

Accepted Manuscript

New aziridine-based inhibitors of cathepsin L-like cysteine proteases with selectivity for the *Leishmania* cysteine protease LmCPB2.8

Philipp Fey, Roula Chartomatsidou, Werner Kiefer, Jeremy C. Mottram, Christian Kersten, Tanja Schirmeister



PII: S0223-5234(18)30568-3

DOI: [10.1016/j.ejmech.2018.07.012](https://doi.org/10.1016/j.ejmech.2018.07.012)

Reference: EJMECH 10547

To appear in: *European Journal of Medicinal Chemistry*

Received Date: 7 March 2018

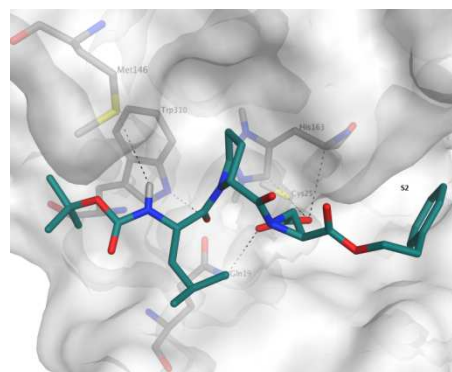
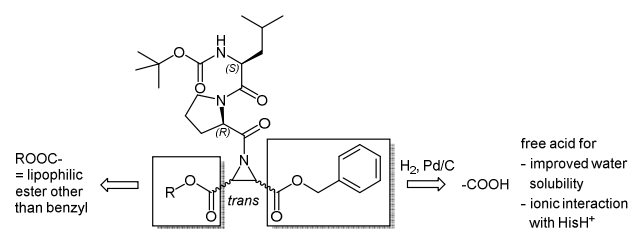
Revised Date: 3 July 2018

Accepted Date: 6 July 2018

Please cite this article as: P. Fey, R. Chartomatsidou, W. Kiefer, J.C. Mottram, C. Kersten, T. Schirmeister, New aziridine-based inhibitors of cathepsin L-like cysteine proteases with selectivity for the *Leishmania* cysteine protease LmCPB2.8, *European Journal of Medicinal Chemistry* (2018), doi: 10.1016/j.ejmech.2018.07.012.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Graphical abstract



New aziridine-based inhibitors of cathepsin L-like cysteine proteases with selectivity for the *Leishmania* cysteine protease LmCPB2.8

Philipp Fey,[§] Roula Chartomatsidou,[§] Werner Kiefer,[§] Jeremy C. Mottram,[#] Christian Kersten,[§] Tanja Schirmeister^{§*}

[§] Institut für Pharmazie und Biochemie, Johannes Gutenberg-Universität Mainz, Staudingerweg 5, 55268 Mainz, Germany

[#] Centre for Immunology and Infection, Department of Biology, University of York, Wentworth Way, Heslington, York, YO10 5DD, UK.

Keywords: cysteine protease, inhibitor, aziridine, rhodesain, cruzain, *Leishmania*

Abbreviations:

DCC (Dicyclohexylcarbodiimide), DMAP (4-(Dimethylamino)pyridine), DPSI ((*S,S*)-Diphenylsulfimine), GB/VI methodology (Generalized Born/Volume Integral methodology), LE (ligand efficiency), LLE (lipophilic ligand efficiency), MMFF94 (Merck Molecular Force Field),

MOE (Molecular Operating Environment), ni (no inhibition), nd (not determined), PDB (Protein Data Bank), PPA (Propylphosphonicanhydride), RMSD (root-mean-square deviation)

Highlights

- Development of new aziridine-based inhibitors of cathepsin L-like cysteine proteases with selectivity for the *Leishmania mexicana* LmCPB2.8
- Analysis of binding sites for LmCPB2.8 and cruzain reveal differences in the S2 pockets
- Predicted ionic interaction with the histidinium ion of the active site of LmCPB2.8 leads to improved affinity and to time-dependent inhibition

Graphical abstract



Abstract

In the present work a series of aziridine-2,3-dicarboxylate inhibitors of papain-like cysteine proteases was designed, synthesized and tested. The compounds displayed selectivity for the parasitic protozoon *Leishmania mexicana* cathepsin L-like cysteine protease LmCPB2.8. The computational methods of homology modelling and molecular docking predicted some

significant differences in the S2 pocket of LmCPB2.8 and cruzain, a related enzyme from *Trypanosoma cruzi*. Due to the presence of Tyr209 in LmCPB2.8 rather than Glu208 in cruzain sterically demanding, lipophilic ester groups (inhibitor **7d**, **9d**, **12d** and **14d**) are predicted to occupy the S2 pocket of the *Leishmania* protease, but do not form favorable interactions in cruzain, which is in common with our experimental results. Further, inhibitor **18** bearing a free carboxylic acid attached to the aziridine moiety showed a time-dependent inhibition of LmCPB2.8 ($K_i = 0.41 \mu\text{M}$; $k_{2\text{nd}} = 190,569 \text{ M}^{-1} \text{ min}^{-1}$). Docking results suggested a strong ionic interaction with the positively charged His163 of the active site. Biological and theoretical data confirm that the novel selective aziridine-based inhibitors are promising candidates for further optimization as LmCPB2.8 inhibitors.

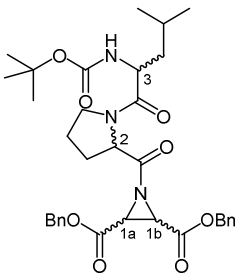
1. Introduction

Cysteine proteases play essential roles in the life cycles of various protozoan organisms making them attractive targets for new antiparasitic drugs. Especially for the neglected diseases HAT (Human African Trypanosomiasis) and the various forms of leishmaniasis, but also for malaria and Chagas disease, new drugs are urgently needed. The parasites causing these diseases express numerous related cysteine proteases that belong to the papain family; according to the MEROPS database system [1,2] they are classified as clan CA, family C1 proteases and share structural similarity to human cathepsins. In detail, these are the cathepsin L-like proteases rhodesain from *Trypanosoma brucei rhodesiense*, falcipains 2 and 3 from *Plasmodium falciparum*, cruzain from *Trypanosoma cruzi*, and various cathepsin L-like proteases from *Leishmania* species, e.g. LmCPB2.8 from *Leishmania mexicana*. There are also some cathepsin B-like proteases expressed by the parasites, e.g. TbCatB from *T. b. rhodesiense* or the *Leishmania* CPCs. Some of

these proteases have already been shown to be suitable targets for new antiparasitic drugs with the vinyl sulfone K11777 which is in preclinical studies against Chagas disease as one of the most promising compounds.[3-5]

In previous studies we identified aziridine-based peptides as inhibitors selective for cathepsin L-like cysteine proteases.[6-9] These consist of a *trans*-configured dibenzyl aziridine-2,3-dicarboxylate fragment Azi(OBn)₂ (i.e. (*S,S*) or (*R,R*) configuration) and the dipeptide Boc-Leu-Pro attached to the aziridine nitrogen. Among the various stereoisomers of the peptide Boc-Leu-Pro-Azi(OBn)₂ some displayed selectivity for parasite proteases over human cathepsin L (Table 1). Some compounds concomitantly showed antileishmanial activity against *L. major* promastigotes and amastigotes,[8-10] namely compounds **I**, **II**, **V** (Table 1). Among these derivatives the most promising is compound **V** which is selective for the leishmanial cathepsin L-like enzyme over the mammalian cathepsins, and which displays good antileishmanial activity.

Table 1: Inhibition of mammalian cathepsins and parasite cathepsin-like proteases by aziridine-2,3-dicarboxylates, and antileishmanial activity against *L. major* promastigotes and amastigotes.[8-10]

								
No.	Configuration at atoms 1a, 1b, 2, 3:	CL <i>K_i</i> [μM]	CB <i>K_i</i> [μM]	LmCPB2.8 <i>K_i</i> [μM]	RD <i>K_i</i> [μM]	CRZ <i>K_i</i> [μM]	FP-2 <i>K_i</i> [μM]	<i>L. m.</i> (p)/(a) EC ₅₀ [μM]
I	<i>S,S,R,S</i>	6.0±0.8	ni	1.7±0.2	1.2±0.2	0.9	43.9±5.2	39.6/2.2±1.5

II	<i>S,S,S,R</i>	4.0±0.2	ni	2.1±0.4	1.1±0.2	0.79	28.3±12.0	35.4/2.7±0.7
III	<i>S,S,S,S</i>	ni	ni	nd	2.3±0.3	nd	ni	ni/nd
IV	<i>S,S,R,R</i>	ni	ni	nd	2.8±0.3	nd	ni	ni/nd
V	<i>R,R,R,S</i>	ni	ni	3.8±0.1	1.7±0.1	ni	ni	37.4/2.3±0.6
VI	<i>R,R,S,R</i>	ni	ni	3.4±0.2	15.2±1.9	30.9±8.2	ni	95/nd
VII	<i>R,R,S,S</i>	0.4±0.2	114±18	nd	0.5±0.1	nd	0.4±0.1	ni/nd

ni = no inhibition at 20 μ M, nd = not determined, CL = cathepsin L, CB = cathepsin B, LmCPB2.8 = *Leishmania mexicana* cathepsin L-like protease, RD = rhodesain from *Trypanosoma brucei rhodesiense*, CRZ = cruzain from *Trypanosoma cruzi*, FP-2 = falcipain-2 from *Plasmodium falciparum*; *L. m.* (p)/(a) = *L. major* promastigotes / amastigotes.

In an ongoing study derivatives with various dipeptides attached to the aziridine nitrogen were synthesized in which the (*S*)-Leu residue of compound **V** was exchanged against another hydrophobic (*S*)-configured amino acid (Ala, Val, Ile, Nva, Nle, Chg, Cha, Phg, Phe, hPhe, Trp).[10] Most of these derivatives showed inhibition of the *Leishmania* enzyme in the same range as compound **V**, but only the inhibitors with Val and Ile also displayed antileishmanial activity against *L. major* promastigotes and amastigotes: (EC₅₀ [μ M] amastigotes/promastigotes: Val: 34.8/2.2±0.6; Ile: 9.8/2.0±0.6; toxicity against J774.4 macrophages > 250 μ M for both).

