

**Investigations on *lin*-Benzopurines With Respect to  
Dissociation Behavior, Pocket Cross-Talk, Targeting Resistance  
Mutants, Residual Mobility, and Scaffold Optimization**

**Dissertation**

zur Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Pharmazie

der PHILIPPS-UNIVERSITÄT MARBURG

vorgelegt von

**Manuel Neeb**

aus Limburg a. d. Lahn

Marburg/Lahn 2014

Dem Fachbereich Pharmazie der Philipps-Universität Marburg  
als Dissertation eingereicht am: 24.09.2014

Erstgutachter: Prof. Dr. Gerhard Klebe

Zweitgutachter: Prof. Dr. Klaus Reuter

Tag der mündlichen Prüfung: 05.11.2014

Hochschulkennziffer: 1180

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Herrn Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie des Fachbereichs Pharmazie der Philipps-Universität Marburg in der Zeit vom Juli 2010 bis April 2014 durchgeführt.



*Meiner Familie*

## Abbreviations

2xYT	Double amount of yeast extract and tryptone
Å	Angström (1 Å = 10 <sup>-10</sup> m)
A <sub>280nm</sub>	Absorption at 280 nm
ADP	Atomic displacement parameters
Ala	Alanine
APS	Ammonium peroxydisulfate
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
bcc	Bond charge correction
BSA	Bovine serum albumin
<i>B. taurus</i>	<i>Bos taurus</i>
C	Celsius
CAC	Critical aggregation concentration
CaDA	Cationic dummy atom
clog P	Calculated logarithm of the octanol-water partition coefficient
cpm	Counts per minute
CSD	Cambridge Structural Database
CYM	Anionic form of cysteine in MD simulations
Cys	Cysteine
ΔG <sup>0</sup>	Gibbs free energy
ΔH <sup>0</sup> /ΔH <sub>bind</sub>	Enthalpy of binding
ΔH <sub>ion</sub>	Heat of ionization
ΔH <sub>obs</sub>	Observed enthalpy of binding not corrected for heat of ionization
ΔS <sup>0</sup>	Entropy of binding
d	Distance
D	Diffusion coefficient
Da	Dalton
DAQ	2,6-diamino-3H-quinazolin-4-one
dd	Double distilled
DD/AA	Donor-donor/acceptor-acceptor motif
d(H)	Hydrodynamic diameter

DLS	Dynamic light scattering
DMSO	Dimethylsulfoxide
ds	Double stranded
DTT	1,4-Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
ECY2	Unmodified <i>Escherichia coli</i> tRNA <sup>Tyr</sup>
EDTA	Ethylenediaminetetraacetate
f-	femto-
FAE	Follicle-associated epithelium
FBLD	Fragment based lead discovery
F <sub>c</sub>	Calculated structure factor amplitudes
F <sub>o</sub>	Observed structure factor amplitudes
g	gramm
GA	Genetic algorithm
gaff	General amber force field
GE	Group efficiency
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
HABA	2-(4-Hydroxyphenylazo)-benzoic acid
H-bond	Hydrogen bond
Hepes	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethanesulfonic acid
HID	Histidine with its hydrogen at $\delta$ -position in MD simulations
HIE	Histidine with its hydrogen at $\epsilon$ -position in MD simulations
HIN	Anionic form of histidine in MD simulations
His	Histidine
HPLC	High performance liquid chromatography
HTS	High Throughput Screening
Hz	Hertz
IFN	Interferone
IL	Interleukine
Ile	Isoleucine
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
IQR	Interquartile range
ITC	Isothermal titration calorimetry

K	Kelvin
$k_B$	Boltzmann constant ( $1.380\ 650 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1}$ )
kb	Kilo-base pair
kcal	Kilocalorie
$k_{\text{cat}}$	Catalytic constant
$K_d$	Dissociation constant
$K_i$	Inhibitory constant
kJ	Kilojoule
$K_M$	Michaelis-Menten constant
LB	Luria-Bertani
LE	Ligand efficiency
Leu	Leucine
Lys	Lysine
m	Meter
m-	milli-
$\mu$ -	micro-
M	Molarity ( $\text{mol} \cdot \text{l}^{-1}$ )
MD	Molecular dynamics
MES	2-(N-Morpholino)ethanesulfonic acid
Met	Methionine
min	Minute
l	Litre
MR	Molecular replacement
MWCO	Molecular weight cut-off
n-	nano-
NMR	Nuclear magnetic resonance
$\text{OD}_{600}$	Optical density at 600 nm
p-	pico-
PAMPA	Parallel artificial membrane permeability assay
PCR	Polymerase chain reaction
PDB	Protein data bank
PDB ID	Protein data bank identifier
PEG	Polyethylene glycol
PEOE	Partial equalization of orbital electronegativities
pH	Negative decimal logarithm of the hydrogen ion activity



Phe	Phenylalanine
pK <sub>a</sub>	Negative decimal logarithm of the acid dissociation constant
PME	Particle Mesh Ewald
Pro	Proline
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
Q1/Q3	First/third quartile
rpm	Rounds per minute
s/sec	Second
SAR	Structure-activity relationship
SDS page	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	Serine
SPR	Surface plasmon resonance
T	Temperature in Kelvin
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
TGT	tRNA-guanine transglycosylase
Thr	Threonine
TLS	Translation/libration/screw-motion
Tricine	N-(Tris(hydroxymethyl)methyl)glycine
Tris	2-Amino-2-(hydroxymethyl)-propane-1,3-diol
tRNA	Transfer ribonucleic acid
Trp	Tryptophane
Tyr	Tyrosine
UV	Ultraviolet
V	Volume
v <sub>0</sub>	Initial velocity
Val	Valine
v <sub>max</sub>	Maximal velocity
v/v	Volume per volume
W	Watt
WT	Wild type
w/v	Weight per volume
<i>Z. mobilis</i>	<i>Zymomonas mobilis</i>
η	Viscosity
°	Degree

## Table of Contents

Abbreviations .....	6
1. Introduction and Motivation .....	14
1.1 Structure-Based Drug Design .....	14
1.2 Aims of the thesis .....	15
1.3 Shigellosis .....	16
1.4 Therapy of Shigellosis .....	16
1.5 Classification of <i>Shigella</i> Bacteria .....	18
1.6 Molecular and Cellular Pathogenicity of <i>Shigella</i> .....	18
1.7 Biosynthesis of Queuine in Prokaryotes .....	21
1.8 Base Exchange Mechanism of TGT .....	23
1.9 Structure of <i>Z. mobilis</i> TGT .....	24
1.9.1 Secondary Structure .....	24
1.9.2 Dimeric Assembly .....	26
1.9.3 Active Site .....	27
2. Impact of Protein and Ligand Impurities on ITC-derived Protein-Ligand Thermodynamics .....	29
2.1 Introductory Remarks .....	29
2.2 Abstract .....	29
2.3 Introduction .....	30
2.4 Material and Methods .....	32
2.4.1 Protein Preparation and Quantification .....	32
2.4.2 Isothermal Titration Calorimetry .....	32
2.4.3 Dynamic Light Scattering .....	33
2.5 Results .....	35
2.5.1 Errors in Protein Concentration Did Not Affect the Thermodynamic Profiles of High Affinity Protein-Ligand Binding .....	35
2.5.2 Organic Ligand Contaminations Can Affect Thermodynamic Profiles of Protein-Ligand Binding Significantly .....	36
2.5.3 Inorganic Ligand Contaminations Can Also Affect Thermodynamic Profiles of Protein-Ligand Binding .....	40
2.6 Comparative discussion .....	42
2.7 Conclusions .....	46
3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and pK <sub>a</sub> Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme .....	48

3.1	Introductory Remarks.....	48
3.2	Abstract .....	48
3.3	Introduction.....	49
3.4	Results and Discussion .....	51
3.4.1	Selected Ligands and Binding Mode of <i>lin</i> -Benzopurines to <i>Z. mobilis</i> TGT .....	51
3.4.2	Protonation States of <i>lin</i> -Benzopurine-TGT Complexes .....	54
3.4.3	Mutational Studies .....	59
3.5	Conclusion .....	66
4.	Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme .....	68
4.1	Introductory Remarks.....	68
4.2	Abstract .....	68
4.3	Introduction.....	69
4.4	Results and Discussion .....	71
4.4.1	Binding Affinities and Thermodynamic Profiles of the 2-Amino- <i>lin</i> -benzopurines.....	71
4.4.2	Crystal Structures and Binding Modes of 2-Amino- <i>lin</i> -benzopurines in <i>Z. mobilis</i> TGT	74
4.4.3	Correlation of Thermodynamic Signatures and Binding Modes .....	80
4.5	Conclusions.....	87
5.	Creating a Resistance Model for TGT: the Effect of Mutations on Flexible <i>lin</i> -Benzoguanine Substituents.....	89
5.1	Introduction.....	89
5.2	Results .....	90
5.2.1	Mutations of Wild Type TGT.....	90
5.2.2	Binding Affinities .....	93
5.2.3	Exploring the Adaptations of <b>4a</b> Targeting the Mutated TGT Variants.....	94
5.3	Discussion .....	97
5.4	Conclusions.....	99
6.	Addressing a New Subpocket of TGT by Elongated 2-Amino- <i>lin</i> -benzoguanines .....	101
6.1	Introduction.....	101
6.2	Results and Discussion .....	102
6.2.1	Ligand Design.....	102
6.2.2	MD Simulations .....	103
6.2.3	Affinity Data.....	105
6.2.4	Crystal Structure Analysis.....	105
6.3	Conclusions.....	109

7.	5-Azacytosine as a Novel Scaffold to Inhibit <i>Z. mobilis</i> TGT with Expected Improved Bioavailability and Synthetic Accessibility.....	112
7.1	Introduction.....	112
7.2	Results and Discussion .....	114
7.2.1	Ligand Design.....	114
7.2.2	Inhibition Mode and Affinity Data.....	116
7.2.3	Putative Agglomeration Behaviour of the Naphthyl Derivative <b>11</b> .....	117
7.2.4	Crystal Structure Analysis.....	120
7.2.5	MD Simulation.....	124
7.2.6	Docking Solutions .....	125
7.3	Conclusions.....	128
8.	Summary and Outlook.....	130
8.	Zusammenfassung und Ausblick .....	137
9.	Materials and Methods .....	146
9.1	<i>In-silico</i> Methods .....	146
9.1.1	pK <sub>a</sub> Calculations .....	146
9.1.2	Molecular Dynamic Simulation .....	146
9.1.3	Docking.....	147
9.2	Experimental Methods.....	148
9.2.1	Chemicals.....	148
9.2.2	Equipment .....	149
9.2.3	Buffer, Solutions and Media.....	150
9.2.4	Bacterial Strains and Plasmids.....	152
9.2.5	Molecular biological methods.....	153
9.2.5.1	Cloning.....	153
9.2.5.2	Protein Expression and Purification .....	154
9.2.5.2.1	Expression of the <i>Z. mobilis tgt</i> gene via the vector pET9d .....	154
9.2.5.2.2	Expression of the <i>Z. mobilis tgt</i> gene via the vector pPR-IBA2 .....	155
9.2.5.3	Mutagenesis .....	156
9.2.5.4	<i>In-vitro</i> Transcription and Purification of <i>E. coli</i> tRNA <sup>Tyr</sup> .....	158
9.2.5.5	Agarose Gel Electrophoresis.....	158
9.2.5.6	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis.....	158
9.2.5.7	Determination of Protein and DNA Concentrations .....	159
9.2.6	Enzyme Kinetics.....	160
9.2.6.1	Trapping Experiment .....	160

9.2.6.2	Enzyme Assay .....	161
9.2.6.3	Kinetic characterization.....	162
9.2.7	Isothermal Titration Calorimetry.....	163
9.2.8	Dynamic Light Scattering.....	164
9.2.9	X-ray Crystallography .....	165
9.2.9.1	<i>Z. mobilis</i> TGT Crystallization.....	165
9.2.9.2	pH Soaking.....	166
9.2.9.3	Data Collection .....	166
9.2.9.4	Calculation of the Anomalous Map.....	167
9.2.9.5	Structure Determination and Refinement .....	167
9.3	Synthesis.....	168
10.	Appendix.....	169
10.1	Calculation of $pK_a$ Values of Titratable Groups of the TGT Active Site.....	169
10.1.1	TGT Wild Type .....	169
10.1.2	TGT Asp102Asn.....	170
10.1.3	TGT Asp156Asn.....	171
10.2	Anomalous Density Data .....	172
10.3	Thermodynamic Data of <i>lin</i> -Benzopurines.....	174
10.4	Sequencing Results of the Strep-tagged TGT Variants after Cleavage.....	175
10.5	Mass Spectra of the Strep-tagged TGT Variants after Cleavage .....	177
10.6	Enzyme Kinetics of the Strep-tagged TGT Variants after Cleavage.....	178
10.7	Data Collection and Refinement Statistics .....	179
	References.....	184
	Acknowledgement.....	197
	Eidesstattliche Erklärung.....	199
	Publications and Poster Presentations.....	200
	Curriculum Vitae.....	201

# 1. Introduction and Motivation

## 1.1 Structure-Based Drug Design

Structure-based drug design is an important tool in the development of novel drugs and represents the basis for the rational design of new drugs. Thereby, the setup of experiments relies on the knowledge of macromolecular target structures mainly derived by X-ray crystallography or NMR spectroscopy. With their help complex interaction patterns of potential drug candidates within the binding pocket can be visualized in the early stage of lead structure generation as well as in further optimization steps. Metaphorically speaking, this procedure provides a three-dimensional model to find the right key (ligand) that optimally fits the lock (target protein) [Klebe, 2009]. The number of crystal structures available in the protein data bank (PDB) increases drastically. Up to date, nearly 100'000 structures are deposited of which more than 90'000 were determined by X-ray crystallography [PDB, April 2014].

However, information concerning the driving forces for binding affinity cannot be derived from structural data alone. Therefore, the concept of shape-complementarity has to be expanded by thermodynamic data, which have been shown to be an essential complement to crystal structure analyses [Henry, 2001]. In this context, isothermal titration calorimetry (ITC) is not only able to elucidate the potency of a ligand but also its enthalpic contribution and the stoichiometry of the binding process within one experiment.

Taking all these data together, structure-based drug design is able to produce highly potent and selective drug candidates, thus, saving a tremendous amount of time, money and laboratory animals.

The enzyme tRNA-guanine transglycosylase (TGT) is one example to which the method of structure-based drug design was applied. TGT plays a key role in the pathogenicity of *Shigella*, the causative agent of shigellosis [Durand *et al.*, 1994]. The crystal structure of the binary complex containing the closely similar TGT from *Z. mobilis* and preQ1, one of its natural substrates, served as a starting point. After initial computational approaches and subsequent iterative optimization cycles *lin*-benzopurines were found to be highly potent inhibitors of TGT [Hörtner *et al.*, 2007; Stengl *et al.*, 2007; Kohler *et al.*, 2009; Barandun *et al.*, 2012]. Unfortunately, *lin*-benzopurines show unsatisfying membrane permeation ability and therefore, need further optimization [Barandun *et al.*, 2012].

## 1.2 Aims of the thesis

In the framework of this thesis different aspects concerning the optimization of the highly potent *lin*-benzopurines should be analyzed. These comprised the setup of new methodologies as well as subsequent investigations with respect to bioavailability, resistance development, binding mode predictions by computational approaches as well as fragment-based lead discovery of new inhibitor classes:

- A method was established to obtain the thermodynamic signature of *lin*-benzopurines binding to TGT via isothermal titration calorimetry (ITC). Subsequently, the robustness of this method was analyzed with respect to organic and inorganic ligand impurities as well as protein impurities (Chapter 2).
- Based on the observation that *lin*-benzopurines exhibit a poor ability to penetrate through cell membranes, the protonation states of *lin*-benzoguanines and *lin*-benzohypoxanthines in solution and in the TGT-bound state were investigated by  $pK_a$  calculations and ITC measurements. Potential protonation sites were verified using site-directed mutagenesis and crystal structure analysis (Chapter 3).
- A proposed entropic term of a series of 2-amino-*lin*-benzoguanines and 2-amino-*lin*-benzohypoxanthines with varying substituents in 2-position that were shown to exhibit an increased flexibility was analyzed using ITC. Screening for modified crystallization conditions gave insight in the so far unknown binding mode of the 2-substituent and allowed a correlation of thermodynamic properties to the substrate recognition mechanism and pocket cross-talk of TGT (Chapter 4).
- Mutagenesis combined with affinity measurements should show whether a binder with high residual mobility can avoid a loss in binding affinity compared to a binder adopting one ordered binding mode and therefore achieve an advantage as a more competent antibiotic less affected by resistance mutations (Chapter 5).
- MD simulations were used to predict the binding mode of a new series of extended 2-amino-*lin*-benzoguanines. Crystal structure analysis of these ligands revealed a so far novel binding mode occupying further subpockets of TGT (Chapter 6).
- In a fragment-based approach a new class of ligands with expected improved bioavailability and synthetic accessibility based on a 5-azacytosine scaffold was investigated in a comprehensive *in-silico* and *in-vitro* study (Chapter 7).

### 1.3 Shigellosis

Shigellosis is an infectious bowel disease that predominantly occurs in developing countries. The uptake of only 10 to 100 organisms of the causative agent *Shigella* is sufficient to result in a severe infection affecting over 165 million people annually with 1.1 million fatalities [DuPont *et al.*, 1989; Kotloff *et al.*, 1999], although more recent estimates suggest a reduced case-fatality of 90 million cases with 108'000 fatalities each year [WHO, 2009]. A remarkably high number of incidents, namely 70% of all episodes and 60% of all deaths, affect children under the age of five years [Niyogi, 2005]. However, a rising number of episodes is also found in the industrialized world, especially among military troops and travellers.

Humans represent the only natural host for *Shigella*. The disease is transmitted by the faecal-oral route mainly via contaminated food and water or due to person-to-person contact [Niyogi, 2005]. Recently, also the transmission via house flies (*Musca domestica*) has been reported [Levine and Levine, 1991]. Therefore, the occurrence of shigellosis is primarily restricted to poor populations with insufficient hygiene standards and sanitation but also to settings of large crowds, political revolutions or natural disasters .

The etiopathology of the disease depends on several factors, i. e. the infecting species of *Shigella*, the age of the host, the presence of further risk factors and the immune state of the patient. Shigellosis is often self-limiting in healthy humans, who will usually recover within five to seven days, if no further complications manifest. Within this period an initially appearing watery diarrhoea with abdominal cramps is quickly followed by fever, pain and tenesmus, which is characterized by bloody, mucoid and purulent stools [Niyogi, 2005]. With high prevalence shigellosis turns into a life-threatening situation in malnourished infants and children due to the occurrence of intestinal complications, metabolic dysfunctions, anaemia in form of the haemolytic uremic syndrome, pneumonia, and diverse inflammatory reactions [Sack *et al.*, 2001].

