CHEMMEDCHEM

CHEMISTRY ENABLING DRUG DISCOVERY

Accepted Article

Title: Inhibition of the Cysteine Protease Human Cathepsin L by Triazine Nitriles: Amide···Heteroarene π-Stacking Interactions and Chalcogen Bonding in the S3 Pocket

Authors: Maude Giroud, Jakov Ivkovic, Mara Martignoni, Marianne Fleuti, Nils Trapp, Wolfgang Haap, Andreas Kuglstatter, Jörg Benz, Bernd Kuhn, Tanja Schirmeister, and Francois Diederich

To be cited as: *ChemMedChem* 10.1002/cmdc.201600563

Link to VoR: http://dx.doi.org/10.1002/cmdc.201600563

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chemmedchem.org

ChemMedChem Full Paper

Inhibition of the Cysteine Protease Human Cathepsin L by Triazine Nitriles: Amide···Heteroarene π-Stacking Interactions and Chalcogen Bonding in the S3 Pocket

Maude Giroud,[a] Jakov Ivkovic,[a] Mara Martignoni,[a] Marianne Fleuti,[a] Nils Trapp,[a] Wolfgang Haap,[b] Andreas Kuglstatter,[b] Jörg Benz,[b] Bernd Kuhn,[b] Tanja Schirmeister,[c] François Diederich*[a]

[a] *Dr. Maude Giroud, Dr. Jakov Ivkovic, Mara Martignoni, Marianne Fleuti, Dr. Nils Trapp, Prof. Dr. François Diederich*
Laboratorium für Organische Chemie, ETH Zurich,
Vladimir-Prelog-Weg 3,
8093 Zurich (Switzerland)
Fax: (+41)44-623-1109
E-mail: diederich@org.chem.ethz.ch

[b] *Dr. Andreas Kuglstatter, Dr. Jörg Benz, Dr. Bernd Kuhn, Dr. Wolfgang Haap*
F. Hoffmann-La Roche Ltd.
Pharma Research and Early Development (pRED), Therapeutic Modalities,
Roche Innovation Center Basel,
Grenzacherstrasse 124,
4070 Basel (Switzerland)

[c] *Prof. Dr. Tanja Schirmeister*
Institut für Pharmazie und Biochemie, Johannes Gutenberg-Universität Mainz,
Staudinger Weg 5,
55128 Mainz (Germany)

Supporting Information for this article is available on the WWW under xxx.

10.1002/cmdc.201600563

We report an extensive "heteroarene scan" of triazine nitrile ligands of the cysteine protease human cathepsin L (hCatL) to investigate π-stacking on the peptide amide bond Gly67–Gly68 at the entrance of the S3 pocket. This heteroarene···peptide bond stacking was supported by a co-crystal structure of an imidazopyridine ligand with hCatL. Inhibitory constants K_i are strongly influenced by the diverse nature of the heterocycles and specific interactions with the local environment of the S3 pocket. Binding affinities vary by three orders of magnitude. All heteroaromatic ligands feature enhanced binding by comparison with hydrocarbon analogues. Predicted energetic contributions from the orientation of the local dipole moments of heteroarene and peptide bond could not be confirmed. Binding of benzothienyl (K_i = 4 nM) and benzothiazolyl (K_i = 17 nM) ligands was enhanced by intermolecular C–S···O=C interactions (chalcogen bonding) with the backbone C=O of Asn66 in the S3 pocket. The ligands were also tested for the related enzyme rhodesain.

Introduction

In 2013, we reported a computational study which predicted that the π-stacking of heteroarenes on peptide amide bonds,[1–4] approximated by *N*-methylacetamide (NMAC), becomes increasingly favorable with (a) decreasing π-electron density of the heterocycle and (b) stronger dipole moment of the heterocycle, when aligned in an antiparallel orientation to the local dipole moment of the peptide bond.[1b, 5–7] Experimental model host-guest chemistry studies with Rebek-type imide platform receptors[8,9] recently validated the first prediction (a).[10]

We subsequently described a fluorine scan (F-scan) of triazine nitrile inhibitors at the S3 pocket of the cysteine protease human cathepsin L (hCatL),[11,12] a versatile enzyme in our group for biomolecular recognition studies.[13] In this F-scan, we analyzed the interactions of differently fluorinated phenyl rings, which stack onto the planar peptide amide bond Gly67–Gly68 at the entrance of the S3 pocket of hCatL. This π-stacking had previously been confirmed by the co-crystal structure of a triazine nitrile ligand bearing a 4-chlorophenyl substituent occupying the S3 pocket (PDB ID: 4AXM, 2.8 Å resolution).[14] By varying degree and pattern of fluorine substitution, binding affinities (inhibitory constants K_i) could be enhanced by up to a factor 13 (1.8 kcal mol^{-1}) from the weakest to the best binder.[11] We explained the variation in binding affinity by (i) energetically more favorable π-stacking of less electron-rich, fluorinated rings on the peptide amide bond, (ii) favorable quadrupolar interactions of the fluorinated phenyl rings with the π-stacking amide bond, and (iii) specific interactions with the local environment in the S3-pocket of the enzyme. Contributions of dipolar interactions between the local dipole moments of the fluorinated rings and the Gly67–Gly68 peptide bond could not be confirmed, however, which could be due to the close

proximity of a second flat peptide bond fragment, Gly68–Leu69, with an opposite direction of its local dipole moment.

We previously identified triazine nitriles as potent inhibitors of rhodesain,[11,14,15] a central cysteine protease from *Trypanosoma brucei rhodesiense* and a potential drug target against human African trypanosomiasis (sleeping sickness).[16,17]

While π-stacking on peptide amide bonds by heteroaromatic rings is ubiquitous in protein-ligand complexation,[6,18] systematic experimental investigation to identify the best (hetero)aromatic partner for stacking on a peptide amide bond is scarce. Following our computational predictions,[1b] we therefore initiated an extensive heteroarene scan of triazine nitrile inhibitors, addressing the π-stacking of different heterocycles on the peptide amide bonds in the flat dipeptide backbone fragment Gly67–Gly68–Leu69 at the entrance of the S3 pocket of hCatL. Here, we report the synthesis and biochemical evaluation of a larger series of 35 new ligands, which contain different or differently attached heteroarenes to fill the S3 pocket, while the S1 and S2 pocket substituents departing from the central triazine scaffold are kept constant. We compare their complexation to a series of ligands with benzenoid hydrocarbon substituents or lacking the S3 substituent entirely. We analyze their binding affinities to decipher effects of nature, size, and type of attachment of the heterocycles, explore contributions from favorable orientations of local dipole moments in the π-stacking array and specific interactions of the heteroarenes with the local environment of the S3 pocket. In this comprehensive analysis, we obtained strong evidence for intermolecular S···O chalcogen bonding interactions with the backbone C=O of Asn66 at the entrance of the pocket,[19,20] stabilizing the binding of sulfur-containing heteroarenes. A similar analysis of the complexation of the new ligands to the cysteine protease rhodesain (RD) is presented in the Supporting Information (SI).

Results and Discussion

Ligand Design

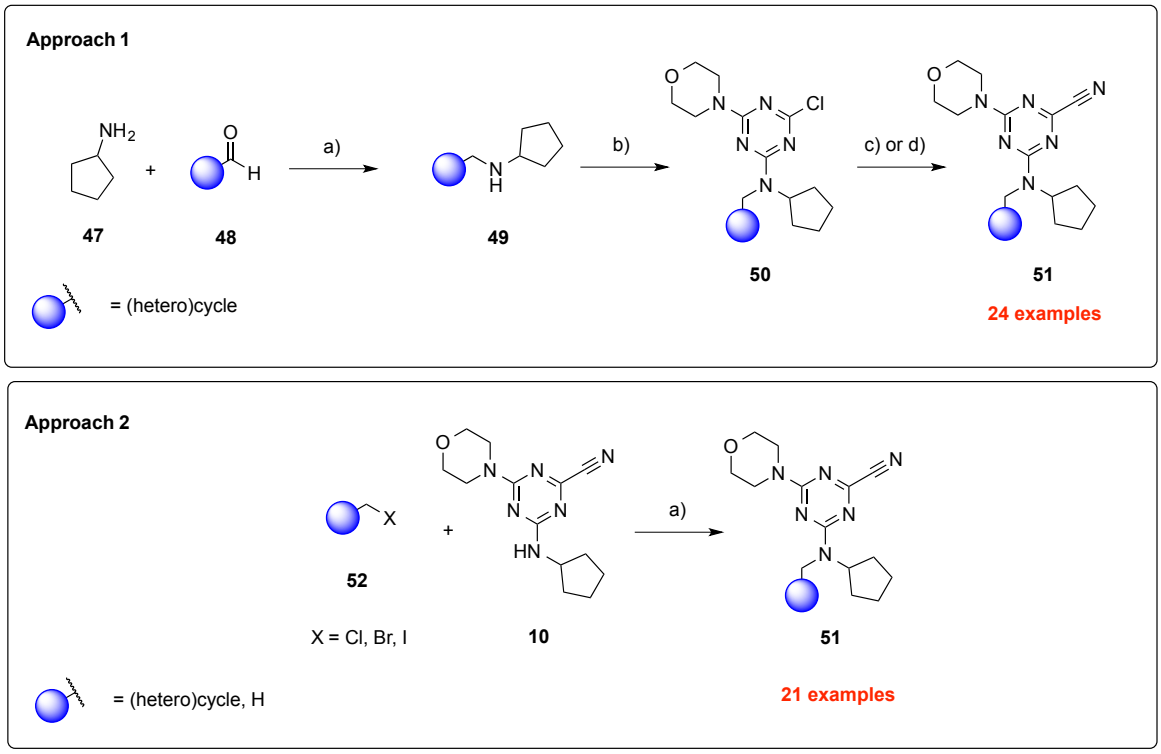
Triazine nitriles (**1**–**46**, Table 1) undergo reversible-covalent binding with the catalytic Cys25 at the active site of hCatL attacking the nitrile group and forming a thioimidate stabilized by the oxyanion hole, as seen in the previously mentioned co-crystal structure (PDB ID: 4AXM, 2.8 Å resolution).[14] In all ligands, the substituents for the S1 and S2 pockets, departing from the central triazine scaffold, were kept identical. A morpholinyl substituent reaches into the S1 sub-pocket, enhancing solubility rather than providing additional binding affinity. The S2 pocket is filled by a cyclopentyl ring, which makes a large contribution to the stability of the complex with hCatL.[11,14,15] The reference inhibitor **1** contains a phenyl substituent reaching into the S3 pocket, with an inhibitory constant K_i (hCatL) = 520 nM.[11] This substituent was systematically changed in the other ligands. Modeling with MOLOC[21] using the MAB force field suggested that the overall binding mode remained similar for all ligands. During optimization, the protein coordinates as well as the position of the cyano group in the ligand, covalently bound to Cys25 in the form of a thioimidate, were kept fixed, except for the flexible Glu63 side chain in the S3 pocket, as previously reported.[11] The phenyl ring of reference system **1** was predicted to stack on the Gly67–Gly68 peptide bond at a distance of 3.6 Å (Figure S1 in the SI),[11] which is the optimal distance according to our initial computational study.[1b] This model system adopts a similar orientation to the previously reported experimental structure (PDB ID: 4AXM)[14] (Figure S1c).

