

Luzi Jakob Barandun,^{a‡} Florian Immekus,^{b‡} Philipp C. Kohler,^{a‡} Tina Ritschel,^{b§} Andreas Heine,^b Pierfrancesco Orlando,^a Gerhard Klebe^{b*} and François Diederich^{a*}

^aLaboratorium für Organische Chemie, ETH Zürich, Hönggerberg, HCI, 8093 Zurich, Switzerland, and ^bInstitut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany

‡ These authors contributed equally to this work.

§ Current address: Computational Discovery and Design Group, CMBI, Radboud University Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands.

Correspondence e-mail:
klebe@mail.uni-marburg.de,
diederich@org.chem.ethz.ch

High-affinity inhibitors of *Zymomonas mobilis* tRNA–guanine transglycosylase through convergent optimization

The tRNA-modifying enzyme tRNA–guanine transglycosylase (TGT) has been recognized as a drug target for the treatment of the foodborne illness shigellosis. The active site of TGT consists of three pockets: the central guanine/preQ₁ recognition site and the ribose-33 and ribose-34 pockets. In previous work, *lin*-benzoguanines and *lin*-benzohypoxanthines, which differ by the presence of an exocyclic NH₂ group in the former and its absence in the latter, were used as central scaffolds that bind to the guanine/preQ₁ recognition site and allow suitable functionalization along exit vectors targeting the two ribose pockets. The substituents for both of these two pockets have been optimized individually. Here, a series of bifunctionalized inhibitors that occupy both ribose pockets are reported for the first time. Dissociation constants K_d down to the picomolar range were measured for the bifunctionalized *lin*-benzoguanine-based ligands and K_d values in the nanomolar range were measured for the corresponding *lin*-benzohypoxanthine-based ligands. The binding mode of all inhibitors was elucidated by X-ray crystal structure analysis. A remarkable influence of the crystallization protocol on the solvation pattern in the solid state and the residual mobility of the bound ligands was observed.

Received 10 April 2013

Accepted 26 May 2013

PDB References: tRNA–guanine transglycosylase–inhibitor complexes, 4gg9; 4gh1; 4gh3; 4gi4; 4giy; 4gkt

1. Introduction

Shigellosis, also known as bacillary dysentery, is an acute inflammatory bowel disease caused by bacteria of the genus *Shigella*. It is endemic throughout the world, but the majority of cases occur in areas with poor hygiene standards. In particular, children under five years of age are very vulnerable (Sansone, 2001a). About 150 million annual cases, with one million deaths, have been estimated (Niyogi, 2005). An exact evaluation of the burden of the disease is difficult, and the available data are probably largely underestimated (Sansone, 2006). Shigellosis is highly contagious, as 10–100 bacteria are sufficient to cause the disease (Sansone, 2001b). To date, no vaccine against shigellosis is available, but polysaccharide-conjugate and attenuated live vaccines are in development (for examples, see Levine *et al.*, 2007; Launay *et al.*, 2009; Kubler-Kielb *et al.*, 2010; Ranallo *et al.*, 2012). Resistance to common antibiotics has been reported for many strains of *Shigella* bacteria (Ashkenazi *et al.*, 2003). Therefore, understanding the pathogenic mechanism and the development of drugs for new targets is of the utmost urgency.

The tRNA-modifying enzyme tRNA–guanine transglycosylase (TGT; EC 2.4.2.29) has been identified as a potential drug target for the treatment of shigellosis, as it is essential for the pathogenicity of *Shigella* bacteria (Durand *et*

al., 1994; Grädler *et al.*, 2001). Prokaryotic TGT exchanges guanine for preQ₁ at the wobble position 34 of tRNA^{Asn, Asp, His, Tyr} (Iwata-Reuyl, 2003), whereas eukaryotic TGT has a different substrate and incorporates queuine (Fig. 1*a*). Its structural features and catalytic activity have been well characterized using the enzyme from *Zymomonas mobilis*, which has an almost identical active site to the *Shigella* enzyme apart from a Tyr/Phe exchange (Brenk *et al.*, 2003; Stengl *et al.*, 2007). Crystal structure analysis of *Z. mobilis* TGT in complex with a tRNA substrate revealed a homodimeric structure as the catalytically active species (Xie *et al.*, 2003), which was confirmed by mass spectrometry (Ritschel, Atmanene *et al.*, 2009). In nearly all studied crystal structures, the homodimer has crystallographic twofold

symmetry owing to the crystal packing. The active site of *Z. mobilis* TGT consists of three pockets (Fig. 1*b*): the guanine/preQ₁ recognition site, where the base exchange takes place, and the ribose-33 and ribose-34 pockets, where the phosphate backbone of the tRNA is accommodated.

In recent years, we have used structure-based design to develop *lin*-benzohypoxanthines and *lin*-benzoguanines as inhibitors of *Z. mobilis* TGT (Fig. 1*a*; Meyer *et al.*, 2006; Hörtner *et al.*, 2007; Ritschel, Hoertner *et al.*, 2009; Kohler *et al.*, 2009; Ritschel, Kohler *et al.*, 2009; Barandun *et al.*, 2012). The two scaffolds differ by the presence of an exocyclic NH₂ at C(6) in the latter, which was found to replace a water cluster in the active site and to form strong ionic hydrogen bonds to the enzyme (Barandun *et al.*, 2012). These additional interactions make *lin*-benzoguanine derivatives stronger inhibitors than the analogous *lin*-benzohypoxanthines. The central *lin*-benzopurine core binds to the guanine/preQ₁ pocket and mimics the complex hydrogen-bonding pattern of the natural substrates (Fig. 1*c*). The *lin*-benzopurine scaffold can be substituted either at C(2) or C(4) to target the ribose-33 or ribose-34 pocket, respectively. The addition of the 2-methylamino substituent revealed a strongly enhanced affinity of 4100 nM in the case of the unsubstituted *lin*-benzoguanine scaffold (Fig. 1*a*) compared with 58 nM in the case of **1a** (Table 1), which was explained by a significant p*K*_a shift resulting in protonation of the imidazole moiety of the parent scaffold (Hörtner *et al.*, 2007; Ritschel, Hoertner *et al.*, 2009). Further extension of the 2-methylamino group towards the ribose-33 pocket yielded *lin*-benzoguanines with inhibitory constants of 6–35 nM (compounds **1b–1d**). The analogous

