Articles

Virtual Screening for Submicromolar Leads of tRNA-guanine Transglycosylase Based on a New Unexpected Binding Mode Detected by Crystal Structure Analysis

Ruth Brenk,[†] Lars Naerum,[‡,§] Ulrich Grädler,[†,||] Hans-Dieter Gerber,[†] George A. Garcia,[⊥] Klaus Reuter,[†] Milton T. Stubbs,[†] and Gerhard Klebe*[,†]

Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35032 Marburg, Germany, NovoNordisk A/S, Novo Allé, 2880 Bagsvaerd, Denmark, and Interdepartmental Program in Medicinal Chemistry, College of Pharmacy, University of Michigan, Ann Arbor, Michigan 48109-1065

Received July 12, 2002

Eubacterial tRNA-guanine transglycosylase (TGT) is involved in the hypermodification of cognate tRNAs, leading to the exchange of G34 by preQ1 at the wobble position in the anticodon loop. Mutation of the tgt gene in *Shigella flexneri* results in a significant loss of pathogenicity of the bacterium due to inefficient translation of a virulence protein mRNA. Herein, we describe the discovery of a ligand with an unexpected binding mode. On the basis of this binding mode, three slightly deviating pharmacophore hypotheses have been derived. Virtual screening based on this composite pharmacophore model retrieved a set of potential TGT inhibitors belonging to several compound classes. All nine tested inhibitors being representatives of these classes showed activity in the micromolar range, two of them even in the submicromolar range.

Introduction

Shigellae are the causative agents of dysentery (shigellosis) and effect more than one million deaths each year.[1] Shigellosis is usually treated by antibiotics,[2,3] but more and more multidrug-resistant strains are reported. Accordingly, there is urgent need for the development of new antibiotics against shigellosis.[4]

Characterization of chromosomal mutants of *S. flexneri* has resulted in the identification of a gene, vacC, that contributes significantly to pathogenicity.[5] The nucleotide sequence of the vacC gene is highly homologous (>98%) to the tgt gene of *Escherichia coli*, coding for tRNA-guanine transglycosylase (TGT, EC 2.4.2.29).[6,7] TGT is involved in the biosynthesis of the highly modified nucleoside queuine (Figure 1) inserted in the anticodon loop of certain tRNAs. In eubacteria, the exchange of guanine in the unmodified tRNA by the queuine precursor preQ$_1$ (7-methylamino-7-deazaguanine, Figure 1) is catalyzed. Recent studies have shown that the reaction follows an associative mechanism.[8] Once incorporated into tRNA, preQ$_1$ undergoes further chemical modification through at least two subsequent enzymatic steps to yield the final queuine modification.[9,10] Although the exact biological function of modified bases in tRNAs is still unclear, it is known that particularly those present in the anticodon region can alter the efficiency and fidelity of tRNAs during translation.[11−13] Blocking the biological function of TGT with specific inhibitors is expected to result in tRNAs lacking queuine in the anticodon. In *S. flexneri*, this leads to significantly reduced expression of the virF gene, which in consequence results in a dramatic reduction of virulence.[6,14]

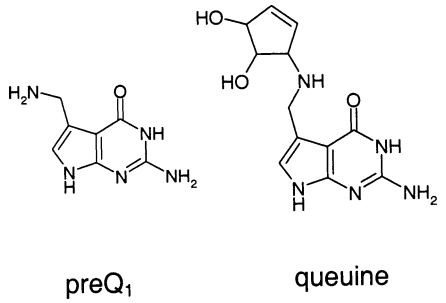


Figure 1. Chemical structures of preQ$_1$ and queuine.

The crystal structure of *Zymomonas mobilis* TGT has been solved at 1.85 Å.[15] The residues participating in substrate binding and catalysis of *Z. mobilis* TGT are identical with those of the *S. flexneri* enzyme apart from the replacement of Tyr106 by a phenylalanine residue.[16] Therefore, the crystal structure of *Z. mobilis* TGT provides an ideal platform for the rational design of potent inhibitors against shigellosis.

Structure-based drug design is an iterative cycle.[17] The process starts with a detailed analysis of the binding site of the target protein, which is preferably complexed with a ligand. This complex unravels the binding mode and conformation of a ligand under

* To whom correspondence should be addressed. Phone: 0049 6421 2821313. Fax: 0049 6421 2828994. E-mail: klebe@mailer.uni-marburg.de.
† Philipps-Universität Marburg.
‡ NovoNordisk A/S.
§ Present address: Combio A/S, Gamle Carlsberg Vej 10, 2500 Valby, Denmark.
|| Present address: ALTANA-Pharma, Byk-Gulden-Strasse 2, 78467 Konstanz, Germany.
⊥ University of Michigan.

10.1021/jm0209937 CCC: $25.00
Published on Web 02/22/2003

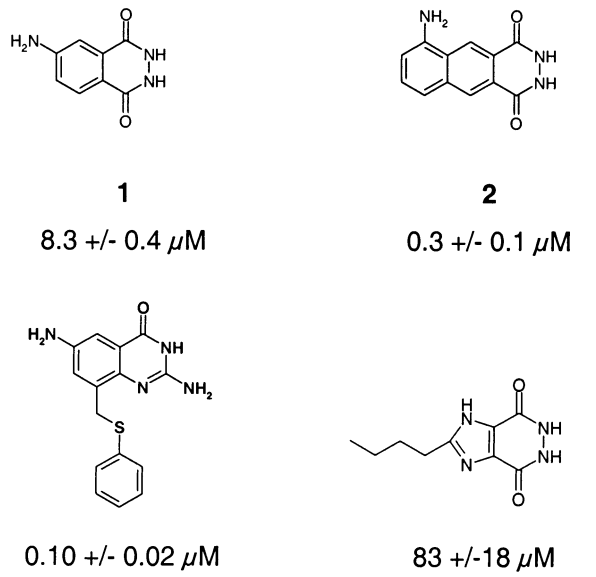


Figure 2. Compounds used in the test set with K_i values.

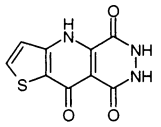
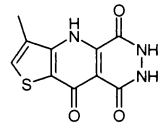
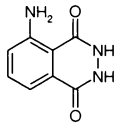
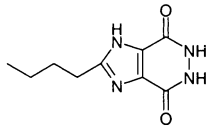
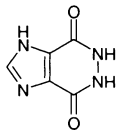
investigation and indicates the essential aspects determining its binding affinity. It is then used to generate new ideas about ways of improving an existing ligand or of developing new alternative bonding skeletons. Computational methods supplemented by molecular graphics are applied to assist this step of hypothesis generation. The features of the protein binding pocket can be translated into queries used for virtual computer screening of large compound libraries or to design novel ligands de novo. These initial proposals must be confirmed experimentally. Subsequently, they are optimized toward higher affinity and selectivity.

In a recent design study, 4-aminophthalhydrazide (**1**) was found to be an inhibitor of TGT.[18] Its binding mode was determined crystallographically. Starting from this structure, here we present the next design cycle using X-ray crystallography, "hot spot" analysis, and virtual screening techniques.