### 1.4 Therapy of Shigellosis

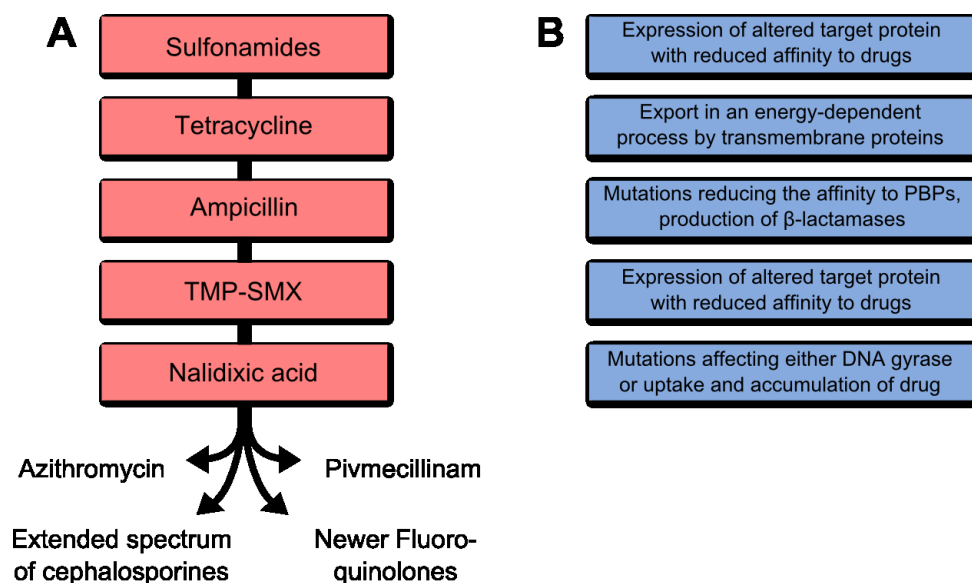
As a life-saving first step in the treatment of shigellosis the World Health Organization (WHO) recommends oral rehydration and an overall improvement in nutrition [Niyogi, 2005]. The prescription of antibiotics depends on several factors like the severity of the disease, the age of the patient and the likelihood of further transmission. For risk groups an



early microbial therapy is advised in order to reduce the duration of illness and thus the risk of complications and transmission. Disappointingly, the treatment becomes more and more challenging due to the increasing resistance of *Shigella* to current antibiotics (Figure 1.1) [Sack *et al.*, 1997]. Strains resistant to sulfonamides, tetracycline, ampicillin, and TMP-SMX exist worldwide. It has been shown that drug resistance is mediated by mobile plasmids encoding “resistance factors”, which are transferred by conjugation [Sack *et al.*, 2001].

Currently, the WHO recommends the use of ciprofloxacin as first-line antibiotic for patients of all ages [WHO, 2005]. Major concerns about the use of fluoroquinolones in children exist due to the potential side-effects of arthropathy. However, in view of life-threatening complications caused by shigellosis this risk is accepted. Alternatives to ciprofloxacin include cephalosporins (cefixime, ceftriaxone), azithromycin, and pivmecillinam but the use of these alternatives exhibits problems concerning costs, formulation, and efficacy [Mandal *et al.*, 2012].

Vaccination offers an additional alternative to the antimicrobial therapy. Thereby, immunity is directed to the O-somatic antigen, which is narrowly type specific. Up-to-date, there is no vaccine available on the market. Vaccines under development include polysaccharide conjugate and live-attenuated vaccines focusing on the *Shigella* species *S. flexneri* type 2a, *S. sonnei* and *S. dysenteriae* type 1 [WHO, 2009].



**Figure 1.1** Resistant antibiotics in the antimicrobial therapy of shigellosis. A) Antibiotics affected by resistance development of *Shigella* (red) and current alternatives. B) Mechanism of resistance in strains of *Shigella*.

## 1.5 Classification of *Shigella* Bacteria

*Shigella* bacteria belong to the family of *Enterobacteriaceae* [Sansone, 2001]. They are characterized as nonsporulating, facultative anaerobic Gram-negative rods that are unencapsulated and non-motile. The name *Shigella* results from its discoverer Kiyoshi Shiga, who first reported about this bacterium in Japan, 1898 [Niyogi, 2005]. Based on DNA hybridization analyses, *Shigella* cannot be distinguished from *Escherichia coli* on the polynucleotide level, however, it can be differentiated on the virulence phenotype. Biochemically, both species share a high similarity and thus, *Shigella* is often referred to as enteroinvasive *E. coli* (EIEC) [Escobar-Páramo *et al.*, 2004]. The genus *Shigella* comprises four different species whose classification is based on biochemical and serological variations of their O-polysaccharide portion of their LPS (Table 1.1) [Sansone, 2001].

**Table 1.1** The four *Shigella* species and their characteristics in order of their frequency.

Species	# serotypes	Geographical distribution	Etiopathology
<i>S. flexneri</i>	6	Worldwide, mainly in the developing world	Less severe
<i>S. sonnei</i>	1	Central Europe	Less severe
<i>S. boydii</i>	8	Indian and North African subcontinent	Harmless
<i>S. dysenteriae</i>	16	Tropic and subtropic regions	Most pathogenic due to shigatoxin (subtype 1)

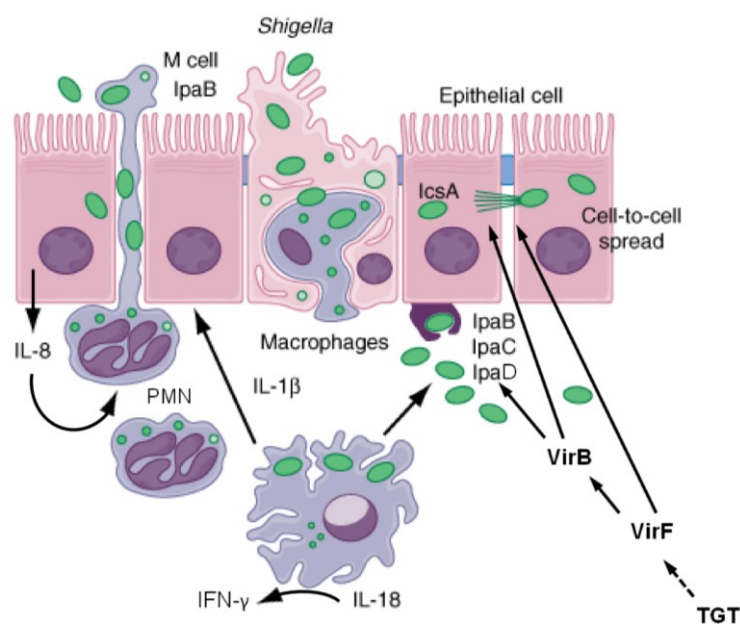
## 1.6 Molecular and Cellular Pathogenicity of *Shigella*

*Shigella* bacteria are unable to enter the epithelial cells of the colon directly from the apical side (Figure 1.2). Contrary, the basolateral side can be efficiently invaded by the bacteria. Thus, they have to choose specific areas to translocate the epithelium in order to infect the host cells [Sansone, 2001].

M-cells belong to the follicle-associated epithelium (FAE). Their task is to transport luminal antigens across the epithelium and present them to mucosa-associated lymphoid follicles filled with lymphocytes and macrophages, which then initiate a mucosal immune reaction. These cells offer *Shigella* the opportunity to translocate to the basolateral pole.

Once invaded into M-cells the bacteria escape from the endocytic vacuole without damage and reach the associated macrophages, where it gets phagocytosed. Within the macrophage *Shigella* bacteria activate the cysteine protease caspase 1, which in return causes apoptosis of the invaded immune cells. In its second function caspase 1 activates the proinflammatory signaling by hydrolyzing (IL)-1 $\beta$  and pro-IL-18 to their mature forms IL-1 $\beta$  and IL-18, respectively [Sansonetti, 2001]. The release of the bacteria and mature interleukins out of the macrophages leads to a severe inflammation. IL-1 $\beta$  triggers the loss of integrity of the epithelial barrier causing further diffusion and invasion of *Shigella*. IL-18 acts as an interferon (IFN)- $\gamma$  inducer targeting natural killer cells and T lymphocytes.

At the basolateral side of the epithelium free *Shigella* bacteria enter the host cells via macropinocytosis. After having escaped from the phagocytic vacuole into the cytoplasm, the bacteria are able to spread intra- and intercellularly and thus, colonize the epithelium. The infected cells release IL-8 that in combination with IL-1 $\beta$  recruits polymorphonuclear leukocytes (PMNs) in subepithelial tissues [Sansonetti, 2001; Jennison and Verma, 2004]. On the one hand, PMNs are, contrary to macrophages, able to phagocytose and kill *Shigella* bacteria. On the other hand, however, they simultaneously increase their M-cell independent invasion to the basolateral side by facilitating transmigration through the epithelial cells due to tissue destruction.



**Figure 1.2** Invasion mechanism of *Shigella* into epithelial cells of the colon and involved virulence factors modified according to Sansonetti and Bergounioux, 2008.

For the apoptotic death of macrophages, the invasion of *Shigella* bacteria into the eukaryotic cells of the colon and the cell-to-cell spread several virulence factors are required [Nhieu *et al.*, 2000; Fernandez and Sansonetti, 2003].

The macrophage-induced apoptosis depends on the secretion of IpaB, which is responsible for the activation of caspase 1. For the invasion of epithelial host cells the bacteria form a pore into the cytoplasmic membrane via a *Shigella* type III secretion [Sansonetti, 2001]. This tube-like apparatus contains a hydrophobic complex composed of IpaB and IpaC proteins. Its function is firstly the secretion of various proteins into the cytoplasm of the host cell and secondly the induction of internalization of shigellae due to a signalling cascade leading to polymerization and depolymerization of the actin tubuli via the C-terminal domain of IpaC. The maturation of the thus formed cell extension underlies then the control of IpaA and IpgD. Cell-to-cell spread of *Shigella* bacteria also relies on actin polymerization and is achieved by the surface-exposed virulence factor IcsA (VirG).

The complete sequence of the *S. flexneri* virulence plasmid is known. Most of the invasion genes are present within a 30-kb region of the 214-kb virulence plasmid, which is named *Shigella* pathogenicity island (PAI) [Dorman and Porter, 1998]. The expression of its chief components is kept under the control of the *virF* and *virB* gene products (Figure 1.2). VirB is directly involved in the transcription of a large number of pathogenicity genes, while VirF regulates this process indirectly by controlling the concentration of VirB. In addition, VirF directly activates transcription of the *icsA* (*virG*) gene. Therefore, VirF exhibits a major role in the pathogenicity cascade of *Shigella*.

Interestingly, a certain threshold level of VirF is a prerequisite for the transcription of virulence factors. This level within the bacterial cell is influenced by several factors like nutrition factors, temperature, pH value, and osmolarity [Durand *et al.*, 2000]. Optimal conditions comprise a temperature of 37°C at pH 7.4 and a moderate osmotic level similar to physiological saline as well as the presence of the free amino acids arginine and methionine. However, the expression of VirF can also be effected on the translational level since for the ribosomal translation of VirF-mRNA modified tRNA is required. These modifications comprise two positions within the tRNA: Firstly, guanine-34 in the wobble position of the anticodon has to be exchanged against the highly modified base queuosine. Secondly, the modified nucleoside 2-methylthio-N<sup>6</sup>-isopentenyladenosine has to be incorporated at position 37. The relevance of the thus derived tRNA in respect to the VirF concentration has

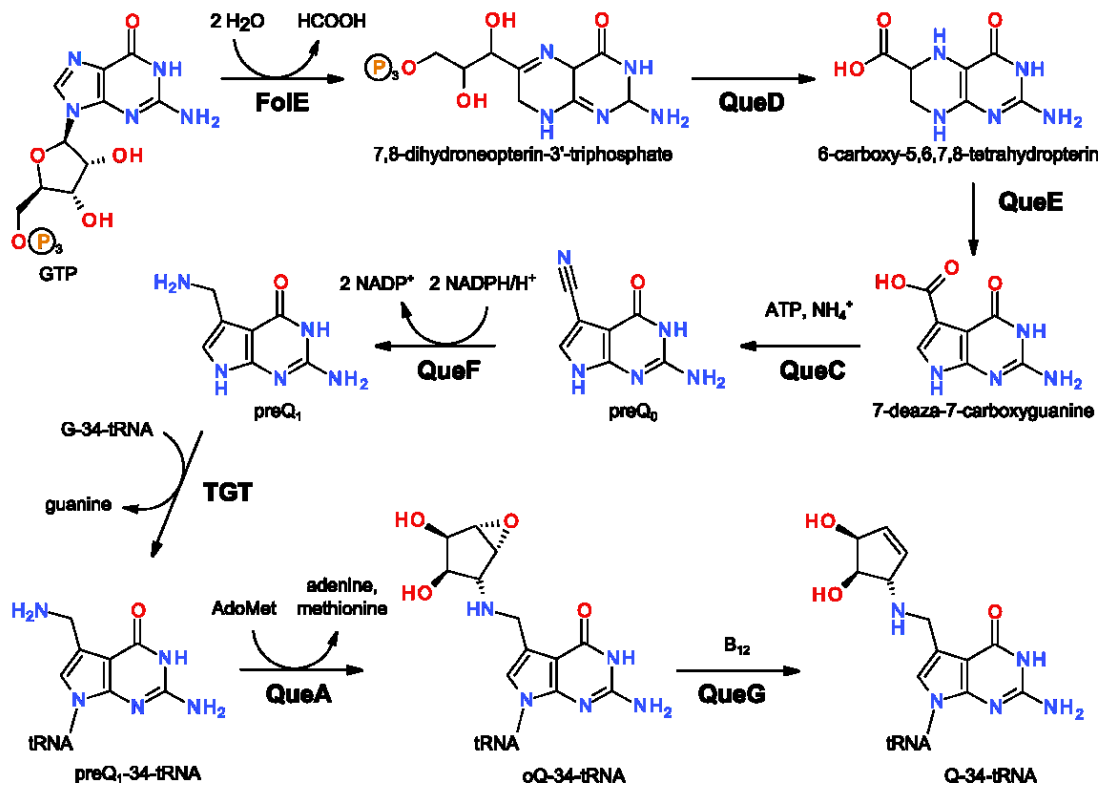
been tested in gene-knock-out experiments. The modification in position 37 depends on the *miaA* gene product. Mutational changes of the *miaA* gene decrease the intracellular concentration of VirF to 10% resulting in a drop of haemolytic activity to 10 – 20% compared to the wild type. The incorporation of a queuine precursor in position 34 of the tRNA is catalyzed by the *tgt* (*vacC*) gene product. Mutation of the corresponding gene reduces the VirF concentration and the haemolytic activity to 50 – 60% compared to the wild type [Durand *et al.*, 1994; Durand *et al.*, 1997; Durand *et al.*, 2000].

These investigations demonstrate that the above-named tRNA-modifying enzymes represent promising drug targets leading to a significant decrease in the pathogenicity of *Shigella*. Potent inhibitors exhibit the potential to act as selective antibiotics potentially less affected by resistance development. The *tgt* gene product tRNA-guanine transglycosylase isolated from *E. coli* is well characterized in diverse studies. *Zymomonas mobilis* TGT serves as a model system sharing a sequence identity of 60.4% with the orthologue of *S. flexneri* [Reuter and Ficner, 1995]. As expression and crystallization protocols are available for this system, it can be used as the basis for structure-based drug design [Reuter and Ficner, 1995; Romier *et al.*, 1996; Romier *et al.*, 1996].

## 1.7 Biosynthesis of Queuine in Prokaryotes

Until today, over 70 modifications in nucleobases of transfer RNA have been identified [Hoops *et al.*, 1995]. One of these modified nucleobases is queuine, which replaces the genetically encoded guanine in position 34 (the “wobble” position) of particular tRNAs. Remarkably, queuine is the sole RNA base in which the parent basic ring system is modified. It is present in almost all eucaryotes and bacteria. Rare exceptions are *Saccharomyces cerevisiae* and some Actinobacteria (*Mycobacterium*, *Corynebacterium*, *Streptomyces*, *Bifidobacterium*) [Reader *et al.*, 2004]. The base exchange reaction in which guanine is substituted against the queuine precursor preQ<sub>1</sub> is catalyzed by the enzyme tRNA-guanine transglycosylase (Figure 1.3) [Hoops *et al.*, 1995]. In only four different tRNA variants the incorporation of queuine can be observed. As a prerequisite they have to exhibit the anticodon sequence U<sub>33</sub>G<sub>34</sub>U<sub>35</sub>N<sub>36</sub> where N represents adenine, cytosine, guanine, or uracil. This sequence is found in tRNAs specific for the amino acids histidine, tyrosine, aspartate,

and asparagine. Recognition of this motif is achieved by interactions to Arg286 and Arg289 [Xie *et al.*, 2003].



**Figure 1.3** Biosynthesis of Q-34-tRNA out of GTP and involved enzymes according to Biela *et al.*, 2013.

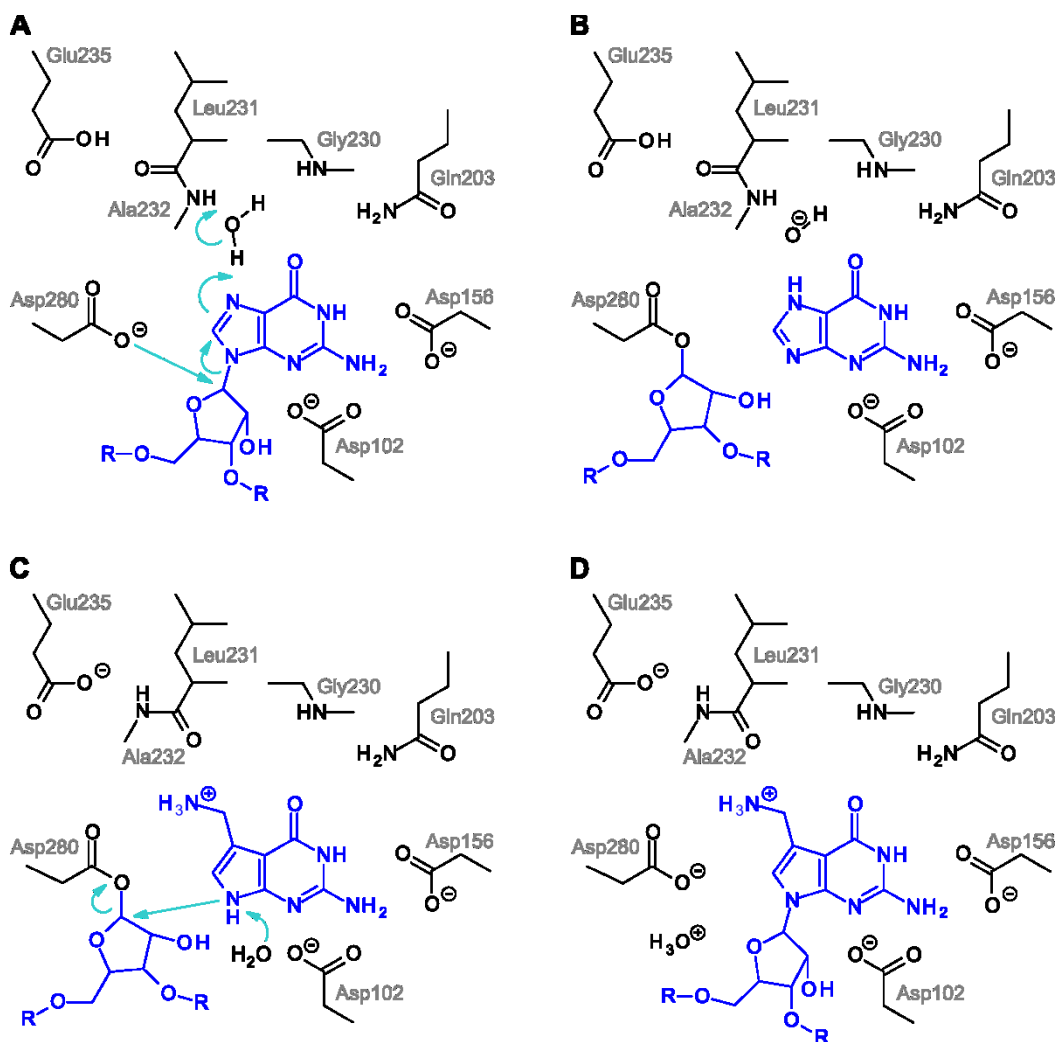
For this purpose preQ<sub>1</sub> is obtained in a multistep reaction out of guanosine 5'-triphosphate (GTP). Its synthesis involves diverse enzymes, namely the GTP cyclohydrolase I (FoIE) [Phillips *et al.*, 2008], 6-carboxy-5,6,7,8-tetrahydropterin synthase (QueD) [McCarty *et al.*, 2009], an *S*-adenosyl-L-methionine-dependent organic radical-generating enzyme (QueE), preQ<sub>0</sub> synthetase (QueC) [McCarty *et al.*, 2009] and a nitrile reductase (QueF) [Lee *et al.*, 2007]. After preQ<sub>1</sub> is incorporated it is further modified to queuine in two following reactions: Firstly, a ribosyl moiety is added to preQ<sub>1</sub> by the *S*-adenosylmethionine:tRNA ribosyltransferase/isomerase (QueA) [Van Lanen *et al.*, 2003]. Secondly, the generated epoxide is reduced by the coenzyme B<sub>12</sub>-dependent epoxyqueuosine reductase (QueG) [Frey *et al.*, 1988].

