Bioorganic & Medicinal Chemistry xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Occupying a flat subpocket in a tRNA-modifying enzyme with ordered or disordered side chains: Favorable or unfavorable for binding?

Manuel Neeb [a,†], Christoph Hohn [b,†], Frederik Rainer Ehrmann [a], Adrian Härtsch [b], Andreas Heine [a], François Diederich [b], Gerhard Klebe [a,*]

[a] Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany
[b] Laboratorium für Organische Chemie, ETH Zurich, Vladimir-Prelog-Weg 3, 8093 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 27 May 2016
Accepted 24 July 2016
Available online xxxx

Keywords:
Drug design
Protein–ligand complexes
Crystallography
Flat-binding pockets
Order-disorder phenomena
Residual mobility

ABSTRACT

Small-molecule ligands binding with partial disorder or enhanced residual mobility are usually assumed as unfavorable with respect to their binding properties. Considering thermodynamics, disorder or residual mobility is entropically favorable and contributes to the Gibbs energy of binding. In the present study, we analyzed a series of congeneric ligands inhibiting the tRNA-modifying enzyme TGT. Attached to the parent *lin*-benzoguanine scaffold, substituents in position 2 accommodate in a flat solvent-exposed pocket and exhibit varying degree of residual mobility. This is indicated in the crystal structures by enhanced B-factors, reduced occupancies, or distributions over split conformers. MD simulations of the complexes suggest an even larger scatter over several conformational families. Introduction of a terminal acidic group fixes the substituent by a salt-bridge to an Arg residue. Overall, all substituted derivatives show the same affinity underpinning that neither order nor disorder is a determinant factor for binding affinity. The additional salt bridge remains strongly solvent-exposed and thus does not contribute to affinity. MD suggests temporary fluctuation of this contact.

1. Introduction

Concepts in rational drug design try to optimize a given ligand scaffold to accommodate and interact optimally with a certain target protein. Our hypotheses and thus the followed design strategies are usually strongly biased by the idea to achieve perfect complementarity between protein-binding pocket and ligand in terms of geometry and experienced interaction patterns. We assume that this strategy leads to energetically favored interactions and a strong binding affinity. By contrast, any residual mobility or intrinsic flexibility of the bound ligand are intuitively assumed as unfavorable for binding. However, binding affinity is a Gibbs energy value and thus composed by an enthalpic and entropic component. With respect to a favorable entropic component to binding, residual mobility or pronounced ligand disorder can be beneficial as a smaller number of degrees of freedom are lost by the ligand upon complex formation.

Residual mobility of ligand side chains or attached substituents is usually observed in flat, solvent-exposed binding pockets where significant contacts to the neighboring solvent environment are given. In a congeneric series of thermolysin inhibitors with varying P_2' substituents of increasing size and hydrophobicity, we observed on the one hand a growing influence imposed by the structural arrangement of the surrounding first and second solvent layer but on the other hand by an enhanced residual flexibility particularly given for larger substituents.[1,2]

As a second example, we studied the binding of potent inhibitors blocking the function of the tRNA-modifying enzyme tRNA-guanine transglycosylase (TGT) by attaching hydrophobic substituents of growing size at the parent *lin*-benzoguanine scaffold.[3–7] The target protein performs a complete nucleobase exchange at the wobble position of specific tRNAs.[8] Inhibition of bacterial TGT may serve as a putative therapeutic concept for drug development as its function has been linked to the pathogenicity of *Shigella*, the causative agent of bacterial dysentery.[9] In developing countries, Shigellosis is a major threat as infections with these bacteria are responsible for more than 100,000 lethal cases every year.[10–12] In previous studies, we attached substituents at the 2- and 4-position of the parent scaffold and in combination of both.[7]

* Corresponding author.
E-mail address: klebe@mailer.uni-marburg.de (G. Klebe).
† Equally contributing authors.

http://dx.doi.org/10.1016/j.bmc.2016.07.053
0968-0896/

The 2-substituents orient into an open, bowl-shaped pocket that recognizes uracil 33 with its adjacent ribose ring of the natural substrate tRNA.[13] In addition, we have developed derivatives, exhibiting substituents at the 4-position, designed to occupy the ribose-34 pocket. In the subsequently determined crystal structures, the position of the 4-substituent could be assigned unequivocally to the difference electron density, while the substituents, anchored at C2, indicated in most of the determined crystal structures enhanced mobility or at least a scatter across multiple conformational states to varying extend.[4,7,13,14] This observation suggests that the 2-substituent, depending on its chemical composition, experiences pronounced flexibility in the bound state.

Closer inspection of the architecture of the uracil-33 pocket indicates the exposure of polar and non-polar residues in this binding cavity. In consequence, *lin*-benzoguanines with alternative 2-substituents were synthesized following the design hypothesis to fix the regarded ligand portion in the uracil-33 pocket by addressing the available polar contacts of the protein. In the present study, we investigated ligands exhibiting an extended phenethyl substituent in 2-position by molecular dynamics (MD) simulations, crystal structure analysis, isothermal titration calorimetry (ITC), and by a radioactive washout-binding assay. At their terminal end, the phenethyl substituents were decorated with different polar functional groups, principally capable to establish the desired contacts with the protein. The consequences of these modifications are described with respect to the obtained binding poses, dynamics, and energetic properties.

2. Results

2.1. Ligand design

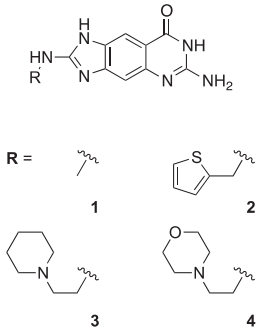
We recently reported on the crystal structures of the parent 2-methylamino-*lin*-benzoguanine (**1**), the 2-thienylmethyleneamino (**2**), 2-piperidylethylenamino- (**3**), and 2-morpholinoethylenamino (**4**) derivatives (Scheme 1).[13,14]

All crystal structures showed the tricyclic ring system in ordered state with full occupancy. In the complex crystal structure with **1**, a network of five water molecules is found wrapping around the terminal methyl group. The two derivatives **3** and **4** with an ethylene linker show the 2-substituent in at least two conformations whereas the thienylmethyl substituent in **2** is scattered over multiple arrangements. Possibly, the ethylene linker is better suited to place the substituent in ordered fashion. We therefore selected **5** decorated with a phenethylamino moiety as a starting point for our ligand design (Table 1). Different to **3** and **4**, the phenethyl substituent does not bear an additional protonation site that may potentially lower the permeation through cell membranes at a physiological pH value and would thus contribute unsatisfactorily to bioavailability.[14,15]

This aspect is of some importance as we planned to attach in *para*-position of the terminal phenyl ring substituents of different polar and hydrogen-bonding character to interact with the polar functional groups of amino acid residues exposed to the uracil-33 pocket.

2.2. Synthesis

The synthesis route of 2-substituted *lin*-benzoguanines has been described by us[15,16] and was successfully applied to the preparation of **6a–d** (Table 1). The last step of the synthesis requires relatively high temperature (150 °C) to cyclize the sulfamoyl-protected aminobenzimidazoles with chloroformamidinium hydrochloride (Schemes 2 and 3). Simultaneously, this temperature and the generation of two equivalents of HCl led to



Scheme 1. Chemical formulas of the investigated ligands **1–4**.

Table 1
Chemical formulae and binding affinities of the extended 2-amino-*lin*-benzoguanines **5**, **6a–d**

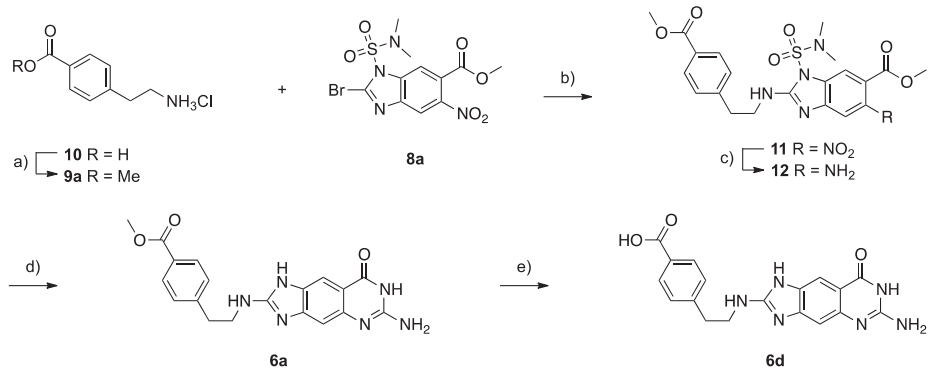
Compound		K_i [nM]
5		10 ± 3[a]
6a		2 ± 2
6b		7 ± 3
6c		21 ± 8
6d		30 ± 7

[a] K_i value according to Lit. 6.