Table 1

Synthesis

Triazine nitriles **1**–**46** were prepared according to two known synthetic routes (Scheme 1).[11,14,15] These synthetic approaches were chosen for their robustness and reliability. In the first approach, by which 24 compounds were prepared, reductive amination[22] of cyclopentylamine (**47**) and an aromatic aldehyde (**48**), in the presence of $NaHB(OAc)_3$ and molecular sieves (4 Å), generated a secondary amine of general structure **49**. Sequential addition of amine **49** and morpholine to cyanuric chloride, in the presence of Hünig base, gave triazine chloride **50**. Cyanodehalogenation of **50** with KCN in Me_2SO or KCN/DABCO in Me_2SO/H_2O 9:1 at 80–120 °C provided triazine nitrile **51**. In the second approach, by which 21 compounds were prepared, addition of benzylic bromides or chlorides **52** to triazine nitrile **10**,[11] pretreated with NaH, gave the desired final compounds of general formula **51**. Similarly, ligand **10** was methylated with MeI to afford product **11**. A complete description of synthesis and analytical characterization of the new ligands and their precursors is included in Section S6 in the SI. Small-molecule X-ray crystallographic structures were obtained for triazine nitriles **10**, **43**, and **45** upon slow evaporation of $EtOH/CH_2Cl_2$ 2:1 solutions, confirming their assigned structures (Section S5 in the SI).

10.1002/cmdc.201600563



Scheme 1. Overview of the approaches used to prepare ligands **2**–**46**. Approach 1: a) 1. CH_2Cl_2, 4 Å MS, 1 h at 23 °C, 2. $NaHB(OAc)_3$, CH_2Cl_2, MS (4 Å), 4–20 h at 23 °C; b) 1. Cyanuric chloride, *i*Pr_2NEt, CH_2Cl_2, 2–6 h at 0 °C, 2. Morpholine, *i*Pr_2NEt, CH_2Cl_2, 2–16 h at 23 °C; c) KCN, Me_2SO, 18–20 h at 120 °C; d) KCN, DABCO, 9:1 Me_2SO/H_2O, 18–20 h at 80–120 °C. Approach 2: a) 1. **10**, NaH, DMF, 30 min to 1 h at 23 °C, 2. **52**, 3–18 h at 23 °C. DABCO = 1,4-Diazabicyclo[2.2.2]octane, MS = molecular sieves, DMF = *N*,*N*-dimethylformamide.

Biochemical and Physicochemical Studies

The binding affinities of ligands **1**–**46** were tested against hCatL in a fluorimetric assay[14,23] in aqueous buffer (50 mM tris pH 6.5, 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35) at 298 K (Table 1). Physicochemical data were also acquired: distribution coefficients log$D_{7.4}$ were measured in a Carrier Mediated Distribution System (CAMDIS©) assay[24] and the solubility in a Lyophilisation Solubility Assay (LYSA).[25] In addition, the *c*log$D_{7.4}$ values were predicted for all reported compounds with a Roche in-house developed machine-learning tool based on a tree-based ensemble method.

Exploring Halogen and Chalcogen Bonding to the Backbone C=O of Gly61 in the S3 Pocket

In a previous systematic study with a structurally very different class of reversible-covalently binding nitrile inhibitors,[13] we had observed strong gains in binding potency (up to a factor 74, translating into a gain in free enthalpy of $-\Delta\Delta G$ = 2.6 kcal mol^{-1}) by establishing halogen bonding (XB)[26] to the conformationally locked backbone C=O of Gly61 in the S3 pocket.[13] We therefore investigated whether 4-halophenyl substituents of triazine nitrile inhibitors would also establish such intermolecular interactions. The comparison between ligand **1** and the halogenated derivatives **3**–**5** however showed that XB is not very effective in this system yielding a maximum 7-fold gain in K_i when adding a bromine atom in *para* position of the phenyl ring (Table 1). Molecular modeling suggested (Figure S2) that the substituted triazine nitrile ligand scaffold prevents the halophenyl ring from penetrating sufficiently deep into the pocket to establish XB with the C=O at an O···X distance (X = Cl, Br, I) significantly below the sum of the van der Waals radii.[27] Also, in the formerly reported class of hCatL ligands, the halophenyl ring was much more pre-organized for XB than the corresponding ring in the triazine nitriles. We therefore abstained from synthesizing ligands with halogenated six-membered heteroarenes as S3 pocket vectors.

Stimulated by the recent interest in chalcogen bonding,[19,20] we also prepared the 4-methylsulfanyl derivative **6**. With a K_i-value of 13 nM (LipE = 3.4), **6** is among the most potent of the ligands shown in Table 1. MOLOC proposes a binding geometry in which the phenyl ring stacks in the middle of the flat dipeptide Gly67–Gly68–Leu69 fragment, orienting the S-atom of **6** to both chalcogen-bonding distance and geometry with the backbone C=O of Gly61 (d(S···O=C$_{Gly61}$) = 3.6 Å, α(O$_{Gly61}$···S–C) = 162°) (Figure S3a). The methoxy derivative **7**, which is not capable of undergoing chalcogen bonding, is however a similarly strong binder (K_i =18 nM, LipE = 3.9) (Figure S3b). We therefore conclude that the origin of

the enhanced affinity of **6** and **7** is the interaction of the ether/thioether methyl group with the local lipophilic environment of the S3 pocket: both groups undergo apolar contacts at van der Waals distance with the phenol ring of Tyr72 and the isobutyl group of Leu69 lining one side of the S3 pocket (Figure S3). Probably, enthalpically strained water molecules in this lipophilic environment are also displaced by the S3 pocket vector in an energetically beneficial way.[26a]

π-Stacking Interactions of 5- and 6-Membered Ring Heteroarenes with the Gly67–Gly68 Peptide Bond at the Entrance of the S3 Pocket of hCatL

In agreement with the earlier co-crystal structure (PDB ID: 4AXM),[14] modeling with MOLOC consistently proposed π-stacking of the heteroarene substituents in ligands **12**–**26** with the Gly67–Gly68 peptide amide bond at the entrance of the S3 pocket of hCatL. Based on new co-crystal structure data, however (see below), we cannot exclude additional, different orientations of the monocyclic heteroarene rings, binding outside the S3 pocket. Conformational flexibility of the ligands in the free state is also supported by the small-molecule X-ray crystal structures in which the S2 and S3 pocket vectors do not always adopt the same orientation as in the predicted bound state (Section S5 in the SI, for similar findings, see [11]).

The triazine nitrile ligands with five- (**12**–**16**) or six-membered ring (**17**–**26**) heteroarenes as S3 pocket vector bind to hCatL with similar or higher affinity to the reference phenylated ligand **1**. All K_i-values are in or near the triple-digit nanomolar range (pK_i values between 6.4 and 7.2). We prepared fluoropyridine derivatives **24**–**26** to lower the pK_a[28] and prevent possible protonation of pyridines in the negative electrostatic potential area of the flexible anionic side chain of Glu63 in the S3 pocket. They possess very similar activity to their fluorophenyl analogues,[11] and are slightly less potent than their pyridine analogues **17**–**19**.

The larger series of six-membered *N*-heteroarenes **17**–**23** was initially prepared to evaluate the predicted[1b] energetic contributions from different local dipole moment orientations of the heteroarene stacking on a peptide amide bond. The dipole moments of the five- and six-membered heterocycles were calculated at the B3LYP/6-31+G* or MP2/3-21G level (see Section S1.3 in the SI). Red arrows in Table 1 represent these dipole moment vectors schematically for the six-membered N-heterocyclic ligands. Because inhibitors with opposite local dipole moments possess similar binding affinities (e.g., **20** *vs*. **22**), we could exclude significant effects of the dipole moment. From this and the previous work on a fluorine scan of S3 pocket substituents,[11] we can conclude that the Gly67–Gly68 peptide amide bond in hCatL is not suitable for experimentally testing the computational predictions of dipolar alignment contributing to heteroarene···amide bond π-stacking. The main reason is the close proximity of a second, coplanar Gly68–Leu69 peptide bond with opposite direction of its local dipole moment (see Figure S1).

We also did not find any correlation of hCatL binding strength with the quadrupole moments of the heteroarenes[29] in the ligand series **17**–**23** (Figure S4). We finally examined contributions of the experimentally determined lipophilicity of the ligands to binding affinity, but no correlation between pK_i and log$D_{7.4}$ values was observed (Figure S5).