Also in eukaryotes, tRNAs harboring the above-mentioned recognition sequence are modified in position 34. Contrary to bacteria, eukaryotes do not have the ability to

synthesize queuine or its precursor molecule *de novo*. Therefore, its incorporation mechanism differs significantly. After queuine is assimilated as a nutrient from diet, it is directly inserted into tRNA within a single reaction performed by the eukaryotic TGT [Okada and Nishimura, 1979; Chen *et al.*, 2011]. It is known that queuosine is further transformed by glycosylation of the hydroxyl groups of the pentenyl moiety but the related enzymes of these reactions are not yet sufficiently explored [Iwata-Reuyl, 2003].

### 1.8 Base Exchange Mechanism of TGT

The base exchange of guanine in position 34 of the tRNA against the modified base preQ<sub>1</sub> follows a *ping-pong* mechanism (Figure 1.4) [Xie *et al.*, 2003].



**Figure 1.4** Base exchange mechanism of G-34 against preQ<sub>1</sub> catalyzed by TGT according to Biela *et al.*, 2013.

In the first step of the reaction Asp280 attacks nucleophilically C(1) of ribose-34 after tRNA<sup>His,Tyr,Asp,Asn</sup> is bound to the enzyme (Figure 1.4A). Thereby, a covalent TGT-tRNA intermediate is formed (Figure 1.4B). The cleaved guanine molecule is released from the binding site after picking up a proton. The origin of this proton is not finally clear. While Xie *et al.* discuss Asp102 to release the proton, recent considerations of Biela *et al.* suggest a close-by water molecule as the likely proton donor. As an argument for the latter serves the irreversibility of the base exchange: By-products of the above-described reaction are a hydroxide ion (Figure 1.4B) and an oxonium ion (Figure 1.4D). Mutual neutralization detracts both ions from equilibrium and thus, prevents a reaction in the reversionary direction.

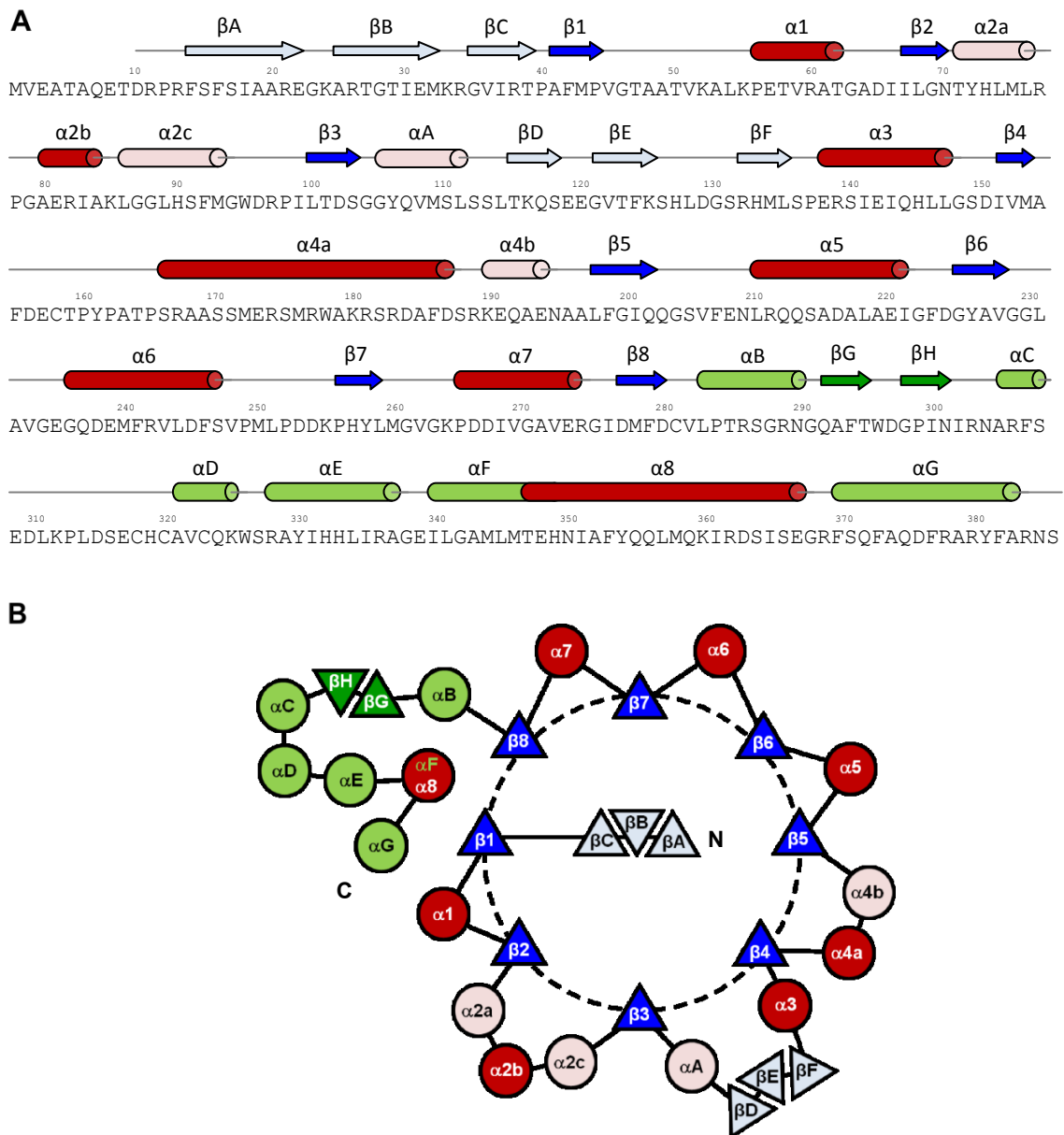
Subsequently, preQ<sub>1</sub> enters the empty binding site (Figure 1.4C). Interestingly, this step induces a backbone flip of the carboxamide group between Leu231 and Ala232 changing its donor functionality to that of an acceptor. This orientation of the backbone is stabilized by Glu235, which is able to interact with both backbone orientations depending on its protonation state [Brenk *et al.*, 2003]. The exocyclic amino function of preQ<sub>1</sub> displaces the previously described water molecule within the recognition site and experiences a hydrogen bond to the backbone C=O function of Leu231, which in its new position faces the binding pocket. Finally, the distance between the covalently bound ribose-34 located within the TGT-tRNA complex and preQ<sub>1</sub> is lowered due to repulsive interactions between Val45, Leu68 and the ribose moiety enabling the nucleophilic attack of preQ<sub>1</sub> to C(1) of ribose-34. Most likely, the released proton originating from N(9) of guanine's imidazole moiety is accepted by a neighbouring water molecule. The covalent bond to TGT is subsequently broken and the modified tRNA leaves the catalytic site (Figure 1.4D).

## 1.9 Structure of *Z. mobilis* TGT

### 1.9.1 Secondary Structure

The crystal structure of *Z. mobilis* TGT exhibits the typical sequence of  $\alpha$ -helices and  $\beta$ -sheets classifying it as member of the family of the  $(\beta/\alpha)_8$ -barrel enzymes, also known as TIM-barrels (Figure 1.5A and 1.6) [Brenk *et al.*, 2003]. A prominent member of this family is triose-phosphate isomerase for which this fold was first discovered.





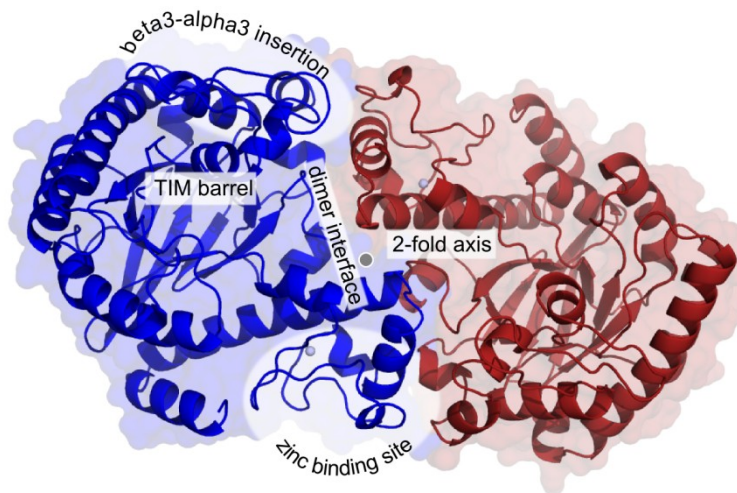
**Figure 1.5** TGT's primary and secondary structure (A) as well as a topology diagram of secondary structure elements (B). The typical TIM-barrel structure elements are colored in blue ( $\beta$ -sheets) and red ( $\alpha$ -helices).

Contrary to the commonly found folding pattern, various deviations in the sequence of secondary structure elements can be observed for *Z. mobilis* TGT (Figure 1.5B) [Brenk *et al.*, 2003; Immekus *et al.*, 2012]. These comprise an antiparallel three-stranded  $\beta$ -sheet located at the N-terminus of the enzyme, a beta3-alpha3 insertion as well as a subdomain offering a binding site for a zinc ion by tetrahedral coordination to the amino acids Cys318, Cys320, Cys323, and His349. Thereby, the zinc binding site involves helix  $\alpha$ 8 of the TIM-barrel. Both, the beta3-alpha3 insertion and the zinc binding domain are involved in the

recognition of tRNA<sup>His,Tyr,Asp,Asn</sup>. Helix  $\alpha$ B acts as a linker connecting the TIM-barrel domain with the zinc binding domain. The active site of *Z. mobilis* TGT is located at the C-terminus of the barrel motif [Okada and Nishimura, 1979; Curnow and Garcia, 1995].

### 1.9.2 Dimeric Assembly

From detailed studies it is known that TGT is only active as a homodimer in which the two monomers are related by a 2-fold symmetry (Figure 1.6) [Xie *et al.*, 2003; Stengl *et al.*, 2007].



**Figure 1.6** Secondary structure and dimeric assembly shown for TGT apo structure 1PUD [Romier *et al.*, 1996]. The structure contains one monomer in the asymmetric unit. Thus, the crystal mate was generated using the program PyMOL.

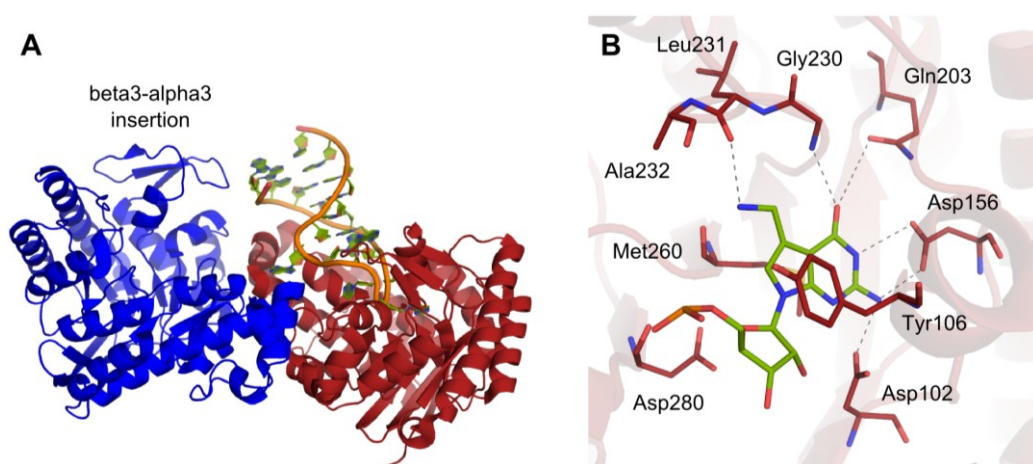
Upon dimerization an interface with the size of  $1667 \text{ \AA}^2$  is formed [Ritschel *et al.*, 2009]. Contrary to apo structures and TGT-ligand complexes where the complementary monomer has to be generated by symmetry operations, the deposited structures of TGT-tRNA complexes (PDB ID: 1Q2R, 1Q2S [Xie *et al.*, 2003]) exhibit the dimeric state in the asymmetric unit. The homodimer is stabilized by diverse hydrogen bonds, salt bridges and van der Waals interactions occurring twice in the homodimer due to the observed C2 symmetry. Interestingly, the amino acid residues involved in dimer formation are highly conserved among various prokaryotic organisms [Stengl *et al.*, 2007]. While initially the focus was set on the strong polar interactions Lys52 ... Glu339', Lys55 ... Glu348', and

Lys85 ... Glu309', recent studies reveal the aromatic cluster Trp326, Tyr330, Phe92', and His333 to be the key motif for dimer stabilization [Ritschel *et al.*, 2009; Jakobi, 2013].

Based on crystal structure analysis as well as non-denaturing mass spectrometry experiments a 1:1 binding stoichiometry between the tRNA and the TGT homodimer is suggested (Figure 1.7A) [Xie *et al.*, 2003; Ritschel *et al.*, 2009]. While the base exchange of guanine against the modified base preQ<sub>1</sub> is performed at the active site of the first monomer, the second monomer stabilizes the ternary complex most likely via its beta3-alpha3 insertion. In this part of the protein several positively charged amino acid residues are present at the protein surface offering the ability to form strong polar interactions to the negatively charged phosphate backbone of the tRNA.

### 1.9.3 Active Site

TGT specifically recognizes the tRNA nucleobases uracil-34, guanine-34, and uracil-35 [Okada and Nishimura, 1979; Curnow and Garcia, 1995]. While both uracil nucleobases bind to rather undefined, largely solvent-exposed pockets, guanine-34 is buried in a tunnel-shaped binding site offering two openings of which one is closed upon substrate binding. Thereby, guanine-34 is sandwiched between Tyr106 and Met260 forming  $\pi$ -stacking interactions to both residues (Figure 1.7B).



**Figure 1.7** Crystal structure of the 1:1 ternary complex of TGT-tRNA (PDB ID: 1Q2S [Xie, Liu *et al.*, 2003]). The crystal structure contains two TGT monomers in the asymmetric unit. A) While the base exchange reaction is performed at the guanine-34 binding site of the first monomer (red), the second monomer (blue) stabilizes the complex via its beta7-alpha6 insertion. B) Binding mode of preQ<sub>1</sub> within the guanine-34 recognition site.

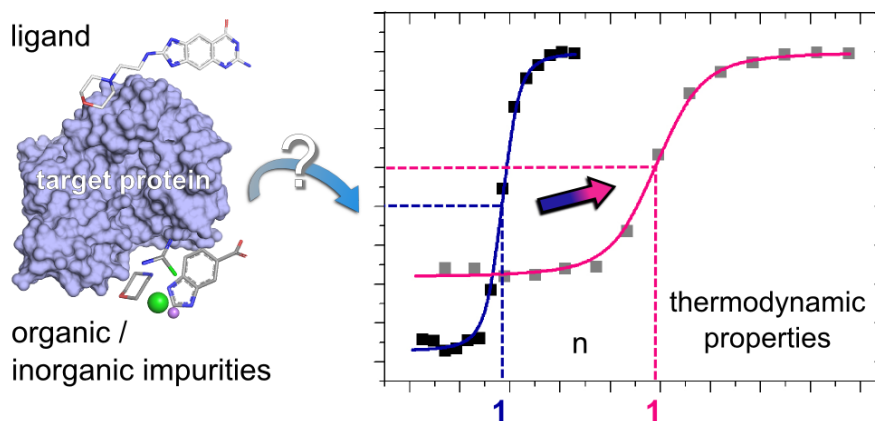
Furthermore, the substrate is recognized by hydrogen bonds to Asp102, Asp156, Gln203, and a water-mediated one to Leu231. Except of Tyr106, which is substituted by phenylalanine in TGT of *S. flexneri*, all residues are identical between the TGTs of *S. flexneri* and *Z. mobilis* [Grädler *et al.*, 2001].

Depending on the pH value and the ligand bound in the active site, its constitution alters: Firstly, the carboxamide group located between Leu231 and Ala232 is able to flip its donor/acceptor functionality. Secondly, Asp102 resting in the apo state within the protein rotates towards the binding site once a substrate is bound [Brenk *et al.*, 2003].

## 2. Impact of Protein and Ligand Impurities on ITC-derived Protein-Ligand Thermodynamics

### 2.1 Introductory Remarks

The following chapter has been published in the scientific journal *Biochimica et Biophysica Acta – General Subjects*.<sup>1</sup> TGT compounds were synthesized by Luzi Jakob Barandun and Christoph Hohn (ETH Zürich). Studies on trypsin were performed by Stefan Grüner (Philipps-Universität Marburg). Trypsin ligands were prepared by Frank Sielaff (Philipps-Universität Marburg). All studies on TGT binding were executed by the author of this thesis. The paper was drafted by Grüner and Neeb.



### 2.2 Abstract

The thermodynamic characterization of protein-ligand interactions by isothermal titration calorimetry is a powerful tool in drug design, giving valuable insight into the interaction driving forces. Isothermal titration calorimetry is thought to require protein and ligand solutions of high quality, meaning both the absence of contaminants as well as accurately determined concentrations. Ligands synthesized to deviating purity and protein of different pureness were titrated by isothermal titration calorimetry. Data curation was attempted also considering information from analytical techniques to correct stoichiometry. We used

<sup>1</sup> Grüner, Neeb *et al.* (2014). Impact of Protein and Ligand Impurities on ITC-derived Protein-Ligand Thermodynamics. *Biochimica et Biophysica Acta* 1840, 2843-2850.

trypsin and tRNA-guanine transglycosylase (TGT), a tRNA modifying enzyme, together with high affinity ligands to investigate both the effect of errors in protein concentration as well as the impact of ligand impurities on the apparent ligand binding thermodynamics. We found that errors in protein concentration of up to 33% did not change the thermodynamic properties obtained significantly. However, most ligand impurities led to pronounced changes in binding enthalpy. In cases where protein binding of the respective impurity would not have been expected, the actual ligand concentration was corrected for and the thus revised data compared to thermodynamic properties obtained with the respective pure ligand. Even in these cases, we observed differences in binding enthalpy of about  $4 \text{ kJ} \cdot \text{mol}^{-1}$ , which is considered as significant. In summary, our results indicate that the ligand purity is the critical parameter to monitor if accurate thermodynamic data of a high affinity protein-ligand complex are to be recorded. Furthermore, artificially changing fitting parameters to obtain a sound interaction stoichiometry in the presence of uncharacterized ligand impurities may lead to thermodynamic parameters significantly deviating from the accurate thermodynamic signature.