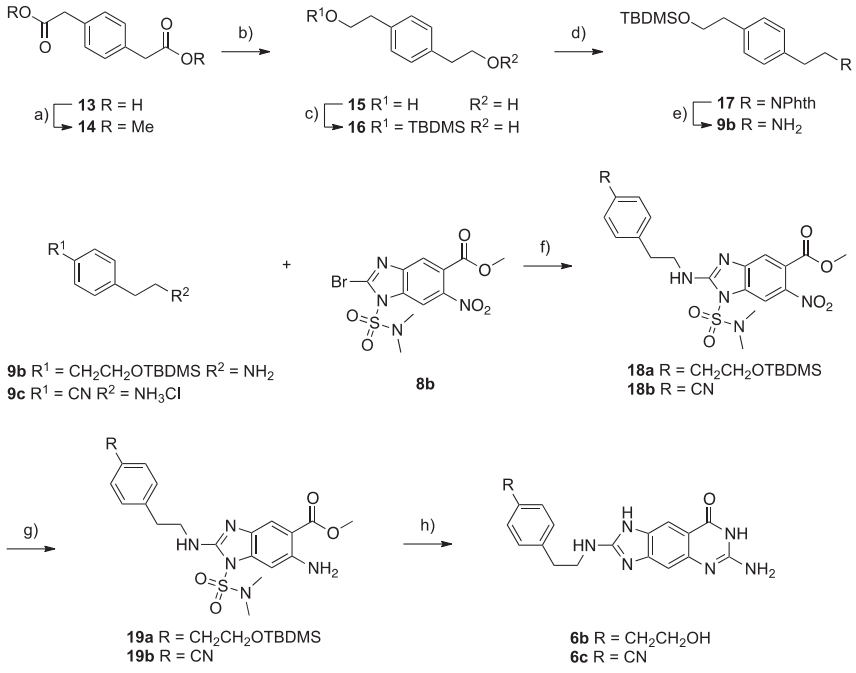
the cleavage of the sulfamoyl protection group. One possible explanation for the poor-yielding last step (7–31%) is the observed low solubility. The *lin*-benzoguanines as extended version of guanine are nearly insoluble in most organic solvents (e.g. THF, MeCN, acetone, CH_2Cl_2, ethyl acetate) and H_2O, which results in a difficult purification process. Canonical nucleobases are known to exhibit poor solubility due to their ability to form a strong hydrogen bond network. In the case of *lin*-benzoguanines, the extension of the ring system probably further decreases the solubility by strong intermolecular π–π stacking interactions. Purification of *lin*-benzoguanines either by silica-based normal-phase HPLC or recrystallization in several solvent systems failed. RP-HPLC with C18-coded silica and H_2O/MeCN with either 0.1 vol % NH_4OAc or NH_4OH as solvent system enabled the isolation of pure samples of the products.

Starting from the commercially available benzimidazole-5-carboxylic acid (**7**, not shown), the synthesis of the regioisomeric 2-bromobenzimidazoles **8a** and **8b** (Scheme S1 in the Supplementary material) followed the sequence esterification, nitration, sulfamoyl protection, and bromination (Schemes 2 and 3). Both brominated regioisomers **8a** and **8b** were used for the synthesis of the *lin*-benzoguanines **6a–d**.

Substitution of the 2-bromobenzimidazole-6-carboxylate **8a** with primary ammonium chloride **9a**, which was prepared by acid-catalyzed esterification of the corresponding acid **10** in methanol,[17] gave the methylated 2-amino-5-nitrobenzimidazole **11** (Scheme 2). The solubility of the bromide in EtOAc was increased by the addition of methanol, enhancing the velocity of the reaction. The 5-nitrobenzimidazole **11** was subsequently reduced to the corresponding amine **12**, which showed some



Scheme 2. Synthesis of elongated *lin*-benzoguanines **6a**,**d** starting from methyl 2-bromobenzimidazole-6-carboxylate **8a**: (a) HCl, MeOH, 1,4-dioxane, reflux, 16 h, 68%; (b) **9a**, NEt_3, EtOAc, MeOH, 23 °C, 16 h, 86%; (c) Zn powder, AcOH, H_2O, 23 °C, 3 h; (d) $ClC(NH)NH_3Cl$, $(CH_3)_2SO_2$, 150 °C, 5 h, 13% (from **11**); (e) NH_4OH, 120 °C, 1.5 h, 16%.



Scheme 3. Synthesis of the *lin*-benzoguanines **6b** and **6c** starting from methyl 2-bromobenzimidazole-5-carboxylate **8b**: (a) MeOH, H_2SO_4, 80 °C, 94 h, 90%; (b) $LiAlH_4$, THF, 23 °C, 16 h, 94%; (c) TBDMSCl, I_2, 1-methylimidazole, MeCN, 23 °C, 1 h, 37%; (d) HNPhth, PPh_3, DIAD, THF, 23 °C, 16 h, 83%; (e) N_2H_4, EtOH, 90 °C, 16 h, 69%; (f) $EtN(iPr)_2$, EtOAc, 23 °C, 16 h, **18a**: 49%, **18b**: 74%; (g) Zn powder, AcOH, H_2O, 23 °C, 3–5 h; **19a**: 50%; **19b**: 44%; (h) $ClC(NH)NH_3Cl$, $(CH_3)_2SO_2$, 150 °C, 4–5 h, **6b**: 7%, **6c**: 31%. TBDMS: *tert*-butyldimethylsilyl; DIAD: diisopropyl azodicarboxylate; Phth: phthaloyl.

instability upon light and air exposure, and was directly used without purification for the cyclization with chloroformamidinium hydrochloride ($ClC(NH)NH_3Cl$). At elevated temperatures, *lin*-benzoguanine **6a** was prepared in 13% yield over two steps. Treatment of the ester **6a** with aqueous ammonia at 120 °C gave 16% of the acid **6d** after RP-HPLC purification.

The 2-bromobenzimidazole-5-carboxylate **8b** was used as the starting material for the preparation of the *lin*-benzoguanines **6b** and **6c**, which contain the ethanol and nitrile functionalities, respectively (Scheme 3). Coupling partners were the commercially available cyanophenethylammonium chloride (**9c**) and the protected amino alcohol **9b**, which was prepared from 1,4-phenylenediacetic acid (**13**) by acid-catalyzed esterification to bis-ester **14**,[18] reduction to the diol **15**,[19] partial silylation to **16**,[20] substitution to the phthalimide **17**, and dephthaloylation (Gabriel synthesis).

The coupling of **8b** with **9b** and **9c** was done in the presence of *N*,*N*-diisopropylethylamine in EtOAc at room temperature and led to a complete consumption of **8b** affording 49% of **18a** and 74% of **18c** (Scheme 3). The reduction of the nitro derivatives **18a** and **18b** with zinc powder in AcOH gave the amines **19a** and **19b**, which were sensitive to light and air exposure; the higher stability of **19a** allowed its purification, a longer storage time, and a full characterization. The above-mentioned cyclization with chloroformamidinium hydrochloride provided the *lin*-benzoguanines **6b** (7%) and **6c** (31%).

2.3. Binding affinity

Binding affinities were determined by a radioactive washout assay. Thereby, the incorporation of [8-^{3}H]guanine into $tRNA^{Tyr}$ (ECY2) at position G34 is measured by liquid scintillation counting at pH 7.3 and 37 °C (see Section 5). Inhibition constants are calculated by the comparison of the initial velocities of the base-exchange reaction in absence and presence of the ligand. The ligands **1**–**5** and **6a**–**d** show a binding affinity (K_i value) in the low one- to two-digit nanomolar range (Table 1). Compared to

the unsubstituted reference ligand **1**, the phenethyl derivative **5** is about five times more potent and exhibits a binding affinity of about 10 nM.[6] Interestingly, the binding affinity of methyl ester **6a** is enhanced by a factor of five. In **6b**, further decoration of the phenyl moiety does not yield a significant change in K_i and in contrary, **6c** and **6d** are slightly decreased in potency.