π-Stacking Interactions of Bicyclic Heteroarenes with the Gly67–Gly68 Peptide Bond at the Entrance of the S3 Pocket of hCatL

A large difference in binding affinity of three orders of magnitude was observed when the heteroarene was extended to fused bicycles in ligands **27**–**46**. One of the weakest ligand with a 2-quinolyl substituent (**45**) only bound in the micromolar range (K_i = 1.45 μM, LipE = 1.6) whereas 2-benzothienyl ligand **27** is the strongest prepared, with K_i = 4 nM (LipE = 4.1). For two heteroarene ligands, 3-benzothienyl derivative **40** and 8-quinolyl derivative **43**, as well as

for 1-naphthyl ligand **8**, the inhibition constants could even not be measured with confidence as they are in the higher micromolar range. Modeling suggested that all these ligands fit sterically into the S3 pocket in the active site of hCatL. As the stacking on the flat peptide bond fragment Gly67–Gly68–Leu69 was not much affected by the position of the heteroatoms in the monocyclic ligands, specific interactions with the local S3 pocket environment should explain the differences in binding affinity, besides proper orientation of the bicycle in the pocket, as well as conformational effects (see below).

The high affinity of the piperonyl substituent (in **36**, K_i = 41 nM, LipE = 3.6) for the S3 pocket of the related cysteine proteases rhodesain and falcipain-2 had already been observed in our earlier work on triazine nitrile ligands,[14] but remains poorly understood in the absence of a co-crystal structure. Binding in the double-digit nanomolar range is also exhibited by the quinoxaline ligand **46** (K_i = 56 nM, LipE = 3.2) and the indazole derivative **34** (K_i = 72 nM, LipE = 3.4). The latter could posssibly form a H-bond with its N–H to the flexible side chain of Glu63, as suggested by modeling.

The strongest binding, however, was observed for four ligands with similar topology and attachment of the S3 pocket vector to the triazine core: 2-benzothienyl (**27**, K_i = 4 nM, LipE = 4.1), 2-benzofuranyl (**28**, K_i = 11 nM, LipE = 3.9), 2-benzoxazolyl (**29**, K_i = 17 nM, LipE = 3.7), and 2-imidazopyridine (**30**, K_i = 35 nM, LipE = 3.5). Gratifyingly, a co-crystal structure of the latter with hCatL was solved.

X-Ray Co-Crystal Structure Analyses with hCatL

A co-crystal structure of imidazopyridine ligand **30** bound to hCatL was solved at 1.0 Å resolution (PDB ID: 5MAJ, Figure 1). The complex crystallized in the $P2_12_12_1$ space group, with one chain of mature enzyme in the asymmetric unit of the crystal. The observed binding mode of ligand **30** is in accordance with our design. The ligand is covalently bound as a

thioimidate, formed by nucleophilic attack of Cys25 on the triazine nitrile. The morpholine moiety addresses the S1 pocket, and the cyclopentyl ring is located within the S2 pocket. The imidazopyridine moiety is directed into the S3 pocket and stacks on top of the Gly67–Gly68 peptide amide bond, at a distance of 3.3–3.9 Å (Figure 1b). The calculated local dipole of the heterocycle and the peptide bond Gly67–Gly68 are oriented in an antiparallel fashion (Figure 1c), but as already stated, the presence of the second amide bond Gly68–Leu69 with opposite local dipole orientation in close proximity complicates the analysis of dipolar contributions to the π-stacking interactions.

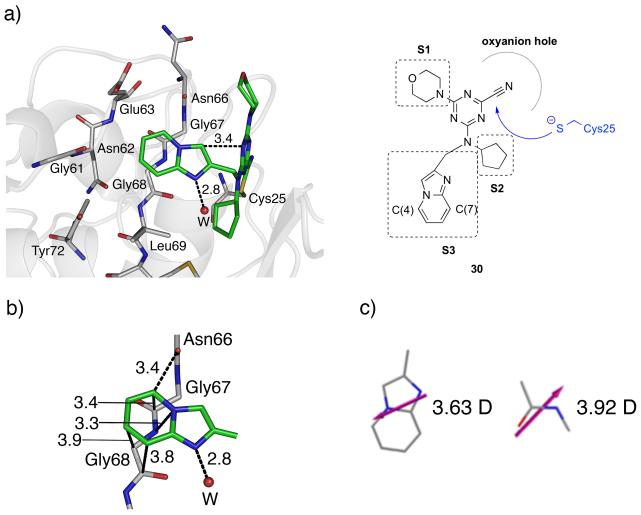


Figure 1. a) Co-crystal structure of **30** with hCatL (PDB ID: 5MAJ, resolution 1.0 Å). b) Position of the imidazopyridine moiety on top of the Gly67–Gly68 peptide fragment. c) Representation of the calculated dipole moments (at the B3LYP/6-31+G* level) for the imidazopyridine ring and NMAC. The arrows are proportional to the strength of the dipole moment. Color code: gray C_{enzyme}, green C_{ligand}, red O, blue N, yellow S. Distances given in Å. Torsion angle τ (N–CH_2–C_{sp}–C_{sp}) = 94°. W = water molecule.

An interaction analysis of the co-crystal structure of **30** in complex with hCatL was performed using the Scorpion network approach.[30] The Scorpion software tries to identify different types of favorable receptor–ligand interactions and uses as network the analysis of

these contacts to estimate atom-based energy contributions to ligand binding (see Section S2 in the SI). This analysis identified carbon atoms C(4) and C(7) of the imidazopyridine ring as binding hot-spots (Figure S6). Most interactions are indeed emerging from the contact of the imidazopyridine ring with the S3 pocket of hCatL. Dispersive interactions with Gly67, Gly68, and Leu69, as well as π-stacking interactions with the backbone C=O and NH of Gly68 and the backbone C=O of Gly67 are detected as binding contributions. Additionally, a C–H···O H-bond between H–C(4) and the backbone C=O of Asn66 was identified, as well as a N···H–O H-bond between N(1) of the imidazopyridine ring and a water molecule (Figures 1 and S6).

The thioimidate moiety is stabilized by a H-bond with the side chain NH_2 of Gln19 in the oxyanion hole, and the thioimidate C=NH bond engages in a dipolar interaction with the backbone C=O of Gly23 (Figure S6b). The central triazine ring stacks on the backbone C=O of Asp162. The triazine core should engage in O–H···π interactions with resolved water molecules. The cyclopentyl ring addressing the S2 pocket engages in dispersive interactions with the Met70 and Ala135 side chains. In the S1 pocket, no favorable or unfavorable interactions were identified. Some crystal contacts between the imidazopyridine ring in the S3 pocket and a second adjacent protein unit were observed (see below).

Another co-crystal structure was obtained for the pyrazine ligand **23**, which is bound to hCatL in a reversible-covalent manner (PDB ID: 5MAE, 1.0 Å resolution) (Figure 2). This complex also crystallized in the $P2_12_12_1$ space group. Two morpholine conformations in the S1 pocket were observed with about equal occupancy. The cyclopentyl moiety again fills the S2 pocket. The pyrazine ring, however, does not stack on the peptide bond at the entrance of the S3 pocket but adopts another orientation.

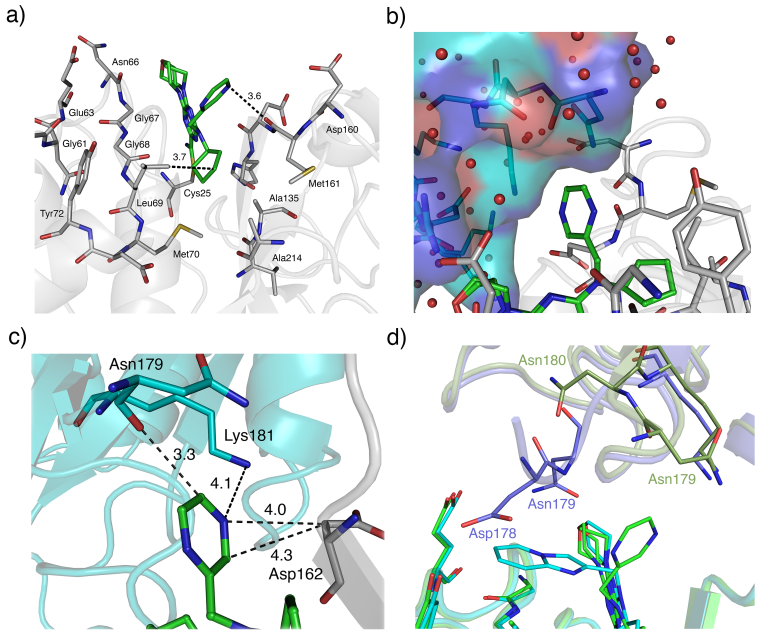


Figure 2. a) Co-crystal structure of **23** in hCatL (PDB ID: 5MAE, resolution 1.0 Å). b) Binding of the pyrazine ring in the interstitial void. c) Contacts of the pyrazine ring with the second protein. Color code: gray $C_{monomer\_1}$, cyan $C_{monomer\_2}$ green C_{ligand}, red O, blue N, yellow S, red dots: water molecules. Distances given in Å. Two conformations of the morpholine were observed with about equal occupancy. d) Overlay of the co-crystal stuctures of **23** (PDB ID: 5MAE) and **30** (PDB ID: 5MAJ) in hCatL, showing their two different cavities with a second protein monomer. Color codes: green C_{5MAE}, dark green $C_{5MAE\_monomer2}$ cyan C_{5MAJ}, violet $C_{5MAJ,monomer\_2}$ red O, blue N.

