ChemPubSoc Europe

DOI: 10.1002/cmdc.201500151

CHEMMEDCHEM Full Papers

# Active Site Mapping of Human Cathepsin F with Dipeptide Nitrile Inhibitors

Janina Schmitz,[a, b] Norbert Furtmann,[a, c] Moritz Ponert,[a] Maxim Frizler,[a] Reik Löser,[a, d] Ulrike Bartz,[b] Jürgen Bajorath,[c] and Michael Gütschow*[a]

Cleavage of the invariant chain is the key event in the trafficking pathway of major histocompatibility complex class II. Cathepsin S is the major processing enzyme of the invariant chain, but cathepsin F acts in macrophages as its functional synergist which is as potent as cathepsin S in invariant chain cleavage. Dedicated low-molecular-weight inhibitors for cathepsin F have not yet been developed. An active site mapping with 52 dipeptide nitriles, reacting as covalent–reversible inhibitors, was performed to draw structure–activity relationships for the non-primed binding region of human cathepsin F. In a stepwise process, new compounds with optimized fragment combinations were designed and synthesized. These dipeptide nitriles were evaluated on human cysteine cathepsins F, B, L, K and S. Compounds **10** (*N*-(4-phenylbenzoyl)-leucylglycine nitrile) and **12** (*N*-(4-phenylbenzoyl)leucylmethionine nitrile) were found to be potent inhibitors of human cathepsin F, with $K_i$ values $<10$ nM. With all dipeptide nitriles from our study, a 3D activity landscape was generated to visualize structure–activity relationships for this series of cathepsin F inhibitors.

## Introduction

Cysteine cathepsins are lysosomal proteases that play crucial roles in various physiological processes, such as bone remodeling, osteoporosis, neurological disorders, autoimmune response and cancer, and represent the largest and best-characterized group of cathepsins.[1,2] One of its representatives, cathepsin F, is predominantly expressed in macrophages, but not in dendritic cells or B cells.[3,4] As other cysteine cathepsins, cathepsin F is translated as an inactive precursor, targeted to the lysosomes and activated by cleavage and dissociation from its pro-region. It is an active endopeptidase and contains a cystatin-like domain in the elongated N-terminal pro-region, a unique feature among the cysteine cathepsins.[5]

The major histocompatibility (MHC) class II pathway allows the presentation of exogenous antigens via the endosomal route. MHC class II is an $\alpha/\beta$ heterodimer synthesized in the rough endoplasmic reticulum (RER) and is assembled with a type II transmembrane protein referred to as the chaperone invariant chain. MHC class II mediated antigen presentation requires the participation of endosomal and lysosomal proteases for two proteolytic events: 1) cleavage of the antigens to small antigenic peptides and 2) degradation of the invariant chain. The MHC class II binding site is blocked by the invariant chain, prohibiting preterm peptide loading. The complex traffics from the RER through the trans-Golgi network to the early endosomes. During transport, the invariant chain is degraded by endosomal and lysosomal proteases. The remaining fragment of the invariant chain in the MHC II binding groove, lip10, is cleaved to a class II-associated invariant chain peptide (CLIP) by certain cysteine proteases. These CLIP fragments can be replaced by antigenic peptides. Loaded MHC class II molecules are transported to the cell surface, allowing the display of antigenic peptides to T cells. When the processing of lip10 to CLIP is discontinued, the transport of MHC II to the cell surface is delayed due to its retention in the lysosomes. Thus, the cleavage of lip10 is thought to be the key event in the antigen-presentation pathway of MHC class II complexes. The major processing enzyme in dendritic cells, macrophages, and B cells in this pathway is cathepsin S, but cathepsin L can carry out this function in thymic epithelial cells.[4,6] Notably, it was observed that macrophages of mice, deficient in cathepsins S and L, were still capable of processing the invariant chain and loading peptides onto MHC class II complexes. It was shown for purified cathepsin F that the enzyme is as potent as cathepsin S in invariant chain processing and CLIP formation.[6] Moreover, with an inhibitor of the cysteine cathepsin family (**I**, Figure 1),[7,8] invariant chain processing could be fully blocked in $catS^{-/-}/catL^{-/-}$ mice. Vinyl sulfone **I** is selective for

[a] J. Schmitz, N. Furtmann, M. Ponert, Dr. M. Frizler, Dr. R. Löser, Prof. Dr. M. Gütschow
Pharmaceutical Institute, Pharmaceutical Chemistry I
University of Bonn, An der Immenburg 4, 53121 Bonn (Germany)
E-mail: guetschow@uni-bonn.de

[b] J. Schmitz, Prof. Dr. U. Bartz
Department of Natural Sciences
University of Applied Sciences Bonn-Rhein-Sieg
von-Liebig-Straße 20, 53359 Rheinbach (Germany)

[c] N. Furtmann, Prof. Dr. J. Bajorath
Department of Life Science Informatics, B-IT
LIMES Program Unit Chemical Biology and Medicinal Chemistry
University of Bonn, Dahlmannstraße 2, 53113 Bonn (Germany)

[d] Dr. R. Löser
Helmholtz-Zentrum Dresden-Rossendorf
Institute of Radiopharmaceutical Cancer Research
Bautzner Landstraße 400, 01328 Dresden (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cmdc.201500151.

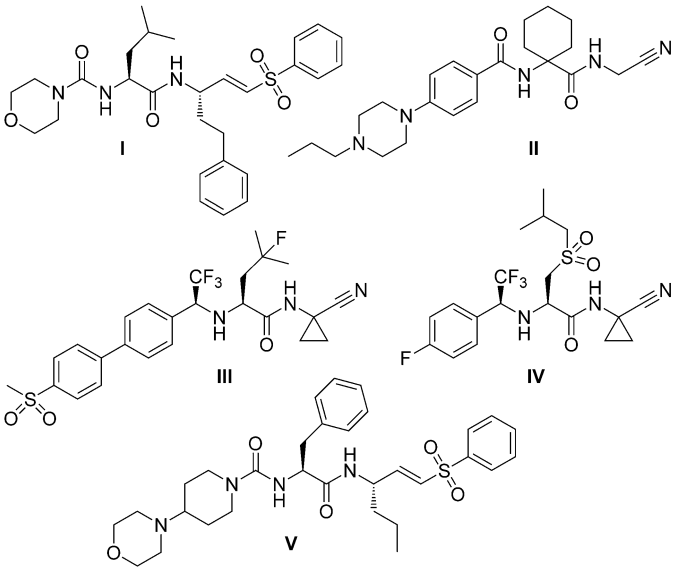



**Figure 1.** Structures of peptidomimetic cathepsin inhibitors.

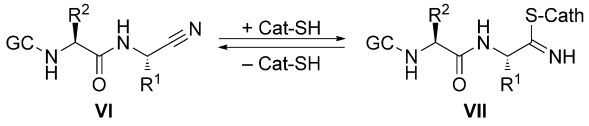
cathepsin S at low concentrations (10 nM), whereas at the higher concentration used, additional cysteine proteases were inhibited, which demonstrated a further cysteine protease, namely cathepsin F, to be involved in the degradation process.[3]

In *cat F*$^{-/-}$ mice, severe neurological disorders were observed starting after 12–16 months, involving incoordination, weakness, and premature death, as well as lysosomal storage defects. Large amounts of lipofuscin, characteristic of neuronal ceroid lipofuscinosis (NCL), were found to accumulate in the central nervous system of cathepsin F-deficient mice.[9] Such storage defects in humans may be induced by cathepsin F gene mutations leading to storage disorders such as Type B Kufs disease, an adult-onset NCL.[10]

Furthermore, cathepsin F has been implicated in atherogenesis. The enzyme appeared to be responsible for the degradation of low-density lipoprotein (LDL) and an enhancement of the binding of LDL to proteoglycans. The resulting generation and accumulation of extracellular lipid droplets in the arterial intima may yield to atherogenesis.[11]

The implication of cathepsin F in various physiological conditions should provide the impetus for the development of low-molecular-weight inhibitors or activity-based probes as possible therapeutic agents or tool compounds, respectively, to further elucidate the physiological role(s) of this protease. Given that cathepsins F and S are the most important processing enzymes in antigen-presenting macrophages, their proteolytic synergism might be addressed in the design of inhibitors as potential therapeutics for macrophage-related disorders. Much effort has been made to develop inhibitors for human cathepsin S, which hold potential for the treatment of autoimmune diseases,[8,12] but little is known about inhibition of cathepsin F by low-molecular weight compounds.

The core of the active site of cathepsin F is characterized by Cys25 and His159, with the nucleophilic thiolate group of the former amino acid attacking the carbonyl atom of the scissile peptide bond, leading to the hydrolytic cleavage of the peptidic substrate in the course of an acyl transfer reaction. Nitrile-based inhibitors can be designed by starting from the specific substrate structure of the target cysteine protease by introducing a nitrile group as reactive warhead in place of the carbonyl group and additional peptidomimetic modifications. Such peptide nitriles and azapeptide nitriles have attracted much attention as highly potent inhibitors.[2,8,13–20] The covalent reaction of peptide nitriles with cysteine cathepsins involves attack of the active-site thiolate at the inhibitor's nitrile carbon atom and reversible formation of a thioimidate adduct (Scheme 1).[2]




**Scheme 1.** Reaction of dipeptide nitriles with cysteine cathepsins and reversible formation of thioimidates (CG = capping group).

It has been shown for several cysteine cathepsins that dipeptide nitriles represent promising therapeutics. Cathepsin K is the major processing enzyme in osteoclastic bone resorption and is responsible for the degradation of bone matrix. Balicatib (**II**), a basic and hence lysosomotropic cathepsin K inhibitor, shows the characteristic features of dipeptide nitrile inhibitors (Figure 1). The replacement of the P3–P2 amide bond by a trifluoroethylamine fragment and the introduction of a 1,1-cyclopropane ring at the P1 position led to the front-runner odanacatib (**III**) with decreased metabolic liabilities. Odanacatib is currently being developed as a once-weekly treatment for osteoporosis.[15,21]

The inhibitor interactions with the S2 and S3 binding sites are essential to achieve selectivity toward other cathepsins. This was also taken into account in the design of cathepsin S inhibitors, among which **IV** (Figure 1) exhibits high potency, selectivity, and bioavailability.[13] This compound was developed by the incorporation of an isobutylsulfonylcysteine building block in P2, and a *para*-fluoro-substituted phenyl residue at the P3 position, advantageous for occupying the deep S2 pocket and the rather small S3 binding site, respectively.

The crystal structure of human cathepsin F has been determined in complex with the vinyl sulfone inhibitor **V**.[22] The structure of cathepsin F adopts a typical fold of papain-like enzymes with some unique features in the S2 and S3 binding sites. The S2 pocket of cathepsin F is defined by Leu67 and Met205; in cathepsin S these amino acids are replaced by two phenylalanines. The methionine residue in cathepsin F points toward the leucine residue, decreasing the size of the binding pocket, while the S3 pocket is wider and shallower than the S2 pocket.[5,22]

In this study, a library of dipeptide nitriles was applied for active-site mapping of the non-primed binding region of human cathepsin F. Stepwise structural modifications were aimed at developing optimized inhibitors and at elucidating structure–activity relationships.

## Results and Discussion

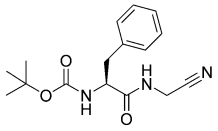
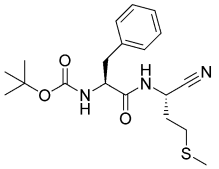
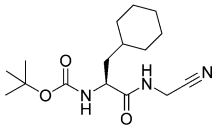
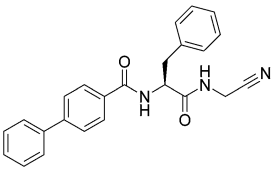
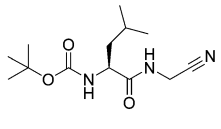
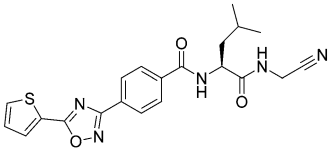
As a starting point of this project, we investigated a library of 52 dipeptide nitriles with various aminonitrile moieties at position P1, different amino acids at position P2, and various capping groups at the N terminus. The general inhibitor structure **VI** is depicted in Scheme 1, and the individual formulae are given in the Supporting Information (Table S1). These dipeptide nitriles have already been evaluated as inhibitors of cathepsins L, S, K, and in some cases cathepsin B,[14,18–20] but have not been investigated as inhibitors of cathepsin F.

