ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Launching Spiking Ligands into a Protein−Protein Interface: A Promising Strategy To Destabilize and Break Interface Formation in a tRNA Modifying Enzyme

Florian Immekus,[⊥,†] Luzi Jakob Barandun,[⊥,‡] Michael Betz,[†] François Debaene,[§] Stéphanie Petiot,[§] Sarah Sanglier-Cianferani,[§] Klaus Reuter,[†] François Diederich,[‡] and Gerhard Klebe*[,†,∥]

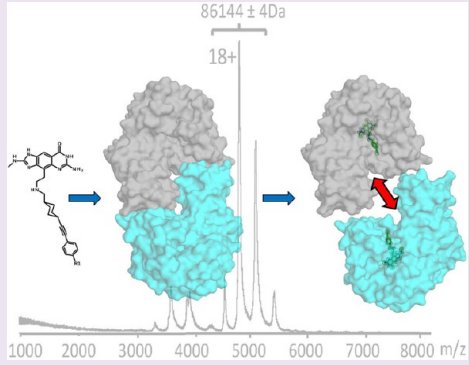
[†]Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany

[‡]Laboratorium für Organische Chemie, ETH Zurich Hönggerberg, HCI, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland

[§]Laboratoire de Spectrométrie de Masse Bio-Organique de Strasbourg, 25, rue Becquerel, 67087 Strasbourg Cedex 2, France

Supporting Information

ABSTRACT: Apart from competitive active-site inhibition of protein function, perturbance of protein−protein interactions by small molecules in oligodomain enzymes opens new perspectives for innovative therapeutics. tRNA−guanine transglycosylase (TGT), a potential target to treat shigellosis, is active only as the homodimer. Consequently, disruption of the dimer interface by small molecules provides a novel inhibition mode. A special feature of this enzyme is the short distance between active site and rim of the dimer interface. This suggests design of expanded active-site inhibitors decorated with rigid, needle-type substituents to spike into potential hot spots of the interaction interface. Ligands with attached ethinyl-type substituents have been synthesized and characterized by K_d measurements, crystallography, noncovalent mass spectrometry, and computer simulations. In contrast to previously determined crystal structures with nonextended active-site inhibitors, a well-defined loop-helix motif, involved in several contacts across the dimer interface, falls apart and suggests enhanced flexibility once the spiking ligands are bound. Mass spectrometry indicates significant destabilization but not full disruption of the complexed TGT homodimer in solution. As directed interactions of the loop-helix motif obviously do not determine dimer stability, a structurally conserved hydrophobic patch composed of several aromatic amino acids is suggested as interaction hot spot. The residues of this patch reside on a structurally highly conserved helix-turn-helix motif, which remains unaffected by the bound spiking ligands. Nevertheless, it is shielded from solvent access by the loop-helix motif that becomes perturbed upon binding of the spiking ligands, which serves as a possible explanation for reduced interface stability.



In summer 2011, Germany witnessed the sudden development of a highly pathogenic *Shiga*-toxin-producing strain of enterohemorrhagic *Escherichia coli* (EHEC). Fast epidemic spread of the infection with more than 2000 cases within six weeks underlines the potential hazard of bacteria-triggered gastrointestinal infections.[1] The closely related bacteria of the genus *Shigella* exhibit similar properties, causing acute inflammation of the colon along with bloody, mucopurulent diarrhea, known as shigellosis. The disease occurs predominantly in developing countries, responsible annually for more than one million lethal cases.[2]

The tRNA-modifying enzyme tRNA−guanine transglycosylase (TGT, EC 2.4.2.29), involved in the infection pathway of *Shigella* bacteria, was identified as potential drug target for the treatment of shigellosis.[3,4] Bacterial TGT catalyzes incorporation of the premodified nucleobase $preQ_1$ at position 34 ("wobble position") of $tRNA^{Asp,Asn,His,Tyr}$, replacing the genetically encoded guanine in due course of the exchange reaction. Subsequently, in TGT-independent reactions, the hypermodified tRNA nucleoside queuosine is formed. This tRNA processing is a prerequisite for the biosynthesis of invasion factors and essential to create pathogenicity of *Shigella*. The latter proteins are pivotal for bacterial invasion of human colon mucosa cells.[5,6] Structural properties of TGT have been studied by crystallography of the related *Zymomonas mobilis* enzyme, deviating in the active site from *Shigella* by only one Phe/Tyr-exchange.[7]

In addition to the design of potent inhibitors, insights into the quaternary structure of TGT, its oligomerization state in solution, and the stoichiometry of the catalytic transformation have attracted notice in recent research.[8,9] Crystal structure analysis of *Z. mobilis* TGT revealed a conserved homodimer with ca. 1600 $Å^2$ buried contact area between the two monomers,[8] independent of whether the apo protein, TGT−ligand complexes, or a TGT−tRNA anticodon stem-loop complex have been studied.[10] In most crystal structures, the

Received: January 9, 2013
Accepted: March 27, 2013
Published: March 27, 2013

homodimer coincides with a crystallographic 2-fold axis; however, also the anticodon stem-loop complex shows homodimeric arrangement with two monomers and one bound tRNA molecule. This complex does not coincide with crystallographic symmetry. Apparently, one monomer binds the substrate, whereas the other stabilizes the ternary complex.[8,10] This crystallographic finding appears obvious as the active sites of the two monomers are closely adjacent and cannot host simultaneously two tRNA substrates. Similar results are found in solution. Noncovalent ESI mass spectrometry clearly reveals 2:1 TGT:tRNA complexes as solution species without any evidence for concentration-dependent higher-order complexes. Thus, homodimer formation is a clear functional prerequisite, and single-site mutations can take impact on dimer stability.[8]

Visual inspection of the crystal structures referred to several directed interactions across the dimer interface. At first glance, the salt bridge between Lys52 and Glu339′ and the hydrogen bond between Ala49 and Tyr330′ have been selected as putative key interactions (Figure 1). Straightforward mutations by structurally related amino acids have been performed: the TGT(Tyr330Phe) and TGT(Lys52Met) mutants exhibited significantly reduced k_{cat} values (factor 10 and 50), while for the double mutant TGT (Lys52Met/Tyr330Phe) no residual catalytic activity was recognized. A crystal structure of TGT(Lys52Met) unraveled pronounced disorder of complete secondary structure elements that contribute major parts to the dimer interface (helix α1 and the adjacent β1α1-loop, Figure 1). Surprisingly, mass spectrometry of TGT(Tyr330Phe) and TGT(Lys52Met) showed only a slight concentration-dependent increase of the monomeric form instead of the anticipated complete dimer disruption. It is therefore in question whether this finding in solution correlates with crystallographic evidence as TGT(Lys52Met) still forms a crystallographic dimer in the solid state.

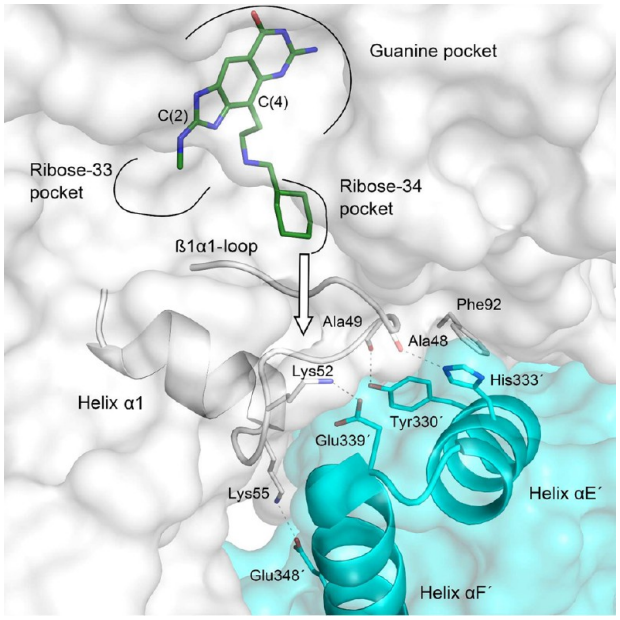


Figure 1. View of the binding pocket with the adjacent part of the dimer interface of *Z. mobilis* TGT. Color coding: monomer A, gray; monomer B, cyan; C_{ligand}: green, O: red, N: blue, S: yellow. Hydrogen bonds are shown as dashed lines. These characteristics apply to all figures unless stated differently. Both monomers are shown with transparent surfaces. Involved secondary structures are depicted as cartoon. Residues involved in directed interactions across the monomers are displayed as sticks. Glu339′ is located at the apex of helix αE′ and helix αF′ (*helix-turn-helix* motif, cyan). Between β1α1-loop/helix α1 (*loop-helix* motif, gray) and the corresponding helix-turn-helix motif, two hydrogen bonds and three salt bridges are formed (Glu57′−Lys325′ not shown for clarity reasons). Ligand **1m** is displayed in the complex structure (PDB ID: 3EOS[11]). This ligand exhibits the parental *lin*-benzoguanine scaffold that was further modified in this study. The white arrow indicates the direction at the terminal cyclohexyl ring used to attach spiking substituents to possibly interfere with the helix-loop and the subjacent helix-turn-helix motif.

More as an incidental observation, two cocrystal structures of TGT with *lin*-benzoguanine active-site inhibitors indicated ligand-induced conformational changes within the dimer interface.[9,11] These changes are produced by ligands with extended C(4)-substituents pointing into the ribose-34 pocket that is found in close vicinity of the dimer interface (Figure 1 and Table 1). However, binding of these ligands had no obvious impact on dimer disruption in solution.[8] Nevertheless, the reported ligand-induced effects paved the ground for a novel strategy to disassemble the interface: ligands with extended C(4)-substituents should be able to directly spike into the interface region, thereby perturbing and subsequently preventing dimer formation. This design concept deviates from most strategies followed in literature,[12−16] which aim at the binding of a small molecule to one of the dimer interface mates mimicking a structural motif such as a helical stretch of the interface partner.

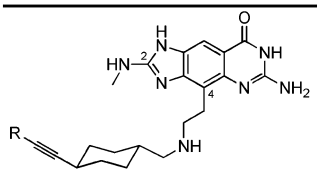
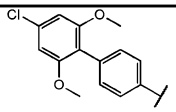
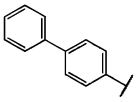
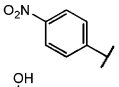
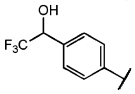
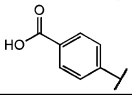
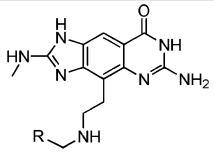
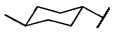
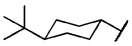
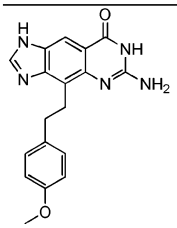
Our approach implicates the following three perspectives: (1) As TGT dimerization is obviously prerequisite for catalytic activity, ligands perturbing the interface could provide a novel mechanism for enzyme inhibition apart from competitive active-site inhibition. (2) Typically, binding sites within a dimer interface are flat and, in the monomeric state, mostly solvent-exposed surfaces with indentations. Hence, structure-based design of interface inhibitors is complicated by their most likely low binding affinity. In consequence, it is difficult to determine a crystal structure of the complex. In contrast, applying reasonably decorated C(4)-substituents attached to potent active-site binders, spiking into the dimer interface, can suggest new interaction sites without dramatic loss of binding affinity. In consequence, this strategy provides better opportunities to determine multiple crystal structures with bound ligands. (3) This approach opens the perspective to design the C(4)-substituents to selectively interfere with the structural arrangement next to the dimer interface. Thus, the most promising interactions with these local features can be analyzed with well-established experimental effort. Usually, mutation studies are performed to obtain insights into interface formation. However, problems regarding protein purification and spontaneous crystallization are frequently experienced.[8] These will be avoided following the above outlined ligand-based approach requiring solely the wild-type protein.

In this contribution, a comprehensive study including X-ray crystallography, K_d measurements, noncovalent ESI mass spectrometry, structure-based design, and molecular dynamic simulations is presented. The multiplicity of methods opens various perspectives to characterize the TGT homodimer interface.