Compound **V**, in contrast to its Val and Ile analogs, inhibits both the cathepsin L-like CPB2.8 and the *L. major* cathepsin B-like enzyme CPC (K_i = 18.2 μ M).[10] Also in *L. major* lysates, inhibition of both cathepsin L- and cathepsin B-like *Leishmania* proteases had been observed with this compound [10] in contrast to compounds **I** and **II**. Thus, we decided to extend our studies towards *Leishmania* protease inhibitors based on this most promising inhibitor **V**. Previous studies had already shown that dibenzyl esters are superior to diethyl esters in terms of enzyme inhibition and antiparasitic activity.[22] On the other hand, a free acid group at the

aziridine ring has several advantages over the diesters. First, the water solubility is enhanced in comparison to diesters. Second, a free acid improves affinity to cysteine proteases by forming an ionic interaction with the histidinium residue of the active site.[7,11] Furthermore, the free acids show fast covalent inhibition of cysteine proteases since this ionic interaction leads to a positioning of the electrophilic aziridine ring, which is favorable for the nucleophilic ring opening by the active site cysteine residue.[18] The new target compounds thus should consist of the dipeptide sequence Boc-(*S*)-Leu-(*R*)-Pro attached to the aziridine nitrogen, and either of two hydrophobic ester groups or of a free acid and a hydrophobic ester group at the aziridine ring (Figure 1).

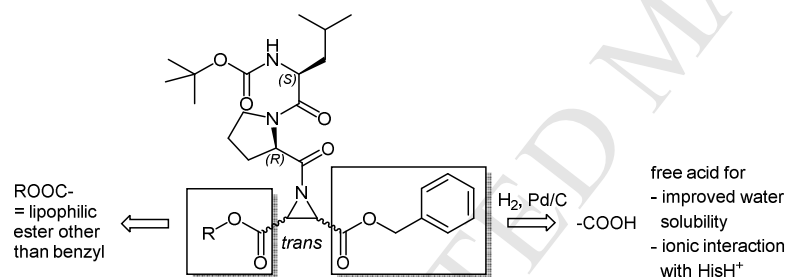


Figure 1. General structure of the target cysteine protease inhibitors.

2. Results and discussion

2.1. Syntheses

The synthesis of such compounds with an ester and also an acid group at the aziridine building block (Figure 1) requires an unsymmetric diester as intermediate. From this diester the free acid

is released in a final synthetic step which, due to the instability of the aziridine and aziridide structure, should not involve strong alkaline or acidic media. In this regard, a benzyl ester is ideal since it can be removed by mild hydrogenolysis (Figure 1). The enantioselective synthesis of an unsymmetric aziridine-2,3-dicarboxylate building block starting from a tartrate needs at least seven steps including protection of the aziridine nitrogen, transesterification of a symmetric aziridine diester by selective hydrolysis and re-esterification, and finally deprotection of the aziridine nitrogen.[12] Thus, we decided to search a new, shorter synthetic route with introduction of the two different esters at an earlier step of the synthesis. Besides the multi-step synthesis starting from a tartrate, aziridine-2,3-dicarboxylates can also be synthesized by a one-step Michael-addition of DPSI to fumarates.[11,13–15] Interestingly, also the *cis*-configured maleates yield the racemic *trans*-configured aziridine-2,3-dicarboxylates due to steric hindrance of the intermediate DPSI adduct T2b (Figure 2). Thus, also the maleates can be used as starting materials.

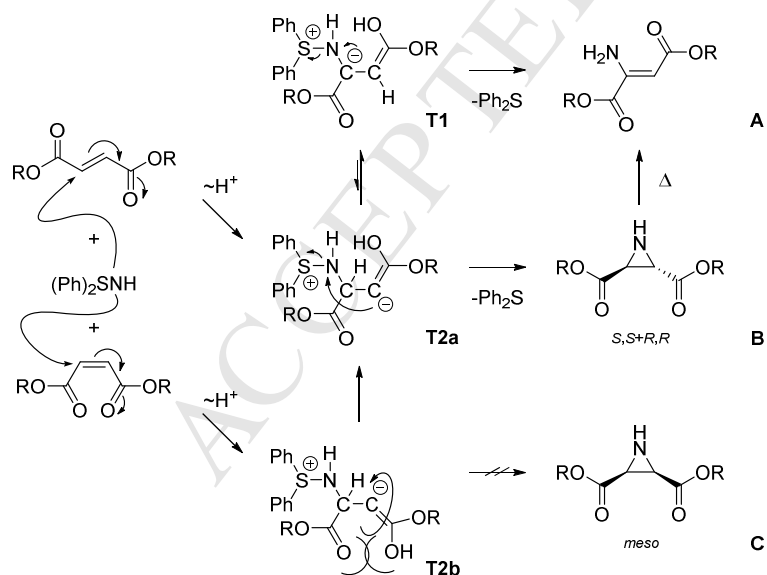


Figure 2. Mechanism of aziridine synthesis by Michael addition of DPSI to either fumarates or maleates. Due to steric hindrance the adduct T2b rotates to the thermodynamically more stable adduct T2a which yields the *trans*-configured aziridine B or –after tautomerization to T1– the enamine A. At higher temperatures (> 90 °C) the aziridine B is converted into enamine A.

The unsymmetric maleates or fumarates were synthesized as follows. Reaction of maleic anhydride with benzyl alcohol yielded the maleic acid benzyl ester (Figure 3). Steglich esterification of this half ester with various alcohols yielded the unsymmetric fumarates and / or maleates (Figure 3). Depending on reaction time, reaction temperature and amount of DMAP as catalyst either the *trans*- or the *cis*-configured unsymmetric diesters could be obtained. With lower amounts of DMAP (0.1 eq), low temperature (0 °C), and shorter reaction times (1 - 2 d), the maleates are the main products, while with higher amounts of DMAP (0.8 eq), room temperature and reaction times of several days the fumarates were formed exclusively due to *Z/E* isomerization catalyzed by DMAP [16,17] (Table 2, Figure 3).

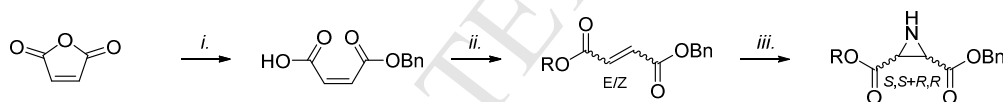


Figure 3. Synthetic route to unsymmetric, racemic *trans*-configured aziridine-2,3-dicarboxylates. *i.* 1.0 eq. BnOH, 1.1 eq. NEt₃, 1 h RT; *ii.* 1.0 eq. ROH, 1.1 eq. DCC, DMAP; *iii.* 1.2 eq. DPSI, 3 mL toluene/mmol DPSI, 24 h 80 °C

Table 2: Synthesized fumarates (a) and maleates (b) starting from maleic acid benzyl ester (I) or fumaric acid (II), reaction conditions leading to either the <i>E</i> - or <i>Z</i> -configured olefin.							
<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>I:</p> </div> <div style="text-align: center;"> <p>II:</p> </div> </div>							
Cpd.	Synthesis	R	<i>E</i> or <i>Z</i>	Eq. DMAP	React. time 0 °C	React. time RT	Yield %

1a	- ^a	Ethyl	<i>E</i>	0.8	1 h	17 h	71
1b	I	Ethyl	<i>Z</i>	0.1	16 h	-	42
2a	I	<i>tert</i> -Butyl	<i>E</i>	0.8	1 h	19 d	54
3a	I	Cyclohexyl methyl	<i>E</i>	0.8	1 h	14 d	55
3b	I	Cyclohexyl methyl	<i>Z</i>	0.1	30 h	-	31
4a	I	Phenyl	<i>E</i>	0.8	1 h	48 h	38
5b	I	Benzyl	<i>Z</i>	0.1	48 h	-	36
6a	I	Naphthalene-2-yl	<i>E</i>	0.8	1 h	72 h	25
7b	I	Naphthalene-2-yl-methyl	<i>Z</i>	0.1	22 h	-	52
8a	I	2-Phenylethyl	<i>E</i>	0.8	1 h	14 d	39
8b	I	2-Phenylethyl	<i>Z</i>	0.1	48 h	-	50
9a	I	3-Phenylpropyl	<i>E</i>	0.8	1 h	14 d	52
10a	I	Furan-2-yl-methyl	<i>E</i>	0.8	1 h	72 h	35
11a^b	I	Pyridine-4-yl-methyl	<i>E</i>	0.1	24 h	-	27
12a	II	Cyclohexyl methyl	<i>E</i>	- ^c	-	24 h reflux	99
5a	II	Benzyl	<i>E</i>	- ^c	-	24 h reflux	99
13a	II	2-Phenylethyl	<i>E</i>	- ^c	-	24 h reflux	96
14a	II	3-Phenylpropyl	<i>E</i>	- ^c	-	24 h reflux	99

^a synthesized from commercially available monoethylfumarate; ^b *E*-product obtained under *Z*-conditions due to catalysis by 4-pyridinemethanol (in analogy to the *Z/E* isomerization catalyzed by DMAP); ^c 2.5 eq. alcohol, *p*-TosOH as catalyst, 24 h reflux in benzene.

Finally the unsymmetric *trans*-configured aziridine-2,3-dicarboxylates which contain one benzyl ester group were synthesized by Michael addition of DPSI to the fumarates or maleates (Figure 3). Only with a few fumarates (with phenyl (**4a**), naphthalene-2-yl (**6a**), furan-2-ylmethyl (**10a**), and pyridine-4-ylmethyl (**11a**) moiety) the aziridine synthesis failed.

Table 3: Synthesized *trans*-configured, racemic aziridine-2,3-dicarboxylate building blocks.