Results and Discussion

In a previous de novo design study, we discovered pyridazinedione derivatives, e.g, **1** and **2** (Figure 2), as possible inhibitors of TGT in the low micromolar to submicromolar range.[18] Attempts to improve binding affinity by placing additional substituents of polar nature into the neighborhood of the two facing Asp102 and Asp280 residues did not result in better binding. We explained this finding by repulsive interactions of an uncharged triazole substituent bound to the parent benzpyridazinedione skeleton. This substituent has been found to be disordered in the crystal structure of the protein−ligand complex. Furthermore, the assumed repulsive interactions lead to a rupture of the hydrogen bond formed between the exocyclic NH_2 group of the parent structure and the backbone carbonyl group of Leu231. On the basis of these findings, we concluded that this hydrogen bond is an important feature for strong ligand binding. This assumption is further supported by a recent study based on substituted 2-amino-3*H*-quinazoline-4-ones to accommodate the guanine binding site. Synthesis of derivatives possessing or lacking the described exocyclic NH_2 group results in a more than 10-fold affinity difference.[19] To further study the structural properties of pyridazinediones, we screened the proprietary NovoNordisk Compound Database (NNCD) for corresponding analogues using traditional 2D substructure searching and a combined search strategy.[20] From these computer searches, the compounds shown in Table 1 were selected and tested. First attempts to dock these compounds into the binding site of TGT suggested a structural inadequacy apart from **5**. None of these inhibitors appeared capable of hydrogen-bonding to the important backbone carbonyl of Leu231. The more we were surprised that all compounds showed inhibitory activity in the micromolar range. **3** and **4** even inhibited the enzyme with low micromolar affinity (Table 1).

Table 1. List of Five Pyridazinediones Retrieved in the NNCD and Tested for TGT Inhibition

No	Compound	K_i [μM]
3		5.0 +/- 1.2
4		4.7+/- 1.0
5		36 +/- 11
6		83 +/-18
7		200 +/- 70

For crystals of the *Z. mobilis* TGT, a soaking system with a well solvent-accessible binding site is established.[15,21] Therefore, we tried to soak the different compounds into crystals of TGT to collect evidence for this rather unexpected good binding affinity. Successful soaking could be achieved in the case of the well-soluble **6**. Unexpectedly, the crystal analysis of this complex shows a flip of the peptide bond at Ala232 to Leu231 (Figures 3 and 4). This flip rotates the carbonyl group of Leu231 out from the binding pocket. The adjacent NH of Ala232 is now facing the ligand binding site. The side chain of Leu231 virtually remains in the same position. Furthermore, an interstitial water molecule (W1) is incorporated, thus mediating a firm contact to the backbone NH of Ala232 and the imidazole nitrogen of **6** (Figure 4). In addition, this water molecule is hydrogen-bonded to the sulfur of Cys158. A water molecule at this site has not yet been observed in any other ligand complex with *Z. mobilis* TGT or the apo structure. It exhibits a B factor of only 34.8 Å^2, which

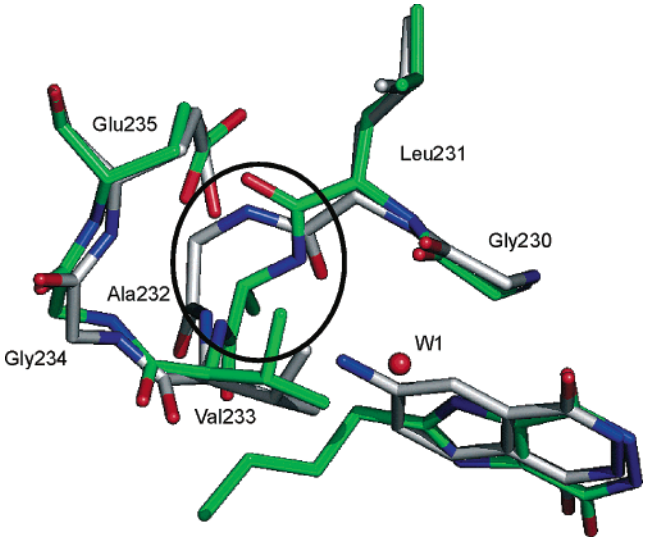


Figure 3. Superpositon of a part of the protein structure of TGT in complex with **1** (green) and **6** (gray). The peptide bond at Leu231 to Ala232 (circle) in the structure of TGT complexed with **6** is flipped compared to that with **1**. The flip rotates the carbonyl group of Leu231 off from the binding pocket. The adjacent NH of Ala232 is now facing the ligand-binding site. The side chain of Leu231 remains at virtually the same position. (Picture has been produced using PyMOL.)

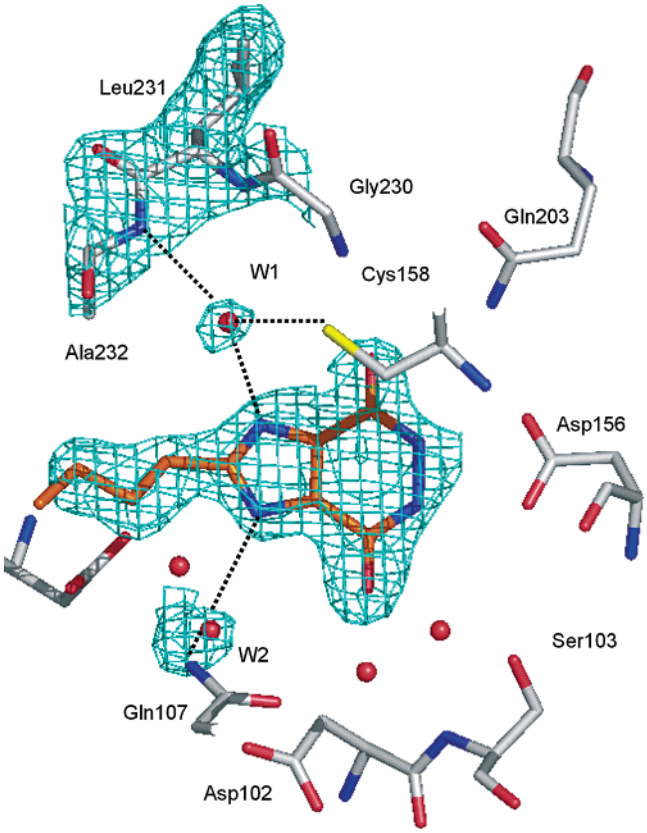


Figure 4. Representation of the 2.1 Å resolution $|F_o| - |F_c|$ simulated annealing omit map contoured at 3.0 σ of the *Z. mobilis* TGT in complex with **6**. (Picture has been produced using PyMOL.)

is in the same range as the B factors for the adjacent ligand atoms (Table 4). This may be taken as a crude indicator for full occupancy and rather strong binding. In addition, fewer dramatic changes are also observed. The side chain of Gln107 moves slightly toward the

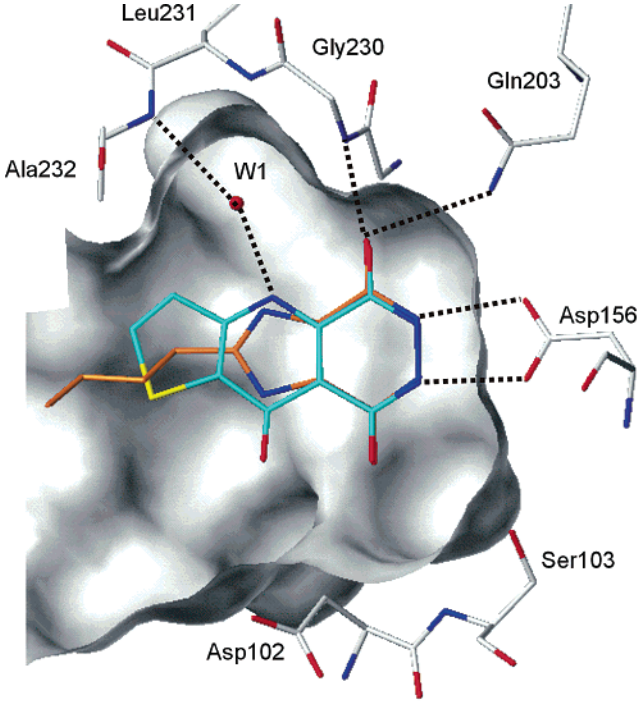


Figure 5. Modeled binding mode of **3** (cyan). The crystallographically determined binding mode of **6** (orange) is shown for comparison. Both ligands can form hydrogen bonds to Asp156, Gln203, and Gly230 as well as to the interstitial water molecule W1, which mediates the contact to Leu231. (Picture has been produced using SYBYL.)

ligand, now giving rise to a hydrogen bond to a second interstitial water molecule (W2), which mediates a further contact to the ligand. The quite unexpected coincidence of two ligand-induced structural rearrangements of the binding site would have been impossible to predict on the basis of the previous structural knowledge of TGT binding. It clearly demonstrates the inherent limitations of predicting binding modes even if several ligand−protein complexes are already characterized. It also underlines the importance of crystal structure analysis as a prerequisite for successful iterative structure-based design.