Based on the highly potent *lin*-benzoguanines as scaffold, C(4)-substituents were designed to spike from the active site into the adjacent dimer interface (Figure 1). Obtained TGT−ligand cocrystal structures provided information about the

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

Table 1. Affinity Data and Calculated log*P* and log*C* Values of the Inhibitors

	R =	K_d / nM	clog*P*[a]	clog*D*[a]	ligand efficiency/[b] kcal mol^{-1}
	**1a** 	32 ± 9	6.26	3.92	0.22
	1b 	24 ± 14	5.96	3.61	0.25
	1c 	5 ± 2	2.78	0.44	0.32
	1d 	20 ± 6	4.98	2.64	0.29
	1e 	14 ± 2	3.63	1.29	0.28
	1h 	36 ± 15	3.86	1.51	0.25
	1i 	4 ± 1	3.80	1.02	0.30
	**1j** 	1.1 ± 0.3	3.25	0.92	0.39
	1k 	1.2 ± 0.3	2.85	0.50	0.43
	1l 	25 ± 10	4.12	1.78	0.33
	1m[11,17] 	4 ± 2[c]	2.31	–0.23	0.42
	1n[11,17] 	2 ± 1[c]	1.78	–0.87	0.46
	**2**[9]	3700 ± 900[c]	1.90	–1.89	0.30

[a]log*P* and log*C* values calculated using ACD/Labs software. [b]Ligand efficiency = $RT \ln[K]$/(no. of non-hydrogen atoms). [c]Competitive inhibition constant (K_i) measured in a kinetic assay.[46]

importance of directional interactions within the interface. Supported by computational chemistry, the stability of the involved secondary structures is discussed. Noncovalent mass spectrometry and K_d measurements relate structural data provided by the crystal structures with the ligand-induced effects in solution. To our surprise, a loop-helix motif in the protein–protein interface that features most of the residues involved in H-bonding and salt bridges across the interface was substantially perturbed by ligand binding; however, the dimer contact is only moderately destabilized. Instead a cluster of hydrophobic residues seems to be determinant for stability, as will be discussed comparatively in the conclusion section.

RESULTS AND DISCUSSION

Analysis of the Dimer Interface Architecture. To perform a detailed analysis of the structural features of the dimer interface, 20 nonredundant TGT crystal structures, denoted as TGT_{20} in the following, have been randomly chosen and compared (see Methods). All structural features reported below were highly conserved in TGT_{20} (see Supplementary Figures 1SI and 2SI). The presented distance values are based on the apo structure (PDB ID: 1P0D[7]) assessed as representative for TGT_{20}.

The TGT homodimer spans an interface region of 1667 Å^2, and each monomer contributes 43 contacting residues.[8] Regarding the directional interactions across the overall dimer interface, 10 salt bridges and 8 hydrogen bonds are involved, and the largest portion of them resides on an α-helix composed of 9 residues (helix α1: Lys55–Gly63). It is preceded by a $\beta1\alpha1$-loop (Val45–Leu54) formed by 10 residues (the entire Val45–Gly63 stretch is named *loop-helix* motif in the following, Figure 1). Due to 2-fold symmetry, any structural feature will

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163–1178

contribute twice to the dimer interface. Each $\beta1\alpha1$-loop comprises one residue to form the Lys52···Glu339′ salt bridge and two hydrogen bonds (Ala48···His333′, Ala49···Tyr330′), while each helix $\alpha1$ contributes to two salt bridges (Lys55···Glu348′, Glu57···Lys325′). The binding partners of the corresponding monomer (indicated by a prime) are located on a *helix-turn-helix* motif comprising helix αE′ (Ser327′−Arg336′) and helix αF′ (Ile340′−Glu367′). The latter helix stretches with 42 Å length across the entire monomer. Both helices are linked via a three-residue turn motif (Ala337′, Gly338′, Glu339′), with Glu339′ localized at the apex between both helices (Figure 1). Analyzing helix αE′ with respect to hydrophobicity, Tyr330′ and His333′ attracted attention apart from their above-mentioned role as hydrogen-bond donors. Complemented by Trp326′, they form a hydrophobic, aromatic triad that experiences several van der Waals contacts to the adjacent monomer unit (Figure 1 and 2). Altogether, a rather

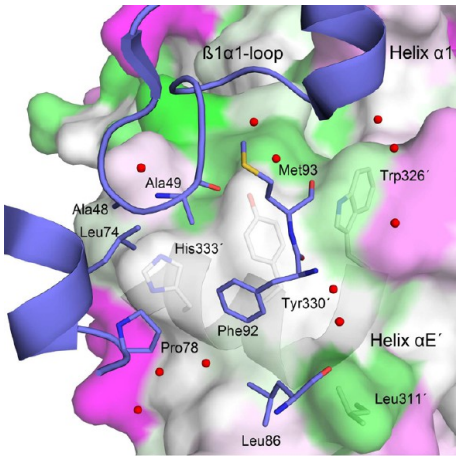


Figure 2. Hydrophobic contact area within the TGT dimer interface as observed in the apo structure (PDB ID: 1P0D[7]) and found to be conserved across a data set of 20 structures (TGT_{20} data set). Color coding: monomer A, blue; monomer B, gray. Water molecules are shown as red spheres and represent archetypes of the local water clusters found in the TGT_{20} data set. The solvent-accessible surface of monomer B is colored according to a normalized consensus hydrophobicity scale[51] ranging from hydrophobic (green) to hydrophilic (pink). An aromatic triad composed by Trp326′, Tyr330′, and His333′ is indicated as part of a hydrophobic contact patch exposed to the corresponding monomer mate of the dimer. The $\beta1\alpha1$-loop of the mate (blue) shields as a kind of lid this hydrophobic patch from the solvent. This lid is further accomplished by two hydrophobic residues: Phe92 and Met93 from the dimer mate. In all analyzed TGT_{20}, no water molecules penetrate into this contact area. All side chains and secondary structures are highly conserved in the crystal structures (see Supporting Information).

hydrophobic, solvent-shielded contact area is generated between both monomers. His333′ is in close contact with Leu74 (4.5 Å) and Pro78 (3.6 Å) of the other dimer mate, whereas Tyr330′ provides hydrophobic interactions to Phe92 (3.7 Å) and Met93 (4.2 Å). The latter residue exhibits further contacts to Trp326′ (4.6 Å), which additionally interacts with the adjacent Pro56 on helix $\alpha1$ (3.7 Å, not shown in Figure 2). Whereas the interaction between Leu86 and Leu311′ (3.9 Å) blocks the contact patch on one face, the $\beta1\alpha1$-loop shields the hydrophobic interface from above. Ala49 is located closely adjacent to Tyr330′ (4.1 Å) and His333′ (4.3 Å), and the neighboring Ala48 interacts with His333′ (4.2 Å). The TGT_{20} data set has been further analyzed regarding conserved water molecules accommodated in this hydrophobic contact region (Figure 2). Interestingly, several highly conserved water molecules surround the hydrophobic triad, whereas only a single water molecule is found within the hydrophobic interface patch.

Design of Interface-Spiking Ligands. We developed *lin*-benzoguanines as a tricyclic scaffold to occupy the guanine/preQ$_1$ recognition pocket with low nanomolar affinity.[11,17] It mimics the natural substrates by replacing their hydrogen-bonding pattern to Asp102, Asp156, Gln203, and Gly230 (see crystallography part).[11,18] Furthermore, the flat tricyclic scaffold intercalates between the side chains of Tyr106 and Met260 similarly to the natural substrates. Design of potent active-site inhibitors exploited the ribose-33 pocket via C(2)-substitution[18] and the ribose-34 pocket via C(4)-substitution at the parent scaffold (Figure 1, Table 1).[11,17]

The design of our interface-spiking ligands was not entirely driven by affinity considerations but particularly focused toward the far end of the ribose-34 pocket that borders the dimer interface. We therefore considered extended C(4)-substituents to selectively perturb molecular portions of the interface and to explore putative interaction hot spots between both monomers (Figure 1). As previously reported, **1m**, **1n**, and **2** (Table 1) exert already some structural impact on the loop-helix motif.[9,11] In particular, **1m** served as a lead for the designed ligand series, as its terminal cyclohexyl ring is deeply buried in the ribose-34 pocket and is suitable as a versatile vector to launch attached substituents targeting the $\beta1\alpha1$-loop and the capping Glu339′ located at the apex of the helix-turn-helix motif of the dimer mate (Figure 1). Ligands **1j−l** were designed to structurally expand the cyclohexyl moiety, whereas **1a−e,h,i** were decorated with a rigid, linear ethinyl linker to launch additional substituents as a kind of warhead into the interface region (Figures 1 and 2).

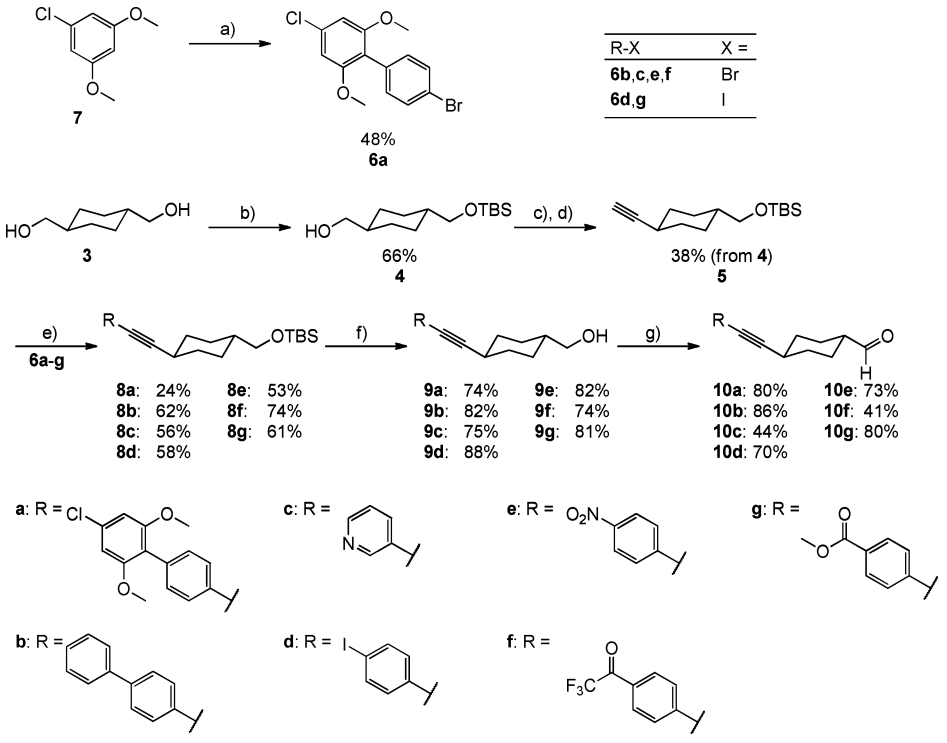
Synthesis. The synthetic strategy followed, as previously published by us, convergence of the substituents for the ribose-34 pocket and a benzimidazole building block by reductive amination.[17] The substituents for the ribose-34 pockets were prepared starting from *trans*-cyclohexane-1,4-dimethanol (**3**) (Scheme 1). Monoprotection to **4**, oxidation, and Seyferth−Gilbert homologation using the Ohira−Bestmann modification gave alkyne **5**.[19] The aryl halides **6b−g** for the following cross-coupling reaction were commercially available, and **6a** was accessible from **7** by treatment with *n*-BuLi followed by addition of B(O*i*Pr)$_3$ and Suzuki cross-coupling reaction. After Sonogashira cross-coupling reaction, the resulting alkynylated cyclohexanes **8a−g** were deprotected using *n*-Bu$_4$NF, giving alcohols **9a−g**, which were oxidized to the corresponding aldehydes **10a−g**.