2.4. Crystal structure analysis

For the investigated ligands, crystal structures of the complexes have been determined with a resolution of 1.14–1.40 Å. In all structures, the fully occupied *lin*-benzoguanine scaffold is well-defined in the difference electron density (Fig. 1). It is placed into the guanine-34 recognition pocket undergoing π-stacking with the side chains of Tyr106 and Met260 and establishing the same hydrogen-bonding interactions to neighboring amino acids as previously described.[3–7]

The difference electron density $|F_o| - |F_c|$ fully defines the binding mode of **5** (Fig. 1A) including the phenethyl substituent. Nevertheless, in this complex larger B-factors are observed for the 2-substituent compared to the tricyclic parent core scaffold ($B_{substituent} = 34.1$ Å^2 vs $B_{core} = 17.8$ Å^2) (for definition of core and substituent, see Table S1 in the Supplementary material) indicating enhanced residual mobility of this moiety. The phenethyl substituent occupies the uracil-33 pocket adopting an all-*trans* conformation. Weak hydrophobic interactions are experienced between the side chains of Ala232, Cys281, Val282, Leu283, and the 2-substituent which covers in lid-like fashion the hydrophobic residues in this pocket.

The 2-substituent of **6a** refines to a reduced occupancy of 75% indicating higher flexibility of this moiety compared to that observed for **5** ($B_{substituent} = 30.7$ Å^2 vs $B_{core} = 15.1$ Å^2). The C–CH_2–CH_2–NH linker adopts an energetically most likely less favorable *gauche* conformation with a torsion angle of −61.3° (Fig. 1B). Contrary to the *gauche* conformation of the 2-substituent in **4**,[13] the one in **6a** adopts a conformation which orients the methyl benzoate moiety out of the uracil-33 pocket facing the backbone of Val233 and Gly234. Obviously, the extended all-*trans* conformation of this substituent is not adopted, as the ester moiety would interfere with Arg286. Instead, the substituent prefers to fold back into the described upwards conformation moving the ligand out of the uracil-33 pocket. Weak van-der-Waals interactions are formed to the side chain of Ala232. Additionally, a weak hydrogen-bond contact between the backbone NH group of Gly234 and the carbonyl oxygen of the ester group in **6a** is formed (3.5 Å). Remarkably, this back-folded conformation is observed for the first time for an attached 2-substituent.

For **6b** two data sets collected on different crystals show deviating results with respect to the substituent's binding mode (Fig. 1C). Refinement of the first structure suggests the 2-substituent to be present in two conformations with a summed overall occupancy of 100%. With increasing distance from the parent scaffold, the difference electron density becomes more blurred and accordingly the assigned temperature factors increase. In consequence, for both conformers, the substituent is not sufficiently resolved in the difference electron density due to the residual mobility of that portion of the molecule resulting in a partly defined difference electron density for the hydroxyethyl moiety. In the first conformer, the C–CH_2–CH_2–NH linker adopts an all-*trans* conformation with an occupancy of 56% and enhanced residual mobility ($B_{substituent} = 21.6$ Å^2 vs $B_{core} = 12.3$ Å^2). The adopted orientation is similar to that observed for **5**. Interestingly, the neighboring side chain of Arg286 is shifted out of the uracil-33 pocket presumably to create sufficient space to accommodate the hydroxyethyl moiety of the ligand, now extending its binding pose even deeper into the pocket. In the second conformer, the linker exhibits a *gauche* conformation resulting in an analogous binding mode as observed exclusively for **6a** ($B_{substituent} = 21.0$ Å^2 vs $B_{core} = 12.3$ Å^2). In this orientation, the substituent refines to 44% occupancy.

We studied a second, independently collected data set using another crystal of this complex. In the diffraction data taken from the second crystal, the ligand is visible in the difference electron density in only the all-*trans* conformation, however, a slightly reduced occupancy for the 2-substituent of 84% results from refinement. Also in this structure enhanced mobility of the 2-substituent is experienced ($B_{substituent} = 27.3$ Å^2 vs $B_{core} = 13.4$ Å^2) and the position of the hydroxyethyl portion can hardly be assigned to the density. Further analysis of the difference density does not allow reliable assignment of the substituent to a second orientation as in the first data set.

The 2-substituent of **6c** also adopts an all-*trans* conformation penetrating into the adjacent ribose-32 subpocket (Fig. 1D). The entire ligand is fully occupied, nevertheless, its 2-substituent shows increased temperature factors compared to the parent scaffold ($B_{substituent} = 18.3$ Å^2 vs $B_{core} = 10.2$ Å^2). The terminal nitrile group does not experience specific interactions with any of the amino acid residues found in the ribose-32 subpocket. Only a single water molecule is located in close neighborhood (3.0 Å) to the nitrile functional group. The side chain of Arg286 has to reshuffle its orientation and it is shifted out of the subpocket adopting the previously observed conformation.

The binding mode of **6d** is fully defined in the difference electron density $|F_o| - |F_c|$ (Fig. 1E). The C–CH_2–CH_2–NH linker adopts the all-*trans* conformation pointing into the uracil-33 subpocket. Under the applied pH conditions (pH 7.8), the terminal carboxylate group is most likely deprotonated. In this state, it experiences a bidentate salt-bridge to the guanidinium moiety of Arg286. In order to establish this contact, the polar group of Arg286 bends toward the ribose-32 subpocket. Despite of the strong electrostatic interaction between the carboxylate group of the ligand and the side chain of Arg286, the 2-substituent exhibits similarly increased B-factors like the other studied derivatives ($B_{substituent} = 22.4$ Å^2 vs $B_{core} = 11.5$ Å^2). For comparison, the binding mode of the natural substrate tRNA is shown in Figure 1F. The bound substrate induces the same conformation of Arg286 as found in the complexes TGT·**6b** and TGT·**6c**. It opens an additional pocket (highlighted in light blue) to accommodate the ribose moiety of base at position 32 and forms two charge-assisted hydrogen bonds to the guanidinium group of Arg286.

2.5. MD simulations

The crystal structures of **6a–d** indicate distinct residual mobility of the considered substituents directed into the uracil-33 pocket. As indicated by the two data sets recorded for two distinct crystals of TGT·**6b**, the observed disorder may also depend on the crystallization protocol used to grow the crystals. We therefore performed molecular dynamic simulations using the program *AMBER*[21] under NTP conditions (constant number of particles, constant pressure of p = 1 bar, constant temperature of $T = 300$ K) to assess whether qualitatively the same structural properties are indicated by the computer simulations as in the crystalline state. All four complexes of **6a–d** were simulated over a productive simulation time to 100 ns. Subsequently, the ligand conformations visited along the trajectory were hierarchically clustered into families with maximal RMSDs of 2 Å using the program *ptraj* (AMBER program suite) and the representative archetypes were visually inspected. The number of observed conformation families is listed in Table 2 and the spatial structure of some family representatives are given in Figure 2. In the Supplementary material, images of all conformer family members can be found (Fig. S1).

