Supporting Information

Azanitrile inhibitors of the SmCB1 protease target are lethal to *Schistosoma mansoni*: structural and mechanistic insights into chemotype reactivity

Adéla Jílková[1‡], Martin Horn[1‡], Jindřich Fanfrlík[1], Jim Küppers[2], Petr Pachl[1], Pavlína Řezáčová[1], Martin Lepšík[1], Pavla Fajtová[1], Petra Rubešová[1], Marta Chanová[3], Conor R. Caffrey[4], Michael Gütschow[2*], Michael Mareš[1*]

[1]Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo n. 2, 16610 Prague 6, Czech Republic

[2]Pharmaceutical Institute, Pharmaceutical & Medicinal Chemistry, University of Bonn, An der Immenburg 4, 53121 Bonn, Germany

[3]Institute of Immunology and Microbiology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Studničkova 2028/7, 12800 Prague 2, Czech Republic

[4]Center for Discovery and Innovation in Parasitic Diseases (CDIPD), Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

[‡]These authors contributed equally to this work.

*Corresponding authors: mares@uochb.cas.cz, guetschow@uni-bonn.de

Table of Contents

Supplementary methods

Electrostatic potential surfaces calculations

The molecular electrostatic potential surfaces of the inhibitor molecules were computed in the implicit SMD solvent environment on the 0.001 a.u. molecular surfaces at the HF/cc-pVDZ level using the Gaussian09 and Molekel4.3 programs.[1-3]

Supplementary results

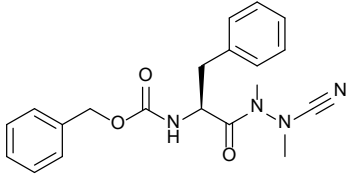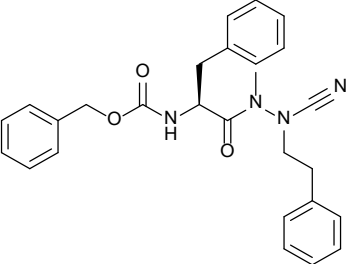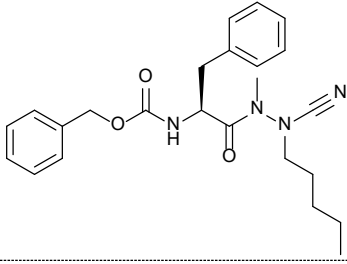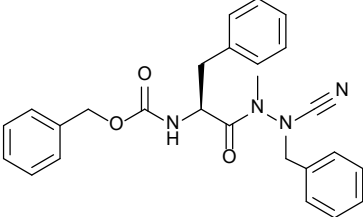### Computed energy differences in inhibitor *E*- and *Z*-configurations
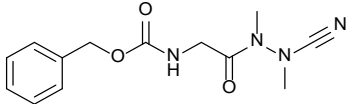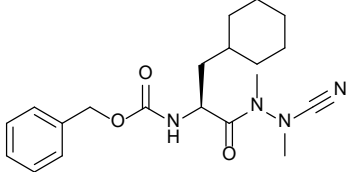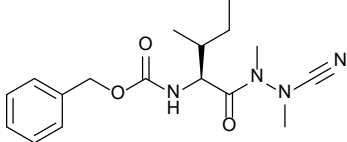
Quantum mechanical 'free' energy calculations of **3a** and **3c** inhibitors in implicit solvent demonstrated that the former is more stable in the *E*-configuration by 1.9 kcal mol^{-1}, while the latter is more stable in the Z-configuration by 2.0 kcal mol^{-1} (Figure S3). The computational analysis of SmCB1 covalent complexes revealed that bound **3a** and **3c** are both more stable in the *Z*-configuration by 39.7 and 3.1 kcal mol^{-1}, respectively. The modeling of the SmCB1 noncovalent complex of **3a** showed that the *Z*-configuration is more stable than the *E*-configuration by 3.5 kcal mol^{-1}; moreover, the S to C_{AB} distance was much larger for the *E*-configuration (5.8 Å as compared to 3.2 Å for *Z*-configuration), which would render the covalent bond formation unfeasible.

Molecular electrostatic potential surfaces of azanitrile and carbanitrile inhibitors

The molecular electrostatic potential (ESP) surfaces were computed for **3a** and **3c** (Figure S4). This identified three distinct regions where the inhibitors differ in their ESP pattern. First, the terminal N_{AA} atom of the nitrile moiety of **3a** has a more negative ESP surface than **3c**, which might result in stronger hydrogen bonds to the backbone amide of Cys100 and the side chain amide of Gln94 (Figure 4D) as well as higher affinity to proton acquired during covalent bond formation (Figure 5C). Second, the H atom bonded to the N_{AE} of **3c** has a more positive ESP surface than the methyl substituent in **3a**. It enables **3c** to form a hydrogen bond with the Gly269 backbone oxygen in the SmCB1-**3c** model, although the contribution to inhibitor binding is rather weak due to the longer distance between the oxygen of Gly269 and N_{AE} (3.5 Å). Third, **3c** has a more negative ESP surface of the O_{AW} atom, which however did not form any close contact with the enzyme. This analysis suggests that the azanitrile warhead forms a stronger noncovalent interaction with the protease active site than the carbanitrile warhead, whereas the interactions of the remaining parts of both inhibitors are of similar strength.