There are driving forces for the preference of the pyrazine ring to bind into this void between proteins. (i) The torsional angle τ for fixation of the pyrazine to the triazine core ($\tau(C_{sp2}–C_{sp2}–C_{sp3}–N) = 26°$) is close to the corresponding one seen in the crystal structure for benzene derivative **1** ($\tau(C_{sp2}–C_{sp2}–C_{sp3}–N) = 11°$). A substantial investment of conformational energy is required to change this angle to 90° for π-stacking on the peptide bond at the entrance of the S3 pocket (see below); which does not have to be paid upon binding into the void in the solid state. (ii) If the pyrazine stacks on the Gly67–Gly68 peptide amide bond at the entrance of the S3 pocket, which can occur in two conformations by rotation of the heteroarene by 180°, the pyrazine N-atoms undergo repulsive intermolecular lone-pair interactions, according to MOLOC modeling (Figure S7). In one conformation, one pyrazine

N-atom approaches the backbone C=O of Asn66; in the second, this N-atom comes close to the anionic side chain of Glu63, with additional desolvation of the second N-atom by the lipophilic side chain of Leu69.

In addition, we analyzed the crystal packing to find a rationale for this unexpected binding geometry in the solid state and observed that two neighboring proteins form an interstitial cavity in which the pyrazine ring is located and benefits from attractive interactions (Figure 2b,c). In this interstitial pocket, the pyrazine ring undergoes favorable C–H···O interactions with the backbone C=O of Asn179 (d(C···O=C_{Asn179}) = 3.3 Å), weak ionic H-bonding with a side chain NH of Lys181 (d(N···H–N_{Lys181}) = 4.1 Å), and C–H···π interactions with the CH_2 of Asp162 ring (d(C···H–C_{Asp162} = 4.0–4.3 Å). The cavity at the interface between the two proteins has shape-complementarity to the six-membered pyrazine but is too small to accommodate bicyclic heteroarene substituents. The interstitial binding is reminescent of clathrate formation,[31] often observed in supramolecular and other chemistries. This pyrazine conformation would not be observed in the hCatL complex in solution, as the heterocycle would be exposed to bulk water. For the complex of **30** with hCatL, such cavity does not exist in the same way. In this case, the Asn179 side chain is moved over to the heteroarene (Figure 2d). The imidazopyridine ring undergoes contacts with both Asp178 and Asn179, which stabilize the ring in the S3 pocket. In addition, the preferred torsional angle (τ(C_{sp2}–C_{sp2}–C_{sp3}–N) ~ 110°) for the imidazopyridine ring makes it pre-organized for binding in the S3 pocket.

Intermolecular S···O Interactions

The sulfur-containing bicyclic heteroarenes benzothiophene **27** (K_i = 4 nM, LipE = 4.1) and benzothiazole **29** (K_i = 17 nM, LipE = 3.7) are among the strongest binders. Sulfur possesses unusual stereoelectronic properties,[32] and has been reported to undergo chalcogen bonding,[19,20] an interaction similar to halogen bonding.[13] Chalcogen-containing

heterocycles possess a positive electrostatic potential adjacent to the sulfur atom consistent with the "σ-hole" concept.[19b] The magnitude and volume of this σ-hole increases from thiophene to thiazole.[20a] A comprehensive summary of intermolecular O···S–C n → σ* interactions reported in the literature has been provided by Meanwell and co-workers.[19b]

We modeled benzothiophene ligand **27** into the active site of hCatL by modifying the imidazopyridine **30** in the co-crystal structure and optimizing the new complex with MOLOC. The resulting structural proposal clearly revealed a chalcogen bond from the benzothiophene donor to the backbone C=O of Asn66 as acceptor (Figure 3). The position of this backbone C=O is highly conserved in co-crystal structures of hCatL, which makes it entropically a favorable interaction partner.[13,33] The interaction occurs at a distance d(S···O=C) = 3.5 Å and at an angle α(O···S–C) = 158°, well within the range of chalcogen bonds observed in other protein-ligand complexes (Figure 3).[19] The 2-benzothiazolyl ligand **29** shows a similar chalcogen bond to the C=O of Asn66 (Figure S8a), but complexation is weaker (K_i = 17 nM, LipE = 3.7). A weak chalcogen bonding may also explain the good binding of thiazole **16** (K_i = 69 nM, LipE = 3.2) in the monocyclic heteroarene ligand series (Table 1). Chalcogen bonding is generally stronger for thiazoles than for thiophenes.[19] While thiazolyl **16** binds better than thienyl **12** (3-fold), the trend was inversed between benzothiazolyl **29** and benzothienyl **27**. The predicted binding mode of **29** in hCatL (Figure S8a) reveals that the C–S···N(4) angle (141°) is not ideal for an intramolecular N···S interaction. In the unbound state, ligand **29** is in a more planar conformation (Figure S8b), enabling a better intramolecular N···S interaction (α(C–S···N(4)) = 164°). This might cost energy upon reorganization for binding of **29**, and explain the weaker binding by comparison with ligand **27**. The benzofuran ligand **28** (K_i = 11 nM, LipE = 3.9) cannot undergo chalcogen bonding and most probably forms a weak C–H···O H-bond with the C=O of Asn66 (Figure S9**)**, with

the O-atom becoming solvated by a water molecule as seen for the N-atom in the imidazopyridine crystal structure.

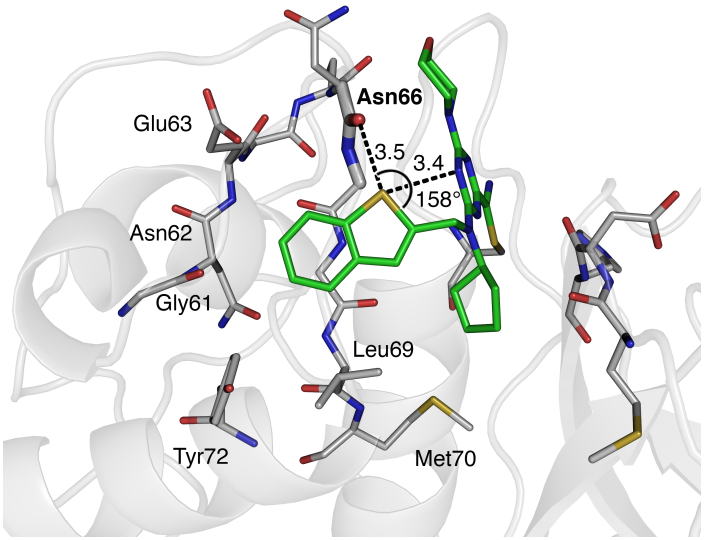


Figure 3. Predicted binding mode of **27** with hCatL (protein coordinates taken from PDB ID: 5MAJ, resolution 1.0 Å), displaying the chalcogen bond (d(O···S) = 3.5 Å, α(O···S–C) = 158 °, d(S···N(4)) = 3.4 Å). Color code: gray C_{enzyme}, green C_{ligand}, red O, blue N, yellow S. Distances given in Å.

Apart from the analysis of intermolecular interactions in the complex, it is important to consider potential ligand strain when trying to rationalize the observed SAR. When comparing the ligand conformations in the protein-bound state with the one found in small molecule crystal structures, we noticed a significant difference in the torsional angle along the benzylic CH_2–C_{sp2} bond of the S3 pocket vector, τ (N–CH_2–C_{sp}–C_{sp}). In both the complex of the 4-chlorophenyl ligand with hCatL (PDB ID: 4AXM)[14] and in the structure of compound **30** (Figure 1), the dihedral angles are 94°, while in the small-molecule crystal structure of triazine nitrile **1** it is τ = 11° (CCDC 1449741).[11] A torsion angle of around 90° is required for stacking onto the peptide amide backbone at the entrance of the S3 pocket. To better assess the energetic implications of this angle difference, we performed quantum mechanical

torsion scans for ligands **1** and **27–30** at the B3LYP/cc-pVDZ level using Gaussian 09[35] (Figure 4).

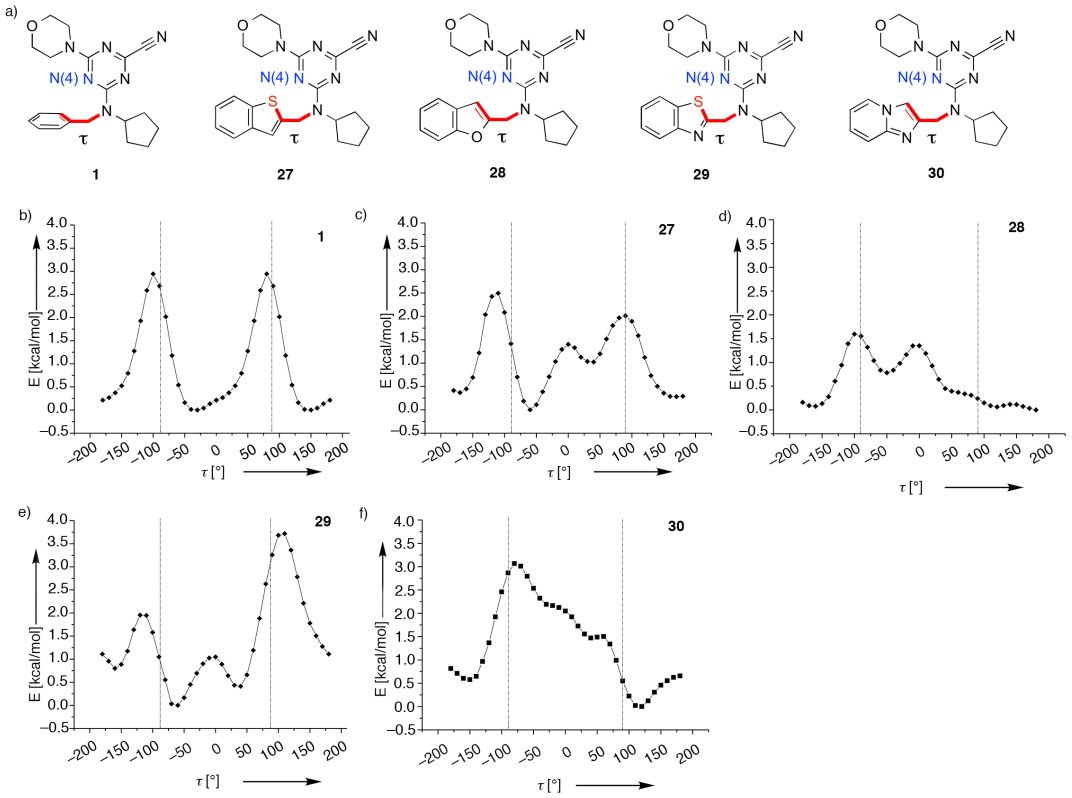


Figure 4. Torsional angle scans of triazine nitriles **1**, **27–30** calculated at the B3LYP/cc-pVDZ level using Gaussian 09.[35] The definition of the torsion angle τ is shown in red, and the 2D ligand depictions show the preferred orientation of the S3 pocket vector relative to the triazine core based on the torsional energy profiles.