The amine building block **11** was prepared in 12 steps from benzimidazole-5-carboxylic acid (**12**) as described in the literature (Scheme 2).[17,18] Reductive amination using aldehydes **10a−g** (Scheme 1) and **10j**,[20] **10k**,[21] and **10l** [22] gave amines **13a−e,g,h,j**. For trifluoroketone **10f**, the product of the reductive amination was the corresponding alcohol **13h** instead of ketone **13f**. The obtained amines were subjected to cyclization with chloroformamidinium chloride, yielding *lin*-benzoguanines **1a−e,g,h,j**, whereby the *N*,*N*-dimethylsulfamoyl group was cleaved. The crude compound **1g** was directly used for saponification, giving **1i.**

Both precursors **11**, **13a−g,j−l** and target ligands **1a−e,h,i,l** were isolated and characterized as free amines. For the

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

Scheme 1. Synthesis of Alkyne Vectors 10a−g[a]



[a]Reagents and conditions: (a) (i) *n*-BuLi, THF, −78 °C, 1 h; (ii) $B(OiPr)_3$, −78 to 25 °C, 4 h; (iii) C_6H_4BrI, $[Pd(PPh_3)_4]$, Na_2CO_3, toluene, EtOH, H_2O, 90 °C, 5 h; 48%. (b) TBS-Cl, Et_3N, DMAP, DMF, 0 to 25 °C, 22 h; 66%. (c) DMP, CH_2Cl_2, 0 to 25 °C, 1.5 h; ca. 65%. (d) Dimethyl (2-oxopropyl)phosphonate, tosyl azid, K_2CO_3, MeCN/MeOH, 25 °C, 12 h; 58%. (e) R−Br or R−I, $[PdCl_2(PPh_3)_2]$, CuI, DIPA/THF, 86 °C, 3.5−18 h; 24−74%. (f) TBAF, THF, 25 °C, 2.5−7 h; 74−88%. (g) TPAP, NMO, CH_2Cl_2, 25 °C, 0.5−5 h; 41−86%. DIPA = diisopropylamine, DMAP = *N*,*N*-dimethylpyridin-4-amine, DMF = *N*,*N*-dimethylformamide, DMP = Dess−Martin periodinane, *n*-BuLi = *n*-butyllithium, NMO = 4-methylmorpholine *N*-oxide, TBAF = tetrabutylammonium fluoride, TBS = *tert*-butylchlorodimethylsilane, THF = tetrahydrofuran, TPAP = tetrapropylammonium perruthenate(VII).

purification of **1j** and **1k**, a previously reported protocol was applied to yield the tris-hydrochloride salts.[17] To isolate the other target ligands **1a−e,h,i,l** as free amines, their formate salts were initially formed (see experimental details in the Supporting Information). During final lyophylization, formic acid was evaporated, yielding the free amines, as clearly evidenced by the analytical characterization.

Evaluation of Binding Data. Dissociation constants K_d were obtained at pH 7.3 in 100 mM HEPES buffer by microscale thermophoresis. Application of a local temperature gradient induces protein motion that depends on size, charge, and hydration shell. These properties are modulated by the presence of bound ligands, which therefore allows determination of dissociation constants K_d.[23]

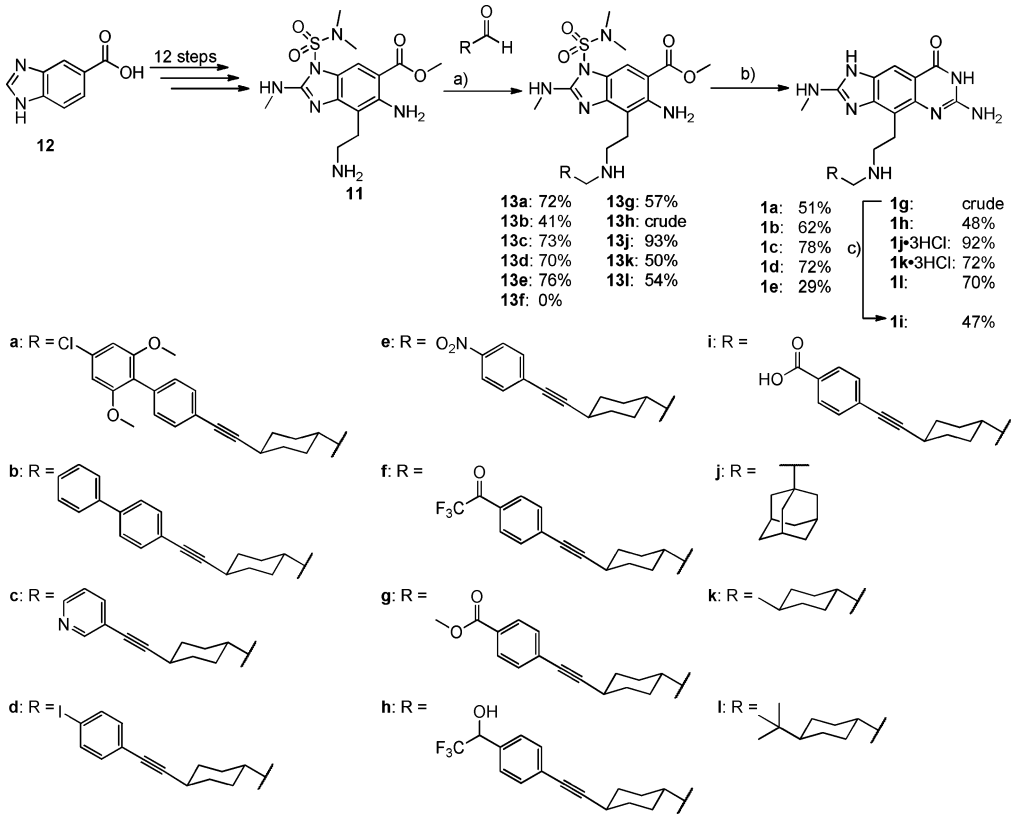
Compared to the parent scaffold **1m**, compound **1k** expands the terminal cyclohexyl ring by a methyl group. It achieves an affinity of 1.2 nM (Table 1). Further expansion by the sterically demanding *tert*-butyl group (compound **1l**) reduces affinity by 20-fold compared to **1k**. Surprisingly, the bulky adamantyl moiety of **1j** shows significantly improved affinity (K_d = 1 nM) compared to **1l** (K_d = 25 nM). An explanation of this affinity gain is provided by the cocrystal structure (see below). Regarding the series of alkyne ligands (**1a−e,h,i**), no dramatic affinity breakdown is experienced compared to the parent compound **1m**, although ligand efficiency is reduced by the large attached substituents. Within the alkyne series, ligands with shorter side chains (**1c,e,i**) tend to exhibit better affinity (4−14 nM), whereas longer and bulkier decorations (**1a,b,h**) result in an affinity drop toward the two-digit nanomolar range (24−36 nM).

Crystal Structure Analysis. To analyze ligand binding modes and perturbance of the dimer interface, we performed cocrystal structure analyses with **1b,c,j** at 1.45−1.74 Å resolution. Furthermore, **1c** has been soaked into protein crystals, and a data set has been collected at 1.62 Å resolution. Detailed information about the experimental conditions are given in the Methods section.

The binding mode of the tricyclic aromatic 2-amino-*lin*-benzoguanine moiety has been extensively described.[18,24] Also here, the *lin*-benzoguanine core of **1b,c,j** forms H-bonds to Asp102, Asp156, and Gln203 and the backbone of Gly230, Leu231, and Ala232 of the guanine recognition pocket (Figures 3 and 4). The interaction to Asp102 is charge-assisted, as N(5) of the tricycle is assumed to be protonated, while Asp102 is, most likely, deprotonated (Table 1).[25] As the 2-aminoimidazole is assumed to be protonated, the hydrogen bond to the Leu231 backbone C=O group has also to be classified as charge-assisted.[18,24] Furthermore, the ligand undergoes favorable π-stacking with Tyr106 and Met260 both flanking the tricycle.[18] The C(4)-substituent points into the ribose-34 pocket (Table 1). A C(4)-ethylamino linker has been established as most promising to serve this purpose with high binding affinity.[11] A conserved water cluster between the two facing Asp residues 102 and 280 is partly displaced, and binding of *lin*-benzoguanines induces a *cis*-peptide bond flip between Val262 and Gly263.[24]

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

Scheme 2. Synthesis of Ligands 1a−e,g−l[a]



[a]Reagents and conditions: (a) RCHO, $NaBH(OAc)_3$, CH_2Cl_2, 0 to 25 °C, 16−19 h; 41−93%. (b) Chloroformamidinium chloride, Me_2SO_2, 130 °C, 1−2 h; 29−92%. (c) $LiOH·H_2O$, H_2O, 100 °C, 30 min; 47% (from **13g**).

Cocrystal Structure of 1j. Well-defined electron density is apparent for the entire adamantyl derivative **1j** (Figure 3A). The adamantyl moiety fills the ribose-34 pocket with several significant ligand-induced rearrangements. This leads to extended hydrophobic contacts compared to the parent structure **1m** (Figure 3B, PDB ID: 3EOS[11]). The adamantyl moiety of **1j** cannot occupy the same position as the cyclohexyl portion of **1m** due to spatial restrictions imposed by Asn70 and Gln107. Hence it is shifted by 0.8 Å and tilted by 74° compared to the cyclohexyl portion of **1m**. In addition, the bulky adamantyl moiety pushes Gln107 about 3 Å off from the rim of the pocket, which translates into a rearrangement of residues 108−115. This region is known to be flexible.[9] Furthermore, the side chain of His73 adopts an orientation that expands the hydrophobic ribose-34 pocket. Additional pocket enlargements are accomplished by reorientations of Thr47 and Asn70. The former moves its hydrophilic carboxamide group away from the ligand and presents the more apolar methylene portion toward the adamantyl moiety. Thr47 adopts a not yet described orientation by being shifted ca. 2 Å compared to the apo structure (PDB ID: 1P0D[7]). Whereas its side chain OH group is oriented toward the ribose-34 pocket in the apo structure, in the complex with **1j**, the apolar methyl group experiences hydrophobic interactions with the adamantyl moiety (Figure 3). In consequence, the OH side chain of Thr47 is now directed toward the dimer interface and forms a hydrogen bond to Lys52. The latter contact resulting from the reorientation is the only structural impact created by **1j** on the targeted $\beta1\alpha1$-loop. Especially the Lys52···Glu339′ salt bridge is not affected by this ligand (Figure 3A).

Cocrystal Structure of 1c. The cocrystal structure with **1c** was obtained at 1.59 Å resolution (Figure 4A). The cell parameters deviate in the *a* axis about 6 Å and in the monoclinic β angle by ca. 2° from the value usually observed for TGT crystals in space group *C*2 (Table 2). A similar shift has been reported for the TGT(Lys52Met) mutant showing disorder in the loop-helix motif.[8] After refinement of the protein portion, the $|F_o| - |F_c|$ difference electron density clearly indicates the tricyclic scaffold and major parts of the C(4)-substituent. However, the terminal, ethyne-linked pyridyl moiety is, even at lower contour level, not detectable in the difference map. We therefore assumed either high residual mobility or pronounced scatter over multiple configurations. The central cyclohexyl ring of the C(4)-substituent adopts a chair conformation with the ethyne attached in favorable equatorial position. With respect to the *B*-values of the ligand atoms, the cyclohexyl ethyne portion refines to nearly 2-fold larger values compared to the remaining ligand part. This indicates enhanced flexibility of the C(4)-substituent with increasing distance from the parent scaffold (Figure 4A). This finding appears reasonable taking into account that the C(4)-substituent of **1c** induces massive conformational changes within the ribose-34 pocket, involving helix $\alpha1$ and the preceding $\beta1\alpha1$-loop. No clearly defined electron density can be assigned to the 16 residues forming this loop-helix motif (Ala48−Gly63). Also here, enhanced flexibility or scatter over multiple conformations is assumed, and none of the directed interactions between the two monomers (Ala48···His333′, Ala49···Tyr330′, Lys52···Glu339′, Lys55···Glu348′, Glu57···Lys325′) are visible. Interestingly, the residues residing