Supplementary tables
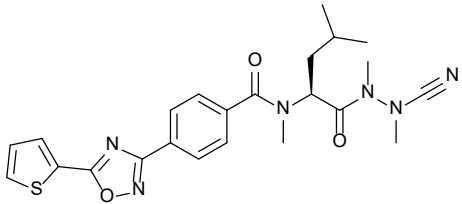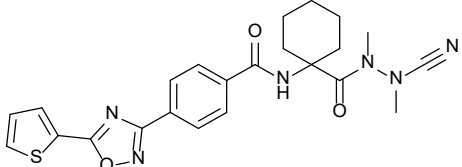
Table S1. Inhibition of SmCB1 by azadipeptide nitriles.

Compound[a]	Inhibitor structure	SmCB1 inhibition[b]	
		K_i (nM)	k_{2nd} (10^3 M^{-1} s^{-1})
P1 substitution			
2a		19.4 ± 1.4	15.7 ± 1.1
8a		4.8 ± 0.2	5.2 ± 0.4
9a		6.2 ± 0.5	2.1 ± 0.3
10a		48.9 ± 4.7	1.5 ± 0.1
P2 substitution			
1a		31 100 ± 1500	n.d.
11a		4.6 ± 0.5	32.7 ± 2.8
12a		71.7 ± 4.4	5.1 ± 0.9

Compound	Structure		
13a		80.6 ± 7.1	12.7 ± 1.9
14a		88.3 ± 10.4	1.5 ± 0.8
15a		714 ± 58	n.d.

See above for the values for compound **2a** (**Gü1303**).