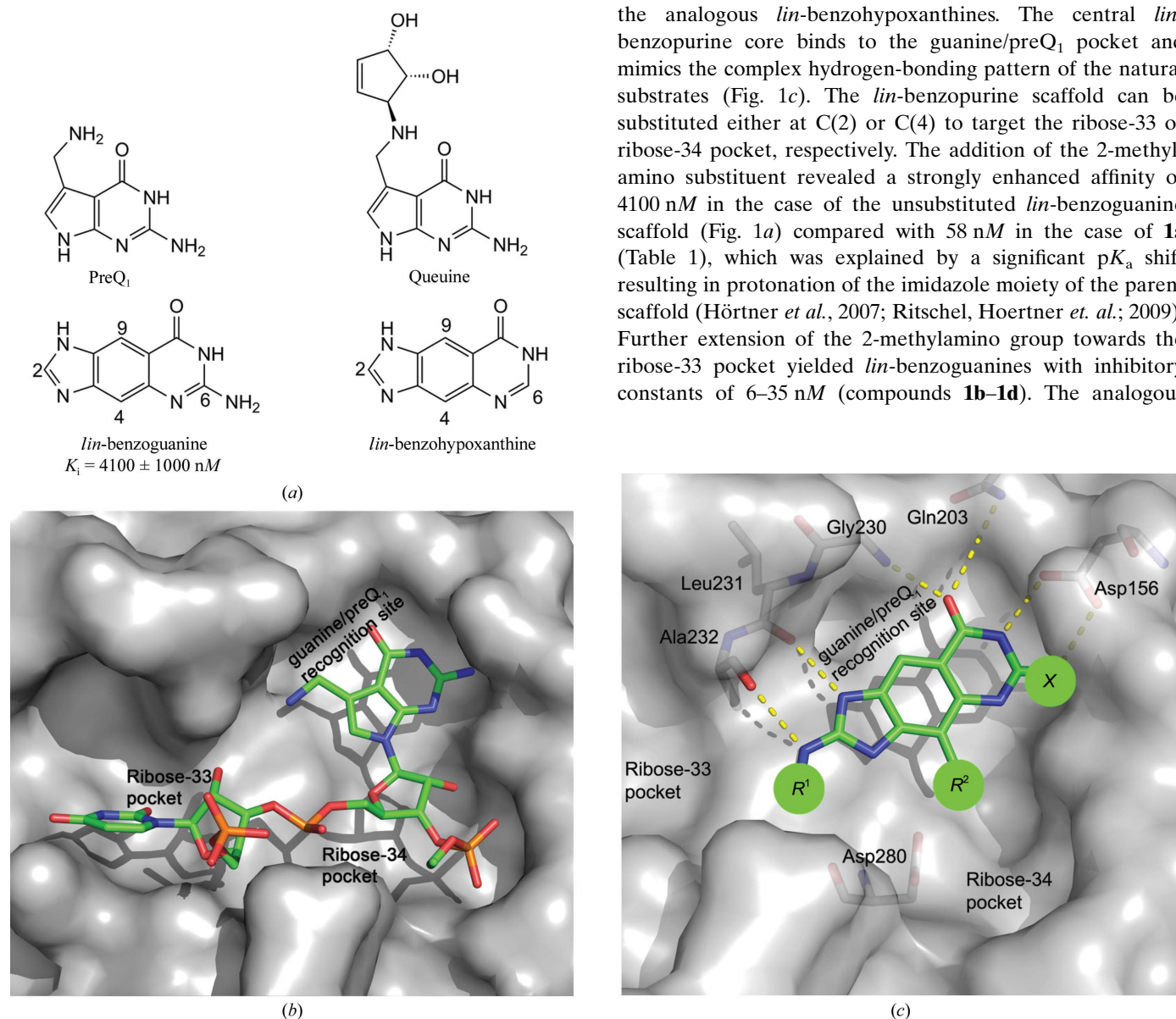
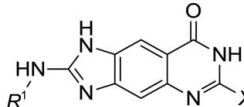


Figure 1
(a) Chemical structures of preQ₁, queuine, *lin*-benzoguanine and *lin*-benzohypoxanthine. (b) Active site of *Z. mobilis* TGT consisting of three major pockets: the guanine/preQ₁ recognition site and the ribose-33 and ribose-34 pockets (image based on PDB entry 1q2s; Xie *et al.*, 2003). A fragment of a tRNA substrate is shown in the active site. (c) Schematic binding mode of the *lin*-benzopurines with vectors for the ribose-33 and/or ribose-34 pockets. Colour code: C_{enzyme}, grey; C_{ligand}, green; O, red; N, blue; enzyme surface, grey. These characteristics apply to all figure captions unless stated otherwise.

Table 1

Competitive inhibition constants K_i , $\text{clog}D$ and $\text{clog}P$ values and ligand efficiencies of selected *lin*-benzopurines with vectors for the ribose-33 pocket.



X	R^1	K_i (nM)	$\text{clog}D^\dagger$ (pH 7.4)	$\text{clog}P^\ddagger$	Ligand efficiency ‡ (kcal mol $^{-1}$)
1a	NH $_2$	58 ± 36 §	-0.33	-0.05	0.58
2a	H	6500 ± 2900 ¶	0.18	-0.19	0.44
1b	NH $_2$	35 ± 9 §	1.09	1.41	0.46
2b	H	2900 ¶	1.76	1.27	0.36
1c	NH $_2$	6 ± 6 §	-0.89	-0.45	0.47
2c	H	4100 ¶	-0.22	-0.59	0.32
1d	NH $_2$	10 ± 3 §	1.76	2.04	0.45
2d	H	3700 ¶	2.37	1.90	0.32

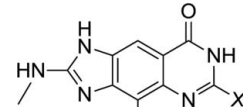
† Calculated using *ACD/Labs* v.12.01 (ACD/Labs, Toronto, Canada). ‡ Ligand efficiency = $RT\ln[K/(\text{number of non-H atoms})]$. 1 cal = 4.186 J. § Hörtner *et al.* (2007). ¶ Barandun *et al.* (2012).

lin-benzohypoxanthines had inhibition constants in the range 2900–6500 nM (Barandun *et al.*, 2012; compounds **2a–2d**). For these inhibitors with extended substituents at C(2), crystal structure analysis showed well resolved difference electron density of the *lin*-benzopurine scaffold but only partly defined difference density characterizing the binding pose of the substituent in the ribose-33 pocket. This was attributed to either high residual mobility or to pronounced disorder with a static scatter over multiple orientations of the substituent in the wide solvent-exposed pocket. The ribose-34 pocket was targeted using 4-ethylamino substituents. Surprisingly, the addition of an unfunctionalized ethylamino substituent at C(4) (compound **3**; $K_i = 55 \pm 3$ nM; Table 2) gave no improvement in affinity. Only further derivatization led to *lin*-benzoguanines **4a–4c** (Kohler *et al.*, 2009; Ritschel, Kohler *et al.*; 2009) with K_i values of 2–25 nM and the corresponding *lin*-benzohypoxanthines **5a** and **5b** (Barandun *et al.*, 2012) with K_i values of 740–1100 nM. In this case, the binding mode of the entire ligand was elucidated by well resolved difference density in the crystal structures obtained by cocrystallization.

Here, we investigate two series of inhibitors that simultaneously target both ribose pockets with extended substituents departing from C(2) and C(4) of the *lin*-benzopurine scaffold. Assuming additivity of the attached substituents, the combination of the independently optimized substituents for occupation of both the ribose-33 and ribose-34 pockets should lead to highly potent ligands. Both *lin*-benzopurine scaffolds were studied, as it was anticipated that the generally more potent *lin*-benzoguanines would have binding affinities in the sub-nanomolar range, which might become difficult to measure with high accuracy in a biochemical enzyme assay. The molecular-recognition effects of simultaneously filling both

Table 2

Competitive inhibition constants K_i , $\text{clog}D$ and $\text{clog}P$ values and ligand efficiencies of selected *lin*-benzopurines with vectors for the ribose-34 pocket.



X	R^2	K_i (nM)	$\text{clog}D$ (pH 7.4) †	$\text{clog}P^\ddagger$	Ligand efficiency ‡ (kcal mol $^{-1}$)
3	NH $_2$	55 ± 3 §	-2.54	-0.69	0.50
4a	NH $_2$	4 ± 2 §	-0.87	1.78	0.44
5a	H	1100 ± 370 ¶	-0.57	1.65	0.33
4b	NH $_2$	2 ± 1 §	-0.23	2.31	0.44
5b	H	740 ± 170 ¶	-0.10	2.17	0.32
4c	NH $_2$	25 ± 2 ¶	-0.24	1.46	0.38

† Calculated using *ACD/Labs* v.12.01 (ACD/Labs, Toronto, Canada). ‡ Ligand efficiency = $RT\ln[K/(\text{number of non-H atoms})]$. 1 cal = 4.186 J. § Kohler *et al.* (2009). ¶ Barandun *et al.* (2012).

ribose pockets were investigated in six X-ray crystal structures of *Z. mobilis* TGT.