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

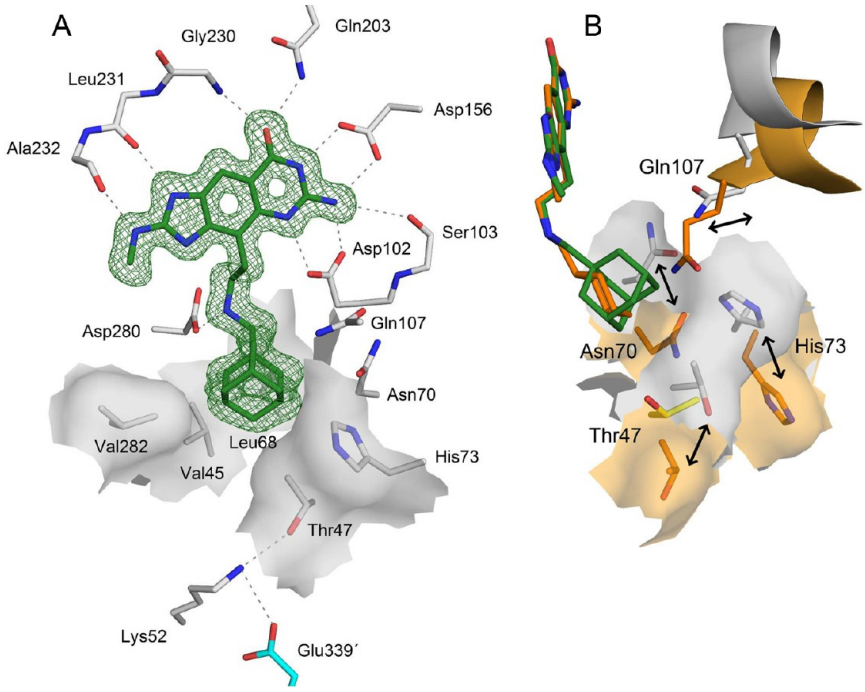


Figure 3. Cocrystal structure of **1j** bound to the active site of TGT (PDB ID: 4FPS). (A) The ligand is well-defined in the difference electron density contoured at the 2σ level (green mesh). This characteristic applies to all figures unless stated otherwise. The hydrophobic bottom of the ribose-34 pocket (shown as sticks and gray surface) encompasses the adamantyl moiety extensively. Thr47 and Asn70 expose their hydrophobic side chain portions toward the adamantyl moiety. The aromatic face of the imidazole moiety of His73 forms hydrophobic contacts to the adamantyl substituent. Lys52 is hydrogen-bonded to Thr47 and Glu339′. Lys52 and Glu339′ occupy crystallographically conserved positions and mutually interact. For reasons of clarity, the geometrically conserved secondary structures across the dimer interface are not shown. Indicated hydrogen bonds exhibit distances of 2.7−3.7 Å. Hydrophobic contact distances between the adamantyl moiety and the surrounding hydrophobic pocket fall between 3.3 and 4.2 Å. (B) Superposition of the crystal structure with **1j** and **1m** (PDB ID: 3EOS[11]) and the apo structure (PDB ID: 1P0D[7]). Residues 107−115 are shown as cartoon. Protein and surface of the complex with **1j** are shown in gray; that with **1m** is in orange. For clarity reasons, only the Thr47 side chain of the apo structure is shown in yellow. The hydrophobic contact surface interacting with the ligand is increased for the complex with **1j** by inducing altered orientations of Thr47, Asn70, and His73. Differences to the complex with **1m** are highlighted by black arrows. Gln107 is pushed backward by ligand **1j**, leading to a shift of the adjacent helix. The orientation of Thr47 in the complex with **1j** also differs from the apo structure (yellow).

on the corresponding dimer mate, except Lys325′ and Glu348′, are all well-defined (Figure 4A). Obviously, the integrity of the helix-turn-helix motif on the contacting monomer is not affected when the geometry on the interacting dimer mate is perturbed. Particularly Glu339′, located directly adjacent to the ligand's C(4)-substituent, is only marginally affected as indicated by slightly increased *B*-values of its carboxylate oxygen atoms (B_{OE1} = 51 Å^2, B_{OE2} = 35 Å^2). The side chain of the neighboring Ile340′ is not visible in the electron density. Apparently, this is the only impact on the dimer mate induced by the ligand. The hydrophobic contact area formed between the aromatic triad (Trp326′, Tyr330′, His333′) and Leu74, Pro78, Phe92, and Met93 is not affected (Figure 5). No water molecules are detectable within the hydrophobic interface. The $3|F_o| - 2|F_c|$ electron density map explicitly reveals Val45, Gly46, and Thr47 at the beginning of the $\beta1\alpha1$-loop in altered conformation compared to the apo structure (Figure 4A). This transition is probably caused by a too close contact between the C(4)-substituent of **1c** and the side chain of Val45. In the novel orientation, Val45 undergoes hydrophobic interactions with Val282. The side chains of Asn70 and Gln107 flank the ligand's cyclohexyl ring. Most likely induced by the reorientation of Gln107, His73 is rotated downward and undergoes a hydrophobic interaction to His333′. Furthermore, **1c** induces conformational changes of the residues Gly69−His73. The backbone atoms are shifted about 2 Å in space, thereby slightly expanding the ribose-34 pocket (Figure 4B). The above-mentioned *a*-cell axis reduction corresponds to a movement of the dimer mate into the emerging space resulting from the structural collapse of the $\beta1\alpha1$-loop (Supplementary Figure 3SI). This occurs without any structural impact on the integrity of the helix-turn-helix motif residing on the shifted dimer mate.

Soaked Crystal Structure of 1c. A further structure with **1c** was obtained by soaking the ligand into the protein crystal instead of cocrystallization. A comparison with the above presented cocrystal structure reveals significant differences (Figure 4B). The cell dimensions of the soaked structure do not deviate from the values usually found (Table 2). In the $|F_o| - |F_c|$ electron density map, an even smaller portion of the C(4)-substituent of the soaked ligand **1c** is visible. Its cyclohexyl ring and ethyne-linked pyridyl moiety are, even at lower contour level, not detectable in the difference map. In contrast to the cocrystal structure, helix $\alpha1$ and the adjacent $\beta1\alpha1$-loop are almost completely defined, and only Thr47 and Ala48 lack clearly visible electron density. The alignment of the soaked structure of **1c** with the apo protein (PDB ID: 1P0D[7]) demonstrates that, independent of the penetrating ligand, the loop-helix motif adopts an unchanged orientation. The conformation of Val45 is clearly defined in the soaked structure and matches well with that found in the apo structure. Nevertheless, it is incompatible with the placement of the ligand found in the cocrystal structure, as a too close contact

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

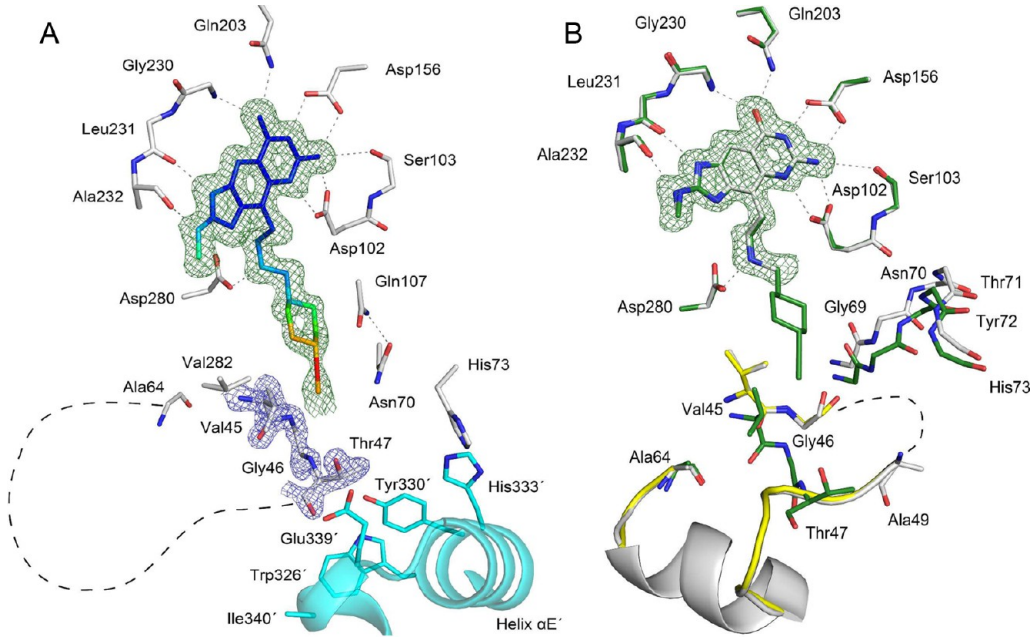


Figure 4. (A) Cocrystal structure of **1c** bound in the active site of TGT (PDB ID: 4FR6). The displayed ligand portion is well-defined by the difference electron density at 2σ; superimposed the assigned *B*-values are indicted by blue (low *B*-values) to red color (high *B*-values). The average *B*-factors indicate an increasing flexibility toward the cyclohexyl ethinyl moiety compared to the *lin*-benzoguanine portion of the ligand ($B_{\text{cyclohexyl-ethinyl}}$ = 28 Å^2, B_{ligand} = 16 Å^2). Observed hydrogen bond distances are between 2.6−3.7 Å. The ligand induces conformational changes within the $\beta1\alpha1$-loop and helix $\alpha1$. Val45, Gly46, and Thr47 are pushed to the side by the ligand ($3|F_o|-2|F_c|$ electron density map is shown at 1.5σ as blue mesh). The adjacent part of the $\beta1\alpha1$-loop as well as the entire helix $\alpha1$ up to Ala64 become disordered (arbitrarily indicated by a black dashed line). The corresponding helix-turn-helix motif is well-defined (shown as cyan cartoon). The side chains of Lys325′, Ile340′, and Glu348′ are disordered, while Trp326′, Tyr330′, His333′, and Glu339′ are well-defined. Asn70 and Gln107 flank the ligand's cyclohexyl ring and induce a rotation of His73, thus enabling hydrophobic interactions with His333′. Val45 and Val282 form a hydrophobic contact (4.2 Å). (B) Soaking structure of **1c** (carbon atoms and cartoon in gray, PDB ID: 4FSA) aligned with the cocrystal structure of **1c** (green) and the apo structure (yellow, PDB ID: 1P0D[7]). The difference electron density defines the part of the ligand indicated in the soaked structure of **1c**. The ligand's *lin*-benzoguanine portion adopts an unchanged binding pose independent of the crystallization protocol. The portion of the C(4)-substituent, visible in the electron density of the soaked structure, is further limited compared to that in the cocrystallized structure. The orientations of Val45 and Gly46 are virtually identical in the soaked and apo structure (gray/yellow). Except for a disorder of Thr47 and Ala48 in the soaked structure of **1c** (implied by the black dashed line) an ordered arrangement of the $\beta1\alpha1$-loop and helix $\alpha1$ superimposed well with that in the apo structure (for clarity only shown as cartoon). However, the orientation of Val45−Thr47 differs strongly in the cocrystal structure of **1c** (green), where also the adjacent $\beta1\alpha1$-loop/helix $\alpha1$ motif is fully disordered (left). The placement of Val45 in the soaked structure would clash with the orientation of the cyclohexyl moiety of **1c** in the cocrystallized complex. The backbone atoms of Gly69−His73 are about 2 Å shifted between the two structures of different protocols. For reasons of clarity, the Gly69−His73 portion as found in the apo structure is not shown.

between the ligand's cyclohexyl moiety and Val45-side chain (2.8 Å) would result. Thus, an altered conformation of Val45 is enforced and an enhanced disorder of the C(4)-substituent in the soaked structure must be given (Figure 4B). Remarkably, the soaked structure lacks well-defined electron density for the residues Gln107−Ser112, which are visible in the density of the cocrystal structure (not shown in Figure 4). Most likely, this is also a consequence of the increased disorder of the ligand's C(4)-substituent. Altogether, the impact of the bound ligand **1c** on the protein structure differs according to whether the complex is formed in solution and then crystallized or the ligand penetrates subsequently into prepacked crystals in the solid state.

Cocrystal structures of 1a,b,d,e,h,i,k,l,. The remarkable impact of ligand binding on the geometry of the loop-helix motif and the surprisingly small changes in the integrity of the contacting helix-turn-helix motif is further evidenced by the cocrystal structure with **1b** (for a detailed description, see Supporting Information). Additional data sets collected for cocrystal structures with **1a,d,e,h,i,k,l** confirm the ligand-induced effects on the dimer interface and, therefore, will not be discussed further.