	P3 substitution		
3a		6.1 ± 0.7	20.4 ± 1.9
4a		36.7 ± 2.6	13.9 ± 2.9
5a		4.8 ± 0.3	29.0 ±1.4
16a		30.0 ± 4.1	9.2 ± 0.8
17a		58.3 ± 2.8	3.1 ± 0.6
18a		112 ± 11	13.5 ± 1.1
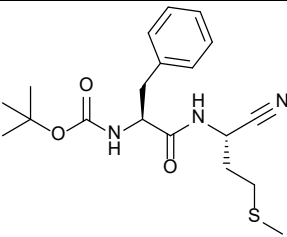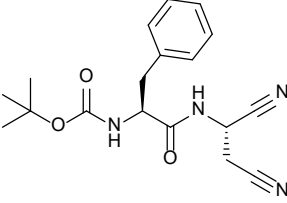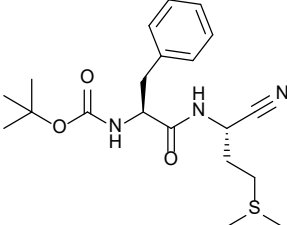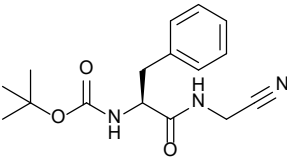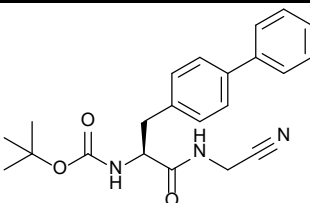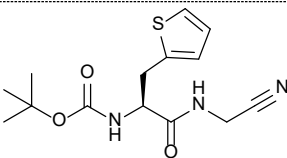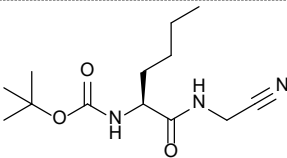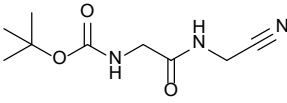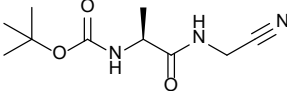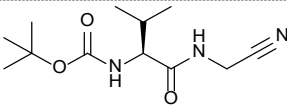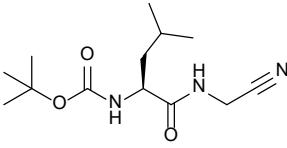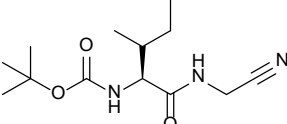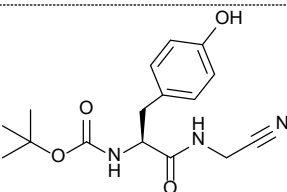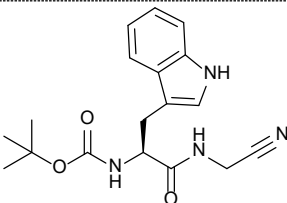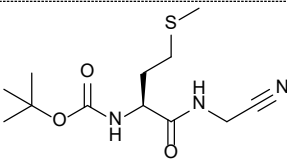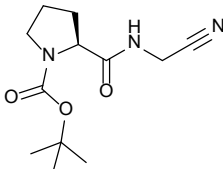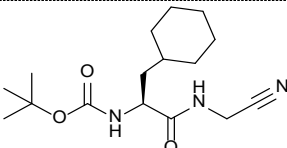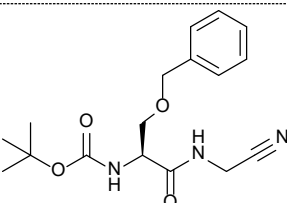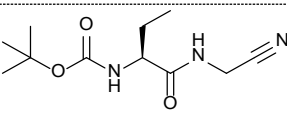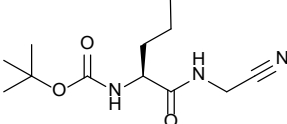
	Miscellaneous		
6a		1540 ± 250	0.5 ±0.1
7a		229 ± 7	1.1 ± 0.2
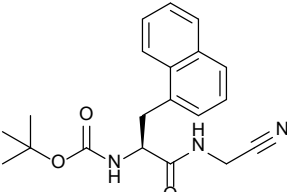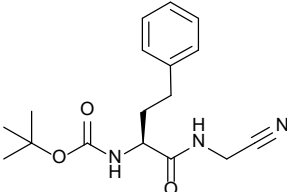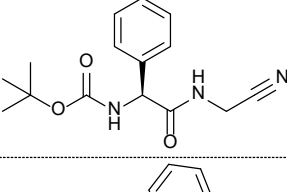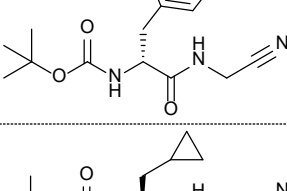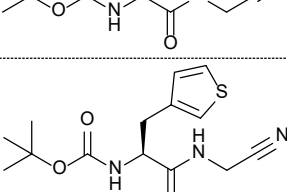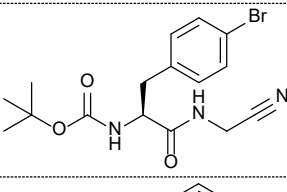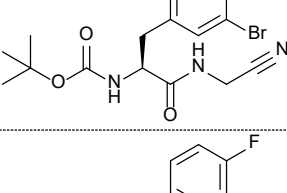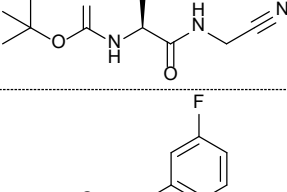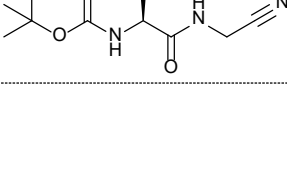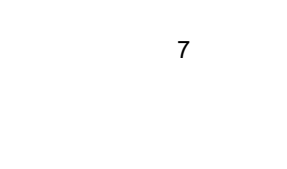
[a]The compounds have been prepared as described.[4-7]
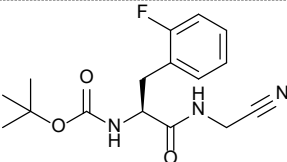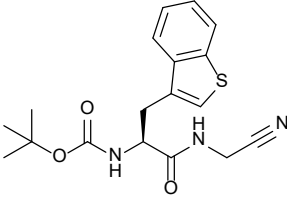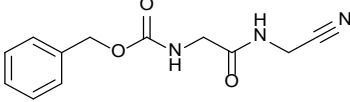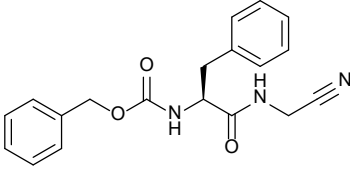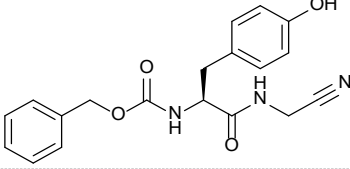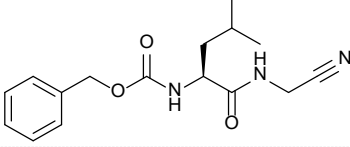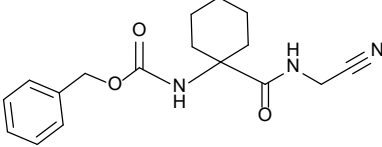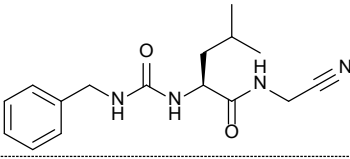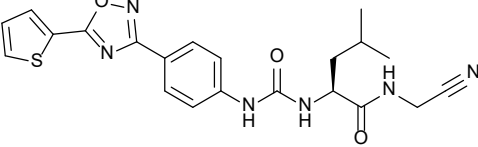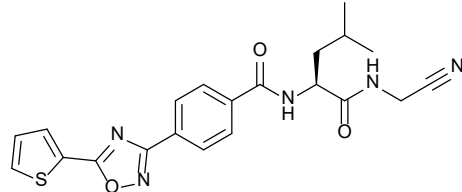
[b]The inhibition parameters were measured in a kinetic activity assay with the fluorogenic peptide substrate Cbz-Phe-Arg-AMC. See Methods for details.