For the phenyl derivative **1** the dihedral angle at the minimum of the energy profile is close to the value observed in the small molecule structure. In contrast, a relative energy of 2.7 kcal mol^{-1} is calculated for $\tau = 90°$ suggesting substantial ligand strain for this compound at the torsional angle required for stacking in the S3 pocket. Relative energies for at least one of the torsional angles $\tau = -90°, 90°$ are lower for all four heterocycles in **27–30**, indicating that reduced ligand strain for these compounds contributes to their improved hCatL activity compared to **1**. The torsional profiles in Figure 4 provide further information on the preferred orientation of the heterocycles in a stacked binding conformation in the S3 pocket. For ligands **27–30**, a significant energy difference > 1 kcal mol^{-1} exists for one of the two

orientations at $\tau = -90°$, 90°. In all these cases, the preferred orientation corresponds to an *anti* alignment of the O and N lone pairs in the heterobicycle with the N-atom (N(4)) of the triazine ring thereby avoiding electrostatic repulsion. For the imidazopyridine **30** this is indeed observed in the co-crystal structure (Figure 1). In the two sulfur-containing heterocycles benzothiophene **27** and benzothiazole **29**, the torsional profiles suggest that the S atom of the ligand is preferentially *syn* to the N(4) of the triazine nitrile. This is in line with the postulated chalcogen-bonding geometries of benzothiophene **27** (Figure 3) and benzothiazole **29** (Figure S8). In these orientations, the ligand sulfur atom would then interact intermolecularly with the C=O O-atom of Asn66 and intramolecularly with the π-orbital on the triazine N-atom using its second σ-hole. The torsional ligand strain to adopt the stacked binding mode in the S3 pocket of hCatL is predicted to be 1.1 kcal mol^{-1} for benzothiazole **29**, and 1.4 kcal mol^{-1} for benzothiophene **27**. Interestingly, this is significantly higher than for the benzofuran **28** (0.2 kcal mol^{-1}) and imidazopyridine **30** (0.6 kcal mol^{-1}). It suggests that there are compensating attractive intermolecular interactions, which we attribute mainly to the $C{=}O_{Asn66}\cdots S{-}C_{sp2}$ chalcogen bonding. Similar conformational strain analyses are undoubtedly needed when trying to rationalize the large affinity differences – by three orders of magnitude – between the entire set of 20 ligands with the various heterobicyclic S3 pocket vectors included in Table 1.

Conclusions

We prepared and investigated a large series of triazine nitrile ligands featuring mono- and bicyclic heteroarene substituents for binding to the S3 pocket of the cysteine protease human cathepsin L (hCatL). A similar analysis for the binding of these ligands to the related enzyme rhodesain (RD) is also reported in the Supporting Information (SI). The ligands undergo reversible-covalent binding, with the nitrile reacting with the catalytic cysteine under

formation of a thioimidate moiety. Their substituents for the S1 and S2 pockets of the enzyme were kept constant. Molecular modeling proposed for all ligands that the heteroarene substituents could undergo π-stacking with the Gly67–Gly68, and partially also with the neighboring Gly68–Leu69 peptide amide bond at the entrance of the S3 pocket. Ligands featuring monocyclic five- or six-membered heteroarenes bind with inhibitory constants K_i within or near the triple-digit nanomolar range. They show higher potency than the reference ligand with a phenyl substituent, providing another example for the preference of peptide amide bonds to undergo π-stacking with heteroaromatic rings. The predicted contributions from the orientation of the local dipole moments of the heteroarenes and the peptide amide bond could not be confirmed in this study, which could be due to the close proximity of two flat peptide bonds, with oppositely oriented local dipole vectors, at the entrance of the S3 pocket.[6a]

An unexpectedly large difference of three orders of magnitudes (K_i values between 4 nM and higher micromolar) was measured for the binding affinity of ligands featuring bicyclic heteroaromatic rings. The highest binding affinity was measured for four ligands with similar topology and attachment of the S3 pocket vector, 2-benzothienyl (**27**, K_i = 4 nM, LipE = 4.1), 2-benzofuranyl (**28**, K_i = 11 nM, LipE = 3.9), 2-benzothiazolyl (**29**, K_i = 17 nM, LipE = 3.7), and 2-imidazopyridinyl (**30**, K_i = 35 nM, LipE = 3.5). The X-ray co-crystal structure of the imidazopyridine ligand in complex with hCatL was determined and confirmed the predicted binding mode. The binding of these ligands was subjected to a more thorough analysis, including an interaction network analysis of the co-crystal structure and a detailed conformational analysis, analyzing the torsional ligand strain when adopting the binding conformation required for π-stacking with the flat dipeptide fragment at the entrance of the S3 pocket. Molecular modeling based on the crystal structure revealed that the complexes of the benzothiophenyl and the benzothiazolyl ligands are additionally stabilized by chalcogen

bonding to the backbone C=O of Asn66. The interaction features suitable geometric parameters for chalcogen bonding, with the distance $d(S\cdots O{=}C_{Asn66})$ = 3.5 Å and the angle $\alpha(O_{Asn66}\cdots S{-}C)$ = 158°. The conformational analysis further supports this interaction, which compensates for the higher torsional strain in the S-containing ligands as compared to the benzofuranyl and imidazopyridinyl ligands. This study shows that the analysis of intermolecular interactions needs to be accompanied by conformational strain analysis, when trying to rationalize the large affinity differences between the entire set of ligands featuring various heterobicyclic S3 pocket vectors. It confirms that optimal design of the S3 pocket vector can strongly enhance affinity to the cysteine proteases hCatL and RD.

Experimental Section

Since the synthetic protocols had been previously developed and applied,[11,14,15] all synthesis and characterization of the new ligands is found in Section S6 of the SI.

Enzymatic Assays

The determination of the activity of the inhibitors against hCatL and RD was performed in fluorescence-based assays. In the following, the procedure for the hCatL inhibition assay is described in detail. The assays with rhodesain were carried out similarly.

The biological activities against hCatL were determined using Cbz-Phe-Arg-AMC as substrate, which releases AMC (7-amino-4-methylcoumarin) after amide bond cleavage by the enzyme.[14,23] The proteolytic activity of the enzyme can be monitored spectrophotometrically by the increase of fluorescence intensity by release of AMC (emission at 460 nm) upon hydrolysis.

An initial screen at an inhibitor concentration of 20 μM was performed to identify ligands with an inhibition of hCatL and RD higher than 80%. For active compounds, continuous assays at different inhibitor concentrations were carried out. The residual enzyme activities were plotted against the inhibitor concentrations to obtain the experimentally determined IC_{50} values. The inhibitory constants (K_i) were calculated according to the Cheng-Prusoff equation,[34] as triazine nitriles have previously been determined to be competitive inhibitors and do not show any time-dependent inhibition:[15,36]

$$K_i = \frac{IC_{50}}{1 + \left(\frac{S}{K_m}\right)}$$

with S as substrate concentration and K_m as Michaelis-Menten constant[34] (K_m = 6.5 μM for Cbz-Phe-Arg-AMC).

Determination of hCatL Inhibition. The protein was purchased at Calbiochem. The assay buffer consisted of 50 mM Tris pH 6.5, 5 mM EDTA, 200 mM NaCl, and 0.005% Brij35, and the enzyme buffer of 50 mM Tris pH 6.5, 5 mM EDTA, 200 mM NaCl, and 2 mM DTT. The substrate Cbz-Phe-Arg-AMC (K_m = 6.5 μM, Bachem, purchased as HCl salt) was diluted from a 1 mM stock solution to a final concentration of 25 μM. The assay mixtures had a total volume of 200 μL and consisted of 180 μL assay buffer, 5 μL enzyme, 10 μL Me_2SO (as negative control) or inhibitor solution (prepared from a stock solution with c = 2 mM), and 5 μL of the substrate solution. An initial screening was performed at an inhibitor concentration of 20 μM (final inhibitor concentration) over 5 min in two independent measurements to identify compounds with an inhibition > 80%. Enzyme was added to the substrate dilutions, and the reaction was initiated upon addition of the substrate. For active inhibitors, dilution series starting from a 20 mM stock solution in Me_2SO at seven different concentrations were prepared in duplicate for each compound at 25 °C. Fluorescence emission was measured over 30 min on a Cary Eclipse Fluorimeter (Varian, Darmstadt, Germany, excitation 380 nm, emission 460 nm) equipped with a heating device in 96-well microtiter plates. Kinetic constants are the mean values of at least two independent assays, each performed in duplicate. The GraFit® software (version 5.0.13, 2006, Erithacus Sofware Ltd., UK) was used for nonlinear regression, and the kinetic constants K_i were calculated based on the IC_{50} and the K_m values using the Cheng-Prusoff equation.

