

Article

## Fragment Screening Hit Draws Attention to a Novel Transient Pocket Adjacent to the Recognition Site of the tRNA-Modifying Enzyme TGT

Engi Hassaan, Christoph Hohn, Frederik R. Ehrmann, Friedrich Wieland Goetzke, Levon Movsisyan, Tobias Hübner-Wulsdorf, Maurice Sebastiani, Adrian Hartsch, Klaus Reuter, François Diederich, and Gerhard Klebe

*J. Med. Chem.*, **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.0c00115 • Publication Date (Web): 09 Jun 2020

Downloaded from [pubs.acs.org](https://pubs.acs.org) on June 19, 2020

### Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.

# Fragment Screening Hit Draws Attention to a Novel Transient Pocket Adjacent to the Recognition Site of the tRNA-Modifying Enzyme TGT

Engi Hassaan,<sup>†,#</sup> Christoph Hohn,<sup>‡,#</sup> Frederik R. Ehrmann,<sup>†</sup> F. Wieland Goetzke,<sup>‡</sup> Levon Movsisyan,<sup>‡</sup> Tobias Hufner-Wulsdorf,<sup>†</sup> Maurice Sebastiani<sup>†</sup>, Adrian Härtsch,<sup>‡</sup> Klaus Reuter<sup>†</sup>, François Diederich,<sup>‡</sup> Gerhard Klebe<sup>†\*</sup>

[†] Institute of Pharmaceutical Chemistry, University of Marburg, Marbacher Weg 6, 35032 Marburg, Germany

[‡] Laboratorium für Organische Chemie, ETH Zurich, Vladimir-Prelog-Weg 3, 8093 Zurich, Switzerland

[#] These authors contributed equally to this work.

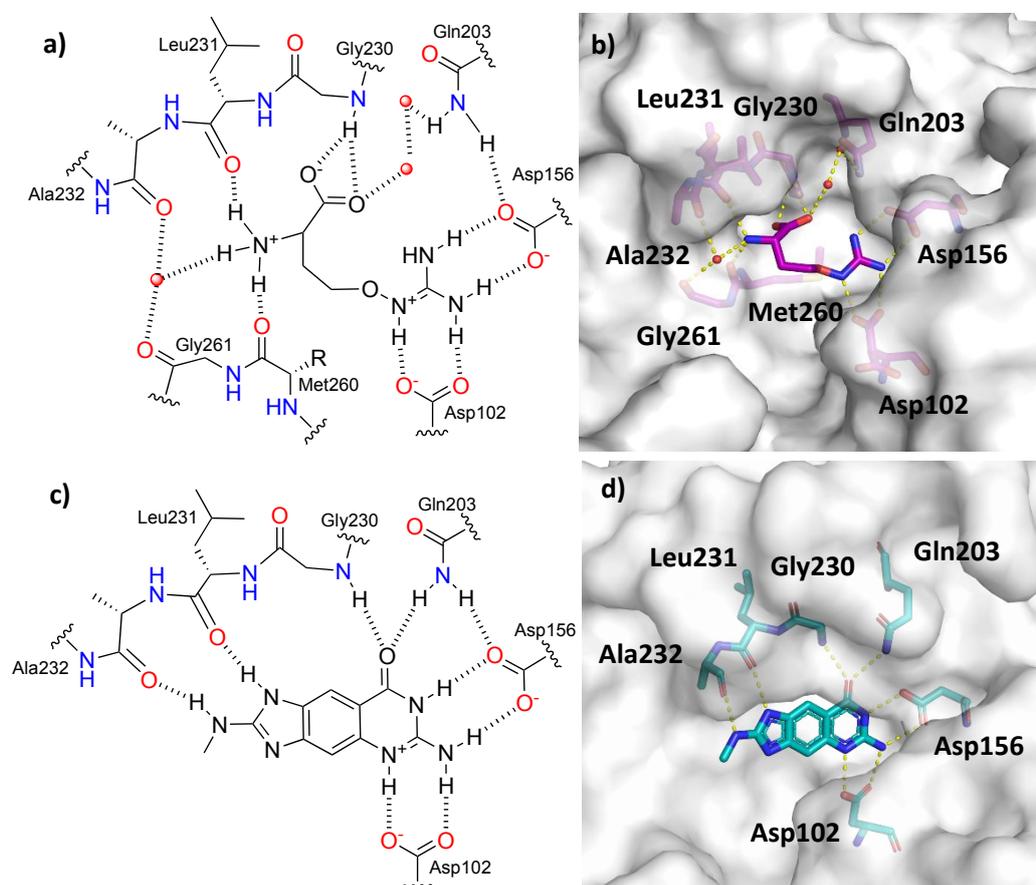
## ABSTRACT

Fragment-based lead discovery was applied to tRNA-guanine transglycosylase, an enzyme modifying post-transcriptionally tRNAs in *Shigella*, the causative agent of Shigellosis. TGT inhibition prevents translation of *Shigella*'s virulence factor, hence reducing pathogenicity. One discovered fragment opens a transient sub-pocket in the preQ<sub>1</sub>-recognition site by pushing back an aspartate residue. This step is associated with reorganization of further amino acids structurally transforming a loop adjacent to the recognition site by duplicating the volume of the preQ<sub>1</sub>-recognition pocket. We synthesized 6-carboxamido-, 6-hydrazido-, and 4-guanidino-benzimidazoles to target the opened pocket, including a dihydro-imidazoquinazoline with a propyn-1-yl exit vector pointing into the transient pocket, and displacing a conserved water network. MD simulations and hydration-site analysis suggest water displacement to contribute favorably to ligand binding. A cysteine residue, exclusively present in bacterial TGTs, serves as gatekeeper of the transient sub-pocket. It becomes accessible upon pocket opening for selective covalent attachment of electrophilic ligands in eubacterial TGTs.

**Keywords:** Transient binding pocket, tRNA-modifying enzyme, Fragment-based lead discovery, Crystallography, Ligand synthesis, Hydration-site analysis

## INTRODUCTION

Fragment-based lead discovery is an increasingly popular approach to generate first hits primarily as promising starting points for drug development projects. In order to generate these starting points, general-purpose fragment libraries have emerged that usually compile a broad range of chemotypes and can therefore probe protein-binding sites efficiently. As opposed to tailored fragment libraries, general-purpose libraries easily allow for the exploration of unprecedented binding motifs or scaffold hopping. They can provide surprising insights into unexpected features of the studied protein.<sup>[1]</sup> Amongst the recent famous success stories of fragment-derived drug candidates progressing in clinic trials is the discovery of the Bcr-Abl Kinase allosteric inhibitor ABL001 by Novartis.<sup>[2]</sup> We recently performed a crystallographic fragment screening on tRNA-guanine transglycosylase (TGT), an enzyme well studied in our laboratory.<sup>[3]</sup> It catalyzes the exchange of the genetically encoded guanine at the wobble position 34 of the anticodon loop of tRNAs specific for Asn, Asp, His and Tyr with the modified base preQ<sub>1</sub> *via* breaking and regenerating the covalent bond between the ribose sugar and the nucleobase.<sup>[4]</sup> To properly recognize the nucleobases, the TGT enzyme features two adjacent aspartates (Asp102, Asp156 in the *Z. mobilis* enzyme, Figure 1a) in its active center. These aspartates along with Asp280 are responsible for carrying out the catalysis of the base-exchange reaction.<sup>[5]</sup> In addition, the carbonyl group of the substrate's pyrimidone ring is hydrogen-bonded to the backbone NH of Gly230 and the terminal carboxamide of Gln203.



1  
2  
3 **Figure 1.** a) Schematic binding mode of the Arg-related fragment as bound to *Zymomonas (Z.) mobilis*  
4 TGT; b) Crystal structure of this fragment in the catalytic center of TGT (carbon atoms magenta, PDB  
5 code: 5UTI, 1.36 Å resolution). The fragment succeeds in opening a transient binding pocket next to  
6 Gly230 and Gln203, which is solvated by water molecules; c) Schematic binding mode of *lin*-  
7 benzoguanine (**1.1**) as bound to *Z. mobilis* TGT; d) Crystal structure of **1.1** (carbon atom cyan, PDB code:  
8 4PUK, 1.49 Å resolution) in the catalytic center of TGT.  
9

10  
11 Our screen identified a small fragment that is chemically related to the amino acid arginine, which  
12 opens a new, transient sub-pocket when bound to the recognition pocket of the target protein (PDB  
13 code: 5UTI).<sup>[3]</sup> *Via* its guanidinium group, the discovered fragment binds to the Asp102/Asp156 motif  
14 forming two bidentate hydrogen-bond interactions and perturbs the spatial arrangement of the  
15 remaining recognition pattern next to Gly230 and Gln203 (Figure 1b). As a result, a small transient  
16 pocket opens, which increases the volume of the recognition site significantly (for further illustration  
17 s. Supporting Information, Figure S49). Two water molecules, stabilized by the nearby carboxylate  
18 group of the fragment, are found in the transient pocket. The amino nitrogen of the fragment also  
19 forms a hydrogen bond with the amide oxygens of Leu231 and Met260. The observed structure  
20 immediately provokes some interesting questions. Is the opening of the transient pocket just an  
21 uncommon and reinforced but functionally useless adaptation of the protein or does it indicate a  
22 functionally essential feature of the enzyme? Furthermore, can a ligand successfully accommodate this  
23 transient pocket and does this support the development of more potent or even more selective  
24 ligands?  
25

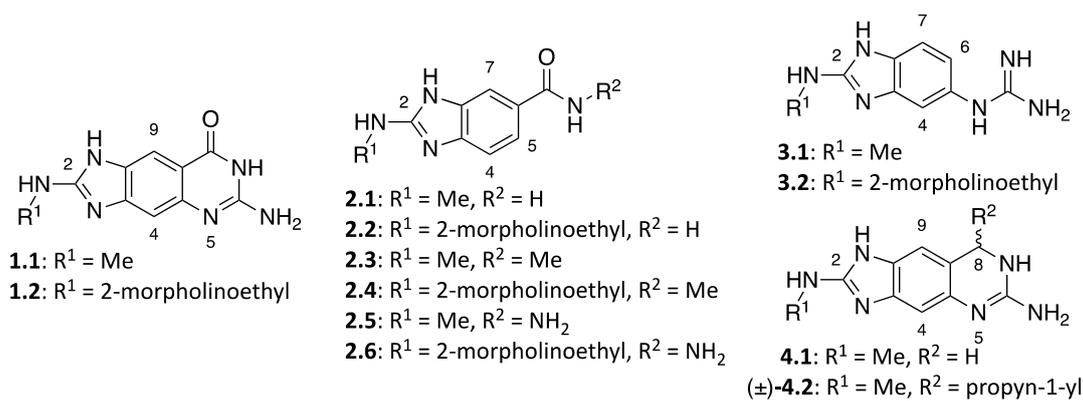
26  
27  
28  
29  
30  
31  
32  
33  
34  
35 The studied enzyme TGT is a potential target to fight Shigellosis.<sup>[6,7]</sup> This infection is a world-wide  
36 endemic with 165 million cases reported per year, including 1.1 million deaths globally, mostly in  
37 children under five years of age.<sup>[8]</sup> It is an acute bacterial infection of the intestine caused by the gram-  
38 negative bacterium *Shigella*. Symptoms of the disease include sudden abdominal cramping, nausea,  
39 fever, vomiting, and bloody stool. Sources of infection include contaminated food or water and  
40 therefore the disease predominantly affects low income countries where poor sanitation and  
41 overcrowding are prevalent.<sup>[8]</sup> Furthermore, this disease can also rapidly spread in countries of higher  
42 life standards.<sup>[9]</sup> The usual treatment of shigellosis is an antibiotics therapy; however, resistance is  
43 rapidly developed. *Shigella* pathogenicity depends on virulence factors, which are required to invade  
44 epithelial cells. The expression of these virulence factors is modulated by the enzyme TGT (EC 2.4.2.29).  
45 This enzyme occurs in all domains of life, however, it differs in substrate specificity and the pre-  
46 modified base to be inserted.<sup>[10]</sup> Eubacterial TGTs catalyze the exchange of the base guanine by preQ1  
47 which is, once incorporated, further transformed to queuine by other enzymes.<sup>[4]</sup> Eukaryotic TGTs,  
48 which show a high sequence identity to eubacterial TGTs (e.g. 43.2% between *Homo sapiens* and  
49 *Shigella flexneri*) directly incorporate queuine at the wobble position of tRNAs. These enzymes use a  
50 subpocket to accommodate the additional substituent of queuine, which is located in a similar region  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 as the transient pocket opened by the above-mentioned fragment.<sup>[5]</sup> Archaeobacterial TGTs on the  
4 other hand, show only about 20–25% sequence identity to eubacterial TGTs and address guanine at a  
5 different site, namely position 15 of the D-arm in the majority of archaeal tRNAs.<sup>[10]</sup>  
6  
7

8  
9 In the present contribution, we study parameters which feature the above-described opening of the  
10 transient recognition pocket and elucidate whether the observed adaptations relate to any functional  
11 properties of the enzyme family that could be exploited for the development of potent and selective  
12 anti-infective drugs. Starting with the well-characterized *lin*-benzoguanine scaffold **1**,<sup>[11]</sup> appropriate  
13 to inhibit eubacterial TGT, we first designed some ligands capable to perturb the recognition pattern  
14 next to the residues Gly230/Gln203 as suggested by our fragment structure. By molecular design and  
15 subsequent chemical synthesis of suitable model compounds, determination of their binding affinity  
16 to the enzyme and crystal structure analyses along with a molecular dynamics-based hydration site  
17 analysis, we succeeded to trace the role and trigger of the opening of the transient pocket and its  
18 potential impact on the development of selective inhibitors of eubacterial TGTs.  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## RESULTS AND DISCUSSION

**Ligand Design.** In the past, the design of our ligands, aimed at inhibiting bacterial TGT, was focused on *lin*-benzoguanine derivatives (Scheme 1, scaffold **1**) as this parent scaffold perfectly occupies the recognition pocket (Schematic binding mode Figure 1c). Without further decoration, it already shows an inhibition of  $K_i = 58 \pm 36$  nM<sup>[11]</sup> and provides multiple options for exit vectors to efficiently address the various sub-pockets surrounding the central recognition site. Thus, sub-nanomolar potency could be achieved.<sup>[12]</sup> Stimulated by the binding geometry of the above-mentioned fragment complex and the surprising opening of a transient pocket, we decided to either open the aminopyrimidone ring of the *lin*-benzoguanine scaffold (scaffolds **2** and **3**) or to reduce the carbonyl group in the pyrimidone ring of the parent scaffold (scaffold **4**). The resulting 7,8-dihydro-imidazoquinazolines in **4** allow the attachment of substituents at C8 that may penetrate into the opened transient pocket (s. Scheme 1). Thereby, the carbonyl function of the pyrimidone ring, interacting with the backbone NH of Gly230 and the carboxamide nitrogen of Gln203, would be lost. As indicated by the fragment complex, some inherent flexibility with respect to the latter residues might be given. In order to start our ring opening exercise conservatively and to study the impact of the abandoned hydrogen-bonding interactions of the core scaffold to Gly230 and Gln203, we decided to also investigate benzimidazole derivatives (scaffolds **2** and **3**) with substituents attached to positions 5 or 6. For this purpose, we considered amides, hydrazides, and guanidines, respectively. With respect to R<sup>1</sup> in Scheme 1, we studied all ligands with an attached *N*-methyl group at position 2. For some selected examples, also the morpholino ethyl derivatives were investigated.

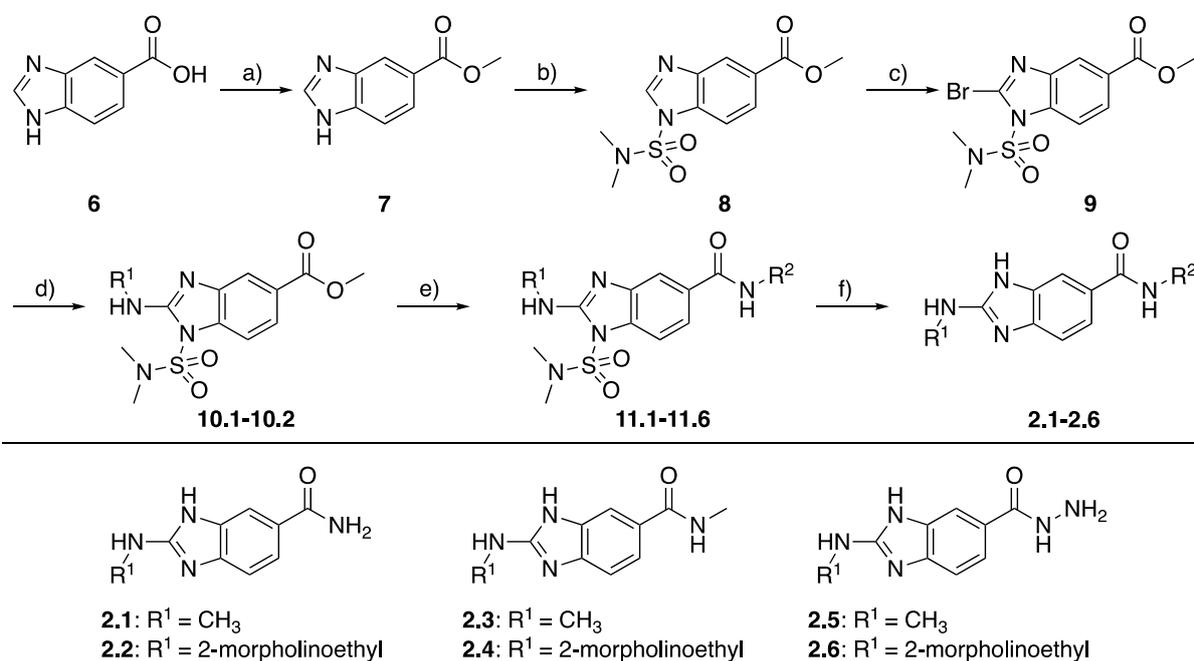


**Scheme 1.** Design of pyrimidone ring-opened benzimidazole (**2** and **3**) derivatives and tricyclic 7,8-dihydroimidazoquinazoline analogues (**4**) of the parent *lin*-benzoguanine scaffold (**1**).

**Synthesis.** The synthesis of 2-substituted *lin*-benzoguanine **1**<sup>[13]</sup> and *lin*-benzohypoxanthines **5**<sup>[14]</sup> (for the structure, see Table 1) was previously reported by our group. The 6-substituted benzimidazole ligands **2** were synthesized in 6 steps from 5-benzimidazole carboxylic acid (**6**) (Scheme 2).<sup>[13,15]</sup> The synthesis commenced with an esterification to give **7**. The benzimidazole nitrogen was sulfamoyl-

protected to give an inseparable, tautomeric mixture of **8** and **8'** (for the structure of **8'**, see Table S1 in the Supporting Information), followed by selective bromination at the 2-position. Partial separation of the tautomers was possible by flash column chromatography, and the enriched tautomer **9** was subjected to nucleophilic aromatic substitution with primary amines to afford 2-aminobenzimidazoles **10.1** and **10.2**. Functional group interconversion of the ester with aqueous ammonia, methanolic methylamine, or aq. hydrazine under microwave irradiation afforded the protected carboxamides and hydrazines **11.1-11.6**, respectively. Final deprotection with HCl (4 M solution in dioxane) and reverse phase flash chromatography purification yielded the ligands **2.1-2.6**.