On the basis of the crystal coordinates of the complex with **6**, the binding modes of **3**, **4**, and **7** have been modeled accordingly. The suggested binding mode of **3** as a representative example is shown in Figure 5.

The discovery of this new binding geometry resulting from an obvious plasticity of the protein in this region inspired us to develop a structure-based pharmacophore hypothesis for virtual screening comprising the different binding modes, as evidenced by crystal structure analysis (Figure 6). First, we considered the hydrogen-bonding pattern as discovered in the complexes with **1** and **2**, in particular since the protein conformer with the exposed backbone carbonyl of Leu231 that accepted a hydrogen bond from the ligand is supposed to contribute strongly to binding.[18] In the newly discovered protein conformer with **6**, an interstitial water operates as mediator. Because this water molecule W1 is obviously tightly bound in the complex, it was included in the search query, thus allowing for a higher variability of potential ligands. Owing to their ambivalent hydrogen-bonding properties, water molecules can operate in principle either as donors or as acceptors. Therefore, to

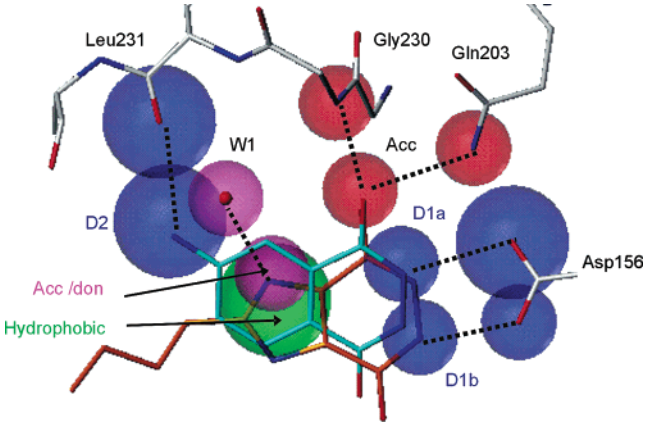


Figure 6. Structure-based pharmacophore hypothesis. The interaction to the carbonyl group of Leu231 is considered as an alternative to interaction with the water molecule W1. Requested donor features and the corresponding acceptor-site features of the protein are in blue, acceptor features and the corresponding donor-site features of the protein are in red, the acceptor/donor feature and the corresponding donor/acceptor site feature of the water molecule are in magenta, and the hydrophobic feature is in green. (Picture has been produced using SYBYL.)

complement the hydrogen-bonding facilities of this water, potential ligands can exhibit a donor or acceptor group at this site (Figure 6). Most of our presently discovered tight binding inhibitors expose one acceptor facility toward Gly230 and Gln203 and two donor properties toward Asp156 in a bifurcated hydrogen-bonding pattern.[18,22] Therefore, these features have also been incorporated in our pharmacophore hypothesis.

To manifest these qualitative considerations of binding features on a more quantitative basis, we performed a "hot-spot" analysis using several probe functional groups in SuperStar[23] and DrugScore.[24] We have already successfully applied this strategy in the search for new inhibitors of human carbonic anhydrase II.[25,26] SuperStar is a method for identifying interaction sites in protein binding pockets based entirely on experimental information about nonbonded interactions derived from the statistical analysis of packing patterns in crystal structures of small molecules. DrugScore is based on data derived from crystal structures of ligand−protein complexes and translates the frequency of particular protein−ligand contacts into atom/atom-pair preferences. In SuperStar, which explicitly considers the orientation of hydrogen atoms, the water molecule W1 was handled as a hydrogen-bond acceptor or donor. In DrugScore, which ignores the actual protonation states of functional groups of both the protein and the ligand, the water molecule was assigned to an O.3-type oxygen. The results of these analyses are shown in Figure 8. Figure 8a shows the "hot spots" calculated with DrugScore for an N.3 probe (ligand donor group) for the binding pocket complexed with **1** (Figure 7a). Figure 8b depicts similar information for the binding pocket with the flipped backbone conformation present in the TGT·**6** complex (Figure 7b). Direct comparison of the two maps clearly demonstrates the shift of the donor "hot spot" from the region where the amino group of **1** is found toward a site where the imidazole nitrogen of **6** is located after the water molecule has been incorporated. Figure 8c shows the "hot spots" for a ligand

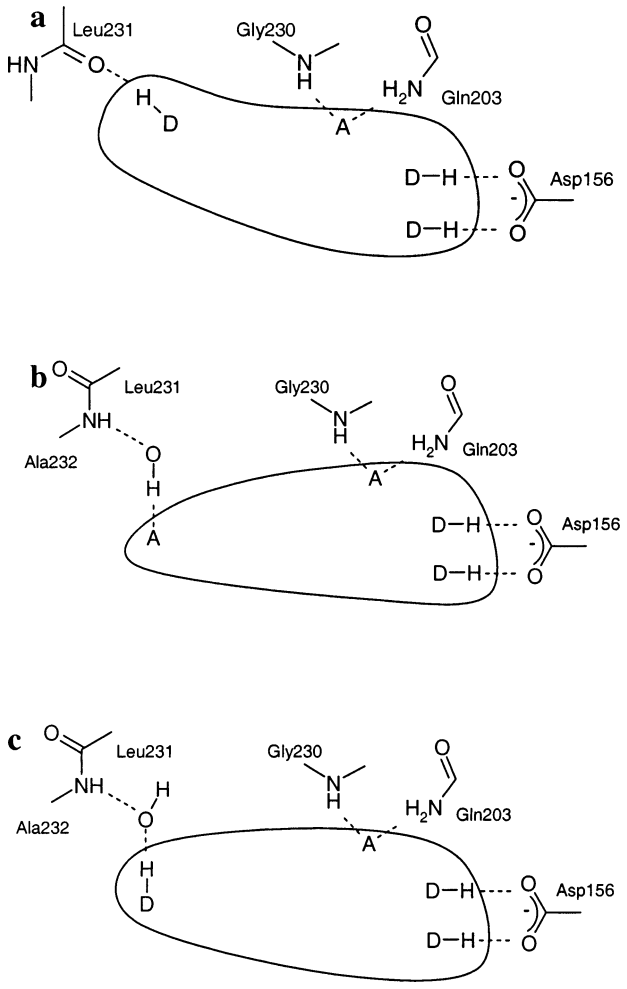


Figure 7. Schematic view of the different backbone conformations in the binding site of TGT, and the hydrogen-bond pattern considered in the pharmacophore hypothesis: (a) backbone conformation of the binding pocket complexed with **1**; (b) backbone conformation of the binding pocket complexed with **6**, with the interstitial water molecule exposing its donor property toward the ligand; (c) backbone conformation of the binding pocket complex with **6**, with the water molecule now exposing its acceptor property toward the ligand. "A" indicates an acceptor group, and "D" indicates a donor group.

acceptor functional group using the pocket with the flipped backbone conformation (Figure 7c). The map has been calculated using SuperStar. In this analysis, the water W1 has been considered as a hydrogen-bond donor with respect to the ligand probe. Both situations, parts b and c of Figure 7, take the ambivalent donor/acceptor property of the interstitial water molecule into account. In consequence, a complementary acceptor or donor functionality is required in a putative ligand at this site. Figure 8d presents the "hot spots" for a hydrophobic C.ar probe atom calculated with DrugScore. Again, the "hot spot" coincides well with the placement of the aromatic ring in **6** found in the crystal structure.