Mass Spectrometry. Noncovalent ESI mass spectrometry enables accurate mass determination by maintaining non-covalently composed assemblies such as protein oligomers or protein−ligand complexes in the gas phase. We applied automated, chip-based nanoelectrospray mass spectrometry (nanoESI-MS) to analyze the influence of the crystallographically studied ligands on the oligomerization state of TGT (Figure 6). Experiments were performed at different protein concentrations (5 and 1 μM, calculated as TGT monomer). Studies of the noncomplexed wild-type protein exhibit a mass that can be assigned to a dimer with one Zn^{2+} bound per monomer.[8] The measured mass of 85608.9 ± 3.7 Da is in good agreement with the expected mass of 85607 Da, and no significant concentration dependency of the monomer/dimer ratio was experienced (Figure 6a,b). The influence of binding **1b** and **1c** was analyzed by preincubating TGT with a 10-fold molar excess of the ligands. Both ligands show no significant effect on the monomer/dimer ratio at 5 μM TGT concentration (Figure 6c,e). In contrast, both ligands are able to shift the ratio slightly toward the monomer applying lower protein concentrations (Figure 6d,f). Ligand **1b** induces a monomer portion of 20%, while ligand **1c** induces 24%

Table 2. X-ray Crystal Structures: Data Collection and Refinement Statistics

	1b	1c	1c	1j
PDB ID	4FR1	4FR6	4FSA	4FPS
Data Collection and Processing				
no. crystals used	1	1	1	1
λ [Å]	0.91841	0.91841	0.91841	0.91841
space group	*C*2	*C*2	*C*2	*C*2
unit cell parameters [Å]				
a [Å]	84.8	84.8	91.1	90.0
b [Å]	64.8	64.8	64.9	64.7
c [Å]	71.4	71.3	70.8	70.7
β [deg]	94.0	93.0	96.3	96.0
Diffraction Data				
resolution range [Å]	30−1.74	30−1.59	50−1.62	30−1.45
unique reflections	39417 (1864)[a]	51448 (2576)[a]	51849 (2587)[a]	70951 (3478)[a]
$R(I)_{sym}$ [%]	7.9 (27.2)[a]	3.5 (9.1)[a]	5.0 (34.6)[a]	6.4 (46.1)[a]
completeness [%]	99.2 (93.0)[a]	99.4 (99.3)[a]	99.5 (99.6)[a]	99.1 (97.6)[a]
redundancy	3.0 (2.6)[a]	3.3 (3.3)[a]	2.9 (2.9)[a]	2.8 (2.7)[a]
I/σ (*I*)	13.4 (3.3)[a]	26.1 (11.9)[a]	19.2 (3.4)[a]	16.9 (2.1)[a]
Refinement				
program used for refinement	SHELXL	SHELXL	SHELXL	SHELXL
resolution range [Å]	30−1.74	30−1.59	50−1.62	30−1.45
reflections used in refinement	36165	48267	47606	67344
final *R* values				
R_{free} (F_o; $F_o > 4\sigma F_o$)[c]	23.3 (21.9)[a]	19.5 (19.2)[a]	20.7 (19.8)[a]	19.7 (18.6)[a]
R_{work} (F_o; $F_o > 4\sigma F_o$)[d]	18.4 (17.4)[a]	16.3 (16.1)[a]	17.1 (16.3)[a]	16.7 (15.7)[a]
no. of atoms (non-hydrogen)				
protein atoms	2556	2611	2626	2736
water molecules	216	271	242	259
ligand atoms	27	29	21	31
RMSD, angle [deg]	1.9	2.2	2.1	2.3
RMSD, bond [Å]	0.007	0.010	0.008	0.010
Ramachandran plot[e]				
most favored regions [%]	94.3	94.4	95.9	94.8
additionally allowed regions [%]	5.3	5.3	3.8	4.9
generously allowed regions [%]	0.4	0.3	0.3	0.3
mean *B*-factors [Å^2]				
protein atoms	18.0	15.9	18.8	15.7
water molecules	27.9	28.4	30.0	27.7
ligand atoms	16.7	15.9	19.4	12.9

[a]Values in parentheses are statistics for the highest resolution shell. [b]$R(I)_{sym} = [\sum_h \sum_i |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_i I_i(h)] \times 100$, where $\langle I(h)\rangle$ 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 5% of data. [e]Calculated with *Procheck*.[45]

monomer. Experiments beyond 1 μM TGT concentration drastically reduced spectra quality and affected mass precision measurements. This concentration dependent ratio shift has already been reported for the TGT(Lys52Met) mutant[8] and underlines that **1b** and **1c** cause destabilization of the TGT homodimer without complete disruption to fully separated monomers. Regarding the binding stoichiometries, **1c** binds predominantly in 2:2 ratio to the TGT dimer. No fully uncomplexed TGT dimer is detectable, and only traces of a TGT dimer:ligand complex with just one bound ligand molecule are observed. The mass spectra with ligand **1b** exhibit, apart from uncomplexed TGT dimer, signals for a 2:1 TGT dimer:ligand complex and only traces of a 2:2 TGT dimer:ligand complex, which correlates with the higher K_d value of **1b** compared to that of **1c** (Table 1).

Modeling and MD Simulation. Based on the cocrystal structures with **1b** and **1c**, computer simulations were performed using the TGT dimer to complement the reported crystallographic findings and to rationalize the observed disorders. One aspect was to screen for alternative conformations of the loop-helix motif that is evidenced by crystallography to be either highly mobile or scattered over multiple arrangements. The other issue was to suggest for **1b** and **1c** possible orientations of the ligand's C(4)-substituent that were not traced in the electron density of the crystallized complexes.

To identify alternative loop conformations, several geometries of the target sequence Ala48−Gly63 were modeled on the basis of the cocrystal structure of **1c**. Putative start structures of the loop were generated using the protocol of ModLoop.[26] This protocol starts with initial random conformations, generates a number of independently optimized geometries, and finally reports for each run the one with the lowest energy score.[27] Accordingly, a set of loop conformers has been generated and subjected to restrained molecular dynamic simulations. The residues Ala48−Gly63, which were

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

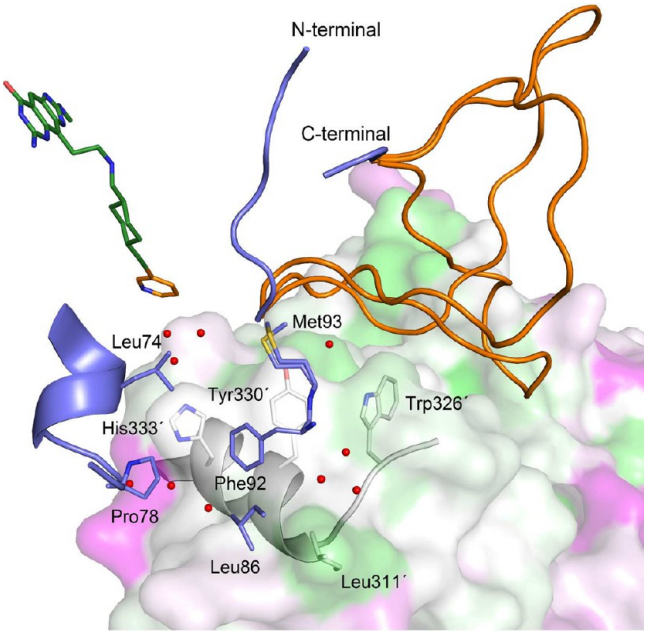


Figure 5. Hydrophobic contact area (hydrophobic (green) to hydrophilic (pink)) across the TGT dimer interface of the cocrystal structure of **1c** (PDB ID: 4FR6) with several model-built conformations for the crystallographically invisible loop-helix motif (Ala48−Gly63). In addition, the model-built geometry of the terminal pyridyl moiety of **1c** is shown. Color coding: monomer A, blue; monomer B, gray; model-built portions in orange. Waters are shown as red spheres. Trp326′ and Tyr330′ of the dimer mate are well-shielded by the only slightly scattered N-terminal part of the modeled loops, whereas His333′ remains rather solvent-exposed. The C-terminal part of the modeled loops orients toward the surrounding solvent and scatters over multiple conformations. The aromatic triad (Trp326′, Tyr330′, and His333′) forms hydrophobic contacts across the interface with the well-defined residues of monomer A (Leu74, Pro78, Phe92, Met93; blue). In addition Leu86 undergoes hydrophobic interactions with Leu311′. All hydrophobic contacts are between 3.9−4.6 Å. No waters are observed within this hydrophobic contact area. The shortest distance between the residues of the aromatic triad and those of modeled loops are ca. 4 Å for Trp326′, ca. 5 Å for Tyr330′, and ca. 6 Å for His333′.

disordered in the crystal structures, were allowed to move freely, whereas the adjacent crystallographically ordered residues were fixed (for details, see Methods). During the MD simulation, the terminal pyridyl moiety of **1c**, also not visible in the crystallographically determined electron density, was covalently attached in approximate position at the end of the ethyne linker still seen in the electron density (Figure 4A).

Starting off with residue 51 (Figure 1), the loop fans out in space in the simulations and adopts a large variety of conformations (Figure 5). This suggests rather unrestricted movement, which appears reasonable as the loop protrudes directly into the solvent. The first three loop residues 48−50 remain spatially rather restricted and contact other residues of the adjacent protein. The reduced spatial mobility of Ala48−Val51 translates across the dimer interface to the adjacent hydrophobic triad Trp326′, Tyr330′, and His333′ on the dimer mate.

The actual binding poses of the terminal pyridyl and biphenyl moiety of **1c** and **1b**, respectively, are not visible in the electron density (Figure 4A and Supplementary Figure 6SI), and unexpectedly, the bound ligands do not perturb the spatial position of Glu339′. Our initial design hypothesis was to hit this residue. However, as the ligands are not decomposed upon binding (cf. mass spectrometry) and their *lin*-benzoguanine moiety up to the cyclohexyl substituent is well-detected in the electron density, our spiking warheads must be accommodated in the dimer complexes. Visual inspection of consecutive frames along the MD trajectory suggests placement of the extended ethinyl substituent with a slight overall bending. Thereby, the warhead does not crash, as planned, into Glu339′ but slightly passes this residue, which is obviously strictly stabilized by its local protein environment (see below). Consulting the geometry of open-chain ethyne derivatives in small molecule crystal structures (CSD)[28] confirms that deviations from linearity up to 10° are observed (Supplementary Figure 7SI). These findings support our evidence from the simulations and suggest accommodation of the spiking ligands with slightly overall bent, however still favorable, geometry (cf. **1a−e,h,i**) bind with no major loss in affinity (Table 1, Supplementary Figure 8SI). The binding of our ligands makes a clear impact on dimer packing as obvious from the expansion of the crystal packing (cf. shifts along the *a*-axis). These changes could even increase in solution where molecular motion enhances the adaptability of proteins.

Comparative Analysis of the Helix-Turn-Helix Motif. Our initial design hypothesis was stimulated by the ligand-induced collapse of an α-helix followed by a subsequent dimer dissociation reported by McMillan et al. for the iNOS oxygenase homodimer interface.[29] Apart from the perturbance of the $\beta1\alpha1$-loop, our spiking ligands were designed to interfere and possible destroy the contact formed to Glu339′. However, in our TGT case, the cocrystal structures of **1b** and **1c** show Glu339′ in well-defined geometry at the apex of a helix-turn-helix motif, whereas the spiking ligand substituents are disordered and scattered over multiple configurations (Figure 4A and Supplementary Figure 6SI). Obviously, neither the position of this glutamate nor the geometry of the adjacent helices are affected. We therefore assume that the orientation of Glu339 must be significantly stabilized by the local protein architecture.

In order to evidence this hypothesis, we analyzed the TGT_{20} data set. Glu339 is part of a turn, linking helix αE and helix αF. The relative orientation of both helices is stabilized by hydrophobic interactions (Figure 7A). Mutually intertwined, Leu334 interacts with Leu341 and Leu345, and Ile331 forms a close van der Waals contact to Met346. This arrangement reminds to some degree of a leucine zipper motif used to stabilize a helix-turn-helix element in DNA recognition.[30] The running directions of both helices are opposed. Whereas helix αE faces the dimer interface with its C-terminus, helix αF orients its N-terminus in this direction.