### Scheme 2. Synthesis of 6-Substituted Benzimidazole Ligands **2**.<sup>a</sup>



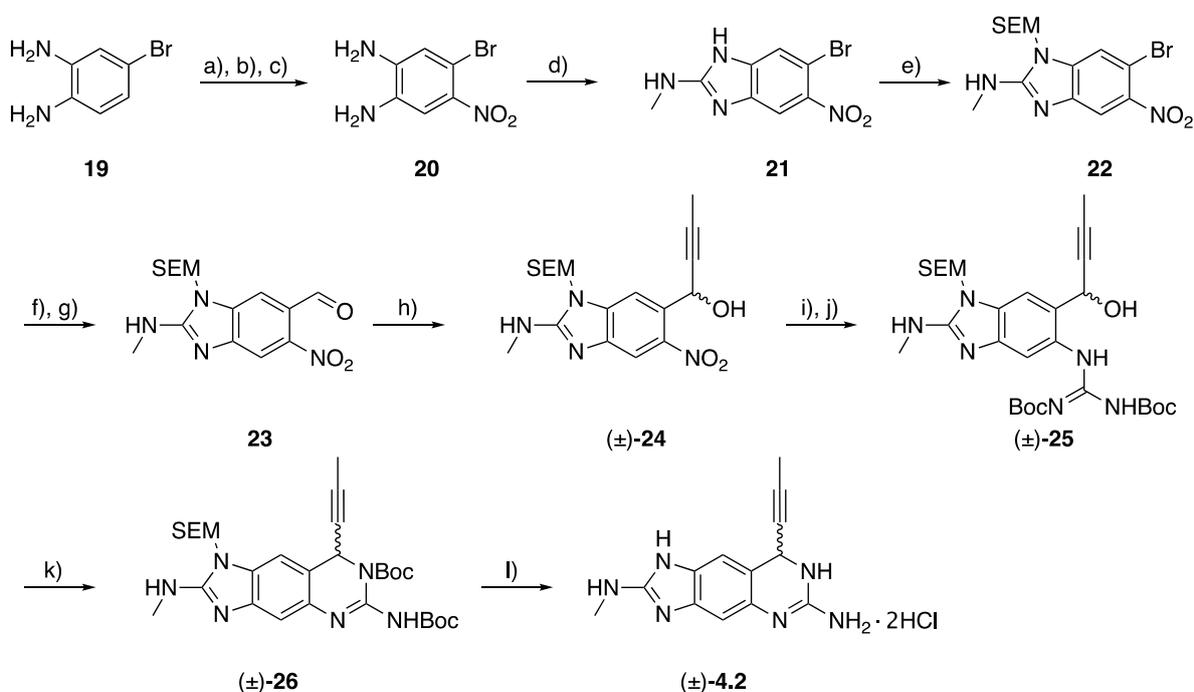
<sup>a</sup>Reaction conditions: a) H<sub>2</sub>SO<sub>4</sub>, MeOH, 65 °C, 18 h, quant.; b) Me<sub>2</sub>NSO<sub>2</sub>Cl, NEt<sub>3</sub>, toluene, 23 °C, 24 h, 66 % (1/0.6 mixture with tautomer **8'**; **8'** not shown); c) LiN(SiMe<sub>3</sub>)<sub>2</sub>, THF, -78 °C, 1h, then NBS, THF, -78 °C to 23 °C, 15 min, 56 % (1/0.05 mixture with tautomer **9'**; **9'** not shown); d) for **10.1**: MeNH<sub>2</sub>, EtOAc/MeOH, 23 °C, 14 h, quant.; for **10.2**: 2-morpholinoethylamine, *i*Pr<sub>2</sub>NEt, EtOAc/MeOH, 23 °C, 15 h, 54%; e) for **11.1** and **11.2**: aq. NH<sub>4</sub>OH, MeOH, 100 °C (MW), 15-30 min, **11.1**: 37%; **11.2**: 31%; for **11.3** and **11.4**: MeNH<sub>2</sub>, MeOH 140 °C (MW), 1.5-2.5 h, **11.3**: 67%; **11.4**: 75%; for **11.5** and **11.6**: N<sub>2</sub>H<sub>4</sub>, H<sub>2</sub>O, 100 °C (MW), 30 min, **11.5**: 80%; **11.6**: 47%; f) HCl (4 M solution in dioxane), dioxane, reflux, 14-45 h, **2.1**: 23%; **2.2**: 64%; **2.3**: 14 %; **2.4**: 19%; **2.5**: 9%; **2.6**: 52%. THF = tetrahydrofuran; NBS = *N*-bromosuccinimide; MW = microwave.

For the 5-substituted benzimidazole ligands **3.1** and **3.2**, we pursued an alternative synthetic strategy (Scheme 3).<sup>[15]</sup> 4-Nitroaniline **12** was converted into 5-nitrobenzimidazoles **13.1** and **13.2** *via* a two-step protocol with isothiocyanates, followed by treatment with EDC.<sup>[16]</sup> The nitro group was reduced using PtO<sub>2</sub> and Pd/C under H<sub>2</sub>, and the obtained aniline was guanylated with *N,N'*-di-Boc-1*H*-pyrazole-1-carboxamide to afford **14.1** and **14.2**, respectively. The *N*-Boc protecting groups were cleaved with HCl (4 M ethereal solution), and recrystallization afforded the ligands **3.1** and **3.2**.



regioisomeric mixture of **22** and **22'** (for the structure of **22'**, see Table S1 in the Supporting Information), and carbonylation *via* a sequence of a palladium catalyzed Heck reaction with ethyl acrylate, followed by oxidative cleavage with catalytic OsO<sub>4</sub>, afforded a mixture of *N*-tautomeric nitrobenzaldehydes **23** and **23'** (for the structure of **23'**, see Table S1 in the Supporting Information), which was separated by careful column flash chromatography. Alkynylation with *in situ* generated propynyllithium gave the racemic propargylic alcohol (±)-**24**. The nitro group was reduced selectively with zinc and aq. NH<sub>4</sub>Cl, and subsequently the activated guanylation reagent *N,N'*-di-Boc-1*H*-pyrazole-1-carboxamide was used to prepare (±)-**25**. Intramolecular cyclization of the guanidine in (±)-**25** on the propargylic alcohol was again achieved by activation with tosyl chloride under basic conditions to yield (±)-**26**. Final deprotection with SnCl<sub>4</sub> and subsequent purification by reverse-phase HPLC afforded the bis-hydrochloride salt of racemic (±)-**4.2**.

### Scheme 5. Synthesis of the Tricyclic Dihydro-imidazoquinazoline Ligand **4.2**.<sup>a</sup>



<sup>a</sup> Reaction conditions: a) TsCl, pyridine, 65 °C, 17 h, 83%; b) HNO<sub>3</sub>, AcOH, 50 °C, 1 h, 67%; c) H<sub>2</sub>SO<sub>4</sub>, H<sub>2</sub>O, 85 °C, 2 h, 87%; d) MeNCS, (CH<sub>3</sub>)<sub>2</sub>SO, 70 °C, 21 h, then EDC, 70 °C, 1 h, 56%; e) SEM chloride, NaH, DMF, 0 °C, 1.5 h, 77% (1.2/1 mixture with tautomer **22'**; **22'** not shown); f) CH<sub>2</sub>=CHCO<sub>2</sub>Et, [Pd(OAc)<sub>2</sub>], PPh<sub>3</sub>, toluene, 110 °C, 48 h; g) OsO<sub>4</sub>, NaIO<sub>4</sub>, THF/H<sub>2</sub>O, 23 °C, 36 h, 52% of **23** and 32% of **23'** (not shown) (over 2 steps). h) H<sub>3</sub>CCH=CHBr, *n*-BuLi, THF/*n*-hexane, -78 °C to 23 °C, 3 h, 83%; i) Zn, aq. NH<sub>4</sub>Cl, MeOH, 0 °C, 2.5 h; j) *N,N'*-di-Boc-1*H*-pyrazole-1-carboxamide, *i*Pr<sub>2</sub>NEt, CH<sub>2</sub>Cl<sub>2</sub>, 21 h, 50 °C, 57% (over 2 steps); k) TsCl, Cs<sub>2</sub>CO<sub>3</sub>, 23 °C, 4 h, 58%; l) SnCl<sub>4</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 0 °C to 23 °C, 6 h, 34%.

**Table 1.** Chemical Formulas, Inhibition ( $K_i$ ) or Dissociation ( $K_D$ ) Constants in  $\mu\text{M}$ , PDB codes, and resolution ( $\text{\AA}$ ) of the Complexes Formed by the Studied Compounds with *Z. mobilis* TGT.<sup>a</sup>

Ligand	Parent scaffold	R	$K_i$ [ $\mu\text{M}$ ]	PDB code	Resolution $\text{\AA}$
1.1		R = Me	$0.058 \pm 0.036^{[8]}$	4PUK <sup>[6]</sup>	1.49
1.2		R =	$0.006 \pm 0.006^{[8]}$	4PUJ <sup>[6]</sup>	1.42
2.1		R = Me	$300 \pm 37$	5J9M	1.33
2.2		R =	$544 \pm 43$	5JT5	1.21
2.3		R = Me	$270 \pm 50$	---	---
2.4		R =	$264 \pm 40$	---	---
2.5		R = Me	$283 \pm 40$	5J9N	1.64
2.6		R =	$282 \pm 18$	5JT6	1.54
3.1		R = Me	$19 \pm 2$	5J9O	1.41
3.2		R =	$58 \pm 2$	5JT7	1.70
4.1		R = H	$17.2 \pm 0.2$	6RKT	1.75
(±)-4.2		R =	$47.9 \pm 6.6$	6RKQ	1.67
5.1		R = Me	$6.5 \pm 2.9^{[9]}$	3S1G <sup>[9]</sup>	1.82
5.2		R =	$4.1^{[9]}$	4Q4R	1.45

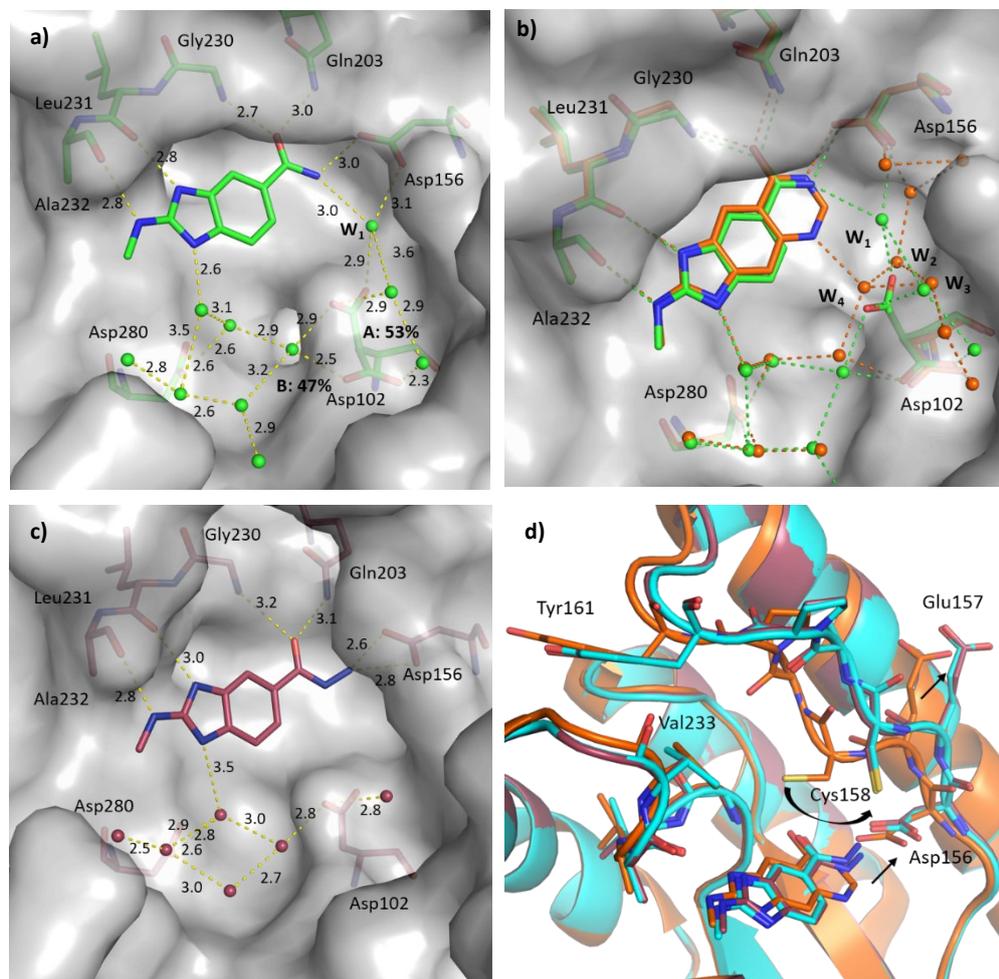
<sup>a</sup> Binding constants were determined *via* a radioactive assay.

1  
2  
3  
4  
5  
6 **Affinity Determination.** The affinities of the studied compounds were determined using a radioactive  
7 assay. This assay is a functional biochemical assay that measures the incorporation of radioactively  
8 labeled guanine into tRNA. The potency of an inhibitor for TGT is quantified by the amount of  
9 decreasing the rate of guanine incorporation into tRNA resulting in a  $K_i$  value.<sup>[19]</sup>  
10  
11

12  
13 **Crystal Structure Determinations.** The crystal structures were obtained by co-crystallizing the  
14 corresponding ligands with the enzyme. In all cases, structures with the common twofold dimer  
15 packing in space group C2 could be determined at a resolution between 1.21 – 1.82 Å. In some cases,  
16 both a structure with  $R^1 = \text{Me}$  and  $R^1 = \text{morpholino ethyl}$  could be resolved. However, as previously  
17 observed, the morpholino substituent showed pronounced positional disorder in all examples apart  
18 **1.2**<sup>[13]</sup> and no difference electron density could be attributed to this part of the ligand. The structural  
19 comparison of the corresponding complexes with  $R^1 = \text{Me}$  and morpholino ethyl showed identical  
20 binding modes, which are adopted in all cases. Accordingly, in the following section only crystal  
21 structures of the  $R^1 = \text{Me}$  derivatives will be described. The crystallographic tables of all here  
22 investigated complexes can be found in the Supporting Information and the coordinates of all  
23 complexes have been deposited with the Protein Data Bank (PDB; www.rcsb.org).  
24  
25  
26  
27  
28  
29  
30  
31

32 **Co-Crystal Structures with the Benzimidazoles of 6-Carboxamide 2.1 and 6-Hydrazide 2.5.** As  
33 observed in previous studies, the bicyclic benzimidazole core scaffold binds to the nucleobase  
34 exchange pocket and forms the expected H-bonding pattern to the backbone atoms of the amino acids  
35 Ala232, Leu231 and Gly230 (Figure 2a,b).<sup>[12,13,20]</sup> The carboxamide terminus of Gln203 acts as H-bond  
36 donor to the carbonyl groups of the attached 6-amide or 6-hydrazide substituents of **2.1** and **2.5**.  
37 Furthermore, the bicyclic core structure is sandwiched between the amino acids Tyr106 and Met260,  
38 establishing  $\pi$ -stacking interactions (not shown). The side chain of the catalytically active Asp102  
39 adopts two alternative, equally populated conformations in the complex with **2.1** (Figure 2a), whereas  
40 in the complex with the hydrazide substituent **2.5**, only the conformer with the carboxylate group  
41 oriented into the recognition pocket is observed (Figure 2c). The geometry with the side chain rotated  
42 off the catalytic center is also found in the structure of the apo protein<sup>[21]</sup> and complexes  
43 accommodating *lin*-benzohypoxanthines (**5.1**, **5.2**).<sup>[14]</sup> The alternative orientation towards the ligand-  
44 binding site is generally found for complexes with *lin*-benzoguanines. In TGT-**2.1**, the carboxylate group  
45 of Asp102 binds to the primary amide function of the ligand (2.9 and 3.0 Å) *via* the interstitial water  
46 molecule W1 (Figure 2b). The second oxygen of this carboxylate is solvated by additional water  
47 molecules found in the ribose-34 pocket (2.9 - 3.1 Å). The carboxylate group of Asp156 also forms an  
48 H-bond to the terminal amide nitrogen of **2.1** (3.0 Å) and water W1 (Figure 2b) additionally solvates  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

the neighboring oxygen of the Asp156 carboxylate (3.1 Å). Apart from these changes in the local water structure, the complex with **2.1** is very similar to that with **5.1** (Figure 2b).

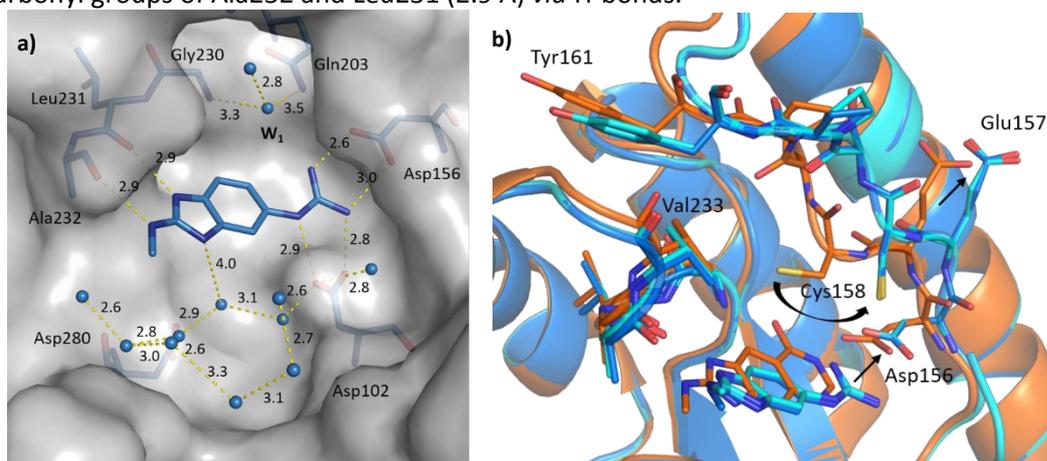


**Figure 2:** a) Crystallographically determined binding mode of **2.1** (carbon atoms green, PDB code: 5J9M) in the recognition pocket of TGT, heteroatoms type-coded, water molecules as small spheres, all distances in Å, hydrogen-bonds as dashed lines; b) Superposition of **2.1** (carbon atoms green PDB code: 5J9M) and **5.1** (carbon atoms orange, PDB code: 3S1G<sup>[14]</sup>); c) Crystallographically determined binding mode of **2.5** (carbon atoms dark red, PDB code: 5J9N); d) Superposition of **2.5** (carbon atoms dark red, PDB code: 5J9N), **2.6** (carbon atoms cyan, PDB code: 5JT6) and **5.1** (carbon atoms orange, PDB code: 3S1G<sup>[14]</sup>), the push back of Asp156 and the conformational transformation of the protein opening the transient pocket in the complexes with **2.5** and **2.6** in comparison to **5.1** is indicated by arrows. For electron density around the bound ligands and water molecules, s. Figure S50, Supporting Information).

In the complex with the hydrazide-based benzimidazole **2.5** (Figure 2c), the terminal nitrogen of the hydrazide protrudes directly toward the carboxylic acid carbon of the Asp156 side chain and forms two H-bonds to its oxygens (2.6 - 2.8 Å). The hydrazide group adopts the *trans* geometry (179°) and pushes Asp156 by approx. 1.2 Å ( $C_{\gamma} \cdots C_{\gamma}$ ) out of the recognition pocket compared to the complex with **5.1**. This push back of Asp156 is associated with a reorganization of further neighboring amino acids and leads to a conformational change of the loop between the amino acids Glu157 and Tyr161 (Figure 2d and

Supporting Information Figure S49). As mentioned above, the carboxylate group of Asp102 in **2.5** refines to one fully populated orientation toward the ligand. Nevertheless, its oxygens do not build a water-mediated interaction network with the ligand, as found for **2.1**. No sufficiently well-defined difference electron density for a water molecule could be detected. In contrast, the interactions to the water molecules in the ribose-34 pocket are also established here. The affinities of the carboxamide **2.1** ( $K_i = 300 \pm 37 \mu\text{M}$ ) and the morpholino analog **2.2** ( $K_i = 544 \pm 43 \mu\text{M}$ ) decrease by a factor of 46 and 132, respectively, compared to the analog *lin*-benzohypoxanthines **5.1** ( $K_i = 6.5 \pm 2.9 \mu\text{M}$ ) and **5.2** ( $K_i = 4.1 \mu\text{M}$ ). In the crystal structures with the latter ligands a well-defined water network between the ligand and Asp102 is formed (Figure 2b).<sup>[14]</sup> A similar network is impossible to establish with **2.1**, which likely explains the reduced affinity of the 5-amide derivatives **2.1** and **2.2**. The affinity loss compared to the parent *lin*-benzoguanine scaffold (**1.1**, **1.2**) is even more dramatic and accounts for nearly five orders of magnitude with respect to **2.1** (Table 1). The structural rearrangement of the loop extending from Glu157 to Tyr161 along with the loss of interactions and the reorganization of water molecules makes it difficult to compare the hydrazide-based ligands **2.5** and **2.6** with the 6-amide analogs **2.1** and **2.2**. Nevertheless, it appears surprising that for **2.1** ( $K_i = 300 \pm 37 \mu\text{M}$ ) and **2.5** ( $K_i = 283 \pm 40 \mu\text{M}$ ) the affinity falls into the same range. This suggests that only minimal energetic costs are involved in the structural rearrangement of the protein, which may represent the transformation to a conformation linked to a protein function that stabilizes the bound ligands.

**Crystal structures with the ring opened 5-guanidino-benzimidazole 3.1.** Compared to the *lin*-benzoguanine scaffold, the 5-guanidino-benzimidazole **3.1** lacks the endocyclic carbonyl function to establish direct H-bonding interactions with Gly230 and Gln203. The corresponding crystal structure shows (Figure 3a) that these amino acids are now solvated and stabilized (3.3 - 3.5 Å) *via* hydrogen bonds to an interstitial water molecule (W1, Figure 3a). The carboxylate groups of Asp102 and Asp156 each adopt an orientation which is analogously observed with the bound *lin*-benzoguanine scaffold **1.1**.<sup>[11]</sup> The basic guanidino group ( $\text{p}K_a > 12$ ) likely binds in charged state. Consequently, salt bridges are formed, as expected, with both aspartates. Furthermore, the 2-aminoimidazole interacts with the carbonyl groups of Ala232 and Leu231 (2.9 Å) *via* H-bonds.