2. Materials and methods

Synthetic procedures, analytical data and NMR spectra are collected in the Supplementary Material¹. The modelling of ligands prior to synthesis and structural analysis was performed using the *MOLOC* molecular-modelling package (Gerber & Müller, 1995).

2.1. *Z. mobilis* TGT crystallization

TGT gene cloning, protein overexpression and purification were performed as described in detail elsewhere (Reuter & Ficner, 1995; Romier *et al.*, 1996). Crystals of TGT suitable for data collection were obtained using the hanging-drop vapour-diffusion method at 288 K. TGT was crystallized in the presence of the inhibitors. A protein solution (12 mg ml $^{-1}$ TGT, 10 mM Tris-HCl pH 7.8, 2 M NaCl, 1 mM EDTA) was incubated with inhibitor stock solution in dimethyl sulfoxide (DMSO) with final concentrations of 1.5 mM inhibitor and 15% (v/v) DMSO. 2 μ l of this solution was mixed with 2 μ l reservoir solution [100 mM MES pH 5.5, 1 mM DTT, 10% (v/v) DMSO, 13% (w/v) PEG 8000]. After three weeks of crystal growth, the specimen reached dimensions of about 0.7 × 0.7 × 0.2 mm. Instead of cocrystallization, compound **6a** was soaked at a final concentration of 5 mM for 1 d into premanufactured wild-type crystals owing to poor cocrystal

¹ Supplementary material has been deposited in the IUCr electronic archive (Reference: BE5233). Services for accessing this material are described at the back of the journal.

Table 3

X-ray cocrystal structures: data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Crystal data	6a	6b	6c	7a	7b	7c
PDB code	4gi4	4gkt	4giy	4gg9	4gh1	4gh3
Data collection and processing						
No. of crystals used	1	1	1	1	1	1
Wavelength (Å)	0.91841	0.91841	0.91841	0.91841	0.91841	0.91841
Space group	C2	C2	C2	C2	C2	C2
Unit-cell parameters						
<i>a</i> (Å)	89.9	90.4	85.0	91.2	90.9	91.8
<i>b</i> (Å)	64.7	64.9	64.7	65.0	64.9	65.0
<i>c</i> (Å)	70.7	70.7	71.3	70.7	71.1	70.9
β (°)	96.1	96.1	94.1	96.2	96.4	96.6
Diffraction data						
Resolution range (Å)	50–1.97	50–1.53	50–1.75	30–1.48	30–1.45	30–2.06
Unique reflections	28479 (1386)	61168 (3062)	38282 (1952)	67886 (3191)	72023 (3368)	25082 (1226)
$R_{\text{merge}}(I)^{\dagger}$ (%)	10.6 (39.0)	5.9 (38.6)	10.1 (39.1)	4.3 (26.9)	3.3 (19.9)	5.9 (22.0)
Completeness (%)	99.9 (100)	99.7 (99.9)	98.2 (99.8)	98.5 (94.7)	99.0 (92.6)	97.1 (93.8)
Multiplicity	3.5 (3.4)	3.3 (3.3)	2.9 (2.9)	2.9 (2.6)	2.5 (2.7)	2.4 (2.0)
Average $I/\sigma(I)$	11.8 (2.8)	18.0 (2.7)	11.2 (3.3)	19.5 (3.5)	25.3 (3.8)	15.9 (4.0)
Refinement						
Program used for refinement	<i>SHELXL</i>	<i>SHELXL</i>	<i>SHELXL</i>	<i>SHELXL</i>	<i>SHELXL</i>	<i>SHELXL</i>
Resolution range (Å)	50–1.97	50–1.53	50–1.75	30–1.48	30–1.45	30–2.06
Reflections used in refinement	27339	58711	36905	64852	69890	24251
Final <i>R</i> values						
$R_{\text{work}} [F_o; F_o > 4\sigma(F_o)]^{\ddagger}$	18.8; 17.2	17.2; 6.2	19.5; 18.4	14.0; 13.4	17.4; 16.7	18.7; 17.5
$R_{\text{free}} [F_o; F_o > 4\sigma(F_o)]^{\S}$	24.8; 23.4	20.0; 18.8	23.2; 22.0	18.8; 18.1	20.6; 19.9	26.5; 24.4
No. of atoms (non-H)						
Protein atoms	2655	2666	2572	2699	2731	2644
Water molecules	165	255	176	303	302	188
Ligand atoms	27	33	34	31	26	20
R.m.s.d., angles (°)	1.8	2.1	1.9	2.2	2.2	1.8
R.m.s.d., bonds (Å)	0.006	0.009	0.007	0.010	0.011	0.006
Ramachandran plot \P						
Most favoured regions (%)	94.7	95.6	94.5	96.0	94.4	94.4
Additionally allowed regions (%)	4.9	4.0	5.2	3.7	5.0	5.3
Generously allowed regions (%)	0.4	0.4	0.3	0.3	0.6	0.3
Mean <i>B</i> factors (Å ²)						
Protein atoms	16.6	17.0	19.8	17.6	16.8	19.4
Water molecules	25.9	28.3	27.9	30.2	28.5	26.7
Ligand atoms	15.3	25.7	18.8	24.3	18.0	19.2

$\dagger R_{\text{merge}}(I) = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the mean of the $I_i(hkl)$ observation of reflection hkl . $\ddagger R_{\text{work}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$. $\S R_{\text{free}}$ was calculated as shown for R_{work} but on a refinement-excluded 5% of data. \P Calculated with *PROCHECK* (Laskowski *et al.*, 1993).

quality. Wild-type crystals were obtained as described above, only differing in the composition of the protein solution, which was free of DMSO.

2.2. Data collection

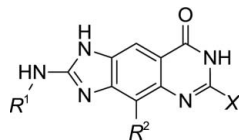
The crystals were transferred for 10 s into a cryoprotectant-containing buffer [50 mM MES pH 5.5, 300 mM NaCl, 0.5 mM DTT, 2% (v/v) DMSO, 4% (w/v) PEG 8000, 30% (v/v) glycerol]. Subsequently, the crystals were flash-cooled in liquid nitrogen. All data sets were collected under cryoconditions (100 K) on BESSY-PSF beamline 14.2 in Berlin at a wavelength of 0.91841 Å. A Rayonix MX225 CCD detector was used for data collection. The TGT crystals belonged to the monoclinic space group *C2* with one monomer per asymmetric unit. Data processing and scaling was performed using *HKL-2000* (Otwinowski & Minor, 1997). Crystal dimensions and data-collection and processing statistics are given in Table 3.