### 2.3 Introduction

Thermodynamic characterization of protein-ligand interactions by isothermal titration calorimetry (ITC) has become a routinely used method in understanding interactions of biomolecules with naturally occurring binding partners as well as a powerful tool in drug design [Freyer and Lewis, 2008]. The obtained thermodynamic data of an interaction are most valuable to complement structural information resulting from X-ray crystallography or NMR spectroscopy methods in order to rationally improve a lead compound in structure-based drug design [Chaires, 2008]. However, ITC experiments usually require substantial amounts of the interacting partners, both being of high purity. Successful production of high amounts and purity might be possible with well-selected model systems for which an efficient expression and purification protocol is available and ligands result from simple synthesis. This becomes, however, increasingly difficult if real drug targets and ligands from multistep synthesis are considered. To investigate the impact of impurities present in both the protein and ligand solution focusing on relevant drug discovery cases, we characterized binding of several competitive inhibitors to *B. taurus* trypsin and *Z. mobilis* tRNA-guanine

transglycosylase (TGT) by ITC, taking into account different amounts of protein impurities and using ligand preparations containing organic and inorganic impurities, respectively. Trypsin is a well-studied serine protease involved in digestive processes and serves frequently as a surrogate for actual drug targets in the design of protease inhibitors or to learn about specificity and selectivity discrimination [Rauh *et al.*, 2002; Rauh *et al.*, 2003; Rauh *et al.*, 2004]. However, being an extracellular mammalian protein, it is difficult to obtain in high amounts in its active form by heterologous production in simple expression systems such as *E. coli*. Similar issues arise for other proteins relevant to drug development such as thrombin [Yonemura *et al.*, 2004]. For some of these proteins, natural sources are abundantly available but material might be contaminated with impurities and is usually supplied in lyophilized form which is known to contain hydrate water [Wang, 2000]. Protein concentration is then routinely determined by spectrophotometric measurements, which are often calibrated via an easily calculated extinction coefficient [Gill and Hippel, 1989]. However, the provided extinction coefficient may differ from the actual one, thereby falsifying the protein concentration systematically. This difference is especially pronounced for trypsin's precursor trypsinogen and amounts to 11% [Gill and Hippel, 1989]. Therefore, trypsin was an ideal candidate to investigate how errors in protein concentration and possible impurities affect the thermodynamic characterization of ligand binding determined by ITC.

The tRNA modifying enzyme tRNA-guanine transglycosylase is known to play a key role in the pathogenicity of *Shigella*, the causative agent of Shigellosis [Sansonetti, 2001]. It catalyzes a base exchange of guanine by a modified base in position 34 of the tRNA-anticodon loop [Xie *et al.*, 2003]. This exchange is essential to produce virulence factors necessary for cell invasion. Thus, blocking the enzymatic activity of TGT prevents invasion of *Shigella* [Durand *et al.*, 2000]. The expression protocol in *E. coli* is well established and yields due to favorable solubility features protein of high purity as confirmed by mass spectrometry, SDS page, and dynamic light scattering measurements [Reuter and Ficner, 1995; Romier *et al.*, 1996]. However, the highly potent 6-amino-1,7-dihydro-8*H*-imidazo[4,5-*g*]quinazolin-8-ones (*lin*-benzoguanines) exhibit unfavorable solubility and polarity issues complicating the synthesis of compounds with high purity [Hörtner *et al.*, 2007]. Hence, TGT was selected as a second real case example to study the influence of organic and inorganic impurities in ligand preparations on ITC measurements.

## 2.4 Material and Methods

**2.4.1 Protein Preparation and Quantification.** Bovine pancreas trypsin from natural source was purchased as ethanol precipitate from Sigma (product number: T8003). Dry weight determination was performed in duplicate to estimate the amount of associated volatile substances, i.e. mainly hydrate water. The protein precipitate was dried at 378 K at normal atmospheric pressure and recurrently weighted after cooling in a desiccator until stable weight was reached [Kupke and Dorrier, 1978]. Protein concentration was determined by two methods. Firstly, the micro-biuret method was used, measuring absorbance at 545 nm after trichloroacetic acid/desoxycholate precipitation and using bovine serum albumin as reference protein [Goa, 1953; Bensadoun and Weinstein, 1976]. Secondly, UV spectrophotometry at 280 nm ( $A_{280\text{nm}}$ ) was employed, using an experimental absorbance coefficient of  $1.54 \text{ mg} \cdot \text{mL}^{-1}$  at 280 nm and a molecular weight of 23305 Da as calculated by ProtParam for cationic bovine trypsin (UniProt-ID: P00760), giving  $\epsilon_{280} = 35890 \text{ M}^{-1} \cdot \text{cm}^{-1}$  [Robinson *et al.*, 1971; Koeppe and Stroud, 1976; Gasteiger *et al.*, 2005]. The extinction coefficient of trypsin at 280 nm was  $\epsilon_{280} = 37650 \text{ M}^{-1} \cdot \text{cm}^{-1}$  as directly calculated by ProtParam. Percentage errors in protein concentration refer to the deviation from the concentration determined by the micro-biuret method.

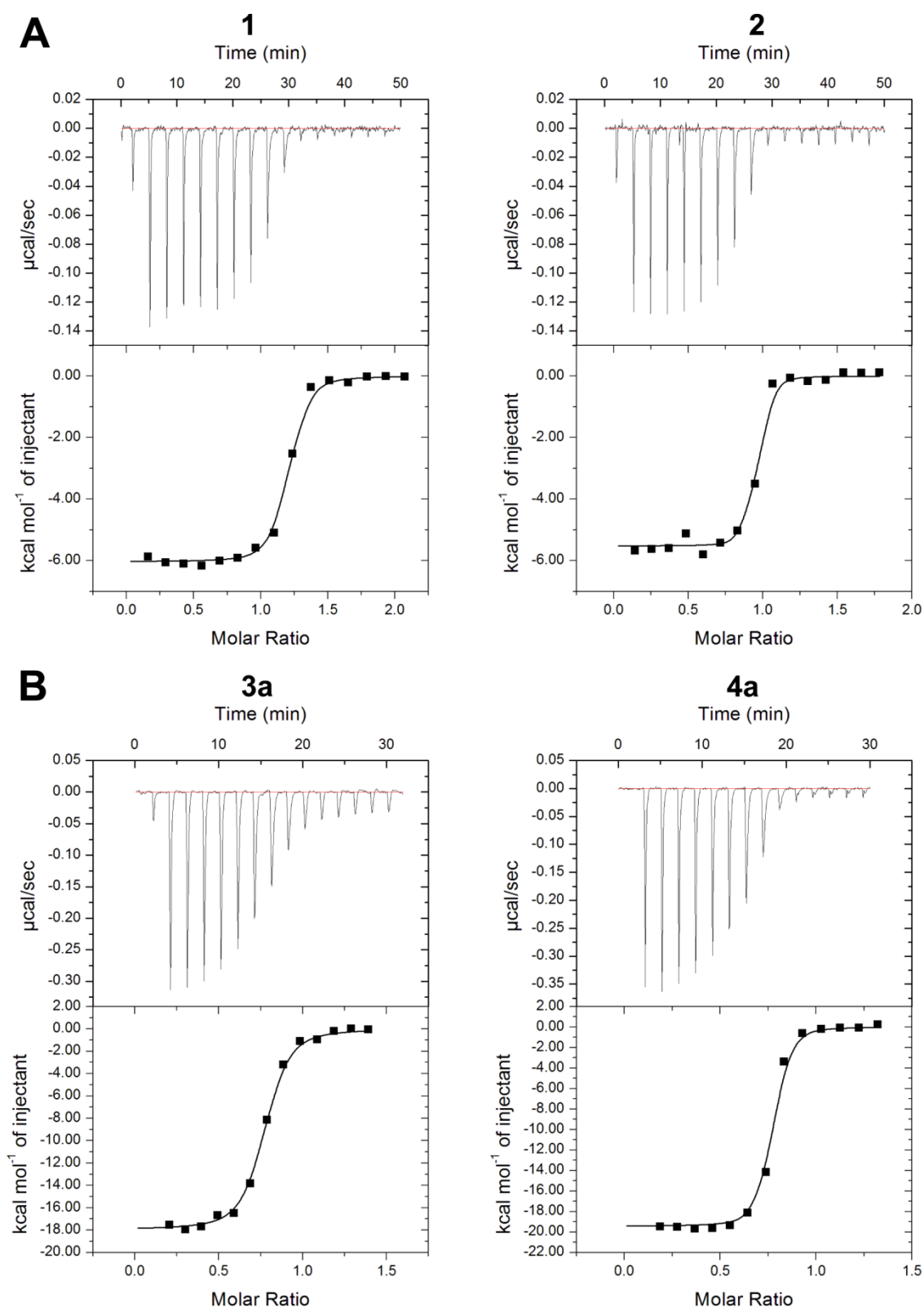
*Z. mobilis* TGT was expressed and purified as described in detail elsewhere [Reuter and Ficner, 1995; Romier *et al.*, 1996]. The protein concentration was determined firstly by UV spectrophotometry at 280 nm ( $A_{280\text{nm}}$ ). An absorption of 0.778 (10 mm path) corresponds to  $1 \text{ mg} \cdot \text{mL}^{-1}$  (23.4  $\mu\text{M}$ ) as suggested by ProtParam [Gasteiger *et al.*, 2005]. Secondly, a Bradford assay was applied on the basis of protein-dye binding using bovine serum albumin as a standard [Bradford, 1976]. Both methods resulted in a closely similar protein concentration.

**2.4.2 Isothermal Titration Calorimetry.** ITC experiments were performed using an iTC200 microcalorimeter (GE Healthcare Europe GmbH, Freiburg, Germany). All experiments were performed at 25°C using filtered solutions only. The reference cell contained demineralized water. The trypsin precipitate was dissolved in ITC buffer (50 mM Tris/HCl, 100 mM NaCl, 0.1% (w/v) polyethylene glycol 8000, pH 7.8) supplemented with 3% (v/v) DMSO and stored on ice. The protein solution was freshly prepared daily. No measurable trypsin autodigestion



occurred as judged by the interaction stoichiometry remaining stable for at least 12 hours under the applied protein concentrations. The final protein concentration in the sample cell was 15  $\mu\text{M}$  based on weight, 12.3  $\mu\text{M}$  based on dry weight measurements and 11.3  $\mu\text{M}$  based on micro-biuret and spectrophotometric measurements. Ligand stock solutions of **1** and **2** of 10 mM were prepared by weight in 100% (v/v) DMSO, subsequently diluted to concentrations ranging from 250 to 400  $\mu\text{M}$  with ITC buffer and adjusted to 3% (v/v) DMSO prior to the experiment. TGT was dissolved in the experimental buffer containing 50 mM HEPES, 200 mM NaCl, and 0.037% (v/v) Tween 20, pH 7.8 to a final concentration of 10  $\mu\text{M}$  containing 3% (v/v) DMSO. Due to their low solubility, ligands **3a** and **4a** were first dissolved in 100% (v/v) DMSO and diluted with buffer solution to a final DMSO concentration of 3%. Ligand concentration in the syringe was adjusted to 200 – 245  $\mu\text{M}$  with experimental buffer. The ligand solution containing syringe was stirred at 1000 rpm and injection started after a stable baseline had been established. A first injection of 0.3  $\mu\text{L}$  was followed by 15 injections of 1.1 to 2.2  $\mu\text{L}$ . All injections were spaced by 120 s to 180 s intervals. Raw data were collected as released heat per time. The collected data were analyzed using ORIGIN Software 7.0 (Microcal Inc.). The area under each peak was integrated, followed by correction for heats of dilution and mixing by subtracting the final baseline, which consisted of small peaks of comparable size. The data point resulting from the first injection was deleted after integration as it is error prone [Mizoue and Tellinghuisen, 2004]. A single-site-binding isotherm that yields the enthalpy of binding  $\Delta H^0$  and the dissociation constant  $K_d$  was fitted to the data [Wiseman *et al.*, 1989]. Each experiment was performed at least in triplicate. For representative curves see Figure 2.1.

**2.4.3 Dynamic Light Scattering.** Potentially aggregation behavior of *Z. mobilis* TGT in the absence and presence of an inhibitor was determined by Dynamic Light Scattering (DLS) using a Zetasizer Nano ZS (Malvern Instruments, Herrenberg, Germany) equipped with a 10 mW HeNe laser at a wavelength of 633 nm at 25°C. Scattered light was detected at a 173° angle with laser attenuation and measurement position adjusted automatically by the software. The same conditions as in ITC experiments were used (10  $\mu\text{M}$  protein in ITC buffer, 3% DMSO, addition of 28  $\mu\text{M}$  (= 20 injections of 2  $\mu\text{L}$ ) of ligand **4a**). The given data comprised three measurements of at least 10 runs.

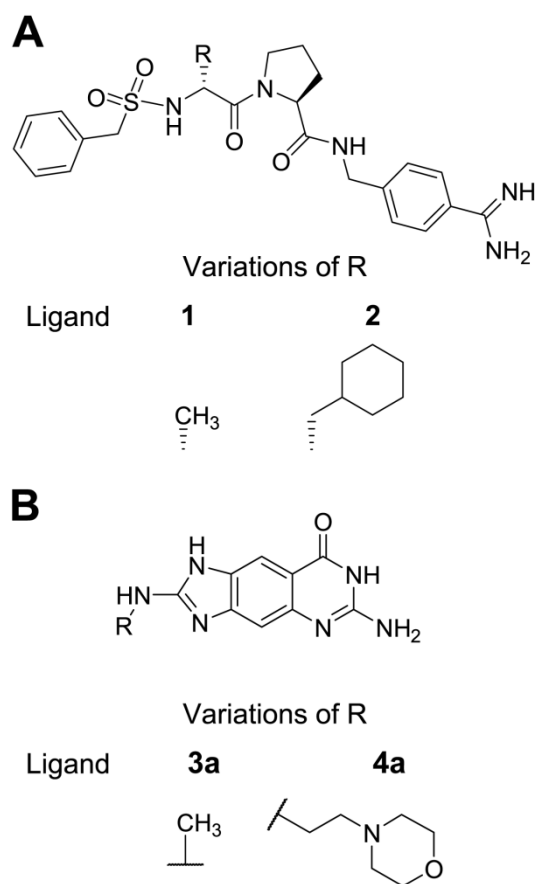


**Figure 2.1** Representative ITC thermograms for **1** and **2** directly titrated to trypsin (A) as well as **3a** and **4a** directly titrated to TGT (B). Shown are titrations of impure preparations, which were partially re-fitted for ligands **2**, **3a** and **4a** to a more reasonable stoichiometry according to data from pure ligand **2** and to elementary analysis for **3a** and **4a**, respectively.

## 2.5 Results

### 2.5.1 Errors in Protein Concentration Did Not Affect the Thermodynamic Profiles of High Affinity Protein-Ligand Binding

To investigate the impact of impurities present in the protein preparation on the thermodynamic characterization of a protein-ligand interaction, binding of two competitive inhibitors (Figure 2.2A) to trypsin was studied by ITC.



**Figure 2.2** Chemical structures of the studied ligands. A) Ligands binding to trypsin. B) Ligands binding to TGT.

First, varying degrees of impurities present in the trypsin preparation were taken into account during modeling of thermodynamic data to study protein impurity-related deviations. The amount of hydrate water present in the protein precipitate was estimated by using the dry weight method. It revealed a hydrate water content of 18% (w/w). Therefore, the actual protein concentration would have been overestimated by 22%, if it was solely determined by weighing. However, all thermodynamic parameters only showed subtle, if

any, changes on considering hydrate water content in data modeling (Table 2.1). The only fitting parameter that changed was the interaction stoichiometry  $n$ . It increased to values closer to 1, the expected value for the interactions investigated. This supports the idea that considering the water content reveals the more correct model for the interaction, even in the absence of other changes in thermodynamic parameters. Hence, ignoring the water content of the used solid protein preparation did not influence the thermodynamic characterization of ligand binding. Next, the actual protein concentration was determined by a colorimetric and a spectrophotometric method, the micro-biuret and  $A_{280\text{nm}}$ -assay, respectively. Both techniques yielded very similar values, if the experimentally determined extinction coefficient of trypsin was used for the  $A_{280\text{nm}}$ -assay. If the determination of protein concentration was based on weight, it deviated from the actual protein concentration by 33%. The difference between the calculated and previously experimentally determined extinction coefficient of trypsin resembled the already known difference for trypsinogen [Gill and Hippel, 1989]. The difference in protein concentration to dry weight determination is probably due to low molecular weight substances such as salts. Nucleic acids were not present in the precipitate. Despite this rather large deviation of 33%, modeling based on the actual protein content also did not significantly change the values of thermodynamic parameters obtained by ITC experiments but interaction stoichiometries approximated much closer to 1, confirming that a better model was fitted (Table 2.1).

### 2.5.2 Organic Ligand Contaminations Can Affect Thermodynamic Profiles of Protein-Ligand Binding Significantly

In order to study how impurities in ligand preparations impact the obtained thermodynamic parameters, two preparations of each trypsin-binding ligand **1** and **2** of different purities were used, referred to as the “impure” and “pure” preparation, respectively. The impure preparation of ligand **1** contained a by-product of its synthesis with similar retention behavior in HPLC, constituting about 20% of the ligand preparation based on HPLC-coupled UV-spectrophotometry. Thermodynamic characterization revealed a high-affinity interaction with an interaction stoichiometry larger than one, indicating that less ligand than assumed was present (Table 2.2, Figure 2.3). ITC experiments using a purified ligand preparation resulted in a slightly improved standard free binding energy as well as standard binding

enthalpy ( $\Delta\Delta G^0 = -1.9 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $\Delta\Delta H^0 = -2.3 \text{ kJ} \cdot \text{mol}^{-1}$ ) and an interaction stoichiometry  $n$  of reasonable value (1.0).

**Table 2.1** Thermodynamic parameters of ligand binding as determined by ITC and modeling with different protein impurity levels.

Ligand	Correction <sup>[a]</sup>	$n^{[b]}$	$\Delta G^0(\text{kJ} \cdot \text{mol}^{-1})^{[b]}$	$\Delta H^0(\text{kJ} \cdot \text{mol}^{-1})^{[b]}$	$-T\Delta S^0(\text{kJ} \cdot \text{mol}^{-1})^{[b]}$
<b>1</b>	none	$0.75 \pm 0.02$	$-45.9 \pm 1.5$	$-27.5 \pm 1.1$	$-18.4 \pm 1.8$
	hydrate	$0.92 \pm 0.02$	$-45.9 \pm 1.6$	$-27.5 \pm 1.1$	$-18.4 \pm 1.9$
	protein	$1.00 \pm 0.02$	$-46.0 \pm 1.6$	$-27.5 \pm 1.2$	$-18.4 \pm 2.0$
<b>2</b>	none	$0.68 \pm 0.01$	$-46.3 \pm 0.9$	$-20.4 \pm 0.7$	$-26.0 \pm 1.1$
	hydrate	$0.83 \pm 0.01$	$-46.2 \pm 0.7$	$-20.4 \pm 0.7$	$-25.8 \pm 1.0$
	protein	$0.90 \pm 0.01$	$-46.1 \pm 0.7$	$-20.4 \pm 0.7$	$-25.7 \pm 1.0$

<sup>a</sup>Correction states the method by which the protein concentration used for thermogram modeling was determined: none – concentration as calculated by protein weight (33% deviation); hydrate – hydrate water content was considered (22% deviation); protein – concentration as determined by micro-biuret/A<sub>280nm</sub>. Deviations are relative to concentration measured by micro-biuret assay. <sup>b</sup>Given errors are standard deviations of at least triplicates.

Ligand **2** was contaminated in its impure form with an organic compound. It was mass spectrometrically identified as 1,1',1''-phosphinylidtrispyrrolidin (CAS-no. 6415-07-2), which is a reaction product of benzotriazol-1-yl-oxytrypyrrolidinophosphonium hexafluorophosphate (PyBOP), a peptide coupling agent used during synthesis. Characterizing ligand binding with the impure preparation showed a high-affinity interaction with trypsin ( $K_d = 11 \text{ nM}$ , Table 2.2), which was apparently dominated by its entropic component, contributing almost 70% to the overall standard free binding energy. However, an abnormally high interaction stoichiometry of about 1.5 indicated considerable ligand purity issues. Further purification of the ligand and additional ITC experiments showed that the actual affinity of the pure ligand was slightly stronger ( $K_d = 8 \text{ nM}$ ). Therefore, the ligand purity did not strongly influence the measured binding affinity. In contrast, the obtained interaction stoichiometry approached values closer to the theoretically expected one (0.9).