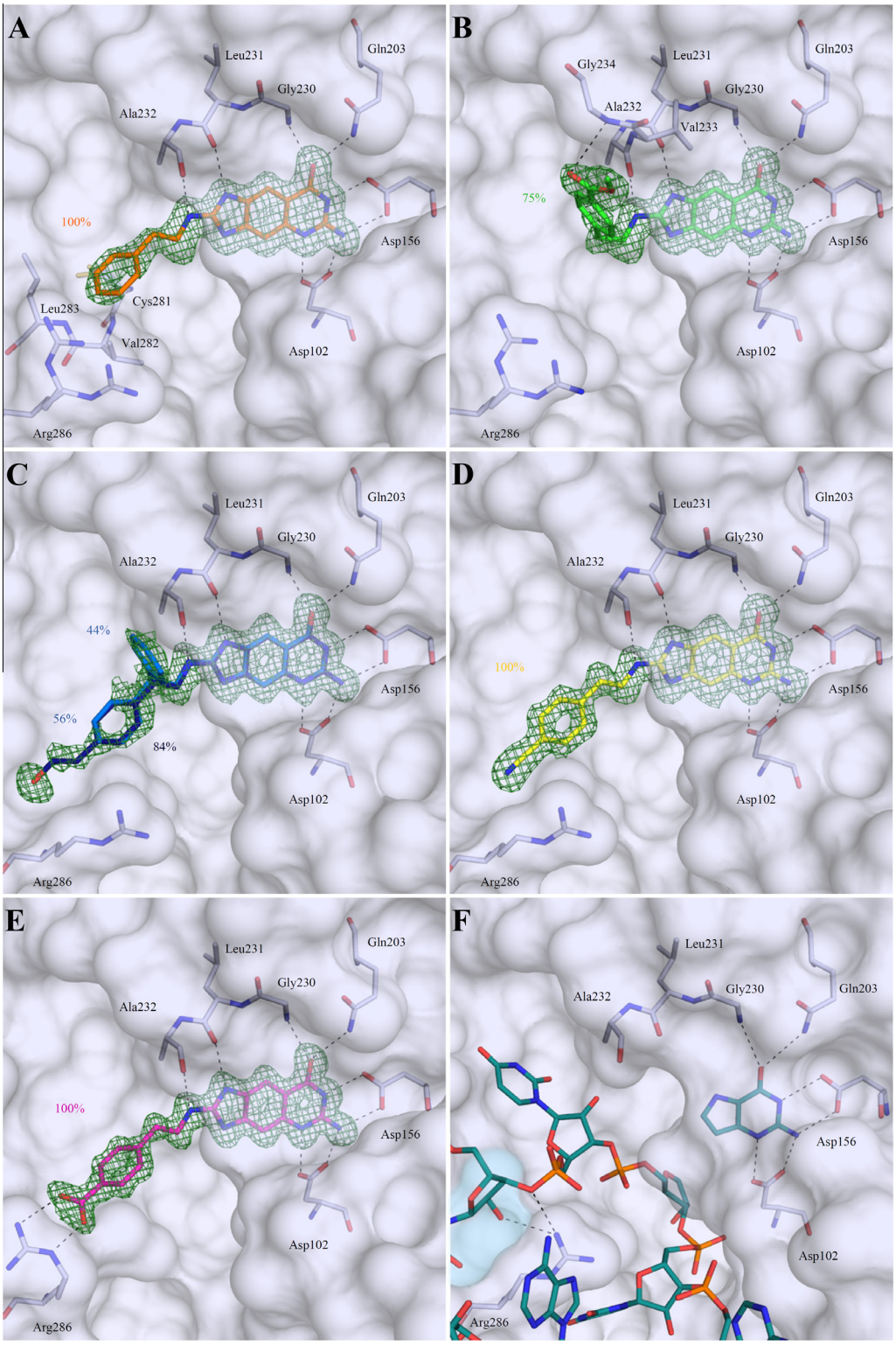


Figure 1. Binding modes of the analyzed *lin*-benzoguanines as well as a tRNA analogue. The solvent-accessible protein surface is shown in white. The ligands and interacting residues are displayed as sticks (nitrogen = blue, oxygen = red, phosphorus = orange). Hydrogen bonds are visualized as dashes. The $|F_o| - |F_c|$ difference electron density is illustrated at a *sigma*-level of $\sigma = 2.5$ as green meshes. For clarity, the π-stacking residues Tyr106 and Met260 as well as water molecules are not shown. A–E) The tricyclic parent scaffold is well-defined in the difference electron density experiencing hydrogen bonds to residues Asp102, Asp156, Gln203, Gly230, Leu231 and Ala232 within the guanine-34 recognition site. Depending on its substitution pattern, the phenethyl substituent in 2-position adopts either an orientation within the uracil-33 pocket or is rotated out of this position now facing the backbone of Val233 and Gly234. Only **6d** (E) forms additional polar interactions to the guanidinium head group of Arg286. While in the complex structures TGT·**5** (A), TGT·**6a** (B) and TGT·**6d** (E) the side chain of Arg286 closes the preceding pocket at position 32, **6b** (C) and **6c** (D) bind to the site by shifting the side chain of Arg286 apart. For TGT·**6b** (C), the results collected at two independent crystals are shown. In one crystal structure (ligand C-atoms dark blue), an expended conformation with an occupancy of 84% of the substituent was found. In the second crystal (ligand C-atoms light blue), a scatter over two orientations (extended and back-folded) with occupancies of 56% and 44% could be refined. (F) Binding of the natural substrate tRNA (PDB ID 1Q2R) induces the same conformation of Arg286 as found in the complexes TGT·**6b** and TGT·**6c** opening the pocket accommodating the ribose moiety of base at position 32 (light blue). The ribose moiety at position 32 interacts with the side chain of Arg286 via two charge-assisted hydrogen bonds.

Please cite this article in press as: Neeb, M.; et al. *Bioorg. Med. Chem.* (2016), http://dx.doi.org/10.1016/j.bmc.2016.07.053

Table 2
Occupancy of the different conformer families in percentage found as clusters along the MD trajectory of **6a–d** (calculated by *ptraj* after hierarchical clustering)

Ligand	# of cluster	Occurrence [%]	Ligand	# of cluster	Occurrence [%]
6a	1[a]	**16.4**[b]	6c	1	2.5
	2	6.6		**2**	**76.1**
	3	0.8		3	3.1
	4	0.4		4	5.9
	5	1.0		5	8.5
	6	8.1		6	3.8
	7	0.8			
	8	**19.0**	6d	**1**	**12.2**
	9	**11.6**		**2**	**11.7**
	10	5.8		3	5.0
	11	7.8		**4**	**10.6**
	12	9.6		5	1.6
	13	5.9		6	3.5
	14	6.0		7	4.7
				8	4.1
6b	**1**	**35.6**		**9**	**17.3**
	2	2.8		10	1.8
	3	0.2		11	4.2
	4	5.0		12	6.1
	5	0.3		13	1.0
	6	4.3[c]		14	0.3
	7	**46.8**[d]		15	0.9
	8	5.1		16	1.1
				17	**14.0**

[a]Conformer clusters approaching the conformers in the crystal structures most closely are highlighted in gray.
[b]Conformers with a relative occupancy of more than 10% are highlighted in bold.
[c]Conformer approaching closest the 84% populated conformer in the first and 56% populated in the second X-ray structure.
[d]Conformer approaching closest the 44% populated conformer in the second X-ray structure.

All simulations found among others a binding pose of the 2-substituent approaching the one observed in the crystal structure within an RMSD between 0.4 Å and 1.6 Å. Nevertheless, in all cases the MD simulations suggest an additional number of possible conformers. The substituent conformation of **6a** matches in the second most populated conformer well with the geometry (RMSD = 0.4 Å) found in the crystal structure, the most populated one agrees better with a *trans* orientation of the substituent. For the other ligands **6b**, **6c**, and **6d**, a conformer close to the crystal structure is also observed, however, to a minor fraction along the trajectory. Also here, the occurrence of additional conformers with *gauche* orientation is indicated, in agreement to the disorder indicated in the first data set collected from a crystal of TGT·**6b**.

2.6. Thermodynamic analysis

To reveal a more detailed picture of the thermodynamic properties of ligand binding, we performed ITC titration experiments. The novel data for **5** and **6d** are listed in Table 3 together with those of **1**–**4**, reported previously.[13]

3. Discussion

In the present study, we investigated the conformational properties and the residual mobility of extended substituents oriented into a flat bowl-shaped binding pocket that recognizes the nucleobase uracil-33 with its attached ribose sugar ring of the natural tRNA substrate. Crystal structure analysis suggests residual mobility of the placed substituents in the protein bound state. This residual dynamic property is indicated by enhanced B-factors attributed to the atoms of the substituents. Not in all cases could the population of the 2-substituent be refined to 100%, thus we assume that minor populated conformers are present that scatter over several configurations. In the crystallographic refinement of a high-resolution crystal structure, assignment of an additional conformer appears appropriate if it is indicated by additional electron density and its occupancy refines to at least 15–20%.

The enhanced mobility of the 2-substituents involves in case of **6a** and **6b** a back folding of this part of the ligand combined with a partial abandonment of the pocket. In one data set, collected on another independent crystal of TGT·**6b**, a second conformation could be attributed to the difference electron density. Molecular dynamics simulations also assign enhanced mobility to this part of the ligand. They suggest an even larger scatter over multiple configurations extending those observed in the crystal structures. Many of these, however, show a population well below 15–20%. In a crystallographic diffraction experiment, an additional conformer will only be indicated in the density if a periodic arrangement from unit cell to unit cell is given.