2.3. Structure determination and refinement

For all complexes, the *CNS* program package was used to perform the initial rigid-body refinement and a cycle of conjugate-gradient energy minimization, simulated annealing and *B*-factor refinement (Brünger *et al.*, 1998). The coordinates of the apo TGT structure (PDB entry 1p0d; Brenk *et al.*, 2003) were used as a starting model. Owing to the high resolution of all data collections, a further refinement using *SHELXL-97* was performed (Sheldrick & Schneider, 1997). For each refinement step, at least 20 cycles of conjugate-gradient minimization were performed with default restraints on bond distances, angles and *B* values. 5% of all data were used to calculate R_{free} . Intermittent model building was performed using *Coot* (Emsley & Cowtan, 2004). Ligand, water and glycerol molecules were placed into the difference electron density and, once assigned, included in the further refinement cycles. Riding H atoms were added to the protein in a final refinement cycle without using additional parameters. Model analysis was performed using

Table 4

Dissociation constants K_d , $\text{clog}D$ and $\text{clog}P$ values and ligand efficiencies of the bifunctionalized *lin*-benzohypoxanthines and *lin*-benzoguanines.



	X	R^1	R^2	K_d (nM)	$\text{clog}D$ (pH 7.4) [†]	$\text{clog}P$ [‡]	Ligand efficiency [‡] (kcal mol ⁻¹)
6a	NH ₂			2.3 ± 0.3	1.82	3.55	0.35
6b	NH ₂			0.8 ± 0.3	1.29	3.87	0.38
6c	NH ₂			15.1 ± 5.3	1.84	4.40	0.31
7a	H			800 ± 200	1.59	3.63	0.27
7b	H			800 ± 300	-0.58	1.77	0.25
7c	H			2000 ± 500	2.15	4.26	0.24

[†] Calculated using *ACD/Labs* v.12.01 (*ACD/Labs*, Toronto, Canada). [‡] Ligand efficiency = $RT \ln[K_d / (\text{number of non-H atoms})]$. 1 cal = 4.186 J.

PROCHECK (Laskowski *et al.*, 1993). Data-collection and refinement statistics are summarized in Table 3. Coordinates and structure factors have been deposited in the Protein Data Bank [PDB entries 4gi4 (complex with **6a**), 4gkt (**6b**), 4giy (**6c**), 4gg9 (**7a**), 4gh1 (**7b**) and 4gh3 (**7c**)].

2.4. Microscale thermophoresis measurements

Z. mobilis TGT was adjusted to a concentration of 10 μM and subsequently labelled with Alexa Fluor 647 succinimidyl ester at a concentration of 40 mg l^{-1} at room temperature for 30 min. The labelling reaction was performed in 500 mM NaCl solution buffered with 50 mM HEPES pH 8.1 (dye: protein molar ratio of ~3:1). Unreacted Alexa Fluor 647 was removed using an NAP5 Sephadex column (GE Healthcare) equilibrated with 2 M NaCl, 1 mM EDTA solution buffered with 10 mM Tris-HCl pH 7.8. The label:protein ratio was determined using photometry at 650 and 280 nm. A ratio of 0.8 was typically achieved. The Alexa Fluor 647-TGT solution was adjusted to 100 nM with 100 mM HEPES pH 7.3 buffer containing 5% DMSO and 323 μM Tween-20 (Roth). The final solution contained NaCl at a concentration of 300 mM. The ligand was dissolved in the same buffer (without NaCl) at a concentration of 1.25 μM . A series of 12 1:1 dilutions of ligand solution/buffer solution was prepared, producing ligand concentrations ranging from 1.25 μM to 305 pM (ligands **6a–6c**). In the case of ligands **7a–7c**, a series of 15 1:1 dilutions of ligand solution/buffer solution was prepared, producing ligand concentrations ranging from 10 μM to 305 pM. For thermo-

phoresis, each of these solutions was mixed with one volume of Alexa Fluor 647-TGT solution, giving a final concentration of fluorescence-labelled TGT of 10 nM (ligands **6a–6c**) or 100 nM (ligands **7a–7c**) and final ligand concentrations ranging from 153 pM to 625 nM and from 153 pM to 5000 nM, respectively. Ligand and protein concentrations were adjusted with respect to the approximate K_d of the ligand to achieve a 100% ligand-bound and a 100% ligand-unbound level. After 10 min incubation followed by centrifugation at 10 000g for 10 min, ~2 μl of each solution was placed into Monolith NT Standard treated capillaries (NanoTemper Technologies GmbH). Thermophoresis (including temperature jump) was measured at room temperature for 10 s using a Monolith NT.015 instrument (NanoTemper Technologies GmbH). Instrument parameters were adjusted with 100% LED power and 60% infrared laser power. Data from three independent measurements were averaged and analyzed using *Origin 7* (OriginLab). Curve fitting and K_d

calculations were performed based on the equation (Wienken *et al.*, 2010)

$$T = U + (B - U) \frac{(c_p + c_i + K_d) - [(c_p + c_i + K_d)^2 - 4c_p c_i]^{1/2}}{2c_p}, \quad (1)$$

where T is the thermophoresis signal, U is the minimal signal (unbound protein), B is the maximal signal (protein saturated with ligand), c_p is the concentration of labelled protein, c_i is the concentration of ligand and K_d is the dissociation constant.

3. Results and discussion

3.1. Binding affinities and physicochemical properties

The affinities of the previously reported inhibitors with substituents for either the ribose-33 or the ribose-34 pocket (Tables 1 and 2) have been described by an inhibition constant K_i , which originated from an enzymatic radiochemical assay based on the catalytic incorporation of 8-³H-guanine into tRNA (Meyer *et al.*, 2006). In the present case, it was not possible to measure the affinities of the *lin*-benzoguanines **6a–6c** with high confidence using this assay owing to their high potency. The affinity for *Z. mobilis* TGT of the ligands **6a–6c** and **7a–7c** was, however, determined using microscale thermophoresis. This method is based on the induced directed movement of particles along a local microscopic temperature gradient. A ligand-enzyme complex has different motion properties from the uncomplexed enzyme owing to differ-

ences in size, charge and hydration shell. This allows the determination of a dissociation constant K_d (Wienken *et al.*, 2010). It has been shown in previous work that the two assays provide quite comparable results and that the ratio between K_d and K_i differs maximally by a factor of 2.6 (Immekus *et al.*, 2013).

The *lin*-benzoguanines **6a–6c** have higher affinities compared with the *lin*-benzohypoxanthines **7a–7c** (Table 4),