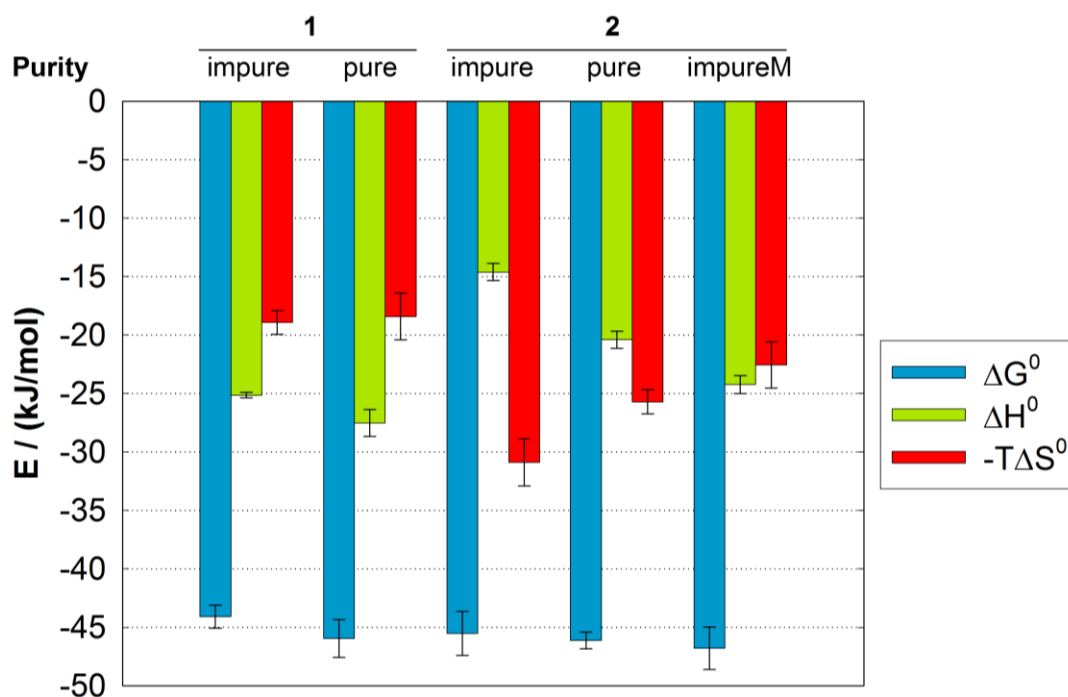
**Table 2.2** Thermodynamic parameters of ligand binding and impact of ligand impurities.

Ligand	Purity	<i>n</i>	$\Delta G^0$ (kJ $\cdot \text{mol}^{-1}$ )	$K_d$ (nM)	$\Delta H^0$ (kJ $\cdot \text{mol}^{-1}$ )	$-T\Delta S^0$ (kJ $\cdot \text{mol}^{-1}$ )
<b>1</b>	pure <sup>[a]</sup>	1.00 ± 0.02	-46.0 ± 1.6	9 ± 6	-27.5 ± 1.2	-18.4 ± 2.0
	impure	1.15 ± 0.03	-44.1 ± 1.0	19 ± 8	-25.2 ± 0.3	-18.9 ± 1.0
<b>2</b>	pure <sup>[a]</sup>	0.90 ± 0.01	-46.1 ± 0.7	8 ± 2	-20.4 ± 0.7	-25.7 ± 1.0
	impure	1.48 ± 0.05	-45.5 ± 1.9	11 ± 8	-14.6 ± 0.7	-30.9 ± 2.0
	impureM <sup>[b]</sup>	0.89 ± 0.02	-46.8 ± 1.8	6 ± 5	-24.2 ± 0.8	-22.6 ± 2.0
<b>3a</b>	pure	0.73 ± 0.05	-41.6 ± 0.1	52 ± 2	-74.8 ± 1.8	33.2 ± 1.8
	Impure	1.16 ± 0.01	-40.3 ± 0.3	86 ± 12	-48.3 ± 0.4	8.0 ± 0.5
	impureM-C <sup>[c]</sup>	0.84 ± 0.00	-41.2 ± 0.2	60 ± 8	-66.1 ± 0.5	24.9 ± 0.6
	impureM-N <sup>[d]</sup>	0.76 ± 0.00	-41.2 ± 0.4	61 ± 9	-74.5 ± 0.6	33.3 ± 0.7
<b>4a</b>	pure	0.68 ± 0.00	-42.2 ± 0.4	41 ± 6	-78.3 ± 1.2	36.1 ± 1.2
	impure	2.23 ± 0.01	-42.4 ± 0.4	38 ± 5	-27.3 ± 0.1	-15.1 ± 0.4
	impureM-C <sup>[c]</sup>	0.73 ± 0.00	-45.6 ± 0.7	11 ± 3	-81.4 ± 0.7	35.8 ± 1.0

<sup>a</sup>Results from pure ligand preparations as in Table 1 shown for ease of comparability. <sup>b</sup>Thermograms resulting from impure preparation were used and fitted to reach the same interaction stoichiometry *n* as thermograms obtained with pure preparation. <sup>c</sup>Thermograms resulting from impure preparation were used and fitted based on the corrected ligand concentration deduced from the carbon value of elementary analysis. <sup>d</sup>Thermograms resulting from impure preparation were used and fitted based on the corrected ligand concentration deduced from the nitrogen value of elementary analysis.

Furthermore, the standard binding enthalpy drastically improved ( $\Delta\Delta H^0 = -5.8 \text{ kJ} \cdot \text{mol}^{-1}$ ). Consequently, the entropic contribution to binding decreased by a similar magnitude ( $-T\Delta\Delta S^0 = 5.2 \text{ kJ} \cdot \text{mol}^{-1}$ ) and afforded then only about 55% of the total standard free binding energy. Re-fitting of the thermograms obtained with the impure ligand was done with a corrected ligand concentration giving similar interaction stoichiometries as the pure ligand (“impureM”). Correcting the ligand concentration yielded a concentration 40% less than originally assumed. The obtained thermodynamic data of these corrected fittings were compared to the data resulting from the pure ligand preparation. Whereas the standard free binding energies were in close agreement ( $\Delta\Delta G^0 = -0.7 \text{ kJ} \cdot \text{mol}^{-1}$ ), the enthalpic and entropic contributions still differed significantly by  $-3.8 \text{ kJ} \cdot \text{mol}^{-1}$  and

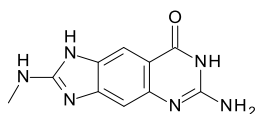
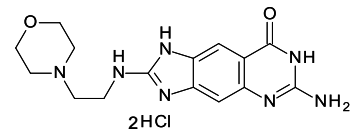
$3.1 \text{ kJ} \cdot \text{mol}^{-1}$ , respectively [Holdgate, 2001]. Hence, the thermodynamic profile of the ligand resulting from the impure, newly modeled preparation did not reasonably agree with the profile obtained from a pure preparation.



**Figure 2.3** Thermodynamic properties for ligands **1** and **2** from impure and pure preparations as determined by ITC. Shown are mean values and standard deviations as error bars of at least triplicates. Purity: impure – ligand preparations containing a contaminant were used for experiments with no correction for actual ligand content; pure – repurified ligand preparations were used; impureM – ligand preparations containing a contaminant were used for experiments and actual ligand content was taken into account for modeling.

Additionally, the impact of an organic impurity on the thermodynamic profile of TGT ligand **3a** (Figure 2.2B) was investigated by analyzing the ligand before and after HPLC purification. Unfortunately, the exact chemical composition of the impurity could not be identified. Elementary analysis revealed that the values for carbon and nitrogen are affected differently (Table 2.3). Considering the determined carbon value, an actual purity of approximately 73% was estimated. In the case of the found value for nitrogen, purity appeared to be further lowered to around 66%. ITC measurements characterized **3a** in its impure form as a potent inhibitor of TGT ( $K_d = 86 \text{ nM}$ ) with a mainly enthalpic contribution to binding (Table 2.2).

**Table 2.3** Elemental analysis of the impure *lin*-benzguanines after precipitation.

Ligand		C:	H:	N:	
<b>3a</b>		calcd:	52.17	4.38	36.50
		found:	37.86	4.48	23.92
<b>4a</b>		calcd: <sup>[a]</sup>	44.78	5.26	24.37
		found:	14.69	2.29	7.78

<sup>a</sup>The calculation assumes **4a** to be present as dihydrochloride.

Accompanied by the described deviations found in elementary analysis, the interaction stoichiometry is increased to a value of 1.16. Re-fitting of the data for the impure ligand following the elementary analysis resulted in small changes of the standard free binding energy ( $\Delta\Delta G^0 = -0.9 \text{ kJ} \cdot \text{mol}^{-1}$ ). However, the thermodynamic profile shows a more pronounced enthalpy term of  $\Delta\Delta H^0 = -17.8$  to  $-26.2 \text{ kJ} \cdot \text{mol}^{-1}$ . The difference between the two newly fitted data sets is significantly increased, especially for  $\Delta H^0$  ( $\Delta\Delta H^0 = 8.4 \text{ kJ} \cdot \text{mol}^{-1}$ ). As a consequence, the entropic term shifts to more unfavorable values. The interaction stoichiometry improved to reasonable values of 0.84 and 0.76, respectively, taking into account that expression and purification did not yield fully active protein. To get a hint about the impurity's composition, the values were compared to those of **3a** in its pure state, after further purification steps via HPLC.  $\Delta G^0$  remained the same within the range of error compared to the re-fitted value ( $\Delta\Delta G^0 = -1.3 \text{ kJ} \cdot \text{mol}^{-1}$ ), similar to refitting data of the impure ligand. The enthalpic term drastically increased compared to the batch of **3a** before HPLC purification ( $\Delta\Delta H^0 = 26.5 \text{ kJ} \cdot \text{mol}^{-1}$ ). Contrary to that, the values are in excellent agreement with the re-fitted values for a purity of approximately 66% ( $\Delta n = 0.03$ ,  $\Delta\Delta G^0 = -0.4 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $\Delta\Delta H^0 = -0.3 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $-T\Delta\Delta S^0 = -0.1 \text{ kJ} \cdot \text{mol}^{-1}$ ).

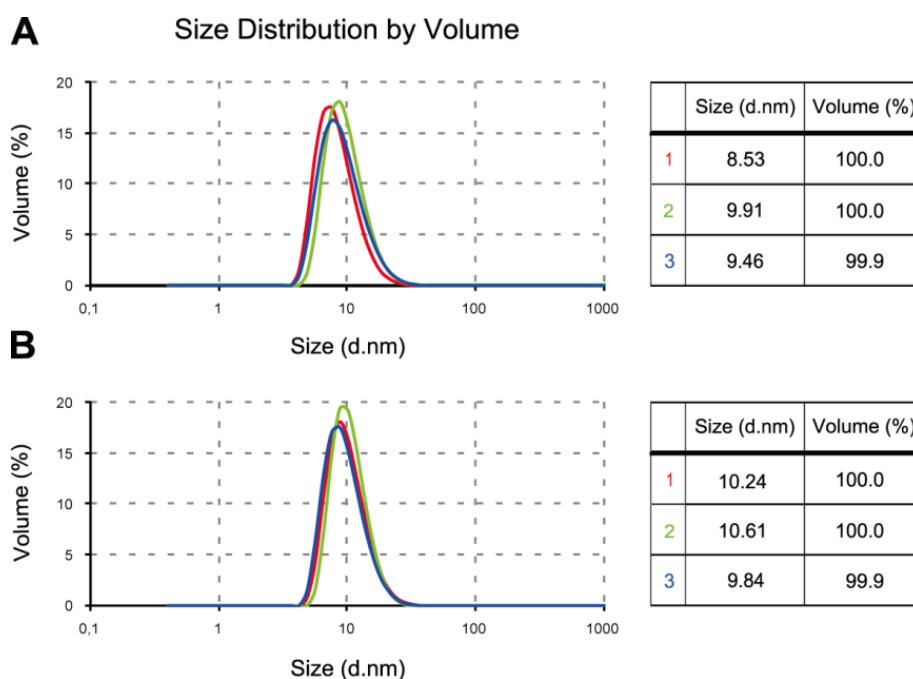
### 2.5.3 Inorganic Ligand Contaminations Can Also Affect Thermodynamic Profiles of Protein-Ligand Binding

A second TGT inhibitor was included in our studies. Compound **4a** shows low purity after synthesis without any further purification steps. According to elementary analysis, the



powder contains the desired compound **4a** to approximately one-third. Contrary to the results of **3a**, the yield resulting from the ratio found to theoretical percentage of carbon and nitrogen, respectively, shows only subtle differences ( $\Delta yield = 0.9\%$ ). Therefore, most likely an inorganic impurity originating from synthesis is present in the ligand preparation. The binding isotherm yielded a binding affinity of  $K_d = 38 \text{ nM}$  for the impure sample of **4a** and  $K_d = 41 \text{ nM}$  for the pure one ( $\Delta\Delta G^0 = 0.2 \text{ kJ} \cdot \text{mol}^{-1}$ ). Again, the largest changes between pure and impure preparations are noticeable in the thermodynamic partitioning: The enthalpic term increases remarkably from  $-27.3 \text{ kJ} \cdot \text{mol}^{-1}$  to  $-78.3 \text{ kJ} \cdot \text{mol}^{-1}$  ( $\Delta\Delta H^0 = -51.0 \text{ kJ} \cdot \text{mol}^{-1}$ ), the entropic contribution decreases from  $-15.1 \text{ kJ} \cdot \text{mol}^{-1}$  to an unfavorable term of  $36.1 \text{ kJ} \cdot \text{mol}^{-1}$  ( $-T\Delta\Delta S^0 = 51.2 \text{ kJ} \cdot \text{mol}^{-1}$ ). In addition, stoichiometry is diminished by a factor of about three from 2.23 to a more reasonable value for this inhibitor of 0.68. In the observed case, re-fitting of data could be performed easier since the results from elementary analysis indicated a purity level of approximately 33% due to an inorganic salt. The obtained values were in good agreement with the values derived from the pure ligand preparation, even if close to the significance level. Stoichiometry differed by 0.05,  $\Delta H^0$  by  $-3.1 \text{ kJ} \cdot \text{mol}^{-1}$  and  $-T\Delta S^0$  by  $-0.3 \text{ kJ} \cdot \text{mol}^{-1}$ . Changes in  $\Delta G^0$  were slightly higher than for the other described compounds **1** – **3a** ( $\Delta\Delta G^0 = -3.4 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $K_d = 41 \text{ nM}$  vs.  $11 \text{ nM}$ ).

It is known that salts can significantly influence the activity and stability of proteins in aqueous solution both in favorable and unfavorable means according to the Hofmeister series [Kunz *et al.*, 2004]. Indeed, TGT needs high salt conditions to be soluble in the long term. To determine possibly occurring protein or ligand aggregation due to an increased salt concentration, dynamic light scattering (DLS) measurements were performed additionally. While inactive aggregates formed by the protein are not expected to change the thermodynamic profile drastically as shown in our studies, aggregates formed by ligands are able to bind to the protein unspecifically leading to false positive signals [McGovern *et al.*, 2002]. Measurements were carried out in the absence and presence of **4a** (Figure 4.4) using the same conditions as in the ITC experiments. Neither the obtained mean peak size ( $\Delta d = 0.9$ ) nor the average peak volume ( $\Delta V = 0.0$ ) deviated significantly. Therefore, the difference in enthalpy of  $\Delta\Delta H^0 = -3.1 \text{ kJ} \cdot \text{mol}^{-1}$  between re-fitted data originating from impure ligand preparation and data from repurified ligand preparation was unexpected.



**Figure 2.4** Influence of **4a** on the size distribution by volume of *Z. mobilis* TGT. In both cases a 10  $\mu$ M protein solution was analyzed. A) A total of three DLS measurements of TGT in the absence of any ligand show a monomodal size distribution with an average diameter of 9.3 nm. The mean peak volume is 99.97%. B) A total of three DLS measurements of TGT in presence of **4a** with a concentration of 28  $\mu$ M (equivalent to 20 injections of 2  $\mu$ l of a 200  $\mu$ M inhibitor solution) show no influence on the monomodal size distribution. The average diameter of the protein is 10.2 nm with a mean peak volume of 99.97%.

## 2.6 Comparative discussion

In biophysical investigations, one is urged to work with protein and in the field of protein-ligand interactions also with ligands of high purity [Daviter and Fronzes, 2013]. ITC is a key technique to thermodynamically characterize chemical and biochemical binding processes. Before analyzing the interaction of interest, particular attention needs to be paid to carefully calibrate the microcalorimeter using well-known standard chemical reactions as well as to the achievable precision and systematic differences reported for various microcalorimeters [Wadsö, 2000; Tellinghuisen, 2007; Baranauskienė *et al.*, 2009] When it comes to analyzing the thermodynamics of protein-ligand interactions, ITC is a central method but frequently proteins and ligands are difficult to prepare at high purity. Therefore, we investigated the effects of errors in protein and ligand concentration and impurities on the obtained thermodynamic property values. We found that, for high-affinity interactions with  $c$  values over 100 [Wiseman *et al.*, 1989], neither considering hydrate water nor further uncertainties

in protein concentration resulted in differences of thermodynamic property values compared to values obtained with the simple protein weight concentration. Hence, ignoring a concentration error of up to 33% and thereby hydrate water and contaminants such as salts in the protein preparation did not affect the apparent thermodynamic data of a protein-ligand interaction. This finding is in agreement with a study of metal ions binding to crown ethers [Turnbull and Daranas, 2003]. Errors in the receptor (crown ether) concentration of up to 15% lead neither to changes in the standard free binding energy nor the binding enthalpy [Turnbull and Daranas, 2003]. The  $c$  value is defined as  $c = n \cdot [Prot]_{total} / K_d$ . Even a more than twice as large error of 33% was found to be acceptable in this study. The tolerance can be rationalized by the idea that for high affinity ligands, the amount of ligand binding to its receptor for each injection is almost solely dependent on the actual amount of ligand injected and therefore not sensitive to changes in receptor concentration and errors in its determination are compensated for during modeling by the stoichiometry parameter  $n$  [Tellinghuisen, 2005; Tellinghuisen and Chodera, 2011]. Changes of the latter mainly lead to shifting the binding isotherm to different interaction stoichiometries accompanied by small changes in  $\Delta G^0$ . Still it is important for experimental accuracy to determine the actual protein concentration as precisely as possible, especially if the interaction stoichiometry is unknown. We showed that routine use of a calculated extinction coefficient might not accurately reflect the actual absorption property of a protein as it is the case for trypsin, hinting at differences in absorption profiles between its native and denatured form. In the absence of a known experimental extinction coefficient, the micro-biuret assay proved to be a valuable method safeguarding against inaccurately quantifying protein concentration.

Considering the ligand, a larger impact of concentration inaccuracies and impurities on thermodynamic properties is to be expected. The fitting process involves normalization of observed heat per injection by the amount of ligand added and the normalized heat is then plotted against the molar ratio of ligand per receptor. Supporting this reasoning, variability in thermodynamic parameters of a benchmarking protein-ligand interaction in an interlaboratory study was found to be mainly caused by varying errors in determining the ligand concentration [Myszka *et al.*, 2003; Tellinghuisen and Chodera, 2011]. Indeed, impure ligand preparations of ligands **1–4a** used in this study all showed a reduced apparent binding

enthalpy compared to binding enthalpy of the pure ligand. Impurities present in a ligand preparation can be thought of as principally belonging to three major classes:

Firstly, side products of synthesis, which contain fragments of the actual ligand and are capable of binding to the same protein, thereby acting as competitive ligand. The contaminant of ligand **1** showed a similar HPLC-retention and UV-absorption behavior as the ligand itself, pointing toward possible structural similarities to the ligand. Therefore, the contaminant itself might also exhibit protein-binding activity. A mixture of ligand and protein-binding by-product is difficult to model in ITC data, as quantification of the ligand content alone is not sufficient for accurate correction. In those cases, ligand repurification is inevitable. Due to the minor impurity content of ligand **1**, the deviations in thermodynamic parameters  $\Delta G^0$  and  $\Delta H^0$  were in the range of  $2 \text{ kJ} \cdot \text{mol}^{-1}$ .