Thus, in this study, the library was used to provide initial structure–activity relationship information for dipeptide nitrile inhibitors of cathepsin F. Human cathepsin F was assayed with a fluorogenic peptide substrate, and $K_i$ values were obtained with the Cheng–Prusoff equation. The observed linearity of the progress curves in the presence of the library compounds indicated a 'fast binding' behavior, reflecting rapid formation of the covalent and reversible enzyme–inhibitor thioimidate adduct **VII** (Scheme 1). However, inhibition constants of the 52 dipeptide nitriles span five orders of magnitude (Table S1 in the Supporting Information). For example, Boc-glycylglycine nitrile (**35**, Table S1) was a particularly poor inhibitor, with a $K_i$ value of 1.2 mM, whereas the strongest dipeptide nitrile (**6**, Table 1) exhibited a $K_i$ value of 12 nM. The most significant results of this screening campaign are depicted in Table 1. We considered the $K_i$ values of the six compounds to deduce the influence of structural fragments on the inhibitory activity against human cathepsin F. The data illustrate that cyclohexylalanine or leucine at position P2 resulted in $K_i$ values in the low-micromolar range. These amino acids appeared somewhat more advantageous than phenylalanine (**2** and **3** versus **1**, Table 1). Bulky capping groups at position P3 as present in compounds **5** and **6** clearly increased the potency (**5** versus **1**; **6** versus **3**, Table 1). Moreover, introduction of the methionine side chain caused a small improvement in activity (**4** versus **1**, Table 1).

The following syntheses were carried out to provide further cathepsin F inhibiting dipeptide nitriles in a stepwise optimization process. First, to corroborate that leucine is more favorable at the P2 position than cyclohexylalanine (**3** versus **2**, Table 1), a new pair of Boc-protected dipeptides, both with methionine nitrile at the P1 position, were designed. The preparation of these compounds, **7** and **8**, is depicted in Scheme 2. Boc-methionine (**13**) was treated with ammonia in a mixed-anhydride procedure, and the resulting carboxamide **14** was deprotected and subsequently coupled with Boc-protected cyclohexylalanine or leucine to obtain dipeptides **16** or **17**, respectively. The nitrile group of the desired products **7** and **8** was formed upon treatment with cyanuric chloride.[23] The kinetic data of **7** and **8** are listed in Table 2. The corresponding $K_i$ values for cathepsin F inhibition confirmed cyclohexylalanine to be a more suitable P2 building block than phenylalanine (**7**, Table 2 versus **4**, Table 1), but to be less advantageous than leucine (**7** versus **8**, Table 2). The two new methionine derivatives were more active than their glycine counterparts (**7** and **8**, Table 2 versus **2** and **3**, Table 1).

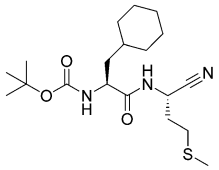
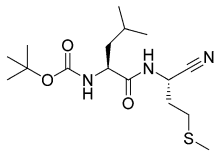
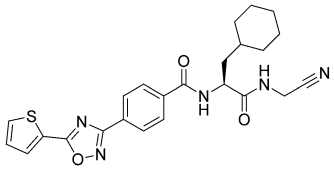
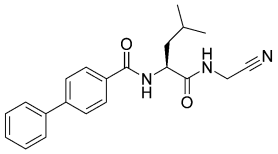
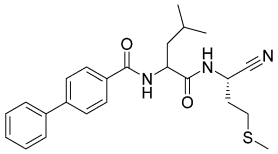
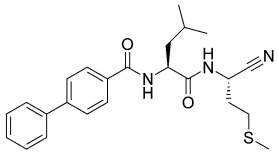
To obtain new dipeptide inhibitors with the bi- or triaryl capping groups present in **5** and **6**, the required carboxylic acids **21** and **24** were prepared (Scheme 3). Synthesis of building block **21** involved acylation of an intermediate amidoxime to **20** and a cyclocondensation to produce the internal oxadiazole

**Table 1.** Inhibition of human cathepsin F by dipeptide nitriles **1**–**6**.

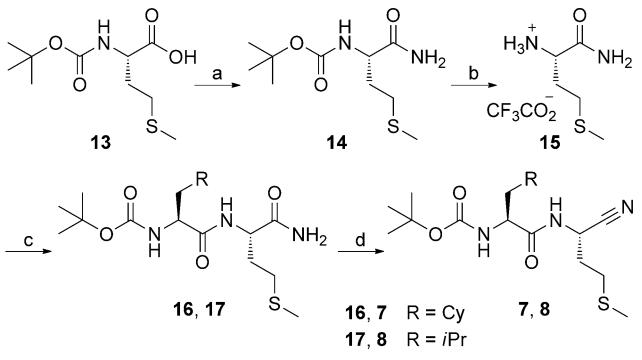
| Compd | | $K_i$ [μM][a] | Compd | | $K_i$ [μM][a] |
|---|---|---|---|---|---|
| **1** |  | 8.48 ± 0.35 | **4** |  | 4.36 ± 0.19 |
| **2** |  | 4.40 ± 0.15 | **5** |  | 0.687 ± 0.046 |
| **3** |  | 1.50 ± 0.04 | **6** |  | 0.0124 ± 0.0006 |

[a] $IC_{50}$ values were obtained from duplicate measurements in the presence of five different inhibitor concentrations. Progress curves were followed over 20 min and analyzed by linear regression. $IC_{50}$ values were determined by nonlinear regression using the equation: $v_s = v_0/(1 + [I]/IC_{50})$, in which $v_s$ is the steady-state rate, $v_0$ is the rate in the absence of inhibitor, and [I] is the inhibitor concentration. Standard error of the mean (SEM) values refer to this nonlinear regression. $K_i$ values ± SEM were calculated from $IC_{50}$ values by applying the equation $K_i = IC_{50}/(1 + [S]/K_M)$, in which [S] is the substrate concentration.

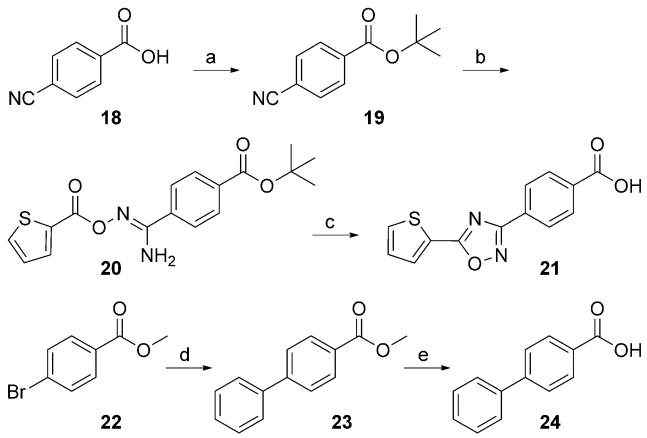
**Table 2.** Inhibition of human cathepsins by dipeptide nitriles **7**–**12**.

| | Compd | $K_i$ [μM][a] cat F | cat B | cat L | cat K | cat S |
|---|---|---|---|---|---|---|
| **7** |  | 1.33 ± 0.03 | 16.5 ± 0.8 | 1.88 ± 0.13 | 2.62 ± 0.35 | 0.203 ± 0.008 |
| **8** |  | 0.510 ± 0.027 | 42.6 ± 1.6 | 0.854 ± 0.085 | 0.0288 ± 0.0014 | 0.0806 ± 0.0052 |
| **9** |  | 0.139 ± 0.016 | 4.94 ± 1.31 | 0.781 ± 0.116 | 0.154 ± 0.011 | 0.106 ± 0.009 |
| **10** |  | 0.00779 ± 0.00074 | 1.79 ± 0.14 | 1.36 ± 0.07 | 0.000561 ± 0.000048 | 0.0387 ± 0.0032 |
| **11** |  | 0.0118 ± 0.0010 | 0.801 ± 0.043 | 1.03 ± 0.05 | 0.000479 ± 0.000035 | 0.0347 ± 0.0016 |
| **12** |  | 0.00728 ± 0.00063 | 0.333 ± 0.025 | 0.563 ± 0.067 | 0.000261 ± 0.000019 | 0.0171 ± 0.0021 |

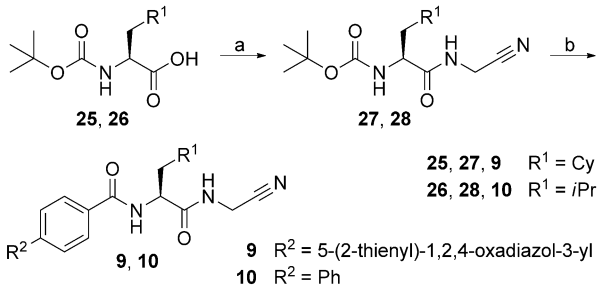
[a] See Table 1 footnotes for details.



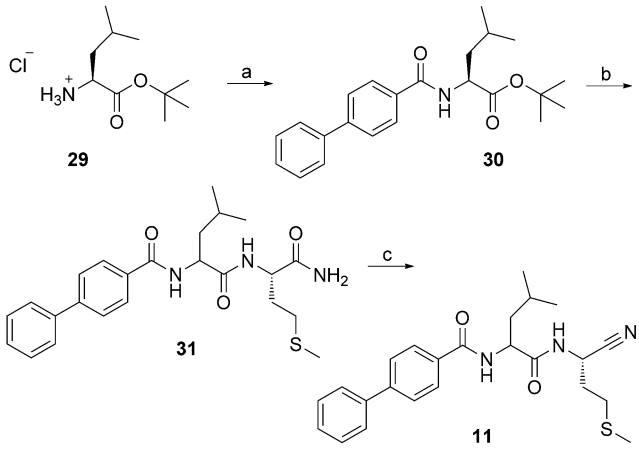

**Scheme 2.** Synthesis of **7** and **8**. *Reagents and conditions*: a) 1. *N*-methylmorpholine (NMM), $ClCO_2iBu$, THF, 2. $NH_3$, −25 °C→RT, overnight, 79%; b) trifluoroacetic acid (TFA), $CH_2Cl_2$, 0 °C, 2 h, quant.; c) Boc-Cha-OH or Boc-Leu-OH, *N*,*N*-diethylisopropylamine (DIPEA), *N*-[(dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminiumhexa-fluoro-phosphate *N*-oxide (HATU), $CH_2Cl_2$, RT, overnight, 99% (**16**), 56% (**17**); d) cyanuric chloride, DMF, RT, 4 h, 62% (**7**), 28% (**8**).



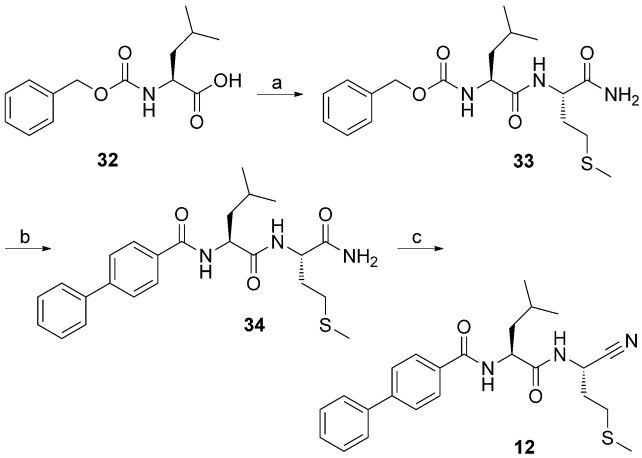

**Scheme 3.** Synthesis of **21**[18] and **24**. *Reagents and conditions*: a) 1. oxalyl chloride, $CH_2Cl_2$, DMF, RT, 1 h, 2. *tert*-butanol, pyridine, RT, 6 h, 77%; b) 1. $NH_2OH{\cdot}HCl$, $Et_3N$, EtOH, reflux, 3 h, 2. 2-thenoyl chloride, $Et_3N$, MeCN, RT, 5 h, 76%; c) 1. AcOH, 80 °C, overnight, 2. TFA, $CH_2Cl_2$, RT, 6 h, 83%; d) 1. $Pd(PPh_3)_4$, 1,2-dimethoxyethane (DME), RT, 10 min, 2. phenylboronic acid, $Na_2CO_3$, 120 °C, 4 h, 52%; e) LiOH, THF/$H_2O$ (2:1), 0 °C→RT, 3 h, 99%.