Determination of RD Inhibition. Assay buffer: 50 mm sodium acetate pH 5.5, 5 mm EDTA, 200 mM NaCl, and 0.005% Brij35. Enzyme: expressed and purified in the group of Prof. Caroline Kisker (Rudolf-Virchow-Zentrum, University of Würzburg, Germany). Enzyme buffer: 50 mM sodium acetate pH 5.5, 5 mM EDTA, 200 mM NaCl, and 2 mM DTT. Substrate: Cbz-Phe-Arg-AMC (K_m = 825 nM, Bachem, purchased as HCl salt).

Co-Crystallization of hCatL in Complex with 23 and 30

hCatL (produced as previously described)[13a] at a concentration of 2.5 μM in 100 mM sodium acetate pH 5.5, 5 mM DTT, 5mM EDTA, 0.02% NaN_3 was incubated with ligands **23** and **30** in a 12 to 15-fold molar excess overnight at 4 °C under argon. Prior to crystallization experiments, the protein-ligand mixtures were concentrated to 27–32 mg mL^{-1} and centrifuged at 20,000 g. The crystallization droplets were set up at 22 °C by mixing 0.15 μL of protein solution with 0.3 μL reservoir solution and 0.02 μL seed solution in sitting drop vapour diffusion experiments. For compound **23** crystals were obtained within two days out of 25% PEG 3350, 0.1 M Bis-Tris pH 6.5. The complex with compound **30** crystallized within one day out of 25% PEG3350, 0.2 M ammonium acetate, 0.1 M HEPES pH 7.5. Details on data collection, processing, and refinement statistics for derivatives **23** and **30** are listed in Table 2.

Table 2. Statistics for X-ray data processing and model refinement for co-crystal structures of triazine nitriles **23** and **30** with hCatL.

	30	**23**
PDB accession number	5MAJ	5MAE
Data Processing[a]		
Space Group	$P2_12_12_1$	$P2_12_12_1$
Unit cell axes [Å]	45.7/57.1/75.3	45.5/56.6/73.9
Resolution limits [Å]	45.51–1.00 (1.09–1.00)	45.00–1.00 (1.09–1.00)
Completeness [%]	96.1 (90.3)	99.6 (98.6)
$I/\sigma(I)$	13.8 (2.0)	11.2 (2.1)
Multiplicity	6.1 (5.4)	6.1 (5.4)
Refinement		
R/R_{free} [%]	13.6/17.0	14.7/18.2
Rmsd bond length [Å]	0.028	0.030
Rmsd bond angles [°]	2.4	2.5

[a] Number in parenthesis are values for the highest of ten resolution shells.

log$D_{7.4}$ Determinations[24]

This protocol was taken from reference [24]. Distribution coefficients are determined using the CAMDIS© (CArrier Mediated DIstribution System, EP2005102211A) method, which is derived from the conventional 'shake flask' method. CAMDIS© is carried out in 96-well

microtiterplates in combination with DIFI©-tubes (Weidmann Plastics Technology AG, Rapperswil, Switzerland), which provide a hydrophobic layer for the 1-octanol phase. The hydrophobic layer (0.45 μm PVDF membranes) fixed on the bottom of each DIFI©-tube is coated (Microfluidic Dispenser BioRAPTR, Bechman Coulter) with 1.0 μL of 1-octanol. Next, the filter membranes are dipped into a 96-well plate, which has been prefilled with 150 μL of aqueous buffer solution (25 mM Phosphate, pH 7.4) containing the compound of interest at a starting concentration of 100 μM. The plate is sealed and shaken for 24 h at room temperature (23 °C) to ensure that the partition equilibrium is reached. The next day, the DIFI©-tubes are removed from the 96-well plate and an aliquote of the aqueous solution is analyzed by LC/MS. The distribution coefficient is calculated from a control experiment without 1-octanol and the remaining compound concentration in the aqueous phase, which was in equilibrium with 1-octanol. Sample preparation is carried out using a TECAN robotic system (RSP 100, 8 channels).

Lyophilisation Solubility Assay (LYSA)[25]

This procedure was taken from reference [25]. Samples were prepared in duplicate from 10 mM Me_2SO stock solutions. After evaporation (1 h) of Me_2SO with a centrifugal vacuum evaporator (Genevac Technologies), the compounds were dissolved in 0.05 M phosphate buffer (pH 6.5), stirred for one hour, and shaken two hours. After 18 h, the solutions were filtered using a microtiter filter plate (Millipore MSDV N65) and the filtrate and its 1/10 dilution were analyzed by direct UV measurement or by HPLC-UV. In addition, a four-points calibration curve is prepared from the 10 mM stock solutions and used for the solubility determination of the compounds. The results are expressed in $\mu g\ mL^{-1}$.

Acknowledgements

This research was supported by F. Hoffmann-La Roche Ltd., Basel, and the ETH Research Council. We thank Ulrike Nowe, Sabine Maehrlein, and Nicole Heindl from the Johannes-Gutenberg Universität Mainz for support with the biological assays, and Mattia Valentini for the preparation of two final compounds. Martine Stihle is acknowledged for support with the co-crystallization experiments. We are grateful to Dr. Bruno Bernet for correcting the Experimental Section in the Supporting Information. J. I. acknowledges the doctoral program W901-B05 DK Molecular Enzymology, funded by the Austrian Science Fund (FWF).

[1] a) S. Cotesta, M. Stahl, *J. Mol. Model.* **2006**, *12*, 436–444; b) M. Harder, B. Kuhn, F. Diederich, *ChemMedChem* **2013**, *8*, 397–404; c) Y. Wang, J. Wang, L. Yao, *J. Phys. Chem. A* **2015**, *119*, 3471–3478.

[2] a) W. H. James, III, C. W. Müller, E. G. Buchanan, M. G. D. Nix, L. Guo, L. Roskop, M. S. Gordon, L. V. Slipchenko, S. H. Gellman, T. S. Zwier, *J. Am. Chem. Soc.* **2009**, *131*, 14243–14245; b) W. H. James, III, E. G. Buchanan, L. Guo, S. H. Gellman, T. S. Zwier, *J. Phys. Chem. A* **2011**, *115*, 11960–11970.

[3] a) G. Duan, V. H. Smith, Jr., D. F. Weaver, *J. Phys. Chem. A* **2000**, *104*, 4521–4532; b) L. Bendová, P. Jurečka, P. Hobza, J. Vondrášek, *J. Phys. Chem. B* **2007**, *111*, 9975–9979; c) Y. N. Imai, Y. Inoue, I. Nakanishi, K. Kitaura, *J. Comput. Chem.* **2009**, *30*, 2267–2276.

[4] C. He, J. Chen, L. An, Y. Wang, Z. Shu, L. Yao, *J. Chem. Inf. Model.* **2015**, *55*, 104–113.

[5] L. M. Salonen, D. D. Medina, E. Carbó-Argibay, M. G. Goesten, L. Mafra, N. Guldris, J. M. Rotter, D. G. Stroppa, C. Rodriguez-Abreu, *Chem Commun.* **2016**, *52*, 7986–7989.

[6] a) A. M. Hartman, M. Mondal, N. Radeva, G. Klebe, A. K. H. Hirsch, *Int. J. Mol. Sci.* **2015**, *16*, 19184–19194; b) B. S. Lauber, L. A. Hardegger, A. K. Asraful, B. A. Lund, O. Dumele, M. Harder, B. Kuhn, R. A. Engh, F. Diederich, *Chem. Eur. J.* **2016**, *22*, 211–221; c) L. Fang, X. Wang, M. Xi, T. Liu, D. Yin, *Mol. BioSyst.* **2016**, *12*, 1174–1182; d) Z. Qiu, B. Kuhn, J. Aebi, X. Lin, H. Ding, Z. Zhou, Z. Xu, D. Xu, L. Han, C. Liu, H. Qiu, Y. Zhang, W. Haap, C. Riemer, M. Stahl, N. Qin, H. C. Shen, G. Tang, *ACS Med. Chem. Lett.* **2016**, *7*, 802–806; e) M. Volgraf, B. D. Sellers, Y. Jiang, G. Wu, C. Q. Ly, E. Villemure, R. M. Pastor, P.-w. Yuen, A. Lu, X. Luo, M. Liu, S.

Zhang, L. Sun, Y. Fu, P. J. Lupardus, H. J. A. Wallweber, B. M. Liederer, G. Deshmukh, E. Plise, S. Tay, P. Reynen, J. Herrington, A. Gustafson, Y. Liu, A. Dirksen, M. G. A. Dietz, Y. Liu, T.-M. Wang, J. E. Hanson, D. Hackos, K. Scearce-Levie, J. B. Schwarz, *J. Med. Chem.* **2016**, *59*, 2760–2779; f) M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2016**, *55*, 9422–9426; *Angew. Chem.* **2016**, *128*, 9569–9574.

[7] O. Korb, B. Kuhn, J. Hert, N. Taylor, J. Cole, C. Groom, M. Stahl, *J. Med. Chem.* **2016**, *59*, 4257–4266.

[8] a) R. K. Castellano, V. Gramlich, F. Diederich, *Chem. Eur. J.* **2002**, *8*, 118–129; b) R. Faraoni, M. Blanzat, S. Kubicek, C. Braun, W. B. Schweizer, V. Gramlich, F. Diederich, *Org. Biomol. Chem.* **2004**, *2*, 1962–1964; c) R. Faraoni, R. K. Castellano, V. Gramlich, F. Diederich, *Chem. Commun.* **2004**, *4*, 370–371.