The second class of impurities are organic compounds originating from the synthesis itself, such as coupling agent reaction products, or remaining from so called “leaking” columns used in purification. These compounds are unlikely to exhibit specific protein binding but might otherwise influence the protein-ligand interaction. The contaminants of ligands **2** and **3a** belong to that class. In both cases, an impurity content of up to 40% led to a drastically decreased apparent binding enthalpy whereas  $\Delta G^0$  was much less affected. Thereby, a much more entropically driven ligand binding is suggested. As ligands **1** and **2** belong to the same congeneric ligand series where the latter contains a more hydrophobic substituent in position R, one could have expected a trend toward increased entropic contribution. However, characterization using the pure ligand preparation showed a less pronounced trend of such an entropic contribution. This case underlines the importance of considering the observed binding stoichiometry as a parameter for quality control. An interesting question arose after the pure ligand **2** was characterized: Are thermodynamics profiles of impure and pure ligand preparations of **2** comparable if corrected for the actual ligand concentration by fitting to the same interaction stoichiometry? Identical values indicate that solely adjusting ligand concentration during modeling might be sufficient. However, a difference in standard binding enthalpy of  $3.8 \text{ kJ} \cdot \text{mol}^{-1}$  remained between re-fitted data and data of pure ligand, pointing toward the contaminant influencing the protein-ligand interaction. The underlying mechanism remained unclear. The magnitude of the observed difference is close to the threshold of  $4 \text{ kJ} \cdot \text{mol}^{-1}$ , above which a difference in binding enthalpy is considered to be significant [Holdgate, 2001].

In contrast, comparison of the re-fitted data of the contaminated ligand **3a** and its pure form showed that it is possible to obtain the actual thermodynamic parameters under consideration of elementary analysis results. The data fitted to a purity of 66%, taking into account the ratio between the found and theoretical nitrogen value, give the same parameters like the pure compound does, within the error limits. However, the fit to a purity of 73%, as suggested by the ratio between found and theoretical values for carbon, failed to yield the same thermodynamic signature as for the pure ligand. Therefore, a prerequisite for the successful re-fitting is the knowledge of the composition of the impurity contained in the ligand preparation. Furthermore, even if no structural similarities are present, one has to reassure that the contamination does not interact with the protein or influence its stability or interaction with a ligand during the measurement period.

Hence, even when correcting for the actual ligand concentration, one has to be aware that the resulting apparent thermodynamics may not accurately reflect the actual thermodynamics. This deviation may go undetected if ligand concentration is adjusted during the fitting process in order to give a reasonable interaction stoichiometry and no revalidation is performed to what extent that adjustment is sufficient. Therefore, this manual adjustment is poor practice.

Thirdly, the ligand preparation may contain a varying proportion of salts resulting from synthesis and/or purification, which may change the thermodynamic properties similarly like an organic impurity. As described for **2** and **3a**, also **4a** shows significant changes in its enthalpic and entropic properties while  $\Delta G^0$  remains largely unaffected. Re-fitting according to elementary analysis yielded improvements only toward the pure compound close to the significance limit. Influences of ions binding to the protein surface might be imaginable. Besides that, the solubility of compound and protein at high concentrations needed for ITC measurements might be negatively affected. The impact of ligand as well as protein aggregation can be easily monitored by Dynamic Light Scattering.

Within the error range of the method, elementary analysis gave the identical yield of **4a** regarding analyzed values for carbon and nitrogen. As a consequence, it can be assumed that an inorganic salt resulting from synthesis as part of the ligand powder. When including the actual yield of **4a** into the fitting procedure of the binding isotherm, the derived data showed a good approximation to the thermodynamic properties obtained with the further purified ligand. Despite this and similar to the modeled data of **1** and **2**, deviations close to

the significance level could be observed ( $\Delta\Delta G^0 = 3.4 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $\Delta\Delta H^0 = -3.1 \text{ kJ} \cdot \text{mol}^{-1}$ ). Also DLS measurements did not give any explanations for those deviations. Protein and ligand aggregation could be excluded according to the conducted experiments. This result was expected since the salt of the ligand powder increases the concentration of the salt contained in the experimental buffer by only 0.1% (w/v calculated as NaCl). However, also if no measurable protein or ligand aggregation could be observed in dynamic light scattering measurements, an influence of increased salt concentrations by other means cannot be totally excluded.

To further verify thermodynamic results, a reverse experimental setup can be devised. In such a reverse titration, protein solution is titrated into the ligand containing cell and is especially useful to check interaction stoichiometry [Velazquez-Campoy and Freire, 2006]. For a simple 1:1 interaction, the results should be invariant to the normal experimental setup but may change if more complex interaction modes are investigated [Liang, 2008; Brown, 2009]. However, protein is required in higher amounts and concentrations in such experiments, making additional validation of the absence of possible protein aggregation necessary. The herein introduced DLS method is ideally suited for this purpose. Unfortunately, this setup is hardly feasible for our model proteins: The self-digestion of trypsin is increased at higher concentrations. Thus, no reliable results can be expected over the period of an ITC experiment. Similarly, TGT's solubility is drastically lowered at low salt concentrations applied during the titration. Therefore, precipitation of the protein occurred during a reversed experimental setup at the needed concentrations.

Additionally to systematic errors discussed above, statistical errors and their correct treatment should also be considered [Tellinghuisen, 2003; Tellinghuisen, 2004].

## 2.7 Conclusions

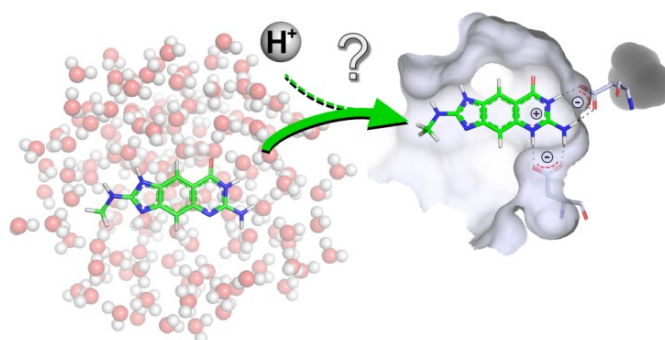
It can be concluded that impurities present in protein preparations used for thermodynamic characterization of a protein-ligand interaction by ITC do not necessarily translate into deviations from the actual thermodynamic data. Under special circumstances, they can even be ignored as it was the case for the protein preparation in the study of high affinity interactions. More caution has to be paid to the preparation of high affinity ligands. Even contaminants without apparent protein-binding capabilities may distort the thermodynamic

properties of a protein-ligand interaction, which goes undetected by sole fitting adjustments. Our studies dealing with high  $c$  value titrations demonstrated that it is possible to derive thermodynamic data from impure ligands, which are in good accordance to data obtained with the pure ligand. The basis for successful fitting is the knowledge about the composition of the contaminant and its potential influence on the analyzed system, which is unfortunately often difficult to estimate. Fitting the binding isotherm of an impure ligand following elementary analysis results yielded a good approximation to the actual value in one case, proving elementary analysis to be a well-suited complementary analysis to NMR and MS methods when it comes to ligand synthesis for ITC purposes. In another case, the same data as with a pure compound were obtained. Nonetheless, this procedure is limited since only a small data range is accepted before deviations get significant. For ligands showing large differences in their thermodynamic profile among each other, this might be negligible. However, the practice of sole fitting adjustments may mislead correlation to structural data of protein-ligand complexes or interpretation of thermodynamic trends observed in a congeneric series of ligands. The latter is especially deleterious as thermodynamic characteristics of such stepwise ligand modifications may not show big differences between the individual components of the series but are frequently used in conjunction with structural data to establish basic mechanisms underlying molecular interactions [Snyder *et al.*, 2011; Biela *et al.*, 2013]. In summary, this study pointed out that thermodynamic characterization of high affinity protein-ligand interactions by ITC shows different robustness against possible experimental errors and highlights the experimental parameters which must be rigorously monitored to obtain accurate and reliable thermodynamic data. If especially ligand purity cannot be assured, the actual value might be at least approximated. However, accurateness of these assumptions and considered values may be difficult to assess. Therefore, the efforts should be rather put in reliable purification protocols, both for proteins as well as ligands, to provide the basis to obtain accurate data by ITC.

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and $pK_a$ Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme

#### 3.1 Introductory Remarks

The following chapter was published in the scientific *Journal of Medicinal Chemistry*.<sup>2</sup>  $pK_a$  Calculations were carried out by Paul Czodrowski (Merck Discovery Technologies, Darmstadt). TGT compounds were synthesized by Luzi Jakob Barandun and Christoph Hohn (ETH Zürich). All crystal structures and ITC measurements were performed by the author of this thesis along with drafting the article.



#### 3.2 Abstract

A drug molecule will exhibit more desirable properties, if it remains uncharged while it travels through the body, is transported across membranes, and adopts a charged state only upon protein target binding as usually higher affinity results when charged species interact with one another. Such a strategy requires careful design of  $pK_a$  properties and methods to elucidate whether and where protonation state changes occur. We investigated the protonation inventory in a series of *lin*-benzoguanines binding to tRNA-guanine transglycosylase, which shows pronounced buffer dependency during ITC measurements. Chemical modifications of the parent ligand scaffold along with ITC measurements,  $pK_a$  calculations, and site-directed mutagenesis allow elucidating the site of protonation. The

<sup>2</sup> Neeb *et al.*, (2014). Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis, and  $pK_a$  Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme. *Journal of Medicinal Chemistry* 57, 5554-5565.



parent scaffold exhibits two guanidine-type portions and both appear as likely candidates for proton uptake. Even mutually compensating effects resulting from proton release of the protein and simultaneous trapping by the ligand can be excluded. A cluster of two adjacent Asp residues induces a strong  $pK_a$  shift at the ligand resulting in a transition to the protonated state. Furthermore, an array of two parallel H-bonds avoiding secondary repulsive effects contributes to the high affinity binding of the *lin*-benzoguanine scaffold.

### 3.3 Introduction

In recent time, isothermal titration calorimetry (ITC) has become a powerful tool in life sciences with a broad scope of application, e.g. to obtain insight into the energetics or kinetics of binding, permeation through lipids as well as the characterization of low or high affinity binding ligands that are not accessible by direct titrations [Zhang and Zhang, 1998; Velazquez-Campoy and Freire, 2006; Tsamaloukas *et al.*, 2007; Burnouf *et al.*, 2012]. In structure-based drug design, this method is of utmost importance to obtain not only information about binding affinities but also to access the entire thermodynamic signature of a ligand within one experiment [Ladbury *et al.*, 2010]. In addition, it gives valuable insight into changes of protonation states upon protein–ligand binding when the measurements are carried out in different buffer systems showing deviating contributions to the heat of ionization  $\Delta H_{ion}$  [Baker and Murphy, 1996].

Numerous studies have underlined the importance to investigate the protonation state of bound ligands to their target enzyme [Dullweber *et al.*, 2001; Raffa *et al.*, 2004; Czodrowski *et al.*, 2007; Steuber *et al.*, 2007; Baum *et al.*, 2009; Biela *et al.*, 2012; Biela *et al.*, 2012]. In fact, the ITC method quantifies the molar amount of protons transferred from or released to the buffer environment. Unfortunately, it does not provide any information about the exact protonation site. Supporting the ITC measurements by  $pK_a$  calculations and mutational studies can help to detect the residue or ligand functional group responsible for the protonation reaction.  $pK_a$  calculations suggest the likely candidate residues of interest and subsequent titrations of the specifically generated mutants can verify the protonation sites experimentally [Czodrowski *et al.*, 2007; Steuber *et al.*, 2007]. In an unfortunate situation no overall change in protonation can be observed, even though different buffers have been used and protons are transferred. Recently, Baum *et al.* observed the

simultaneous superposition of proton uptake and release reaction leading overall to a balanced proton inventory during ITC titrations [Baum *et al.*, 2009]. This study underlines the importance to investigate congeneric ligand series, which differ only by small gradual changes varying the basic or acidic properties, respectively.

On the one hand, understanding the properties of titratable groups that can change their protonation state upon binding, both on the protein and the ligand sites, is an important prerequisite to correctly assign and subsequently interpret the thermodynamic signature of a compound. For instance, the enthalpic signal recorded in a titration comprises the heat contributions of all interactions that are formed between the ligand and the amino acids of the target protein as well as those to trapped or released water molecules or ions. Depending on the type of interaction ranging from simple hydrogen bonds between uncharged species to charge-assisted ones and hydrogen bonds with salt-bridge character, their strength and thus their contribution to the enthalpic signal of the binding event can vary strongly. Taking continuum electrostatics as a reference, this fact can be easily understood. If charge-assisted contacts are formed in an environment of low local dielectric conditions, which are usually given in deeply buried binding pockets of significant hydrophobic character, they will exhibit large contributions. This results from the fact that electrostatic forces, described by Coulomb's law, show the product of the charges in the numerator and, apart for the squared contact distance (if the potential is considered, it is inverse-linear in distance), the dielectricity constant  $\epsilon$  in the denominator. Thus, high dielectricity as given in an aqueous environment ( $\epsilon \approx 80$ ) attenuates charge-assisted contacts strongly. In hydrophobic environment, where  $\epsilon$  is small ( $\epsilon \approx 1 - 4$ ), an increasingly exothermic signal is experienced as there the involved interacting groups are shielded from the solvent environment. The thermodynamic signature of protein–ligand binding can become quite complex. For the interpretation of such data, it is therefore of utmost importance to know whether and where possible changes in protonation states might create charge-assisted contacts. If these contacts are planned and subsequently established in buried binding pockets they can contribute significantly to the exothermic binding signature and thus enhance affinity in a tailored fashion.

On the other hand, proper adjustment of protonation states plays an important role for the permeation and distribution of the potential drug molecule to obtain access to a target cell under consideration. It is well known that permeation depends on multiple

factors such as lipophilicity, hydrogen bonding properties, which are important for the transfer from the water phase into the lipid bilayer, or the shape of a drug molecule and last but not least its charge [Schanker, 1962; Spencer *et al.*, 1979; Conradi *et al.*, 1991; El Tayar *et al.*, 1991; Testa *et al.*, 1996]. Drug molecules often contain functional groups of acidic and basic character and thus, occur dependent on their  $pK_a$  values and the local pH conditions of the surrounding compartment as charged or neutral species in equilibrium. The uncharged form of a drug will permeate to a greater extent through biological membranes than its corresponding charged form. Accordingly, it would be highly desirable that a drug molecule remains uncharged as long as it is transported via membranes and adopts a charged state only upon binding to the target protein as usually higher affinity results when charged species interact with one another (see above).

In the present study we investigated a series of ligands comprising a 6-amino-1,7-dihydro-8*H*-imidazo[4,5-*g*]quinazolin-8-one (*lin*-benzoguanine) and a 1,7-dihydro-8*H*-imidazo[4,5-*g*]quinazolin-8-one (*lin*-benzohypoxanthine) scaffold with respect to their inhibitory potency of the tRNA modifying enzyme tRNA-guanine transglycosylase (TGT) from *Zymomonas mobilis* [Stengl *et al.*, 2007; Barandun *et al.*, 2012]. TGT has been shown to be a putative target to fight the pathogenicity of shigellosis, a severe diarrheal disease [Durand *et al.*, 1997; Grädler *et al.*, 2001]. Previous comparisons of the *lin*-benzoguanines and *lin*-benzohypoxanthines already showed surprising differences concerning their binding modes and their binding affinities [Barandun *et al.*, 2012], which prompted us to take a closer look at the features of the two scaffolds in the TGT-bound state.

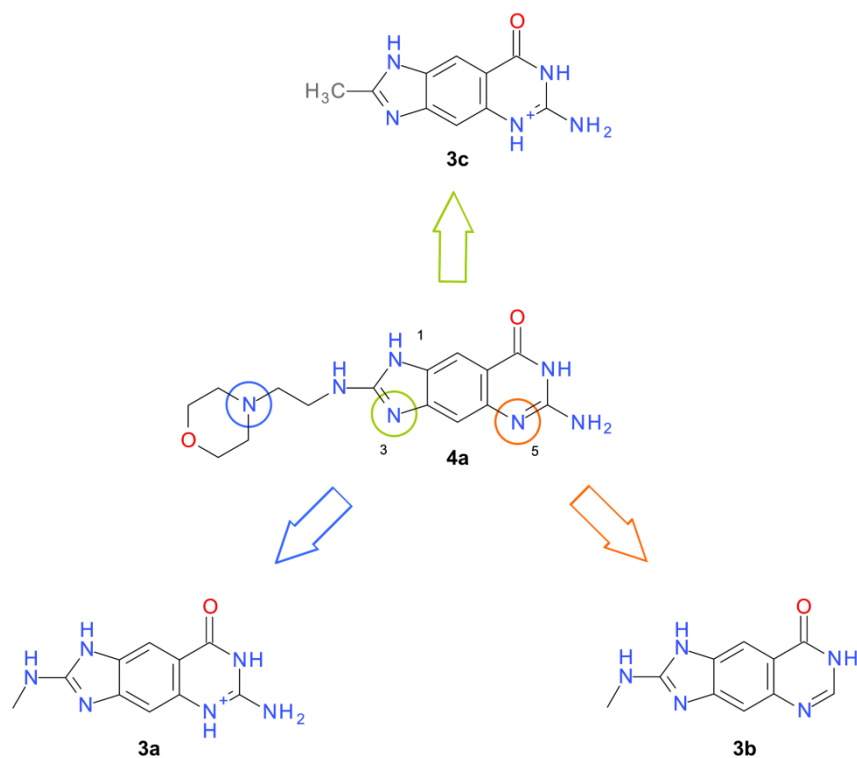
## 3.4 Results and Discussion

### 3.4.1 Selected Ligands and Binding Mode of *lin*-Benzopurines to *Z. mobilis* TGT

In order to study putative changes in the protonation states of *lin*-benzopurines, **4a** has been selected for our initial measurements (Scheme 3.1).

This compound exhibits several basic functional groups appropriate to pick-up a proton, namely N(5) of the aminopyrimidinone substructure (encircled in orange), N(3) of the attached aminoimidazole ring (encircled in green), and the nitrogen of the morpholino substituent (encircled in blue).