4
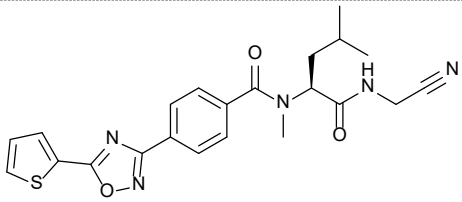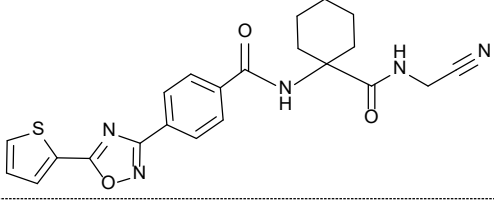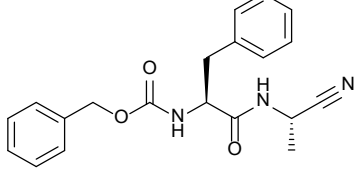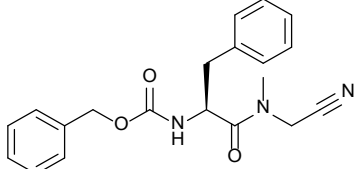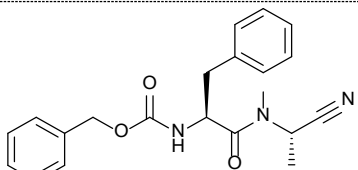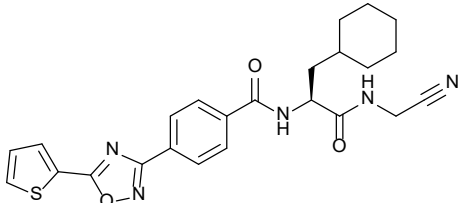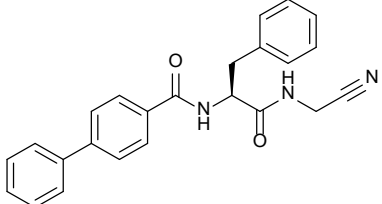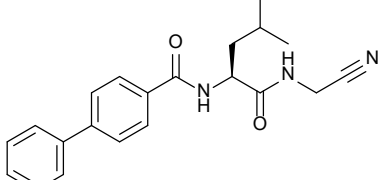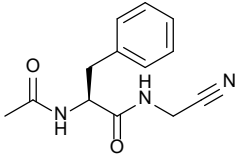
Table S2. Inhibition of SmCB1 by carbadipeptide nitriles.

Compound[a]	Inhibitor structure	SmCB1 inhibition[b] K_i (µM)
P1 substitution		
Gü968		>10
Gü970		>10
Gü1353		>10
Gü809		>10
P2 substitution of Boc-capped compounds		
Gü1124		0.27 ± 0.05
Gü969		0.62 ± 0.05
Gü967		0.87 ± 0.10
Gü807		>10
Gü933		>10

Gü931		>10
Gü914		>10
Gü932		>10
Gü915		>10
Gü916		>10
Gü935		>10
Gü936		>10
Gü929		>10
Gü930		>10
Gü1132		>10
Gü1131		>10

Compound	Structure	Value
Gü1127		>10
Gü1129		>10
Gü1130		>10
Gü1133		>10
Gü1228		>10
Gü1232		>10
Gü1121		>10
Gü1122		>10
Gü1237		>10
Gü1239		>10

Compound	Structure	IC50
Gü1242		>10
Gü1246		>10
P2 substitution of Cbz-capped compounds		
1c		>10
2c		2.92 ± 0.32
Gü2031		0.76 ± 0.22
Gü928		>10
Gü2195		>10
Miscellaneous		
3c		>10
4c		6.62 ± 0.91
5c		1.22 ± 0.20

ID	Structure	Value
6c		>10
7c		>10
8c		>10
9c		>10
10c		>10
Gü2421		0.012 ± 0.004
Gü1136		0.053 ± 0.012
Gü2420		0.056 ± 0.013
Gü1134		>10

| Gü1135 | | >10 |

Table S3. Antischistosomal activity of azanitrile and carbanitrile inhibitors.

