish this interaction. This observation is in agreement with the concept that TSA measurements require similar binding enthalpies for correct affinity ranking.[9] However, how do TSA data perform for ligands of identical chemotype that fall in the same affinity range, but exhibit different binding geometries?

Only two of our four binding mode clusters qualify for such a comparison, as the singletons **1** and **7** are slightly off-

Table 1: TSA data ordered by decreasing ΔT_m values, as well as assignment to the binding mode clusters (color coded) and K_i values.

Inhibitor [100 μM]	T_m [°C]	ΔT_m [°C]	K_i
DMSO	58.7 ± 0.2		
4	63.3 ± 0.2	4.6	460 ± 24 nM
9	62.5 ± 0.0	3.8	545 ± 35 nM
6	61.9 ± 0.1	3.2	1.8 ± 0.2 μM
3	61.5 ± 0.2	2.8	4.0 ± 0.7 μM
5	60.9 ± 0.1	2.2	1.1 ± 0.2 μM
1	59.4 ± 0.2	0.7	18.2 ± 7.4 μM
8	59.1 ± 0.1	0.4	1.5 ± 0.5 μM
7	58.8 ± 0.1	0.1	13.1 ± 3.0 μM

range with respect to their affinity. However, the affinities of the green and the purple clusters cover a similar affinity range thus permitting such a comparison. Interestingly, for these two clusters, ligands of rather identical affinity deviate with respect to their associated TSA profile: The two most potent ligands **4** and **9** possess affinities of 460 nM and 545 nM, respectively, but differ by 0.8 °C in their TSA profiles (4.6 versus 3.8 °C) which corresponds to a deviation of more than threefold of the standard deviation of the measurement. Analogously, **5** and **6** inhibit EP with a virtually identical affinity of 1.1 and 1.8 μM, respectively, and **8** is equally potent (1.5 μM). However, the latter exhibits a TSA shift of only 0.4 °C, while the first two ligands significantly enhance the stability of the protein as indicated by a shift of ΔT_m of 2.2 and 3.2 °C, respectively (with **5** being the abovementioned internal outlier within the green cluster). These observations suggest that protein stabilization by a particular ligand depends not only on affinity, but also on its adopted interaction pattern. This is also true for the comparison of **1** and **8**, both of which exhibit a similar TSA shift (0.7 vs. 0.4 °C, respectively) but possess a substantial difference in affinity (18 vs. 1.5 μM), thus suggesting that they stabilize the target protein by deviating interaction geometries.

Hence, discrepancies between affinity data based on enzyme kinetics and TSA results within one inhibitor series may be exploited as a hint for putative changes in the adopted binding geometry, which may qualify the seeked "outlier"-type ligands as promising candidates for revalidation of the assumed binding mode.

The presented example impressively points out that minor structural changes within one lead series, such as the replacement of an ester by an amide or the introduction of an additional carbon atom, may result in entirely different binding modes. An inconsistent affinity–TSA correlation may serve as an easily available indicator to qualify candidates for structural revalidation, even if the interaction mode is believed to be known. In the absence of such structural data, a reasonable interpretation of substituent effects or assumed bioisosteric replacements at a given chemotype will be rather meaningless and will likely drive scaffold optimization in the wrong direction. We expect that our case study and the application of the suggested validation protocol will encourage drug designers to further pursue and study chemically attractive scaffolds albeit an at first glance seemingly inconclusive SAR. This is likely to be rewarded with a more comprehensive understanding and an increasing number of starting options for further lead optimization.

Experimental Section

Protein Databank codes for this manuscript: **1** (3T7Q), **3** (4L6B), **4** (3PSY), **5** (3WZ6), **6** (3WZ7), **7** (3T7X), **8** (3WZ8), **9** (4LAP).

Received: November 19, 2014
Published online: January 28, 2015

Keywords: drug design · enzyme inhibitors · ligand–enzyme binding · structure determination · thermal shift assay (TSA)

[1] a) *Structure-Based Drug Discovery* (Ed.: R. E. Hubbard), Royal Society of Chemistry, Cambridge, **2006**; b) A. C. Anderson, *Chem. Biol.* **2003**, *10*, 787 – 797.
[2] a) A. Wlodawer, J. Vondrasek, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 249 – 284; b) M. von Itzstein, *Nat. Rev. Drug Discovery* **2007**, *6*, 967 – 974; c) S. Venkatraman, *Trends Pharmacol. Sci.* **2012**, *33*, 289 – 294; d) A. K. Ghosh, *Structure-based Design of Drugs and Other Bioactive Molecules. Tools and Strategies*, Wiley-VCH, Weinheim, **2014**.
[3] a) T. L. Blundell, J. Cooper, S. I. Foundling, D. M. Jones, B. Atrash, M. Szelke, *Biochemistry* **1987**, *26*, 5585 – 5590; b) S. Geschwindner, L.-L. Olsson, J. S. Albert, J. Deinum, P. D. Edwards, T. de Beer, R. H. A. Folmer, *J. Med. Chem.* **2007**, *50*, 5903 – 5911.
[4] K. Gewald, E. Schinke, H. Böttcher, *Chem. Ber.* **1966**, *99*, 94 – 100.
[5] a) A. Martin, http://www.bioinf.org.uk/software/profit/; b) A. D. McLachlan, *Acta Crystallogr. Sect. A* **1982**, *38*, 871 – 873.
[6] a) G. Chi, A. Manos-Turvey, P. D. O'Connor, J. M. Johnston, G. L. Evans, E. N. Baker, R. J. Payne, J. S. Lott, E. M. M. Bulloch, *Biochemistry* **2012**, *51*, 4868 – 4879; b) S. H. Reich, M. Melnick, J. F. Davies II, K. Appelt, K. K. Lewis, M. A. Fuhry, M. Pino, A. J. Trippe, D. Nguyen, H. Dawson, B.-W. Wu, L. Musick, M. Kosa, D. Kahil, S. Webber, D. K. Gehlhaar, D. Andrada, B. Shetty, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3298 – 3302; c) H. Steuber, M. Zentgraf, C. La Motta, S. Sartini, A. Heine, G. Klebe, *J. Mol. Biol.* **2007**, *369*, 186 – 197; d) C. P. Mpamhanga, D. Spinks, L. B. Tulloch, E. J. Shanks, D. A. Robinson, I. T. Collie, A. H. Fairlamb, P. G. Wyatt, J. A. Frearson, W. N. Hunter et al., *J. Med. Chem.* **2009**, *52*, 4454 – 4465.
[7] J. Boström, A. Hogner, S. Schmitt, *J. Med. Chem.* **2006**, *49*, 6716 – 6725.
[8] a) M. W. Pantoliano, E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B. A. Springer, P. Lane et al., *J. Biomol. Screening* **2001**, *6*, 429 – 440; b) R. Zhang, F. Monsma, *Curr. Opin. Drug Discovery Dev.* **2010**, *13*, 389 – 402; c) E. V. Bobkova, M. J. Weber, Z. Xu, Y.-L. Zhang, J. Jung, P. Blume-Jensen, A. Northrup, P. Kunapuli, J. N. Andersen, I. Kariv, *J. Biol. Chem.* **2010**, *285*, 18838 – 18846; d) P. Śledź, S. Lang, C. J. Stubbs, C. Abell, *Angew. Chem. Int. Ed.* **2012**, *51*, 7680 – 7683; *Angew. Chem.* **2012**, *124*, 7800 – 7803.
[9] M.-C. Lo, A. Aulabaugh, G. Jin, R. Cowling, J. Bard, M. Malamas, G. Ellestad, *Anal. Biochem.* **2004**, *332*, 153 – 159.