In the following, the above-described "hot spot" areas for the different probes, "hydrogen-bond donor", "hydrogen-bond acceptor", and "hydrophobic aromatic property", have been translated into a search query appropriate for Unity[27] following the protocol described by Grüneberg et al.[26] The various search-tolerance spheres have been placed by picking the appropriate atoms of **1**, compared with the atoms of **6**, and their diameters

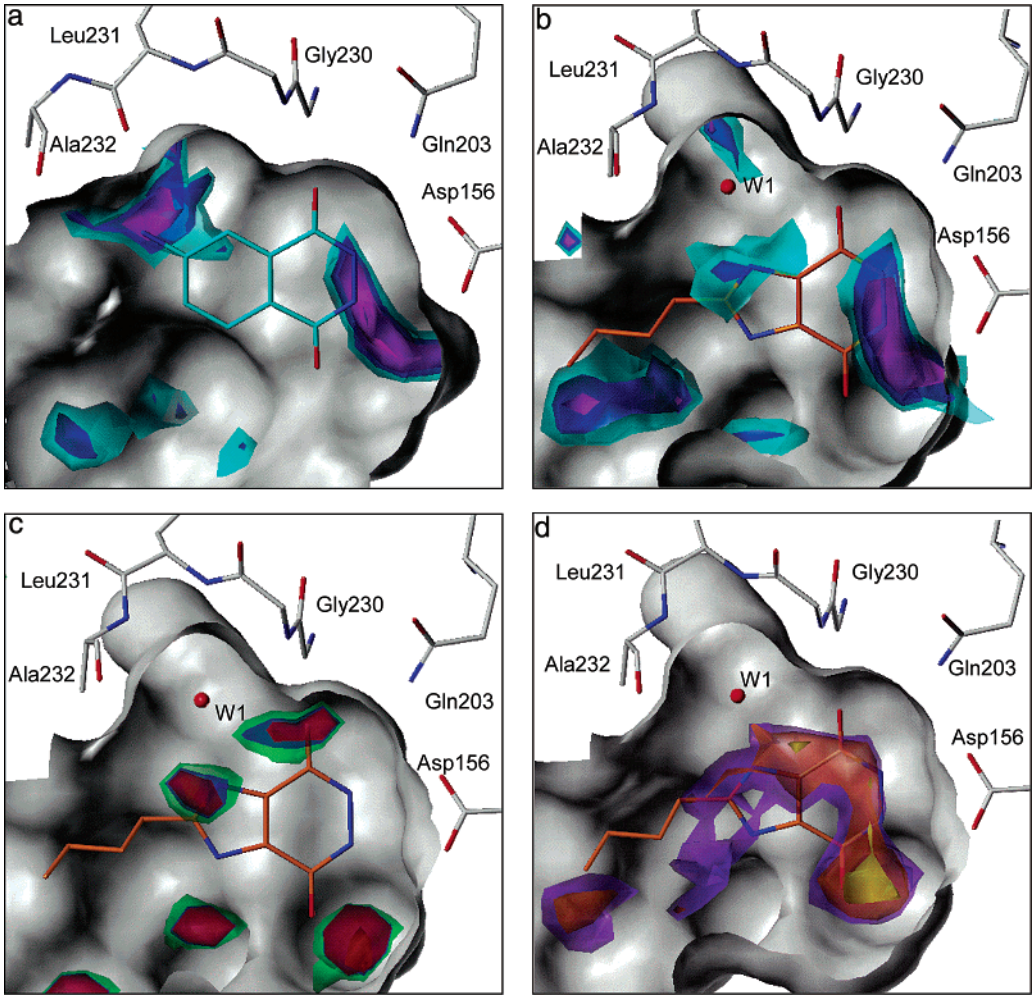


Figure 8. Mapping of putative binding "hot spots" in the active-site of TGT. For orientation, the binding geometry of the inhibitor **1** or **6**, studied crystallographically, is also shown: (a, b) highlighting the properties of a putative hydrogen-bond donor group in a ligand, calculated using DrugScore with a N.3 probe, contoured at 80% (cyan), 84% (blue), and 88% (magenta) levels with respect to the global minimum, for (a) the binding pocket of TGT in complex with **1** and (b) the binding pocket of TGT in complex with **6**; (c) highlighting the properties of a putative hydrogen-bond acceptor group in a ligand, calculated using SuperStar with a carbonyl oxygen probe, contoured at relative propensity levels of 4 (green), 8 (blue), and 10 (red), in which a propensity level of 1 corresponds to random occurrence; (d) highlighting the properties of a putative hydrophobic probe in a ligand, with DrugScore using C.ar, contoured at 89% (magenta), 91% (orange), and 95% (yellow). (Pictures have been made using SYBYL.)

have been adjusted to the spread of the indicated "hot spots" (Figure 6).

In our previous study, we limited the virtual screening to a subset of the ACD. In the present study, we extended our retrieval to eight different databases containing in total over 800 000 candidate molecules.

The screening has been performed in a stepwise fashion using Selector,[28] Unity, and FlexX[29] and included several hierarchical filters of increasing complexity with respect to their computational requirements.[26] In a fast initial step, only compounds with up to seven rotatable bonds and a molecular weight of less than 450 Da have been considered. The reasons for these criteria are (a) to avoid highly flexible ligands and (b) to retrieve hits small enough to allow for further optimization, thus focusing on "leadlike" hits.[30,31] Almost 50% of the compounds were eliminated by this rather unspecific and target-independent filter. In a second step, limiting the database to compounds comprising at least (a) two hydrogen-bond donors, (b) one hydrogen-bond acceptor, and (c) one hydrophobic moiety further reduced the candidate molecules to about 20% of the initial set. The protein-based pharmacophore hypothesis has been used in the third step to constrain the directionality and mutual spatial arrangement of the required hydrogen bonds and the hydrophobic moiety. In a fourth step, in addition to the criteria of the previous one, the approximate shape of the binding site has been considered in terms of excluded volume constraints. The final hit list contained 872 compounds. Accordingly, the hierarchical filtering procedure reduced the databases to only 0.11% of their original size (Table 2). Sixteen duplicates have been discarded because of multiple occurrences in the eight databases of deviating origin that obviously contain slightly overlapping information.

Out of these 856 compounds, 130 passed all the filters because they can form a direct hydrogen bond with the carbonyl oxygen of Leu231, as found in the crystal structure of TGT·**1**. An amount of 726 hits were retrieved because of their ability to form a hydrogen bond with the interstitial water molecule W1, underlying the importance of including this water in the query.

The hit list contained compounds of several different chemical classes, e.g, 120 carboxylic acid hydrazides

Table 2. Statistical Overview of the Results from Sequential Application of a Series of Hierarchical Filters on the Eight Considered Databases

	ACD[a]		Ambinter[a]		Ambinter Nat[a]		AEGC[a]	
filter step	no. of compds	%	no. of compds	%	no. of compds	%	no. of compds	%
	215212	100.00	114855	100.00	960	100.00	182485	100.00
1. rotatable bonds/MW	135502	62.96	59580	51.87	297	30.94	91677	50.24
2. requested no. of hydrophobic donor and acceptor properties	41626	19.34	26164	22.78	120	12.50	39068	21.41
3. pharmacophore hypothesis	1223	0.57	496	0.43	0	0.00	591	0.32
3.1. discarded because of timed out[b]	202	0.09	316	0.28	1	0.10	392	0.21
4. excluded volumes	478	0.22	118	0.10	0	0.00	75	0.04
4.1. discarded because of timed out[b]	128	0.06	55	0.05	0	0.00	115	0.06