**Scheme 4.** Synthesis of **9** and **10**. *Reagents and conditions*: a) 1. NMM, $ClCO_2iBu$, THF, 2. $H_2NCH_2CN{\cdot}0.5\,H_2SO_4$, 2 N NaOH, −25 °C→RT, overnight, 32% (**27**), 42% (**28**); b) 1. methanesulfonic acid, THF, 0 °C→RT, overnight, 2. *N*-4-[5-(2-thienyl)-1,2,4-oxadiazol-3-yl]benzoic acid (**21**) or biphenyl-4-carboxylic acid (**24**), EDC, DMAP, DIPEA, THF, RT, overnight, 22% (**9**), 30% (**10**).




**Scheme 5.** Synthesis of **11**. *Reagents and conditions*: a) biphenyl-4-carboxylic acid (**24**), DIPEA, HATU, $CH_2Cl_2$, RT, overnight, 73%; b) 1. TFA, $CH_2Cl_2$, 0 °C, 2 h, 2. compound **15**, DIPEA, HATU, $CH_2Cl_2$, RT, overnight, 81%; c) cyanuric chloride, DMF, RT, 4 h, 33%.




**Scheme 6.** Synthesis of **12**. *Reagents and conditions*: a) compound **15**, DIPEA, HATU, $CH_2Cl_2$, RT, overnight, 50%; b) 1. $H_2$, 10% Pd/C, MeOH, RT, 7 h, 2. biphenyl-4-carboxylic acid (**24**), DIPEA, HATU, $CH_2Cl_2$, RT, overnight, 95%; c) cyanuric chloride, DMF, RT, 4 h, 31%.

ring,[18] and **24** was produced through Suzuki coupling.[24] Boc-cyclohexylalanine (**25**) and Boc-leucine (**26**) were reacted with aminoacetonitrile to form dipeptide nitriles **27** and **28** (Scheme 4), which were then deprotected with TFA and finally coupled with the aforementioned carboxylic acids **21** or **24**. By combination of the triaryl capping group and the cyclohexylalanine P2 building block, we obtained the dipeptide nitrile **9** and, by combination of the biphenyl group and leucine, the dipeptide nitrile **10**. With **9** and **10** in hand and after evaluating their cathepsin F inhibitory activities (Table 2), further structure–activity relationships could be drawn. Importantly, the corresponding $K_i$ values revealed that the biphenyl group is more favorable than the triaryl capping group (**10**, Table 2 versus **6**, Table 1). A comparison of the two triaryl-containing analogues confirmed the leucine–cyclohexylalanine exchange to be disadvantageous, this time leading to a 10-fold loss in potency (**6**, Table 1 versus **9**, Table 2). Obviously, when accommodating an inhibitor with a more voluminous P3 moiety, cathepsin F became more susceptible to a change in the P2 substructure. Such a trend was also observed for the biphenyl pair, in which the expected improvement due to replacement of phenylalanine at the P2 position with leucine led a 100-fold lower $K_i$ value (**10**, Table 2 versus **5**, Table 1). Moreover, **9** and **10** (Table 2) are clearly more active than the Boc-capped analogues **2** and **3** (Table 1).

Next, we intended to examine whether glycine nitrile might be favorably replaced by methionine nitrile. Leucine *tert*-butyl ester was acylated with biphenyl-4-carboxylic acid to obtain **30** (Scheme 5). This amino acid derivative was deprotected and activated with HATU. The activated intermediate could easily undergo oxazolone formation because of the nucleophilic potential of the acyl oxygen, leading to racemization. Thus, the coupling reaction with methionine amide trifluoroacetate (**15**) produced the diastereomeric mixture of carboxamide **31**, which was finally dehydrated to the dipeptide nitrile **11**. After verifying the biological activity of **11** (Table 2), the diastereomerically pure compound **12** was synthesized (Scheme 6). In comparison with the aforementioned sequence (introduction of the capping group, peptide coupling, dehydration), the order of the two first steps was reversed. Reaction of the carbamate-protected leucine **32** with **15** yielded the dipeptide **33** without racemization. Hydrogenolytic removal of the Cbz group followed by introduction of the biphenyl capping group and treatment of the resulting **34** with cyanuric chloride afforded the desired dipeptide nitrile **12**. Whereas the diastereomeric mixtures of compounds **31** and **11** (Scheme 5) exhibited double sets of NMR signals, this was not the case for the counterparts **34** and **12** (Scheme 6), indicating their diastereomeric purity. Expectedly, dipeptide **11** had lower activity against cathepsin F and the other cathepsins listed in Table 2 than the (*S*,*S*)-configured analogue **12**.

Dipeptide nitrile **12** showed a $K_i$ value for inhibition of human cathepsin F of 7.3 nM. Although the introduction of the methionine side chain did not further lead to a notable improvement of cathepsin F inhibitory activity (**12** versus **10**, Table 2), it was the most potent cathepsin F inhibitor obtained in the course of this study.

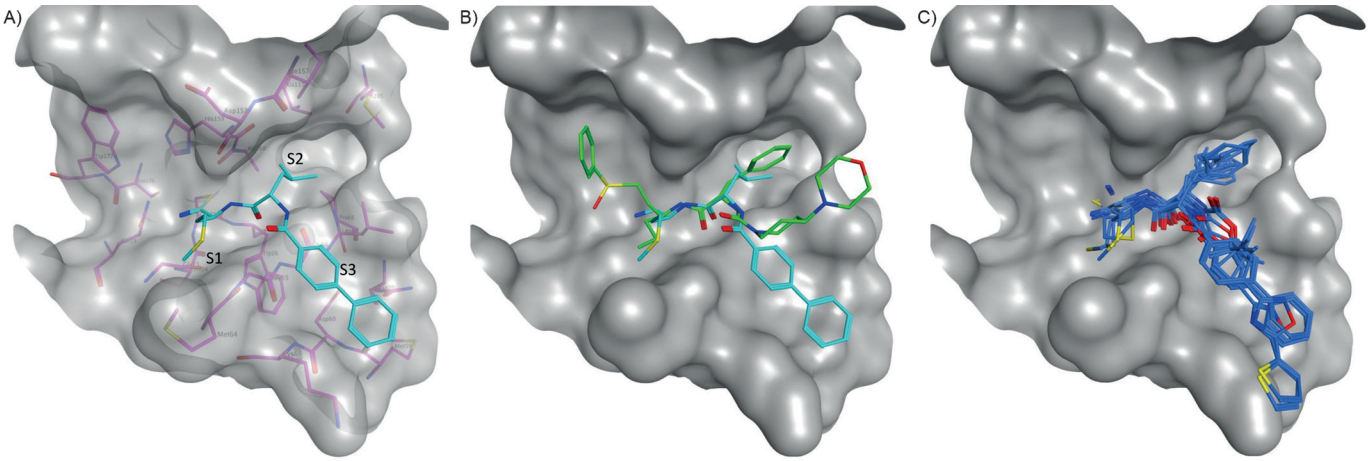



**Figure 2.** Shown are putative binding modes of covalent inhibitors of human cathepsin F. A) The predicted binding mode of compound **12** (cyan) within the active site of cathepsin F. Active site residues are depicted in magenta and the surface is rendered transparent. Binding pockets are annotated according to the Schechter and Berger nomenclature. B) Overlay of the predicted binding mode of **12** (cyan) and the crystallographic vinyl sulfone inhibitor **V** (green). The active site surface is rendered non-transparent. C) Overlay of the predicted binding modes of compounds **1**–**10** and **12** (all in blue) within the active site of cathepsin F using a non-transparent surface representation.

Figure 2A shows the putative binding mode of inhibitor **12** in the active site of human cathepsin F. Through formation of a thioimidate (**VII**, Scheme 1) the nitrile function of the inhibitor is covalently bound to the side chain of residue Cys25. Such a covalent docking approach is not appropriate to provide information about the initial enzyme–inhibitor complex and the orientation of the nitrile warhead relative to the active site cysteine nucleophile, but to model the likely binding pose of the inhibitor within the active site focusing on the complementary occupancy of subsites with substructures of the inhibitor after formation of the covalent enzyme–inhibitor complex. In the case of inhibitor **12**, the thioimidate nitrogen atom is in potential hydrogen bonding distance to the backbone NH group of residue Cys25, while the methionine side chain occupies the S1 pocket. The leucine residue at the P2 position of **12** is directed into the S2 pocket and might form van der Waals contacts with the side chains of residues Leu67, Ala133, and Ala160. The shallow S3 pocket of cathepsin F hosts the large biphenyl moiety of the inhibitor. The side chains of residues Leu67 and Lys61, which determine the width of the S3 pocket,[22] might form lipophilic interactions with the biphenyl moiety. In addition, the backbone of the dipeptide inhibitor is in a favorable position to form several hydrogen bonding interactions to peptide backbone atoms of residues Asp158, Gly65, and Gly66. An overlay of the predicted binding mode of compound **12** and the crystallographic vinyl sulfone inhibitor **V** (Figure 1) is shown in Figure 2B. The backbone peptide structures of both inhibitor compounds display a similar orientation. The propyl substituent at position P1 and the phenyl moiety at position P2 of the crystallographic inhibitor **V** occupy the S1 and S2 pockets of cathepsin F, respectively, in a manner similar to the corresponding side chains in compound **12**. In contrast, the substituents at the P3 position show different orientations. While the biphenyl moiety of inhibitor **12** fits very closely into the S3 pocket, the P3 residue of the crystallographic inhibitor is directed out of the pocket and shows a high degree of solvent exposure. In Figure 2C, an overlay of the putative binding modes of inhibitors **1**–**10** and **12** within the active site of cathepsin F is presented. All compounds adopt a similar orientation such that the S1, S2, and S3 pockets of the enzyme are occupied by the corresponding P1, P2, and P3 substituents of the inhibitor compounds. The methionine side chain at the P1 position is well accommodated in the S1 pocket. This is reflected by the kinetic data; compounds with glycine at this position, which leave the S1 pocket empty, are less active than the methionine analogues. However, these activity differences were not striking. According to the predicted docking poses, the smaller leucine is preferred over phenylalanine and cyclohexylalanine in the narrow S2 pocket, the size of which is determined by the position of the side chains of residues Leu67, Ile157, and Met205.[22] In contrast to the well-defined S2 pocket of cathepsin F, its S3 binding site is spatially less restricted. Accordingly, bulky P3 residues such as the biphenyl substituent of **12** fit nicely into the wide open S3 pocket, which accounted for a further increased binding affinity. Thus, for the compounds presented herein, plausible binding modes were generated that help to rationalize structure–activity relationships. The orientations of the peptide backbones of the inhibitors are very similar to the crystallographic vinyl sulfone inhibitor **V**, and side chains are directed into the S1, S2, and S3 pockets in an equivalent manner.