Wada and Hol suggested the description of α-helices as macro dipoles resulting from multiple parallel-oriented peptide bonds and leading to a more positive dipole end at the N-terminus and a more negative end at the C-terminus.[31,32] Following this idea, Glu339 takes a dual function as a capping residue. With respect to helix αE, the backbone NH of Glu339 forms a hydrogen bond to the otherwise unsaturated backbone C═O of Leu334 (2.7 Å). In contrast, the C═O functionalities of Arg336 and Ala337 are not saturated by any binding partners of the protein. The side chain of Glu339 is oriented toward the apex of helix αF. This agrees with a study of Forsyth et al. where a significantly higher population of negatively charged residues such as Asp and Glu has been observed at N-termini of

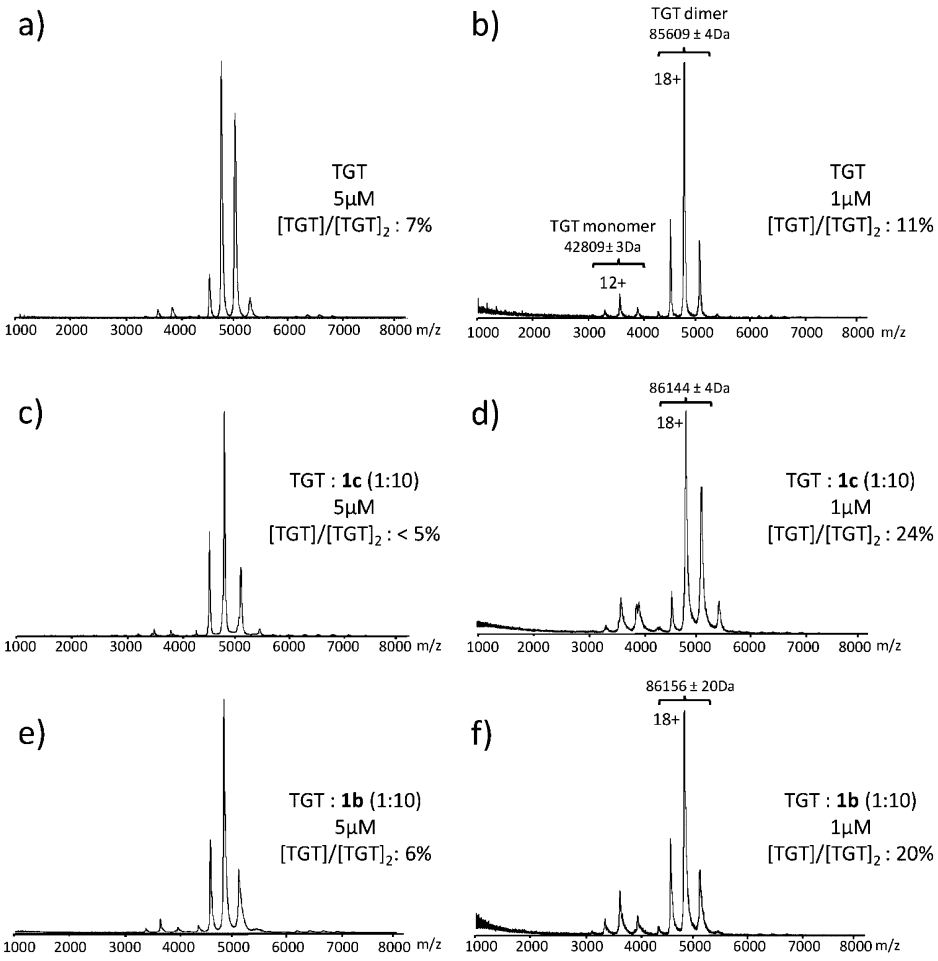


Figure 6. ESI mass spectrometry analysis of TGT wild-type (a,b) and incubated with ligand **1c** (c,d) or **1b** (e,f). TGT was analyzed either at 5 μM (a,c,e) or at 1 μM (b,d,f) in ammonium acetate (1 M, pH 7.5). The monomer/dimer ratio was calculated after integrating over all charged states intensities of each species.

α-helices.[33] Furthermore, it has been reported that Asp and Glu experience a pK_a shift of ca. 0.6 units toward stronger acidity next to helical N-terminal positions, which is caused by a stronger polarization of the acidic group.[34] Supposedly, their placement to the terminal end ("capping position") of the helix results in a stabilization of the helical arrangement. Strong directed H-bond interactions of the Glu339′ carboxylate side chain formed to the uncomplexed terminal NH backbone groups of helix αF′ seem to be lacking. The only direct interaction Glu339 C═O forms to helix αF′ is the weak hydrogen bond to the backbone NH of Ala343 (3.7 Å).

Obviously, the Glu339 side chain is stabilized by several water molecules found to be conserved in the TGT_{20} reference set. The upper part of the helix-turn-helix motif is solvated by four water molecules, which are conserved to different extent in the TGT_{20} data sets (Figure 7B; for water identifiers, see Supplementary Table 1SI). The water molecules in clusters A and B are present in all analyzed structures with minimal positional deviations. The water positions in cluster C are more scattered and only in 45% of all cases populated. Waters in cluster D are present in 75% of all structures and exhibit a small positional variance. Water molecules A and B are involved in the capping of helix αF.

In summary, the highly conserved water network presumably stabilizes the capping positions between the two helices and favors conserved placement of Glu339′. Most likely, the acidity of Glu339 is enhanced at the capping position and this residue is deprotonated. Thus, the described hydrogen bonds are charge-assisted and in consequence of increased stability. The cocrystal structures with **1b** and **1c** reveal a thinned-out water network. The presence of waters B, C, and D is no longer indicated in the electron density, whereas the interstitial water A still maintains a contact between the backbone NH of Leu341 and the carboxylate of Glu339. The latter is assumed as crucial for the spatial stabilization of the Glu339 side chain, which remains virtually unperturbed upon the interference with the spiking ligands.

Comparative Discussion and Conclusion. It has been suggested that the TGT dimer interface formation and stability is strongly determined by a structural motif composed by a 10-residue $\beta1\alpha1$-loop (Val45−Leu54) and a subsequent 9-residue helix $\alpha1$ (Lys55−Gly63).[8,9,11] The rationale to assume that this structural element is crucial for dimer stability was based on the analysis of multiple crystal structures, which show this loop-helix motif to be highly conserved and strongly involved in almost all directed interactions (three salt bridges, two H-bonds) across the interface. Our hypothesis is supported by previous observations in literature. For example, Chako et al.[35] reported on single-site mutations in a protein−protein interface that have dramatic impact on stability. Within a 51 residue-composed heterodimer interface between the Fab HyHEL-5 and chicken lysozyme, the conservative Arg/Lys exchange

dx.doi.org/10.1021/cb400020b | ACS Chem. Biol. 2013, 8, 1163−1178

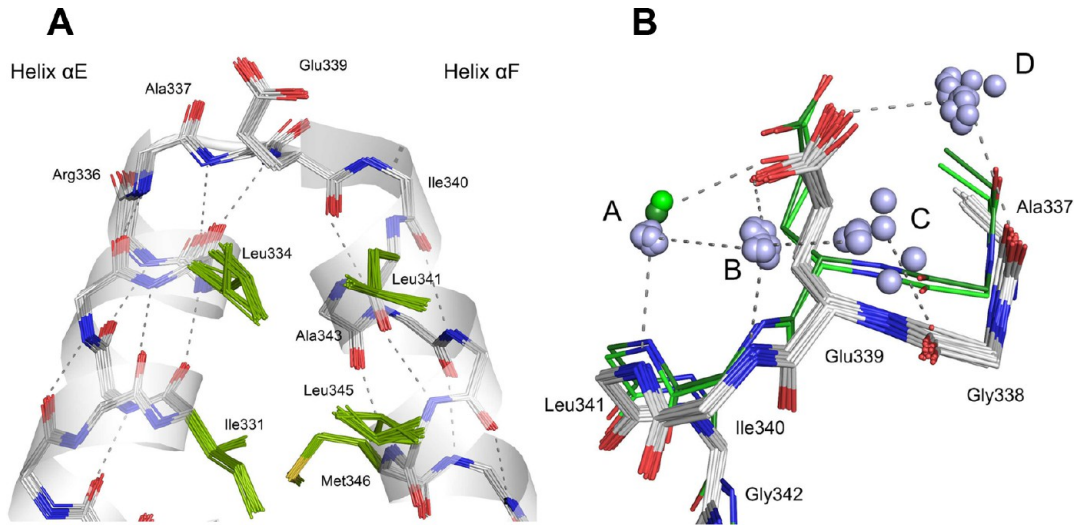


Figure 7. Helix-turn-helix motif and interstitial water molecules found in the TGT_{20} reference set of *Z. mobilis* TGT crystal structures (for PDB IDs see Methods). (A) The turn motif is formed by Ala337, Gly338, and Glu339, preceding the C-terminus of helix αE, whereas the N-terminus of helix αF follows subsequently; secondary structures are indicated as cartoons. For reasons of clarity, side chains are only shown for the residues involved in hydrophobic interactions (green) and for Glu339. Hydrophobic contacts between helix αE (Ile331, Leu334) and helix αF (Leu341, Leu345, Met346) fall between 3.5 and 4.3 Å, and displayed hydrogen bonds are between 2.7−3.7 Å. (B) Conserved water positions (A−D) as found in the TGT_{20} reference set (white, waters: light-blue spheres) superimposed with cocrystal structures of **1b** and **1c** (green, PDB IDs: 4FR1, 4FR6). Only the side chain of Glu339 is displayed. Waters in cluster A form H-bonds to the backbone NH of Leu341 (~2.9 Å), the carboxylate of Glu339 (~2.7 Å), and waters in B (~3.5 Å). Waters in B are involved in H-bonds to the backbone NH of Ile340 (~2.8 Å), the carboxylate of Glu339 (~2.6 Å), and waters in C (~2.6 Å); waters in C interact with the backbone C═O of Gly338 (~3.8 Å); and waters in D are involved in the capping of helix αE, forming hydrogen bonds to the backbone C═O of Ala337 (~3.0 Å) and the carboxylate of Glu339 (~2.8 Å). In the cocrystal structures with **1b** and **1c**, the water network is reduced to solely water A mediating a contact between Leu341 and Glu339, and the other water molecules are obviously repelled from these complexes.

causes interface disruption, presumably triggered by a change in the hydrogen-bonding pattern.

All the more, we were surprised that in our TGT example, the ligands **1j−l** with a curtailed hydrophobic C(4)-substituent solely perturb the assumed crucial loop-helix motif and cannot create massive structural perturbance of the dimer interface. The ligands **1a−e,h,i** were designed in a way to spike and directly interfere with the loop-helix motif to finally hit the salt-bridge connected Glu339′. This residue is exposed on the interface mate in prominent position at the apex between two nearly parallel oriented helices in the contact area. The spiking ligands **1b** and **1c** influence the geometry of the targeted loop-helix motif in the determined cocrystal structures. Noncovalent mass spectrometry confirms that these ligands induce homodimer destabilization in solution. Interestingly enough, we showed that the structures determined on crystals obtained by soaking and cocrystallization of **1c** do show important differences that shed light on the structural properties of the loop-helix motif in solution. If the ligand is bound to the solvated protein prior to the crystallization process (cocrystallization), it induces or sustains disorder of the loop-helix motif likely present in solution. In consequence, the crystallized complex exhibits this disorder also in the solid state. In contrast, if the uncomplexed protein is crystallized first, the loop-helix motif accommodates unperturbed and thus adopts a well-ordered geometry in the solid state. This geometry is hardly affected once the ligand is soaked into the premanufactured crystals. Instead, the ligand has to arrange with the given geometry of the fixed loop in the crystal. It remains speculative whether the loop-helix motif is scattered over multiple arrangements in solution prior to ligand binding or whether it adopts this state only once the spiking ligand is bound. It might possibly be that the observed order/disorder transition is related to the monomer/dimer equilibrium. At least we can conclude that the uncomplexed dimeric protein exhibits ordered geometry in the solid state, whereas with bound ligand, the loop adopts multiple geometries in solution that are manifested in the crystal. In any case, ordered and disordered state must be energetically close together as the ligands **1a−e,h,i** hardly show any loss in binding affinity compared to the analogs with much shorter C(4)-substituents.