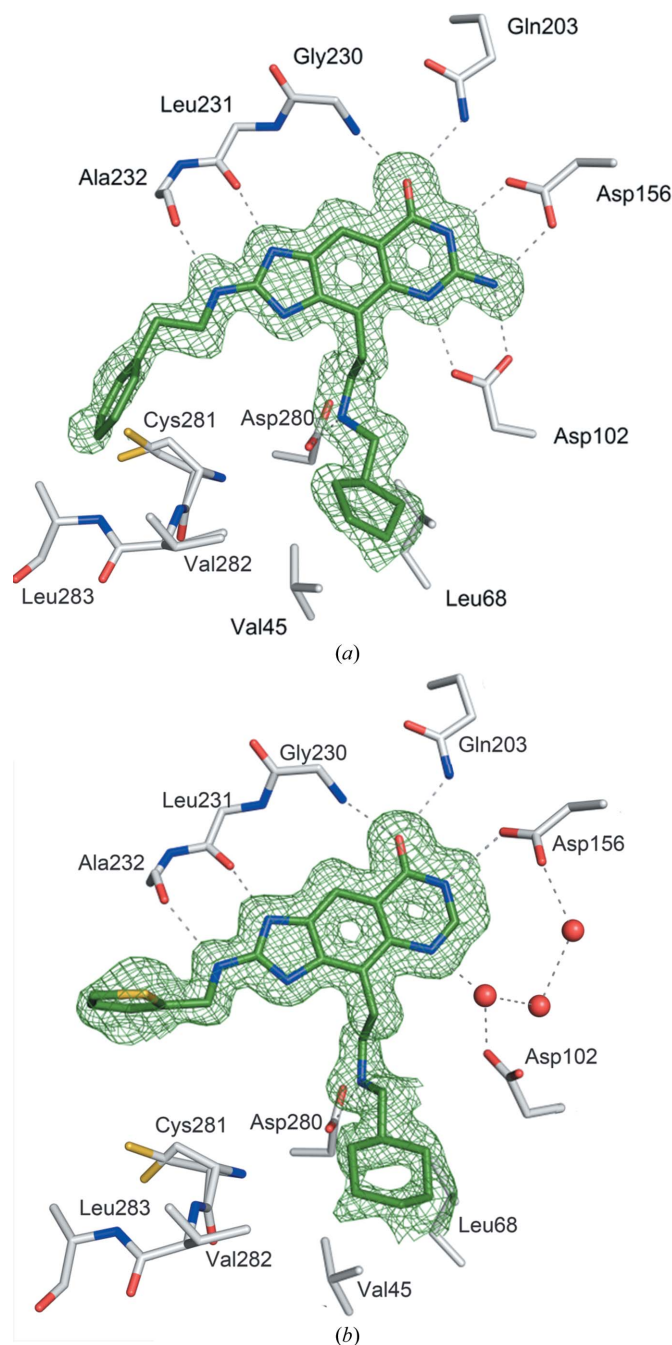


Figure 2
X-ray crystal structures of (a) **6b** (PDB entry 4gkt) and (b) **7a** (PDB entry 4gg9) bound to the active site of TGT. Hydrogen bonds are indicated as dashed lines and the difference electron density contoured at 3σ around the ligand is shown as a green mesh. The water molecules in the ribose-33 pocket are omitted for clarity (see Fig. 3 for details).

which is in agreement with our previous findings (Barandun *et al.*, 2012). Compound **6b** has a dissociation constant K_d of 0.8 ± 0.3 nM (Table 4) and represents the most potent inhibitor of *Z. mobilis* TGT that has been described in the literature to date. Replacing the cyclopentyl by a phenyl substituent (*cf.* **6b** and **6a**) reduces the affinity by a factor of three. Based on the radiochemical enzyme assay, an affinity reduction by a factor of six was observed for the similarly 4-substituted *lin*-benzoguanines filling the ribose-34 pocket and exhibiting a smaller methylamino group at C(2) (compare **4a** and **4c**; Table 2; Kohler *et al.*, 2009; Ritschel, Kohler *et al.*, 2009; Riniker *et al.*, 2012). A more pronounced affinity decrease was found for the cyclohexyl compound **6c**, where the K_d value was reduced by a factor of about 18 compared with **6b**. This trend was not observed in the previously studied 2-methylamino *lin*-benzoguanine series. The altered structure–activity relationship becomes rational when the binding modes in the cocrystal structure are analyzed (see below).

For the series of analogous *lin*-benzohypoxanthines **7a–7c** the affinity was equal or slightly better compared with their monofunctionalized analogues. The thiophene and morpholine compounds **7a** and **7b** have K_d values of 800–900 nM, whereas the phenyl derivative **7c** has a dissociation constant in the micromolar range. Obviously, the combination of the extended 2- and 4-substituents individually optimized for the series of monofunctionalized derivatives does not reveal the desired huge affinity gain; only a slight additivity of the side-chain contributions is actually experienced. In both series, however, the ligand efficiency of the bifunctionalized inhibitors is decreased compared with the monofunctionalized inhibitors. The hydrophilicity is dominated in all inhibitors by the 2-aminoimidazole moiety (Barandun *et al.*, 2012). The monofunctionalized *lin*-benzopurines **4a–4c** and **5a** and **5b** have negative $\text{clog}D$ values (calculated logarithmic distribution constant for *n*-octanol/water at pH 7.4) and $\text{clog}P$ values (calculated logarithmic partition constant for *n*-octanol/water) of 1.46–2.31 (Table 2). These values become more positive on the hydrophobic expansion of the 2-methylamino group resulting in the bifunctionalized *lin*-benzopurines **6a–6c** and **7a–7c** (Table 3).

3.2. Crystal structures obtained by cocrystallization

For all of the bifunctionalized inhibitors **6a–6c** and **7a–7c**, X-ray crystal structures in complex with *Z. mobilis* TGT were determined at resolutions between 1.45 and 2.06 Å (Table 3). The overall binding mode of the bifunctionalized *lin*-benzoguanines and *lin*-benzohypoxanthines is in good agreement with their previously reported analogues bearing an extended substituent at either C(2) or C(4) (Hörtner *et al.*, 2007; Ritschel, Kohler *et al.*, 2009; Barandun *et al.*, 2012). In the base-recognition pocket, the tricyclic 2-amino-*lin*-benzopurine core forms several hydrogen bonds to the side chains of Asp156 and Gln203 and to the peptide backbone of Gly230, Leu231 and Ala232 (Fig. 2; for hydrogen-bond distances, see Supplementary Figs. S1–S6). In addition, the *lin*-benzopurine core makes favourable π – π interactions with Tyr106 and

Met260 (not shown in Fig. 2). Based on the experimentally determined and *in silico* calculated pK_a values of the inhibitor core, it is assumed that the 2-aminoimidazole moiety is protonated (Ritschel, Hoertner *et al.*, 2009; Barandun *et al.*, 2012). The *lin*-benzoguanines differ from the *lin*-benzohypoxanthines by the presence of an exocyclic NH_2 group, which has a major influence on the pK_a value of $N(5)H^+$ and the interaction of the ligand with Asp102. A study with monofunctionalized *lin*-benzopurines indicated that $N(5)H^+$ is protonated (pK_a of 4.0–4.4) in the *lin*-benzoguanines and is unprotonated in the *lin*-benzohypoxanthines (pK_a of 1.8–2.0) (Barandun *et al.*, 2012). Consequently, the *lin*-benzoguanines

6a–6c form strong ionic hydrogen bonds to Asp102 (Fig. 2a), whereas this interaction is not present in the *lin*-benzohypoxanthines (*c.f.* Fig. 2b and Supplementary Fig. S7). In the structures with **7a–7c**, Asp102 is tilted by 90° , making space for a cluster of several water molecules that replaces the lost ligand–protein interactions (Fig. 2b).

In all crystal structures, the substituents at C(4) occupy the ribose-34 pocket, in which the protonated ethylamino linker replaces a conserved water cluster around Asp102 and Asp280 and forms a strong ionic hydrogen bond to Asp280. The adjacent hydrophobic groove of the ribose-34 pocket is filled by a cyclopentyl, cyclohexyl or phenyl substituent, which makes van der Waals interactions with Val45, Leu68 and Val282. Apart from the structure with **7c** (Supplementary Fig. S6), the substituent in the ribose-34 pocket is well resolved in all complexes. The binding modes of the ribose-34 substituents match those of the analogues that exhibit a short 2-methylamino substituent well (Supplementary Fig. S8). Only in the structures with **6a** and **6c** is a small shift of the substituent observed (*c.f.* §3.5).