As the binding affinity is not strongly altered across the ligands showing deviating residual mobility, this property cannot be detrimental to binding. This underlines the thermodynamic consideration that residual ligand mobility takes a beneficial contribution to the entropic portion of the Gibbs binding energy. The number of conformational clusters found for the studied ligands by MD simulations differs, however; in all cases, the conformer observed in the crystal is also traversed along the trajectory with a similar geometry. In the packing environment of a crystal structure, molecules are accommodated under restricted conditions. The local crystal environment limits the available space to evolve the dynamic properties of a protein-ligand complex. In case of trypsin, we experienced a remarkable selection of deviating crystal forms.[23] Remarkably, ligand binding was only successful in a particular crystal form if the ligand could establish sufficient residual mobility in the bound state. It accommodated only successfully in the crystal packing if enough unoccupied space next to the binding pocket was accessible so that the ligand could experience the residual mobility necessary for its binding. We even observed

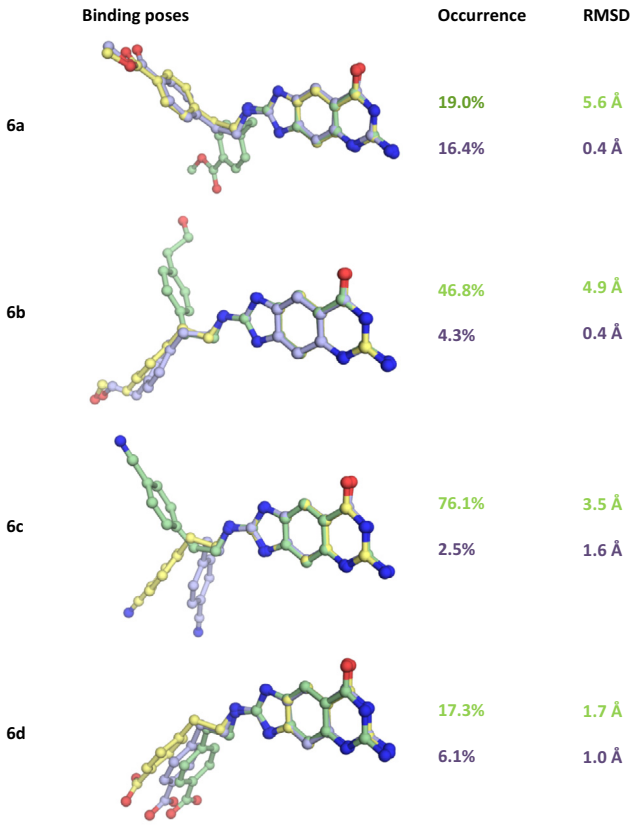


Figure 2. Results of the hierarchical clustering of the frames observed along the trajectory of the 100 ns MD simulations. The conformation observed in the crystal structure (carbon atoms yellow) is superimposed with the most populated conformer (carbon atoms light green) and the conformer approximating most closely the crystal structure (carbon atoms light blue) together with the occupancy in percentage of the simulation time and the RMSD with respect to the geometry in the crystal structure. RMSD values (in Å) were calculated using *fconv* after alignment on the tricyclic scaffold without hydrogens.[22]

decomposition of the crystals if ligands requiring binding with residual mobility were soaked into crystals not offering a large enough volume. Vice versa, the conformationally restricted ligands did not populate in this crystal form. In another study, we evaluated in total 17 data sets of a complex of aldose reductase with zopolrestat.[24] Crystals were obtained under varying soaking and co-crystallization conditions. Depending on the applied protocol, a flip of a peptide bond was observed, accompanied by a rupture of an H-bond formed to the bound ligand.

Protein crystals are usually studied under cryo conditions at 100 K to limit radiation damage.[25] Unfortunately, it is difficult to define a temperature corresponding to the situation seen in a frozen crystal structure. The flash freezing of a crystal will clearly take an impact on residual molecular motions and the probability distribution of putative conformers actually seen in the studied crystal.[26] The applied cooling protocol can be crucial for the results observed in the determined crystal structure.[27] Systematic studies have shown that molecular motions such as the adjustment of side-chain rotamers will find sufficient time to equilibrate according to the conformational distribution given at low temperature.[26] However, adjustment involving larger movements and transformations such as the flipping of entire ring systems or larger rearrangements of ligand portions may be too slow to adjust with flash cooling, thus the situation at ambient in solution is probably reflected in the solid state. Supposedly, the time needed to cool below the glass transition temperature of water in the solvent channels spanning through a protein crystal is crucial for the fixing of molecular motions. Finally, as mentioned above, periodicity in the arrangement across the crystal must be given to detect a molecular portion in the diffraction pattern, whereas complete random distribution will remain unresolved.

Table 3
Raw data of the investigated ligands measured at 25 °C in Hepes buffer

Ligand	K_d [nM]	ΔG^0 [kJ mol^{-1}]	ΔH_{obs} [kJ mol^{-1}]	$-T\Delta S^0$ [kJ mol^{-1}][a]
1	52.4 ± 6.9	−41.6 ± 0.3	−74.8 ± 1.8	33.2 ± 1.9[b]
2	111.6 ± 11.8	−39.7 ± 0.3	−50.9 ± 0.6	11.2 ± 0.6
3	34.3 ± 6.5	−42.6 ± 0.5	−66.7 ± 0.5	24.1 ± 0.7
4	35.0 ± 6.9	−42.6 ± 0.5	−78.3 ± 1.2	35.7 ± 1.3
5	56.6 ± 14.3	−41.4 ± 0.6	−45.3 ± 1.8	3.9 ± 1.9
6d	29.8 ± 8.9	−43.0 ± 0.8	−47.0 ± 1.4	4.0 ± 1.6

[a] $-T\Delta S^0$ was calculated according to the Gibbs–Helmholtz equation.
[b] Errors were estimated by means of standard deviation. The error for $-T\Delta S^0$ was calculated according to error propagation.

For **6b**, we could determine two independent crystal structures and the conformational multiplicity of the substituent differs between the two structures. The back-folded conformation is only observed in one crystal, whereas the extended conformer is found in both determinations. Presumably, the latter conformer corresponds to the energetically more stable situation. In case the adjustment of this conformer into the most stable configuration is slow, a dependence on the cooling rate of the crystals below the glass transition temperature of water could be given. This can lead to deviating populations of conformational states in the measured crystal structures. It remains speculative that in the crystal showing only one predominant conformer, the ligands had more time to properly equilibrate upon cooling.

Facing the MD trajectory with the different crystal structures suggests that residual mobility is significantly reduced in the solid state (even though populations beyond 15–20% are difficult to recognize in the diffraction pattern). Nonetheless, scatter over multiple states and enhanced B-factors serve as a qualitative indicator for residual mobility of the ligand in its bound state.

Most likely, the cooling protocol, the quality and size of the crystal, as well as the rate by which the experimentalist performs the plunging of the crystal into liquid nitrogen and the thickness of cold-gas layer formed above the liquid cryogen (the transfer takes several milliseconds where major part of the crystal cooling already happens) take an impact on the resulting picture in the crystal structure.[27] Even the liquid cryogen used for cooling can influence the results, as different cryo media (e.g. nitrogen, ethane, helium) exhibit different heat capacities and heat conductivities to absorb and conduct heat through the crystal upon flash freezing. As these parameters are difficult to control by the experimentalist in a systematic fashion, a reproducible protocol will hardly be possible to accomplish. However, for the interpretation of crystal structure data this aspect has to be taken into account. In case of aldose reductase, we observed a peptide bond flip in the refined structures of a protein-ligand complex, recorded for data sets taken from multiple crystals.[24] The amount by which this flip had taken place was dependent on the protocol how the crystals were grown and manipulated prior to measurement.

Different terminal functional groups were attached to the *para* position of the phenyl ring at the 2-substituent to provide this ligand portion with a facility to undergo directional interactions with the protein. In case of the attached carboxylate group, the ligand vector departing from the 2-position recruits the guanidinium group of Arg286 to form a salt-bridge and it adopts an ordered conformation. As the assay data show, the formed solvent-exposed salt-bridge contributes hardly anything to the binding affinity of the ligand. ITC confirms this observation. A virtually unchanged thermodynamic profile is observed for **5** and **6d**. This example underlines the fact that a salt-bridge, formed between ligand and protein, can only enhance binding if it forms in a deeply

buried binding pocket exhibiting low dielectric conditions and excluding access of solvent molecules. In a previous study, we could show that similar salt-bridges, formed at the opposing end of the *lin*-benzoguanine scaffold, which becomes entirely buried in the pocket, make a huge enthalpic contribution to binding.[13] The molecular dynamics simulations even indicate that the salt-bridge found in the crystal structure is not stable along the trajectory and forms only temporarily.