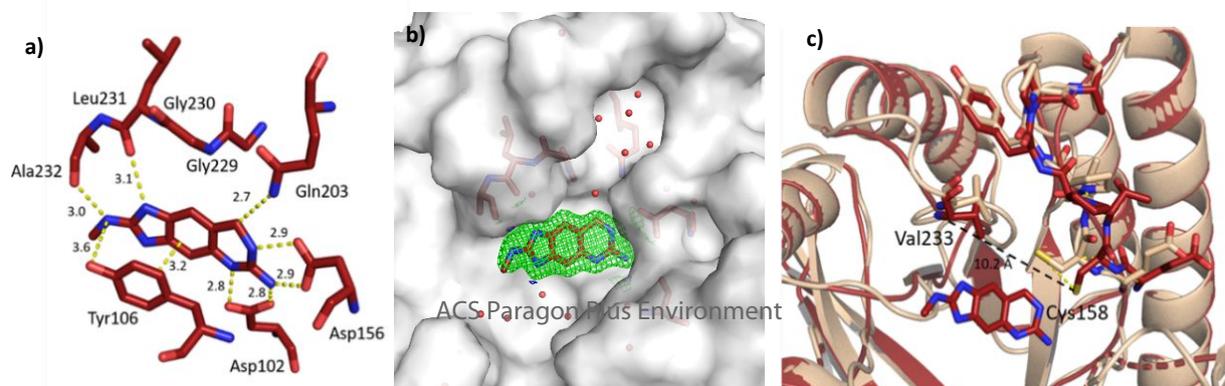


**Figure 3:** a) Crystallographically determined binding mode of **3.1** (PDB code: 5J9O) in the recognition pocket of TGT, heteroatoms type-coded, water molecules small spheres, all distances in Å, hydrogen-bonds as dashed lines; b) Superposition of **3.1** (carbon atoms cyan, PDB code: 5J9O), **3.2** (carbon atoms light blue, PDB code: 5JT7) and **5.1** (carbon atoms orange, PDB code: 3S1G)<sup>[14]</sup> shows the push back of Asp156 and the triggered conformational transformation of the protein opening the transient pocket in the complexes with **3.1** and **3.2** in comparison to **5.1** (indicated by the arrows). The latter ligand has only one H-bond donor to interact with Asp156, therefore a slightly tilted geometry is found, however, no opening of the transient pocket is initiated. For electron density around the bound ligands and water molecules, s. Figure S50, Supporting Information).

At first sight, the loss of the H-bonds to Gly230 and Gln203 also leads to a drop in affinity as observed for the 6-amide analogs **2.1** and **2.5** with respect to the corresponding *lin*-benzoguanine **1.1** or *lin*-benzohypoxanthine **5.1**. With  $K_i = 19 \pm 2 \mu\text{M}$  for **3.1** and  $K_i = 58 \pm 2 \mu\text{M}$  for **3.2**, the 5-guanidine-substituted ligands bind slightly stronger to TGT than the studied 6-benzimidazoles. This suggests that the gained charge-assisted interactions to Asp102 and Asp156 can only partially compensate for the lost interactions to the nitrogen atoms of Gly230 and Gln203. However, a perhaps even more important effect must be considered. By opening the pyrimidone ring of the *lin*-benzoguanine scaffold, the ligand scaffold loses its correct preorganization. Instead, the guanidine portion, attached to the benzimidazole, must adopt one single conformation, which supposedly is not the most favored one. This, along with the loss of conformational degrees of freedom, will have a detrimental effect on binding affinity. In other studies we saw that this effect can reduce the binding constant by two to three orders of magnitude.<sup>[22]</sup>

Interestingly, as with hydrazide-substituted benzimidazoles **2.5** and **2.6**, the structural transformation of the secondary structural element (Glu157 to Tyr161) is also observed with the guanidine-substituted ligands **3.1** and **3.2** (Figure 3b). Analogously, a spatial shift of Asp156 is observed, as seen in TGT·**2.5** ( $C_\gamma\text{---}C_\gamma = 0.9 \text{ \AA}$ ) compared to the complex with **1.1**. However, a second aspect may be of importance in this complex to transfer the enzyme into the alternative state with the opened transient pocket. The above-mentioned water W1 (Figure 3a), which binds in the vicinity of the original position of the carbonyl O-atom of the *lin*-benzoguanine scaffold, solvates the polar groups of Gly230 and Gln203, but in addition, it fills the created void and exerts steric pressure onto the flexible loop of Glu157 to Tyr161 (Figure 3a). Accordingly, W1 supports the expansion of the opening pocket.

To study the impact of steric pressure in this region of the binding pocket by a bound ligand in more detail, we designed chemically reduced tricyclic analogs of the *lin*-benzoguanine scaffold bearing an

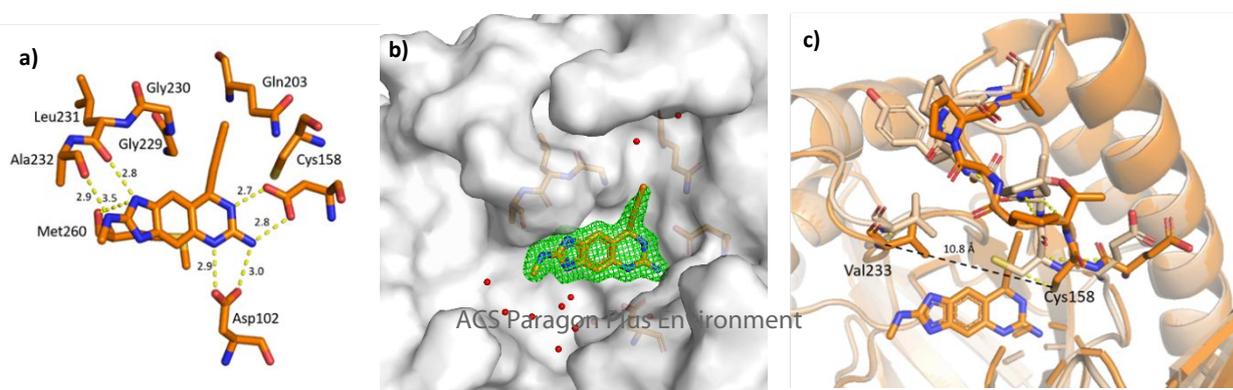


1  
2  
3  $sp^3$ -carbon atom at the former spatial position of the carbonyl group in pyrimidone ring (**4** in Scheme  
4 1). Through the attachment of substituents at this carbon atom, a stereogenic center is created.  
5 Therefore, we decided to first study the achiral unsubstituted derivative **4.1** and subsequently the  
6 propyn-1-yl analog (( $\pm$ )-**4.2**). Modeling considerations strongly supported the hypothesis that only the  
7 *R*-enantiomer will orient the attached substituents in a way to fit into the possibly opened transient  
8 pocket, as modeling the *S*-enantiomer would create major steric conflicts with the protein. The  
9 compounds were synthesized (Schemes 4 and 5) and subsequently subjected to co-crystallization.

15  
16 **Figure 4:** a) Crystallographically determined binding mode of **4.1** (carbon atoms dark red, PDB code:  
17 6RKT) in the recognition pocket of TGT, heteroatoms type-coded, all distances in Å, hydrogen-bonds  
18 as dashed lines; b) **4.1** difference electron density ( $|F_o| - |F_c|$  omit map) contoured at  $3\sigma$ , protein  
19 displayed with the solvent accessible surface showing the recognition site together with the opened  
20 transient pocket; c) Conformational transformation of the protein induced by the binding of **4.1**  
21 (carbon atoms dark red, PDBcode: 6RKT). For comparison, the complex with **4.1** is superimposed  
22 with the crystal structure of the apo protein (carbon atoms in wheat color). For electron density around the  
23 bound ligands and water molecules, s. Figure S50, Supporting Information).

26 **Co-Crystal Structures with Dihydro-imidazoquinazolines **4.1** and ( $\pm$ )-**4.2** featuring the Tricyclic *lin*-  
27 Benzoguanine Scaffold.** Ligand **4.1** binds with an occupancy of 87% to the recognition pocket. The  
28 guanidine moiety in the six-membered ring forms two bidentate hydrogen bonds with the carboxylate  
29 groups of Asp156 and Asp102 (Figure 4a). The second guanidine moiety in the imidazole ring interacts  
30 *via* a hydrogen bond with the carbonyl oxygen of Leu231, while the exocyclic amino group hydrogen-  
31 bonds the carbonyl oxygen of Ala232. Furthermore, **4.1** establishes a  $\pi$ - $\pi$  stacking by being sandwiched  
32 between the side chains of Tyr106 and Met260. Thus, a very similar recognition pattern as seen for *lin*-  
33 benzoguanine **1.1**<sup>[11]</sup> is found in this area. Nevertheless, the ligand also induces the structural  
34 transformation of the secondary structural element (loop Glu157 to Tyr161), which leads to an opening  
35 of the transient sub-pocket (Figure 4b, c) and the carboxamide group of Gln203 forms a van der Waals  
36 contact to C8 of the dihydro-imidazoquinazoline scaffold. Due to the pocket opening, the distance  
37 between Val233 C $\alpha$  and Cys158 S $\gamma$  expands in TGT-**4.1** to 10.2 Å, which amounts in the complex with  
38 the *lin*-benzoguanine parent scaffold **1.1** to 6.5 Å.

39  
40 The complex formed with ( $\pm$ )-**4.2** shows a very similar geometry to TGT-**4.1** with nearly identical  
41 recognition pattern. In contrast to the latter complex, Gln203 adopts an altered conformation  
42 providing additional space to accommodate the needle-like propyn-1-yl side chain. Even though the  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54



1  
2  
3 complex was co-crystallized from a racemic mixture, the enzyme only picks the *R*-enantiomer. During  
4 structure determination, placement of the *S*-enantiomer into the electron density lead to negative and  
5 positive sigma values in the  $m|Fo| - |Fc|$  map after refinement, which indicates an incorrect solution.  
6 Thus, in agreement with our modeling considerations described above, the *S*-enantiomer does not  
7 bind as it would have to poke into the protein and create substantial steric clashes. The difference  
8 density allows placement of all atoms of the propyn-1-yl side chain, even though the less-well defined  
9 density next to the terminal methyl group indicates enhanced residual mobility of the side chain  
10 terminus. As for the other complex, ( $\pm$ )-**4.2** induces a spatial shift of the loop comprising Glu157 to  
11 Tyr161. Asp156 is, compared to TGT-**1.1**, pushed by 0.6 Å out of the recognition pocket and helps to  
12 trigger the structural transformation of the loop (Figure 5c). The opening of the transient pocket results  
13 in a mutual distance between Val233 C $\alpha$  and Cys158 S $\gamma$  of 10.8 Å.  
14  
15  
16  
17  
18  
19  
20  
21

22 **Figure 5:** a) Crystallographically determined binding mode of (*R*)-**4.2** (carbon atoms orange, PDB code:  
23 6RKQ) in the recognition pocket of TGT, heteroatoms type-coded, all distances in Å, hydrogen-bonds  
24 as dashed lines; b) (*R*)-**4.2** together with the difference electron density ( $|Fo| - |Fc|$  omit map)  
25 contoured at  $3\sigma$ , protein displayed with the solvent accessible surface showing the recognition site  
26 together with the opened transient pocket; c) Conformational transformation of the protein induced  
27 by the binding of (*R*)-**4.2** (carbon atoms orange, PDB code: 6RKQ). For comparison, the complex with  
28 **4.2** is superimposed with the crystal structure of the apo protein (carbon atoms in wheat color). For  
29 electron density around the bound ligands and water molecules, s. Figure S50, Supporting  
30 Information).  
31  
32  
33

34 **Expansion of the Recognition Pocket by a Transient Sub-Pocket.** The enzyme TGT can perform large  
35 conformational adaptations upon ligand binding.<sup>[21]</sup> The opening of the small transient pocket is not  
36 only observed in our initially described fragment structure<sup>[3]</sup> (Figure 1) but also in the here described  
37 subset of complexes. These complexes push Asp156 slightly out of the recognition pocket and create  
38 sufficient steric pressure onto residues Gly230 and Gln203, thus initiating the opening of the transient  
39 binding cleft. Transient pockets may populate opened or closed states within the conformational  
40 dynamics of the protein under thermodynamic equilibrium conditions. Alternatively, transient pockets  
41 may open only in the presence of ligating species such as a drug or substrate molecule, a situation that  
42 we could observed for the specificity pocket found in the enzyme aldose reductase.<sup>[23]</sup> Under either  
43 scenario, conformational selection or induced fit, energetically favorable ligand binding is possible.  
44 However, in both cases differing thermodynamic consequences accompanied by fundamentally  
45 different binding mechanisms are at play. We therefore performed MD simulations in order to obtain  
46 some insights into the distribution of opened and closed states of the binding pocket. We found that  
47 the binding pocket of the *apo* enzyme is transiently adopting opened and closed states during the MD  
48 simulation (3x 200 ns simulations). We first used the mutual Cys158S $\gamma$ -Val233C $\alpha$  distance (see analysis  
49 of the crystal structures above) but furthermore the adopted volume of the binding pocket to separate  
50 both states. The distance descriptor shows a broad distribution suggesting multiple states for the  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

transition between both states (see Figure S1). Nonetheless, a large portion of conformations during the MD simulation adopts distances between 6 - 8 Å, thus indicating a closed binding pocket. The volume descriptor records a bimodal distribution with maxima at 150 Å<sup>3</sup> and 650 Å<sup>3</sup>. The ratio of the integrated densities of the two volume distributions suggests that 13% and 86% of the time the lower and upper distributions are populated (see Figure S1). The mutual correlation of both descriptors (see Figure S2) reveals that the small volumes are only present at lower distance range. This is in agreement with the crystal structures and we assign distances between 6 – 8 Å to the closed state. Using the volume distribution as molecular descriptor to characterize the opened state is somewhat exaggerating as it also contains states that depict an expanded volume but do not properly correspond to the protein in opened states of the transient binding pocket. In summary, the MDs support the transient nature of the binding pocket. However, they suggest a more complex distribution of conformations along the Cys158S<sub>γ</sub>-Val233C<sub>α</sub> distance descriptor of the *apo* protein under dynamic conditions than being observed in the ligand-bound crystal structures. We therefore also performed simulations starting with TGT and its bound natural substrate preQ<sub>1</sub>. This ligand is hydrogen-bonded to Asp156 within 95% of all MD frames. However, the interactions hardly trigger any of the above-described steric pressure on Asp156 since the binding pocket is exclusively observed in closed state (see Figures S1 and S4). This finding thus indicates that once the interaction to Asp156 is stably formed by the substrate, the transient behavior of the binding pocket is suppressed.

**Table 2.** Mutual Cys158S<sub>γ</sub>-Val233C<sub>α</sub> distances and backwards push of Asp156 C<sub>γ</sub> in different TGT·ligand complexes with respect to TGT·1.1.

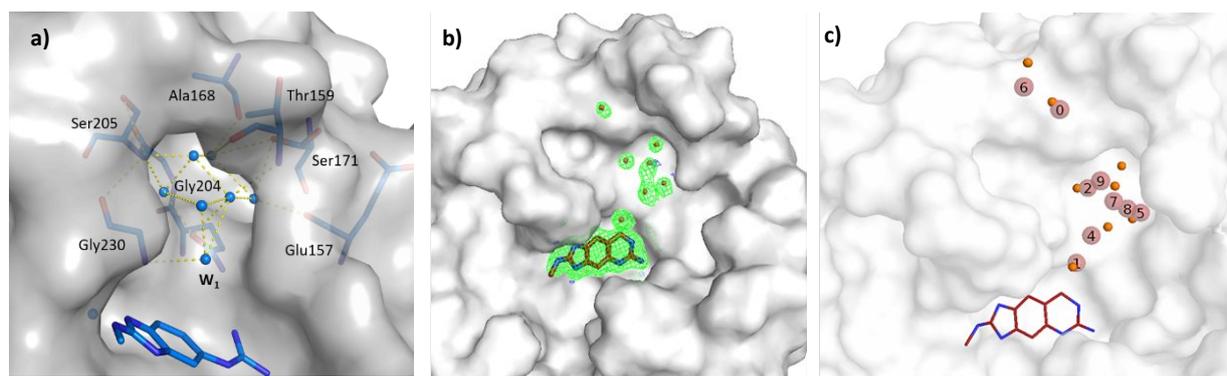
complex	Cys158S <sub>γ</sub> -Val233C <sub>α</sub> distance [Å]	Asp156 push compared to TGT·1.1 [Å] <sup>a)</sup>	pdb-code
TGT·1.1	6.5	0.0	4PUK
TGT·frag	9.8	0.4	5UTI
TGT·2.1	6.4	0.2	5J9M
TGT·2.2	6.2	0.1	5JT5
TGT·2.5	10.2	0.9	5J9N
TGT·2.6	10.2	0.6	5JT6
TGT·3.1	10.5	0.7	5J9O
TGT·3.2	10.6	0.7	5JT7
TGT·5.1	6.2	0.3	3S1G
TGT·4.1	10.2	0.3	6RKT
TGT·(R)-4.2	10.8	0.6	6RKQ

<sup>a)</sup> The amount by which Asp156 is pushed out of the recognition pocket was determined by performing an overall RMSD C<sub>α</sub> atom superposition of the studied complex with TGT·1.1 and measuring the mutual distances of Asp156C<sub>γ</sub>s in both structures.

The transformation of the protein, leading to the expansion of the binding pocket, correlates with the properties of the substituents in **2.1** – **2.6**, **3.1**, and **3.2** attached to the benzimidazole or the dihydroimidazoquinazoline scaffold (**4.1** and (*R*)-**4.2**). The primary amide stabilizes the closed conformation found in most TGT·ligand complexes. The extension of the amide by an additional amino group to a

1  
2  
3 hydrazide leads to the structural transformation. The attached terminal amino group of the hydrazide  
4 exerts steric pressure on Asp156 and thus induces the protein rearrangement leading to its  
5 stabilization. The guanidine-substituted ligands show the same reorganization as a combined effect of  
6 steric pressure and altered solvation of the amino acids Gly230 and Gln203. The water molecule W1  
7 (Figure 3a) and the slight spatial shift of Asp156 are responsible for this effect. In the complexes with  
8 the dihydro-imidazoquinazolines **4.1** and (*R*)-**4.2** the steric pressure is exerted by the attachment at  
9 C8. Obviously, solely a hydrogen as substituent is sufficient to trigger the pocket opening.

10  
11 The newly formed cavity is occupied and stabilized by several water molecules, which indicates the  
12 hydrophilic character of the created pocket (Figure 6). A complex water network with short distances  
13 between 2.5 and 3.5 Å is detected involving the carbonyl groups of the Glu157, Thr159, Ala168, and  
14 Gly230 and the backbone amide nitrogens of Gly204, Ser205, and Thr159, which protrude toward the  
15 expanded pocket. In addition, the side chains of Gln203, Ser171, and Ser205 form H-bonds to the  
16 accommodated water molecules.



17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Figure 6:** a) Binding pose of **3.1** together with the opened transient pocket, which is filled with a cluster of water molecules. The protein is displayed by its solvent accessible surface, and the residues involved in interactions with the water cluster are indicated, hydrogen bonds as dashed lines. b) Binding mode of **4.1** (PDB code: 6RKT) together with the difference electron density ( $|F_o| - |F_c|$  omit map) contoured at  $3\sigma$ , protein displayed with the solvent accessible surface showing the recognition site together with the opened transient pocket which is filled by water molecules; c) Comparison of the spatial arrangement of computed hydration sites (numbered pink balls, Table 3) together with the crystallographically determined waters in TGT-**4.1** (orange spheres, cf. b)).

Depending on the achieved resolution, the amount of water molecules captured by crystallography will differ.<sup>[25]</sup> Furthermore, crystal structures only provide implicit information about the thermodynamic properties of the water molecules found in such a transient pocket. However, upon ligand binding to the protein, the solvent molecules occupying the pocket will be displaced to the surrounding bulk solvent phase. To estimate the contribution of the water displacement to the overall inventory of ligand binding, the contributions resulting from the displacement of these waters are important. We usually assume that the hydration of the transient pocket occurs instantly as soon as the pocket undergoes the transition into the opened state. However, the solvation thermodynamic

properties of the water molecules (i.e. free energy, enthalpy and entropy of solvation) in the opened pocket must ultimately be a contribution to the thermodynamic properties of the opened transient pocket itself. Thus, the distribution of opened and closed pocket conformations depends on the water molecules as well as the internal molecular interactions of the protein. Waters displaced from hydration sites that are energetically less favorable than in the bulk, will contribute favorably to ligand binding. While displacement from hydration sites that are more favorable in the protein than in the bulk, will counteract binding. The energetic contributions are usually counterbalanced by a favorable entropic contribution resulting from the displacement of bound water molecules into bulk water phase. In order to investigate the thermodynamic inventory of the water molecules in the transient binding pocket in conjunction with possible consequences for ligand binding, we performed MD simulations and hydration site analysis (HSA).<sup>[26]</sup> We carried out 30 ns MD simulations based on the bound protein structure of TGT-4.1 but with the ligand removed from the binding site. Then, we processed the MD trajectories using the HSA implementation of the SSTMap package in order to determine the enthalpic and entropic contributions of the bound water molecules with respect to pure bulk water (Table 3). Our calculations suggest that individual water molecules cluster at representative sites, which reasonably well match with the crystallographically determined water molecules (Figure 6b, c). Thus, the crystal structure of TGT-4.1 reveals the presence of eight water molecules in the transient binding pocket, while MD simulations suggests a similar cluster of nine hydration sites populating this region. The distances between the assigned crystallographic water molecules and the centers of the computed hydration sites are listed in Table 3.

**Table 3:** Hydration site analysis of the 9 most prominent hydration sites found in the transient recognition pocket by HSA, ranked by  $\Delta G_{\text{solv}}$ .