The ribose-33 pocket was targeted with thiophene, morpholino and phenyl substituents departing from the methylamino group at C(2) of the *lin*-benzopurine core. While these substituents were never found with a well ordered binding pose in the monofunctionalized analogues (Ritschel, Hoertner *et al.*, 2009; Barandun *et al.*, 2012), the entire ligand was crystallographically located for the first time in the structures with **6b** (Fig. 2a and Supplementary Fig. S2), **6c** (Supplementary Fig. S3) and **7a** (Fig. 2b and Supplementary Fig. S4). In these cases, the *lin*-benzopurines with the extended substituents at C(2) and C(4) show distinct $|F_o| - |F_c|$ electron density, although the B values of the substituents are higher compared with the entire ligand (for **6b** $B_{\text{phenyl}} = 54.3 \text{ \AA}^2$, $B_{\text{ligand}} = 25.7 \text{ \AA}^2$, for **6c** $B_{\text{phenyl}} = 35.9 \text{ \AA}^2$, $B_{\text{ligand}} = 18.8 \text{ \AA}^2$ and for **7a** $B_{\text{thiophene}} = 47.2 \text{ \AA}^2$, $B_{\text{ligand}} = 24.3 \text{ \AA}^2$). This indicates an increased residual mobility or pronounced static scatter of the substituent in the ribose-33 pocket. In the crystal structures in complex with **6b** and **6c** several van der Waals interactions between the phenyl substituent and the side chains of Ala232, Cys281, Val282 and Leu283 are observed (Fig. 2a). For the thiophene substituent in the structure with **7a**, a directional S–O interaction with Tyr106 [$d(S_{\text{thiophene}} \cdots O_{\text{Tyr106}}) = 3.8 \text{ \AA}$; Supplementary Fig. S4] and a weak hydrophobic contact to Ala232 are formed (Fig. 2b).

3.3. Water cluster in the ribose-33 pocket

A crystallographically conserved water network in the ribose-33 pocket (water molecules W1–W4; for residue identifiers, see Table 5) was observed in the crystal structures with the *lin*-benzopurines (Fig. 3). While water molecules W3 and W4 were conserved in all crystal structures, W1 and W2 were not detected in every structure. In this water network, W1 forms an ionic hydrogen bond to Asp267 and a hydrogen bond to W2, which interacts with the carbonyl O atom of Gly261 (*cf.* Fig. 3 for the water network in the structures with **6b** and **7a** and Supplementary Figs. S1–S6 for hydrogen-bond distances

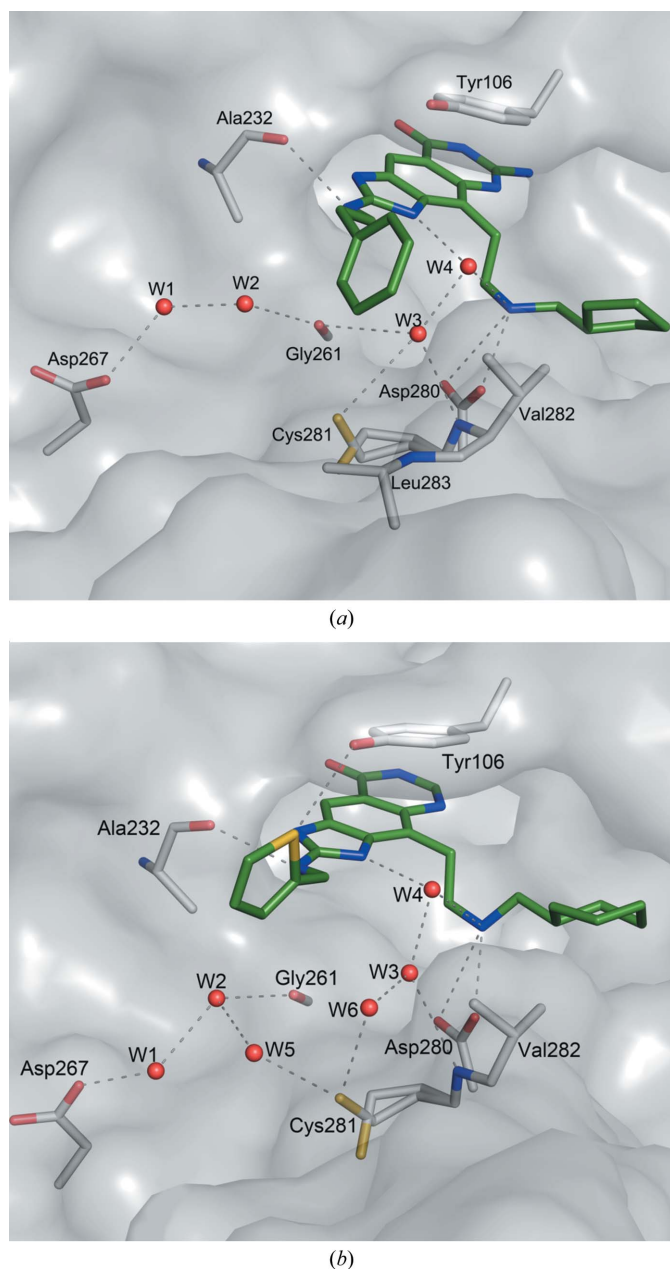


Figure 3
Selected water molecules (shown as red spheres and labelled W1–W6) in the ribose-33 pocket in the X-ray crystal structures of the complexes with (a) **6b** (PDB entry 4gkt) and (b) **7a** (PDB entry 4gg9).

Table 5

Residue identifiers in deposited coordinate files for the water molecules mentioned in the text and figures.

	PDB code	W1	W2	W3	W4	W5	W6
6a	4gi4	O627	—	O710	O671	—	—
6b	4gkt	O652	O616	O598	O565	—	—
6c	4giy	O670	O582	O642	O505	—	—
7a	4gg9	O1793	O1731	O1595	O1647	O1782	O1772
7b	4gh1	O766	O650	O514	O697	—	—
7c	4gh3	—	—	O548	O546	—	—

and the other structures). In addition, W2 is in close contact to the substituent in the ribose-33 pocket, forming weak C—H...O interactions $\{d[C_{sp2}(\text{ligand})\cdots O_{W2}] = 3.1\text{--}3.7 \text{ \AA}$ for **6b**, **6c** and **7a**). In case of TGT-**6a** and TGT-**6b**, water molecule W3 undergoes four interactions: hydrogen bonds to Cys281, to the carbonyl O atom of Gly261, to Val282 NH and to water molecule W4. The latter, W4, solvates the two protonated and charged ammonium groups $[N(3)H^+$ and $CH_2CH_2NH_2^+]$ of the ligand. For the crystal structure in complex with **7a**, two additional water molecules were localized: W5, which forms hydrogen bonds to W2 and the side chain of Cys281, and W6, which is bound to Cys281 and water molecule W3. In consequence, a slightly altered interaction pattern is observed for W3.