[9] a) J. Rebek, Jr., B. Askew, P. Ballester, C. Buhr, S. Jones, D. Nemeth, K. Williams, *J. Am. Chem. Soc.* **1987**, *109*, 5033–5035; b) B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones, K. Parris, K. Williams, J. Rebek, Jr., *J. Am. Chem. Soc.* **1989**, *111*, 1082–1090; c) K. Williams, B. Askew, P. Ballester, C. Buhr, K. S. Jeong, S. Jones, J. Rebek, Jr., *J. Am. Chem. Soc.* **1989**, *111*, 1090–1094; d) J. Rebek, Jr., *Acc. Chem. Res.* **1990**, *23*, 399–404; e) J. Rebek, Jr., *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 245–255; *Angew. Chem.* **1990**, *102*, 261–272.

[10] M. Harder, M. A. Carnero Corrales, N. Trapp, B. Kuhn, F. Diederich, *Chem. Eur. J.* **2015**, *21*, 8455–8463.

[11] M. Giroud, M. Harder, B. Kuhn, W. Haap, N. Trapp, W. B. Schweizer, T. Schirmeister, F. Diederich, *ChemMedChem* **2016**, *11*, 1042–1047.

[12] a) D. R. Sudhan, M. B. Rabaglino, C. E. Wood, D. W. Siemann, *Clin. Exp. Metastasis*, **2016**, *33*, 461–473; b) D. R. Sudhan, C. Pampo, L. Rice, D. W. Siemann, *Int. J. Cancer* **2016**, *138*, 2665–2677.

[13] a) L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, D. W. Banner, W. Haap, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 314–318; *Angew. Chem.* **2011***, 123,* 329–334; b) L. A. Hardegger, B. Kuhn, B. Spinnler, L. Anselm, R. Ecabert, M. Stihle, B. Gsell, R. Thoma, J. Diez, J. Benz, J.-M. Plancher, G. Hartmann, Y. Isshiki, K. Morikami, N. Shimma, W. Haap, D. W. Banner, F. Diederich, *ChemMedChem* **2011**, *6*, 2048–2054.

[14] V. Ehmke, E. Winkler, D. W. Banner, W. Haap, W. B. Schweizer, M. Rottmann, M. Kaiser, C. Freymond, T. Schirmeister, F. Diederich, *ChemMedChem* **2013**, *8*, 967–975.

[15] a) V. Ehmke, C. Heindl, M. Rottmann, C. Freymond, W. B. Schweizer, R. Brun, A. Stich, T. Schirmeister, F. Diederich, *ChemMedChem* **2011**, *6*, 273–278; b) V. Ehmke, J. E. Q. Quinsaat, P. Rivera-Fuentes, C. Heindl, C. Freymond, M. Rottmann, R. Brun, T. Schirmeister, F. Diederich, *Org. Biomol. Chem.* **2012**, *10*, 5764–5768.

[16] a) I. D. Kerr, J. H. Lee, C. J. Farady, R. Marion, M. Rickert, M. Sajid, K. C. Pandey, C. R. Caffrey, J. Legac, E. Hansell, J. H. McKerrow, C. S. Craik, P. J. Rosenthal, L. S. Brinen, *J. Biol. Chem.* **2009**, *284*, 25697–25703; b) I. D. Kerr, P. Wu, R. Marion-Tsukamaki, Z. B. Mackey, L. S. Brinen, *PLoS Negl. Trop. Dis.* **2010**, *4*, e701.

[17] a) R. Ettari, L. Tamborini, I. C. Angelo, N. Micale, A. Pinto, C. De Micheli, P. Conti, *J. Med. Chem.* **2013**, *56*, 5637–5658; b) R. Ettari, A. Pinto, S. Previti, L. Tamborini, I. C. Angelo, V. La Pietra, L. Marinelli, E. Novellino, T. Schirmeister, M. Zappalà, S. Grasso, C. De Micheli, P. Conti, *Bioorg. Med. Chem.* **2015**, *23*, 7053–7060; c) R.

Ettari, S. Previti, S. Cosconati, J. Kesselring, T. Schirmeister, S. Grasso, M. Zappalà, *J. Enzyme Inhib. Med. Chem.* **2016**, *31*, 1184–1191.

[18] a) L. M. Salonen, M. Ellermann, F. Diederich, *Angew. Chem. Int. Ed.* **2011**, *50*, 4808–4842; *Angew. Chem.* **2011**, *123*, 4908–4944; b) L. M. Salonen, M. C. Holland, P. S. J. Kaib, W. Haap, J. Benz, J.-L. Mary, O. Kuster, W. B. Schweizer, D. W. Banner, F. Diederich, *Chem. Eur. J.* **2012**, *18*, 213–222.

[19] For biological recognition studies on chalcogen bonding, see: a) X. Huang, C. C. Cheng, T. O. Fischmann, J. S. Duca, X. Yang, M. Richards, G. W. Shipps, Jr., *ACS Med. Chem. Lett.* **2012**, *3*, 123–128; b) X. Huang, C. C. Cheng, T. O. Fischmann, J. S. Duca, M. Richards, P. K. Tadikonda, P. Adulla Reddy, L. Zhao, M. Arshad Siddiqui, D. Parry, N. Davis, W. Seghezzi, D. Wiswell, G. W. Shipps, Jr., *Bioorg. Med. Chem. Lett.*, **2013,** *23*, 2590–2594; c) A. Lange, M. Günther, F. M. Büttner, M. O. Zimmermann, J. Heidrich, S. Hennig, S. Zahn, C. Schall, A. Sievers-Engler, F. Ansideri, P. Koch, M. Laemmerhofer, T. Stehle, S. A. Laufer, F. M. Boeckler, *J. Am. Chem. Soc.* **2015**, *137*, 14640–14652; d) B. R. Beno, K.-S. Yeung, M. D. Bartberger, L. D. Pennington, N. A. Meanwell, *J. Med. Chem.* **2015**, *58*, 4383–4438; e) R. J. Fick, G. M. Kroner, B. Nepal, R. Magnani, S. Horowitz, R. L. Houtz, S. Scheiner, R. C. Trievel, *ACS Chem. Biol.* **2016**, *11*, 748–754.

[20] For computational and crystallographic studies on chalcogen bonding, see: a) J. S. Murray, P. Lane, P. Politzer, *Int. J. Quantum Chem.* **2008**, *108*, 2770–2781; b) L. Junming, L. Yunxiang, Y. Subin, Z. Weiliang, *Struct. Chem.* **2011**, *22*, 757–763; c) M. E. Brezgunova, J. Lieffrig, E. Aubert, S. Dahaoui, P. Fertey, S. Lebègue, J. G. Ángyán, M. Fourmigué, E. Espinosa, *Cryst. Growth Des.* **2013**, *13*, 3283–3289; d) Y. Fang, A. Y. Li, F. Y. Ma, *J. Mol. Model.* **2015**, *21*, 61; e) M. S. Pavan, A. Kumar Jana, S. Natarajan, T. N. Guru Row, *J. Phys. Chem. B* **2015**, *119*, 11382–11390; f) M.

Pyziak, J. Pyziak, M. Hoffmann, M. Kubicki, *Cryst. Growth Des.* **2015**, *15*, 5223–5232; g) X. Zhang, Z. Gong, J. Li, T. Lu, *J. Chem. Inf. Model.* **2015**, *55*, 2138–2153; h) F. Zhou, R. Liu, P. Li, H. Zhang, *New J. Chem.* **2015**, *39*, 1611–1618; i) M. R. Koebel, A. Cooper, G. Schmadeke, S. Jeon, M. Narayan, S. Sirimulla, *J. Chem. Inf. Model* **2016**, DOI: 10.1021/acs.jcim.6b00236.

[21] P. R. Gerber, K. Müller, *J. Comput.-Aid. Mol. Des.* **1995**, *9*, 251–268.

[22] A. F. Abdel-Magid in *Comprehensive Organic Synthesis*, *Vol. 8.02* (Eds.: P. Knochel, G. A Molander), Elsevier, Amsterdam, **2014**, pp. 85–150.

[23] S. Ludewig, M. Kossner, M. Schiller, K. Baumann, T. Schirmeister, *Curr. Top. Med. Chem.* **2010**, *10*, 368–382.

[24] M. Kansy, F. Senner, K. Gubernator, *J. Med. Chem.* **1998**, *41*, 1007–1010.

[25] A. Joliton, J.-M. Plancher, E. M. Carreira, *Angew. Chem. Int. Ed.* **2016**, *55*, 2113–2117; *Angew. Chem.* **2016**, *6*, 2153–2157.

[26] a) E. Persch, O. Dumele, F. Diederich, *Angew. Chem. Int. Ed.* **2015**, *54*, 3290–3327; *Angew. Chem.* **2015**, *127*, 3341–3382; b) L. C. Gilday, S. W. Robinson, T. A. Barendt, M. J. Langton, B. R. Mullaney, P. D. Beer, *Chem. Rev.* **2015**, *115*, 7118–7195; c) G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati, G. Terraneo, *Chem. Rev.* **2016**, *116*, 2478–2601.

[27] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441–451.

[28] M. Morgenthaler, E. Schweizer, A. Hoffmann-Röder, F. Benini, R. E. Martin, G. Jaeschke, B. Wagner, H. Fischer, S. Bendels, D. Zimmerli, J. Schneider, F. Diederich, M. Kansy, K. Müller, *ChemMedChem* **2007**, *2*, 1100–1115.

[29] R. J. Doerksen, A. J. Thakkar, *J. Chem. Phys. A* **1999**, *103*, 10009–10014.

[30] B. Kuhn, J. E. Fuchs, M. Reutlinger, M. Stahl, N. R. Taylor, *J. Chem. Inf. Model.* **2011**, *51*, 3180–3198.

[31] a) E. Weber, *Top. Curr. Chem.* **1987**, *140*, 1–20; b) J. L. Atwood, *Inclusion Compounds* in Ullmann's Encyclopedia of Industrial Chemistry VCH, Weinheim, **2012**, pp. 1–8; c) R. Bishop, *Austr. J. Chem.* **2012**, *65*, 1361–1370; d) S. A. Bourne, M. R. Caira, *Inclusion Compounds* in Ullmann's Encyclopedia of Industrial Chemistry, VCH, Weinheim, **2014**, pp. 1–20.