4. Conclusions

Protein-ligand binding occurs overwhelmingly in surface-exposed depressions and cavities on a protein. The available pockets are of different shape, burial, and physicochemical composition. The affinity of ligand binding to such pockets depends on the strength of the interactions established in these pockets, the change in the dynamic properties and thus in the degrees of freedom of the system, and the modulation of the local solvent structure.

Usually in deep and highly buried binding pockets, major contributions to binding affinity are experienced. Nonetheless, also accommodation of ligand portions to flat solvent-exposed pockets can be essential to optimize affinity and particularly selectivity. However, how much affinity can be gained by filling such flat pockets, and should the binding pose of a ligand be restricted to one orientation or is a residual mobility with a scatter over multiple orientations beneficial for binding? Considering conformational entropy and the restriction to a small number of degrees of freedom, ligand disorder can be of advantage for binding.

In the present study, we investigated the properties of a series of 2-substituted *lin*-benzoguanine inhibitors binding to the tRNA-modifying enzyme TGT. The attached substituents were designed to fill the flat bowl-shaped uracil-33 binding pocket of the protein. Crystal structure analysis revealed either significantly enhanced residual mobility of the attached substituents in the pocket indicated by larger B-factors or a distribution over multiple orientations is found. Molecular dynamics simulations show an even larger scatter over several clusters of deviating side chain conformers. In the crystal environment, molecular motion is restricted and the flash-cooling process of the crystals freezes or at least reduces mobility to a certain amount, which is difficult to quantify. Nonetheless, some configurations experienced along the computed MD trajectory agree with geometries similarly found in the solid state. Most likely, the flash-cooling protocol takes impact on the final scatter seen in the crystal structures. The interpretation of the properties of a particular binding pose seen in a crystal structure must consider this fact. Finally, the attachment of a functional group competent to establish a salt-bridge in the flat solvent-exposed pocket helps to restrict the bound ligand in a fixed orientation; however, the formation of a solvent exposed salt-bridge has no detectable contribution to binding affinity.

5. Experimental section

5.1. Enzyme assay

The enzyme kinetic characterization of TGT is based on a method described by Grädler et al.[28] and Stengl et al.[29] Due to low solubility, the ligands were dissolved in 100% DMSO and subsequently diluted to the desired concentration with 5% DMSO with the assay buffer. The protein dissolved in the same buffer was added to the various ligand samples with a final concentration of 9 nM and incubated for 10 min at 37 °C. Additionally, a reference without the ligand only containing DMSO and buffer was incubated under the same conditions. Subsequently, base exchange was started by the 1:1 addition of a similarly prepared solution containing *E. coli* $tRNA^{Tyr}$ (ECY2; 3 μM) and a mixture of guanine and radioactively labelled [8-^{3}H]guanine (20 μM). 15 μL aliquots were removed from this mixture (76 μL) every hour and pipetted on Whatman GC-F glass microfiber filters. The reaction was immediately quenched in 10% (w/v) trichloroacetic acid solution at 0 °C. In order to separate the tRNA from excess [8-^{3}H]guanine, not incorporated into tRNA, the glass microfiber filters were washed twice in a 5% (w/v) TCA solution for 10 min, followed by an additional washing step using technical grade ethanol over 20 min. The labelled tRNA was captured in the filters during the described steps. Afterwards, the filters were dried at 60 °C for at least 30 min. The resulting count rate was obtained after the addition of 4 mL Rotiszent™ to each filter by liquid scintillation counting. K_i values were determined using the method described by Dixon at least in duplicate.[30] Thereby, a Michaelis–Menten constant of 0.9 μmol·L^{-1} was used for data evaluation.

5.2. Synthesis

The individual synthesis protocols and the analytical characterizations are given in the Supplementary material.

5.3. Isothermal titration calorimetry

A Microcal iTC_{200} microcalorimeter system (Malvern) was used to perform the ITC measurements. The protein was dissolved in the experimental buffer (50 mM Hepes, 200 mM NaCl and 0.037% Tween 20, pH 7.8) to a final concentration of 10 μM containing 3% DMSO. Due to their low solubility, the ligands were first dissolved in 100% DMSO and then diluted in the buffer solution to a final DMSO concentration of 3%. The ligand concentration in the syringe was adjusted to 200 μM with the same experimental buffer.

All ITC experiments were run at 25 °C after a stable baseline had been achieved. The reference cell was filled with filtered demineralized water. The initial delay before the injections were started and the spacing between each injection were adjusted to 180 s. The first injection contained 0.3–0.5 μL of the ligand solution followed by 14–24 injections of 1.0–1.8 μL. A stirring speed of 1000 rpm was adjusted. Raw data were collected as released heat per time.

To analyze the raw data using the Origin 7.0 software, the baseline and integration limits were adjusted manually. After integrating the area under the peaks, the first data point was removed due to a reduced accuracy.[31] The influence of the heat of dilution was corrected considering heat contributions collected after saturation of the protein. K_d (dissociation constant) as well as ΔH^0 (enthalpy of binding) were extracted by applying a single-site binding model as provided by the manufacturer. Subsequently, $-T\Delta S^0$ was calculated according to the Gibbs–Helmholtz equation. For the current protein-ligand system, we have detected a buffer dependence of the heat signal owing to a protonation linkage.[14] Since the groups involved in the change of protonation state are remote from the attached acidic group in **6d**, data collected in one buffer is sufficient to determine the relative differences between the ligands reported in Table 3. The values should not be used for comparisons on absolute scale.

5.4. X-ray data

5.4.1. *mobilis* TGT crystallization and ligand soaking

Z. *mobilis* TGT was overexpressed and purified as described elsewhere.[32,33] Crystals were grown using the sitting drop vapor diffusion method at 291 K in the presence of the inhibitor. The protein solution was adjusted to 12 mg·mL^{-1} by dilution with high salt buffer (10 mM Tris, 2 m NaCl, 1 mM EDTA, pH 7.8) and incubated with the inhibitor previously dissolved in 100% DMSO at a final

Table 4
Data collection, processing, and refinement statistics for the crystallographically investigated TGT-ligand-complexes

Crystal data	TGT·5	TGT·6a	TGT·6b	TGT·6c	TGT·6d
PDB ID	4Q8T	4Q8U	5JXQ	4Q8V	4Q8W
(A) Data collection and processing					
Collection site	BESSY 14.2	BESSY 14.2	PETRA P14	BESSY 14.2	PETRA P14
No. crystals used	1	1	1	1	1
λ [Å]	0.91841	0.91841	0.97627	0.91841	0.97627
Space group	*C*2	*C*2	*C*2	*C*2	*C*2
Unit cell parameters					
a [Å]	90.2	89.8	90.2	91.2	89.8
b [Å]	64.6	64.5	64.8	65.1	64.7
c [Å]	71.0	70.9	70.3	70.8	70.7
β [°]	93.2	93.2	96.2	96.3	96.3
(B) Diffraction data[a]					
Resolution range [Å]	30–1.40 (1.42–1.40)	30–1.31 (1.33–1.31)	80–1.20 (1.24–1.20)	50–1.40 (1.48–1.40)	80–1.14 (1.19–1.14)
Unique reflections	80321 (3968)	96126 (4340)	123477 (11058)	79794 (12619)	134549 (15089)
$R(I)$sym [%][b]	5.8 (48.9)	4.8 (47.4)	2.9 (49.2)	6.2 (48.3)	2.8 (30.9)
Completeness [%]	99.9 (98.3)	99.1 (89.9)	98.3 (94.4)	97.8 (95.8)	92.0 (80.0)
Redundancy	4.2 (3.8)	3.3 (2.9)	3.3 (2.8)	2.9 (2.9)	5.7 (5.0)
I/σ (I)	22.7 (2.4)	21.8 (2.0)	19.8 (2.2)	11.0 (2.0)	27.8 (4.7)
Wilson B-factor [Å^2]	17.1	15.8	12.8	12.1	11.6
Matthews Coefficient [Å^3/Da]	2.4	2.4	2.4	2.4	2.4
(C) Refinement					
PHENIX version	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.4_1496
Resolution range [Å]	29.8–1.40	18.1–1.31	69.9–1.20	21.8–1.40	70.3–1.14
Reflections used for R_{free}	1949	4815	6174	3990	6728
Reflections used for R_{work}	78372	91306	117303	75804	127821
Final R values[a]					
R_{free} [%][c]	15.4	16.2	15.7	16.3	15.0
R_{work} [%][d]	13.8	14.0	13.7	13.1	13.4
No. of atoms (non-hydrogen)					
Protein atoms	2899	2892	2941	2913	2927
Water molecules	311	330	351	384	376
Ligand atoms	24	28	27	26	27
RMSD, angle [°]	1.1	1.1	1.1	1.1	1.1
RMSD, bond [Å]	0.007	0.007	0.005	0.007	0.007
Ramachandran plot[e]					
Most favoured regions [%]	95.3	94.9	94.6	95.0	94.9
Additionally allowed regions [%]	4.4	4.7	5.1	4.7	4.8
Generously allowed regions [%]	0.3	0.3	0.3	0.3	0.3
Mean B-factors [Å^2]					
Protein atoms	20.7	18.8	16.6	14.6	15.1
Water molecules	33.8	32.8	31.7	29.4	28.0
Ligand atoms	22.6	21.2	18.6	13.0	15.5