Inhibitor	Descriptors[a]				Severity score[b,c]			
	Day 1	Day 2	Day 3	Day 4	Day 1	Day 2	Day 3	Day 4
Azanitriles								
1a					0	0	0	0
2a	Deg	Deg	D	D	4	4	4	4
3a	R, Unc, Dark	Deg	D	D	3	4	4	4
4a	R, Unc, Dark	Deg	D	D	3	4	4	4
5a	Deg	Deg	D	D	2	4	4	4
6a	Deg	Deg	D	D	4	4	4	4
7a	R, Unc, Dark	Deg	D	D	3	4	4	4
8a	Deg	D	D	D	4	4	4	4
9a	Deg	Deg	D	D	4	4	4	4
10a	R, Unc	Deg	D	D	2	4	4	4
11a	R, Unc	Deg	Deg	D	2	4	4	4
12a	Deg	D	D	D	4	4	4	4
13a	Deg	Deg	D	D	4	4	4	4
14a	R, Unc, Dark	Deg	D	D	3	4	4	4
15a	R, Unc	Deg	D	D	2	4	4	4
16a		R, Unc	R, Unc, Dark	Deg	0	2	3	4
17a					0	0	0	0
18a	R, Unc	Deg	D	D	2	4	4	4
Carbanitriles								
1c					0	0	0	0
2c					0	0	0	0
3c					0	0	0	0
4c	R, S	R		R, S	2	1	0	2
5c		R	R, Dark	Deg	0	1	2	4
6c		Dark		R	0	1	0	1
7c					0	0	0	0
8c					0	0	0	0
9c					0	0	0	0
10c					0	0	0	0
Gü967					0	0	0	0
Gü969					0	0	0	0
Gü1124					0	0	0	0
Gü1136					0	0	0	0
Gü2031					0	0	0	0
Gü2420		R, Dark	R, Dark	R, Dark	0	2	2	2
Gü2421	R, S	R, S	R, S, Dark	R, S, Dark	2	2	3	3

[a]The induction of phenotypic changes in newly transformed schistosomula (NTS) of *S. mansoni* by 10 μM inhibitors was recorded every day for 4 days. Phenotypic changes are reported as descriptors, which are: R, rounded; S, slow; Unc, uncoordinated; Deg, degenerated; D, dead.
[b]Each descriptor was awarded a value of 1 (except for Deg and D, which were given the maximum of 4). Values were then added to yield a severity score ranging from 0 (no effect) to 4 (severe), as described previously.[11]
[c]The correlation between the severity scores generated by the azanitriles and carbanitriles (for individual days) and their potency of inhibition of SmCB1 (Tables S1 and S2) was highly significant as evaluated by the Spearman correlation test (with non-zero value coefficients, 20 000 permutations, $p < 0.001$).

Table S4. Cytotoxicity of the most potent azanitrile inhibitors of SmCB1.

| Inhibitor | Cell line viability after 72 h (%)[a] | | | |
| | HepG2 | | HUVEC | |
	1 µM	10 µM	1 µM	10 µM
3a	82	83	106	90
5a	91	101	91	66
8a	102	83	101	44
9a	85	78	104	81
11a	89	109	92	85

[a]Cytotoxicity of the inhibitors towards human cell lines are expressed as % viability vs. untreated cells. Cells were treated with the indicated concentration of the compounds for 72 h and assayed by means of the XTT test (as described previously [12]). Standard deviation values were within 10% of the mean of tetraplicates.

Table S5. X-ray data collection and refinement statistics.