Inhibitors **7**–**12** were also evaluated at the human cathepsins B, L, K and S (Table 2). None of these dipeptide nitriles was selective for cathepsin F over cathepsin K. In fact, $K_i$ values for cathepsin K inhibition by compounds **8** and **10**–**12** were more than 10-fold lower. These results show that the intended optimization of cathepsin F inhibitors was accompanied by a parallel improvement of their affinity for cathepsin K and reflect a high similarity of the addressed binding region in both cysteine cathepsins. However, cathepsin K is mainly expressed in osteoclasts and hence has a different tissue distribution than cathepsin F.[4] Compounds **7**–**12** exhibited cathepsin L in-

hibitory activity with $K_i$ values that differ by only a factor of 3. By contrast, a clear successive improvement of the inhibitory activity toward cathepsins B and S was observed in ascending order from **7** to **12**. Selectivity for cathepsin F over S is of particular interest, because these two proteases share the catalytic activity to cleave the invariant chain of the MHC II complex, a process which enables antigen presenting cells to display an antigen fragment on their membrane, leading to recognition by T cells and the subsequent events in the immune response. Whereas compounds **7** and **8** with the Boc capping group were more active at cathepsin S, nitrile inhibitors **10** and **12** already exhibited a preference for cathepsin F over S. Further investigations are needed for the development of cathepsin F-selective nitrile-based inhibitors. Such tool compounds are expected to be valuable for investigating the overlapping roles of cathepsins F and S in macrophages.

Three dimensional (3D) activity landscape representations are suitable tools to illustrate structure–activity relationships of large compound data sets. The activity cliff concept has attracted much interest in medicinal chemistry because it can reveal whether small chemical changes lead to significant differences in potency.[25] However, activity landscape modeling and cliff analysis have thus far been rarely applied to peptidic drugs. In the course of this study, a small library of enzyme inhibitors of the same chemotype and hence with a comparably high similarity was analyzed. Figure 3 shows the 3D activity landscape representation for the 57 cathepsin F inhibitors (compound **11** was not included).

The front part of the landscape represents an area of continuous structure–activity relationships. Compounds in this area have low to medium potency against cathepsin F, which is indicated by the green color, and structural changes among these compounds only lead to moderate changes in the biological activity. For example, the phenylalanine residue at the P2 position of compound **1** is replaced by cyclohexylalanine in compound **2**, which leads to a minor change in potency against cathepsin F of about factor 2. Furthermore, the compounds are arranged in clusters that are defined by substituents at the P2 and P3 positions. Compounds in the area of inhibitors **1** and **4** all contain aromatic side chains at the P2 position and a *tert*-butyl residue of the Boc capping group at position P3. In contrast, the region defined by inhibitors **2**, **3**, **7**, and **8** represents compounds with aliphatic and non-aromatic residues at the P2 position and again a P3 *tert*-butyl group. The set of compounds on the lower left side of the landscape with low potency values displays a trace of structure–activity relationship progression to a peak region with the highly active Cbz-leucylglycine nitrile (**40**, Table S1 in the Supporting Information). All compounds involved in this progression have the benzyl residue of the N-terminal Cbz capping group. An area of discontinuous structure–activity relationships is present in the upper part of the landscape. This can be seen, for example, if the N-terminal *tert*-butyloxy group of **3** is replaced by a biphenyl moiety in **10**, improving the potency by a factor of ~200. Compounds in the upper part of the landscape contain large aromatic P3 substituents such as a biphenyl capping group and belong to the most potent inhibitors within this series including **6**, **9**, **10**, and **12**. In this area of the activity landscape, most of the compounds contain leucine at the P2 position and have been synthesized in the course of this study, attempting to structurally optimize cathepsin F inhibitors.

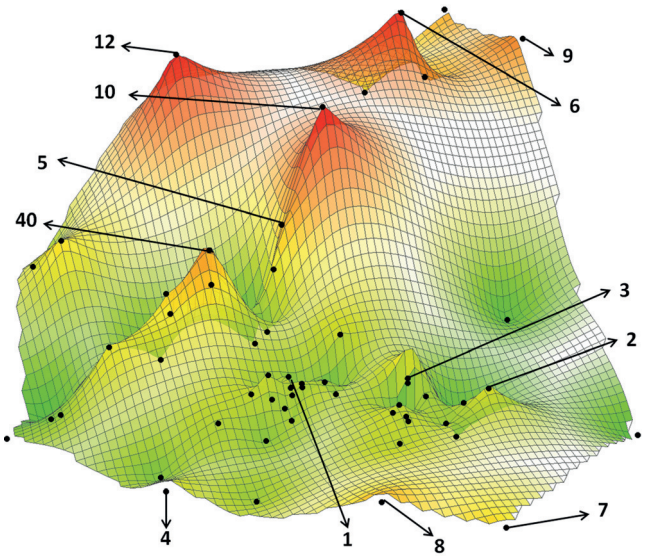



**Figure 3.** 3D activity landscape representation for the cathepsin F inhibitor series. Euclidean fingerprint distances were calculated for the 57 inhibitors using the ECFP2 fingerprint. The surface is colored according to interpolated surface elevation using a continuous spectrum from green (lowest potency) to red (highest potency). Black spheres represent compound data points and are labeled with compound numbers for inhibitors **1**–**10**, **12**, and **40**.

## Conclusions

Despite the involvement of human cathepsin F in important physiological and pathophysiological processes, this enzyme has thus far not been a focal point of drug development efforts. This prompted us to perform an active site mapping by employing a library of peptide nitriles known to interact with cysteine cathepsins in a covalent and reversible manner. A stepwise structural optimization afforded two highly potent cathepsin F inhibitors, **10** and **12**, with $K_i$ values $<10$ nM. The structure–activity relationships obtained for 57 dipeptide nitriles have been illustrated by 3D activity landscape representations. The inhibitors' affinities clearly reflect the characteristics of the non-primed binding region of human cathepsin F. Our data and, in particular, the structures of inhibitors **10** and **12** might represent a starting point for the future design of therapeutics against autoimmune diseases. Moreover, these peptidomimetic inhibitors are expected to be valuable tools for investigations on the role of cathepsin F in macrophage-related disorders, atherogenesis, or lysosomal storage defects.

## Experimental Section

Amino acid derivatives **13**, **25**, **26**, **29**, **32** as well as compounds **18** and **22** were obtained from Bachem (Bubendorf, Switzerland), Acros (Geel, Belgium) and Aldrich (Steinheim, Germany). Solvents were used without additional purification. Preparative column

chromatography was performed on silica gel 60, 0.060–0.200 mm (Acros Organics). Mass spectra were recorded on an API 2000 mass spectrometer (electrospray ion source, Applied Biosystems, Darmstadt, Germany) coupled with an Agilent 1100 HPLC system using a Phenomenex Luna HPLC $C_{18}$ column (50×2.00 mm, particle size: 3 μm). The purity of the tested compounds was determined by HPLC-DAD obtained on an LC–MS instrument (HPLC Agilent 1100) using the procedure as follows: dissolving the compounds at a concentration of 1.0 mg mL$^{-1}$ in MeOH and, if necessary, sonication to complete dissolution. Then, 10 μL of the substance solution was injected into the Phenomenex Luna $C_{18}$ column, and elution was carried out with a gradient of $H_2O$/MeOH containing 2 mM ammonium acetate from 90:10 up to 0:100 for 30 min at a flow rate of 250 μL min$^{-1}$, starting the gradient after 10 min. UV absorption was detected from $\lambda$ 220 to 400 nm using a diode array detector. All tested compounds possessed a purity of no less than 95%. $^1$H and $^{13}$C NMR spectra were recorded on a Bruker Avance 500 MHz NMR spectrometer with [$D_6$]DMSO as solvent. NMR spectra were recorded at 30 °C. Chemical shifts are given in parts per million (ppm) relative to the remaining protons of the deuterated solvent used as internal standard. Coupling constants *J* are given in Hertz, and spin multiplicities are given as s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and br (broad). Melting points were determined on a Büchi 510 oil bath apparatus and were uncorrected. Compounds **19**–**21** were prepared as described.[18]

### *Chemistry*

**(*S*,*S*)-*N*-(*tert*-Butoxycarbonyl)cyclohexylalanylmethionine nitrile (7)**: Compound **16** (0.40 g, 1.00 mmol) was dissolved in dry DMF (20 mL). Cyanuric chloride (0.18 g, 1.00 mmol) was added, and the mixture was stirred for 4 h at RT. The solvent was evaporated under reduced pressure. The resulting solid was treated with 10% $NaHCO_3$ (20 mL) and stirred for 10 min at RT. The aqueous suspension was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL), and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was purified by column chromatography on silica gel using EtOAc/petroleum ether (PE) (3:7) as eluent to obtain **7** as a white solid (0.24 g, 62%); mp: 74–75 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO): $\delta$ = 0.80–0.91 (m, 2H), 1.09–1.18 (m, 3H), 1.23–1.26 (m, 1H), 1.36 (s, 9H), 1.38–1.42 (m, 2H), 1.59–1.69 (m, 5H), 1.99–2.06 (m, 5H), 2.51–2.56 (m, 2H), 3.92–3.95 (m, 1H), 4.83–4.86 (m, 1H), 6.95 (d, *J* = 6.6 Hz, 1H), 8.62 ppm (d, *J* = 6.6 Hz, 1H); $^{13}$C NMR (125 MHz, [$D_6$]DMSO): $\delta$ = 14.7, 25.9, 26.0, 26.2, 28.3, 28.9, 31.5, 33.1, 33.2, 33.8, 38.9, 40.3, 52.2, 78.3, 119.4, 155.6, 173.1 ppm; LC/ESI-MS (*m*/*z*): negative mode 382 [*M*−H]$^-$, positive mode 384 [*M*+H]$^+$; purity, 100%.

**(*S*,*S*)-*N*-(*tert*-Butoxycarbonyl)leucylmethionine nitrile (8)**: Compound **17** (0.90 g, 2.50 mmol) was dissolved in dry DMF (20 mL). Cyanuric chloride (0.46 g, 2.50 mmol) was added, and the mixture was stirred for 4 h at RT. The solvent was evaporated under reduced pressure. The resulting solid was treated with 10% $NaHCO_3$ (20 mL) and stirred for 10 min at RT. The aqueous suspension was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was purified by column chromatography on silica gel using EtOAc/PE (1:2) as eluent to obtain **8** as a white solid (0.24 g, 28%); mp: 68–69 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO): $\delta$ = 0.85 (d, 3H, *J* = 6.6 Hz), 0.87 (d, 3H, *J* = 6.7 Hz), 1.36 (s, 9H), 1.40–1.49 (m, 2H), 1.55–1.60 (m, 1H), 1.99–2.03 (m, 2H), 2.04 (s, 3H), 2.51–2.57 (m, 2H), 3.89–3.93 (m, 1H), 4.82–4.87 (m, 1H), 6.96 (d, *J* = 8.2 Hz, 1H), 8.62 ppm (d, *J* = 7.6 Hz, 1H); $^{13}$C NMR (125 MHz, [$D_6$]DMSO): $\delta$ = 14.6, 21.7, 23.0, 24.4, 28.3, 28.8, 31.5, 38.9, 52.8, 78.2, 119.3, 155.5, 173.0 ppm, one carbon signal is obscured by the DMSO signal; LC/ESI-MS (*m*/*z*): negative mode 342 [*M*−H]$^-$, positive mode 344 [*M*+H]$^+$; purity, 96%.