[a] Values in parentheses are statistics for the highest resolution shell.
[b] $R(I)_{sym} = [\sum_h \sum_i |I_i(h) - I(h)| / \sum_h \sum_i I_i(h)] \times 100$, in which $I(h)$ is the mean of the $I(h)$ observation of reflection h.
[c] $R_{work} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|$.
[d] R_{free} was calculated as shown for R_{work} but on refinement-excluded 2.4–5% of data.
[e] Statistics from *PROCHECK*.[36]

concentration up to 1.5 mm depending on the solubility of the corresponding ligand. This solution was mixed with 1.5 µL reservoir solution (100 mM MES, pH 5.5, 10% (v/v) DMSO, 11–13% (w/v) PEG 8000) to a 3 µL droplet. The reservoir contained 1.0 mL of the above-mentioned solution. Within one week, crystals showing an appropriate size for data collection were obtained.

Crystals grown in the presence of ligand **5** did not show an appropriate size or form sufficient for diffraction experiments. Thus, a soaking protocol was applied. For this purpose, crystals were grown in the absence of the ligand according to the previously described protocol. Instead of mixing the protein solution with the inhibitor before crystallization, pre-grown *apo* crystals were transferred to a 3 µL droplet of reservoir solution mixed with the stock solution containing the ligand to a final concentration of 1 mM. The droplet was sealed against 1.0 mL reservoir solution and soaked into the crystal overnight.

5.5. Data collection

For data collection, the soaked crystals were transferred into a cryoprotectant solution consisting of 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 for 20 s followed by immediately flash-freezing in liquid nitrogen. Data sets TGT·**5**, TGT·**6a** and TGT·**6c** were collected at the BESSY II (Helmholtz-Zentrum, Berlin, Germany) beamline 14.2 at a wavelength of $\lambda = 0.91841$ Å using a Rayonix MX225 CCD detector. Complex structures for TGT·**6b** and TGT·**6d** were collected at PETRA III (EMBL, Hamburg, Germany) beamline P14 at a wavelength of $\lambda = 0.97627$ Å using a PILATUS 6 M-F detector. To minimize radiation damage, all data sets have been collected at cryo-conditions (100 K).

All TGT crystals corresponded the monoclinic space group C2 containing one monomer in the asymmetric unit. Data sets TGT·**5**

and TGT·**6a** were processed and scaled with the HKL2000 package.[34] Data processing and scaling for TGT·**6b**, TGT·**6c**, and TGT·**6d** were performed with XDS and XSCALE, respectively.[35] Cell dimensions, data collection, and processing statistics are given in Table 4.

5.6. Structure determination and refinement

The coordinates of the TGT *apo* structure 1PUD served as a starting model for molecular replacement using the program Phaser MR of the ccp4 program suite.[37] Structures were refined using the program Phenix (phenix.refine 1.8.4_1496)[38] starting with a first cycle of simulated annealing using default parameters. Further refinement cycles comprised the coordinates (xyz), occupancy, and individual B-factor refinement as well as applying metal restraints for the zinc ion. The individual atomic displacement parameter (ADP) weights were refined for all structures. The temperature factors of all structures were refined anisotropically. The calculation of the R_{free} value comprised a 2.4–5% fraction of the data.

For all structures, amino acid side chains were fitted according to their σA-weighted $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ electron density obtained in the program *Coot*.[39] After the initial refinement cycles, the zinc ion as well as water and glycerol molecules were implemented in the model. For adding water molecules, the option 'update waters' included in Phenix was used after increasing the hydrogen-bond length threshold for the solvent-model and solvent–solvent contacts to 2.3 Å. The inserted molecules were visually inspected afterwards. Ligand restraints were generated by the CSD based gradeWebServer [http://grade.globalphasing.org] in case of **6b**, **6c**, and **6d**. Ligands were built and minimized using the program *MOE* [*MOE 2012.10*] and geometric restraints calculated subsequently by the program *Monomer Library Sketcher*[40] in case of **5** and **6a**. Multiple protein-residue conformations were assigned in case a reasonable electron density was observed and were kept during refinement if the side chain with the lowest occupancy showed a value of at least 20%. In case of **6a** and **6b**, the occupancy of the substituent was refined due to elevated B-factors observed for this portion. Ramachandran plots have been generated with the program *PROCHECK*.[36] For the analysis of temperature factors the program *Moleman* has been used.[41] The burial of molecular ligand portions in the protein binding pocket has been computed using the program MS.[42]

5.7. Accession codes

The atomic coordinates and structure factors of the TGT complex with ligands **5, 6a–d** have been deposited in the RCSB Protein Data Bank (PDB) with accession codes: 4Q8T, 4Q8U, 5JXQ, 4Q8V, 4Q8W.

5.8. Docking

Docking was performed using *GOLD* Suite v5.1 [The Cambridge Crystallographic Data Centre: Cambridge, U.K.], the default values for the genetic algorithm (GA)[43] and the scoring function *ChemScore*.[44] For each ligand, 30 GA runs were performed with 100,000 operations. As a protein model, the coordinates of the crystal structure 2QZR were used. The binding site included all atoms within a distance of 10 Å to the sulfur of Met260. Ligands were drawn using the *MOE* builder.

5.9. Molecular dynamic simulation

Molecular dynamic (MD) simulations were performed with the program *AMBER11*.[45] Docking results of the corresponding ligands served as starting coordinates. Prior to the simulation, the pdb-file had to be modified: All crystallographically determined water molecules were extracted. Missing amino acids as well as the most probable conformer of missing side chains were added. In case of more than one visible conformation, the highest occupied one was kept during the simulation. Protonation states of histidines were inspected visually and set as HID (hydrogen at δ-position, HIS#: 90, 257, 319, 332) or HIE (hydrogen at ε-position, HIS#: 73, 145, 333). The zinc ion was mimicked by four massless dummy atoms, each with a charge of +0.5 using the CaDA approach by Pang.[46] Cysteine residues involved in the binding of the zinc ion were set as CYM (CYS#: 308, 310, 313), histidines as HIN (HIS#:349).

Due to the required computational time, the calculations were based on only one TGT monomer. Parameters for the ligands were generated with the program *antechamber* using the general amber force field (gaff),[47] its charges were calculated via bond charge correction (bcc).[48,49] Addition of hydrogen atoms to the protein, neutralization of the system by adding two sodium ions, and solvation of the complex in a TIP3P water box was done with the *tleap*.[50]

After a minimization of the water box comprising 100 steps and the whole system comprising 500 steps performed with a generalized Born solvent model, all following simulations included periodic boundary conditions, the Particle Mesh Ewald procedure (PME)[51] and the SHAKE algorithm[52] using the ff99SB force field[53] with a cut-off of 10 Å. Thereby, the system is heated up to 300 K stepwise (0 K···150 K···225 K···300 K) over a period of 150 ps fixing the TGT monomer with weak restraints (25 kcal · mol^{-1} · $Å^{-2}$). Subsequently, the pressure is adjusted to 1 bar over a time scale of 50 ps followed by a productive simulation for 10 to 100 ns using 2 fs time steps under NPT conditions. The trajectory derived under these conditions was further analyzed with the program *ptraj*, whereby every second frame was included into the analyses.