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and pK<sub>a</sub> Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme



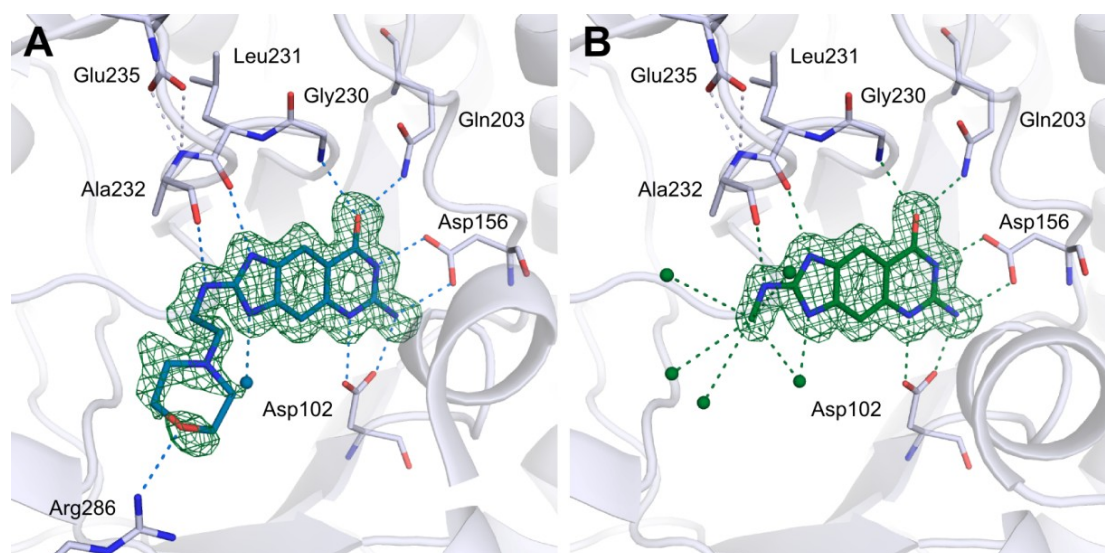
**Scheme 3.1** Chemical formulae of the investigated ligands. *lin*-Benzoguanine **4a** changes protonation state upon binding and exhibits three potential protonation sites (colored circles). Each of these sites is altered and possibly removed by chemical means in the ligands **3a** – **3c**.

A crystal structure of TGT in complex with **4a** was determined to a resolution of 1.42 Å (Figure 3.1A). The *lin*-benzoguanine scaffold is well-defined in the  $|F_o| - |F_c|$  difference electron density and a  $\pi$ -stacking interaction between the side chains of Tyr106 and Met260 can be observed. Polar interactions of the aminopyrimidinone ring are experienced with the side chains of Asp102, Asp156, Gln203, and the backbone NH of Gly230 as analogously found for other members of this compound series [Stengl *et al.*, 2007]. The guanidinium moiety of the aminoimidazole ring forms interactions to the closely adjacent backbone carbonyl oxygens of Ala232 and Leu231 and contacts a conserved water molecule, which itself is part of a crucial cluster of water molecules [Kohler *et al.*, 2009]. The two guanidine-type portions of the parent scaffold are highly buried in the protein and thus fully shielded from solvent access (aminopyrimidinone portion: 100%, aminoimidazole portion:  $\approx$  96%, if the conserved water molecule is excluded from the calculation).

We observed a reasonably well-defined electron density for the entire ligand. Although the average B-factor of the morpholino moiety is approximately three-times larger than that of the tricyclic scaffold ( $B_{morpholine} = 21.5 \text{ \AA}^2$  vs.  $B_{tricycle} = 9.1 \text{ \AA}^2$ ), the

$N^+-CH_2-CH_2-NH$  linker to the morpholine ring is observed in a *gauche*-conformation partly occupying the flat ribose-33 subpocket and forming a weak charge-assisted hydrogen bond (3.3 Å) via its morpholino oxygen to the guanidinium group of Arg286. The morpholine ring adopts a chair conformation and was refined in the shown orientation with an occupancy of 80%. To a minor extent the residual density indicates the fully extended substituent with an *anti*-conformation of the linker.

Since a broad series of *lin*-benzoguanines has been synthesized during our studies [Hörtner *et al.*, 2007; Stengl *et al.*, 2007; Kohler *et al.*, 2009; Barandun *et al.*, 2012], we considered selected ligands in the present work, either avoiding or decreasing the basic character of the scaffold and the substituents under consideration (Scheme 3.1).



**Figure 3.1** Binding modes of TGT·4a (PDB ID: 4PUJ) and TGT·3a (PDB ID: 4PUK). The protein is represented as cartoon. The ligand and interacting residues are represented as sticks (carbon<sub>Protein</sub> = gray, carbon<sub>Ligand</sub> = blue/green, nitrogen = blue, oxygen = red). For clarity the  $\pi$ -stacking residues Tyr106 and Met260, which flank the tricyclic core, are not shown. The *lin*-benzoguanine scaffold located in the guanine 34 binding pocket is well defined in the  $|F_o| - |F_c|$  difference electron density (green) at a sigma level of 2.5. The ligands form several interactions to Asp102, Asp156, Gln203, Gly230, Leu231, and Ala232 (dashed lines). The hydrogen bond to Leu231 is enabled by a ligand induced backbone-flip, which is stabilized by Glu235. A) The morpholinoethyl substituent shows residual mobility in the bound state indicated by the less well-defined electron density. Nevertheless, it was refined to an occupancy of 80%. The linker to the morpholine substituent adopts a *gauche*-conformation and forms a weak charge-assisted hydrogen bond of 3.3 Å between the oxygen of the morpholine ring and the guanidinium group of Arg286. To minor extent, the residual density indicates the fully extended *anti*-conformer. B) Due to the lack of the extended 2-substituent, the methyl group interacts with the solvent (green spheres) showing C $\cdots$ O distances of up to 4.1 Å.

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and pK<sub>a</sub> Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme

In ligand **3a** the basic morpholinoethyl substituent is replaced by a simple methyl group. The binary complex with *Z. mobilis* TGT was obtained with a maximum resolution of 1.49 Å (Figure 3.1B). The scaffold adopts exactly the same binding mode forming similar interactions to **4a**. Due to the lack of the extended 2-substituent, the methyl group is in direct contact with five water molecules showing C···O distances up to 4.1 Å.

In addition to the changes found in *lin*-benzoguanine **3a**, the exocyclic amino function at the pyrimidinone ring has been removed in *lin*-benzohypoxanthine **3b**. Barandun *et al.* recently reported on the binding mode of this class of *lin*-benzopurines [Barandun *et al.*, 2012]. In contrast to binary complexes observed with the *lin*-benzoguanines, Asp102 adopts a downwards-rotated orientation pointing away from the ligand and the binding pocket. The observed rotamer undergoes hydrogen bonds to the side chain of Asn70 and the backbone NH of Thr71. A similar geometry is found in apo TGT [Brenk *et al.*, 2003]. The created empty space between ligand and protein is filled by a cluster of six water molecules. All other interactions are established similarly as in the TGT complexes with *lin*-benzoguanines.

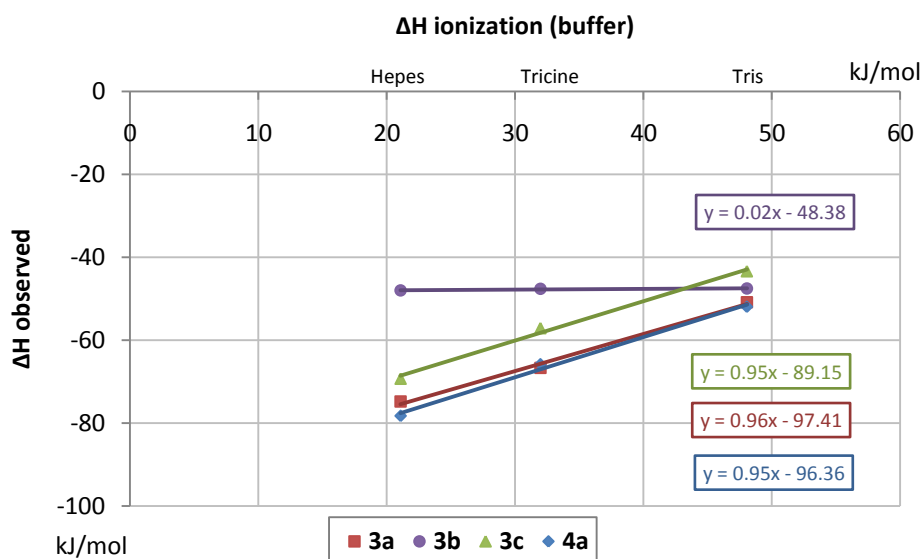
The strongly basic character of the aminoimidazole moiety in 2-amino-*lin*-benzopurines **3a**, **3b**, and **4a** has been changed to a less basic amidinium portion in **3c**. Except for the hydrogen bond to the backbone oxygen of Ala232, this compound is able to undergo the same hydrogen bonding pattern as described for the above-mentioned crystal structure (PDB ID: 3C2Y) [Ritschel *et al.*, 2009].

#### 3.4.2 Protonation States of *lin*-Benzopurine·TGT Complexes

To obtain insights into the potentially superimposed protonation effects that occur upon ligand binding, ITC measurements were performed using Hepes, Tris, and Tricine buffer at pH 7.8. The three buffers differ in their enthalpy of ionization [Christensen, 1976; Fukada and Takahashi, 1998]. The observed enthalpies for the complexation of the studied ligands were plotted against the heat of ionization of the buffers (Figure 3.2), and the recorded thermodynamic profiles are listed in Table 3.1. Ligand **4a** picks up 0.95 protons per mole formed complex, which either protonate the protein or the bound ligand (Figure 3.2, blue).

Apart from the basic functionalities of **4a**, there are several amino acid residues that could also be involved in the proton uptake. Within a 12 Å radius centered at Cy of Tyr106 (Figure 3.3) around the binding pocket, Asp102, Asp156, Glu157, Glu235, and Asp280 are

putative proton acceptor groups. According to Poisson-Boltzmann calculations these residues can be assumed as fully deprotonated at the applied pH conditions between 5.5 and 8.5 [Ritschel *et al.*, 2009]. Asp102 and Asp156 are directly involved in ligand binding (Figure 3.1).



**Figure 3.2**  $\Delta H_{obs}$  is plotted against  $\Delta H_{ion}$  for ligands **3a** – **4a** binding to TGT. The icons indicate the mean of the measured enthalpy values for each buffer system in kJ/mol. A linear fit was applied to the data points illustrated as solid lines. The positive slope shows a buffer dependency of **3a**, **3c**, and **4a** suggesting an uptake of approximately one proton per mole by the complex. Only binding of **3b** does not experience any changes in protonation state. The corrected value for  $\Delta H_{bind}$  can be extracted from the intercept, the number of captured protons results from the slope, which was calculated separately taking the standard deviations of each data point into account.

We decided to first investigate the ligand series exhibiting different basic characters. The *lin*-Benzoguanines **3a** (0.96 protons per mole; Figure 3.2, red) and **3c** (0.95 protons per mole; Figure 3.2, green) also suggest the uptake of approximately one proton. Only *lin*-benzohypoxanthine **3b** (0.02 protons per mole; Figure 3.2, purple) is obviously not capable to share these properties.

Our findings agree with the experimentally determined  $pK_a$  values reported for the *lin*-benzohypoxanthines and *lin*-benzoguanines in aqueous solution (Table 3.2) [Hörtner *et al.*, 2007; Barandun *et al.*, 2012].

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and pK<sub>a</sub> Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme

**Table 3.1** Thermodynamic profiles of the investigated ligands in the applied buffer systems.

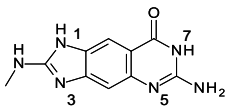
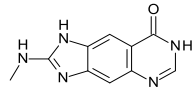
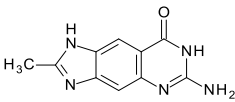
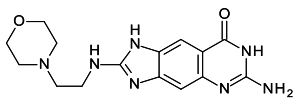
Ligand	TGT variant	$K_d$ (nM)	$\Delta G^0$ (kJ · mol <sup>-1</sup> )	buffer	$\Delta H_{obs}$ (kJ · mol <sup>-1</sup> ) <sup>[a]</sup>	$-T\Delta S^0$ (kJ · mol <sup>-1</sup> ) <sup>[b]</sup>
<b>3a</b>	Wild type	52.4 ± 6.9	-41.6 ± 0.3	Hepes	-74.8 ± 1.8	33.2 ± 1.9
				Tricine	-66.7 ± 0.1	25.1 ± 0.3
				Tris	-50.9 ± 0.9	9.3 ± 0.9
					<b>-97.4 ± 2.2</b>	<b>55.8 ± 2.2</b>
<b>3a</b>	D102N	146.1 ± 25.6	-39.0 ± 0.4	Hepes	-33.5 ± 0.4	-6.1 ± 0.5
				Tricine	-35.0 ± 0.4	-4.0 ± 0.5
				Tris	-33.9 ± 0.3	-4.8 ± 0.4
					<b>-33.9 ± 2.1</b>	<b>-4.9 ± 2.1</b>
<b>3a</b>	D156N	188.1 ± 59.1	-38.5 ± 0.8	Hepes	-51.4 ± 0.2	13.0 ± 0.4
				Tricine	-52.6 ± 2.5	13.4 ± 2.5
				Tris	-49.2 ± 0.4	11.4 ± 0.6
					<b>-53.2 ± 0.4</b>	<b>12.6 ± 0.9</b>
<b>3b</b>	Wild type	411.0 ± 63.8	-36.5 ± 0.4	Hepes	-48.0 ± 0.2	11.5 ± 0.4
				Tricine	-47.7 ± 0.2	11.2 ± 0.5
				Tris	-47.6 ± 0.3	11.1 ± 0.5
					<b>-48.4 ± 0.3</b>	<b>11.9 ± 0.5</b>
<b>3c</b>	Wild type	287.5 ± 52.9	-37.4 ± 0.5	Hepes	-69.3 ± 0.8	31.9 ± 0.9
				Tricine	-57.1 ± 1.6	19.7 ± 1.6
				Tris	-43.4 ± 0.1	6.0 ± 0.5
					<b>-89.2 ± 1.4</b>	<b>51.8 ± 1.5</b>
<b>4a</b>	Wild type	35.0 ± 6.9	-42.6 ± 0.5 <sup>[c]</sup>	Hepes	-78.3 ± 1.2	35.7 ± 1.3 <sup>[c]</sup>
				Tricine	-65.8 ± 0.3	23.2 ± 0.6
				Tris	-51.9 ± 1.1	9.3 ± 1.2
					<b>-96.4 ± 4.0<sup>[d]</sup></b>	<b>53.8 ± 4.0<sup>[d]</sup></b>
<b>4a</b>	E235Q	1275.6 ± 297.3	-33.8 ± 0.6	Hepes	-51.0 ± 0.7	17.9 ± 0.7
				Tricine	-39.9 ± 0.5	6.0 ± 0.5
				Tris	-28.6 ± 0.3	-5.7 ± 0.5
					<b>-66.2 ± 3.2</b>	<b>32.4 ± 3.3</b>

<sup>a</sup>The corresponding plot to obtain net  $\Delta H_{bind}$  of wild type TGT corrected for ionization effects is shown in Figure 2. <sup>b</sup> $-T\Delta S^0$  was calculated according to the Gibbs-Helmholtz equation. <sup>c</sup>Errors were estimated by means of standard deviation for  $K_d$  and  $\Delta G^0$  comprising at least six measurements and for  $\Delta H_{obs}$  at least two. The error for  $-T\Delta S^0$  was calculated according to error propagation. <sup>d</sup>Buffer corrected data are displayed in bold.



Unfortunately, no values are available for the morpholinoethyl substituent in **4a**; we therefore assume a value similar to that of parent morpholinoethyl. The value for the latter has been reported to be  $pK_a = 7.7$  [Hall, 1957]. We suppose that this assumption is justified as the morpholino portion is attached via an aliphatic ethyl linker which largely avoids transmission of electronic effects across molecular portions. For 2-amino-*lin*-benzoguanines a value of 4.4 has been assigned to  $N(5)H^+$  and 5.7 to  $N(3)H^+$  in *lin*-benzoguanine **3a** [Barandun *et al.*, 2012]. Remarkably, the acidity of  $N(5)H^+$  is shifted by about two logarithmic units to  $pK_a = 1.8$  in *lin*-benzohypoxanthine **3b**.

**Table 3.2** Measured  $pK_a$  values for the scaffolds of *lin*-benzoguanines **3a**, **3c**, **4a**, and *lin*-benzohypoxanthine **3b**.

		$pK_a$ measurements		
		N(7)H	N(3)H <sup>+</sup>	N(5)H <sup>+</sup>
<b>3a</b>		> 10	5.7 6.2 <sup>[b]</sup>	4.4
<b>3b</b>		> 10	5.6	1.8
<b>3c</b>		n.d. <sup>[a]</sup>	5.4	n.d. <sup>[a]</sup>
<b>4a</b>		n.d. <sup>[a]</sup>	n.d. <sup>[a]</sup>	n.d. <sup>[a]</sup>

<sup>a</sup>n.d.: not determined. <sup>b</sup>Data of Barandun *et al.*, 2012 supersede Hörtner *et al.*, 2007.

In order to confirm the experimentally derived evidence, Poisson-Boltzmann calculations were consulted to analyze shifts of  $pK_a$  values of the active site residues and the ligand functional groups upon binding [Czodrowski *et al.*, 2006]. Within the 12 Å sphere centered around Cy of Tyr106 all titratable residues were selected and the corresponding  $pK_a$  values were calculated prior and after complex formation. The calculations reveal only significant shifts in basic or acidic properties for Asp102, Asp156, and N(3) or N(5) of ligand **4a** (Table 3.3). Accordingly as a first working hypothesis, either N(3), embedded into an aminoimidazole substructure, or N(5), being part of an aminopyrimidinone moiety, could be likely candidates for the proton uptake. The shift suggested for N(5) appeared larger

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and $pK_a$ Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme

( $\Delta pK_a = 2.6$ ) compared to N(3) ( $\Delta pK_a = 1.0$ ), however, as N(3) shows the more basic character in aqueous solution, both sites appear appropriate to pick up a charge.

**Table 3.3.** Calculated  $pK_a$  values of the crucial amino-acid residues within the binding pocket and ligand **3a** before and after complex formation.

Residue <sup>[a]</sup>	$pK_a$ prior to complexation <sup>[b]</sup>	$pK_a$ in complexed state	
		N3	N5
<b>Asp102</b>	1.63	1.64	-2.12
<b>Asp156</b>	5.44	4.40	3.02
<b>3a<sup>[c]</sup></b>		6.65	7.00

<sup>a</sup>All other side chains exhibited only marginal changes in their  $pK_a$  values and are hence not listed. <sup>b</sup>For the Poisson-Boltzmann calculations a dielectric constant of  $\epsilon = 20$  at a pH of 7.8 was applied. <sup>c</sup> $pK_a$  values in aqueous solution according to Barandun *et al.*, 2012.

A previous study on the  $pK_a$  properties of TGT·2-amino-*lin*-benzoguanine inhibitor complexes was based on the hypothesis that the ligands bind with N(3) in protonated state [Hörtner *et al.*, 2007; Ritschel *et al.*, 2009]. Accordingly, the 2-aminoimidazole moiety of the ligand scaffold would be positively charged and form charge-assisted hydrogen bonds to the backbone carbonyl oxygens of Leu231 and Ala232. This interaction would induce a backbone flip, which is stabilized by a further charge-assisted hydrogen bond formed between Glu235 and the backbone NH of Leu231 [Hörtner *et al.*, 2007; Ritschel *et al.*, 2009]. The nearly 30-fold increase in binding affinity suggested by a biochemical assay between **3a** and **3c** was considered as clear indication of this protonation event at the imidazole moiety of **3a**. However, recent results also support the hypothesis that N(5) is the likely candidate to pick-up a proton. A newly published crystal structure of *Z. mobilis* TGT in complex with 2-amino-*lin*-benzohypoxanthine-type inhibitors shows that Asp102 adopts a significantly different orientation in the complexes compared to the corresponding *lin*-benzoguanine complexes [Barandun *et al.*, 2012]. This residue is rotated away from the ligand towards the ribose 34 subpocket as similarly found in the apo-enzyme [Brenk *et al.*, 2003; Barandun *et al.*, 2012]. In the new position, Asp102 forms an interaction with the carboxamide group of Asn70 and the backbone NH of Thr71 [Barandun *et al.*, 2012]. By capturing a proton, the *lin*-benzoguanine moiety obviously shifts its  $pK_a$  properties and induces an alternative conformation of Asp102 to interact with the ligand's guanidinium group.

In order to assess whether N(3) or N(5) becomes protonated upon protein binding or whether an even more complex situation with several superimposed and mutually compensating effects is given, we embarked onto a systematic mutational study combined with subsequent ITC experiments and  $pK_a$  calculations.