	AEPC[a]		ChemStar[a]		IBS[a]		LeadQuest[a]		Σ	
filter step	no. of compds	%	no. of compds	%	no. of compds	%	no. of compds	%	no. of compds	%
	44549	100.00	57927	100.00	158942	100.00	52022	100.00	826952	100.00
1.	9417	21.14	28712	49.57	76321	48.02	18231	35.04	419737	50.76
2.	5890	11.65	12466	21.52	35661	22.44	8093	15.56	168387	20.36
3.	67	0.15	169	0.29	699	0.44	64	0.12	3309	0.40
3.1[b]	18	0.04	0	0.00	328	0.21	19	0.04	1276	0.15
4	31	0.07	39	0.07	128	0.08	3	0.01	872	0.11
4.1[b]	6	0.01	45	0.08	152	0.10	7	0.01	508	0.06

[a] Release dates: ACD, 2000; AMBINTER, AECG, AEPC, ChemStar, and IBS, 2001; Leadquest, 2000. [b] The main reasons that the compounds could not be placed within the time limit were that they had many rotatable bonds and/or many functional groups, e.g., sugars.

directly attached to an aromatic ring moiety, 33 carboxylic acid hydrazides attached to an aliphatic carbon, 75 guanine derivatives, 61 hydroxyaminopiperidines, 43 pyrazolylamines or pyrazololes, and 20 pterins. Some selected examples are given in Figure 9.

Subsequently, the obtained hit list was inspected visually. Prospective hits were then docked into one of the two alternative binding site conformations of the TGT as found in the complex with **1** or **6**, using FlexX. Ligands that still matched the pharmacophore hypothesis after docking were minimized with the MAB force field[32] while keeping the binding pocket rigid. A final selection for purchase and enzyme testing included the following criteria: (a) the overall matching of the requested hydrogen-bonding network, (b) complementarity between ligand and protein surfaces in terms of spatial occupancy and matched contacts in hydrophobic/hydrophilic surface patches, and (c) the absence of unfavorable van der Waals interactions after minimization.

For the retrieved hits **8**−**16**, enzyme inhibition has been characterized. Interestingly enough, all compounds displayed activities in the micromolar range or higher (Table 3).

For the pterin-derivatives (**8**−**10**), a binding mode very similar to that of **6**, considering the mediating water molecule, can be assumed. 6-Methylpterin (**8**) is the most potent compound discovered, with a K_i value of 0.2 μM. Pterin (**9**) falls into the same range ($K_i = 0.6$ μM). The latter hit has already been tested in eucaryotic TGT, with a K_i of 90 nM being reported.[33] The third pterin derivative (**10**, 7-phenylpterin) exhibits a K_i value of 3.8 μM. The reduced affinity of **10** compared to **8** and **9** is probably due to the fact that the molecule has to adopt an unfavorable conformation in the binding pocket. The optimal torsion angle between the phenyl ring and the pterin moiety is about 40°. However, if this conformation is adopted in the binding pocket, the planar stacking between the pterin moiety and Tyr106 would be disturbed because of close contacts with the adjacent 7-phenyl ring of the ligand. Accordingly, the ligand is most likely forced to adopt a less favorable virtually planar conformation.

Pyrazolylamines (**11** and **12**) are a newly detected class of TGT inhibitors. Compared to the natural substrate preQ$_1$, which exhibits a K_m of 0.2 μM,[34] they are rather weak binders. This is probably a consequence of the fact that, compared to preQ$_1$, the acceptor group is not an exocyclic carbonyl function but an aromatic nitrogen. Depending on the tautomeric form (Figure 10), this nitrogen can either be an acceptor (**12**) or carry a hydrogen and thus act as a donor (**12a**), which would be unfavorable for binding. This "unfavoured" tautomer (**12a**), however, appears to be the significantly more stable tautomeric form, as indicated by quantum chemical calculations carried out for compound **12**. Accordingly, the compound may predominantly exist in a form that precludes tight binding and would actually have been discarded by the pharmacorphore filter once considered as this tautomeric form in the search. In addition, the pyrazolylamine tautomer **12** providing the correct hydrogen-bonding pattern does not ideally fulfill the requirements for tight interaction. The distance between the acceptor group and the two donor groups, which are exposed to Gly230-NH, Gln203-NH$_2$, or Asp156, appears much shorter, thus supposedly not ideally matching the required hydrogen-bonding pattern defined by the protein. In consequence, extended hydrogen bonds to Glu203 and Gly230 might result. These two aspects taken together may explain the reduced binding affinity.

Open-chain carboxylic acid hydrazides constitute a second new compound class. The K_i value of **13** falls into the same range as those of **5** and **6**. Compared to **5** and **6**, the carboxylic acid hydrazides are easily accessible via substitution of the free carboxylic acids or esters.[35]