Hydration Site	Distance to closest crystal water [ $\text{\AA}$ ]	$\Delta H_{\text{solv}}$ [ $\text{kJ}\cdot\text{mol}^{-1}$ ]	$T\Delta S_{\text{solv}}$ [ $\text{kJ}\cdot\text{mol}^{-1}$ ]	$\Delta G_{\text{solv}}$ [ $\text{kJ}\cdot\text{mol}^{-1}$ ]
9	0.5	12.51	-16.40	28.90
1	0.89	5.23	-22.64	27.86
4	1.37	1.42	-19.00	20.43
0	0.68	-1.17	-21.34	20.13
2	0.72	-3.18	-22.38	19.18
5	0.73	-2.85	-20.29	17.47
7	1.08	-3.31	-16.95	13.65
8	2.3	-5.27	-18.24	12.97
6	1.55	0.63	-12.30	12.92

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

The HSA calculations suggest that all the waters in the transient pocket have unfavorable Gibbs free binding energies with respect to the bulk phase. Hydration sites 9 and 1 are energetically highly depleted, suggesting unfavorable interactions in the transient binding pocket compared to bulk solvent. We therefore expect that their displacement favorably contributes to ligand binding. This hypothesis is supported by the observation that the ligands, which open the transient pocket and displace some of the water molecules in the pocket (e.g. **4.1** and (*R*)-**4.2**, note, the affinity was determined for the racemic mixture, likely only the *R*-enantiomer shows inhibitory potency), are by a factor of 10 more potent compared to the members of series **2** and equipotent to those of series **3**.

**Functional Role of the Transient Pocket and Outlook for Further Inhibitor Design.** From the crystallographic as well as the computational perspective, the solvation structure in the studied binding pocket has quite a critical role for the understanding of the binding behavior as well as the design of successive ligand molecules. For the transient pocket studied in the present contribution, it is most remarkable that the hydration sites computed with our HSA calculations are quite unfavorable in free energy compared to their bulk solvent state. We already observed similar behavior in the case of human Aldose Reductase,<sup>[23]</sup> in which a transient pocket was occupied with an energetically depleted solvent pattern. Although the present case of TGT most likely demonstrates a different mechanism of pocket opening and closing, we reason that the very unfavorable solvation patterns are key for the understanding of the stability of transient binding pockets. From the crystallographic point of view, the present case of TGT does not indicate any obviously unfavorable interactions between protein atoms at this pocket, such that the role of the solvation thermodynamics must be a critical factor for the population of the different states of the pocket.

We inspected the *Z. mobilis* TGT complex structures, deposited in the Protein Data Bank (PDB, [www.rcsb.org](http://www.rcsb.org)).<sup>[27]</sup> The multiple structures show that even though the loop between Glu157 and Tyr161 is structurally in most complexes highly conserved, the side chain of Cys158 occurred already in a few structures with alternative conformations. In these examples, the Cys158<sub>S<sub>v</sub></sub> - Val233C<sub>α</sub> distances vary from 6.6 Å to 9.8 Å. However, the examples with expanded distances are predominantly mutated variants of TGT in the guanine/preQ<sub>1</sub> recognition pocket. These variants were produced to investigate the differences of the eubacterial and eukaryotic TGTs.<sup>[5]</sup> The latter enzymes recognize queuine as significantly expanded substrate. In the human enzyme as in most other eukaryotic species, Cys158 is replaced by Val and the opposing Val233 is exchanged to Gly.<sup>[28]</sup> Remarkably, the crystal structure of queuine (PDB code: 3BLO)<sup>[5]</sup> with this variant shows a similar expansion of the recognition pocket as observed in the complexes studied here. Possibly, the latter complexes indicate an intrinsic

1  
2  
3 conformational flexibility the architecture of TGT enzymes exhibits in this area that is functionally  
4 important for the eukaryotic species to recognize its extended substrate.  
5

6  
7 The pronounced similarity of the recognition site in the eubacterial variant, which hosts the  
8 nucleobases guanine and preQ<sub>1</sub>, and the eukaryotic one means a real challenge developing species-  
9 selective inhibitors. However, one observation with our expanded inhibitors might provide a promising  
10 opportunity to approach this problem. Remarkably, the Cys158 residue is highly conserved across all  
11 eubacterial enzymes but lacking in eukaryotic species. The thiol group of a cysteine residue is a  
12 potential anchor to covalently attach an inhibitor, e.g. *via* a Michael acceptor or a potentially alkylating  
13 group.<sup>[29–31]</sup> For the development of a potent anti-infective, irreversible inhibitors may constitute a  
14 desired strategy to follow. The ligands studied here based on the 4-guanidino benzimidazoles or the  
15 C(8)-functionalized dihydro-imidazoquinazoline scaffold induce the opening of the transient pocket  
16 and make the thiol group of the crucial cysteine accessible for chemical modification. *Via* appropriate  
17 design, it may be possible to develop inhibitors, preferentially based on the latter likely more potent  
18 dihydro-imidazoquinazoline scaffold, that selectively and irreversibly block the bacterial TGT isoforms  
19 via an electrophilic warhead capable to covalently bind to the thiol group of Cys158.  
20  
21  
22  
23  
24  
25  
26  
27  
28

## 29 CONCLUSION

30  
31  
32 Crystallographic fragment screening not only provides structural information useful for lead  
33 development but can also unveil surprising hits pointing towards intrinsic functional properties of, e.g.,  
34 enzymes important for some members of a given family. The screening of a 96-fragment library against  
35 the shigellosis target TGT led to the discovery of a fragment molecule that opens a transient sub-pocket  
36 in the preQ<sub>1</sub> recognition site of the enzyme by pushing back Asp156 in the catalytic center. The latter  
37 movement is associated with a reorganization of further neighboring amino acids and leads to a  
38 conformational change of the loop between the amino acids Glu157 and Tyr161. The transient pocket,  
39 which had been predicted before by MD simulations, causes an expansion of the preQ<sub>1</sub> recognition  
40 pocket volume between 250 and 750 Å<sup>3</sup>. Inspired by these observations, we designed and synthesized  
41 a series of model compounds to target the opening of the pocket, including a dihydro-  
42 imidazoquinazoline derivative with a propyn-1-yl exit vector that points directly into the transient  
43 pocket and causes the displacement of a conserved water network. The latter compound, which is a  
44 reduced analog of our initial parent ligand **1.1** is still far from being optimized and shows reduced  
45 potency compared to **1.1**. Our MD simulations and hydration site analysis used to determine the  
46 thermodynamic contribution of the water network to the binding of the reduced ligand scaffold  
47 showed that the displacement of the involved water molecules will contribute favorably to ligand  
48 binding. Additionally, the Cys158 residue of the TGT binding pocket was traced to be the gatekeeper  
49 of the transient sub-pocket. The conservation of this Cys158 in 99% of eubacterial TGT and the  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 presence of Val145 in its eukaryotic counterpart opens the possibility of developing inhibitors based  
4 on the dihydro-imidazoquinazoline scaffold, that selectively and irreversibly block the bacterial TGT  
5 isoforms via a warhead capable to covalently bind to the thiol group of Cys158 as the latter functional  
6 group becomes accessible upon pocket opening.  
7  
8  
9

## 10 11 12 **EXPERIMENTAL SECTION**

13  
14 **Expression and Purification of the *Z. mobilis* TGT.** The expression and purification protocol relates to  
15 work published earlier by us.<sup>[32]</sup> The *E. coli* cells BL21-CodonPlus (DE3)-RIPL (Cam<sup>r</sup>), transformed with  
16 the plasmid vector pPR-IBA2-ZM10 (Amp<sup>r</sup>), were incubated in a pre-culture of 100 mL LB medium  
17 containing 100 mg · L<sup>-1</sup> ampicillin and 34 mg · L<sup>-1</sup> chloramphenicol for 17 h at 37°C and 220 rpm. In  
18 addition to ampicillin and chloramphenicol resistance, the plasmid contains a sequence encoding an  
19 N-terminal Strep-tag II<sup>®</sup> separated from the *tgt* start codon by a spacer sequence and a sequence  
20 encoding a thrombin cleavage site. The pre-culture was added to 2 × 2 L main culture (LB medium  
21 including 100 mg × L<sup>-1</sup> ampicillin and 34 mg · L<sup>-1</sup> chloramphenicol) which is incubated at 37°C and 220  
22 rpm until the OD<sub>600</sub> = 0.7. This main culture was then cooled to 15°C and protein expression induced  
23 by addition of IPTG (final concentration = 1 mM). The main culture was then incubated at 15°C and 220  
24 rpm for a further 16-18 h. Afterwards, the cell pellets were harvested through centrifugation (10,000  
25 rpm at 4°C). The cell pellets were re-suspended in 100 mL lysis buffer (20 mM TRIS pH 7.8, 10 mM  
26 EDTA, 1mM DTT and 2 cComplete™-Protease Inhibitor Cocktail Tablets (Roche) per 4 L of bacterial  
27 culture) and cell disruption was achieved via 3 rounds of sonification using Branson Sonifier 250, with  
28 90 seconds intervals (Duty cycle 30%, output control 7). Alternatively, cell disruption was achieved via  
29 an EmulsiFlex-C5™ high-pressure homogenizer (Avestin Europe GmbH). The soluble protein in the  
30 supernatant was then separated by centrifugation from the insoluble cell constituents in the pellet  
31 (centrifugation speed 19,000 rpm, 45 min, 4°C).  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

45 Purification of the protein was achieved at room temperature by two FPLC steps using an ÄKTA Purifier  
46 LC system. A Q-Sepharose Fast Flow Anion Exchange Column (XK 26/15; GE Healthcare) was  
47 equilibrated with buffer A (10 mM TRIS pH 7.8, 1 mM EDTA, 1mM DTT) and the clear cell lysate was  
48 passed through the column. The protein was then eluted by buffer B (10 mM TRIS pH 7.8, 1 mM EDTA,  
49 1mM DTT, 1M NaCl) through gradient elution (a linear increase in the proportion of buffer B (from 0 -  
50 100% at 4 mL · min<sup>-1</sup> column flow) and fractionation of the eluted portion. Fractions containing  
51 the target protein *Z. mobilis* TGT with the Strep tag II were determined by SDS-PAGE. A Strep-Tactin<sup>®</sup>  
52 Superflow<sup>®</sup> column (XK 16/10, IBA) was equilibrated with buffer W (100 mM TRIS pH 7.8, 1MNaCl, 1  
53 mM EDTA) and the corresponding fractions were passed through the column. The target protein was  
54 eluted by buffer E (100 mM TRIS pH 7.8, 1MNaCl, 1 mM EDTA, 2.5 mM D-desthiobiotin).  
55  
56  
57  
58  
59  
60

The fractions containing TGT protein were then concentrated in a VIVASPIN<sup>®</sup>20 centrifugal concentrator (Sartorius, MWCO = 30,000) to a concentration of approximately 2 mg · mL<sup>-1</sup> in a high salt buffer (10 mM TRIS pH 7.8, 2MNaCl, 1mM EDTA). Subsequently, the Strep-tag<sup>®</sup> II was cleaved off and separated from the TGT protein via a Thrombin Cleavage Capture Kit (Novagen<sup>®</sup>) following the manufacturer's instructions where 2.5 U of biotinylated thrombin per mg TGT protein was incubated with the TGT protein for 16-18 h at 20°C. The cleaved TGT protein was separated from the Strep-tag<sup>®</sup> II, the biotinylated thrombin from the kit, and the streptavidin-agarose beads by filtration using the filters of the kit. The separated TGT protein was then dialyzed against high salt buffer and concentrated via VIVASPIN<sup>®</sup>20 centrifugal concentrator until a final protein concentration of 12 mg · mL<sup>-1</sup>. Finally, the protein was flash frozen into aliquots of 70 µL and stored at -80°C.

### Crystallization and Structure Determination of *Z. mobilis* TGT

**Crystallization.** *Z. mobilis* TGT co-crystals were grown at 18°C using the sitting-drop vapor diffusion method. A solution of 0.23 mM of *Z. mobilis* TGT (stored in high salt buffer) was mixed with 6.6 mM of the respective inhibitor (stock solution in 100% DMSO) and incubated at 18°C for 1 hour. The solution was then centrifuged and 1.5 µL of this mixture was mixed with 1.5 µL of reservoir solution (100 mM MES pH 5.5, 1 mM DTT, 13% (w/v) PEG8000, 10% (v/v) DMSO) in the wells of a crystallization plate containing 650 µL reservoir solution. Co-crystals could be seen within three days up to one week. They were transferred into cryo-buffer (50 mM MES pH 5.5, 0.5mM DTT, 300mM NaCl, 2% (v/v) DMSO, 4% (w/v) PEG 8000, 30% (v/v) glycerol) for several seconds and flash-frozen in liquid nitrogen.

**Data Collection and Structural Refinement.** Crystallographic data and data collection and refinement results are summarized in Tables S1 in the Supporting Information. The diffraction data were indexed, scaled, and merged using XDS<sup>[33]</sup> and XDSAPP.<sup>[34]</sup> Molecular replacement was performed with the program PHASER MR<sup>[35]</sup> from the CCP4 suite<sup>[36]</sup> to determine all crystal structures. The structure 4LBU was used as a search model. During refinement, a 5% subset of all reflections was omitted to be used for  $R_{\text{free}}$  calculation. Model building was achieved in COOT<sup>[37]</sup> and refinement using PHENIX.refine version 1.10.1-2155.<sup>[38]</sup> Cartesian simulated annealing with default parameters was used as a first refinement step for all the structures. This was followed by the refinement of XYZ coordinates and occupancies of protein residues and fragments (with the exception of water molecules whose occupancies were fixed). In the case of protein residues that gave additional density, they were refined in double conformation and kept if their refined occupancy was  $\geq 20\%$ . Chemicalize<sup>[39]</sup> developed by ChemAxon<sup>[40]</sup> was used for name to structure generation and SMILES code generation. The ligand PDB and restraint files were generated with the Grade Web Server.<sup>[41]</sup>

**Enzyme kinetic investigation.** The inhibition constants ( $K_i$ ) of the ligands with respect to the *Z. mobilis* TGT were determined according to the method described by Meyer et al. (2006) <sup>[19]</sup> with minor

1  
2  
3 modifications. The concentration of *Z. mobilis* TGT used depended on the expected ligand affinity and  
4 amounted to 9 – 150 nM. The concentration of tRNA<sup>Tyr</sup> was kept at 1.5 μM, while the concentration of  
5 [8-<sup>3</sup>H]-guanine used was a (saturating) concentration of 10 μM. The assay buffer consisted of 100 mM  
6 HEPES pH 7.3, 20 mM MgCl<sub>2</sub> and 0.037% (v/v) Tween 20. A total of four aliquots of 15 μL were  
7 withdrawn every 60 minutes from the sample and scintillation counting gave for the liquid scintillation  
8 counting. The final K<sub>i</sub> value reported is the average of three measurements. The preparation of *E. coli*  
9 tRNA<sup>Tyr</sup> (ECYC2) [42] was carried out via in vitro transcription using the RiboMAX™ Large Scale RNA  
10 Production System-T7 (Promega) following the manufacturer's instructions.  
11  
12

13  
14  
15 **MD Simulations.** We first performed MD simulations of the *apo* and preQ<sub>1</sub>-bound protein (using the  
16 crystal structures 1PUD and 1POE) to record the transient opening and closing of the binding pocket.  
17 Overall, 3 x 200 ns and 3 x 100 ns were simulated for the *apo* and the preQ<sub>1</sub>-bound protein,  
18 respectively. The applied simulation protocol is described in the Supporting Information (section S2.2  
19 and S2.3).  
20  
21  
22  
23  
24  
25

26 The hydration site analysis also afforded molecular dynamics simulations and we used the protein-  
27 ligand complex of TGT-4.1 (PDB code 6RKT) of this investigation with the following protocol. The  
28 TGT-4.1 structure was loaded into MOE<sup>[43]</sup> and the ligand molecule was removed. Then, missing atoms  
29 were built and protonation states were assigned using the protonate3D utility of MOE. The cysteine  
30 residues coordinating the Zn<sup>2+</sup> ion, were modelled in their deprotonated form. Then, the protonated  
31 protein structure, crystallographic water molecules and the zinc ion were loaded into *tLEaP* from the  
32 AmberTools17 package.<sup>[44]</sup> In *tLEaP*, parameters from the amber *FF99SB*<sup>[45]</sup> force field were assigned  
33 to the protein atoms and TIP4P-Ew<sup>[46]</sup> parameters were assigned to the water molecules. Two sodium  
34 counter ions were added using the *addions2* utility in order ensure net neutrality through the MD  
35 simulation. Finally, the protein was embedded into a truncated octahedron simulation box filled with  
36 TIP4P-Ew water molecules. The box was build such that the minimum distance between each solute  
37 or crystallographically determined water molecule and any box edge was at minimum 10 Å, resulting  
38 in a simulation box with 12756 water molecules in total.  
39  
40  
41  
42  
43  
44  
45  
46  
47

48 In the following, all minimization operations were carried out using *pmemd* and all MD simulations  
49 were calculated using the *pmemd.cuda*<sup>[47-49]</sup> for use with GPUs. Both programs were, as well as all  
50 other programs from the Amber program package, used from Amber16 together with  
51 AmberTools17.<sup>[44]</sup>  
52  
53  
54

55 Initially, the system energy was minimized with 2500 steps of steepest descent minimization followed  
56 by 2500 steps of conjugate gradient minimization. During this minimization operation, all solute heavy  
57 atoms were restrained to their crystallographically determined positions using a harmonic restraining  
58 potential with a force constant of 25 kcal·mol<sup>-1</sup>·Å<sup>-2</sup>. Subsequently, the force constant was set to  
59  
60

2 kcal·mol<sup>-1</sup>·Å<sup>-2</sup> and the minimization was repeated analogously to the first one. In the next step, the system was heat to 300 K within 25 ps using an integration time step of 1 fs. At this point, the solute heavy atoms are again positionally restrained with a force constant of 25 kcal·mol<sup>-1</sup>·Å<sup>-2</sup>. Now, the integration time step was switched to 2 fs and the system equilibrated under NPT conditions for 5 ns using the Berendsen barostat<sup>[50]</sup> and a target pressure of 1 bar. The temperature was regulated at 300 K using a Langevin dynamics thermostat with a collision frequency value of  $\gamma = 2 \text{ ps}^{-1}$ . In the final equilibration step, the system is run under NVT conditions for an additional 5 ns.

A final production run was carried out for 30 ns and coordinates were saved to disk every 2 ps.

During all MD runs, the system was treated using periodic boundary conditions and the particle-mesh Ewald technique together with a 9.0 Å real-space distance cutoff for the electrostatic interactions. Furthermore, all bonds involving hydrogen atoms were constrained using the SHAKE algorithm.<sup>[50]</sup>

**Hydration Site Calculations.** For the HSA calculations, we used the program SSTMap<sup>[51,52]</sup> (Version 1.1.1). We used ligand molecule TGT-4.1 in its bound pose in order to define the binding site for the HSA analysis. After the HSA calculation, the hydration site energy values were manually referenced to the TIP4P-Ew bulk energy value.

### Synthesis of the Ligands

For SMILES codes, general methods and <sup>1</sup>H and <sup>13</sup>C NMR spectra, see Supporting Information.

**Chemical Synthesis.** Materials and suppliers, the analytical instrumentation used, and the chromatographic equipment are described in the General Methods in the Supporting Information. The main manuscript contains the General Procedures, the complete synthesis of carboxamide **2.1**, hydrazide **2.5**, guanidine **3.1**, and dihydro-imidazoquinazoline ( $\pm$ )-**4.2** and the characterization of intermediates and target products. The complete synthesis of all other target compounds is reported in the Supporting Information. The purity of the compounds was achieved by chromatography and was found to be above 95% as confirmed by HPLC and elementary analysis.

**General Procedure 1 (GP1).** The tautomericly enriched ester (1.0 equiv) was suspended in 25% aq. ammonia (6–18 mL), and the mixture was treated in a microwave at 100 °C for 15–30 min. An aq. sat. NaCl solution (10 mL) was added, and the aq. layer was extracted with EtOAc or CH<sub>2</sub>Cl<sub>2</sub> (4x 20 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was evaporated. FC yielded the desired amide.