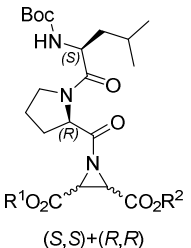
Cpd.	R ¹	R ²	Yield %
1c	Ethyl	Benzyl	14
2c	<i>tert</i> -Butyl	Benzyl	14
3c	Cyclohexyl methyl	Benzyl	56
5c	Benzyl	Benzyl	68
7c	Napthalene-2-yl-methyl	Benzyl	10 ^a
8c	2-Phenylethyl	Benzyl	19
9c	3-Phenylpropyl	Benzyl	17
12c	Cyclohexyl methyl	Cyclohexyl methyl	18
13c	2-Phenylethyl	2-Phenylethyl	18
14c	3-Phenylpropyl	3-Phenylpropyl	9

^a synthesis starting from the corresponding maleate; in all other cases the synthesis started from the fumarate.

The *N*-acylated aziridines with two different ester groups (Table 4, Figure 4, compounds **1d**, **2d**, **3d**, **7d**, **8d**, **9d**) were obtained as diastereomeric mixtures by acylation with the dipeptide Boc-(*S*)-Leu-(*R*)-Pro using PPA as coupling reagent (Table 4, Figure 4). The halfesters were finally

synthesized by hydrogenolysis of the benzyl ester part of the unsymmetric diesters (Table 4, Figure 4, compounds **15-19**). Additionally, several symmetric diesters were also synthesized, starting from fumaric acid, esterification with excess of the respective alcohol, aziridine synthesis with DPSI and finally *N*-acylation with the dipeptide to yield the symmetric *N*-acylated aziridine-2,3-dicarboxylates (Table 2, synthesis II; Table 3, compounds **5c**, **12c**, **13c**, **14c**; Table 4, cpds. **5d**, **12d**, **13d**, **14d**).

Table 4: Synthesized *N*-acylated aziridine-2,3-dicarboxylates.

 (<i>S,S</i>)+(<i>R,R</i>)				
Cpd.	R ¹	R ²	Reaction time at rt [h]	Yield diester/acid ^a [%]
1d	Ethyl	Benzyl	72	15
2d	<i>tert</i> -Butyl	Benzyl	48	26
3d	Cyclohexyl methyl	Benzyl	24	13
5d	Benzyl	Benzyl	24	58
7d	Napthalene-2-yl-methyl	Benzyl	48	23
8d	2-Phenylethyl	Benzyl	48	43
9d	3-Phenylpropyl	Benzyl	48	24
12d	Cyclohexyl methyl	Cyclohexyl methyl	24	22
13d	2-Phenylethyl	2-Phenylethyl	48	31
14d	3-	3-	72	35

	Phenylpropyl	Phenylpropyl		
15	Ethyl	H		85
16	<i>tert</i> -Butyl	H		72
17	Cyclohexyl methyl	H		74
18	2-Phenylethyl	H		95
19	3-Phenylpropyl	H		98

^a step i. / ii. of the synthesis shown in Figure 4.

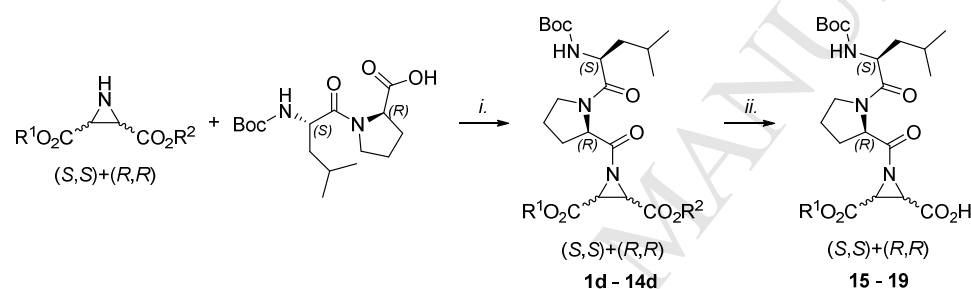


Figure 4. Synthesis of *N*-acylated aziridine-2,3-dicarboxylates. *i.* 1.2 eq. dipeptide, 3.0 eq. PPA, EA, 1 h 0 °C, then several h RT; *ii.* R² = benzyl, 103 mg/mmol Pd/C, MeOH, 30 min, 5 bar H₂.

The previous studies [10] on these compounds (Table 1) showed that the diastereoisomers which only differ in the configuration of the aziridine ring (*R,R* or *S,S*) are similarly active against LmCPB2.8 (**I** and **V**; **II** and **VI**; see Table 1), indicating that the activity is mainly dependent on the configuration of the dipeptide and to a lesser extent on the configuration of the aziridine ring. Thus, we refrained from the rather difficult separation of the diastereoisomers.

2.2. Biological activities

The synthesized *N*-acylated aziridine-2,3-dicarboxylates (Table 4) were tested against the cysteine proteases LmCPB2.8 (Table 5), against mammalian cathepsins L (Table 5) and B (not shown), cruzain, falcipain-2 and rhodesain (Table 5).

Table 5: Inhibition of cysteine proteases by *N*-acylated aziridine-2,3-dicarboxylates and LE and LLE values for inhibition of LmCPB2.8

Cpd.	LmCPB2.8 K_i [μ M]	LE	LLE	Cruzain	CL		Falcipain-2	Rhodesain
1d	ni	-	-	ni	ni		ni	ni
2d	ni	-	-	ni	ni		ni	ni
3d	1.8 \pm 0.4	0.174	-1.62	7.6 \pm 0.7	ni		ni	ni
5d	2.8 ^a	0.169	-0.86	1.8 ^a	12.0 ^a		nd	1.5 ^a
7d	0.8	0.170	-1.50	ni	ni		ni	ni
8d	1.92 \pm 0.06	0.170	-1.09	5.5 \pm 0.2	ni		ni	ni
9d	1.36 \pm 0.05	0.171	-1.32	ni	ni		ni	ni
12d	0.77 \pm 0.05	0.186	-2.19	ni	ni		ni	ni
13d	1.6 \pm 0.1	0.169	-1.40	7.7 \pm 0.7	ni		ni	ni
14d	0.9 \pm 0.1	0.169	-1.90	ni	ni		ni	ni
15	ni	-	-	ni	ni		ni	ni
16	ni	-	-	ni	ni		ni	ni
17^b	1.79 \pm 0.43	0.207	0.35	0.71 \pm 0.015	7.3 \pm 2.4		0.45 \pm 0.015	0.12 \pm 0.002
18^c	0.44 \pm 0.022	0.223	1.53	ni	ni		1.26 \pm 0.22	0.16 \pm 0.003
19^d	ni	-	-	ni	ni		28.6 \pm 3.7	ni

^a see Table 1, cpd. **5d** is a mixture of the two diastereomers **I** and **V** (see Table 1); ni, inhibition at 20 μ M < 50%; ^b

k_{2nd} [$M^{-1} min^{-1}$] for inhibition of cruzain = 322,531, LmCPB2.8 = 53,590, rhodesain = 494,038, falcipain-2 = 76,789,

cathepsin L = 13,939; ^c k_{2nd} [$M^{-1} min^{-1}$] for inhibition of LmCPB2.8 = 190,569, rhodesain = 178,388, falcipain-2 = 11,100; ^d k_{2nd} [$M^{-1} min^{-1}$] for inhibition of falcipain-2 = 3,349.

None of the new *N*-acylated aziridine-2,3-dicarboxylates inhibited mammalian cathepsin B (20 μM final concentration, not shown), only one inhibited mammalian cathepsin L (the acid **17**, K_i = 7.3 μM), only three falcipain-2 (the acids **17-19**, K_i = 0.45, 1.26, 62.0 μM), and only two compounds rhodesain (the acids **17**, **18**, K_i = 0.12, 0.16 μM). Cruzain was inhibited by the diesters **3d**, **8d**, **13d** (K_i = 7.6, 5.5, 7.7 μM) and the acid **17** (K_i = 0.71 μM).

These results show that the acids, especially the acid **17** with cyclohexylmethyl ester moiety, are the less selective compounds. Only these acids showed time-dependent inhibition of the enzymes, which is in agreement with experimental and theoretical studies on oxirane- and aziridine-2,3-dicarboxylates.[18] These studies had shown that an acid at the electrophilic three-membered ring leads to a much faster ring opening reaction since the strong ionic interaction between the inhibitor's carboxylate and the histidinium ion of the active center of the protease pushes the electrophilic ring onto the negatively charged active site's cysteinate residue.[18] For these acids the second-order rate constants of inhibition k_{2nd} were determined (Table 5).

For the inhibition of LmCPB2.8 the ligand efficiencies (LE) and the lipophilic ligand efficiencies (LLE) were calculated using the following equations (Table 5): [19–21]

$$LE = \left(\frac{pK_i}{HA} \right) * 1.37 \quad \text{and} \quad LLE = pK_i - clogP$$

with HA as heavy atom count and clogP as calculated logP values.

The suitability of the calculated clogP values was verified by comparing the calculated values to the retention times obtained in RP-HPLC runs (see supplementary data). The results show a good correlation (Figure 5A) rendering the clogP values suitable for calculation of LLE values.