[32] R. E. Rosenfield, R. Parthasarathy, J. D. Dunitz, *J. Am. Chem. Soc.* **1977**, *99*, 4860–4862.

[33] B. Kuhn, W. Guba, J. Hert, D. Banner, C. Bissantz, S. Ceccarelli, W. Haap, M. Körner, A. Kuglstatter, C. Lerner, P. Mattei, W. Neidhart, E. Pinard, M. G. Rudolph, T. Schulz-Gasch, T. Woltering, M. Stahl, *J. Med. Chem.* **2016**, *59*, 4087–4102.

[34] Y.-C. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.

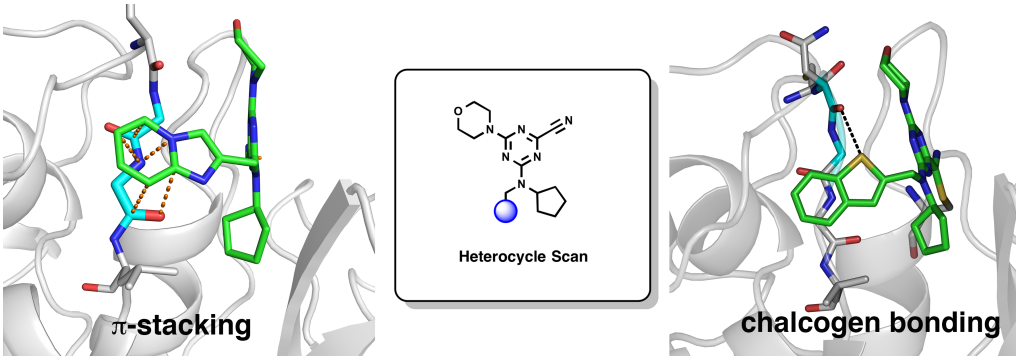
[35] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Revision A.1; Gaussian, Inc.: Wallingford, CT, **2009**.

[36] T. Schirmeister, J. Schmitz, S. Jung, T. Schmenger, L. R. Krauth-Siegel, M. Guetschow, *Bioorg. Med. Chem. Lett.* **2017**, *27*, 45–50.

Text for Table of Contents

Inhibition of the Cysteine Protease Human Cathepsin L by Triazine Nitriles: Amide···Heteroarene π-Stacking Interactions and Chalcogen Bonding in the S3 Pocket

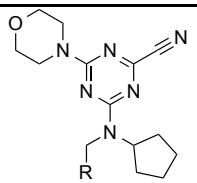
Heteroarene scan: The introduction of heteroarene substituents in the S3 pocket of human Cathepsin L (hCatL) enhances the binding affinity of triazine nitrile inhibitors by comparison to their corresponding aromatic hydrocarbon analogues. The heterocycles undergo π-stacking on peptide amide bonds located at the entrance of the pocket. Inhibitory activities vary over three orders of magnitude in a series of ligands bearing bicyclic heteroarenes as S3 pocket vectors. Intermolecular chalcogen bonding is suggested to be responsible for the enhanced activity of sulfur-containing heterocycles. The insights into biomolecular recognition are supported by co-crystal structures, an interaction network analysis, and conformational analyses.

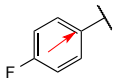
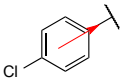
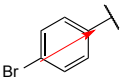
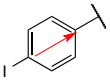
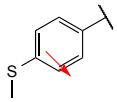
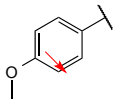
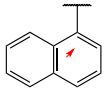
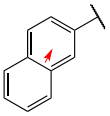


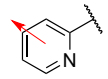
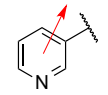
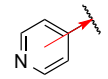
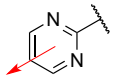
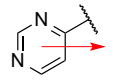
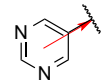
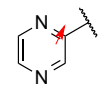
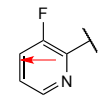
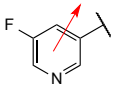
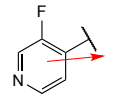
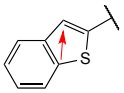
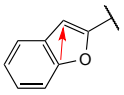
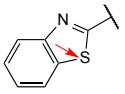
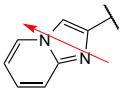
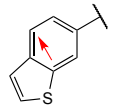
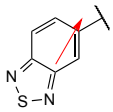
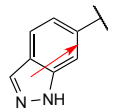
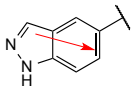
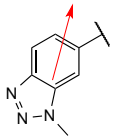
Keywords:

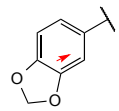
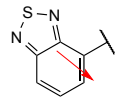
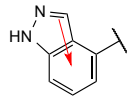
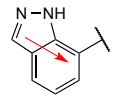
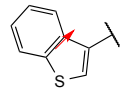
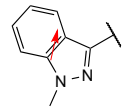
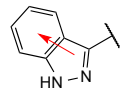
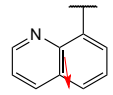
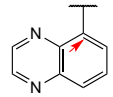
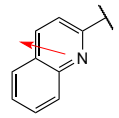
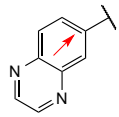
Heteroarene scan – peptide amide bonds – π-stacking – cysteine protease inhibitors — molecular recognition – chalcogen bonding – conformational analysis

Table 1. Predicted and experimental lipophilicity (*c*)log$D_{7.4}$, LYSA solubility, inhibition of hCatL, respective pK_i of triazine nitriles **1–46**, and lipophilic efficiency (LipE). The red arrows represent the dipole moments of the corresponding methyl derivatives, and their size is proportional to the dipole strength (physics convention used, see Section S1.3 in the SI).



Compound	R =	*c*log$D_{7.4}$[a]	log$D_{7.4}$[b]	LYSA[c]	K_i[d] [nM]	pK_i	LipE[e]
1[11]		3.6	n.d.	n.d.	520	6.3	2.4
2[11]		3.8	n.d.	< 0.5	485	6.3	2.2
3		4.2	n.d.	n.d.	549	6.3	1.7
4		4.2	n.d.	< 0.5	77	7.1	2.6
5		4.2	n.d.	< 0.5	177	6.8	2.3
6		4.2	n.d.	< 0.5	13	7.9	3.4
7		3.6	n.d.	< 0.5	18	7.7	3.9
8		3.9	n.d.	n.d.	10 % inh. at 20 μM inh.	n.a.	n.a.
9		3.9	n.d.	n.d.	204	6.7	2.5
10[11]	H	3.1	3.20 ± 0.01	2.6 ± 0.2	910	6.0	2.6
11	Me	3.3	3.39 ± 0.04	25 ± 1	504	6.3	2.7
12		3.5	n.d.	< 0.5	200	6.7	2.9
13		3.4	n.d.	< 0.5	160	6.8	3.1
14		3.8	3.58	< 0.5	210	6.7	2.6
15		3.9	4.01 ± 0.04	< 0.5	170	6.8	2.6
16		3.7	3.78 ± 0.04	7.2 ± 0.1	69	7.2	3.2

17		3.6	3.74 ± 0.01	15 ± 0	420	6.4	2.5
18		3.5	n.d.	n.d.	95	7.0	3.2
19		3.5	n.d.	n.d.	165	6.8	3.0
20		2.6	n.d.	n.d.	439	6.4	3.5
21		2.8	n.d.	n.d.	267	6.6	3.5
22		2.9	2.91 ± 0.03	234 ± 14	273	6.6	3.4
23		2.9	3.06 ± 0.02	171 ± 15	146	6.8	3.64
24		3.6	3.61 ± 0.02	5 ± 0.1	388	6.4	2.5
25		3.8	3.79 ± 0.00	5.5 ± 0.1	270	6.6	2.5
26		3.4	3.58 ± 0.03	7.1 ± 0	276	6.6	2.9
27		4.0	n.d.	n.d.	4	8.4	4.1
28		3.8	n.d.	< 0.5	11	8.0	3.9
29		3.8	n.d.	< 0.5	17	7.8	3.7
30		3.7	3.90 ± 0.01	12 ± 0	35	7.5	3.5
31		4.0	n.d.	n.d.	130	6.9	2.6
32		3.5	n.d.	n.d.	66	7.2	3.4
33		3.5	3.44	<0.5	266	6.6	2.8
34		3.5	n.d.	n.d.	72	7.1	3.4
35		3.7	n.d.	n.d.	460	6.3	2.3

36		3.5	n.d.	n.d.	41	7.4	3.6
37		3.5	n.d.	< 0.5	183	6.7	2.9
38		3.6	n.d.	1.4	195	6.7	2.8
39		3.6	n.d.	0.6	216	6.7	2.8
40		4.2	n.d.	n.d.	35% inh. at 20 μM inh.	n.a.	n.a.
41		3.6	n.d.	< 0.5	120	6.9	3.0
42		4.2	4.46	< 0.5	550	6.3	1.8
43		3.6	n.d.	n.d.	45% inh. at 20 μM inh.	n.a.	n.a.
44		3.4	n.d.	< 0.5	334	6.5	2.8
45		3.9	n.d.	< 0.5	1450	5.8	1.6
46		3.8	3.91 ± 0.06	0.5 ± 0.1	56	7.3	3.2

[a] *c*log$D_{7.4}$ calculated with an in-house developed machine-learning tool based on a tree-based ensemble method. [b] log$D_{7.4}$ measured in a CAMDIS© assay. [c] Solubility measured in a LYSA assay. [d] The reported K_i values are the average of at least two independent measurements, each performed in duplicate. [e] LipE = pIC_{50} – *c*log*D*.