3.4. Effects of the crystallization protocol on the water cluster and ligand dynamics

The crystal structures with compounds **6b**, **6c** and **7a–7c** were obtained by cocrystallization, whereas the complex with **6a** was only obtained by soaking. In our previous work, it was shown that the protocol followed for crystallization has an influence on the conformation of a flexible loop–helix portion adjacent to the ribose-34 pocket (Immekus *et al.*, 2013). It adopts a well ordered geometry in the crystalline apo form of the protein (Brenk *et al.*, 2003). Ligands soaked into thus premanufactured crystals adjust their binding mode to the geometry of the ordered protein structure; in these crystals, only minor changes are observed. In contrast, if ligands with extended side chains are bound prior to crystallization and they interfere significantly with the placement of the loop–helix portion, this perturbation is also observed in the crystalline state. Thus, a remarkable influence of the crystallization protocol on the solvation pattern in the solid state and the residual mobility of the bound ligands, particularly of the phenyl substituent in the ribose-33 pocket, is observed for the bifunctionalized *lin*-benzoguanines **6a–6c**. In the complexes obtained by cocrystallization with **6b** and **6c**, the residues flanking the ribose-33 pocket adjust to the binding pose of the phenyl substituent and as a result this portion is much better defined in the difference electron density. This pronounced fixation is further enhanced through the import of water molecule W2, which additionally defines and limits the complementary space. This water molecule forms a C—H...O interaction [for **6b**, $d(C\cdots O) = 3.3 \text{ \AA}$; for **6c**, $d(C\cdots O) = 2.7 \text{ \AA}$] with the phenyl ring and makes any wiggling motion of the

phenyl substituent rather unfavourable. In contrast, the available space is larger in the soaked crystal structures and in consequence the ligand can adopt several conformations (Supplementary Fig. S9). This makes detection of the phenyl substituent and W2 in the difference electron density impossible in the soaked TGT-**6a** complex. This is also in agreement with results from our previous studies, in which the ribose-33 substituent was never located in crystal structures obtained by soaking (Hörtner *et al.*, 2007; Ritschel, Hoertner *et al.*, 2009; Barandun *et al.*, 2012).

In the case of cocrystallized complexes with the *lin*-benzohypoxanthines, only the thiophene substituent (compound **7a**), which forms an interaction with Tyr106, was observed in the difference density. It is possible that the additional contact restricts the residual mobility of the thiophene portion compared with the morpholino (compound **7b**) and phenyl (compound **7c**) substituents, which both indicate enhanced flexibility. Nevertheless, as the *lin*-benzohypoxanthines exhibit much weaker inhibitory potency, the overall fixation, and as a result the accuracy of the structure determination, is reduced.

3.5. Modulated structure–activity relationships possibly caused by conformational changes of Val45

The previously reported *lin*-benzoguanines **4a–4c** with a short 2-methylamino substituent (Kohler *et al.*, 2009) show a different affinity trend compared with the similar derivatives **6a–6c** with an extended (2-phenylethyl)amino substituent. With the small group placed into the ribose-33 pocket, the cyclopentyl derivative **4a** ($K_i = 2 \text{ nM}$) and the cyclohexyl derivative **4b** ($K_i = 4 \text{ nM}$) are the strongest, nearly equipotent inhibitors, whereas the phenyl analogue **4c** ($K_i = 25 \text{ nM}$) is significantly less potent (Table 2). This trend was rationalized by the binding modes observed in crystal structures obtained from soaking. The aliphatic cycloalkyl moieties in **4a** and **4b** penetrate deeply into the ribose-34 pocket, whereas the phenyl group in **4c** remains as a kind of lid that is less buried on top of the pocket. Remarkably, in the series **6a–6c** with the larger (2-phenylethyl)amino substituent, the cyclopentyl (**6b**; $K_d = 0.8 \pm 0.3 \text{ nM}$) and phenyl (**6a**; $K_d = 2.3 \pm 0.3 \text{ nM}$) derivatives are the most potent and the cyclohexyl derivative loses affinity substantially ($K_d = 15.1 \pm 5.3 \text{ nM}$; Table 4). A direct comparison of the binding modes of the two series **4a–4c** and **6a–6c** must take the differing crystallization protocols into consideration. The two complexes with phenyl substituents (**4c** and **6a**), which were both obtained by soaking, are directly comparable. In the case of the more potent **6a**, the phenyl ring penetrates about 1 Å deeper into the ribose-34 pocket (Supplementary Fig. S10). The enhanced contacts between the protein and ligand possibly contribute to the higher potency of this ligand.

More interesting is the observed affinity difference between the cyclopentyl and cyclohexyl derivatives in both series. The complexes with **4a** and **4b** show virtually the same affinity and highly comparable binding poses. In contrast, the slightly larger cyclohexyl ring in **6c** requires more space in the ribose-34 pocket compared with **6b** with the smaller cyclopentyl ring.

This has consequences with respect to the spatial arrangement of the adjacent Val45 residue. This amino acid is pushed out of position by the larger cyclohexyl moiety and consequently the subsequent $\beta 1\alpha 1$ loop is moved into a deviating orientation compared with the complex with the cyclopentyl derivative (Fig. 4). Presumably, the close proximity of the cyclohexyl substituent to the original position of Val45 (Fig. 4) enforces the conformational transformation, which seems to be energetically less favourable. The perturbation of the $\beta 1\alpha 1$ loop with Val45 in a toggle position by *lin*-benzoguanines bearing bulky groups at the 4-position of the cyclohexyl group has previously been observed in another study (Immekus *et al.*, 2013). In §3.3 (Fig. 3a), the conserved water molecules W3 and W4 were discussed. They fill the gap between the substituents at C(2) and C(4) and mediate interactions with the protein. It

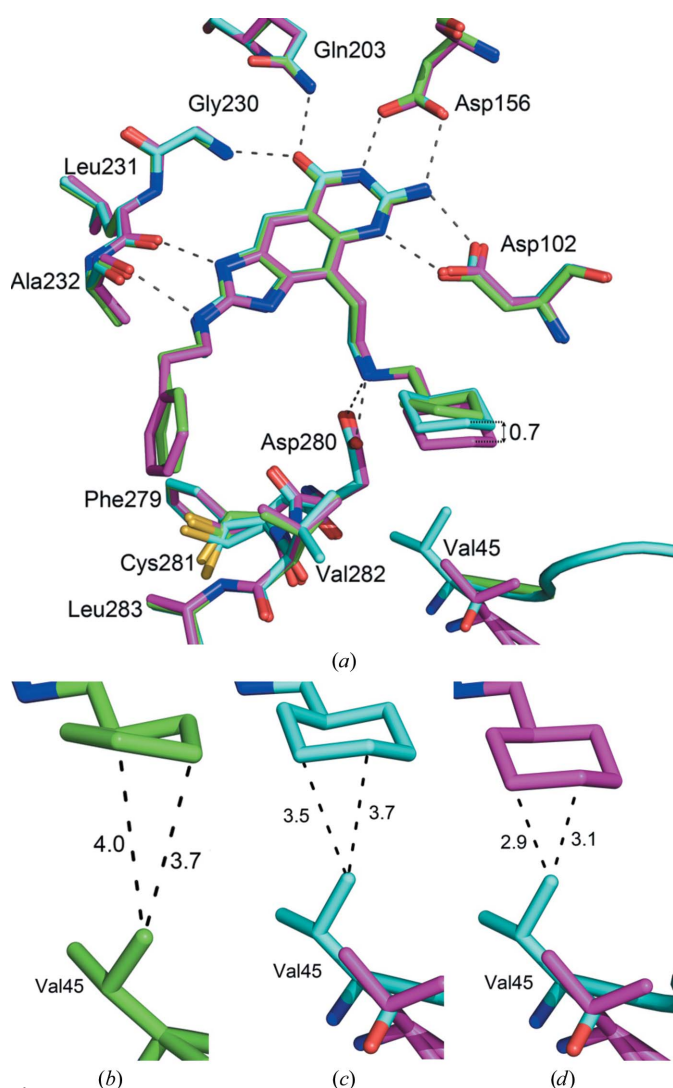


Figure 4
 (a) Overlay of the X-ray cocrystal structures of *Z. mobilis* TGT in complex with **4b** (cyan; PDB entry 3e0s; Ritschel, Kohler *et al.*, 2009), **6b** (green; PDB entry 4gkt) and **6c** (magenta; PDB entry 4giy). The loop area after Val45 is shown as a cartoon representation; for **6b**, not all amino acids are resolved in the structure. (b, c) van der Waals contacts of (b) **6b** and (c) **4b** with Val45. (d) Unfavourable proximity of **6c** to the original position of Val45 (cyan). Distances are given in Å.

remains speculative whether these interstitial water molecules contribute to a kind of cross-talk between substituents of deviating size at the two positions. Clearly, a 2-methylamino substituent leaves a larger space than a (2-phenylethyl)amino substituent. Small structural differences could translate into larger conformational changes also involving the adjacent protein residues. However, the resulting packing effects are difficult to compare as two different crystallization protocols have been applied that make an impact on the molecular interactions in this region (Supplementary Fig. S9).