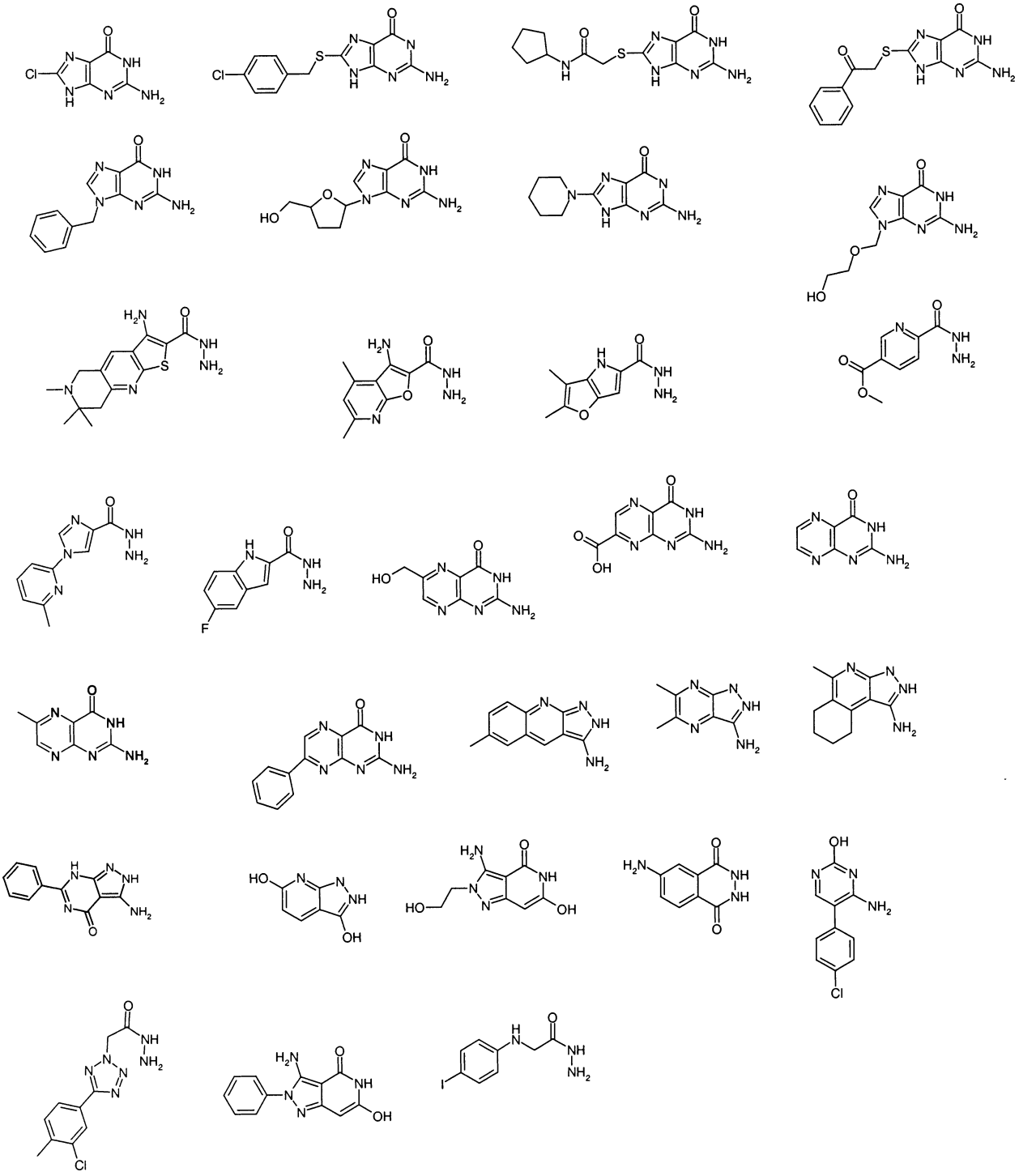


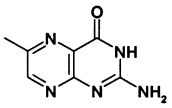
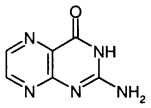
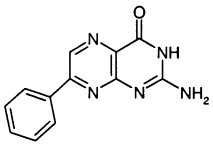
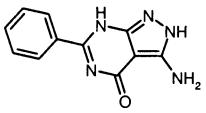
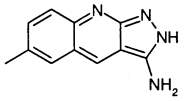
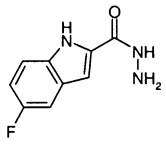
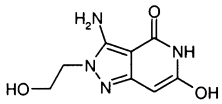
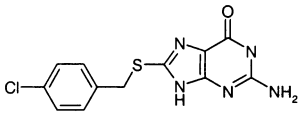
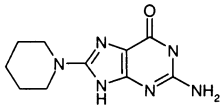
Figure 9. Selected hits discovered by virtual screening.

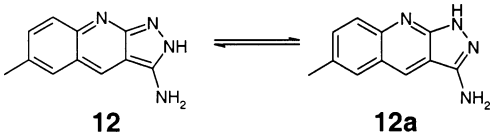
Numerous esters and acids are commercially available. Accordingly, **13** might serve as a prospective lead structure for further optimization. In addition, open-chain hydrazides are already known as drugs; for example, isoniazid is a well-known drug for the treatment of tuberculosis. It is even administered to children without observation of serious side effects over a period of up to half a year.[36]

Hit **14** is the first inhibitor discovered that most likely exposes a hydroxyl group to Asp156. The second OH group in the side chain appears suitable as a linker for the attachment of molecular portions that could possibly interact with additional amino acid residues in the active site distal from the substrate recognition site.

Compounds **15** and **16** are substituted guanines. Guanine itself, which is accepted by the TGT as a substrate, exhibits a K_m of 0.7 μM.[37] The K_i of **16** is about 1.5 orders of magnitude higher, probably because of steric crowding next to the directly connected piperidine moiety, while the K_i for **15** (2.7 μM) is in the range of the K_m for guanine. Up to now, we could report on two studies focused on structure-based design of TGT inhibitors.[18,22] In both studies, attempts were made to improve affinity by adding appropriate side chains that

Table 3. List of the Nine Compounds Discovered by Virtual Screening That Were Subsequently Selected for TGT Inhibition Testing

No	Compound	Database	Label	K_i [μM]
8		ACD	MFCD00012137	0.25 +/- 0.05
9		ACD / IBS	MFCD00010557 / STOCK1N-11489	0.6+/- 0.2
10		ACD	MFCD02091220	3.8 +/- 0.1
11		ACD	MFCD00110262	249 +/-47
12		AEPC	ASN2538273	156 +/- 36
13		ACD	MFCD01566929	72 +/- 5
14		AEGC	BAS0316634	8.1 +/- 1.0
15		AEPC	ASN3578296	2.7 +/- 0.3
16		AEGC	BAS0449619	37 +/- 7



12 **12a**

Figure 10. Tautomeric forms of pyrazolamines.

possibly interact with the polar functional groups of amino acid residues Asp280, Asp102, and Asn70 or fill the adjacent small hydrophobic cleft formed by Val45, Leu68, and Val282. For these studies, molecular skeletons derived from either the benzpyridazinedione- or 2-amino-3*H*-quinazoline-4-one type were selected. In both cases, side chains have been added to the core structure in the 6-position instead of in the 8-position. Compound **15** is the first lead discovered with a side chain attached in a topologically different orientation (Figure 11). It is thus expected that this molecule experiences new sites of interactions in the binding pocket of TGT. A promising affinity in the low micromolar range has been detected for this compound, thus indicating that these additional interactions could favorably contribute to ligand binding.

Conclusions

We have discovered a new, unexpected binding mode of a novel TGT inhibitor. Owing to the flip of a binding site exposed peptide bond, the originally presented backbone carbonyl group becomes buried while the donor functionality of the adjacent backbone NH is now presented for ligand binding. In addition, a water molecule is incorporated, thus mediating a favorable ligand−protein interaction. This hardly predictable behavior underlies the importance of crystal structure analysis as a prerequisite for successful iterative structure-based drug design. The new binding mode has been

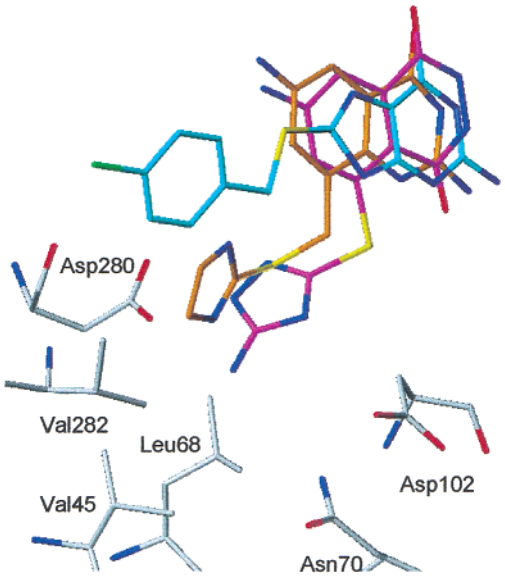


Figure 11. Assumed binding mode of **15** (cyan) compared to the crystallographically determined binding modes of two already characterized inhibitors (orange[22] and magenta[18]). (Picture has been produced using SYBYL.)

considered along with the ambivalent donor/acceptor properties of the interstitial water molecule to perform hydrogen-bonding to a putative ligand in order to define three slightly deviating pharmacophore hypotheses. Virtual screening based on this composite pharmacophore model retrieved a set of potential TGT inhibitors emerging from several compound classes. Out of these, nine inhibitors have been selected for experimental testing. All retrieved compounds show activity in the micromolar range, and two even exhibit submicromolar inhibition. Some of the newly discovered ligand skeletons are currently being investigated as a starting point for further lead optimization.

Experimental Section

Molecular modeling was performed using SYBYL 6.7[38] and MAB/Moloc 01/05/08[39] running on a Silicon Graphics O2 (R10000) workstation.

Structure Determination. The TGT was expressed, purified, and crystallized as described elsewhere.[37,40] Crystals have been soaked overnight with the inhibitor as described by Grädler et al.[18] X-ray data (Table 4) were collected at −173 °C as described.[18] Diffraction data were processed using the programs DENZO and SCALEPACK.[41] The structure was refined through several cycles of least-squares refinement along with an energy minimization using CNS.[42] Manual adjustments to the electron density have been performed using O.[43]

"Hot Spot" Analysis of the Binding Pocket. "Hot Spots" have been calculated using SuperStar[23] and DrugScore[24] as alternative methods. Both programs have been applied with default settings. Calculations have been performed using either the unoccupied binding pocket of TGT in its conformation found with the bound ligand **1** (1ENU) or in the conformation with **6**. For SuperStar, which in contrast to DrugScore explicitly considers hydrogen atoms, the hydrogens of the water molecule W1 have been orientated in the binding pocket to expose their hydrogen-bond acceptor or hydrogen-bond donor property toward the ligand. For DrugScore, the water molecule W1 has been assigned to an O.3-type oxygen. In SuperStar, a carbonyl oxygen probe was selected to analyze hydrogen-bond acceptor interactions, the uncharged amino nitrogen probe was selected to analyze hydrogen-bond donor interactions, and the aromatic CH carbon probe was selected to analyze hydrophobic interactions. In DrugScore, the O.2, N.3, and C.ar probe atoms have been selected accordingly.