Nonetheless, mass spectrometry analyses emphasize that structural interference of **1b** and **1c** with the interface region results in a gradual destabilization of the TGT homodimer (Figure 6). Under equilibrium conditions in solution about 20−25% dimer disruption is observed. A strictly required preorganization of this tentatively flexible loop-helix motif is obviously not required to form the TGT dimer interface. On first glance, this fact is surprising as most of the interface-forming amino acids, involved in H-bonds and salt bridges, reside on this motif. We therefore conclude that the adopted geometry of the loop-helix motif operates as a kind of toggle switch in the protein−protein interface formation but is of reduced importance with respect to the establishment of directed interactions between both monomers.

Given that the dimer stability is not determined by the interactions of the loop-helix motif, we turned our focus on a hydrophobic area that is shielded by the $\beta1\alpha1$-loop. This region is formed by a triad of three hydrophobic residues (Trp326′, Tyr330′, His333′) and complemented by Phe92. The aromatic triad residing on the amphipathic helix αE is supposedly a putative interaction hot spot on the TGT dimer interface (Figure 2). Also this hypothesis is supported by results reported in literature. Kussie et al. observed a comparable arrangement

dx.doi.org/10.1021/cb400020b | *ACS Chem. Biol.* 2013, 8, 1163−1178

as putative protein−protein interaction (PPI) hot spot for the interaction of p53 with MDM2.[36] In their example, a similar hydrophobic triad (Phe, Trp, Leu), lies on an amphiphilic α-helix of p53 and solely undergoes van der Waals interactions in a hydrophobic cleft of MDM2. Furthermore, numerous studies elucidated enhanced occurrence of tryptophan and tyrosine within PPI-hot spots.[37−39] Our previously reported slight destabilization of the TGT dimer resulting from the rather conservative Tyr330Phe point mutation also suggests that this hydrophobic contact area is important for dimer stability.[8] The wild-type Tyr330′ acts as a hydrogen-bond donor in the interface, and possibly its phenolic OH function in the aromatic triad also prevents penetration of water molecules into the hydrophobic contact region. Important enough, accommodation of **1b** and **1c** perturbs the loop-helix motif, but the arrangement of the hydrophobic triad remains unaffected (Figure 5, Supplementary Figure 4SI). This fact underlines the hypothesis that the hydrophobic cluster of aromatic amino acids displays the hot spot of the dimer formation.

Our initial design was stimulated by the hypothesis to launch an expanded C(4)-substituent as a kind of warhead into the interface region concurrently targeting the spatial position of the salt-bridge forming Glu339′ present on the adjacent dimer interface mate. Surprisingly, Glu339′ remains virtually unaffected at its location. First of all, we could evidence by our MD simulations that the long spiking ligands adopt overall a slightly curved geometry avoiding any steric clashes in this region. Furthermore, high structural integrity is given for the helix αE-turn-helix αF motif on the adjacent interface mate in all examples of the TGT_{20} test set. Especially, the high spatial conservation of the Glu339′ side chain is surprising. Several highly conserved water molecules, present in all analyzed crystal structures, mediate interactions between the N-terminal backbone of helix αF and the side chain of Glu339. The pronounced stability of the helix-turn-helix motif is of utmost importance for the geometry of the aromatic triad, assumed to be the putative hot spot for dimer formation. The structural integrity of this arrangement even in the presence of the spiking ligands **1b** and **1c** is most likely the reason for the residual dimer stability. Nevertheless, the observed gradual ligand-induced dimer destabilization possibly results from the perturbance of the adjacent 10-residue $\beta1\alpha1$-loop. This loop shields the hydrophobic contact area from solvent access as a kind of lid.

The presented ligand-approach helps to elucidate the importance of structural motifs with respect to the stability of a dimer interface and provides new insights into the architecture of the protein−protein interaction established between the two TGT monomer units.

METHODS

Chemistry. The synthetic details, experimental data, and NMR spectra for all new compounds are given in the Supporting Information.

***Z. mobilis* TGT Crystallization.** Protein was cloned, overexpressed, and purified as described in detail elsewhere.[40,41] Crystals of TGT suitable for data collection were obtained using the hanging-drop, vapor diffusion method at 288 K. TGT was crystallized in 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, 15% (v/v) Me_2SO) was incubated with 1.5 mM inhibitor. Then, 2 μL of this solution was mixed with 2 μL of reservoir solution (100 mM MES, pH 5.5, 1 mM DTT, 10% (v/v) Me_2SO, 13% (w/v) PEG 8000). After 3 weeks of crystal growth, the specimen reached dimensions of about 0.7 × 0.7 × 0.2 mm^3. Compound **1c** was in addition to cocrystallization soaked at a final concentration of 5 mM for 1 day into premanufactured wild-type crystals.

Data Collection. 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 crystals were flash-frozen in liquid nitrogen. All data sets were collected at cryo-conditions (100 K) at BESSY-PSF Beamline 14.2 in Berlin at a wavelength of λ = 0.91841 Å. A Rayonix MX225 CCD-detector was used for data collection. TGT crystals exhibited a monoclinic space group *C*2 with one monomer per asymmetric unit. Data processing and scaling was done using HKL2000. Crystal dimensions, data collection, and processing statistics are given in Table 2.

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.[42] The coordinates of the TGT apo structure 1P0D were used as starting model. Due to the high resolution of all data collections a further refinement using SHELXL-97 was performed.[43] For each refinement step, at least 20 cycles of conjugate gradient minimization were performed with default restraints on bond distances, angles, and *B*-values, and 5% of all data were used to calculate R_{free}. Intermittent model building was performed using COOT.[44] Ligand, water, and glycerol molecules were placed into the difference electron density and once assigned included in the further refinement cycles. Riding hydrogen atoms were added for the protein in a final refinement cycle without using additional parameters. Model analysis was performed using PROCHECK.[45]

PDB Deposition. Coordinate files were deposited in the PDB with the following access ID: 4FPS, 4FR6, 4FSA, and 4FR1.

Validation Data Sets. A randomly selected, nonredundant set of 20 TGT structures (denoted as TGT_{20}, resolutions between 1.28−2.25 Å) was aligned on the basis of a RMSD minimizing sequence alignment by use of the PyMOL Molecular Graphics System. Structures have been visually inspected. PDB ID of the selected data sets are 2Z7K, 3C2Z, 3EOS, 3EOU, 3S1G, 3V0Y, 3RR4, 4FPS, 1P0D, 1Q66, 1R5Y, 1N2V, 1OZM, 1OZQ, 1PUD, 3BL3, 3BLO, 2OKO, 2PWV, and 2QII.

Microscale Thermophoresis Measurements. *Z. mobilis* TGT was adjusted to a concentration of 10 μM and subsequently labeled with Alexa Fluor647 succinimidyl ester at a concentration of 40 mg L^{-1} at RT for 30 min. Labeling reaction was performed in a 500 mM NaCl solution buffered with 50 mM HEPES pH 8.1 (molar dye:protein ratio ≈ 3:1). Unreacted Alexa Fluor647 was removed with a NAP5 sephadex column (GE Healthcare) equilibrated with 2 M NaCl, 1 mM EDTA solution buffered with 10 mM TrisHCl pH 7.8. The label:protein ratio was determined using photometry at 650 and 280 nm. Thereby, a ratio of 0.8 was typically achieved. The Alexa Fluor647-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 305 pM to 1.25 μM. For thermophoresis, each of these solutions was mixed with one volume of Alexa Fluor647-TGT solution, which leads to a final concentration of fluorescence labeled TGT of 10 nM and final ligand concentrations ranging from 153 pM to 625 nM. After 10 min incubation, followed by centrifugation at 10,000 × *g* for 10 min, ca. 2 μL of each solution was filled into Monolith NT Standard treated capillaries (NanoTemper Technologies GmbH). Thermophoresis (including temperature jump) was measured at RT 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 of three independent measurements were averaged and analyzed using Origin 7 (Origin Lab). Curve fitting and K_d calculation was done based on the following equation:[23]

$$T = U + (B - U)\frac{(c_p + c_i + K_d) - \sqrt{(c_p + c_i + K_d)^2 - 4c_p c_i}}{2c_p}$$

with T = thermophoresis signal; U = minimal signal (unbound protein); B = maximal signal (protein saturated with ligand); c_p = concentration of labeled protein; c_i = concentration of ligand; K_d = dissociation constant.

Comparability between the reported K_d values and the competitive inhibition constant K_i values of the alternatively used [8-^{3}H]guanine kinetic assay[46] were confirmed for a subset of four inhibitors. The K_d/K_i ratio of these test compounds differs maximally by a factor of 2.6 (Supplementary Table 2SI).

Loop Prediction. Several loop orientations for the missing amino acid sequence A(48)ATVKALKPETVRATG(63) were generated using the ModLoop web service.[26,27] For this purpose, the cocrystal structure of **1c** (PDB ID: 4FR6) was prepared using *fconv*[47] and completed with the missing sequence information. A putative conformation of the undetermined pyridyl moiety of the ligand C(4)-side chain was modeled using MOE (see below). The plausibility of the modeled loop structures was validated by running short molecular dynamic simulations using Amber 12.[48]

Ligand Modeling. Possible conformations for missing side chain portions of ligand **1b** and **1c** were modeled and protonated with MOE. The ligand was stepwise assembled using the integrated molecule builder tool. In each step, the modeled parts were locally minimized applying the implemented MMFF94x force field. For the modeled part of the fully assembled ligand, a conformational search has been performed, combining the systematically, the stochastically, and the "LowModeMD" method. For each method, the standard search parameters were used; however, the option 'enforce chair conformation' was disabled.

CCDC Database. The conformational search suggested multiple locally minimized conformations. To further rationalize the suggested geometries, a CSD search[28] with respect to deviations of the ethinyl linker from linearity (for **1b** and **1c**) and the assumed torsion angle between the rings in the biphenyl moiety (**1b**) was performed using ConQuest[49] version 1.14. Organometallic complexes, structures with ions, and structures determined from powder material were ignored, and the search was restricted to acyclic compounds. The hits were analyzed using Mercury.[50]

ACD/Laboratories software was used to estimate clog *P* and clog *C* values: *ACD/Laboratories*, Version 12.01; Advanced Chemistry Development, Inc., Toronto, 2009.

Noncovalent NanoESI-MS Experiments. NanoESI-MS analyses were performed in positive ion mode on an electrospray time-of-flight mass spectrometer (Synapt G2 HDMS, Waters, Manchester, U.K.) equipped with a nanoESI source (Triversa Nanomate, Advion Biosciences, Ithaca, NY). Prior to mass spectrometry analysis, TGT sample was buffer exchanged against 1 M ammonium acetate solution buffered at pH 7.5 with ammonia, using 2 cycles of microcentrifuge gel filtration column (Zeba 0.5 mL, Thermo Scientific, Rockford, IL). Protein concentration was determined by Bradford assay.

Mass measurement under denaturing conditions were carried out by diluting samples to 2 μM in water/acetonitrile/formic acid 50:50:1 (V_c = 40 V; P_i = 2.5 mbar). Experiments in nondenaturing conditions were achieved in 1 M ammonium acetate solution buffered at pH 7.5 (V_c = 80 V, P_i = 6 mbar). External calibration was performed over a mass range m/z 500−12000 using a 2 mg mL^{-1} solution of cesium iodide in water/isopropyl alcohol (1:1). In nondenaturing conditions, optimal accelerating voltages applied on sample cone (V_c) and adapted pressure in the interface (P_i) were used to maintain noncovalent interactions while achieving efficient ion desolvatation and transmission. Reported masses were measured at V_c = 120 V, P_i = 6 mbar. MS data were acquired and processed using MassLynx 4.1 (Waters). Both ligands were solubilized in ethanol and incubated at 10 molar excess to TGT sample during 30 min at RT under gentle agitation with a final concentration of ethanol at 0.1%.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary figures and tables, syntheses, and NMR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: (+49) 6421-282-8994. E-mail: klebe@mailer.uni-marburg.de.