	SmCB1-3a complex[a]
Data collection statistics	
Wavelength (Å)	0.918
Temperature (K)	100
Space group	$P2_12_12_1$
a, b, c (Å)	32.93, 78.97, 89.95
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	44.98–1.29 (1.37–1.29)
Number of unique reflections	58 937 (9351)
Redundancy	3.2 (3.3)
Completeness (%)	98.9 (98.4)
R_{merge}[b] (%)	5.2 (40.4)
Average I/σ (I)	13.53 (2.76)
$CC_{1/2}$ (%)	99.9 (78.5)
Wilson B (Å^2)	16.59
Refinement statistics	
Resolution range (Å)	44.98–1.29 (1.33–1.29)
Number of reflections in working set	57 789 (4144)
Number of reflections in test set	1147 (82)
R value[c] (%)	12.1 (18.9)
R_{free} value[d] (%)	15.7 (24.7)
Number of molecules in AU[e]	1
Number of atoms in AU[e] protein/inhibitor/solvent	2066/24/408
Average ADP[f] for protein/inhibitor/solvent (Å^2)	12.8/29.0/27.7
RMSD bond length (Å)	0.014
RMSD bond angle (°)	1.55
Ramachandran plot statistics[g]	
Favored regions (%)	96.0
Allowed regions (%)	4.0
PDB code	6YI7

[a]The numbers in the parentheses refer to the highest-resolution shell.
[b]$R_{merge}=100\sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is an individual intensity of the ith observation of the reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of the reflection hkl with summation over all data.
[c]R value = $||F_o| - |F_c||/|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.
[d]R_{free} is equivalent to the R value but is calculated for up to 5% of the reflections chosen at random and omitted from the refinement process.[13]
[e]AU, asymmetric unit.
[f]ADP, atomic displacement parameter, formally B-factor.
[g]As determined by Molprobity.[14]

Table S6. Energy calculation of the binding reaction of the inhibitors **3a** and **3c** to SmCB1.

Reaction state[a]	Relative 'free' energy (kcal mol^{-1})[b]	
	Inhibitor	
	3a	**3c**
① Uncomplexed SmCB1 and **3a**(*E*)/**3c**(*Z*)	0.0	0.0
② Transition state 1	15.0	-
③ Uncomplexed SmCB1 and **3a**(*Z*)	1.9	-
④ Noncovalent complex[c]	−28.8	−22.4
⑤ Transition state 2	−17.7	−3.2
⑥ Covalent complex	−38.0	−27.8

[a]The numbering of the reaction states is taken from Figure 5.
[b]The relative 'free' energy (compared to the energy of the separated reactants) of the modeled states upon the binding of **3a** and **3c** to SmCB1.
[c]The relative 'free' energy of the noncovalent complex can be decomposed into individual terms that describe gas-phase interaction energy (ΔE_{int}), interaction desolvation free energy ($\Delta\Delta G_{solv}$) and the change of conformational 'free' energy ($\Delta G'_{conf}$). The ΔE_{int}, $\Delta\Delta G_{solv}$ and $\Delta G'_{conf}$ terms are −89.2, 54.3 and 6.1 kcal mol^{-1} for **3a**, and −78.7, 50.8, and 5.5 kcal mol^{-1} for **3c**, respectively.

Supplementary figures
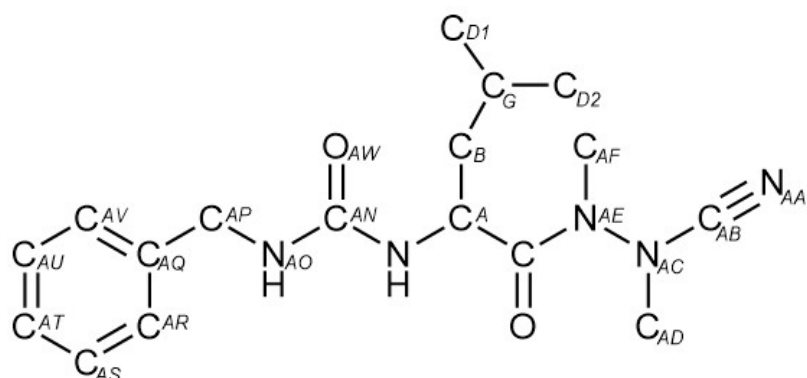
Figure S1. Atom labeling scheme for the inhibitor **3a** (hydrogen atoms on carbon atoms are omitted).