Table 4. Crystallographic Data and Refinement Statistics of the TGT Complex with **6**

space group	*C*2
cell constants	
a (Å)	91.06
b (Å)	64.38
c (Å)	70.77
β (deg)	96.5
resolution (Å)	30−2.1
total no. of reflections	87,053
no. of unique reflections	23,882
completeness of all data (%) (outer shell (%))	99.8 (99.4)
R_{symm} for all data [a] (%) (outer shell (%))	9.5 (30.2)
R_{free} [b] (%)	23.3
R factor[b] (%)	18.8
no. of protein atoms (non-hydrogen atoms)	2902
no. of water molecules	282
rms deviation of angle (deg)	1.209
rms deviation of bond (Å)	0.005
average *B* factor of protein atoms (Å^2)	26.0
average *B* factor of water atoms (Å^2)	33.1
average *B* factor of ligand atoms (Å^2)	32.4

[a] $R_{symm} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity for multiple measurements. [b] The R_{free} (ref 54) was calculated from a random selection of reflections constituting ∼10% of the data. The R factor was calculated with the remaining intensities.

Structure-Based Pharmacophore Generation. The crystal structures of **6** and **1** (1ENU) in complex with TGT have been superimposed by a least-squares fit based on the C_α positions. The donor features D1a and D1b indicated in Figure 6 have been defined using Unity 4.2[27] by picking the appropriate atoms of **6**. The acceptor feature Acc and the donor feature D2 have been defined by picking the appropriate atoms of **1**. The Acc/don feature has been defined by connecting a donor atom and an acceptor atom that have been located by picking the appropriate atom of **6** via the partial match utility. To consider the directionality of the hydrogen bonds, corresponding sites have been attributed to the neighboring atoms of the protein and the interstitial water molecule. The hydrophobic moiety has been spatially characterized as the centroid of the benzoic ring of **1**. D2 and the Acc/don feature as well as the connection of Acc to Gly230 and Glu203 have been connected by the partial match utility (Figure 6). The receptor site has been defined by picking the appropriate atoms of the binding pocket. A spatial constraint has been added to all features. The diameters of the features on the ligand site have been set to the spread of the indicated "hot spots". The diameters for the feature spheres on the protein site have been adjusted in such a way that a test sample composed of already known inhibitors for the TGT (Figure 2) could pass the filter.[18,22]

Construction of 3D Databases. The 2D connection tables of entries in the Available Chemicals Directory (ACD),[44] IBS Library,[45] ChemStar Library,[46] ASINEX Express Gold Collection (AEGC), ASINEX Express Platinum Collection (AEPC),[47] LeadQuest,[48] and AMBINTER[49] have been converted into 3D structures using CORINA.[50] Thereby, missing hydrogens have been added and small disconnected fragments (e.g., counterions) have been discarded. Subsequently, the generated 3D structures were transferred into Unity databases by applying the Unity utility "dpimport" followed by the standard 2D and macrofingerprint generation based on "dbmkscreen".[27]

Hierarchical Filtering. First, the databases have been screened by applying the Selector compound filtering utility[28] to retrieve compounds with less than eight rotatable bonds and a molecular weight of less than 450 Da. The obtained hit lists have been further selected by applying the same utility to retrieve molecules that possess at least two hydrogen-bond donors, one hydrogen-bond acceptor, and one hydrophobic feature (a five- or six-membered ring) in accordance to the above-described requirements of the protein-derived pharmacophore. The remaining hits have been further filtered in a

subsequent step by Unity to match in a flexible fashion the proposed pharmacophore hypothesis, first ignoring but then considering excluded volume constraints assigned to the adjacent active-site residues. During the former step, the computing time maximally spent per structure has been restricted to 180 s, and during the step considering volume constraints, it was restricted to 300 s. To focus on leads that are suitable for optimization,[30,31] Lipinski's "rule of five" check has been superseded. Duplicates, occasionally present in the different databases, were then removed from the hit lists using the Unix shell script "nodbdups" provided with Unity.

Docking. During the Unity searches, different tautomeric forms have been considered. To prepare the preselected hits for a final docking with FlexX,[29] manual corrections to obtain the desired tautomeric form have been performed using SYBYL. With an in-house script, exocyclic guanidino and amidino groups or primary and secondary aliphatic amino groups have been protonated. Similarly, phosphoric, sulfonic, and carboxylic groups have been deprotonated.

Docking calculations were performed utilizing FlexX, version 1.102. The FlexX scoring function has been used during the complex construction phase. The obtained solutions were reranked by DrugScore[51] as implemented in FlexX. The particle concept,[52] which is capable of considering water molecules during the incremental construction phase, was activated. For each compound, 30 docking solutions have been generated. These solutions have been checked using an in-house PYTHON script to find the solution with the best score that matches the original pharmacophore hypothesis. Only compounds passing these criteria have been further considered.

Force-Field Minimization and Quantum Chemical Calculations. Ligands have been minimized in the binding pocket exhibiting the protein conformation as complexed with **1** or **6** using the MAB force-field.[32] The default parameters have been applied, and the binding pocket has been kept rigid.

The quantum chemical calculations for the tautomers of compound **12** have been performed with the HF method using the 6-31G(d) basis set and ligand geometries optimized at the same level. Gaussian 98 was used for these calculations.[53]

Kinetic Analysis. The apparent K_i values have been measured as described elsewhere.[18]

Protein Data Bank Accession Codes. The atomic coordinates and structure factors for TGT·**6** have been deposited in the RCSB Protein Data Bank with ID code 1N2V.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft (Grant KL-1204/1). We thank Dr. C. Sotriffer for his help in quantum chemical calculations and C. Sohn for his assistance in the X-ray measurements. We are grateful to Tripos (Munich, Germany) for making the Tripos software available to us and to MDL (San Leandro, CA) for a copy of the ACD library.

References

(1) Kotloff, K.; Winickoff, J.; Ivanoff, B.; Clemens, J.; Swerdlow, D.; et al. Global burden of Shigella infections: implications for vaccine development and implementation of control strategies. *Bull. WHO* **1999**, *77*, 651−666.

(2) Replogle, M. L.; Fleming, D. W.; Cieslak, P. R. Emergence of antimicrobial-resistant shigellosis in Oregon. *Clin. Infect. Dis.* **2000**, *30*, 515−519.

(3) Khan, W. A.; Seas, C.; Dhar, U.; Salam, M. A.; Bennish, M. L. Treatment of shigellosis: V. Comparison of azithromycin and ciprofloxacin. A double-blind, randomized, controlled trial. *Ann. Intern. Med.* **1997**, *1*, 697−703.

(4) Overcoming Antimicrobial Resistance, Word Health Report on Infectious Diseases. http://www.who.int/infectious-disease-report/2000/index.html (accessed 2000).

(5) Okada, N.; Sasakawa, C.; Tobe, T.; Yamada, M.; Nagai, S.; et al. Virulence-associated chromosomal loci of *Shigella flexneri* identified by random Tn5 insertion mutagenesis. *Mol. Microbiol.* **1991**, *5*, 187−195.

(6) Durand, J. M.; Dagberg, B.; Uhlin, B. E.; Bjork, G. R. 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.* **2000**, *35*, 924−935.

(7) Okada, N.; Nishimura, S. Isolation and characterization of a guanine insertion enzyme, a specific tRNA transglycosylase, from *Escherichia coli*. *J. Biol. Chem.* **1979**, *254*, 3061−3066.