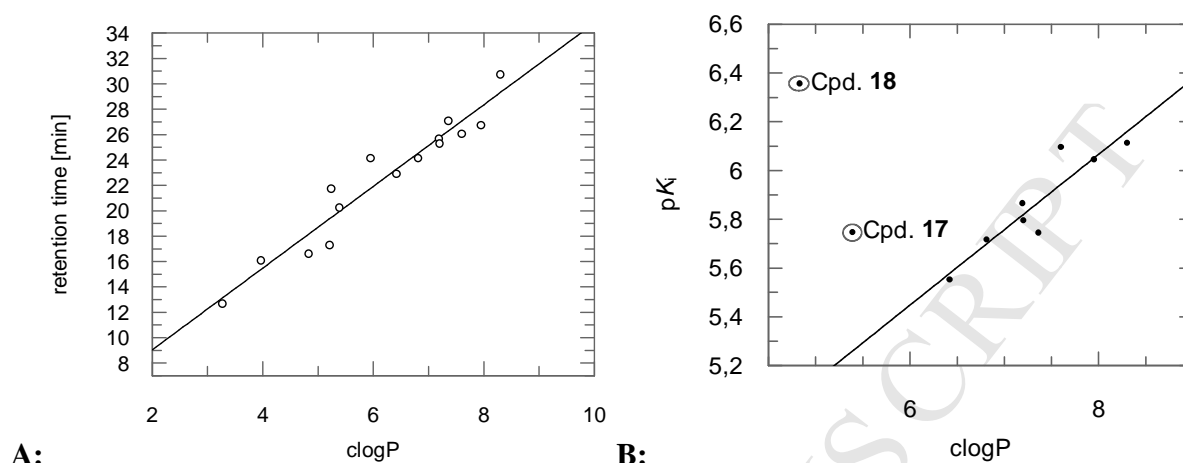


Figure 5. A) Correlation of clogP values and retention times in RP-HPLC runs. B) LLE diagram showing correlation between clogP and pK_i values. The regression line shows that the enhanced affinity of the diesters is correlating to their lipophilicities.

Both parameters, LE and LLE (Table 5), were found to be better for the acids **17** and **18** in comparison to the diesters. For the diesters, the affinities correlate with the lipophilicities of the compounds (Figure 5B). Thus, in terms of suitability as hit-to-lead candidate for further development of LmCPB2.8 inhibitors the acid **18** turned out to be not an optimal but a reasonably good candidate. This compound also displays a quite good selectivity against other cysteine proteases.

2.3. Molecular Docking

The inhibitors can be divided in three groups based on the groups attached to the aziridine ring: symmetric diesters (**5d**, **12d**, **13d**, **14d**), unsymmetric diesters (**1d**, **2d**, **3d**, **7d**, **8d**, **9d**) and acids

(halfesters) (**15** - **19**). For symmetric and unsymmetric diesters selectivity and also affinity for LmCPB2.8 improve, when the aromatic ester chain gets longer. All inhibitors with two large lipophilic ester groups (**3d** – **14d**) display good affinity and selectivity for LmCPB2.8. Only compounds **3d**, **5d** (which is a mixture of inhibitors **I+V**, see Table 1), **8d**, and **13d** are also active against cruzain. These results point to differences in the binding modes between cruzain and LmCPB2.8 on the one hand, but also to differences in the binding modes between diesters and halfesters.

2.4. Comparison of the active sites of LmCPB2.8 and cruzain

In order to explain these observed differences in inhibition of cruzain and LmCPB2.8 we performed docking studies with cruzain (PDB 3I06) [23,24] and LmCPB2.8. Since for the *Leishmania* enzyme no X-ray structure is available we built a homology model based on its protein sequence with the Uniprot identifier P36400 (<http://www.uniprot.org/uniprot/P36400>) (see Material and Methods).

For identification of relevant amino acids in the binding pockets of both enzymes and to explore possible differences, we aligned and superposed the homology model structure of *L. mexicana* CPB2.8 with the X-ray structure of cruzain (PDB 3I06, resolution 1.1 Å) (Figure 6). Although the alignment indicates high similarity between the two proteins with a C α RMSD value of 0.61 Å, some significant differences are found, especially for the S2 subsite. Table 6 shows the key amino acids of each subsite. As expected, the catalytic triad and the subsites S1 and S1' are highly conserved for both parasite cathepsin L-like cysteine proteases.[25,26] Due to the exchange of Glu208 (cruzain) into Tyr209 (LmCPB2.8) in the S2 pockets of the enzymes,

aromatic and lipophilic residues would be more favorable for the leishmanial protease. In agreement with experimental data and the docking studies (see below), the deep hydrophobic S2 pocket of LmCPB2.8 can be occupied by sterically demanding, aromatic residues, such as phenylpropyl or methylnaphthalene ester groups.

Table 6: Key active site and binding pocket amino acids for cruzain and LmCPB2.8 (differences are underlined)

Protease	S2 subsite	S1 subsite	Catalytic triad	S1' subsite
Cruzain	<u>Asn69, Glu208, Ser210</u>	Leu67, Met68, Ala138, Gly163	Cys25, His162, Asn182	Gln19, Trp184
LmCPB2.8	<u>Leu69, Tyr209, Val211</u>	Leu67, Met68, Ala139, Gly164	Cys25, His163, Asn183	Gln19, Trp185

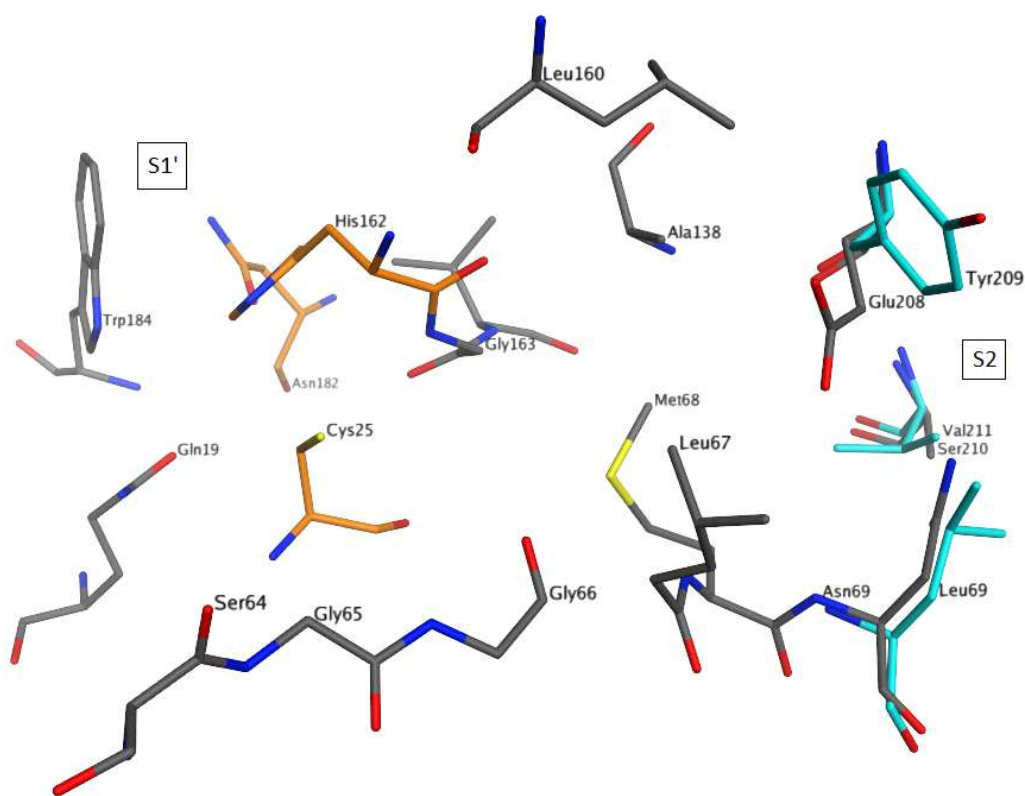


Figure 6. Alignment of the active sites of cruzain (PDB 3I06) and the LmCPB2.8 homology model; amino acids, which differ between two enzymes are highlighted in light blue; catalytic residues are highlighted in orange.

2.5. Docking studies

Docking of the symmetric diesters (-OBn (**5d**) vs. -OEtPh (**13d**) vs. -OPrPh (**14d**)) into the binding site of cruzain and the homology model of the leishmanial enzyme show, that the selective LmCPB2.8 inhibitor **14d** ($K_i = 0.9 \mu\text{M}$) has good FlexX binding scores for LmCPB2.8 (-11,89 kJ/mol for the (*R,R*)- and -10,22 kJ/mol for the (*S,S*)-diastereomer), while for cruzain the binding scores are much higher for both diastereomers (1,26 kJ/mol for the (*R,R*)- and -2,71 kJ/mol for the (*S,S*)-diastereomer). One of the phenylpropyl ester groups is proposed to bind

deeply into the hydrophobic S2 pocket of the leishmanial protease and the second ester group occupies the hydrophobic S1' pocket, whereas in cruzain the ester group does not fit into the S2 subsite and is solvent exposed (see Figure 7). Docking of compound **13d**, with two phenylethyl ester groups, shows similar binding modes for both enzymes, with the ester groups occupying the S2 and S1' pockets. This indicates that the S2 pocket of the leishmanial protease tolerates longer aromatic ester chains than cruzain, which is in agreement with inhibition data. Also the distance between the nucleophilic sulfur atom of the catalytic cysteine residue and the electrophilic aziridine carbon atom is for the leishmanial enzyme in a range, where a covalent bond formation is possible (4.65 Å for (*S,S*)- and 4.92 Å for (*R,R*)-diastereomer).