**General Procedure 2 (GP2).** A solution of tautomericly enriched ester (1.0 equiv) in 1 M methanolic MeNH<sub>2</sub> (2.4–12 mL) was treated in a microwave at 140 °C for 1.5–2.5 h. The solvent was evaporated, the solid was dissolved in sat. NaHCO<sub>3</sub> (20 mL) and CH<sub>2</sub>Cl<sub>2</sub> (20 mL), and the aq. layer was extracted with

1  
2  
3 CH<sub>2</sub>Cl<sub>2</sub> (3x 20 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was  
4 evaporated. FC yielded the desired secondary amide.  
5  
6

7 **General Procedure 3 (GP3).** The tautomericly enriched ester (1.0 equiv) was suspended in H<sub>2</sub>O (6 mL)  
8 and hydrazine monohydrate (60–65%, 6 mL) or hydrazine hydrate (50–60%, 6 mL). The mixture was  
9 treated in the microwave at 100 °C for 0.5–1 h. A white solid was filtered off, and the aq. layer was  
10 extracted with CH<sub>2</sub>Cl<sub>2</sub> (4x 20 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent  
11 was evaporated. FC yielded the desired hydrazide.  
12  
13  
14

15  
16 **General Procedure 4 (GP4).** A suspension of tautomericly enriched primary amide (1.0 eq),  
17 secondary amide (1.0 equiv), or hydrazide (1.0 equiv) in 4 M HCl in 1,4-dioxane (4–5 mL) was heated  
18 to reflux for 16–22 h. The solvent was evaporated. FC and/or reversed-phase FC followed by  
19 lyophilization yielded the desired product.  
20  
21  
22

23  
24 **General Procedure 5 (GP5).** A solution of the aminoaniline (1.0 equiv) in DMF was treated with the  
25 thioisocyanate (1.1–2.4 equiv) at 90 °C and stirred at this temperature for 1–2 h (monitored by TLC).  
26 After addition of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (1.2–1.3 equiv), the solution  
27 was stirred at this temperature for 1–2 h (monitored by LC/MS). The mixture was cooled down to 23 °C,  
28 filtered, and washed with EtOAc (3 x). Evaporation of the solvent, and FC gave the desired 2-  
29 aminobenzimidazole.  
30  
31  
32

33  
34 **General Procedure 6 (GP6).** A suspension of the nitrobenzimidazole (1.0 equiv), Pd/C (10% w/w), and  
35 PtO<sub>2</sub> (10% w/w) in EtOH was stirred under hydrogen atmosphere at 23 °C for 2–4 h (monitored by TLC).  
36 The mixture was filtered over Celite® and washed with CH<sub>2</sub>Cl<sub>2</sub>/MeOH. The solvent was evaporated,  
37 and the crude amine was directly dissolved in dry DMF. *N,N'*-Di-Boc-1H-pyrazole-1-carboximidine  
38 (1.2 eq) and *N,N*-diisopropylethylamine (1.5 eq) were added, and the mixture was stirred at 23 °C for  
39 16 h (monitored by LC/MS). The solution was diluted with H<sub>2</sub>O and EtOAc, and the organic layer was  
40 washed with H<sub>2</sub>O. The organic solvent was evaporated, and the desired protected guanidine was  
41 obtained after FC.  
42  
43  
44  
45  
46  
47

48  
49 **General Procedure 7 (GP7).** The protected guanidine (1.0 eq) was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (2 mL) and 4 M  
50 ethereal HCl (1 mL) was added. The solution was stirred at 23 °C for 10 min. The solvent was  
51 evaporated, and the solid was recrystallized from Et<sub>2</sub>O/MeOH.  
52  
53

54 **Methyl 1H-Benzimidazole-5-carboxylate (7).** A solution of 5-benzimidazole carboxylic acid (**6**) (19.3 g,  
55 119 mmol) in anhydrous MeOH (500 mL) was treated dropwise with conc. H<sub>2</sub>SO<sub>4</sub> (10 mL), and the  
56 mixture was stirred for 18 h at 65 °C. The solvent was evaporated, and the residue was suspended in  
57 EtOAc (50 mL). A sat. aq. NaHCO<sub>3</sub> solution was added until pH ≈ 8 was reached. The aq. layer was  
58 extracted with EtOAc (3x 250 mL), the combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, and the  
59 solvent was evaporated. **7** (21.0 g, assumed quant.) was obtained as a brown crystalline solid. The  
60 product was used in the next reaction without further purification. M.p. 135 °C; <sup>1</sup>H NMR (400 MHz,

(CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 3.86 (s, 3H; COOCH<sub>3</sub>), 7.67 (s, 1H; H-C(6)), 7.84 (d,  $J$  = 8.5 Hz, 1H; H-C(7)), 8.22 (s, 1 H; H-C(4)), 8.40 (s, 1 H; H-C(2)), 12.79 ppm (s, 1H; NH); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO; broad peaks due to tautomeric N=C-NH):  $\delta$  = 52.0 (COOCH<sub>3</sub>), 111.7 (C(7)), 120.8 (C(4)), 123.3 (C(6)), 132.8 (C(5)), 136.8 (C(7a)), 142.6 (C(3a)), 144.3 (C(2)), 166.8 ppm (COOCH<sub>3</sub>); IR (ATR):  $\tilde{\nu}$  = 3095 (w), 2811 (w), 1711 (s), 1625 (m), 1434 (m), 1412 (m), 1300 (s), 1223 (s), 1198 (s), 1128 (m), 1080 (s), 954 (m), 743 cm<sup>-1</sup> (s); HR-ESI-MS:  $m/z$  (%): 177.0674 (100, [M + H]<sup>+</sup> calcd. for C<sub>9</sub>H<sub>9</sub>N<sub>2</sub>O<sub>2</sub><sup>+</sup>: 177.0659).

*Methyl 1-(N,N-Dimethylsulfamoyl)benzimidazole-5-carboxylate (8) and Methyl 1-(N,N-Dimethylsulfamoyl)benzimidazole-6-carboxylate (8')*. A solution of ester **7** (20.6 g, 117 mmol) in toluene (300 mL) was treated with NEt<sub>3</sub> (18.5 mL, 131 mmol) and Me<sub>2</sub>NSO<sub>2</sub>Cl (14.2 mL, 131 mmol) at 23 °C. The solution was stirred for additional 24 h at 23 °C. After evaporation of the solvent, FC (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/EtOAc 4:1) yielded a mixture of both regioisomers **8** and **8'** (21.8 g, 66%; 1:0.6 ratio) as a white crystalline solid.  $R_f$  = 0.56 (SiO<sub>2</sub>; EtOAc/CH<sub>2</sub>Cl<sub>2</sub> 4:1); m.p. 118 °C; for **8**: <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.91 (s, 6H; N(CH<sub>3</sub>)<sub>2</sub>), 3.89 (s, 3H; COOCH<sub>3</sub>), 7.97 (dd,  $J$  = 8.1, 0.6 Hz, 1H; H-C(7)), 8.04 (dd,  $J$  = 8.7, 1.6 Hz, 1H; H-C(6)), 8.33 (dd,  $J$  = 1.6, 0.6 Hz, 1H; H-C(4)), 8.79 ppm (s, 1H; H-C(2)); for **8'** <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.91 (s, 4H; N(CH<sub>3</sub>)<sub>2</sub>), 3.90 (s, 1.8H; COOCH<sub>3</sub>), 7.92 (dd,  $J$  = 8.5, 0.6 Hz, 0.6H; H-C(7)), 8.00 (dd,  $J$  = 6.9, 1.6 Hz, 0.7H; H-C(6)), 8.41 (dd,  $J$  = 1.6, 0.7 Hz, 0.6H; H-C(4)), 8.87 ppm (s, 0.6H; H-C(2)); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 37.86, 37.87, 52.3, 52.4, 113.1, 114.0, 120.6, 121.7, 125.1, 125.8, 126.0, 126.3, 131.4, 134.7, 142.8, 144.4, 145.4, 146.2, 165.98, 166.02 ppm; IR (ATR):  $\tilde{\nu}$  = 3130 (w), 2953 (w), 1709 (s), 1618 (m), 1493 (m), 1430 (m), 1386 (s), 1317 (m), 1281 (s), 1268 (s), 1239 (m), 1166 (s), 1147 (s), 960 (s), 732 cm<sup>-1</sup> (s); HR-ESI-MS:  $m/z$  (%): 284.0704 (100, [M + H]<sup>+</sup> calcd. for C<sub>11</sub>H<sub>14</sub>N<sub>3</sub>O<sub>4</sub>S<sup>+</sup>: 284.0700).

*Methyl 2-Bromo-1-(N,N-dimethylsulfamoyl)benzimidazole-5-carboxylate (9)*. A solution of **8** and **8'** (21.4 g, 75.5 mmol) in anhydrous THF (500 mL) was treated dropwise with a solution of LiN(SiMe<sub>3</sub>)<sub>2</sub> in THF (1 M, 112 mL, 112 mmol) at -70 °C under a N<sub>2</sub> atmosphere. The mixture was stirred at -70 °C for 1 h until a solution of NBS (20.3 g, 114 mmol) in THF (250 mL) was added dropwise. The solution was stirred 15 min at -70 °C, before it was allowed to warm to 23 °C. The solvent was evaporated, and the residue was dissolved in H<sub>2</sub>O (200 mL) and CH<sub>2</sub>Cl<sub>2</sub> (200 mL) and the aq. layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3x 200 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, and evaporated. FC (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/EtOAc 10:0 to 10:1) afforded **9** (15.3 g, 56%) as a pale yellow solid with traces of **9'**.  $R_f$  = 0.76 (SiO<sub>2</sub>; EtOAc/CH<sub>2</sub>Cl<sub>2</sub> 4:1); m.p. 151 °C; <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 3.01 (s, 6H; N(CH<sub>3</sub>)<sub>2</sub>), 3.89 (s, 3H; COOCH<sub>3</sub>), 8.01 (2 s, 2H; H-C(7), H-C(6)), 8.23 ppm (dd,  $J$  = 1.2, 1.2 Hz, 1H; H-C(4)); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO; C(2) not visible):  $\delta$  = 38.2 (2C; N(CH<sub>3</sub>)<sub>2</sub>), 52.4 (COOCH<sub>3</sub>), 114.3 (C(7)), 120.3 (C(4)), 126.0 (C(6)), 128.8, 137.9 (C(7a)), 141.5, 165.8 ppm (COOCH<sub>3</sub>); IR (ATR):  $\tilde{\nu}$  = 2954 (w), 1708 (s), 1464 (m), 1443 (m), 1418 (m), 1395 (m), 1300 (s), 1197 (s), 1165 (m), 1132 (m), 1028 (m), 968 (s), 719 cm<sup>-1</sup> (s); HR-ESI-MS:  $m/z$  (%): 363.9788 (44, [M + H]<sup>+</sup> calcd. for C<sub>11</sub>H<sub>13</sub>N<sub>3</sub>O<sub>4</sub>S<sup>81</sup>Br<sup>+</sup>: 363.97), 361.9806 (40, [M + H]<sup>+</sup> calcd. for C<sub>11</sub>H<sub>13</sub>N<sub>3</sub>O<sub>4</sub>S<sup>79</sup>Br<sup>+</sup>: 361.9805).

*Methyl 1-(N,N-Dimethylsulfamoyl)-2-(methylamino)benzimidazole-5-carboxylate (10.1)*. A solution of tautomericly enriched bromide **9** (3.00 g, 8.28 mmol) in EtOAc/MeOH (1:1; 120 mL) was treated with 1 M methanolic MeNH<sub>2</sub> (12.4 mL, 124 mmol), and the mixture was stirred at 23 °C for 14 h. The solvents were evaporated, and the residue was dissolved in a mixture of H<sub>2</sub>O (150 mL) and CH<sub>2</sub>Cl<sub>2</sub> (150 mL). The aq. layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (2x 150 mL). The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, and the solvent was evaporated to afford **10.1** (2.76 g, assumed quant.) as a crystalline white solid with traces of **10.1'**.  $R_f$  = 0.61 (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:1); m.p. 160 °C; <sup>1</sup>H NMR (400 MHz,

(CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.88 (s, 6H; N(CH<sub>3</sub>)<sub>2</sub>), 2.99 (d,  $J$  = 4.7 Hz, 3H; CH<sub>3</sub>NH), 3.84 (s, 3H; COOCH<sub>3</sub>), 6.87 (q,  $J$  = 4.6 Hz, 1H; CH<sub>3</sub>NH), 7.63 (dd,  $J$  = 8.4, 0.6 Hz, 1H; H-C(7)), 7.68 (dd,  $J$  = 8.4, 1.6 Hz, 1H; H-C(6)), 7.79 ppm (d,  $J$  = 1.1 Hz, 1H; H-C(4)); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 30.0 (CH<sub>3</sub>NH), 38.2 (2C; N(CH<sub>3</sub>)<sub>2</sub>), 52.0 (COOCH<sub>3</sub>), 111.8 (C(7)), 116.5 (C(4)), 121.8 (C(6)), 125.3, 135.5 (C(7a)), 142.1, 154.0 (C(2)), 166.4 ppm (COOCH<sub>3</sub>); IR (ATR):  $\tilde{\nu}$  = 3355 (m), 2960 (w), 1712 (s), 1626 (s), 1582 (m), 1435 (m), 1409 (m), 1379 (s), 1294 (s), 1216 (m), 1163 (m), 1144 (m), 1123 (s), 1085 (m), 1025 (m), 968 (m), 884 (m), 758 cm<sup>-1</sup> (s); HR-ESI-MS:  $m/z$  (%): 313.0972 (100, [M + H]<sup>+</sup>, calcd. for C<sub>12</sub>H<sub>17</sub>N<sub>4</sub>O<sub>4</sub>S<sup>+</sup>: 313.0965).

*1-(N,N-Dimethylsulfamoyl)-2-(methylamino)benzimidazole-5-carboxamide (11.1)*. A suspension of tautomericly enriched ester **10.1** (501 mg, 1.60 mmol) was treated in 25% aq. ammonia (18 mL) according to **GP1** at 100 °C for 15 min. FC (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:0 to 9:1) afforded **11.1** (154 mg, 37%) as a white crystalline solid with traces of **11.1'**. Unreacted **10.1** (59 mg, 12%) was reisolated.  $R_f$  = 0.34 (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:1); m.p. 175 °C; <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.87 (s, 6H; N(CH<sub>3</sub>)<sub>2</sub>), 2.98 (d,  $J$  = 4.7 Hz, 3H; CH<sub>3</sub>NH), 6.80 (q,  $J$  = 4.6 Hz, 1H; CH<sub>3</sub>NH), 7.26 (s, 1H; CONHH), 7.55 (dd,  $J$  = 8.4, 0.6 Hz, 1H; H-C(7)), 7.58 (dd,  $J$  = 8.4, 1.6 Hz, 1H; H-C(6)), 7.80 (dd,  $J$  = 1.6, 0.8 Hz, 1H; H-C(4)), 7.89 ppm (s, 1H; CONHH); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 30.0 (CH<sub>3</sub>NH), 38.2 (2C; N(CH<sub>3</sub>)<sub>2</sub>), 111.3 (C(7)), 115.1 (C(4)), 120.3 (C(6)), 130.2, 133.9 (C(7a)), 141.8, 153.7 (C(2)), 168.0 ppm (CONH<sub>2</sub>); IR (ATR):  $\tilde{\nu}$  = 3418 (m), 1629 (s), 1576 (s), 1434 (m), 1413 (m), 1376 (s), 1260 (m), 1224 (w), 1155 (m), 1027 (m), 936 (m), 830 (w), 715 cm<sup>-1</sup> (m); HR-ESI-MS:  $m/z$  (%): 298.0973 (100, [M + H]<sup>+</sup>, calcd. for C<sub>11</sub>H<sub>16</sub>N<sub>5</sub>O<sub>3</sub>S<sup>+</sup>: 298.0968).

*1-(N,N-Dimethylsulfamoyl)-2-(methylamino)benzimidazole-5-hydrazide (11.5)*. A suspension of tautomericly enriched ester **10.1** (505 mg, 1.62 mmol) in 50–60% hydrazine hydrate (6 mL) was treated according to **GP3** for 0.5 h. The organic layer was evaporated to obtain **11.5** (403 mg, 80%) as a white crystalline powder with traces of **11.5'** and was used without further purification.  $R_f$  = 0.41 (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:1); m.p. 190 °C; <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.87 (s, 6H; N(CH<sub>3</sub>)<sub>2</sub>), 2.98 (d,  $J$  = 4.6 Hz, 3H; CH<sub>3</sub>NH), 4.46 (s, 2H; CONHNH<sub>2</sub>), 6.80 (q,  $J$  = 4.6 Hz, 1H; CH<sub>3</sub>NH), 7.53 (dd,  $J$  = 8.4, 1.6 Hz, 1H; H-C(6)), 7.56 (dd,  $J$  = 8.4, 0.7 Hz, 1H; H-C(7)), 7.72 (dd,  $J$  = 1.6, 0.7 Hz, 1H; H-C(4)), 9.68 ppm (s, 1H; CONHNH<sub>2</sub>); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 30.0 (CH<sub>3</sub>NH), 38.2 (2C; N(CH<sub>3</sub>)<sub>2</sub>), 111.5 (C(7)), 114.4 (C(4)), 119.8 (C(6)), 129.2, 133.8 (C(7a)), 141.9, 153.7 (C(2)), 166.1 ppm (CONHNH<sub>2</sub>); IR (ATR):  $\tilde{\nu}$  = 3324 (m), 2911 (w), 1657 (s), 1621 (s), 1592 (s), 1577 (s), 1493 (m), 1458 (m), 1408 (m), 1376 (s), 1308 (m), 1253 (m), 1160 (s), 1122 (s), 1051 (w), 1025 (s), 967 (s), 867 (m), 715 cm<sup>-1</sup> (s); HR-ESI-MS:  $m/z$  (%): 313.1078 (100, [M + H]<sup>+</sup>, calcd. for C<sub>11</sub>H<sub>17</sub>N<sub>6</sub>O<sub>3</sub>S<sup>+</sup>: 313.1077).

*2-(Methylamino)benzimidazole-6-carboxamide (2.1)*. A suspension of tautomericly enriched **11.1** (96 mg, 320  $\mu$ mol) in a solution of 4 M HCl in 1,4-dioxane (5 mL) was treated according to **GP4** for 45 h. FC (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/7 M NH<sub>3</sub> in MeOH 10:1) followed by reversed-phase FC (C18; H<sub>2</sub>O/THF 10:0 to 10:1) yielded **2.1** (14 mg, 23%) as a white powder.  $R_f$  = 0.06 (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/MeOH/NH<sub>4</sub>OH 10:1:0.1); m.p. 261 °C; <sup>1</sup>H NMR (600 MHz, (CD<sub>3</sub>)<sub>2</sub>SO/CF<sub>3</sub>COOD 5:1; four exchanging NH protons not visible):  $\delta$  = 3.00 (s, 3H; CH<sub>3</sub>NH), 7.37 (dd,  $J$  = 8.4, 0.6 Hz, 1H; H-C(4)), 7.79 (dd,  $J$  = 8.4, 1.6 Hz, 1H; H-C(5)), 7.84 ppm (dd,  $J$  = 1.6, 0.6 Hz, 1H; H-C(7)); <sup>13</sup>C NMR (151 MHz, (CD<sub>3</sub>)<sub>2</sub>SO/CF<sub>3</sub>COOD 5:1):  $\delta$  = 29.4 (CH<sub>3</sub>NH), 110.9 (C(4)), 111.0 (C(7)), 123.4 (C(5)), 129.8, 130.1, 132.5 (C(3a)), 152.0 (C(2)), 167.8 ppm (CONH<sub>2</sub>); IR (ATR):  $\tilde{\nu}$  = 3177 (m), 1638 (s), 1587 (s), 1388 (s), 1283 (s), 1104 (w), 991 (w), 881 (w), 772 cm<sup>-1</sup> (w); HR-ESI-MS:  $m/z$  (%): 191.0942 (100, [M + H]<sup>+</sup>, calcd. for C<sub>9</sub>H<sub>11</sub>N<sub>4</sub>O<sup>+</sup>: 191.0855); elemental analysis calcd (%) for C<sub>9</sub>H<sub>10</sub>N<sub>4</sub>O·0.7H<sub>2</sub>O (202.82): C 53.30, H 5.67, N 27.62; found: C 53.36, H 5.83, N 27.79.