Author Contributions

⊥These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

∥Associated with the Center of Synthetic Microbiology, SYNMIKRO, University of Marburg.

■ ACKNOWLEDGMENTS

We are grateful for the generous support of the Deutsche Forschungsgemeinschaft FO 806, KL1204/13-1. Research at the ETH has been supported by grants from the ETH Research Council and the ETH foundation. We thank Dr. B. Bernet (ETH) for help with the interpretation and correction of the spectroscopic data of the new compounds. We further acknowledge the support from the beamline staff at BESSY II in Berlin and a travel grant from the Helmholtz-Zentrum für Materialien und Energie in Berlin.

■ REFERENCES

(1) Robert Koch Institut (2011) *Epidemiol. Bull. 24*, 215−217.

(2) Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K., and Levine, M. M. (1999) Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull. World Health Organ.*, 651−666.

(3) Durand, J. M. B., Dagberg, B., Uhlin, B. E., and Björk, G. R. (2000) Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of *Shigella flexneri*: the expression of the *virF* gene. *Mol. Microbiol. 35*, 924−935.

(4) Grädler, U., Gerber, H.-D., Goodenough-Lashua, D. M., Garcia, G. A., Ficner, R., Reuter, K., Stubbs, M. T., and Klebe, G. (2001) A new target for shigellosis: Rational design and crystallographic studies of inhibitors of tRNA-guanine transglycosylase. *J. Mol. Biol. 306*, 455−467.

(5) Durand, J. M., Okada, N., Tobe, T., Watarai, M., Fukuda, I., Suzuki, T., Nakata, N., Komatsu, K., Yoshikawa, M., and Sasakawa, C. (1994) *vacC*, a virulence-associated chromosomal locus of *Shigella flexneri*, is homologous to *tgt*, a gene encoding tRNA-guanine transglycosylase (Tgt) of *Escherichia coli* K-12. *J. Bacteriol. 176*, 4627−4634.

(6) Durand, J. M. B., Björk, G. R., Kuwae, A., Yoshikawa, M., and Sasakawa, C. (1997) The modified nucleoside 2-methylthio-N^6-isopentenyladenosine in tRNA of *Shigella flexneri* is required for expression of virulence genes. *J. Bacteriol. 179*, 5777−5782.

(7) Brenk, R., Stubbs, M. T., Heine, A., Reuter, K., and Klebe, G. (2003) Flexible adaptations in the structure of the tRNA-modifying enzyme tRNA-guanine transglycosylase and their implications for substrate selectivity, reaction mechanism and structure-based drug design. *ChemBioChem 4*, 1066−1077.

(8) Ritschel, T., Atmanene, C., Reuter, K., Van Dorsselaer, A., Sanglier-Cianferani, S., and Klebe, G. (2009) An integrative approach combining noncovalent mass spectrometry, enzyme kinetics and X-ray crystallography to decipher Tgt protein-protein and protein-RNA interaction. *J. Mol. Biol. 393*, 833−847.

(9) Stengl, B., Meyer, E. A., Heine, A., Brenk, R., Diederich, F., and Klebe, G. (2007) Crystal structures of tRNA-guanine transglycosylase (TGT) in complex with novel and potent inhibitors unravel pronounced induced-fit adaptations and suggest dimer formation upon substrate binding. *J. Mol. Biol. 370*, 492−511.
(10) Xie, W., Liu, X., and Huang, R. H. (2003) Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. *Nat. Struct. Biol. 10*, 781−788.
(11) Ritschel, T., Kohler, P. C., Neudert, G., Heine, A., Diederich, F., and Klebe, G. (2009) How to replace the residual solvation shell of polar active site residues to achieve nanomolar inhibition of tRNA-guanine transglycosylase. *ChemMedChem 4*, 2012−2023.
(12) Fletcher, S., and Hamilton, A. D. (2006) Targeting protein−protein interactions by rational design: mimicry of protein surfaces. *J. R. Soc. Interface 3*, 215−233.
(13) Chea, Y., Brooksa, B. R., and Marshall, G. R. (2006) Development of small molecules designed to modulate protein−protein interactions. *J. Comput.-Aided Mol. Design 20*, 109−130.
(14) Morelli, X., Bourgeas, R., and Roche, P. (2011) Chemical and structural lessons from recent successes in protein−protein interaction inhibition (2P2I). *Curr. Opin. Chem. Biol. 15*, 475−481.
(15) Arkin, M. R., and Wells, J. A. (2004) Small molecule inhibitors of protein−protein interactions: progressing towards the dream. *Nat. Rev. Drug Discovery 3*, 301−317.
(16) Sillerud, L. O., and Larson, R. S. (2005) Design and structure of peptide and peptidemimetic antagonists of protein-protein interaction. *Curr. Prot. Pept. Sci. 6*, 151−169.
(17) Kohler, P. C., Ritschel, T., Schweizer, W. B., Klebe, G., and Diederich, F. (2009) High-affinity inhibitors of tRNA-guanine transglycosylase replacing the function of a structural water cluster. *Chem.—Eur. J. 15*, 10809−10817.
(18) Hörtner, S. R., Ritschel, T., Stengl, B., Kramer, C., Schweizer, W. B., Wagner, B., Kansy, M., Klebe, G., and Diederich, F. (2007) Potent inhibitors of tRNA-guanine transglycosylase, an enzyme linked to the pathogenicity of the *Shigella* bacterium: Charge-assisted hydrogen bonding. *Angew. Chem., Int. Ed. 46*, 8266−8269. *Angew. Chem. 119*, 8414−8417.
(19) Fotsop, D. F., Roussi, F., Leverrier, A., Bretéché, A., and Guéritte, F. (2010) Biomimetic total synthesis of meiogynin A, an Inhibitor of Bcl-xL and Bak Interaction. *J. Org. Chem. 75*, 7412−7415.
(20) Wright, J. A., Gaunt, M. J., and Spencer, J. B. (2006) Novel anti-Markovnikov regioselectivity in the Wacker reaction of styrenes. *Chem.—Eur. J. 12*, 949−955.
(21) Carcenac, Y., Tordeux, M., Wakselman, C., and Diter, P. (2006) Experimental determination of the conformational free energies (*A* values) of fluorinated substituents in cyclohexane by dynamic ^{19}F NMR spectroscopy. Part 2. Extension to fluoromethyl, difluoromethyl, pentafluoroethyl, trifluoromethylthio and trifluoromethoxy groups. *New J. Chem. 30*, 447−457.
(22) DiLabio, G. A., Ingold, K. U., Roydhouse, M. D., and Walton, J. C. (2004) Axial and equatorial cyclohexylacyl and tetrahydropyranyl-2-acyl radicals. An experimental and theoretical study. *Org. Lett. 6*, 4319−4322.
(23) Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D., and Duhr, S. (2010) Protein-binding assays in biological liquids using microscale thermophoresis. *Nat. Commun.*, DOI: 10.1038/ncomms1093.
(24) Ritschel, T., Hoertner, S., Heine, A., Diederich, F., and Klebe, G. (2009) Crystal structure analysis and in silico pK_a calculations suggest strong pK_a shifts of ligands as driving force for high-affinity binding to TGT. *ChemBioChem 10*, 716−727.
(25) Barandun, L. J., Immekus, F., Kohler, P. C., Tonazzi, S., Wagner, B., Wendelspiess, S., Ritschel, T., Heine, A., Kansy, M., Klebe, G., and Diederich, F. (2012) From *lin*-benzoguanines to *lin*-benzohypoxanthines as ligands for *Zymomonas mobilis* tRNA-guanine transglycosylase: Replacement of protein−ligand hydrogen bonding by importing water clusters. *Chem.—Eur. J. 18*, 9246−9257.
(26) Fiser, A., and Šali, A. (2003) ModLoop: automated modeling of loops in protein structures. *Bioinformatics 19*, 2500−2501.
(27) Fiser, A., Do, R. K. G., and Šali, A. (2000) Modeling of loops in protein structures. *Protein Sci. 9*, 1753−1773.
(28) Allen, F. H. (2002) The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B58*, 380−388.
(29) McMillan, K., Adler, M., Auld, D. S., Baldwin, J. J., Blasko, E., Browne, L. J., Chelsky, D., Davey, D., Dolle, R. E., Eagen, K. A., Erickson, S., Feldman, R. I., Glaser, C. B., Mallari, C., Morrissey, M. M., Ohlmeyer, M. H. J., Pan, G., Parkinson, J. F., Phillips, G. B., Polokoff, M. A., Sigal, N. H., Vergona, R., Whitlow, M., Young, T. A., and Devlin, J. J. (2000) Allosteric inhibitors of inducible nitric oxide synthase dimerization discovered via combinatorial chemistry. *Proc. Nat. Acad. Sci. U.S.A. 97*, 1506−1511.
(30) Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) The Leucine Zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science 240*, 1759−1764.
(31) Wada, A. (1976) The alpha-helix as an electric macro-dipole. *Adv Biophys 9*, 1−63.
(32) Hol, W. G. J. (1985) The role of the α-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol. 45*, 149−195.
(33) Forsyth, W. R., Antosiewicz, J. M., and Robertson, A. D. (2002) Empirical relationships between protein structure and carboxyl pK_a values in proteins. *Proteins 48*, 388−403.
(34) Porter, M. A., Hall, J. R., Locke, J. C., Jensen, J. H., and Molina, P. A. (2006) Hydrogen bonding is the prime determinant of carboxyl pK_a values at the N-termini of α-helices. *Proteins 63*, 621−635.
(35) Chacko, S., Silverton, E., Kam-Morgan, L., Smith-Gill, S., Cohen, G., and Davies, D. (1995) Structure of an antibody−lysozyme complex unexpected Effect of a conservative mutation. *J. Mol. Biol. 245*, 261−274.
(36) Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., and Pavletich, N. P. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science 274*, 948−953.
(37) Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science 267*, 383−386.
(38) Bogan, A. A., and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces. *J. Mol. Biol. 280*, 1−9.
(39) Moreira, I. S., Fernandes, P. A., and Ramos, M. J. (2007) Hot spots−A review of the protein−protein interface determinant amino-acid residues. *Proteins 68*, 803−812.
(40) Reuter, K., and Ficner, R. (1995) Sequence analysis and overexpression of the *Zymomonas mobilis tgt* gene encoding tRNA-guanine transglycosylase: Purification and biochemical characterization of the enzyme. *J. Bacteriol. 177*, 5284−5288.
(41) Romier, C., Ficner, R., Reuter, K., and Suck, D. (1996) Purification, crystallization, and preliminary X-ray diffraction studies of tRNA-guanine transglycosylase from *Zymomonas mobilis*. *Proteins: Structure, Function Genet. 24*, 516−519.
(42) 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., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905−921.
(43) Sheldrick, G. M., and Schneider, T. R. (1997) SHELXL: high-resolution refinement. *Methods Enzymol. 277*, 319−343.
(44) Emsley, P., and Cowtan, K. (2004) *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.
(45) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr. 26* (2), 283−291.
(46) Meyer, E. A., Donati, N., Guillot, M., Schweizer, W. B., Diederich, F., Stengl, B., Brenk, R., Reuter, K., and Klebe, G. (2006) Synthesis, biological evaluation, and crystallographic studies of extended guanine-based (*lin*-benzoguanine) inhibitors for tRNA-guanine transglycosylase (TGT). *Helv. Chim. Acta 89*, 573−597.

(47) Neudert, G., and Klebe, G. (2011) fconv: format conversion, manipulation and feature computation of molecular data. *Bioinformatics 27*, 1021−1022.
(48) Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A. W., Kolossvai, I., Wong, K. F., Paesani, F., Vanicek, J., Wolf, R. M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2012) AMBER 12, University of California, San Francisco
(49) Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. F., McCabe, P., Pearson, J., and Taylor, R. (2002) New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallogr. B58*, 389−397.
(50) Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J., and Wood, P. A. (2008) *Mercury CSD* 2.0−new features for the visualization and investigation of crystal structures. *J. Appl. Crystallogr. 41*, 466−470.
(51) Eisenberg, D., Schwarz, E., Komarony, M., and Wall, R. (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol. 179*, 125−142.