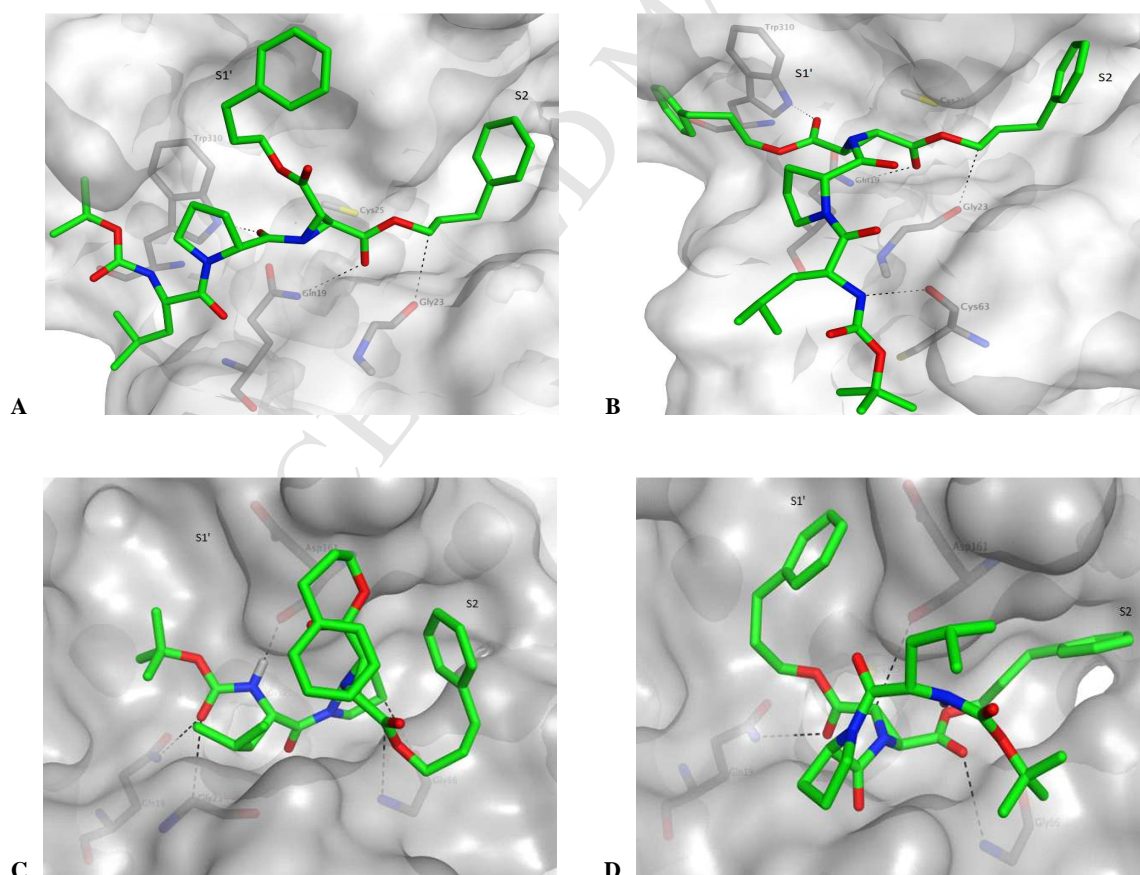


Figure 7. Binding mode of inhibitor **14d**

A Diastereomer with (*S,S*)-configuration at the aziridine ring into the leishmanial protease (FlexX-Score -9,53 kJ/mol, distance to the active site Cys residue: 4.65 Å)

B Diastereomer with (*R,R*)-configuration at the aziridine ring into the leishmanial protease (FlexX-Score -11,83 kJ/mol, distance to the active site Cys residue: 4.92 Å)

C Diastereomer with (*S,S*)-configuration of the aziridine ring) into cruzain (FlexX-Score -2,71 kJ/mol, distance to the active site Cys residue: 5.29 Å)

D Diastereomer with (*R,R*)-configuration at the aziridine ring into cruzain (FlexX-Score 1,26 kJ/mol, distance to the active site Cys residue: 9.54 Å)

For the unsymmetric diesters (-OEtPh (**8d**) vs. -OPrPh (**9d**) vs. -OMeNaph (**7d**)) we obtained similar results. The LmCPB2.8-selective inhibitor **7d** ($K_i = 0.8 \mu\text{M}$) with the sterically most demanding ester group showed the best FlexX scores for the leishmanial protease (-13,65 kJ/mol for the (*R,R*)- and -15,39 kJ/mol for the (*S,S*)-diastereomer). Both diastereomers placed the methylnaphthalene ester group in the hydrophobic S2 pocket (see Figure 8). The configuration of the aziridine ring determines the position of the second ester group. For the (*R,R*)-diastereomer the benzylester is placed in the hydrophobic S1' pocket and for the (*S,S*)-diastereomer the proline residue occupies the S1' pocket. In contrast, no binding mode could be found for cruzain and the (*R,R*)-derivate of **7d**, and for the (*S,S*)-derivate the binding score was much higher for cruzain (-7,99 kJ/mol), which is in accordance to experimental data.

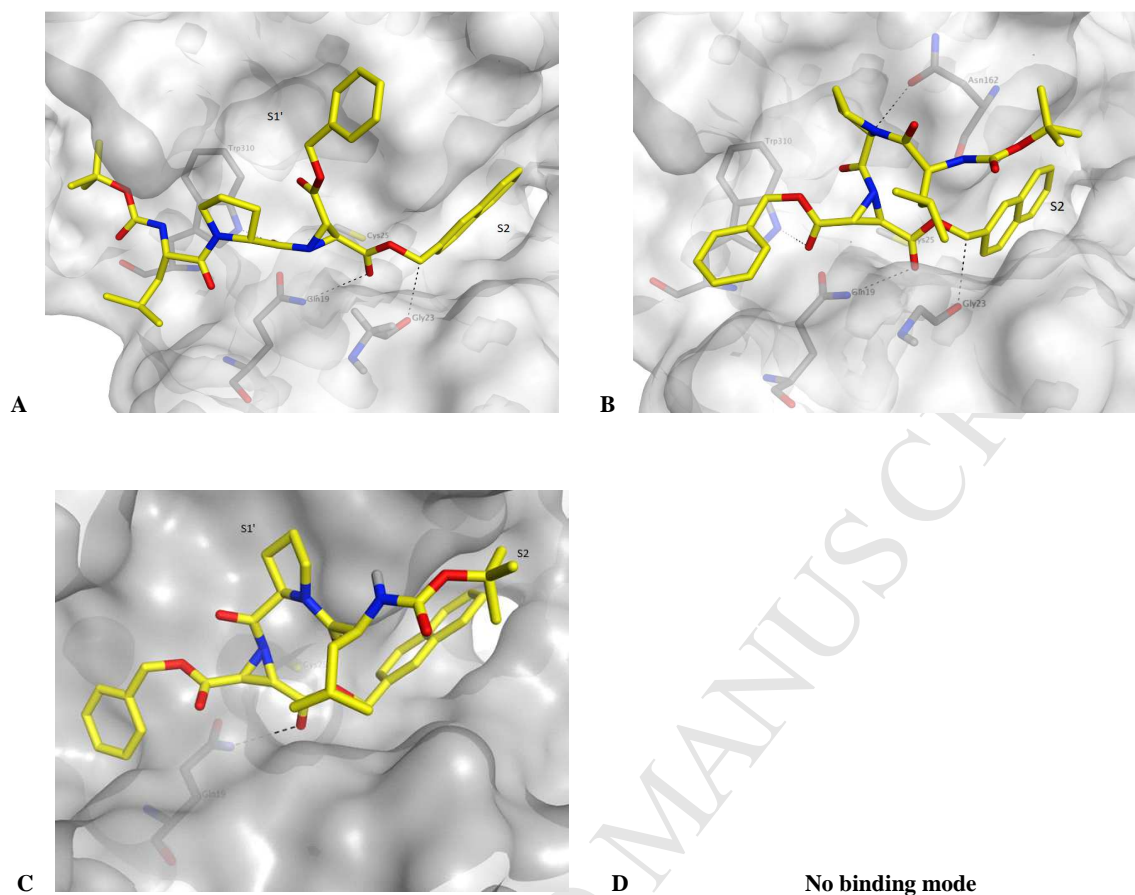


Figure 8. Binding mode of inhibitor **7d**

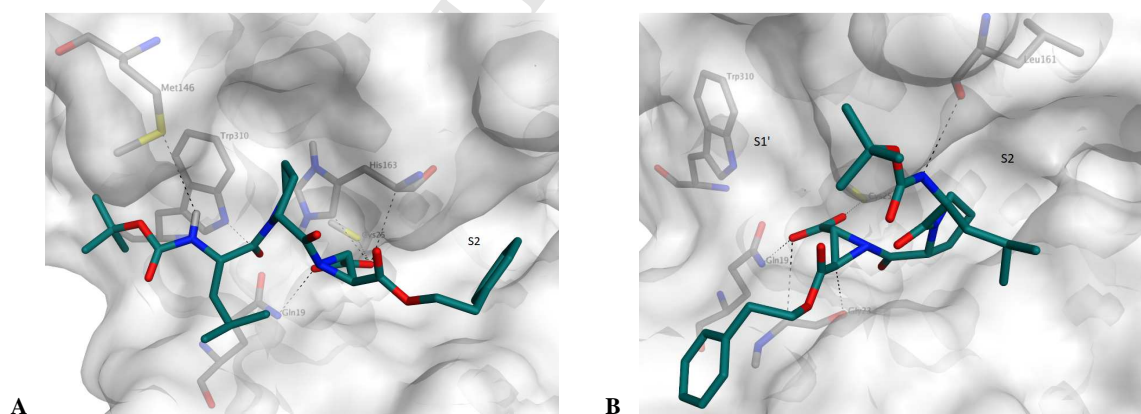
A Diastereomer with (*S,S*)-configuration at the aziridine ring into the leishmanial protease (FlexX-Score -19,92 kJ/mol, distance to the active site Cys residue: 4.90 Å)

B Diastereomer with (*R,R*)-configuration at the aziridine ring into the leishmanial protease (FlexX-Score -14,21 kJ/mol, distance to the active site Cys residue: 4.60 Å)

C Diastereomer with (*S,S*)-configuration of the aziridine ring) into cruzain (FlexX-Score -7,99 kJ/mol, distance to the active site Cys residue: 5.66 Å)

D No binding mode into cruzain was found for the diastereomer with (*R,R*)-configuration of the aziridine ring

For the acids (**15** - **19**), as mentioned above, biological data showed a time-dependent inhibition of the enzymes which means fast irreversible inhibition. In previous studies, this property of aziridine-carboxylic acids had been explained by a strong ionic interaction between the free carboxylate group of the ligand and the positively charged histidinium ion of the active site.[11] This strong ionic interaction may also be responsible for the improved affinity of the acids compared to the corresponding diesters (**17** vs. **3d**, **18** vs. **8d** / **18** vs. **13d**). Docking of inhibitor **18** (see Figure 9) with an affinity of 0.4 μ M for the leishmanial protease showed indeed for both diastereomers an ionic and hydrogen bonding interaction with the catalytic His163. The ester group is placed in the S2 pocket in case of the (*S,S*)-diastereomer, whereas in case of the (*R,R*)-diastereomer the proline moiety occupies this pocket. In contrast, docking into the active site of cruzain indicated that there is no ionic interaction, which is in accordance with the inhibition data.