4. Conclusions

In this study, the design, synthesis, biological evaluation and crystallographic analysis of two new series of inhibitors of tRNA-guanine transglycosylase are reported. The bifunctionalized inhibitors feature a central *lin*-benzoguanine or *lin*-benzohypoxanthine core decorated with two substituents targeting the ribose-33 and ribose-34 pockets. While the filling of the individual pockets has been optimized independently and characterized in previous studies, the joint filling of the two pockets was explored for the first time in this work. An improvement in affinity is observed; however, it is below the extent expected from sole additivity considerations. Also, the structure-activity relationship attributed to the substituents of deviating size is not parallel in the two series. As the structures indicate, a cross-talk between the attached substituents is observed resulting from small differences in steric demand. This translates into induced-fit adaptations of the protein and is also possibly transmitted through interstitial water molecules. Nonetheless, very potent inhibitors with K_d values of 0.8–15.1 nM for the *lin*-benzoguanine derivatives and of 800–2000 nM for the *lin*-benzohypoxanthine derivatives were obtained. For the first time, the substituents filling the larger ribose-33 pocket were fully located in three crystal structures obtained by cocrystallization. Surprisingly, a strong influence of the crystallization protocol was observed. For the cocrystallized structures, the enzyme pocket can adapt to the ligand and incorporates water molecules. In contrast, soaking leads to structures that show only partly defined difference electron density for some ligand portions, resulting either from enhanced residual mobility or pronounced static scatter over multiple orientations in the binding pocket.

We are grateful to the ETH Research Council, F. Hoffmann-La Roche (Basel) and Chugai Pharmaceuticals for their generous support of the work performed at ETH. We are grateful for the generous support of the Deutsche Forschungsgemeinschaft FO 806, KL1204/13-1. We thank Elke Persch (ETH) for correcting the manuscript and Dr Bruno Bernet (ETH) for proofreading the experimental part. We further acknowledge the support of the beamline staff at BESSY II in Berlin and a travel grant from the Helmholtz-Zentrum für Materialien und Energie in Berlin.

References

Ashkenazi, S., Levy, I., Kazaronovski, V. & Samra, Z. (2003). *J. Antimicrob. Chemother.* **51**, 427–429.

- Barandun, L. J., Immekus, F., Kohler, P. C., Tonazzi, S., Wagner, B., Wendelspiess, S., Ritschel, T., Heine, A., Kansy, M., Klebe, G. & Diederich, F. (2012). *Chem. Eur. J.* **18**, 9246–9257.
- Brenk, R., Stubbs, M. T., Heine, A., Reuter, K. & Klebe, G. (2003). *ChemBioChem*, **4**, 1066–1077.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* **D54**, 905–921.
- Durand, J. M., Okada, N., Tobe, T., Watarai, M., Fukuda, I., Suzuki, T., Nakata, N., Komatsu, K., Yoshikawa, M. & Sasakawa, C. (1994). *J. Bacteriol.* **176**, 4627–4634.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Gerber, P. R. & Müller, K. (1995). *J. Comput. Aided Mol. Des.* **9**, 251–268.
- Grädler, U., Gerber, H.-D., Goodenough-Lashua, D. M., Garcia, G. A., Ficner, R., Reuter, K., Stubbs, M. T. & Klebe, G. (2001). *J. Mol. Biol.* **306**, 455–467.
- Hörtner, S. R., Ritschel, T., Stengl, B., Kramer, C., Schweizer, W. B., Wagner, B., Kansy, M., Klebe, G. & Diederich, F. (2007). *Angew. Chem. Int. Ed.* **46**, 8266–8269.
- Immekus, F., Barandun, L. J., Betz, M., Debaene, F., Petiot, S., Sanglier-Cianferani, S., Reuter, K., Diederich, F. & Klebe, G. (2013). *ACS Chem. Biol.*, **8**, 1163–1178.
- Iwata-Reuyl, D. (2003). *Bioorg. Chem.* **31**, 24–43.
- Kohler, P. C., Ritschel, T., Schweizer, W. B., Klebe, G. & Diederich, F. (2009). *Chem. Eur. J.* **15**, 10809–10817.
- Kubler-Kielb, J., Vinogradov, E., Mocca, C., Pozsgay, V., Coxon, B., Robbins, J. B. & Schneerson, R. (2010). *Carbohydr. Res.* **345**, 1600–1608.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Launay, O. *et al.* (2009). *Vaccine*, **27**, 1184–1191.
- Levine, M. M., Kotloff, K. L., Barry, E. M., Pasetti, M. F. & Sztein, M. B. (2007). *Nature Rev. Microbiol.* **5**, 540–553.
- Meyer, E., Donati, N., Guillot, M., Schweizer, W., Diederich, F., Stengl, B., Brenk, R., Reuter, K. & Klebe, G. (2006). *Helv. Chim. Acta*, **89**, 573–597.
- Niyogi, S. K. (2005). *J. Microbiol.* **43**, 133–143.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Ranallo, R. T., Fonseka, S., Boren, T. L., Bedford, L. A., Kaminski, R. W., Thakkar, S. & Venkatesan, M. M. (2012). *Vaccine*, **30**, 5159–5171.
- Reuter, K. & Ficner, R. (1995). *J. Bacteriol.* **177**, 5284–5288.
- Riniker, S., Barandun, L. J., Diederich, F., Krämer, O., Steffen, A. & van Gunsteren, W. F. (2012). *J. Comput. Aided Mol. Des.* **26**, 1239–1309.
- Ritschel, T., Atmanene, C., Reuter, K., Van Dorsselaer, A., Sanglier-Cianferani, S. & Klebe, G. (2009). *J. Mol. Biol.* **393**, 833–847.
- Ritschel, T., Hoertner, S., Heine, A., Diederich, F. & Klebe, G. (2009). *ChemBioChem*, **10**, 716–727.
- Ritschel, T., Kohler, P. C., Neudert, G., Heine, A., Diederich, F. & Klebe, G. (2009). *ChemMedChem*, **4**, 2012–2023.
- Romier, C., Ficner, R., Reuter, K. & Suck, D. (1996). *Proteins*, **24**, 516–519.
- Sansonetti, P. J. (2001a). *FEMS Microbiol. Rev.* **25**, 3–14.
- Sansonetti, P. J. (2001b). *Am. J. Physiol. Gastrointest. Liver Physiol.* **280**, G319–G323.
- Sansonetti, P. J. (2006). *PLoS Med.* **3**, e354.
- Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
- Stengl, B., Meyer, E. A., Heine, A., Brenk, R., Diederich, F. & Klebe, G. (2007). *J. Mol. Biol.* **370**, 492–511.
- Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. (2010). *Nature Commun.* **1**, 100.
- Xie, W., Liu, X. & Huang, R. H. (2003). *Nature Struct. Biol.* **10**, 781–788.