***N*-{4-[5-(2-Thienyl)-1,2,4-oxadiazol-3-yl]benzoyl}cyclohexylalanylglycine nitrile (9)**: Compound **27** (0.19 g, 0.60 mmol) was dissolved in dry THF (5 mL). Under ice cooling, methanesulfonic acid (0.35 g, 3.60 mmol) was added. The resulting mixture was stirred at RT overnight. The solvent was evaporated under reduced pressure and the crude product, cyclohexylalanylglycine nitrile methanesulfonate, was used without further purification. 4-[5-(2-Thienyl)-1,2,4-oxadiazol-3-yl]benzoic acid (**21**, 0.16 g, 0.60 mmol) was dissolved in dry THF (30 mL). EDC (93 mg, 0.60 mmol) and DMAP (4 mg, 0.03 mmol) were added and stirred for 10 min at RT. DIPEA (0.39 g, 0.53 mL, 3.00 mmol) and the cyclohexylalanylglycine nitrile methanesulfonate were added and the solution was stirred overnight at RT. After evaporation of the solvent, the residue was extracted with EtOAc (3×100 mL). The combined organic layers were washed with 10% $KHSO_4$ (60 mL), $H_2O$ (60 mL), sat. $NaHCO_3$ (2×60 mL), $H_2O$ (60 mL) and brine (60 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was purified by column chromatography on silica gel using PE/EtOAc (1:2) as eluent to obtain **9** as a white solid (61 mg, 22% over two steps); mp: 168–170 °C; $^1$H NMR (500 MHz, $CDCl_3$): $\delta$ = 0.90–1.04 (m, 2H), 1.22–1.25 (m, 3H), 1.36–1.44 (m, 1H), 1.62–1.89 (m, 7H), 4.08–4.12 (m, 1H), 4.21 (dd, *J* = 17.4 Hz, *J* = 5.7 Hz, 1H), 4.75–4.79 (m, 1H), 6.86 (d, *J* = 7.9 Hz, 1H), 7.21 (dd, *J* = 5.0 Hz, *J* = 3.8 Hz, 1H), 7.56 (t, *J* = 5.5 Hz, 1H), 7.65 (dd, *J* = 5.1 Hz, *J* = 1.3 Hz, 1H), 7.87 (d, *J* = 8.5 Hz, 2H), 7.94 (dd, *J* = 3.6 Hz, *J* = 1.1 Hz, 1H), 8.19 ppm (d, *J* = 8.6 Hz, 2H); $^{13}$C NMR (125 MHz, $CDCl_3$): $\delta$ = 26.0, 26.1, 26.3, 27.6, 32.8, 33.5, 34.2, 39.0, 51.4, 115.8, 125.6, 127.7, 127.9, 128.6, 130.3, 132.1, 132.3, 135.4, 167.4, 167.9, 171.7, 172.4 ppm; LC/ESI-MS (*m*/*z*): negative mode 462 [*M*−H]$^-$, positive mode 464 [*M*+H]$^+$; purity, 99%.

**(*S*,*S*)-*N*-(4-Phenylbenzoyl)leucylglycine nitrile (10)**: Compound **28** (0.16 g, 0.60 mmol) was dissolved in dry THF (5 mL). Under ice cooling, methanesulfonic acid (0.35 g, 3.60 mmol) was added. The resulting mixture was stirred at RT overnight. The volume of the solvent was reduced, $Et_2O$ was added, the crude product, leucinylglycine nitrile methanesulfonate, was filtered off and used without further purification. Biphenyl-4-carboxylic acid (**24**, 56 mg, 0.30 mmol) was dissolved in dry THF (10 mL). EDC (47 mg, 0.30 mmol) and DMAP (2 mg, 0.02 mmol) were added and stirred for 10 min at RT. DIPEA (0.47 g, 0.63 mL, 3.60 mmol) and the leucinylglycine nitrile methanesulfonate were added and the solution was stirred overnight at RT. After evaporation of the solvent, the residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (2×30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was purified by column chromatography on silica gel using PE/EtOAc (1:2) as eluent to obtain **10** as a white solid (31 mg, 30%); mp: 173–175 °C; $^1$H NMR (500 MHz, $CDCl_3$): $\delta$ = 0.96 (d, *J* = 6.3 Hz, 3H), 0.99 (d, *J* = 6.3 Hz, 3H), 1.71–1.77 (m, 2H), 1.82–1.88 (m, 1H), 4.08 (dd, *J* = 17.5 Hz, *J* = 5.8 Hz, 1H), 4.19 (dd, *J* = 17.3 Hz, *J* = 5.7 Hz, 1H), 4.75 (q, *J* = 7.5 Hz, 1H), 6.72 (d, *J* = 7.9 Hz, 1H), 7.36–7.39 (m, 1H), 7.42–7.45 (m, 2H), 7.56–7.58 (m, 2H), 7.59–7.60 (m, 1H), 7.64 (d, *J* = 8.6 Hz, 2H), 7.83 ppm (d, *J* = 8.5 Hz, 2H); $^{13}$C NMR (125 MHz, $CDCl_3$): $\delta$ = 22.1, 22.9, 24.9, 27.5, 40.3, 51.8, 115.8, 127.2, 127.4, 127.6, 128.2, 129.0, 131.7, 139.7, 145.1, 167.9, 172.4 ppm; LC/ESI-MS (*m*/*z*): negative mode 348 [*M*−H]$^-$, positive mode 350 [*M*+H]$^+$; purity, 98%.

***N*-(4-Phenylbenzoyl)-(*R*,*S*)-leucyl-(*S*)-methionine nitrile (11)**: Compound **31** (0.22 g, 0.50 mmol) was dissolved in dry DMF (10 mL). Cyanuric chloride (92 mg, 0.50 mmol) was added and the mixture was stirred for 4 h at RT. The solvent was evaporated under reduced pressure. The resulting solid was treated with 10% $NaHCO_3$ (20 mL) and stirred for 10 min at RT. The aqueous suspension was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was purified by column chromatography on silica gel using $CH_2Cl_2$/MeOH (9:1) as eluent to obtain **11** as a white solid (70 mg, 33%); mp: 134–135 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO, mixture of diastereomers, ratio approximately 3:4 according to $^1$H NMR, *w*=weak, refers to minor diastereomer, *i*=intense, refers to major diastereomer): $\delta$=0.88–0.90 (*w*+*i* overlapping; m, 3H (*w*), 3H (*i*)), 0.92–0.94 (*w*+*i* overlapping; m, 3H (*w*), 3H (*i*)), 1.51–1.58 (*w*+*i* overlapping; m, 1H (*w*), 1H (*i*)), 1.65–1.77 (*w*+*i* overlapping, m, 2H (*w*), 2H (*i*)), 2.00–2.10 (*w*+*i* overlapping, m, 5H (*w*), 5H (*i*)), 2.51–2.59 (*w*+*i* overlapping, m, 2H (*w*), 2H (*i*)), 4.48–4.53 (*w*+*i* overlapping, m, 1H (*w*), 1H (*i*)), 4.83–4.90 (*w*+*i* overlapping, m, 1H (*w*), 1H (*i*)), 7.39–7.42 (*w*+*i* overlapping, m, 1H (*w*), 1H (*i*)), 7.47–7.50 (*w*+*i* overlapping, m, 2H (*w*), 2H (*i*)), 7.72–7.73 (*w*+*i* overlapping, m, 2H (*w*), 2H (*i*)), 7.76–7.78 (*w*+*i* overlapping, m, 2H (*w*), 2H (*i*)), 7.99–8.02 (*w*+*i* overlapping, m, 2H (*w*), 2H (*i*)), 8.55–8.57 (*w*+*i* overlapping, m, 1H (*w*), 1H (*i*)), 8.77 ppm (d, *J*=7.9 Hz, 1H (*i*)), 8.82 (d, *J*=7.9 Hz, 1H (*w*)); $^{13}$C NMR (125 MHz, [$D_6$]DMSO, *w*=weak, refers to minor diastereomer, *i*=intense, refers to major diastereomer): $\delta$=14.6 (*w*), 14.7 (*i*), 21.5 (*i*), 21.5 (*w*), 23.0 (*w*), 23.1 (*i*), 24.6 (*i*), 24.6 (*w*), 28.9 (*w*+*i*), 31.4 (*w*), 31.4 (*i*), 39.0 (*w*), 39.1 (*i*), 40.3 (*w*+*i*), 51.9 (*i*), 52.0 (*w*), 119.3 (*w*+*i*), 126.5 (*w*+*i*), 127.0 (*w*+*i*), 128.2 (*w*+*i*), 128.4 (*w*+*i*), 129.2 (*w*+*i*), 132.9 (*w*+*i*), 139.3 (*w*+*i*), 143.0 (*w*+*i*), 166.3 (*w*+*i*), 172.7 (*i*), 172.7 ppm (*w*); negative mode 422 $[M-H]^-$, positive mode 424 $[M+H]^+$; purity, 100%.

**(*S*,*S*)-*N*-(4-Phenylbenzoyl)leucylmethionine nitrile (12)**: Compound **34** (0.13 g, 0.30 mmol) was dissolved in dry DMF (10 mL). Cyanuric chloride (55 mg, 0.30 mmol) was added and the mixture was stirred for 4 h at RT. The solvent was evaporated under reduced pressure. The resulting solid was treated with 10% $NaHCO_3$ (20 mL) and stirred for 10 min at RT. The aqueous suspension was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was purified by column chromatography on silica gel using $CH_2Cl_2$/MeOH (50:1) as eluent to obtain **12** as a white solid (39 mg, 31%); mp: 152–153 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO): $\delta$=0.89 (d, *J*=6.3 Hz, 3H), 0.93 (d, *J*=6.7 Hz, 3H), 1.53–1.58 (m, 1H), 1.65–1.77 (m, 2H), 2.03–2.10 (m, 5H), 2.51–2.59 (m, 2H), 4.48–4.52 (m, 1H), 4.86–4.90 (m, 1H), 7.38–7.42 (m, 1H), 7.47–7.50 (m, 2H), 7.71–7.73 (m, 2H), 7.75–7.78 (m, 2H), 7.98–8.01 (m, 2H), 8.57 (d, *J*=7.9 Hz, 1H), 8.77 ppm (d, *J*=7.6 Hz, 1H); $^{13}$C NMR (125 MHz, [$D_6$]DMSO): $\delta$=14.7, 21.5, 23.1, 24.6, 29.0, 31.4, 39.1, 51.9, 119.3, 126.5, 127.0, 128.2, 128.4, 129.2, 132.9, 139.3, 143.1, 166.3, 172.7 ppm, one carbon signal is obscured by the DMSO signal; negative mode 422 $[M-H]^-$, positive mode 424 $[M+H]^+$; purity, 97%.

**(*S*)-*N*-(*tert*-Butoxycarbonyl)methionine amide (14)**: Boc-Met-OH (**13**, 1.99 g, 8.00 mmol) was dissolved in dry THF (20 mL) and cooled at −25 °C. To the stirred solution, NMM (0.81 g, 8.00 mmol) and isobutyl chloroformate (1.09 g, 8.00 mmol) were added consecutively. Concentrated aqueous ammonia solution (2.99 mL, 40.0 mmol) was given to the reaction mixture when the precipitation of *N*-methylmorpholine hydrochloride occurred. The mixture was allowed to warm to RT within 30 min and was stirred overnight. After evaporation of the solvent the resulting aqueous residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (2×30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from EtOAc to obtain **14** as a white solid (1.56 g, 79%); mp: 116–118 °C, lit.[19] mp: 118–119 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO): $\delta$=1.37 (s, 9H), 1.70–1.77 (m, 1H), 1.81–1.87 (m, 1H), 2.02 (s, 3H), 2.40–2.45 (m, 2H), 3.90–3.95 (m, 1H), 6.80 (d, *J*=8.2 Hz, 1H), 6.94 (s, 1H), 7.20 ppm (s, 1H); $^{13}$C NMR (125 MHz, [$D_6$]DMSO): $\delta$=14.8, 28.3, 30.0, 31.9, 53.5, 78.1, 155.5, 173.9 ppm. LC/ESI-MS (*m*/*z*): negative mode 247 $[M-H]^-$, positive mode 249 $[M+H]^+$; purity, 100%.

**(*S*)-Methionine amide trifluoroacetate (15)**: Compound **14** (0.25 g, 1.00 mmol) was dissolved in a mixture of dry $CH_2Cl_2$ and TFA (10 mL, 1:1) and stirred for 2 h at 0 °C. The solvent was evaporated and the oily crude product was used without further purification.