A   **B**

1.0 σ   **0.7 σ**
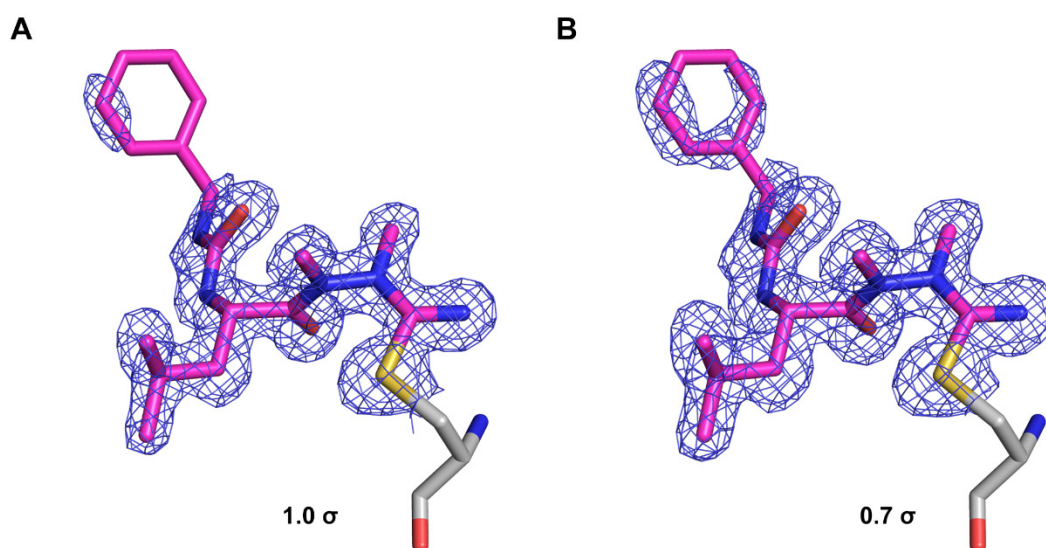
Figure S2. Electron density maps for the inhibitor **3a**. The $2F_o$-F_c electron density map is contoured at 1 σ (A) or 0.7 σ (B), respectively. The inhibitor (magenta) is covalently bound to the catalytic cysteine residue of SmCB1 (gray); heteroatoms have standard color-coding (O, red; N, blue; S, yellow).
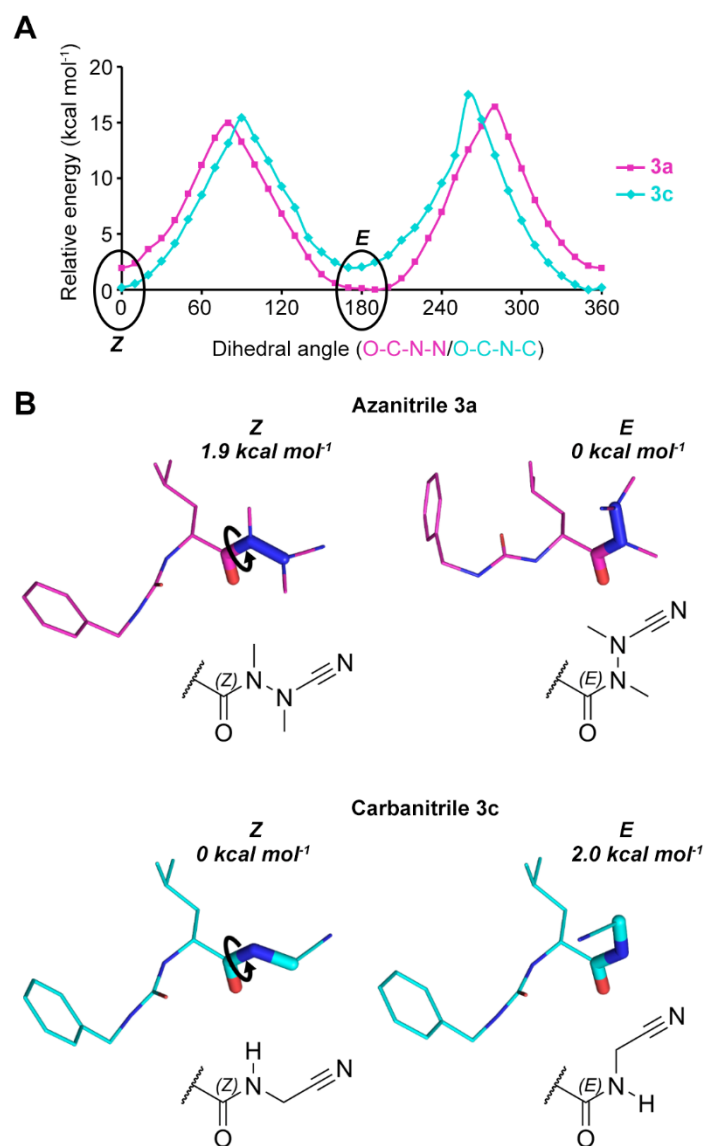
Figure S3. Energy calculation of the conformation change of the inhibitors **3a** and **3c**. (A) The relative 'free' energy of **3a** and **3c** in solution is plotted against the O-C-N-N and O-C-N-C dihedral angles, respectively. The *E*- and *Z*-configurations are marked. (B) A snapshot of the conformation of **3a** (magenta) and **3c** (cyan) in the *E*- and *Z*-configurations from A is shown with the corresponding relative 'free' energy values. The atoms defining dihedral angles are in bold sticks, and the rotating bond is indicated by an arrow. Hydrogen atoms are omitted for clarity.
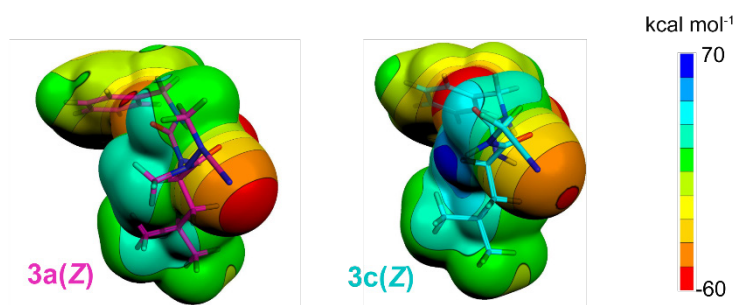
Figure S4. The molecular surface of the electrostatic potential (ESP) of the inhibitors **3a** and **3c**. ESP was computed in solution on the 0.001 a.u. contour of the electron density; the color scale is indicated.