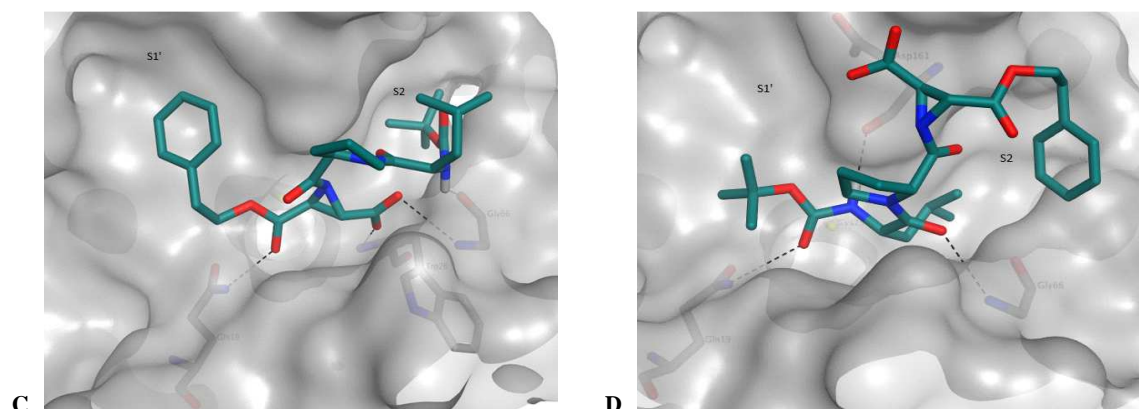


Figure 9. Binding mode of inhibitor **18**

A Diastereomer with (*S,S*)-configuration at the aziridine ring into the leishmanial protease (FlexX-Score -13,60 kJ/mol, distance to the active site Cys residue: 5.8 Å)

B Diastereomer with (*R,R*)-configuration at the aziridine ring into the leishmanial protease (FlexX-Score -11,40 kJ/mol, distance to the active site Cys residue: 4.81 Å)

C Diastereomer with (*S,S*)-configuration of the aziridine ring into cruzain (FlexX-Score -12,94 kJ/mol, distance to the active site Cys residue: 4.6 Å)

D Diastereomer with (*R,R*)-configuration at the aziridine ring into cruzain (FlexX-Score -10,79 kJ/mol, distance to the active site Cys residue: 10.50 Å)

3. Conclusion

Novel and selective aziridine-based inhibitors for the *Leishmania mexicana* protease LmCBP2.8 were discovered. Compounds **7d** ($K_i = 0.8 \mu\text{M}$), **9d** ($K_i = 1.36 \mu\text{M}$), **12d** ($K_i = 0.77 \mu\text{M}$), **14d** ($K_i = 0.9 \mu\text{M}$) and **18** ($K_i = 0.441 \mu\text{M}$) showed the highest potency and selectivity over related parasite and human homologues within the investigated series. For an analysis of the predicted binding modes and binding pockets a homology model for the leishmanial protease was built and compared to the structure of cruzain. This indicated that the S2 pocket of LmCBP2.8 is a deep, lipophilic pocket, favorably being occupied by sterically demanding and lipophilic aromatic

moieties of the ligands, while a tyrosine to glutamate substitution was observed in cruzain. This significant difference between the two enzymes may also be the selectivity-determining feature. Attachment of a carboxylate group at the aziridine ring leads to a time-dependent inhibition and to an improved affinity to LmCPB2.8 (**18** ($K_i = 0.441 \mu\text{M}$) $k_{2\text{nd}} = 190,569 \text{ M}^{-1} \text{ min}^{-1}$) likely to be caused by an ionic interaction with His163 as indicated by the docking studies performed. This strong interaction could be the reason for higher affinity and the fast irreversible inhibition. Furthermore, LE and LLE values were improved, compared to the corresponding diesters.

4. Experimental

4.1. Homology Modelling of LmCPB2.8

A homology model of the *Leishmania mexicana* LmCPB2.8 was generated to obtain the three-dimensional structure of the enzyme in order to perform docking studies and to determine relevant amino acids of the binding pockets. The homology model was built using the Molecular Operating Environment (MOE) 2015.1001 software. The protein sequence of LmCPB2.8 used for modeling has the Uniprot identifier P36400. Based on this target sequence a homology search was performed using MOE Application *PDB-Search*. To find the most suitable template for modeling, the sequence of the first nine results were aligned with the target sequence. PDB-ID 3KKU,[27] solved at 1.28 Å resolution, the crystal structure of cruzain in complex with a ligand, was selected as template showing a sequence similarity of 74.0% and an identity of 59.5% with the target sequence (see supplementary data).

The homology model was built based on the protein sequence P36400 and the template structure 3KKU using the MOE Homology Model application. The calculation was carried out under default conditions with the AMBER10 [28] force field used for energy minimization. Ten independent models were constructed by the Boltzmann-weighted randomized modeling procedure. The final homology model was selected based on the best-scoring intermediate model. The scoring method used was the Generalized Born/Volume Integral methodology [29] (GB/VI methodology). The homology model was subsequently aligned to the template structure, showing a C α RMSD value of 0.54 Å (see supplementary data).

4.2. Molecular Docking

A successful covalent inhibitor of cysteine proteases in its pre-reactive form must first be able to fit non-covalently into the binding site of the enzyme in order to bring the electrophilic center in close proximity to the cysteine sulfur atom. Non-covalent binding interactions, such as hydrogen bonding, hydrophobic, ionic, van der Waals interactions and (de-)solvation processes, are responsible and important for a correct ligand binding. Molecular modeling routinely addresses these non-covalent contributions to ligand binding without taking into consideration the contribution of covalent reaction. In analogy to previous studies,[6] in the present docking study the potential for a possible covalent bond formation was determined for each binding pose by measuring the distance between the nucleophilic sulfur atom of the active site cysteine residue and the electrophilic atoms of the ligand, namely the aziridine ring carbon atoms. Predicted binding modes with a distance less than 5 Å were interpreted as possible start geometry for a covalent bond formation.

The docking program FlexX (LeadIT, version 2.1.6) in the default mode was used to dock the inhibitors flexibly into the active site of cruzain (PDB 3I06) and the leishmanial protease LmCPB2.8 homology model. In the structure of cruzain, solved with a resolution of 1.1 Å, a purine nitrile inhibitor is covalently bound to the enzyme. The binding site was defined as 6.5 Å shell around this reference ligand. The covalent bond and the inhibitor were manually deleted from the structure. Water molecules were removed and hydrogens were added to the protein molecule to simulate physiological conditions at pH 6-7. The catalytic histidine residue was protonated at both imidazole nitrogen atoms. As already successfully employed in previous studies, the active site cysteine residue was mutated to alanine to avoid clashes.[6] After the docking session it is mutated back to the original amino acid for illustration of the pre-covalent

binding mode and measurement of the distance between the cysteine residue and aziridine ring.[6] The aziridine-based inhibitors were each docked as both diastereomers (with (*S,S*)- and (*R,R*)- configured aziridine ring). Energy minimization of the inhibitors prior to docking, using the MMFF94 force field,[30] and visualization were performed with Molecular Operating Environment (MOE 2015.1001).

For validation of the FlexX algorithm for current docking studies, the nitrile inhibitor, extracted from the structure 3I06, was redocked into the enzyme. The highest-scoring docking pose showed good agreement with the native binding mode (RMSD = 0.98 Å, see supplementary data).

4.4. Syntheses

4.4.1. General

Reagents and solvents were obtained from commercial sources (Acros Organics, Alfa Aesar, AllessaChemie, Bachem, Fluka, Iris Biotech, Merck, Roth or Sigma-Aldrich). Solvents were distilled before use, while the other chemicals were used as received. ^1H and ^{13}C NMR spectra were determined on a *Bruker* Fourier 300 or 400 MHz using CDCl_3 or MeOD as solvent. Chemical shifts δ are reported in parts per million (ppm) using residual proton peaks of the solvent as internal standard (CDCl_3 (^1H : 7.26 ppm, ^{13}C : 77.16 ppm), MeOD (^1H : 4.87 ppm, ^{13}C : 49.00 ppm)). The coupling constants (J) are given in Hz, and the splitting patterns are designated as follows: s (singlet); bs (broad singlet); d (doublet); dd (double doublet); t (triplet); m (multiplet). TLC was performed on Merck silica gel (60F254 or RP-C18) precoated plates (0.20

mm). Column chromatography was performed on Merck silica gel 60 (0.040-0.063 mm or 0.063-0.2 mm). Melting points were determined on a SMP3 *Stuart* and are uncorrected. Optical rotation was measured on a P3000 A. *Krüß Optronic* polarimeter. ESI mass spectra were performed on an Ultima 3 Micromass Flash Waters *QTOF* mass spectrometer. HPLC analyses were run on a *Varian* ProStar with an analytical/preparative linear upsacle HPLC system (0.05-50 mL/min at 275 bar back pressure with ScaleUp-mast) and a 2-channel UV detector (detection at 220 nm and 254 nm). For analytical runs a *Phenomenex* Hyperclone 5u 0DS (C18) 120A (250mm x 4.6mm) 5 Micron and for semipreparative runs a *VARIAN DYNAMAX* 250x21.4 MM (LxID) Microsorb 60-8 C18 column were used. The eluent system was millipore V3 water (A) and acetonitrile HPLC grade with 0.1% formic acid (B) (for HPLC method details see supplementary data).

4.4.2. General procedure for the syntheses of fumarates 2a-4a, 6a and 8a-10a of Table 2 (2a as an example)

A solution of maleic acid benzyl ester (20.0 g, 97.0 mmol), *tert*-butanol (7.19 g, 97.0 mmol) and DMAP (9.48 g, 77.6 mmol) in 250 mL dichlormethane was stirred at 0 °C. Subsequently, DCC (22.0 g, 106 mmol) was added. The reaction mixture was stirred 1 h at 0 °C and several hours at room temperature (see Table 2). After filtration, the solution was washed (10% hydrochloric acid solution, saturated NaHCO₃ solution and water), dried over anhydrous Na₂SO₄ and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (petroleum ether/EtOAc) to yield the desired product **2a** as a colorless oil (yield: 54%).