**(*S*,*S*)-*N*-(*tert*-Butoxycarbonyl)cyclohexylalanylmethionine amide (16)**: Boc-Cha-OH (0.27 g, 1.00 mmol) was dissolved in dry $CH_2Cl_2$ (10 mL). Compound **15** (0.26 g, 1.00 mmol), DIPEA (0.65 g, 0.88 mL, 5.00 mmol) and HATU (0.38 g, 1.00 mmol) were added. The solution was stirred overnight at RT. After evaporation of the solvent, the resulting aqueous residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from EtOAc to obtain **16** as a white solid (0.40 g, 99%); mp: 166–167 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO): $\delta$=0.79–0.90 (m, 2H), 1.08–1.18 (m, 3H), 1.24–1.29 (m, 1H), 1.37 (s, 9H), 1.40–1.44 (m, 2H), 1.59–1.70 (m, 5H), 1.73–1.80 (m, 1H), 1.87–1.95 (m, 1H), 2.01 (s, 3H), 2.35–2.43 (m, 2H), 3.91–3.96 (m, 1H), 4.25–4.29 (m, 1H), 6.94 (d, *J*=7.9 Hz, 1H), 7.06 (s, 1H), 7.25 (s, 1H), 7.72 ppm (d, *J*=7.9 Hz, 1H); $^{13}$C NMR (125 MHz, [$D_6$]DMSO): $\delta$=14.8, 25.8, 26.0, 26.2, 28.3, 29.6, 32.1, 32.2, 33.3, 33.7, 39.2, 51.6, 52.6, 78.3, 155.6, 172.5, 173.0 ppm; LC/ESI-MS (*m*/*z*): negative mode 400 $[M-H]^-$, positive mode 402 $[M+H]^+$; purity, 100%.

**(*S*,*S*)-*N*-(*tert*-Butoxycarbonyl)leucylmethionine amide (17)**: Boc-Leu-OH (1.12 g, 4.50 mmol) was dissolved in dry THF (20 mL) and cooled at −25 °C. To the stirred solution, NMM (046 g, 4.50 mmol) and isobutyl chloroformate (0.61 g, 4.50 mmol) were added consecutively. Compound **15** (1.18 g, 4.50 mmol) was dissolved in 2 N NaOH (2.25 mL, 4.50 mmol) and given to the reaction mixture when the precipitation of *N*-methylmorpholine hydrochloride occurred. The mixture was allowed to warm to RT within 30 min and stirred overnight at RT. After evaporation of the solvent, the resulting aqueous residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from EtOAc to obtain **17** as a white solid (0.90 g, 56%); mp: 154–155 °C; $^1$H NMR (500 MHz, [$D_6$]DMSO): $\delta$=0.84 (d, *J*=6.6 Hz, 3H), 0.86 (d, *J*=6.7 Hz, 3H), 1.37 (s, 9H), 1.41 (t, *J*=7.4 Hz, 2H), 1.56–1.61 (m, 1H), 1.73–1.80 (m, 1H), 1.88–1.95 (m, 1H), 2.02 (s, 3H), 2.34–2.43 (m, 2H), 3.88–3.93 (q, *J*=6.6 Hz, 1H), 4.25–4.29 (m, 1H), 6.96 (d, *J*=8.2 Hz, 1H), 7.05 (s, 1H), 7.26 (s, 1H), 7.72 ppm (d, *J*=8.2 Hz, 1H); $^{13}$C NMR (125 MHz, [$D_6$]DMSO): $\delta$=14.8, 21.7, 23.1, 24.4, 28.3, 29.6, 32.2, 40.5, 51.6, 53.2, 78.3, 155.6,

172.5, 173.0 ppm; LC/ESI-MS (*m/z*): negative mode 360 $[M-H]^-$, positive mode 362 $[M+H]^+$; purity, 100%.

**Methyl biphenyl-4-carboxylate (23)**: To a suspension of $Pd(PPh_3)_4$ (0.11 g, 0.09 mmol) in dry DME (20 mL) in an autoclave, methyl-4-bromobenzoate (**22**, 0.65 g, 3.00 mmol) was added. The mixture was stirred for 10 min at RT. Phenylboronic acid (0.55 g, 4.5 mmol) and aqueous $Na_2CO_3$ (0.67 g, 6.3 mmol, in 5 mL $H_2O$) were added. The mixture was stirred at 120 °C for 4 h. The solvent was evaporated under reduced pressure and the residue was suspended in $H_2O$ (50 mL). The aqueous suspension was extracted with EtOAc (3×30 mL). The combined organic layers were washed with 10% $KHSO_4$ (2×30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (2×30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from MeOH to obtain **23** as a white solid (0.33 g, 52%); mp: 115–116 °C, lit.[24] mp: 114–115 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta$ = 3.87 (s, 3H), 7.40–7.44 (m, 1H), 7.48–7.52 (m, 2H), 7.72–7.74 (m, 2H), 7.80–7.83 (m, 2H), 8.02–8.04 ppm (m, 2H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta$ = 52.3, 127.1, 127.1, 128.5, 128.6, 129.2, 129.9, 139.0, 144.8, 166.2 ppm; LC/ESI-MS (*m/z*): positive mode 213 $[M+H]^+$; purity, 99%.

**Biphenyl-4-carboxylic acid (24)**: Compound **23** (0.32 g, 1.50 mmol) was dissolved in a mixture of THF/$H_2O$ 2:1 (30 mL). LiOH (0.16 g, 3.75 mmol) was added under ice cooling and stirred for 3 h. The solvent was evaporated under reduced pressure. The mixture was adjusted with concentrated HCl to pH 1–2. The aqueous solution was extracted with EtOAc (3×30 mL) and washed with brine (30 mL). The crude product was recrystallized from MeOH to obtain **24** as a white solid (0.30 g, 99%); mp: 227–228 °C, lit.[26] mp: 224–225 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta$ = 7.40–7.43 (m, 1H), 7.47–7.51 (m, 2H), 7.71–7.73 (m, 2H), 7.77–7.80 (m, 2H), 8.00–8.03 (m, 2H), 12.92 ppm (bs, 1H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta$ = 126.9, 127.1, 128.4, 129.2, 129.8, 130.1, 139.2, 144.4, 167.3 ppm; LC/ESI-MS (*m/z*): positive mode 199 $[M+H]^+$; purity, 98%.

***N*-(*tert*-Butyloxycarbonyl)-β-cyclohexylalanylglycine nitrile (27)**: Boc-Cha-OH (1.09 g, 4.00 mmol) was dissolved in dry THF (10 mL) and cooled at −25 °C. To the stirred solution, NMM (0.40 g, 4.00 mmol) and isobutyl chloroformate (0.55 g, 4.00 mmol) were added consecutively. Aminoacetonitrile monosulfate (0.62 g, 4.00 mmol) was dissolved in 2 N NaOH (2.00 mL, 4.00 mmol) and given to the reaction mixture. It was allowed to warm to RT within 30 min and stirred overnight at RT. After evaporation of the solvent, the residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (2×30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from EtOAc to obtain **27** as a colorless oil (0.40 g, 32%); mp: 82–83 °C, lit.[20] mp: 81–83 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta$ = 0.78–0.92 (m, 2H), 1.08–1.18 (m, 3H), 1.23–1.29 (m, 1H), 1.37 (s, 9H), 1.41 (t, *J* = 7.1 Hz, 2H), 1.58–1.71 (m, 5H), 3.98 (q, *J* = 7.7 Hz, 1H), 4.09 (dd, *J* = 5.6 Hz, *J* = 2.7 Hz, 2H), 6.93 (d, *J* = 8.2 Hz, 1H), 8.50 ppm (t, *J* = 5.5 Hz, 1H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta$ = 25.8, 26.0, 26.2, 27.2, 28.3, 31.8, 33.2, 33.7, 51.9, 78.2, 117.7, 155.5, 173.5 ppm, one carbon signal is obscured by the DMSO signal; LC/ESI-MS (*m/z*): negative mode 308 $[M-H]^-$; purity, 100%.

**(*S*,*S*)-*N*-(*tert*-Butoxycarbonyl)leucylglycine nitrile (28)**: Boc-Leu-OH (0.93 g, 4.00 mmol) was dissolved in dry THF (10 mL) and cooled at −25 °C. To the stirred solution, NMM (0.40 g, 4.00 mmol) and isobutyl chloroformate (0.55 g, 4.00 mmol) were added consecutively. Aminoacetonitrile monosulfate (0.62 g, 4.00 mmol) was dissolved in 2 N NaOH (2.00 mL, 4.00 mmol) and given to the reaction mixture. It was allowed to warm to RT within 30 min and stirred overnight at RT. After evaporation of the solvent, the residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (2×30 mL), $H_2O$ (30 mL) and brine (30 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from EtOAc to obtain **28** as a white solid (0.45 g, 42%); mp: 113–114 °C, lit.[20] mp: 114–116 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta$ = 0.84 (d, *J* = 6.7 Hz, 3H), 0.86 (d, *J* = 6.7 Hz, 3H), 1.37 (s, 9H), 1.40–1.47 (m, 2H), 1.57 (sept, *J* = 6.6 Hz, 1H), 3.93–3.97 (m, 1H), 4.09 (dd, *J* = 5.7 Hz, *J* = 1.0 Hz, 2H), 6.95 (d, *J* = 7.9 Hz, 1H), 8.51 ppm (t, *J* = 5.5 Hz, 1H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta$ = 21.5, 23.0, 24.4, 27.2, 28.3, 40.6, 52.6, 78.2, 117.7, 155.5, 173.4 ppm; LC/ESI-MS (*m/z*): negative mode 268 $[M-H]^-$, positive mode 270 $[M+H]^+$; purity, 100%.

***tert*-Butyl (*S*)-*N*-(4-phenylbenzoyl)leucine (30)**: Biphenyl-4-carboxylic acid (**24**, 0.20 g, 1.00 mmol) was dissolved in dry $CH_2Cl_2$ (10 mL). H-Leu-O*t*Bu hydrochloride (**29**, 0.22 g, 1.00 mmol), DIPEA (0.35 mL, 2.00 mmol) and HATU (0.38 g, 1.00 mmol) were added. The solution was stirred overnight at RT. After evaporation of the solvent the resulting aqueous residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from EtOAc to obtain **30** as a white solid (0.27 g, 73%); mp: 105–106 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta$ = 0.88 (d, *J* = 6.6 Hz, 3H), 0.93 (d, *J* = 6.3 Hz, 3H), 1.41 (s, 9H), 1.53–1.59 (m, 1H), 1.67–1.78 (m, 2H), 4.37–4.42 (m, 1H), 7.39–7.42 (m, 1H), 7.47–7.51 (m, 2H), 7.71–7.73 (m, 2H), 7.76–7.78 (m, 2H), 7.96–7.98 (m, 2H), 8.59 ppm (d, *J* = 7.9 Hz, 1H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta$ = 21.4, 23.0, 24.7, 27.8, 38.4, 51.8, 80.5, 126.6, 127.0, 128.2, 128.3, 129.1, 133.0, 139.3, 143.0, 166.3, 172.0 ppm; negative mode 366 $[M-H]^-$, positive mode 368 $[M+H]^+$; purity, 88%.