(8) Kittendorf, J. D.; Barcomb, L. M.; Nonekowski, S. T.; Garcia, G. A. tRNA-Guanine Transglycosylase from *Escherichia coli*: Molecular Mechanism and Role of Aspartate 89. *Biochemistry* **2001**, *40*, 14123−14133.

(9) Slany, R. K.; Kersten, H. Genes, enzymes and coenzymes of queuosine biosynthesis in procaryotes. *Biochimie* **1994**, *76*, 1178−1182.

(10) Frey, B.; McCloskey, J.; Kersten, W.; Kersten, H. New function of vitamin B12: cobamide-dependent reduction of epoxyqueuosine to queuosine in tRNAs of *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **1988**, *170*, 2078−2082.

(11) Curran, L. Modified nucleosides in translation. In *Modifiaction and Editing of RNA*; American Society for Microbiology Press: Washington, DC, 1998; pp 493−516.

(12) Björk, G. R. Biosynthesis and function of modified nucleosides in tRNA. In *tRNA: Structure, Biosynthesis, and Function*; American Society for Microbiology Press: Washington, DC, 1995; pp 165−205.

(13) Björk, G. Stable RNA modification. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd ed.; American Society for Microbiology Press: Washington, DC, 1996; pp 861−886.

(14) Durand, J. M.; Okada, N.; Tobe, T.; Watarai, M.; Fukuda, I.; et al. 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.* **1994**, *176*, 4627−4634.

(15) Romier, C.; Reuter, K.; Suck, D.; Ficner, R. Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. *EMBO J.* **1996**, *15*, 2850−2857.

(16) Romier, C.; Meyer, J. E.; Suck, D. Slight sequence variations of a common fold explain the substrate specificities of tRNA-guanine transglycosylases from the three kingdoms. *FEBS Lett.* **1997**, *416*, 93−98.

(17) Klebe, G. Recent developments in structure-based drug design. *J. Mol. Med.* **2000**, *78*, 269−281.

(18) Grädler, U.; Gerber, H. D.; Goodenough-Lashua, D. M.; Garcia, G. A.; Ficner, R.; et al. A New Target for Shigellosis: Rational Design and Crystallographic Studies of Inhibitors of tRNA-guanine Transglycosylase. *J. Mol. Biol.* **2001**, *306*, 455−467.

(19) Meyer, E. A. Unpublished results.

(20) Nærum, L.; Graedler, U.; Klebe, G. Virtual screening based on tRNA-guanine transglycosylase. In *Rational Approaches to Drug Design*; Prous Science: Barcelona, Spain, 2001; pp 400−402.

(21) Grädler, U.; Ficner, R.; Garcia, G. A.; Stubbs, M. T.; Klebe, G.; et al. Mutagenesis and crystallographic studies of *Zymomonas mobilis* tRNA-guanine transglycosylase to elucidate the role of serine 103 for enzymatic activity. *FEBS Lett* **1999**, *454*, 142−146.

(22) Meyer, E. A.; Brenk, R.; Castellano, R. K.; Furler, M.; Klebe, G.; et al. De Novo Design, Synthesis, and in Vitro Evaluation of Inhibitors for Prokaryotic tRNA-guanine Transglycosylase: A Dramatic Sulfur Effect on Binding Affinity. *ChemBioChem* **2002**, *3*, 250−253.

(23) Verdonk, M. L.; Cole, J. C.; Taylor, R. SuperStar: a knowledge-based approach for identifying interaction sites in proteins. *J. Mol. Biol.* **1999**, *289*, 1093−108.

(24) Gohlke, H.; Hendlich, M.; Klebe, G. Predicting binding modes, binding affinities and "hot spots" for protein−ligand complexes using a knowledge-based scoring function. *Perspect. Drug Discov. Des.* **2000**, *20*, 115−144.

(25) Grüneberg, S.; Wendt, B.; Klebe, G. Subnanomolar Inhibitors from Computer Screening: A Model Study Using Human Carbonic Anhydrase II. *Angew. Chem., Int. Ed.* **2001**, *40*, 389−393.

(26) Grüneberg, S.; Stubbs, M.; Klebe, G. Successful Virtual Screening for Novel Inhibitors of Human Carbonic Anhydrase: Stragegy and Experimental Confirmation. In press.

(27) *Unity, Chemical Information Software*, version 4.2.1; Tripos, Inc.: St. Louis, MO.

(28) *Selector*; Tripos, Inc.: St. Louis MO.

(29) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(30) Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856−864.

(31) Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D. Is There a Difference between Leads and Drugs? A Historical Perspective. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1308−1315.

(32) Gerber, P. R. Charge distribution from a simple molecular orbital type calculation and non-bonding interaction terms in the force field MAB. *J. Comput.-Aided. Mol. Des.* **1998**, *12*, 37−51.
(33) Jacobson, K. B.; Farkas, W. R.; Katze, J. R. Presence of queuosine in Drosophila melanogaster: correlation of free pool with queuosine content of tRNA and effect of mutations in pteridine metabolism. *Nucleic Acids Res.* **1981**, *9*, 2351−2366.
(34) Hoops, G. C.; Townsend, L. B.; Garcia, G. A. Mechanism-based inactivation of tRNA-guanine transglycosylase from *Escherichia coli* by 2-amino-5-(fluoromethyl)pyrrolo[2,3-d]pyrimidin-4(3*H*)-one. *Biochemistry* **1995**, *34*, 15539−15544.
(35) Hiremath, S. P.; Hiremath, D. M.; Purohit, M. G. Synthesis of substituted 2-(5′-oxo-thioxo-1′,3′,4′-oxadiazol-2′-yl)-indoles and 2-(5′-oxo-thioxo-1,3,4′-oxadiazol-2′-ylamino)indoles. *Indian J. Chem., Sect. B* **1983**, *22*, 571−576.
(36) Al-Dossary, F. S.; Ong, L. T.; Correa, A. G.; Starke, J. R. Treatment of childhood tuberculosis with a six month directly observed regimen of only two weeks of daily therapy. *Pediatr. Infect. Dis. J.* **2002**, *21*, 91−97.
(37) Reuter, K.; Ficner, R. Sequence analysis and overexpression of the *Zymomonas mobilis* tgt gene encoding tRNA-guanine transglycosylase: purification and biochemical characterization of the enzyme. *J. Bacteriol.* **1995**, *177*, 5284−5288.
(38) *SYBYL, Chemical Information Software*, version 6.7; Tripos, Inc.: St. Louis, MO.
(39) Gerber, P. R.; Muller, K. MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 251−268.
(40) Romier, C.; Ficner, R.; Reuter, K.; Suck, D. Purification, crystallization, and preliminary X-ray diffraction studies of tRNA-guanine transglycosylase from *Zymomonas mobilis*. *Proteins* **1996**, *24*, 516−519.
(41) Otwinowski, Z. *DENZO*; Yale University: New Haven, CT.
(42) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905−921.
(43) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **1991**, *47*, 110−119.
(44) Web site: http://www.mdli.com.
(45) Web site: http://www.ibscreen.com.
(46) Web site: http://www.chemstar.ru.
(47) Web site: http://asinex.com.
(48) Web site: http://leadquest.tripos.com.
(49) Web site: http://www.ambinter.com.
(50) Sadowski, J.; Schwab, C. H.; Gasteiger, J. *CORINA, 3D Structure Generator*, version 2.0; University Erlangen-Nürnberg: Erlangen, Germany.
(51) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge-based scoring function to predict protein−ligand interactions. *J. Mol. Biol.* **2000**, *295*, 337−356.
(52) Rarey, M.; Kramer, B.; Lengauer, T. The particle concept: placing discrete water molecules during protein−ligand docking predictions. *Proteins* **1999**, *34*, 17−28.
(53) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
(54) Bruenger, A. T. The free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **1992**, *355*, 472−474.

JM0209937