1  
2  
3 *2-(Methylamino)benzimidazole-6-hydrazide (2.5)*. A suspension of tautomerically enriched **11.5**  
4 (101 mg, 340  $\mu$ mol) in a solution of 4 M HCl in 1,4-dioxane (4 mL) was treated according to **GP4** for  
5 16 h. FC (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/7 M NH<sub>3</sub> in MeOH 10:1) followed by reversed-phase FC (C18; H<sub>2</sub>O/THF 10:0 to  
6 10:1) yielded **2.5** (6 mg, 9%) as a white powder. *R*<sub>f</sub> = 0.11 (SiO<sub>2</sub>; CH<sub>2</sub>Cl<sub>2</sub>/MeOH/NH<sub>4</sub>OH 10:1:0.1); m.p.  
7 276 °C; <sup>1</sup>H NMR (600 MHz, (CD<sub>3</sub>)<sub>2</sub>SO/CF<sub>3</sub>COOD 5:1; three exchanging NH protons not visible):  $\delta$  = 3.03–  
8 3.05 (m, 3H; CH<sub>3</sub>NH), 7.51 (dd, *J* = 8.3, 1.4 Hz, 1H; H–C(4)), 7.79 (dd, *J* = 8.4, 1.6 Hz, 1H; H–C(5)), 7.85  
9 (d, *J* = 1.6 Hz, 1H; H–C(7)), 9.22 (q, *J* = 4.8 Hz, 0.9H; CH<sub>3</sub>NH), 11.53 ppm (s, 0.7H; NH); <sup>13</sup>C NMR  
10 ((CD<sub>3</sub>)<sub>2</sub>SO/CF<sub>3</sub>COOD 5:1):  $\delta$  = 29.3 (CH<sub>3</sub>NH), 110.6 (C(4)), 111.2 (C(7)), 123.2 (C(6)), 125.1 (C(3a)), 130.1  
11 (C(5)), 133.5 (C(7a)), 152.2 (C(2)), 165.7 ppm (CONHNH<sub>2</sub>); IR (ATR):  $\tilde{\nu}$  = 3124 (m), 1602 (s), 1519 (m),  
12 1464 (m), 1420 (m), 1283 (s), 1153 (m), 992 (m), 816 cm<sup>-1</sup> (w); HR-ESI-MS: *m/z* (%): 206.1041 (25,  
13 [M + H]<sup>+</sup>, calcd. for C<sub>9</sub>H<sub>12</sub>N<sub>5</sub>O<sup>+</sup>: 206.1036); elemental analysis calcd. (%) for C<sub>9</sub>H<sub>11</sub>N<sub>5</sub>O·1.3H<sub>2</sub>O (228.64):  
14 C 47.28, H 6.00, N 30.63; found: C 47.23, H 6.23, N 30.56.  
15  
16  
17  
18  
19

20 *2-(Methylamino)-5-nitrobenzimidazole (13.1)*. A solution of 4-nitro-1,2-diaminobenzene (**12**) (1.50 g,  
21 9.80 mmol) in dry DMF (10 mL) was treated with 2-(4-morpholino)ethyl isothiocyanate (788 mg,  
22 10.8 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (2.3 mL, 13 mmol) according to **GP5**.  
23 FC (SiO<sub>2</sub>; EtOAc/MeOH 100:0 to 90:10) yielded the desired benzimidazole **13.1** (1.30 g, 69%) as an  
24 orange solid. *R*<sub>f</sub> = 0.21 (SiO<sub>2</sub>; EtOAc/MeOH 97:3); m.p. 192–193 °C; <sup>1</sup>H NMR (400 MHz,  
25 (CD<sub>3</sub>)<sub>2</sub>SO/CF<sub>3</sub>COOD):  $\delta$  = 3.04 (s, 3H; CH<sub>3</sub>NH), 7.53 (d, *J* = 8.6 Hz, 1H; H–C(7)), 8.11–8.14 (m, 2H; H–C(4),  
26 H–C(6)), 9.47 ppm (q, *J* = 5.0 Hz, 0.4H; CH<sub>3</sub>NH); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO/CF<sub>3</sub>COOD):  $\delta$  = 29.3 (CH<sub>3</sub>),  
27 107.0 (C(4) or C(6)), 111.3 (C(7)), 119.5 (C(4) or C(6)), 139.3, 135.4, 143.1, 153.0, 153.1 ppm (C(2)); IR  
28 (ATR):  $\tilde{\nu}$  = 3294 (w), 3127 (w), 2875 (m), 1682 (m), 1615 (m), 1587 (m), 1544 (w), 1491 (m), 1476 (m),  
29 1420 (w), 1366 (w), 1315 (m), 1263 (s), 1223 (s), 1152 (m), 1119 (m), 1064 (s), 978 (m), 944 (m), 878  
30 (m), 856 (w), 808 (m), 757 (m), 728 (m), 701 (w), 688 (m), 622 cm<sup>-1</sup> (w); HR-ESI-MS: *m/z* (%): 193.0725  
31 (100, [M + H]<sup>+</sup>, calcd. for C<sub>8</sub>H<sub>9</sub>N<sub>4</sub>O<sub>2</sub><sup>+</sup>: 193.0720).  
32  
33  
34  
35  
36

37 *N,N'*-Bis(*tert*-butoxycarbonyl)-*N''*-[(2-methylamino)benzimidaz-5-yl]guanidine (**14.1**). A solution of  
38 **13.1** (200 mg, 1.04 mmol) in EtOH (10 mL) was treated with Pd/C (10% w/w, 20 mg) and PtO<sub>2</sub> (10%  
39 w/w, 20 mg) under a H<sub>2</sub> atmosphere, followed by *N,N*-diisopropylethylamine (260  $\mu$ l, 1.6 mmol) and  
40 *N,N'*-di-Boc-1H-pyrazole-1-carboxamidine (388 mg, 1.25 mmol) in anhydrous DMF (10 mL) according  
41 to **GP6**. FC (SiO<sub>2</sub>; EtOAc/MeOH/NH<sub>4</sub>OH 100:0:0 to 89:10:1) yielded the desired protected guanidine  
42 **14.1** (230 mg, 55%) as a white solid. *R*<sub>f</sub> = 0.15 (SiO<sub>2</sub>; EtOAc/MeOH/NH<sub>4</sub>OH 89:10:1); m.p. > 280 °C  
43 (decomp.); <sup>1</sup>H NMR (400 MHz, CD<sub>2</sub>Cl<sub>2</sub>; two exchanging NH not visible):  $\delta$  = 1.36 (s, 9H; C(CH<sub>3</sub>)<sub>3</sub>), 1.57 (s,  
44 9H; C(CH<sub>3</sub>)<sub>3</sub>), 2.89 (s, 3H; CH<sub>3</sub>NH), 6.85 (dd, *J* = 8.2, 2.0 Hz, 1H; H–C(6)), 6.94 (d, *J* = 2.0 Hz, 1H; H–C(4)),  
45 7.08 (d, *J* = 8.2 Hz, 1H; H–C(7)), 10.03 (br. s, 0.8H; NH), 11.71 ppm (br. s, 0.8H; NH); <sup>13</sup>C NMR (101 MHz,  
46 CD<sub>2</sub>Cl<sub>2</sub>; one signal are hidden by noise):  $\delta$  = 28.36 (3C; C(CH<sub>3</sub>)<sub>3</sub>), 28.37 (3C; C(CH<sub>3</sub>)<sub>3</sub>), 29.3 (CH<sub>3</sub>NH), 80.4  
47 (C(CH<sub>3</sub>)<sub>3</sub>), 84.5 (C(CH<sub>3</sub>)<sub>3</sub>), 109.7 (C(4)), 112.9 (C(7)), 119.4 (C(6)), 135.0, 138.7 153.7, 156.8, 157.8,  
48 163.6 ppm; IR (ATR):  $\tilde{\nu}$  = 3231 (w), 3044 (m) 2919 (m), 2786 (m) 2664 (m), 2573 (m), 1683 (s), 1582 (w),  
49 1543 (w), 1506 (m), 1484 (m), 1441 (m), 1355 (m), 1307 (m), 1283 (w), 1245 (m), 1226 (m), 1150 (m),  
50 1122 (w) 1094 (m), 988 (m), 954 (w), 869 (m), 852 (w), 816 (m), 780 (w), 740 (m), 678 (m), 620 cm<sup>-1</sup>  
51 (w); HR-MALDI-MS: *m/z* (%): 405.2245 (32, [M + H]<sup>+</sup>, calcd. for C<sub>19</sub>H<sub>29</sub>N<sub>6</sub>O<sub>4</sub><sup>+</sup>: 405.2245), 293.0998 (100,  
52 [M – C<sub>8</sub>H<sub>16</sub>]<sup>+</sup>, calcd. for C<sub>11</sub>H<sub>13</sub>N<sub>6</sub>O<sub>4</sub><sup>+</sup>: 293.1022).  
53  
54  
55  
56  
57  
58

59 *N*-[(2-Methylamino)-benzimidaz-5-yl]guanidium chloride (**3.1**). A solution of **14.1** (30 mg, 74  $\mu$ mol) in  
60 CH<sub>2</sub>Cl<sub>2</sub> (2 mL) was treated with 4 M ethereal HCl (1 mL) according to **GP7**. Recrystallization from  
MeOH/Et<sub>2</sub>O afforded **3.1**·2HCl (12 mg, 59%) as white solid. M.p. > 250 °C (decomp.); <sup>1</sup>H NMR (600 MHz,

(CD<sub>3</sub>)<sub>2</sub>SO; two exchanging NH protons not visible):  $\delta$  = 3.03 (d,  $J$  = 4.8 Hz, 3H; CH<sub>3</sub>NH), 7.07 (dd,  $J$  = 8.4, 2.0 Hz, 1H; H-C(6)), 7.25 (d,  $J$  = 2.0 Hz, 1 H; H-C(4)), 7.41 (dd,  $J$  = 8.4, 0.5 Hz, 1H; H-C(7)), 7.50 (br. s, 4H; C(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>), 9.17 (s, 1H; NH), 10.07 ppm (s, 1H; NH); <sup>13</sup>C NMR (151 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 29.5 (CH<sub>3</sub>NH), 108.8 (C(4)), 111.9 (C(7)), 120.6 (C(6)), 128.8 (C(7a)), 130.1 and 130.7 (2C; C(3a), C(5)), 151.4 (C(NH<sub>2</sub><sup>+</sup>)NH<sub>2</sub>), 156.6 ppm (C(2)); IR (ATR):  $\tilde{\nu}$  = 2933 (br. m), 1662 (s), 1627 (s), 1598 (s), 1507 (m), 1471 (m), 1262 (m), 1187 (w), 1133 (m), 1095 (m), 1015 (w), 906 (w), 870 (w), 806 (w), 692 cm<sup>-1</sup> (w); HR-MALDI-MS:  $m/z$  (%): 205.1196 (51, [M + H]<sup>+</sup>, calcd. for C<sub>9</sub>H<sub>13</sub>N<sub>6</sub><sup>+</sup>: 205.1196).

*N,N'*-(4-Bromo-1,2-phenylene)bis(4-methylbenzenesulfonamide) (**19a**). A solution of 4-bromobenzene-1,2-diamine (**19**) (25.4 g, 136 mmol) in pyridine (30 mL) was treated dropwise with a solution of TsCl (54.7 g, 271 mmol) in pyridine (60 mL) over 30 min at 23 °C. After stirring the mixture for 17 h at 65 °C, the reaction mixture was poured into a mixture of ice-water (400 mL) and aq. HCl (37 wt%, 100 mL). The resulting solid was filtered and washed with water. The crude product was suspended in a mixture of acetic acid (630 mL) and water (70 mL) and was heated to reflux, and after filtration **19a** (56.2 g, 83%) was collected as a white solid.  $R_f$  = 0.54 (SiO<sub>2</sub>; cy-hexane:EtOAc 50:50); <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.36 (s, 3H; CH<sub>3</sub>), 2.37 (s, 3H; CH<sub>3</sub>), 6.93 (d,  $J$  = 8.7 Hz, 1H; H-C(6)), 7.13 (d,  $J$  = 2.3 Hz, 1H; H-C(3)), 7.19 (dd,  $J$  = 8.7, 2.3 Hz, 1H; H-C(5)), 7.37 (dd,  $J$  = 8.4, 8.4 Hz, 4H; H-C(3')), 7.60 (dd,  $J$  = 8.4, 6.8 Hz, 4H; H-C(2')), 9.41 (br. s, 1H; NH), 9.48 ppm (br. s, 1H; NH); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 21.0 (2C; CH<sub>3</sub>), 117.4 (C(4)), 125.0 (2C; C(3), C(6)), 126.89 (2C; C(2')), 126.91 (2C; C(2')), 128.3 (C(5)), 128.6 (C(2)), 129.82 (2C; C(3')), 129.87 (2C; C(3')), 131.3 (C(1)), 135.8 (C(1')), 135.9 (C(1')), 143.8 (C(4')), 144.0 ppm (C(4')); IR (ATR):  $\tilde{\nu}$  = 3289 (w), 3241 (w), 2922 (w), 1910 (w), 1591 (w), 1583 (w), 1492 (m), 1432 (w), 1389 (m), 1336 (s), 1311 (w), 1293 (w), 1260 (m), 1184 (m), 1161 (s), 1120 (m), 1092 (s), 1024 (w), 955 (w), 930 (m), 894 (w), 873 (m), 856 (m), 810 (m), 721 (m), 704 (w), 666 (s), 629 cm<sup>-1</sup> (m); HR-EI-MS:  $m/z$  (%): 497.0022 (72, [M + H]<sup>+</sup>, calcd. for C<sub>20</sub>H<sub>20</sub><sup>81</sup>BrN<sub>2</sub>O<sub>4</sub>S<sub>2</sub><sup>+</sup>: 497.0023), 495.0041 (64, [M + H]<sup>+</sup>, calcd. for C<sub>20</sub>H<sub>20</sub><sup>79</sup>BrN<sub>2</sub>O<sub>4</sub>S<sub>2</sub><sup>+</sup>: 495.0042), 369.9684 (100, calcd. for C<sub>13</sub>H<sub>11</sub><sup>81</sup>BrN<sub>3</sub>O<sub>3</sub>S<sup>+</sup>: 369.9679), 367.9703 (100, calcd. for C<sub>13</sub>H<sub>11</sub><sup>79</sup>BrN<sub>3</sub>O<sub>3</sub>S<sup>+</sup>: 367.9699).

*N,N'*-(4-Bromo-5-nitro-1,2-phenylene)bis-4-methylbenzene-sulfonamide (**19b**). A solution of **19a** (30.0 g, 60.6 mmol) in acetic acid (120 mL) was treated with a mixture of acetic acid (2.7 mL) and conc. nitric acid (2.0 mL). After stirring the mixture for 1 h at 50 °C, the reaction mixture solidified. The solid was filtered and washed with water (10 mL). Recrystallization from acetic acid (1.1 L) gave the nitro compound **19b** (21.8 g, 67%) as a beige solid. Unreacted starting material **19a** (7.20 g, 24%) was reisolated by evaporation of the filtrate of the recrystallization.  $R_f$  = 0.11 (SiO<sub>2</sub>; cy-hexane:EtOAc 50:50); <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 2.37 (s, 6H; CH<sub>3</sub>), 7.39 (dd,  $J$  = 8.1, 8.1 Hz, 4H; H-C(3')), 7.47 (s, 1H; H-C(3)), 7.63 (d,  $J$  = 8.4 Hz, 2H; H-C(2')), 7.67 (d,  $J$  = 8.4 Hz, 2H; H-C(2')), 7.74 (s, 1H; H-C(6)), 9.82 ppm (br. s, 2H; NH); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta$  = 21.1 (2C; CH<sub>3</sub>), 110.8 (C(2)), 120.6 (C(3)), 124.2 (C(6)), 126.89 (2C; C(2')), 126.91 (2C; C(2')), 127.1 (C(1)), 130.0 (2C; C(3')), 130.1 (2C; C(3')), 135.6 (C(1')), 135.7 (C(1')), 135.8 (C(4)), 143.2 (C(5)), 144.2 (C(4')), 144.3 ppm (C(4')); IR (ATR):  $\tilde{\nu}$  = 3253 (m), 1599 (w), 1579 (m), 1526 (m), 1488 (m), 1434 (m), 1410 (w), 1376 (m) 1327 (s), 1289 (m), 1272 (m), 1185 (w), 1162 (s), 1089 (s), 1019 (w), 985 (w), 908 (m), 866 (m), 835 (m), 812 (s), 750 (w), 723 (m), 689 (w), 666 (s), 636 (w), 609 cm<sup>-1</sup> (w); HR-EI-MS:  $m/z$  (%): 563.9692 (51, [M + Na]<sup>+</sup>, calcd. C<sub>20</sub>H<sub>9</sub><sup>81</sup>BrN<sub>3</sub>NaO<sub>6</sub>S<sub>2</sub><sup>+</sup>: 563.9693), 561.9715 (45, [M + Na]<sup>+</sup>, calcd. for C<sub>20</sub>H<sub>9</sub><sup>79</sup>BrN<sub>3</sub>NaO<sub>6</sub>S<sub>2</sub><sup>+</sup>: 561.9713), 541.9872 (26, [M + H]<sup>+</sup>, calcd. for C<sub>20</sub>H<sub>10</sub><sup>81</sup>BrN<sub>3</sub>O<sub>6</sub>S<sub>2</sub><sup>+</sup>: 541.9874), 539.9893 (24, [M + H]<sup>+</sup>, calcd. for C<sub>20</sub>H<sub>10</sub><sup>79</sup>BrN<sub>3</sub>O<sub>6</sub>S<sub>2</sub><sup>+</sup>: 539.9893).

1  
2  
3 **4-Bromo-5-nitrobenzene-1,2-diamine (20)**. A mixture of **19b** (16.5 g, 30.5 mmol) in a mixture of conc.  
4 sulfuric acid (31 mL) and water (3.0 mL) was stirred for 2 h at 85 °C. It was cooled to 0 °C and poured  
5 into water (400 mL). An aq. sat. solution of NH<sub>4</sub>Cl was added until pH 10 was reached. The solid was  
6 filtered and recrystallization from EtOH/H<sub>2</sub>O (3:2; 250 mL) to afford **20** (6.17 g, 87%) as a red solid.  
7  $R_f = 0.50$  (SiO<sub>2</sub>; EtOAc); <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta = 5.20$  (s, 2H; NH<sub>2</sub>), 6.08 (s, 2H; NH<sub>2</sub>), 6.79 (s,  
8 1H; H-C(6)), 7.38 ppm (s, 1H; H-C(3)); <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta = 103.5$  (C(2)), 110.6 (C(3)),  
9 116.6 (C(1)), 133.8 (C(4)), 136.0 (C(6)), 142.5 ppm (C(5)); IR (ATR):  $\tilde{\nu} = 3699$  (br. w), 3476 (w), 3418 (m),  
10 3346 (m), 3078 (w), 2746 (w), 1618 (s), 1566 (m), 1514 (s), 1471 (s), 1404 (w), 1375 (w), 1289 (s), 1235  
11 (s), 1132 (m), 1061 (w), 979 (w), 872 (s), 851 (w), 837 (m), 747 (m), 683 (w), 656 cm<sup>-1</sup> (w); HR-EI-MS:  
12  $m/z$  (%): 233.9697 (30, [M + H]<sup>+</sup>, calcd. for C<sub>6</sub>H<sub>7</sub><sup>81</sup>BrN<sub>3</sub>O<sub>2</sub><sup>+</sup>: 233.9696), 231.9718 (29, [M + H]<sup>+</sup>, calcd. for  
13 C<sub>6</sub>H<sub>7</sub><sup>79</sup>BrN<sub>3</sub>O<sub>2</sub><sup>+</sup>: 231.9716).

14  
15  
16  
17  
18  
19 **6-Bromo-2-(methylamino)-5-nitro-benzimidazole (21)**. A solution of diamine **20** (2.32 g, 10.0 mmol) in  
20 dry (CH<sub>3</sub>)<sub>2</sub>SO (20 mL) was treated with methylisothiocyanate (804 mg, 11.0 mmol), and the mixture  
21 was stirred for 21 h at 70 °C. EDC (2.1 mL, 12 mmol) was added and the mixture was stirred for 1 h at  
22 70 °C. The reaction mixture was frozen with liquid nitrogen, and the solvent was lyophilized under  
23 reduced pressure. FC (SiO<sub>2</sub>; EtOAc:(CH<sub>3</sub>)<sub>2</sub>CO 100:0 to 40:60) afforded the benzimidazole **21** (1.53 g,  
24 56%) as an orange solid. Additionally, unreacted starting material was reisolated by FC (470 mg, 20%).  
25  $R_f = 0.54$  (SiO<sub>2</sub>; (CH<sub>3</sub>)<sub>2</sub>CO); m.p. = 287–288 °C, <sup>1</sup>H NMR (400 MHz, (CD<sub>3</sub>)<sub>2</sub>SO):  $\delta = 2.91$  (d,  $J = 4.8$  Hz, 3H;  
26 CH<sub>3</sub>NH), 7.31 (br. s, 1H; (CH<sub>3</sub>)NH), 7.46 (s, 1H; H-C(7)), 7.81 (s, 1H; H-C(4)), 11.47 ppm (br. s, 1H; NH);  
27 <sup>13</sup>C NMR (101 MHz, (CD<sub>3</sub>)<sub>2</sub>SO:F<sub>3</sub>CCO<sub>2</sub>D 95:5)  $\delta = 29.3$  (CH<sub>3</sub>NH), 107.2, 108.8 (C(4)), 115.9 (C(7)), 129.6,  
28 134.0, 144.75, 153.0 ppm; IR (ATR):  $\tilde{\nu} = 3256$  (br. w), 3125 (w), 2883 (br. m), 1681 (m), 1614 (m), 1572  
29 (m), 1542 (w), 1500 (w), 1470 (s), 1435 (m), 1408 (m), 1375 (w), 1361 (w), 1296 (s), 1248 (s), 1199 (s),  
30 1156 (m), 1115 (m), 1089 (s), 986 (w), 971 (s), 879 (s), 850 (m), 832 (m), 774 (m), 745 (m), 735 (m), 719  
31 (m), 702 (m), 612 cm<sup>-1</sup> (s); HR-EI-MS:  $m/z$  (%): 272.9807 (100, [M + H]<sup>+</sup>, calcd for C<sub>8</sub>H<sub>8</sub><sup>81</sup>BrN<sub>4</sub>O<sub>2</sub><sup>+</sup>:  
32 272.9807), 270.9827 (98, [M + H]<sup>+</sup>, calcd. for C<sub>8</sub>H<sub>8</sub><sup>79</sup>BrN<sub>4</sub>O<sub>2</sub><sup>+</sup>: 270.9825).