1-Benzyl-4-tert-butyl-(2E)-but-2-endioate (2a, Table 2): colorless liquid (54 %); FT-IR: ν [cm⁻¹] = 2974, 1711, 1638, 1499, 1454, 1368, 1290, 1249, 1135, 972, 841, 751; ¹H NMR (300 MHz,

CDCl₃): δ 7.42-7.32 (m, 5H, Ar-CH), 6.80 (s, 2H, CH=CH), 5.23 (s, 2H, CO₂CH₂Ph), 1.50 (s, 9H, C(CH₃)₃); ¹³C NMR (75 MHz, CDCl₃): δ 165.04, 164.02, 136.07, 135.34, 132.25, 128.63, 128.47, 128.32, 81.96, 69.96, 27.96; ESI-MS m/z calculated for C₁₅H₁₈O₄, 262.12, found 285.12 [M+Na]⁺, 301.10 [M+K]⁺.

4.4.3. General procedure for the syntheses of maleates 1b, 3b, 5b, 7b, 8b and of fumarate 11a of Table 2 (1b as an example)

The synthesis was performed according to the general procedure for synthesis of fumarates using 0.1 eq. DMAP instead of 0.8 eq. DMAP.

1-Benzyl-4-ethyl-(2Z)-but-2-endioate (1b, Table 2): colorless oil (42 %); FT-IR: ν [cm⁻¹] = 2982, 1715, 1634, 1405, 1213, 1151, 1025, 808, 730, 694; ¹H NMR (300 MHz, CDCl₃): δ 7.55-7.30 (m, 5H, Ar-CH), 6.26 (s, 2H, CH=CH), 5.18 (s, 2H, CO₂CH₂Ph), 4.17 (q, 2H, CO₂CH₂CH₃, ³J = 14.3, 7.2 Hz), 1.25 (t, 3H, CO₂CH₂CH₃, ³J = 7.1 Hz); ¹³C NMR (75 MHz, CDCl₃): δ 165.28, 165.07, 135.23, 130.58, 129.18, 128.63, 128.59, 128.52, 67.13, 61.33, 14.01; ESI-MS m/z calculated for C₁₃H₁₄O₄, 234.09, found 257.10 [M+Na]⁺.

4.4.4. General procedure for the syntheses of symmetric fumaric acid diesters 5a, 12a, 13a and 14a of Table 2 (5a as an example)

To a solution of fumaric acid (34.8 g, 300 mmol) and benzyl alcohol (68.1 g, 630 mmol) in 180 mL benzene *p*-TosOH (2.85 g, 15.0 mmol) was added. The reaction mixture was refluxed for 24

h. The residue was diluted with EtOAc, washed with saturated NaHCO₃ solution and water and dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure and the crude product was purified by column chromatography on silica gel (methylene chloride) to yield the desired product **5a** as yellow oil (yield: > 99%).

Dibenzyl-(2E)-but-2-endioate (5a, Table 2): colorless solid (96 %); ¹H NMR (400 MHz, CDCl₃): δ 7.42-7.32 (m, 10H, Ar-CH), 6.93 (s, 2H, CH=CH), 5.23 (s, 2H, 2 x CO₂CH₂Ph).

4.4.5. General procedure for the syntheses of *trans*-configured, racemic aziridine-2,3-dicarboxylates **1c-3c**, **5c**, **7c-9c** and **12c-14c** of Table 3 (**1c** as an example)

1-Benzyl-4-ethyl-(2E)-but-2-endioate (908 mg, 3.88 mmol) and DPSI (1.02 g, 4.65 mmol) in toluene (25mL/g DPSI) were stirred at 83 °C for 24 h. Then, the solvent was removed under reduced pressure and the crude product was purified by column chromatography on silica gel (petroleum ether/EtOAc) to yield the product **1c** as a colorless solid (yield: 14%).

(2S,3S)+(2R,3R)-2-Benzyl-3-ethyl-aziridine-2,3-dicarboxylate (1c, Table 3): colorless solid (14 %); mp 50 °C; ¹H NMR (300 MHz, CDCl₃): δ 7.43-7.30 (m, 5H, Ar-CH), 5.31-5.07 (m, 2H, CO₂CH₂Ph), 4.33-4.10 (m, 2H, CO₂CH₂CH₃), 3.02-2.79 (m, 2H, 2 x Azi-CH), 1.85 (s, 1H, NH), 1.29 (t, 3H, CO₂CH₂CH₃, ³J = 7.1 Hz).

4.4.6. General procedure for the syntheses of *N*-acylated aziridine-2,3-dicarboxylates **1d-14d** of Table 4 (**1d** as an example)

To a solution of BOC-(*S*)-Leu-(*R*)-Pro (177 μ g, 539 μ mol, synthesized according to standard peptide chemistry) and Azi(OBn)(OEt) **1c** (112 mg, 449 μ mol), PPA in 50% EtOAc (857 mg, 1.34 mmol) was added stepwise at 0 °C. The reaction mixture was stirred 1 h at 0 °C and further 24-72 h at room temperature (see Table 4). Then, EtOAc was added and the solution was washed with half-saturated NaHCO₃ solution and water. The combined aqueous phases were extracted with EtOAc. Finally, the organic phases were washed with half saturated NaHCO₃ solution and dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure and the crude product was purified by column chromatography (petroleum ether/EtOAc) and semi-preparative HPLC (for details see supplementary data) to yield the desired product **1d** as colorless oil (35 mg, yield: 15%).

(2*S*,3*S*)+(2*R*,3*R*)-2-Benzyl-3-ethyl-1-[*N*-(*tert*-butoxycarbonyl)-(*S*)-leucyl-(*R*)-prolyl]-aziridine-2,3-dicarboxylate (**1d**, Table 4): colorless oil (15 %); FT-IR: ν [cm⁻¹] = 2958, 1736, 1707, 1646, 1495, 1425, 1364, 1245, 1164, 1025, 735; ¹H NMR (400 MHz, CDCl₃): δ 7.41-7.30 (m, 5H, Ar-CH), 5.26-5.11 (m, 3H, CO₂CH₂Ph, NH), 4.71, 4.62 (je dd, 1H, Pro- α -CH, ³*J* = 7.8, 2.6 Hz), 4.55-4.43 (m, 1H, Leu- α -CH), 4.28-4.15 (m, 2H, CO₂CH₂CH₃), 3.87-3.69 (m, 1H, Pro- δ -CH₂), 3.69-3.43 (m, 3H, Pro- δ' -CH₂, 2 x Azi-CH), 2.33-1.81 (m, 4H, Pro- β -CH₂, Pro- γ -CH₂), 1.80-1.64 (m, 1H, Leu- γ -CH), 1.60-1.35 (m, 11H, Leu- β -CH₂, C(CH₃)₃), 1.28, 1.27 (je t, 3H, CO₂CH₂CH₃, ³*J* = 7.1 Hz), 0.99, 0.97 (je d, 3H, Leu- δ -CH₃, ³*J* = 3.6 Hz), 0.93, 0.91 (je d, 3H, Leu- δ' -CH₃, ³*J* = 1.2 Hz); ¹³C NMR (75 MHz, CDCl₃): δ 179.38, 179.15, 172.21, 171.58, 166.14, 166.09, 166.03, 155.74, 155.53, 135.03, 134.99, 128.76, 128.72, 128.69, 128.64, 79.57, 79.54, 67.93, 67.92, 62.34, 62.32, 60.79, 60.32, 50.46, 50.32, 46.83, 46.74, 42.75, 42.37, 40.53, 40.46, 40.28, 40.17, 29.51, 28.82, 28.48, 28.45, 24.81, 24.68, 24.65, 24.54, 23.59, 23.50, 22.06, 21.88, 14.14; ESI-MS *m/z* calculated for C₂₉H₄₁N₃O₈, 559.29; found 1141.59 [2M+Na]⁺; HPLC

(p): Rt = 23.90 min (method P-IIIa); HPLC (a): Rt = 21.73 min (method A), purity > 99 % (HPLC method details see supplementary data).

4.4.7. General procedure for the hydrogenolysis of *N*-acylated aziridine-2,3-dicarboxylates to yield the acids 15-19 of Table 4 (15 as an example)

To a solution of *N*-acylated aziridine-2,3-dicarboxylate **1d** (7 mg) in 4 mL methanol 103 mg/mmol Pd/C (10 %) were added. The mixture was stirred for 5-30 min inside an autoclave at a pressure of 5 bar H₂. Then the catalyst was filtered over Celite® and the solvent was removed under reduced pressure to yield the desired product **15** as a colorless oil (5 mg, yield: 85%).