***N*-(4-Phenylbenzoyl)-(*R*,*S*)-leucyl-(*S*)-methionine amide (31)**: Compound **30** (0.26 g, 0.70 mmol) was dissolved in a mixture of dry $CH_2Cl_2$ and TFA (10 mL, 1:1) and stirred for 2 h at 0 °C. The solvent was evaporated and the oily crude product, (*S*)-*N*-(4-phenylbenzoyl)leucine, was used without further purification and dissolved in dry $CH_2Cl_2$ (10 mL). (*S*)-Methionine amide trifluoroacetate (**15**, 0.18 g, 0.70 mmol), DIPEA (1.21 mL, 7.00 mmol) and HATU (0.27 g, 0.70 mmol) were added. The solution was stirred overnight at RT. After evaporation of the solvent the resulting aqueous residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from MeOH to obtain **31** as a white solid (0.25 g, 81% over two steps); mp: 231–232 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO, mixture of diastereomers, ratio approximately 1:3 according to $^1$H NMR, *w* = weak, refers to minor diastereomer, *i* = intense, refers to major diastereomer): $\delta$ = 0.88–0.90 (*w*+*i* overlapping; m, 3H (*w*), 3H (*i*)), 0.93 (*w*+*i* overlapping; m, 3H (*w*), 3H (*i*)), 1.55–1.60 (*w*+*i* overlapping; m, 1H (*w*), 1H (*i*)), 1.66–1.75 (*w*+*i* overlapping; m, 2H (*w*), 2H (*i*)), 1.76–1.84 (*w*+*i* overlapping; m, 1H (*w*), 1H (*i*)), 1.91–1.98 (*w*+*i* overlapping; m, 1H (*w*), 1H (*i*)), 2.01 (*w*+*i* overlapping; s, 3H (*w*), 3H (*i*)), 2.37–2.46 (*w*+*i* overlapping; m, 2H (*w*), 2H (*i*)), 4.27–4.31 (*w*+*i* overlapping; m, 1H (*w*), 1H (*i*)), 4.47–4.51 (*w*+*i* overlapping; m, 1H (*w*), 1H (*i*)), 7.05 (s, 1H (*i*)), 7.11 (s, 1H (*w*)), 7.26 (s, 1H (*i*)), 7.31 (s, 1H (*w*)), 7.39–7.42 (*w*+*i* overlapping; (m, 1H (*w*), 1H (*i*)), 7.47–7.50 (*w*+*i* overlapping) (m, 2H (*w*), 2H (*i*)), 7.71–7.73 (*w*+*i* overlapping; m, 2H (*w*), 2H (*i*)), 7.76–7.78 (*w*+*i* overlapping; m,

2H (*w*), 2H (*i*)), 7.92 (d, $J=8.2$ Hz, 1H (*i*)), 7.97–7.99 (*w*+*i* overlapping; m, 2H (*w*), 2H (*i*)), 8.29 (d, $J=8.5$ Hz, 1H (*w*)), 8.55 (d, $J=7.9$ Hz, 1H (*i*)), 8.60 ppm (d, $J=7.0$ Hz, 1H (*w*)); $^{13}$C NMR (125 MHz, [D$_6$]DMSO, *w*=weak, refers to minor diastereomer, *i*=intense, refers to major diastereomer): $\delta=14.7$ (*w*), 14.8 (*i*), 21.6 (*i*), 22.0 (*w*), 22.9 (*w*), 23.2 (*i*), 24.6 (*w*+*i*), 29.7 (*i*), 29.9 (*w*), 31.3 (*w*), 32.0 (*i*), 51.7 (*w*), 51.8 (*i*), 52.4 (*i*), 52.8 (*w*), 126.6 (*w*+*i*), 127.0 (*w*+*w*+*i*), 128.2 (*w*+*i*), 128.4 (*i*), 129.2 (*w*+*w*+*i*), 132.9 (*w*), 133.1 (*i*), 139.3 (i), 143.0 (*w*+*i*), 166.4 (*i*), 166.7 (*w*), 172.2 (*i*), 172.5 (*w*), 173.1 (*i*), 173.4 ppm (*w*), one carbon signal is obscured by the DMSO signal; negative mode 440 $[M-H]^-$, positive mode 442 $[M+H]^+$; purity, 87%.

**(*S*,*S*)-*N*-Benzyloxycarbonylleucylmethionine amide (33)**: Cbz-Leu-OH (**32**, 0.53 g, 2.00 mmol) was dissolved in dry $CH_2Cl_2$ (20 mL). Compound **15** (0.52 g, 2.00 mmol), DIPEA (1.29 g, 1.74 mL, 10.00 mmol) and HATU (0.76 g, 2.00 mmol) were added. The solution was stirred overnight at RT. After evaporation of the solvent the resulting aqueous residue was extracted with EtOAc (3× 50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from MeOH to obtain **33** as a white solid (0.39 g, 50%); mp: 199–201 °C, lit.[27] mp: 210–211 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta=0.84$ (d, $J=6.7$ Hz, 3H), 0.86 (d, $J=6.6$ Hz, 3H), 1.43–1.45 (m, 2H), 1.57–1.65 (m, 1H), 1.74–1.81 (m, 1H), 1.87–1.95 (m, 1H), 2.01 (s, 3H), 2.35–2.46 (m, 2H), 4.00–4.04 (m, 1H), 4.25–4.29 (m, 1H), 5.02 (s, 2H), 7.04 (s, 1H), 7.26 (s, 1H), 7.28–7.37 (m, 5H), 7.44 (d, $J=7.9$ Hz, 1H), 7.84 ppm (d, $J=8.2$ Hz, 1H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta=14.8$, 21.6, 23.1, 24.3, 29.7, 32.1, 40.6, 51.7, 53.5, 65.5, 127.8, 128.5, 127.9, 137.2, 156.1, 172.3, 173.0 ppm; negative mode 394 $[M-H]^-$, positive mode 396 $[M+H]^+$; purity, 95%.

**(*S*,*S*)-*N*-(4-Phenylbenzoyl)leucylmethionine amide (34)**: Compound **33** (0.20 g, 0.50 mmol) was dissolved in MeOH (20 mL) and 10% Pd/C (20 mg) was added. The reaction mixture was hydrogenated at 44 psi for 7 h. Pd/C was filtered off and the solvent was evaporated. The crude product, (*S*,*S*)-leucinylmethionine amide, was used without further purification. Biphenyl-4-carboxylic acid (**24**, 0.10 g, 0.50 mmol) was dissolved in dry $CH_2Cl_2$ (10 mL). The (*S*,*S*)-leucinylmethionine amide, DIPEA (65 mg, 0.08 mL, 0.50 mmol) and HATU (0.19 g, 0.50 mmol) were added. The solution was stirred overnight at RT. After evaporation of the solvent the resulting aqueous residue was extracted with EtOAc (3×50 mL). The combined organic layers were washed with 10% $KHSO_4$ (30 mL), $H_2O$ (30 mL), sat. $NaHCO_3$ (30 mL), $H_2O$ (30 mL) and brine (50 mL). The solvent was dried ($Na_2SO_4$) and evaporated. The crude product was recrystallized from MeOH to obtain **34** as a white solid (0.21 g, 95%, over two steps); mp: 245–247 °C; $^1$H NMR (500 MHz, [D$_6$]DMSO): $\delta=0.88$ (d, $J=6.3$ Hz, 3H), 0.93 (d, $J=6.3$ Hz, 3H), 1.54–1.60 (m, 1H), 1.67–1.75 (m, 2H), 1.76–1.83 (m, 1H), 1.92–1.99 (m, 1H), 2.01 (s, 3H), 2.38–2.46 (m, 2H), 4.27–4.31 (m, 1H), 4.47–4.51 (m, 1H), 7.05 (s, 1H), 7.25 (s, 1H), 7.39–7.42 (m, 1H), 7.47–7.50 (m, 2H), 7.72–7.73 (m, 2H), 7.76–7.78 (m, 2H), 7.91 (d, $J=8.2$ Hz, 1H), 7.97–7.99 (m, 2H), 8.54 ppm (d, $J=7.6$ Hz, 1H); $^{13}$C NMR (125 MHz, [D$_6$]DMSO): $\delta=14.8$, 21.6, 23.2, 24.6, 29.7, 32.0, 39.2, 51.8, 52.3, 126.6, 127.0, 128.2, 128.3, 129.2, 133.1, 139.3, 143.0, 166.4, 172.2, 173.1 ppm; negative mode 440 $[M-H]^-$, positive mode 442 $[M+H]^+$; purity, 96%.

### *Enzyme inhibition assays*

Human cathepsins F, K, and S were assayed fluorimetrically on a Monaco Safas spectrofluorimeter flx. The wavelength for excitation was 360 nm and for emission 440 nm. The reactions were followed at 25 °C over 20 min. Human cathepsins B and L were assayed spectrophotometrically on a Cary 50 Bio, Varian, at 405 nm. The reactions were followed at 37 °C over 20 min. The 10 mM inhibitor solutions were prepared in DMSO. The final concentration of DMSO in all assays was 2%. Progress curves were analyzed by linear regression. Results given in Tables 1 and 2 were obtained from experiments performed in duplicate with five different inhibitor concentrations. $IC_{50}$ values were determined by nonlinear regression using the equation $v=v_0/(1+[I]/IC_{50})$, in which $v$ and $v_0$ are the rates in the presence and absence of the inhibitor, respectively, and [I] is the inhibitor concentration. $K_i$ values were calculated from $IC_{50}$ values by applying the equation $K_i=IC_{50}/(1+[S]/K_M)$.

*Cathepsin F inhibition assay*: A human recombinant cathepsin F (Enzo Life Sciences, Lörrach, Germany) stock solution (0.014 mg mL$^{-1}$) in 20 mM NaOAc buffer pH 5.0, 2.5 mM EDTA, 250 mM NaCl, and 20 mM L-cysteine was diluted (1:50) with assay buffer (100 mM sodium phosphate pH 6.0, 100 mM NaCl, 5 mM EDTA, 0.01% Brij 35) containing 5 mM dithiothreitol (DTT) and incubated for 2 h at 37 °C. A 10 mM stock solution of the fluorogenic substrate Cbz-Phe-Arg-AMC was prepared in DMSO. The assay was performed with a final cathepsin F concentration of 22.4 ng mL$^{-1}$. The final concentration of substrate was 40 μM (2.1 $K_M$). Into a cuvette containing 900 μL assay buffer, the inhibitor solution and DMSO in a total volume of 16 μL, and 4 μL of the substrate solution (10 mM) were added and thoroughly mixed. The reaction was initiated by adding 80 μL of the cathepsin F solution (280 ng mL$^{-1}$).

*Cathepsin B inhibition assay*:[18] Assay buffer was 100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, 0.01% Brij 35. A stock solution of human isolated cathepsin B (Calbiochem, Darmstadt, Germany) of 1.81 mg mL$^{-1}$ in 20 mM sodium acetate buffer pH 5.0, 1 mM EDTA was diluted 1:500 with assay buffer containing 5 mM DTT and incubated for 30 min at 37 °C. A 100 mM stock solution of the chromogenic substrate Cbz-Arg-Arg-pNA was prepared with DMSO. The final concentration of the substrate was 500 μM (0.45 $K_M$). Assays were performed with a final cathepsin B concentration of 72 ng mL$^{-1}$. Into a cuvette containing 960 μL assay buffer, inhibitor solution and DMSO in a total volume of 15 μL, and 5 μL of the substrate solution were added and thoroughly mixed. The reaction was initiated by adding 20 μL of the cathepsin B solution.

*Cathepsin L inhibition assay*:[18] Assay buffer was 100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, and 0.01% Brij 35. A stock solution of human isolated cathepsin L (Enzo Life Sciences, Lörrach, Germany) of 135 μg mL$^{-1}$ in 20 mM malonate buffer pH 5.5, 400 mM NaCl, and 1 mM EDTA was diluted 1:100 with assay buffer containing 5 mM DTT and incubated for 30 min at 37 °C. A 10 mM stock solution of the chromogenic substrate Cbz-Phe-Arg-pNA was prepared with DMSO. The final concentration of the substrate was 100 μM (5.88 $K_M$). Assays were performed with a final cathepsin L concentration of 54 ng mL$^{-1}$. Into a cuvette containing 940 μL assay buffer, inhibitor solution and DMSO in a total volume of 10 μL, and 10 μL of the substrate solution were added and thoroughly mixed. The reaction was initiated by adding 40 μL of the cathepsin L solution.

*Cathepsin K inhibition assay*:[18] A human recombinant cathepsin K (Enzo Life Sciences, Lörrach, Germany) stock solution of 23 μg mL$^{-1}$ in 50 mM sodium acetate buffer pH 5.5, 50 mM NaCl, 0.5 mM EDTA, 5 mM DTT was diluted 1:100 with assay buffer (100 mM sodium citrate pH 5.0, 100 mM NaCl, 1 mM EDTA, 0.01% CHAPS) containing 5 mM DTT and incubated for 30 min at 37 °C. A 10 mM stock solu-

tion of the fluorogenic substrate Cbz-Leu-Arg-AMC was prepared with DMSO. The final concentration of the substrate was 40 μM (13.3 $K_M$). Assays were performed with a final cathepsin K concentration of 5 ng mL$^{-1}$. Into a cuvette containing 960 μL assay buffer, inhibitor solution and DMSO in a total volume of 16 μL, and 4 μL of the substrate solution were added and thoroughly mixed. The reaction was initiated by adding 20 μL of the cathepsin K solution.