33  
34  
35  
36  
37  
38 **6-Bromo-2-(methylamino)-5-nitro-1-[2-(trimethylsilyl)ethoxy]methyl-benzimidazole (22) and 5-bromo-**  
39 **2-(methylamino)-6-nitro-1-[2-(trimethylsilyl) ethoxy]methylbenzimidazole (22')**. Solid NaH (60 wt%  
40 dispersion in mineral oil) was washed with *n*-hexane (3x 6 mL). The remaining NaH (430 mg,  
41 17.9 mmol) was suspended in dry DMF (10 mL), treated portionwise with a solution of benzimidazole  
42 **21** (3.80 g, 13.8 mmol) in DMF (20 mL) at 0 °C, and the mixture was then allowed to reach 23 °C. After  
43 stirring the reaction mixture for 30 min at 23 °C, [2-(trimethylsilyl)ethoxy]methyl chloride (3.2 mL,  
44 18 mmol) was added at 0 °C. The reaction mixture was then allowed to reach 23 °C and was stirred for  
45 an additional 30 min. H<sub>2</sub>O (5 mL) was added, and the solvents were evaporated. H<sub>2</sub>O (20 mL) was  
46 added and the aq. layer was extracted with CH<sub>2</sub>Cl<sub>2</sub> (4x 100 mL). The combined organic layers were  
47 dried over MgSO<sub>4</sub>, and the solvents were removed under reduced pressure. FC (SiO<sub>2</sub>; *cy*-  
48 hexane:acetone 80:20) gave the protected benzimidazole as an inseparable tautomeric mixture of **22**  
49 and **22'** (4.26 g, 77%; 1.2:1 ratio) as a yellow solid.  $R_f = 0.55$  (SiO<sub>2</sub>; *cy*-hexane:acetone 50:50); m.p. =  
50 138–139 °C; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta = -0.01$ , 0.00 (2 s, 10.8H and 9H; Si(CH<sub>3</sub>)<sub>3</sub>), 0.93–0.97 (m, 2.4H  
51 and 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 3.17, 3.19 (2 d,  $J = 5.0$  Hz, 3H and 3.6H; CH<sub>3</sub>NH), 3.54–3.62 (m, 2.4H and 2H; OCH<sub>2</sub>N),  
52 4.99 (q,  $J = 5.0$  Hz, 1H; CH<sub>3</sub>NH), 5.15 (q,  $J = 5.0$  Hz, 1.2H; CH<sub>3</sub>NH), 5.27 (s, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 5.29 (s, 2.4H;  
53 CH<sub>2</sub>CH<sub>2</sub>Si), 7.33 (s, 1H; H-C(4')), 7.69 (s, 1.2H; H-C(7)), 7.81 (s, 1.2H; H-C(4)), 8.01 ppm (s, 1H; H-C(7'));  
54 <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta = -1.30$  (3C; Si(CH<sub>3</sub>)<sub>3</sub>), -1.29 (3C; Si(CH<sub>3</sub>)<sub>3</sub>), 17.9, 18.0 (CH<sub>2</sub>CH<sub>2</sub>Si), 29.9  
55 (CH<sub>3</sub>NH), 67.2, 67.3 (CH<sub>2</sub>CH<sub>2</sub>Si), 72.66 (NCH<sub>2</sub>O), 72.72 (NCH<sub>2</sub>O), 105.0 (C(5')), 105.5 (C(4)), 108.6 (C(5)),  
56 112.0 (C(4')), 114.4 (C(7')), 121.8 (C(7)), 133.3 (C(3a)), 138.1 (C(7a')), 141.4 (C(7a)), 142.0 (C(3a')), 144.5  
57 (C(6')), 147.8 (C(5)), 158.0, 159.4 ppm (C(2), C(2')); IR (ATR):  $\tilde{\nu} = 3339$  (w), 3191 (w), 2957 (w), 2870 (w),  
58  
59  
60

2051 (w), 1632 (m), 1609 (m), 1574 (s), 1520 (m), 1470 (m), 1461 (m), 1412 (w), 1321 (s), 1293 (m), 1275 (m), 1259 (m), 1246 (m), 1215 (m), 1186 (w), 1140 (w), 1118 (w), 1090 (s), 1081 (s), 1066 (m), 1045 (m), 911 (m), 862 (s), 834 (s), 804 cm<sup>-1</sup> (s); HR-EI-MS: *m/z* (%): 403.0612 (26, [M + H]<sup>+</sup>, calcd. for C<sub>14</sub>H<sub>22</sub><sup>81</sup>BrN<sub>4</sub>O<sub>3</sub>Si<sup>+</sup>: 403.0620), 401.0630 (25, [M + H]<sup>+</sup>, calcd. for C<sub>14</sub>H<sub>22</sub><sup>79</sup>BrN<sub>4</sub>O<sub>3</sub>Si<sup>+</sup>: 401.0639).

*2-(Methylamino)-5-nitro-1-((2-(trimethylsilyl)ethoxy)methyl)benzimidazole-6-carbaldehyde (23) and 2-(Methylamino)-6-nitro-1-((2-(trimethylsilyl)ethoxy)methyl)benzimidazole-5-carbaldehyde (23')*. A solution of the tautomeric mixture of **22** and **22'** (6.42 g, 16.0 mmol), Pd(OAc)<sub>2</sub> (180 mg, 800 μmol), PPh<sub>3</sub> (420 mg, 1.60 mmol), ethyl acrylate (2.6 mL, 28 mmol), and NEt<sub>3</sub> (3.9 mL, 53 mmol) in toluene (50 mL) was degassed with N<sub>2</sub> and stirred for 48 h at 110 °C under a N<sub>2</sub> atmosphere. The solvents were removed under reduced pressure, and the crude mixture was filtered over a silica plug (cy-hexane:acetone 1:1). The solvents were removed under reduced pressure and the obtained brown solid was directly dissolved in THF (60 mL) and H<sub>2</sub>O (15 mL). *N,N*-Dimethylbenzylamine hydrochloride (370 mg, 2.14 mmol), NaIO<sub>4</sub> (10.3 g, 48 mmol), and an aq. sol. of OsO<sub>4</sub> (4 wt%, 2.0 mL, 0.32 mmol) were added to the solution. After stirring the mixture for 36 h at 23 °C, a sat. aq. solution of Na<sub>2</sub>SO<sub>3</sub> (200 mL) was added, and stirring was continued for additional 30 min. The aq. layer was extracted with EtOAc (4x 200 mL), the combined organic layers were dried over MgSO<sub>4</sub>, and the solvents were evaporated. FC (SiO<sub>2</sub>; cy-hexane:acetone 100:0 to 80:20) gave the tautomeric benzaldehydes **23** (2.87 g, 52%) and **23'** (1.80 g, 32%) as orange solids. For **23**: *R*<sub>f</sub> = 0.52 (SiO<sub>2</sub>; cy-hexane:acetone 50:50); m.p. = 130–131 °C; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 0.00 (s, 9H; Si(CH<sub>3</sub>)<sub>3</sub>), 0.95–0.99 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 3.21 (d, *J* = 5.0 Hz, 3H; CH<sub>3</sub>NH), 3.59–3.63 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 5.23 (q, *J* = 4.9 Hz, 1H; CH<sub>3</sub>NH), 5.37 (s, 2H; NCH<sub>2</sub>O), 7.90 (s, 1H; H-C(4)), 7.90 (s, 1H; H-C(7)), 10.47 ppm (s, 1H; CHO); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = -1.3 (3C; Si(CH<sub>3</sub>)<sub>3</sub>), 18.0 (CH<sub>2</sub>CH<sub>2</sub>Si), 23.0 (CH<sub>3</sub>NH), 67.4 (CH<sub>2</sub>CH<sub>2</sub>Si), 72.9 (NCH<sub>2</sub>O), 104.0 (C(4)), 116.6 (C(7)), 128.6 (C(6)), 136.7 (C(7a)), 143.1 (C(3a)), 147.5 (C(5)), 159.7 (C(2)), 188.9 ppm (CHO); IR (ATR):  $\tilde{\nu}$  = 3112 (br. m), 2950 (w), 2919 (w), 1625 (s), 1602 (m), 1587 (s), 15010 (m), 1477 (m), 1427 (m), 1412 (m), 1334 (m), 1313 (s), 1291 (s), 1247 (m), 1218 (m), 1190 (m), 1137 (m), 1112 (m), 1084 (s), 1032 (m), 943 (m), 882 (m), 861 (s), 834 (s), 817 (s), 787 (m), 751 (m), 722 cm<sup>-1</sup> (m); HR-EI-MS: *m/z* (%): 373.1305 (26, [M + Na]<sup>+</sup>, calcd. for C<sub>15</sub>H<sub>22</sub>N<sub>4</sub>NaO<sub>4</sub>Si<sup>+</sup>: 373.1303), 351.1492 (38, [M + H]<sup>+</sup>, calcd. for C<sub>15</sub>H<sub>23</sub>N<sub>4</sub>O<sub>4</sub>Si<sup>+</sup>: 351.1483). For **23'** *R*<sub>f</sub> = 0.57 (SiO<sub>2</sub>; cy-hexane:acetone 50:50); m.p. = 126–127 °C; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ = -0.02 (s, 9H; Si(CH<sub>3</sub>)<sub>3</sub>), 0.93–0.97 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 3.21 (d, *J* = 5.0 Hz, 3H; CH<sub>3</sub>NH), 3.57–3.61 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 5.26 (q, *J* = 4.9 Hz, 1H; CH<sub>3</sub>NH), 5.37 (s, 2H; NCH<sub>2</sub>O), 7.69 (s, 1H; H-C(7)), 8.11 (s, 1H; H-C(4)), 10.40 ppm (s, 1H; CHO); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = -1.3 (3C; Si(CH<sub>3</sub>)<sub>3</sub>), 18.0 (CH<sub>2</sub>CH<sub>2</sub>Si), 29.9 (CH<sub>3</sub>NH), 67.4 (CH<sub>2</sub>CH<sub>2</sub>Si), 72.8 (NCH<sub>2</sub>O), 106.6 (C(7)), 112.9 (C(4)), 124.3 (C(5)), 137.8 (C(7a)), 146.7 (2C; C(3a), C(6)), 159.6 (C(2)), 188.4 ppm (CHO); IR (ATR):  $\tilde{\nu}$  = 3675 (w), 3191 (br. w), 2967 (m), 2901 (m), 1692 (m), 1623 (m), 1608 (m), 1569 (m), 1512 (m), 1476 (m), 1444 (m), 1412 (m), 1380 (m), 1316 (s), 1293 (m), 1279 (m), 1261 (m), 1248 (m), 1220 (m), 1186 (m), 1157 (m), 1139 (m), 1086 (s), 1066 (s), 1028 (m), 945 (m), 914 (m), 880 (m), 859 (s), 834 cm<sup>-1</sup> (s); HR-EI-MS: *m/z* (%): 373.1304 (22, [M + Na]<sup>+</sup>, calcd. for C<sub>15</sub>H<sub>22</sub>N<sub>4</sub>NaO<sub>4</sub>Si<sup>+</sup>: 373.1303), 351.1487 (27, [M + H]<sup>+</sup>, calcd. for C<sub>15</sub>H<sub>23</sub>N<sub>4</sub>O<sub>4</sub>Si<sup>+</sup>: 351.1483).

*(±)-1-(2-(Methylamino)-5-nitro-1-((2-(trimethylsilyl)ethoxy)methyl)-benzimidazol-6-yl)but-2-yn-1-ol ((±)-24)*. A solution of 1-bromo-1-propene (530 μL, 6.3 mmol) in THF (10 mL) was cooled to -78 °C. A solution *n*-BuLi in *n*-hexane (1.6 M; 7.9 mL, 13 mmol) was added dropwise, and the mixture was stirred for 2 h at -78 °C under an argon atmosphere. A solution of **23** (1.00 g, 2.85 mmol) in THF (3 mL) was cooled to -78 °C, and then added to the reaction mixture. After stirring the mixture for 1 h, H<sub>2</sub>O (1 mL), followed by an aq. sat. solution of NH<sub>4</sub>Cl (20 mL) was added. The aq. layer was extracted with EtOAc (3x 100 mL), the combined organic layers were dried over MgSO<sub>4</sub>, and the solvents were evaporated.

MPLC (SiO<sub>2</sub>; *n*-hexane:EtOAc 50:50 to 20:80) afforded racemic ( $\pm$ )-**24** (931 mg, 83%) as a yellow solid.  $R_f$  = 0.34 (SiO<sub>2</sub>; EtOAc); m.p. = 155–156 °C; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = –0.01 (s, 9H; Si(CH<sub>3</sub>)<sub>3</sub>), 0.93–0.97 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 1.90 (d,  $J$  = 2.2 Hz, 3H; C $\equiv$ CCH<sub>3</sub>), 3.21 (d,  $J$  = 4.9 Hz, 3H; CH<sub>3</sub>NH), 3.56–3.61 (m and br. s, 3H; CH<sub>2</sub>CH<sub>2</sub>Si and OH), 5.15 (br. s, 1H; CH<sub>3</sub>NH), 5.31 (s, 2H; NCH<sub>2</sub>O), 6.07 (q,  $J$  = 2.2 Hz, 1 H; CHOH), 7.85 (s, 1H; H–C(4)), 8.03 ppm (s, 1H; H–C(7)); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>)  $\delta$  = –1.4 (3C; Si(CH<sub>3</sub>)<sub>3</sub>), 3.8 (C $\equiv$ CCH<sub>3</sub>), 17.8 (CH<sub>2</sub>CH<sub>2</sub>Si), 29.8 (CH<sub>3</sub>NH), 62.3 (CHOH), 67.0 (CH<sub>2</sub>CH<sub>2</sub>Si), 72.6 (NCH<sub>2</sub>O), 77.6 (C $\equiv$ CCH<sub>3</sub>), 83.2 (C $\equiv$ CCH<sub>3</sub>), 104.8 (C(4)), 116.9 (C(7)), 132.2 (C(6)), 133.6 (C(3a)), 140.5 (C(7a)), 147.7 (C(5)), 159.3 ppm (C(2)); IR (ATR):  $\tilde{\nu}$  = 3676 (w), 3223 (br. w), 2953 (w), 2923 (w), 2901 (w), 1635 (m), 1608 (s), 1582 (s), 1513 (m), 1479 (m), 1416 (m), 1394 (m), 1381 (m), 1318 (s), 1283 (s), 1245 (s), 1213 (m), 1145 (w), 1110 (m), 1076 (s), 994 (m), 891 (m), 855 (s), 833 cm<sup>–1</sup> (s); HR-EI-MS:  $m/z$  (%): 391.1797 (47, [M + H]<sup>+</sup>, calcd. for C<sub>18</sub>H<sub>27</sub>N<sub>4</sub>O<sub>4</sub>Si<sup>+</sup>: 391.1796).

( $\pm$ )-*N,N'*-Bis(*tert*-butoxycarbonyl)-*N''*-[6-(1-hydroxy-but-2-ynyl)-2-(methylamino)-1-((2-(trimethylsilyl)ethoxy)methyl)-benzimidazol-5-yl)]guanidine (( $\pm$ )-**25**). A solution of ( $\pm$ )-**24** (400 mg, 1.02 mmol) was dissolved in a mixture of MeOH (10 mL) and an aq. sat. solution of NH<sub>4</sub>Cl (5 mL). Zinc powder (999 mg, 15.3 mmol) was added in small portions over 2.5 h at 0 °C. The reaction mixture was filtered over Celite<sup>®</sup> and washed with CH<sub>2</sub>Cl<sub>2</sub>. H<sub>2</sub>O (10 mL) was added. The aq. layer was separated, and extracted with CH<sub>2</sub>Cl<sub>2</sub> (2x 10 mL). The combined organic layers were dried over MgSO<sub>4</sub>, and the solvents were evaporated. The crude product was directly used in the next step without any purification. A solution of the crude aniline, *N,N'*-di-Boc-1*H*-pyrazole-1-carboximidine (447 mg, 1.44 mmol) and *N,N*-diisopropylethylamine (350  $\mu$ L, 2.0 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (5.0 mL) was stirred for 21 h in a sealed microwave vial at 50 °C. The solvents were evaporated and purification by MPLC (SiO<sub>2</sub>; *n*-hexane:EtOAc 10:90 to 0:100) gave the guanidine ( $\pm$ )-**25** (326 mg, 57%) as a pale brownish solid.  $R_f$  = 0.13 (SiO<sub>2</sub>; EtOAc); m.p. = 220–222 °C; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = –0.01 (s, 9H; Si(CH<sub>3</sub>)<sub>3</sub>), 0.91–0.95 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 1.40 (s, 9H; C(CH<sub>3</sub>)<sub>3</sub>), 1.55 (s, 9H; C(CH<sub>3</sub>)<sub>3</sub>), 1.88 (d,  $J$  = 2.2 Hz, 3H; C $\equiv$ CCH<sub>3</sub>), 3.14 (d,  $J$  = 5.0 Hz, 3H; CH<sub>3</sub>NH), 3.54–3.58 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 4.76 (q,  $J$  = 4.9 Hz, 1H; CH<sub>3</sub>NH), 4.83 (br. s, 1H; OH), 5.19 (s, 2H; NCH<sub>2</sub>O), 5.54 (q,  $J$  = 2.2 Hz, 1H; CHOH), 6.98 (s, 1H; H–C(4)), 7.96 (s, 1H; H–C(7)), 10.05 (s, 1H; NH), 11.56 ppm (s, 1H; NH); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  = –1.2 (3C; Si(CH<sub>3</sub>)<sub>3</sub>), 3.9 (C $\equiv$ CCH<sub>3</sub>), 18.0 (CH<sub>2</sub>CH<sub>2</sub>Si), 28.22 (3C; C(CH<sub>3</sub>)<sub>3</sub>), 28.24 (3C; C(CH<sub>3</sub>)<sub>3</sub>), 29.9 (CH<sub>3</sub>NH), 61.8 (CHOH), 66.7 (CH<sub>2</sub>CH<sub>2</sub>Si), 72.6 (NCH<sub>2</sub>O), 79.1 (C $\equiv$ CCH<sub>3</sub>), 80.1 (C(CH<sub>3</sub>)<sub>3</sub>), 82.3 (C $\equiv$ CCH<sub>3</sub>), 84.1 (C(CH<sub>3</sub>)<sub>3</sub>), 105.1 (C(4)), 117.3 (C(7)), 125.7 (C(6)), 131.8 (C(5)), 134.8 (C(7a)), 142.3 (C(3a)), 153.3 (NCO<sub>2</sub>), 156.5 (NCO<sub>2</sub>), 157.0 (C(2)), 162.9 ppm (NHCNNH); IR (ATR):  $\tilde{\nu}$  = 3676 (w), 3252 (br. w), 2987 (s), 2972 (s), 2901 (m), 1724 (w), 1603 (m), 1581 (m), 1476 (m), 1453 (m), 1406 (s), 1394 (s), 1368 (m), 1327 (m), 1279 (m), 1250 (m), 1229 (m), 1147 (s), 1104 (s), 1075 (s), 1066 (s), 1057 (s), 1027 (s), 857 (m), 834 (m), 807 (m), 765 cm<sup>–1</sup> (m); HR-EI-MS:  $m/z$  (%): 603.3314 (100, [M + H]<sup>+</sup>, calcd. for C<sub>29</sub>H<sub>47</sub>N<sub>6</sub>O<sub>6</sub>Si<sup>+</sup>: 603.3321).

( $\pm$ )-*tert*-Butyl 6-((*tert*-Butoxycarbonyl)amino)-2-(methylamino)-8-(prop-1-yn-1-yl)-3-((2-(trimethylsilyl)ethoxy)methyl)-3,8-dihydroimidazo[4,5]quinazoline-7-carboxylate (( $\pm$ )-**26**). A mixture of ( $\pm$ )-**25** (275 mg, 481  $\mu$ mol) and dry Cs<sub>2</sub>CO<sub>3</sub> (345 mg, 1.06 mmol) in DMF (5.0 mL) was stirred for 0.5 h at 23 °C, and then treated with a solution of TsCl (101 mg, 529  $\mu$ mol) in DMF (2.0 mL). The reaction mixture was stirred for additional 2 h. CH<sub>2</sub>Cl<sub>2</sub> (200 mL) was added, the aq. layer was separated, the organic layer was washed with H<sub>2</sub>O (5x 15 mL) and was dried over MgSO<sub>4</sub>. FC (SiO<sub>2</sub>; *cy*-hexane:EtOAc:NH<sub>3</sub> in MeOH (4 M) 50:49:1 to 30:69:1) gave the cyclized compound ( $\pm$ )-**26** (163 mg, 58 %) as a pale yellow solid in yield. Small amounts of ( $\pm$ )-**26** were also tosylated at the *exo*-cyclic amino group. This impurity could not be removed *via* flash chromatography.  $R_f$  = 0.53 (SiO<sub>2</sub>; EtOAc:NH<sub>3</sub> in MeOH (4 M) 99:1); m.p. = 122–123 °C; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = –0.02 (s, 9H; Si(CH<sub>3</sub>)<sub>3</sub>), 0.87–0.95 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 1.53 (s, 9H; C(CH<sub>3</sub>)<sub>3</sub>), 1.53 (s, 9H; C(CH<sub>3</sub>)<sub>3</sub>), 1.70 (d,  $J$  = 2.3 Hz, 3H; C $\equiv$ CCH<sub>3</sub>), 3.12 (d,  $J$  = 5.0 Hz, 3H; CH<sub>3</sub>NH), 3.55 (m, 2H; CH<sub>2</sub>CH<sub>2</sub>Si), 4.74 (q,  $J$  = 5.1 Hz, 1H; CH<sub>3</sub>NH), 5.22 (s, 2H; NCH<sub>2</sub>O), 6.04 (s, 1H; C(8)–H), 7.06 (br. s, 1 H;

H–C(4)), 7.20 (br. s, 1 H; H–C(9)), 8.27 ppm (br. s, 1H; NH); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>, four signals are hidden by noise due to rotameric broadening)  $\delta$  = –1.2 (Si(CH<sub>3</sub>)<sub>3</sub>), 4.0 (C≡CCH<sub>3</sub>), 18.1 (CH<sub>2</sub>CH<sub>2</sub>Si), 28.2 (3C; C(CH<sub>3</sub>)<sub>3</sub>), 28.4 (3C; C(CH<sub>3</sub>)<sub>3</sub>), 29.9 (CH<sub>3</sub>NH), 47.4 (C(8)), 66.7 (CH<sub>2</sub>CH<sub>2</sub>Si), 72.7 (NCH<sub>2</sub>O), 81.5 (C(CH<sub>3</sub>)<sub>3</sub>), 84.2 (C(CH<sub>3</sub>)<sub>3</sub>), 103.4 (C(4)), 112.5 (C(9)), 119.9, 133.7, 135.1, 150.3, 151.5, 156.5 ppm (C(2)); IR (ATR):  $\tilde{\nu}$  = 3675 (w), 2988 (s), 2972 (s), 2901 (s), 1725 (w), 1613 (m), 1573 (w), 1454 (m), 1407 (m), 1394 (m), 1369 (m), 1345 (m), 1249 (m), 1242 (m), 1230 (m), 1148 (m), 1075 (s), 1066 (s), 1057 (s), 1028 (m), 859 (m), 835 cm<sup>–1</sup> (m); HR-EI-MS: *m/z* (%): 585.3206 (92, [M + H]<sup>+</sup>, calcd. for C<sub>29</sub>H<sub>45</sub>N<sub>6</sub>O<sub>5</sub>Si<sup>+</sup>: 585.3215), 529.2584 (100, [M – CH<sub>2</sub>=C(CH<sub>3</sub>)<sub>2</sub> + H]<sup>+</sup>, calcd. for C<sub>25</sub>H<sub>37</sub>N<sub>6</sub>O<sub>5</sub>Si<sup>+</sup>: 529.2584).