(2*S*,3*S*)+(2*R*,3*R*)-1-[*N*-(*tert*-Butoxycarbonyl)-(S)-leucyl-(*R*)-prolyl]-3-ethyloxycarbonyl-aziridine-2-carboxylic acid (**15**, Table 4): colorless oil (85 %); FT-IR: ν [cm⁻¹] = 2953, 1744, 1715, 1634, 1503, 1425, 1258, 1164, 1025, 751, 698; ¹H NMR (300 MHz, MeOD): δ 16.29-14.92 (br s, 1H, CO₂H), 4.75-4.57 (m, 1H, Pro- α -CH), 4.49-4.11 (m, 3H, Leu- α -CH, CO₂CH₂CH₃), 3.88-3.69 (m, 1H, Pro- δ -CH₂), 3.65-3.34 (m, 3H, Pro- δ' -CH₂, 2 x Azi-CH), 2.39-1.84 (m, 4H, Pro- β -CH₂, Pro- γ -CH₂), 1.79-1.56 (m, 1H, Leu- γ -CH), 1.54-1.35 (m, 11H, Leu- β -CH₂, C(CH₃)₃), 1.34-1.17 (m, 3H, CO₂CH₂CH₃), 0.99-0.79 (m, 6H, Leu- δ -CH₃, Leu- δ' -CH₃); ¹³C NMR (75 MHz, MeOD): δ 181.05, 181.01, 180.63, 174.67, 174.25, 173.46, 169.16, 167.89, 167.81, 167.58, 158.08, 157.69, 80.52, 80.50, 80.43, 63.44, 63.22, 62.64, 62.39, 62.03, 51.93, 51.88, 51.80, 48.03, 47.95, 47.74, 42.11, 41.91, 41.57, 41.17, 41.14, 40.94, 40.71, 30.06, 29.58, 28.73, 28.70, 25.83, 25.70, 25.43, 23.87, 23.67, 22.30, 22.02, 21.86, 14.41, 14.38; ESI-MS *m/z* calcul

at-ed for $C_{22}H_{35}N_3O_8$, 469.24; found 492.27 $[M+Na]^+$, 961.58 $[2M+Na]^+$; HPLC (a): $R_t = 12.67$ min (method A), purity 86 % (HPLC method details see supplementary data).

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

Financial support by the DFG (Deutsche Forschungsgemeinschaft) is gratefully acknowledged. JCM is supported by the Medical Research Council (MR/K019384). We thank Sabine Maehrlein, Ulrike Nowe and Nicole Denk for technical support with the enzyme assays.

Appendix A. Supplementary data

The following is the supplementary data related to this article:

Spectroscopic, analytical and purity data for the synthesized compounds, NMR spectra of the compounds tested in the enzyme assays, HPLC methods, description of fluorometric enzyme assays, docking results for cruzain and LmCPB2.8 and homology modelling data.

Supplementary data related to this article can be found at

.....

References

- [1] Rawlings, N. D.; Barrett, A. J. Evolutionary families of peptidases. *Biochem. J.* **1993**, *290*, 205–218.
- [2] Rawlings, N. D.; Barrett, A. J.; Bateman, A. MEROPS: The peptidase database. *Nucleic Acids Res.* **2010**, *38*, D227–33.
- [3] Rosenthal, P. J.; Olson, J. E.; Lee, G. K.; Palmer, J. T.; Klaus, J. L.; Rasnick, D. Antimalarial effects of vinyl sulfone cysteine proteinase inhibitors. *Antimicrob. Agents Chemother.* **1996**, *40*, 1600–1603.
- [4] Kerr, I. D.; Lee, J. H.; Farady, C. J.; Marion, R.; Rickert, M.; Sajid, M.; Pandey, K. C.; Caffrey, C. R.; Legac, J.; Hansell, E. *et al.* Vinyl sulfones as antiparasitic agents and a structural basis for drug design. *J. Biol. Chem.* **2009**, *284*, 25697–25703.
- [5] Olson, J. Antimalarial effects in mice of orally administered peptidyl cysteine protease inhibitors. *Bioorg. Med. Chem.* **1999**, *7*, 633–638.
- [6] Vicik, R.; Busemann, M.; Gelhaus, C.; Stiefl, N.; Scheiber, J.; Schmitz, W.; Schulz, F.; Mladenovic, M.; Engels, B.; Leippe, M. *et al.* Aziridine-based inhibitors of cathepsin L: Synthesis, inhibition activity, and docking studies. *ChemMedChem* **2006**, *1*, 1126–1141.
- [7] Vicik, R.; Busemann, M.; Baumann, K.; Schirmeister, T. Inhibitors of cysteine proteases. *Curr Top Med Chem* **2006**, *6*, 331–353.
- [8] Ponte-Sucre, A.; Vicik, R.; Schultheis, M.; Schirmeister, T.; Moll, H. Aziridine-2,3-dicarboxylates, peptidomimetic cysteine protease inhibitors with antileishmanial activity. *Antimicrob. Agents Chemother.* **2006**, *50*, 2439–2447.
- [9] Schurigt, U.; Schad, C.; Glowa, C.; Baum, U.; Thomale, K.; Schnitzer, J. K.; Schultheis, M.; Schaschke, N.; Schirmeister, T.; Moll, H. Aziridine-2,3-dicarboxylate-based cysteine cathepsin inhibitors induce cell death in *Leishmania major* associated with accumulation of debris in autophagy-related lysosome-like vacuoles. *Antimicrob. Agents Chemother.* **2010**, *54*, 5028–5041.
- [10] Schad, C.; Baum, U.; Frank, B.; Dietzel, U.; Mattern, F.; Gomes, C.; Ponte-Sucre, A.; Moll, H.; Schurigt, U.; Schirmeister, T. Development of a new antileishmanial aziridine-2,3-dicarboxylate-based inhibitor with high selectivity for parasite cysteine proteases. *Antimicrob. Agents Chemother.* **2016**, *60*, 797–805.
- [11] Schirmeister, T.; Peric, M. Aziridinyl peptides as inhibitors of cysteine proteases: Effect of a free carboxylic acid function on inhibition. *Bioorg. Med. Chem.* **2000**, *8*, 1281–1291.
- [12] Breuning, A.; Vicik, R.; Schirmeister, T. An improved synthesis of aziridine-2,3-dicarboxylates via azido alcohols—epimerization studies. *Tetrahedron-Asymmetry* **2003**, *14*, 3301–3312.
- [13] Schirmeister, T. Aziridine-2,3-dicarboxylic acid derivatives as inhibitors of papain. *Arch. Pharm.* **1996**, *329*, 239–244.
- [14] Furukawa, N.; Oae, S. The Michael type addition of free sulfilimine. *Synthesis* **1976**, *1976*, 30–32.
- [15] Furukawa, N.; Ohtsu, M.; Akasaka, T.; OAE, S. One step synthesis of aziridines by the Michael type addition of free sulfimides. *Tetrahedron* **1980**, *36*, 73–80.

- [16] Christensen, M. S.; Pedersen, P. J.; Andresen, T. L.; Madsen, R.; Clausen, M. H. Isomerization of all-(E)-retinoic acid mediated by carbodiimide activation - Synthesis of ATRA ether lipid conjugates. *Eur. J. Org. Chem.* **2010**, 719–724.
- [17] Gaunt, M. J.; Boschetti, C. E.; Yu, J.; Spencer, J. B. Preferential hydrogenolysis of NAP esters provides a new orthogonal protecting group strategy for carboxylic acids. *Tetrahedron Lett* **1999**, 40, 1803–1806.
- [18] Mladenovic, M.; Schirmeister, T.; Thiel, S.; Thiel, W.; Engels, B. The importance of the active site histidine for the activity of epoxide- or aziridine-based inhibitors of cysteine proteases. *ChemMedChem* **2007**, 2, 120–128.
- [19] Leeson, P. D.; Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* **2007**, 6, 881–890.
- [20] Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: A useful metric for lead selection. *Drug Discov. Today* **2004**, 9, 430–431.
- [21] Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* **2014**, 13, 105–121.
- [22] Vicik, R.; Hoerr, V.; Glaser, M.; Schultheis, M.; Hansell, E.; McKerrow, J. H.; Holzgrabe, U.; Caffrey, C. R.; Ponte-Sucre, A.; Moll, H.; Stich, A.; Schirmeister, T. Aziridine-2,3-dicarboxylate inhibitors targeting the major cysteine protease of *Trypanosoma brucei* as lead trypanocidal agents. *Bioorg. Med. Chem. Lett.* **2006**, 16, 2753–2757.
- [23] Berman, H. M.; Bhat, T. N.; Bourne, P. E.; Feng, Z.; Gilliland, G.; Weissig, H.; Westbrook, J. The Protein Data Bank and the challenge of structural genomics. *Nat. Struct. Biol.* **2000**, 7 Suppl, 957–959.
- [24] Mott, B. T.; Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Ang, K. K.-H.; Leister, W.; Shen, M.; Silveira, J. T.; Doyle, P. S.; Arkin, M. R. *et al.* Identification and optimization of inhibitors of trypanosomal cysteine proteases: Cruzain, rhodesain, and TbCatB. *J. Med. Chem.* **2010**, 53, 52–60.
- [25] Alves, L. C.; Melo, R. L.; Cezari, M. H. S.; Sanderson, S. J.; Mottram, J. C.; Coombs, G. H.; Juliano, L.; Juliano, M. A. Analysis of the S2 subsite specificities of the recombinant cysteine proteinases CPB of *Leishmania mexicana*, and cruzain of *Trypanosoma cruzi*, using fluorescent substrates containing non-natural basic amino acids. *Mol. Biochem. Parasit.* **2001**, 117, 137–143.
- [26] Juliano, M. A.; Brooks, D. R.; Selzer, P. M.; Pandolfo, H. L.; Judice, W. A. S.; Juliano, L.; Meldal, M.; Sanderson, S. J.; Mottram, J. C.; Coombs, G. H. Differences in substrate specificities between cysteine protease CPB isoforms of *Leishmania mexicana* are mediated by a few amino acid changes. *Eur. J. Biochem.* **2004**, 271, 3704–3714.
- [27] Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Eidam, O.; Mott, B. T.; Keiser, M. J.; McKerrow, J. H.; Maloney, D. J.; Irwin, J. J.; Shoichet, B. K. Complementarity between a docking and a high-throughput screen in discovering new cruzain inhibitors. *J. Med. Chem.* **2010**, 53, 4891–4905.
- [28] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, 25, 1157–1174.
- [29] Labute, P. The generalized Born/volume integral implicit solvent model: Estimation of the free energy of hydration using London dispersion instead of atomic surface area. *J. Comput. Chem.* **2008**, 29, 1693–1698.
- [30] Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, 17, 490–519.

Highlights

- Development of new aziridine-based inhibitors of cathepsin L-like cysteine proteases with selectivity for the *Leishmania mexicana* LmCPB2.8
- Analysis of binding sites for LmCPB2.8 and cruzain reveal differences in the S2 pockets
- Predicted ionic interaction with the histidinium ion of the active site of LmCPB2.8 leads to improved affinity and to time-dependent inhibition