Acknowledgements

We are grateful to the beamline staff at BESSY II (Helmholtz-Zentrum Berlin) and Petra in Hamburg (Desy, EMBL) for providing outstanding support during data collection. We also thank the Helmholtz-Zentrum Berlin for the travel support. We are grateful to the Chemical Computing Group (Montreal, Canada) for making the MOE program available to us. The usage of the AMBER program suite is kindly acknowledged. The present study was supported by the ERC Advanced Grant no. 268145-DrugProfilBind awarded to G.K. and a grant from the ETH Research Council presented to F.D.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmc.2016.07.053.

References and notes

1. Krimmer, S. G.; Betz, M.; Heine, A.; Klebe, G. *ChemMedChem* **2014**, *9*, 833.
2. Biela, A.; Nasief, N. N.; Betz, M.; Heine, A.; Hangauer, D.; Klebe, G. *Angew. Chem., Int. Ed.* **2013**, *52*, 1822.
3. Stengl, B.; Meyer, E. A.; Heine, A.; Brenk, R.; Diederich, F.; Klebe, G. *J. Mol. Biol.* **2007**, *370*, 492.
4. Kohler, P. C.; Ritschel, T.; Schweizer, W. B.; Klebe, G.; Diederich, F. *Chem. Eur. J.* **2009**, *15*, 10809.
5. Ritschel, T.; Kohler, P. C.; Neudert, G.; Heine, A.; Diederich, F.; Klebe, G. *ChemMedChem* **2009**, *4*, 2012.
6. Ritschel, T.; Hoertner, S.; Heine, A.; Diederich, F.; Klebe, G. *ChemBioChem* **2009**, *10*, 716.
7. Barandun, L. J.; Immekus, F.; Kohler, P. C.; Ritschel, T.; Heine, A.; Orlando, P.; Klebe, G.; Diederich, F. *Acta Crystallogr., Sect. D* **2013**, *D69*, 1798.
8. Xie, W.; Liu, X.; Huang, R. H. *Nat. Struct. Biol.* **2003**, *10*, 781.
9. Kotloff, K. L.; Winickoff, J. P.; Ivanoff, B.; Clemens, J. D.; Swerdlow, D. L.; Sansonetti, P. J.; Adak, G. K.; Levine, M. M. *WHO Bull.* **1999**, *77*, 651.
10. Durand, J. M.; Okada, N.; Tobe, T.; Watarai, M.; Fukuda, I.; Suzuki, T.; Nakata, N.; Komatsu, K.; Yoshikawa, M.; Sasakawa, C. *J. Bacteriol.* **1994**, *176*, 4627.

11. Levine, M. M.; Kotloff, K. L.; Barry, E. M.; Pasetti, M. F.; Sztein, M. B. *Nat. Rev. Microbiol.* **2007**, *5*, 540.
12. Gu, B.; Cao, Y.; Pan, S.; Zhuang, L.; Yu, R.; Peng, Z.; Qian, H.; Wei, Y.; Zhao, L.; Liu, G.; Tong, M. *Int. J. Antimicrob. Agents* **2012**, *40*, 9.
13. Neeb, M.; Betz, M.; Heine, A.; Barandun, L. J.; Hohn, C.; Diederich, F.; Klebe, G. *J. Med. Chem.* **2014**, *57*, 5566.
14. Neeb, M.; Czodrowski, P.; Heine, A.; Barandun, L. J.; Hohn, C.; Diederich, F.; Klebe, G. *J. Med. Chem.* **2014**, *57*, 5554.
15. Barandun, L. J.; Immekus, F.; Kohler, P. C.; Tonazzi, S.; Wagner, B.; Wendelspiess, S.; Ritschel, T.; Heine, A.; Kansy, M.; Klebe, G.; Diederich, F. *Chem. Eur. J.* **2012**, *18*, 9246.
16. Hörtner, S. R.; Ritschel, T.; Stengl, B.; Kramer, C.; Schweizer, W. B.; Wagner, B.; Kansy, M.; Klebe, G.; Diederich, F. *Angew. Chem.* **2007**, *119*, 8414; *Angew. Chem., Int. Ed.* **2007**, *46*, 8266.
17. B. Zacharie, C. Penney, L. Gagnon, B. Grioux, L. Geerts, S. Abbott, WO2008/131547, (2008).
18. S. Hachiya, M. Oku, H. Mukai, T. Shin, K. Matsuura, R. Seo, T. Kamikubo, Y. Terada, M. Sanagi, K. Yoshihara, T. Takahashi, WO2006/123725, (2006).
19. Desai, D.; Madhunapantula, S. V.; Gowdahalli, K.; Sharma, A.; Chandagaludoreswamy, R.; El-Bayoumy, K.; Robertson, G. P.; Amin, S. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2038.
20. Bartoszewicz, A.; Kalek, M.; Nilsson, J.; Hiresova, R.; Stawinski, J. *Synlett* **2008**, 37.
21. Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossvary, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, B.; Wu, X.; Kollman, P. A. *Amber 14*; University of California: San Francisco, 2014.
22. Neudert, G.; Klebe, G. *Bioinformatics* **2011**, *27*, 1021.
23. Brandt, T.; Holzmann, N.; Muley, L.; Khayat, M.; Wegscheid-Gerlach, C.; Baum, B.; Heine, A.; Hangauer, D.; Klebe, G. *J. Mol. Biol.* **2011**, *405*, 1170.
24. Steuber, H.; Zentgraf, M.; Gerlach, C.; Sotriffer, C. A.; Heine, A.; Klebe, G. *J. Mol. Biol.* **2006**, *363*, 174.
25. Garman, E. F. *Acta Crystalogr., Sect. D* **2010**, *D66*, 339.
26. Fraser, J. S.; van den Bedem, H.; Samelson, A. J.; Lang, P. T.; Holton, J. M.; Echols, N.; Alber, T. *Proc. Natl. Acad. Sci.* **2011**, *108*, 16247.
27. Warkentin, M.; Berejnov, V.; Husseini, N. S.; Thorne, R. E. *J. Appl. Crystallogr.* **2006**, *39*, 805.
28. Grädler, U.; Gerber, H.-D.; Goodenough-Lashua, D. M.; Garcia, G. A.; Ficner, R.; Reuter, K.; Stubbs, M. T.; Klebe, G. *J. Mol. Biol.* **2001**, *306*, 455.
29. Stengl, B.; Reuter, K.; Klebe, G. *ChemBioChem* **2005**, *6*, 1926.
30. Dixon, M. *Biochem. J.* **1953**, *55*, 170.
31. Mizoue, L. S.; Tellinghuisen, J. *Anal. Biochem.* **2004**, *326*, 125.
32. Romier, C.; Ficner, R.; Reuter, K.; Suck, D. *Proteins: Struct., Funct., Bioinf.* **1996**, *24*, 516.
33. Reuter, K.; Ficner, R. *J. Bacteriol.* **1995**, *177*, 5284.
34. Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307.
35. Kabsch, W. *Acta Crystallogr., Sect. D* **2010**, *D66*, 125.
36. Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
37. McCoy, A. *Acta Crystallogr., Sect. D* **2007**, *D63*, 32.
38. Adams, P. D.; Afonine, P. V.; Bunkoczi, 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. *Acta Crystallogr., Sect. D* **2010**, *D66*, 213.
39. Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D* **2004**, *D60*, 2126.
40. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr., Sect D* **2011**, *67*, 235.
41. Kleywegt, G. J.; Zou, J. Y.; Kjeldgaard, M.; Jones, T. A.; Around, O. *International Tables for Crystallography*; Kluwer Academic Publishers: Dordrecht, 2001.
42. Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548.
43. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727.
44. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425.
45. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668.
46. Pang, Y.-P. *Proteins: Struct., Funct., Bioinf.* **2001**, *45*, 183.
47. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
48. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.
49. Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623.
50. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
51. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
52. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
53. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinform.* **2006**, *65*, 712.

Please cite this article in press as: Neeb, M.; et al. *Bioorg. Med. Chem.* (2016), http://dx.doi.org/10.1016/j.bmc.2016.07.053