(±)-*N*<sup>2</sup>-Methyl-8-(prop-1-yn-1-yl)-3,8-dihydroimidazo[4,5]quinazoline-2,6-diamine Dihydrochloride Salt ((±)-**4.2**). A solution of (±)-**28** (100 mg, 171 μmol) in CH<sub>2</sub>Cl<sub>2</sub> (2.0 mL) was treated dropwise with a solution of SnCl<sub>4</sub> (1.0 M in CH<sub>2</sub>Cl<sub>2</sub>, 1.7 mL, 1.7 mmol) at 0 °C under a N<sub>2</sub> atmosphere. After 0.5 h, the reaction mixture was allowed to reach 23 °C and was stirred additional 5 h. The solvents were removed under reduced pressure. Reversed-phase HPLC (column: Merck, LiChrospher RP-18e, 7 μm, 250 x 25 mm; flow: 10 mL/min; detection: 254 nm; eluent: MeOH:H<sub>2</sub>O (0.1 % TFA) 30:70; sample dissolved in eluent (ca. 15 mg/mL)) followed by an anion exchange (3x 50 μL HCl in MeOH followed by evaporation) afforded (±)-**4.2**·2HCl (19 mg, 34%) as a pale yellow solid. m.p. >165 °C (decomp.); <sup>1</sup>H NMR (600 MHz, CD<sub>3</sub>OD; seven exchanging NH protons not visible)  $\delta$  = 1.85 (d, *J* = 2.2 Hz, 3H; C≡CCH<sub>3</sub>), 3.11 (s, 3H; CH<sub>3</sub>NH), 5.64 (q, *J* = 1.9 Hz, 1H; H–C(8)), 7.12 (s, 1H; H–C(4)), 7.37 ppm (s, 1H; H–C(9)); <sup>13</sup>C NMR (151 MHz, CD<sub>3</sub>OD)  $\delta$  = 3.0 (C≡CCH<sub>3</sub>), 29.7 (CH<sub>3</sub>NH), 44.9 (C(8)), 77.0 (C≡CCH<sub>3</sub>), 83.7 (C≡CCH<sub>3</sub>), 100.1 (C(9)), 110.2 (C(4)), 117.6 (C(8a)), 128.6 (C(9a)), 130.0 (C(4a)), 132.0 (C(3a)), 153.3 (C(2)), 153.7 ppm (C(6)); IR (ATR):  $\tilde{\nu}$  = 3167 (br. m), 3051 (br. m), 2918 (br. m), 1671 (s), 1584 (s), 1490 (s), 1441 (m), 1392 (m), 1295 (w), 1185 (m), 1140 (m), 986 (m), 797 cm<sup>–1</sup> (w); HR-EI-MS: *m/z* (%): 255.1354 (100, [M + H]<sup>+</sup>, calcd. for C<sub>13</sub>H<sub>15</sub>N<sub>6</sub><sup>+</sup>: 255.1353).

## ASSOCIATED CONTENT

**Supporting Information.** Structures with SMILES CODES; MD Simulations; Crystal Data and Refinement Statistics for TGT Complex Structures; Chemical Synthesis and NMR Spectra; HPLC Traces; Small-Molecule X-Ray Data of **9**; Figure S49: Illustration of the structural rearrangement involved with the opening of the transient pocket; Figure S50: Electron density of the bound ligands and water molecules in the preQ1 recognition site of TGT.

**Accession Codes.** Coordinates and structure factors have been deposited in the Protein Data Bank: **1.1** (4PUK), **fragment** (5UTI), **2.1** (5J9M), **2.2** (5JT5), **2.5** (5J9N), **2.6** (5JT6), **3.1** (5J9O), **3.2** (5JT7), **5.1** (3S1G), **4.1** (6RKT), **4.2** (6RKQ). Authors will release the atomic coordinates and experimental data upon article publication.

## AUTHOR INFORMATION

## Corresponding Author

\*Mail: [klebe@staff.uni-marburg.de](mailto:klebe@staff.uni-marburg.de); phone: +49 6421 28 21313

## ORCID

Francois Diederich: 0000-0002-3277-6570

Engi Hassaan: 0000-0001-7393-7687

Christoph Hohn 0000-0003-2406-7683

Gerhard Klebe: 0000-0002-4913-390X

Klaus Reuter: 0000-0003-3673-7971

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The presented work was funded from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie Grant Agreement No. 675555, Accelerated Early stage drug discovery (AEGIS). Work at ETH Zurich was supported by a grant from the ETH Research Council (ETH-01 13-2). We thank the Helmholtz-Zentrum Berlin for the allocation of synchrotron radiation beamtime and travel support. We thank Dr. Anatol Schwab for help with HPLC separations and Dr. Bruno Bernet for help with the experimental section. We are grateful to Dr. Nils Trapp for solving a small-molecule X-ray crystal structure.

## ABBREVIATIONS

MD, Molecular Dynamics; HSA, Hydration Site Analysis; Glu, Glutamate; Gln, Glutamine, Asp; Aspartate; Thr, Threonine; Cys, Cysteine

## Literature

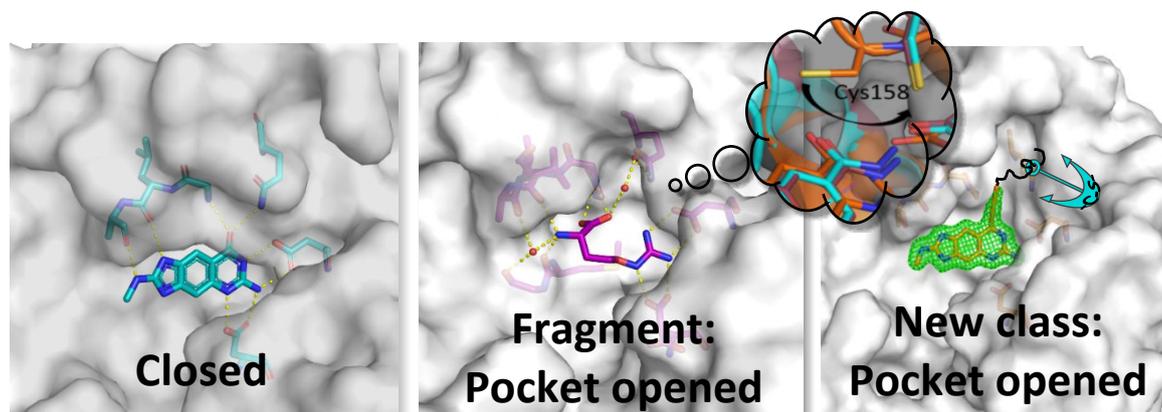
- [1] Joseph-McCarthy, D.; Campbell, A. J.; Kern, G.; Moustakas, D. Fragment-Based Lead Discovery and Design. *J. Chem. Inf. Model.* **2014**, *54*, 693–704.
- [2] Schoepfer, J.; Jahnke, W.; Berellini, G.; Buonamici, S.; Cotesta, S.; Cowan-Jacob, S. W.; Dodd, S.; Druceckes, P.; Fabbro, D.; Gabriel, T.; Groell, J.-M.; Grotzfeld, R. M.; Hassan, A. Q.; Henry, C.; Iyer, V.; Jones, D.; Lombardo, F.; Loo, A.; Manley, P. W.; Pellé, X.; Rummel, G.; Salem, B.; Warmuth, M.; Wylie, A. A.; Zoller, T.; Marzinzik, A. L.; Furet, P. Discovery of Asciminib (ABL001), an Allosteric Inhibitor of the Tyrosine Kinase Activity of BCR-ABL1. *J. Med. Chem.* **2018**, *61*, 8120–8135.

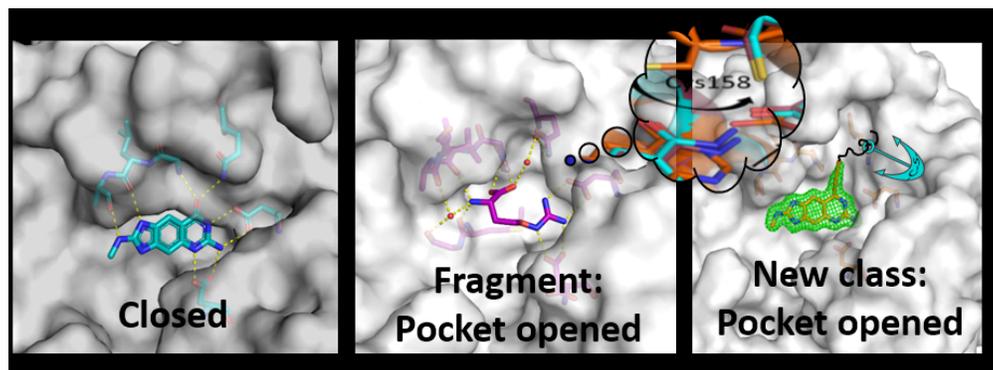
- 1  
2  
3 [3] Hassaan, E.; Eriksson, P.-O.; Geschwindner, S.; Heine, A.; Klebe, G. Fragments as Novel Starting  
4 Points for tRNA-Guanine Transglycosylase Inhibitors Found by Alternative Screening Strategies.  
5 *Chemmedchem*. **2019**, doi.org/10.1002/cmdc.201900604.  
6  
7 [4] Romier, C.; Reuter, K.; Suck, D.; Ficner, R. Crystal Structure of tRNA-Guanine Transglycosylase:  
8 RNA Modification by Base Exchange. *EMBO. J* **1996**, *15*, 2850–2857.  
9  
10 [5] Biela, I.; Tidten-Luksch, N.; Immekus, F.; Glinca, S.; Nguyen, T. X. P.; Gerber, H.-D.; Heine, A.;  
11 Klebe, G.; Reuter, K.; Investigation of Specificity Determinants in Bacterial tRNA-Guanine  
12 Transglycosylase Reveals Queuine, the Substrate of Its Eucaryotic Counterpart, as Inhibitor.  
13 *PLoS ONE* **2013**, *8*, e64240.  
14  
15 [6] Durand, J. M. B.; Dagberg, B.; Uhlin, B. E.; Björk, G. R. Transfer RNA Modification, Temperature  
16 and DNA Superhelicity Have a Common Target in the Regulatory Network of the Virulence of  
17 *Shigella Flexneri*: the Expression of the *virF* Gene. *Mol. Microbiol.* **2000**, *35*, 924–935.  
18  
19 [7] Grädler, U.; Gerber, H.-D.; Goodenough-Lashua, D. M.; Garcia, G. A.; Ficner, R.; Reuter, K.; Stubbs,  
20 M. T.; Klebe, G. A New Target for Shigellosis: Rational Design and Crystallographic Studies of  
21 Inhibitors of tRNA-Guanine Transglycosylase. *J. Mol. Biol.* **2001**, *306*, 455–467.  
22  
23 [8] Christopher, P. R.; David, K. V.; John, S. M.; Sankarapandian, V. Antibiotic Therapy for *Shigella*  
24 *Dysentery*. *Cochrane Database Syst. Rev.* **2010**, CD006784.  
25  
26 [9] Baker, S.; The, H. C. Recent Insights into *Shigella*: A Major Contributor to the Global Diarrhoeal  
27 Disease Burden. *Curr. Opin. Infect. Dis.* **2018**, *31*, 449–454.  
28  
29 [10] Stengl, B.; Reuter, K.; Klebe, G. Mechanism and Substrate Specificity of tRNA-Guanine  
30 Transglycosylases (TGTs): tRNA-Modifying Enzymes from the Three Different Kingdoms of Life  
31 Share a Common Catalytic Mechanism. *ChemBioChem* **2005**, *6*, 1926–1939.  
32  
33 [11] Neeb, M.; Czodrowski, P.; Heine, A.; Barandun, L. J.; Hohn, C.; Diederich, F.; Klebe, G. Chasing  
34 Protons: How Isothermal Titration Calorimetry, Mutagenesis, and  $pK_a$  Calculations Trace the  
35 Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme. *J. Med. Chem.* **2014**, *57*, 5554–  
36 5565.  
37  
38 [12] Kohler, P. C.; Ritschel, T.; Schweizer, W. B.; Klebe, G.; Diederich, F. High-Affinity Inhibitors of  
39 tRNA-Guanine Transglycosylase Replacing the Function of a Structural Water Cluster. *Chem.–*  
40 *Eur. J.* **2009**, *15*, 10809–10817.  
41  
42 [13] Hörtner, S. R.; Ritschel, T.; Stengl, B.; Kramer, C.; Schweizer, W. B.; Wagner, B.; Kansy, M.; Klebe,  
43 G.; Diederich, F. Potent Inhibitors of tRNA-Guanine Transglycosylase, an Enzyme Linked to the  
44 Pathogenicity of the *Shigella* Bacterium: Charge-Assisted Hydrogen Bonding. *Angew. Chem. Int.*  
45 *Ed.* **2007**, *46*, 8266–8269.  
46  
47 [14] Barandun, L. J.; Immekus, F.; Kohler, P. C.; Tonazzi, S.; Wagner, B.; Wendelspiess, S.; Ritschel, T.;  
48 Heine, A.; Kansy, M.; Klebe, G.; Diederich, F. From *lin*-Benzoguanines to *lin*-Benzohypoxanthines  
49 as Ligands for *Zymomonas mobilis* tRNA-Guanine Transglycosylase: Replacement of Protein-  
50 Ligand Hydrogen Bonding by Importing Water Clusters. *Chem.–Eur. J.* **2012**, *18*, 9246–9257.  
51  
52 [15] Hohn, C. Inhibitors of tRNA-Guanine Transglycosylase: From *lin*-Benzopurines to Transition State  
53 and Nucleobase Analogues. ETH Dissertation No. 23162, ETH Zürich, **2015**.  
54  
55 [16] Lecoutey, C.; Fossey, C.; Rault, S.; Fabis, F. Efficient Room-Temperature One-Pot Synthesis of 2-  
56 Amino-3-alkyl(3-aryl)quinazolin-4(3H)-ones *Eur. J. Org. Chem.* **2011**, 2785–2788.  
57  
58 [17] Ehrmann, F. R.; Kalim, J.; Pfaffeneder, T.; Bernet, B.; Hohn, C.; Schäfer, E.; Botzanowski, T.;  
59 Cianféroni, S.; Heine, A.; Reuter, K.; Diederich, F.; Klebe, G. Swapping Interface Contacts in the  
60 Homodimeric tRNA-Guanine Transglycosylase: An Option for Functional Regulation. *Angew. Chem. Int. Ed.* **2018**, *57*, 10085–10090.  
[18] Nasielski-Hinkens, R.; Levêque, P.; Castelet, D.; Nasielski, J. The Four 6-Halo-7-nitroquinoxalines  
*Heterocycles* **1987**, *26*, 2433–2442.  
[19] Meyer, E. A.; Donati, N.; Guillot, M.; Schweizer, W. B.; Diederich, F.; Stengl, B.; Brenk, R.; Reuter,  
K.; Klebe, G. Synthesis, Biological Evaluation, and Crystallographic Studies of Extended Guanine-

- Based (*lin*-Benzoguanine) Inhibitors for tRNA-Guanine Transglycosylase (TGT) *Helv. Chim. Acta* **2006**, *89*, 573–597.
- [20] Ritschel, T.; Hoertner, S.; Heine, A.; Diederich, F.; Klebe, G. Crystal Structure Analysis and in Silico  $pK_a$  Calculations Suggest Strong  $pK_a$  Shifts of Ligands as Driving Force for High-Affinity Binding to TGT. *ChemBioChem* **2009**, *10*, 716–727.
- [21] Brenk, R.; Stubbs, M. T.; Heine, A.; Reuter, K.; Klebe, G. Flexible Adaptations in the Structure of the tRNA-Modifying Enzyme tRNA-Guanine Transglycosylase and Their Implications for Substrate Selectivity, Reaction Mechanism and Structure-Based Drug Design. *ChemBioChem* **2003**, *4*, 1066–1077.
- [22] Rühmann, E. H.; Rupp, M.; Betz, M.; Heine, A.; Klebe, G. Boosting Affinity by Correct Ligand Preorganization for the S2 Pocket of Thrombin: A Study by Isothermal Titration Calorimetry, Molecular Dynamics, and High-Resolution Crystal Structures. *ChemMedChem* **2016**, *11*, 309–319.
- [23] Rechlin, C.; Scheer, F.; Terwesten, F.; Wulsdorf, T.; Pol, E.; Fridh, V.; Toth, P.; Diederich, W. E.; Heine, A.; Klebe, G. Price for Opening the Transient Specificity Pocket in Human Aldose Reductase upon Ligand Binding: Structural, Thermodynamic, Kinetic, and Computational Analysis *ACS Chem. Biol.* **2017**, *12*, 1397–1415.
- [24] Tidten, N. Structural and mutational characterisation of *Shigella* Pathogenicity Factors. Doctoral-Dissertation Philipps-Universität Marburg, **2007**.
- [25] Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. Protein Crystallography for Non-Crystallographers, or How to Get the Best (But Not More) from Published Macromolecular Structures. *FEBS J.* **2008**, *275*, 1–21.
- [26] Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.
- [27] Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- [28] Romier, C.; Meyer, J. E. W.; Suck, D. Slight Sequence Variations of a Common Fold Explain the Substrate Specificities of tRNA-Guanine Transglycosylases from the Three Kingdoms. *FEBS Lett.* **1997**, *416*, 93–98.
- [29] Chaikuad, A.; Koch, P.; Laufer, S. A.; Knapp, S. The Cysteinome of Protein Kinases as a Target in Drug Development *Angew. Chem. Int. Ed.* **2018**, *57*, 4372–4385.
- [30] Zhao, Z.; Liu, Q.; Bliven S.; Xie, L.; Bourne P. E. Determining Cysteins Available for Covalent Inhibition Across the Human Kinome. *J. Med. Chem.* **2017**, *60*, 2879–2889.
- [31] Gehringer, M.; Laufer S. A. Emerging and Re-Emerging Warheads for Targeted Covalent Inhibitors: Applications in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2019**, *62*, 5673–5724.
- [32] Jakobi, S.; Nguyen, P. T. X.; Debaene, F.; Cianféroni, S.; Reuter, K.; Klebe, G. What Glues a Homodimer Together: Systematic Analysis of the Stabilizing Effect of an Aromatic Hot Spot in the Protein–Protein Interface of the tRNA-Modifying Enzyme Tgt. *ACS Chem. Biol.* **2015**, *10*, 1897–1907.
- [33] Kabsch, W. *XDS. Acta Cryst.* **2010**, *D66*, 125–132.
- [34] Krug, M.; Weiss, M. S.; Heinemann, U.; Mueller, U. *XDSAPP: A Graphical User Interface for the Convenient Processing of Diffraction Data using XDS. J. Appl. Cryst.* **2012**, *45*, 568–572.
- [35] McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *Phaser Crystallographic Software. J. Appl. Cryst.* **2007**, *40*, 658–674.
- [36] Collaborative Computational Project, Number 4. The *CCP4* Suite: Programs for Protein Crystallography. *Acta Cryst.* **1994**, *D50*, 760–763.
- [37] Emsley, P.; Cowtan, K. *Coot: Model-building Tools for Molecular Graphics. Acta Cryst.* **2004**, *D60*, 2126–2132.
- [38] Adams, P. D.; Afonine P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner R.; Read, R.

- J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart P. H. *PHENIX: A Comprehensive Python-based System for Macromolecular Structure Solution. Acta Cryst.* **2010**, *D66*, 213–221.
- [39] <https://www.chemicalize.com>, accessed March 2017 and May 2018.
- [40] <http://www.chemaxon.com>, accessed March 2017 and May 2018.
- [41] Smart, O. S.; Womack, T. O.; Sharff, A.; Flensburg, C.; Keller, P.; Paciorek, W.; Vonrhein, C.; Bricogne, G. (2011) grade, version v1.103. Cambridge, United Kingdom; Global Phasing Ltd., <http://www.globalphasing.com>, accessed March 2017 and May 2018.
- [42] Curnow, A. W.; Kung, F. L.; Koch, K. A.; Garcia, G. A. tRNA-Guanine Transglycosylase from *Escherichia Coli*: Gross tRNA Structural Requirements for Recognition. *Biochemistry* **1993**, *32*, 5239–5246.
- [43] Molecular Operating Environment (MOE), version 2015.10 (2016), Chemical Computing Group (CCG) Inc., Montreal, Canada.
- [44] Case, D. A.; Cerutti, D. S.; Cheatham III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu J.; Luchko, T.; Luo, R.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 2017 University of California, San Francisco.
- [45] Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65*, 712–725.
- [46] Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.
- [47] Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed without Compromise – A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374–380.
- [48] Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555.
- [49] Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.
- [50] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- [51] Haider, K.; Cruz, A.; Ramsey, S.; Gilson, M. K.; Kurtzman, T. Solvation Structure and Thermodynamic Mapping (SSTMap): An Open-Source, Flexible Package for the Analysis of Water in Molecular Dynamics Trajectories. *J. Chem. Theory Comput.* **2018**, *14*, 418–425.
- [52] <https://github.com/KurtzmanLab/SSTMap>, accessed 11-08-2018

## ToC Graphic





Graphic for Table of Content

156x57mm (150 x 150 DPI)