References

1. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. *et al.*, *Gaussian 16 Rev. C.01*, Gaussian, Inc.: Wallingford (USA), 2016.
2. Flűkiger, P.; Lűthi, H. P.; Portmann, S.; Weber, J. *MOLEKEL 4.3*, Swiss Center for Scientific Computing: Manno (Switzerland), 2000-2002.
3. Portmann, S.; Lűthi, H. P., MOLEKEL: An interactive molecular graphics tool. *Chimia* **2000,** *54* (12), 766-770.
4. Löser, R.; Frizler, M.; Schilling, K.; Gütschow, M., Azadipeptide nitriles: highly potent and proteolytically stable inhibitors of papain-like cysteine proteases. *Angew. Chem. Int. Ed.* **2008,** *47* (23), 4331-4334.
5. Frizler, M.; Lohr, F.; Furtmann, N.; Kläs, J.; Gütschow, M., Structural optimization of azadipeptide nitriles strongly increases association rates and allows the development of selective cathepsin inhibitors. *J. Med. Chem.* **2011,** *54* (1), 396-400.
6. Frizler, M.; Lohr, F.; Lülsdorff, M.; Gütschow, M., Facing the gem-dialkyl effect in enzyme inhibitor design: preparation of homocycloleucine-based azadipeptide nitriles. *Chem. Eur. J.* **2011,** *17* (41), 11419-23.
7. Laube, M.; Frizler, M.; Wodtke, R.; Neuber, C.; Belter, B.; Kniess, T.; Bachmann, M.; Gütschow, M.; Pietzsch, J.; Löser, R., Synthesis and preliminary radiopharmacological characterisation of an ^{11}C-labelled azadipeptide nitrile as potential PET tracer for imaging of cysteine cathepsins. *J. Labelled Compd. Radiopharm.* **2019,** *62* (8), 448-459.
8. Löser, R.; Schilling, K.; Dimmig, E.; Gütschow, M., Interaction of papain-like cysteine proteases with dipeptide-derived nitriles. *J. Med. Chem.* **2005,** *48* (24), 7688-707.
9. Löser, R.; Gütschow, M., Dipeptide-derived nitriles containing additional electrophilic sites: potentially irreversible inhibitors of cysteine proteases. *J. Enzyme Inhib. Med. Chem.* **2009,** *24* (6), 1245-52.
10. Schmitz, J.; Furtmann, N.; Ponert, M.; Frizler, M.; Löser, R.; Bartz, U.; Bajorath, J.; Gütschow, M., Active Site Mapping of Human Cathepsin F with Dipeptide Nitrile Inhibitors. *ChemMedChem* **2015,** *10* (8), 1365-77.
11. Long, T.; Neitz, R. J.; Beasley, R.; Kalyanaraman, C.; Suzuki, B. M.; Jacobson, M. P.; Dissous, C.; McKerrow, J. H.; Drewry, D. H.; Zuercher, W. J. *et al.*, Structure-Bioactivity Relationship for Benzimidazole Thiophene Inhibitors of Polo-Like Kinase 1 (PLK1), a Potential Drug Target in *Schistosoma mansoni. PLoS Negl. Trop. Dis.* **2016,** *10* (1), e0004356.
12. Tokarenko, A.; Lišková, B.; Smolen, S.; Táborská, N.; Tichý, M.; Gurská, S.; Perlíková, P.; Frydrych, I.; Tloust'ová, E.; Znojek, P. *et al.*, Synthesis and cytotoxic and antiviral profiling of pyrrolo- and furo-fused 7-deazapurine ribonucleosides. *J. Med. Chem.* **2018,** *61* (20), 9347-9359.
13. Brünger, A. T., Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **1992,** *355* (6359), 472-475.
14. Lovell, S. C.; Davis, I. W.; Arendall, W. B., III; de Bakker, P. I.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C., Structure validation by Cα geometry: φ, ψ and Cβ deviation. *Proteins: Struct. Funct. Bioinform.* **2003,** *50* (3), 437-450.