*Cathepsin S inhibition assay*:[28] Assay buffer was 100 mM sodium phosphate buffer pH 6.0, 100 mM NaCl, 5 mM EDTA, and 0.01% Brij 35. A stock solution of human recombinant cathepsin S (Enzo Life Sciences, Lörrach, Germany) of 70 μg mL$^{-1}$ in 100 mM MES buffer, pH 6.5, 1 mM EDTA, 50 mM L-cysteine, 10 mM DTT, 0.5% Triton X-100 and 30% glycerol was diluted 1:100 with a 50 mM sodium phosphate buffer pH 6.5, 50 mM NaCl, 2 mM EDTA, 0.01% Triton X-100 and 5 mM DTT and incubated for 60 min at 37 °C. A 10 mM stock solution of the fluorogenic substrate Cbz-Phe-Arg-AMC was prepared in DMSO. The assay was performed with a final substrate concentration of 40 μM (0.74 $K_M$) and a final cathepsin S concentration of 42 ng mL$^{-1}$. Into a cuvette containing 920 μL assay buffer, inhibitor solution and DMSO in a total volume of 16 μL, and 4 μL of the substrate solution were added and thoroughly mixed. The reaction was initiated by adding 60 μL of the cathepsin S solution.

### Docking studies

The X-ray structure of human cathepsin F (PDB ID: 1M6D)[22] bound to an irreversible vinyl sulfone inhibitor (**V**, Figure 1) was obtained from the RCSB Protein Data Bank.[29] The structure was used as template for covalent ligand docking with AutoDock 4.2.[30] Prior to the docking calculations, the bound inhibitor and crystallographic water molecules were removed from the template structure, and hydrogen atoms were added using Molecular Operating Environment (MOE 2012.10).[31] AutoDock Tools[30] was used to calculate atomic partial charges. By forming a covalent thioimidate bond to their nitrile function, candidate compounds were linked to the active site residue Cys25. The side chain of the residue Cys25 including the attached inhibitor structures was treated flexibly to predict possible binding modes of selected compounds with AutoDock 4.2. All other residues were fixed in their crystallographic conformation for the docking process. Putative compound binding modes were selected after visual inspection of high-scoring docking poses.

### Generation of a 3D activity landscape

For the inhibitor set, a 3D activity landscape was generated as described previously.[32] The ECFP2[33] fingerprint was used to calculate Euclidean distance relationships between compounds (constituting a coordinate-free chemical reference space), which were projected onto a 2D plane by dimension reduction using multidimensional scaling (MDS).[34] As a third dimension, the compound potency of the 57 cathepsin F inhibitors was added. By interpolating with the krige function,[35] coherent potency surface was created. The color of the surface was adjusted to the interpolated potency values (surface elevation) using a color gradient from green (indicating lowest potency values) over yellow to red (highest potency). The density of experimental potency measurements is reflected by the transparency of the activity landscape. Opaque colored grid points represent close proximity to original data points, whereas fully transparent grid points are furthest away from experimental data points. Accordingly, white areas in the landscape are interpolated and not populated with data points.

## Acknowledgements

*The authors thank Tianwei Li, Adela Dudic, Karina Scheiner, Erik Gilberg, and Robert Sellier (all University of Bonn, Germany) for assistance. J.S. is supported by the Gender Equality Center of the Bonn-Rhein-Sieg University of Applied Sciences, and N.F. by a fellowship from the Jürgen Manchot Foundation (Düsseldorf, Germany).*

**Keywords:** 3D activity landscapes · cysteine proteases · human cathepsins · nitrile inhibitors

[1] a) J. Reiser, B. Adair, T. Reinheckel, *J. Clin. Invest.* **2010**, *120*, 3421 – 3431; b) V. Turk, V. Stoka, O. Vasiljeva, M. Renko, T. Sun, B. Turk, D. Turk, *Biochim. Biophys. Acta* **2012**, *1824*, 68 – 88; c) A. Pišlar, J. Kos, *Mol. Neurobiol.* **2014**, *49*, 1017 – 1030.

[2] M. Frizler, M. Stirnberg, M. T. Sisay, M. Gütschow, *Curr. Top. Med. Chem.* **2010**, *10*, 294 – 322.

[3] B. G. Shi, R. A. Bryant, R. Riese, S. Verhelst, C. Driessen, Z. Li, D. Brömme, H. L. Ploegh, H. A. Chapman, *J. Exp. Med.* **2000**, *191*, 1177 – 1185.

[4] a) N. Obermajer, B. Doljak, J. Kos, *Expert Opin. Biol. Ther.* **2006**, *6*, 1295 – 1309; b) M. Perišić Nanut, J. Sabotič, A. Jewett, J. Kos, *Front. Immunol.* **2014**, *5*, 616.

[5] M. Fonovič, D. Brömme, V. Turk, B. Turk, *Biol. Chem.* **2004**, *385*, 505 – 509.

[6] a) J. A. Villadangos, R. A. Bryant, J. Deussing, C. Driessen, A. M. Lennon-Duménil, R. J. Riese, W. Roth, P. Saftig, G. P. Shi, H. A. Chapman, C. Peters, H. L. Ploegh, *Immunol. Rev.* **1999**, *172*, 109 – 120; b) K. Honey, A. Y. Rudensky, *Nat. Rev. Immunol.* **2003**, *3*, 472 – 482; c) L. C. Hsing, A. Y. Rudensky, *Immunol. Rev.* **2005**, *207*, 229 – 241.

[7] J. T. Palmer, D. Rasnick, J. L. Klaus, D. Brömme, *J. Med. Chem.* **1995**, *38*, 3193 – 3196.

[8] A. Lee-Dutra, D. K. Wiener, S. Sun, *Expert Opin. Ther. Pat.* **2011**, *21*, 311 – 337.

[9] C. H. Tang, J. W. Lee, M. G. Galvez, L. Robillard, S. E. Mole, H. A. Chapman, *Mol. Cell. Biol.* **2006**, *26*, 2309 – 2316.

[10] K. R. Smith, H. H. Dahl, L. Canafoglia, E. Andermann, J. Damiano, M. Morbin, A. C. Bruni, G. Giaccone, P. Cossette, P. Saftig, J. Grötzinger, M. Schwake, F. Andermann, J. F. Staropoli, K. B. Sims, S. E. Mole, S. Franceschetti, N. A. Alexander, J. D. Cooper, H. A. Chapman, S. Carpenter, S. F. Berkovic, M. Bahlo, *Hum. Mol. Genet.* **2013**, *22*, 1417 – 1423.

[11] a) K. Öörni, M. Sneck, D. Brömme, M. O. Pentikäinen, K. A. Lindstedt, M. Mäyränpää, H. Aitio, P. T. Kovanen, *J. Biol. Chem.* **2004**, *279*, 34776 – 34784; b) H. D. Urbina, F. Debaene, B. Jost, C. Bole-Feysot, D. E. Mason, P. Kuzmic, J. L. Harris, N. Winssinger, *ChemBioChem* **2006**, *7*, 1790 – 1797.

[12] a) J. J. M. Wiener, S. Sun, R. L. Thurmond, *Curr. Top. Med. Chem.* **2010**, *10*, 717 – 732; b) R. Löser, *Expert Opin. Ther. Pat.* **2011**, *21*, 585 – 591.

[13] J. Y. Gauthier, W. C. Black, I. Courchesne, W. Cromlish, S. Desmarais, R. Houle, S. Lamontagne, C. S. Li, F. Massé, D. J. McKay, M. Ouellet, J. Robichaud, J. F. Truchon, V. L. Truong, Q. Wang, M. D. Percival, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4929 – 4933.

[14] M. Frizler, F. Lohr, M. Lülsdorff, M. Gütschow, *Chem. Eur. J.* **2011**, *17*, 11419 – 11423.

[15] W. C. Black, *Curr. Top. Med. Chem.* **2010**, *10*, 745 – 751.

[16] a) R. Löser, M. Frizler, K. Schilling, M. Gütschow, *Angew. Chem. Int. Ed.* **2008**, *47*, 4331 – 4334; *Angew. Chem.* **2008**, *120*, 4403 – 4406; b) J. Y. Gauthier, N. Chauret, W. Cromlish, S. Desmarais, T. Duongle, J. P. Falgueyret, D. B. Kimmel, S. Lamontagne, S. Léger, T. LeRiche, C. S. Li, F. Massé, D. J. McKay, D. A. Nicoll-Griffith, R. M. Oballa, J. T. Palmer, M. D. Percival, D. Riendeau, J. Robichaud, G. A. Rodan, S. B. Rodan, C. Seto, M. Thérien, V. L. Truong, M. C. Venuti, G. Wesolowski, R. N. Young, R. Zamboni, W. C. Black, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 923 – 928; c) J.

Schmitz, A. M. Beckmann, A. Dudic, T. Li, R. Sellier, U. Bartz, M. Gütschow, *ACS Med. Chem. Lett.* **2014**, *5*, 1076 – 1081.
[17] M. Frizler, J. Schmitz, A. C. Schulz-Fincke, M. Gütschow, *J. Med. Chem.* **2012**, *55*, 5982 – 5986.
[18] M. Frizler, F. Lohr, N. Furtmann, J. Kläs, M. Gütschow, *J. Med. Chem.* **2011**, *54*, 396 – 400.
[19] R. Löser, M. Gütschow, *J. Enzyme Inhib. Med. Chem.* **2009**, *24*, 1245 – 1252.
[20] R. Löser, K. Schilling, E. Dimmig, M. Gütschow, *J. Med. Chem.* **2005**, *48*, 7688 – 7707.
[21] J. Y. Reginster, A. Neuprez, C. Beaudart, M. P. Lecart, N. Sarlet, D. Bernard, S. Disteche, O. Bruyere, *Drugs Aging* **2014**, *31*, 413 – 424.
[22] J. R. Somoza, J. T. Palmer, J. D. Ho, *J. Mol. Biol.* **2002**, *322*, 559 – 568.
[23] G. A. Olah, S. C. Narang, A. P. Fung, B. G. Gupta, *Synthesis* **1980**, 657 – 658.
[24] S. Schneider, W. Bannwarth, *Helv. Chim. Acta* **2001**, *84*, 735 – 742.
[25] D. Stumpfe, J. Bajorath, *J. Med. Chem.* **2012**, *55*, 2932 – 2942.
[26] H. Berger, A. E. Siegrist, *Helv. Chim. Acta* **1979**, *62*, 1411 – 1428.
[27] S. Bajusz, *Acta Chim. Acad. Sci. Hung.* **1964**, *42*, 383 – 391.
[28] M. D. Mertens, J. Schmitz, M. Horn, N. Furtmann, J. Bajorath, M. Mareš, M. Gütschow, *ChemBioChem* **2014**, *15*, 955 – 959.
[29] RCSB Protein Data Bank: www.rcsb.org.
[30] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785 – 2791.
[31] Molecular Operating Environment (MOE), ver. 2012.10, Chemical Computing Group Inc., Montreal, QC (Canada).
[32] L. Peltason, P. Iyer, J. Bajorath, *J. Chem. Inf. Model.* **2010**, *50*, 1021 – 1033.
[33] D. Rogers, M. Hahn, *J. Chem. Inf. Model.* **2010**, *50*, 742 – 754.
[34] I. Borg, P. J. F. Groenen, *Modern Multidimensional Scaling. Theory and Applications*, 2nd ed., Springer, New York, NY, **2005**.
[35] N. Cressie, *Statistics for Spatial Data*, revised ed., Wiley-VCH, New York, NY, **1993**.

Received: April 7, 2015
Published online on June 26, 2015