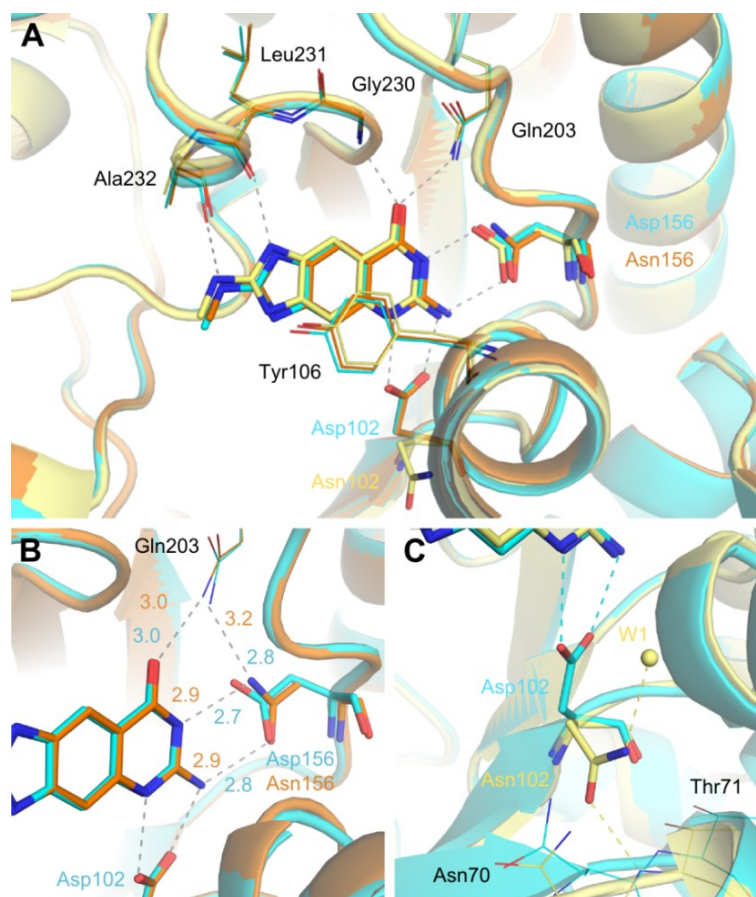
### 3.4.3 Mutational Studies

According to our  $pK_a$  calculations, Asp102 and Asp156 form charge-assisted interactions with the aminopyrimidinone portion of the *lin*-benzoguanine-type inhibitors and might be involved in the protonation of N(5). Therefore, we decided to replace Asp by Asn at both sites to investigate the impact of these residues on ligand binding.

The replacement by Asn avoids the negative charge in this region and introduces a permanently uncharged contact. While the Asp156Asn mutant still exhibits some, however by a factor of three significantly reduced catalytic activity compared to the wild type, the Asp102Asn mutant shows no activity at all [Romier *et al.*, 1996]. For both mutants, the crystal structures with **3a** were determined and ITC measurements carried out in the three buffer systems.

Structural data were recorded to a resolution of 1.65 Å and 1.85 Å. Ligand **3a** adopts in the Asp156Asn mutant a similar binding mode to the wild type (Figure 3.3A and 3.3B). The aminopyrimidinone ring interacts with the side chains of Asp102, Asn156, Gln203, and the backbone NH of Gly230. The orientation of the replaced Asn156 residue is slightly shifted compared to the wild type moving the terminal carboxamide group approximately 16° out of the plane through the *lin*-benzoguanine scaffold. The interaction between the terminal carboxamide oxygen of Asn156 and the ligand's exocyclic amino function remains unchanged compared to the wild type ( $\Delta d = 0.1$  Å). However, the adjacent contact via the NH<sub>2</sub> of the carboxamide group of Asn156 and N(7) of **3a** is extended from 2.7 Å to 2.9 Å. This expansion both results from the larger radius of the nitrogen atom and the replacement of a charge-assisted hydrogen bond by a neutral one. Additionally, the side chain oxygen of residue Gln203 tilts about 13° to the front elongating the distance between its amino function and the same group in Asn156 from 2.8 Å to 3.2 Å. The strength of the interaction formed to the carbonyl oxygen of the complexed ligand and Gln203 remains unaffected as the uncharged contact distance suggests. Also the interactions formed between the ligand

and the side chain of Asp102 do not differ from the wild type. To establish this interaction pattern, the proton at N(7) must be transferred to N(5) to produce the other tautomer. In summary, neutral hydrogen bonds are formed to Asn156 and charge-assisted ones to Asp102 (Scheme 3.2).



**Figure 3.3** The protein is displayed as cartoon with its interacting residues shown in line representation. The mutated side chains and **3a** are highlighted by sticks (nitrogen = blue, oxygen = red). Dashed lines represent hydrogen bonds. Distances are given in Å. A) Superimposition of **3a** as bound to wild type TGT (cyan; PDB ID: 4PUK), mutant Asp102Asn (yellow; PDB ID: 4PUL) and mutant Asp156Asn (orange; PDB ID: 4PUM). The ligands adopt overall the same orientation in all displayed complexes. The side chain of Asn102 adopts a conformation rotated away from the binding site compared to Asp102 in the wild type enzyme. Contrary, Asn156 remains in a similar position as observed for Asp156. B) Close-up of the interactions of the mutated residue Asn156 (orange) compared to those of wild type enzyme (cyan). Asn156 is slightly rotated by 16° extending the interaction to the amino group of Gln203 and N(7) of the ligand. All other connections remain unaffected. C) Close-up of the interactions of the mutated residue Asn102 (yellow) compared to those of wild type enzyme (cyan). Asn102 shows the rotated orientation adopting a similar conformation as found for the *lin*-benzohypoxanthine complexes. Instead of interacting similarly with the ligand as Asp102 in the wild type, Asn102 forms contacts to the backbone NH of Asn70 and a nearby water molecule.

In the complex structure of the Asp102Asn mutant with **3a**, the ligand adopts a position similar to the one in the wild type. It is slightly shifted off from the position of Asp156 (Figure 3.3A). The mutated residue Asn102 does not form hydrogen bonds to the aminopyrimidinone moiety of the ligand but adopts an orientation similar to those of wild type enzymes with the *lin*-benzohypoxanthine-type inhibitors. A hydrogen bond is formed to a nearby water molecule (2.8 Å) and the backbone NH of Thr71 (2.7 Å) [Barandun *et al.*, 2012].

As expected, the binding affinity of **3a** determined by ITC measurements drops for both mutant variants as favorable interactions are lost in both cases. For the Asp156Asn mutant a reduction by a factor of nearly four (wild type:  $49 \pm 5$  nM, mutant:  $188 \pm 59$  nM) is found, whereas a threefold loss is recorded for the Asp102Asn mutant ( $146 \pm 26$  nM).

Surprisingly, **3a** does not show any significant buffer dependency upon binding to the two mutants (Table 3.1). Also, Poisson-Boltzmann calculations performed on these mutant complexes support this hypothesis (Table 3.4). The close contacts to the two negatively charged aspartates in the wild type reinforce protonation of the ligand and cause the strong  $pK_a$  shift.

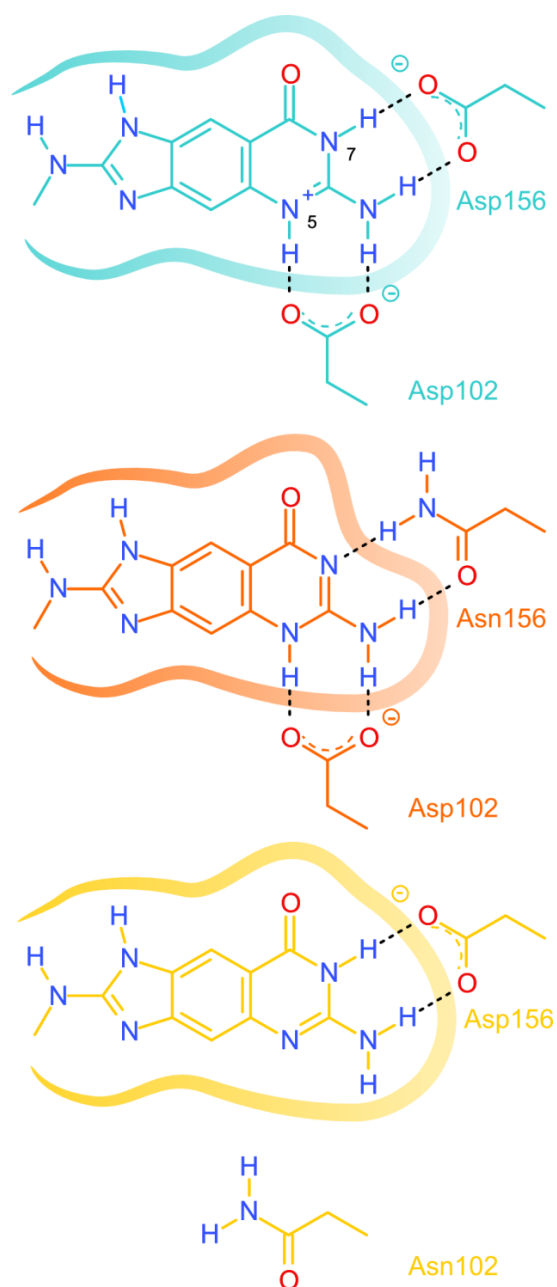
Remarkably, in the Asp102Asn mutant, the presence of Asp156 as charged residue and the neutral Asn102, which is rotated away from the binding site is not sufficient to induce an equally strong  $pK_a$  shift and thus, a protonation of the ligand as experienced in the wild type. This finding also correlates with the expanded interactions of Asn156 and the ligand's aminopyrimidinone moiety in the Asp156Asn mutant where no positive charge is located on this part of the ligand.

**Table 3.4** Calculated  $pK_a$  values of the crucial amino-acid residues within the binding pocket of Asp102Asn and Asp156Asn, respectively, and N(5) of ligand **3a** before and after complex formation.

Mutant	Residue	$pK_a$ prior to complexation <sup>[a]</sup>	$pK_a$ in complexed state
<b>Asp102Asn</b>	Asp156	2.80	2.57
	<b>3a</b>		-1.14
<b>Asp156Asn</b>	Asp102	1.29	-2.31
	<b>3a</b>		5.32

<sup>a</sup>For the Poisson-Boltzmann calculations the same conditions were applied as noted in Table 3.

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and $pK_a$ Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme



**Scheme 3.2** Interactions formed between **3a** and amino acids 102 and 156 of wild type TGT (cyan), the D156N mutant (orange) and the Asp102Asn mutant (yellow). In the TGT·**3a** complex N(5) of the parent scaffold becomes protonated upon binding and forms salt bridges to Asp102 and Asp156. To bind to the Asp156Asn-mutant in a similar way, the proton at N(7) must be transferred to N(5), then forming hydrogen bonds to Asn156 and charge-assisted ones to Asp102. In the Asp102Asn·**3a** complex Asn102 is rotated off the binding pocket. Therefore, N(7) carries the proton to form charge-assisted hydrogen bonds to Asp156.

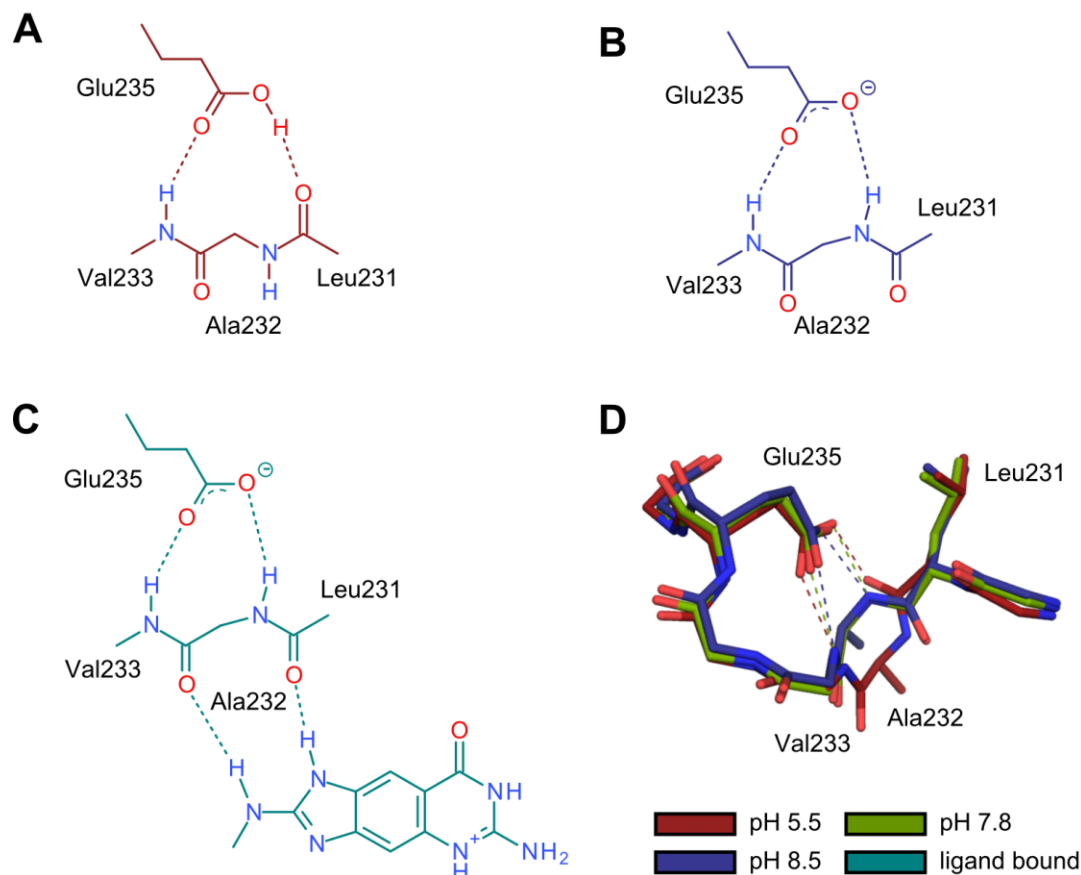
Up to this point all experimental findings indicate a protonation of N(5) in the aminopyrimidinone moiety. Does this, however, rule out the former hypothesis that N(3) in the aminoimidazole portion changes protonation and that the picked-up charge contributes significantly to the enhanced binding affinity? Also, a more complex situation with respect to changes of protonation states might be given in this part of the structure. In this area, Glu235 takes an important role on ligand-binding and protein function. It has been shown that Glu235 adapts its protonation state depending on the orientation of the peptide bond between Leu231 and Ala232 (Scheme 3.3A, B) [Brenk *et al.*, 2003]. This peptide bond is exposed to the binding pocket and experiences a backbone flip upon ligand binding or upon changes in the environmental pH value. If a *lin*-benzopurine binds to the active site, the peptide bond's carbonyl oxygen of Ala232 is oriented into the binding pocket and accepts a hydrogen bond from N(1)H of the ligand (Scheme 3.3C). In consequence, Glu235 becomes deprotonated in order to accept a hydrogen bond of the flipped backbone NH group [Tidten *et al.*, 2007]. Glu235 is the only proximal amino acid involved in ligand binding, which could be able to release a proton.

Based on these considerations, we mutated Glu235 to Gln to experimentally assess the remaining overall protonation inventory. ITC experiments were carried out with the Gln235 variant in the three mentioned buffer systems indicating an overall proton uptake of 0.79 protons per mole. This nearly identical protonation inventory speaks for N(5) in the aminopyrimidinone moiety as sole proton acceptor of the *lin*-benzoguanine scaffold. Nonetheless, these experiments do not completely rule out a possibly more complex situation. Glu235 has to release a proton in consequence of the peptide bond flip. This only makes this bond able to recognize the *lin*-benzopurine scaffold.

It could be imagined that the proton release of Glu235 is combined with the simultaneous protonation of N(3) in the aminoimidazole portion. In this case both effects would mutually compensate and nullify in the inventory. The following experiment makes this consideration unlikely. Previously, we were able to determine the apo structures of TGT wild type at pH 5.5 and 8.5 [Romier *et al.*, 1996; Brenk *et al.*, 2003]. At pH 5.5, the NH group of the Leu231/Ala232 peptide bond is oriented towards the binding pocket and Glu235 must adopt a protonated state to stabilize the peptide bond (Scheme 3.3A). At pH 8.5, the flipped peptide bond orientation is given and Glu235 has to adopt a deprotonated state

### 3. Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and $pK_a$ Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme

(Scheme 3.3B). We therefore took crystals grown at pH 5.5 and transferred them in a gradual stepwise fashion to higher pH finally reaching a value of 7.8.



**Scheme 3.3** Dependence of the peptide bond between Leu231 and Ala232 on the pH value and ligand binding. TGT apo structures of crystal grown at different pH values clearly indicate that a backbone flip between Leu231 and Ala232 occurs dependent on the environmental pH value and the ligand bound. A) Scheme of the backbone orientation at pH 5.5. The backbone NH group of Leu231 is facing the binding pocket. The carboxyl group of Glu235 is present in its protonated state interacting with the carbonyl function of Leu231 and the backbone NH group of Val233. B) Scheme of the backbone orientation at pH 8.5. The backbone carbonyl function of Leu231 is exposed to the binding pocket. Hence, the side chain of Glu235 accepts two hydrogen bonds from the backbone NH groups of Ala232 and Val233 and must occur in its deprotonated form. The crystal structure determined at pH 7.8 and obtained by a gentle pH transition from pH 5.5 shows the same backbone orientation as at pH 8.5. This clearly indicates that Glu235 is already deprotonated in the apo structure at pH 7.8. C) Scheme of the ligand bound backbone orientation. Taking the orientation of the peptide bond between Leu231 and Ala232 in the apo structure at pH 7.8 into account, Glu235 accommodates a bound ligand without release of a proton at this pH value. D) Overlay of TGT apo structures at pH 5.5 (PDB ID: 1P0D), 7.8 (PDB ID: 4PUN) and 8.5 (PDB ID: 1PUD). Amino acid residues are shown in stick representation (nitrogen = blue, oxygen = red). For clarity the side chain of Val233 is not shown.



Subsequently, we subjected the thus treated crystal to a structure determination and observed the geometry of the peptide bond with the backbone carbonyl group oriented towards the binding pocket (Scheme 3.3B). This geometry is only compatible with Glu235 present in the deprotonated state. We thus conclude that the backbone flip must occur already at lower pH and ligand binding studied by our ITC titrations occurs to the enzyme with the peptide bond in the orientation for binding the ligand and with Glu235 in the deprotonated state. Therefore, a simultaneous compensating protonation change of N(3) and Glu235 can be excluded.

Finally, there remains the question why the addition of the exocyclic C(2)-NH group at the imidazole moiety of the *lin*-benzopurines (**3a** compared to **3c**) results in a significant affinity enhancement, which made Ritschel *et al.* to believe that this part of the ligand would bear a charge improving the affinity contribution of the formed additional hydrogen bonds [Ritschel *et al.*, 2009]. The biochemical assay exaggerated this improvement (factor 30) as the ITC experiments suggest a smaller value (factor 5.5). Nonetheless, the enhancement is still remarkable. It has to be regarded that both NH functionalities of the 2-aminoimidazole portion forming the hydrogen bonds to the neighboring carbonyl groups of Leu231 and Ala232 are mutually adjacent (Scheme 3.3C). This dual hydrogen bonding array involves a pattern of two donor and acceptor groups with parallel orientation. Such an arrangement avoids secondary repulsive interactions that, for example, are given in the above-discussed contact of the aminopyrimidinone moiety and Asn156 (Scheme 3.2, center). The separation distance of the hydrogen atoms in the directly adjacent hydrogen bonds is rather short and thus substantial electrostatic interactions will occur [Jorgensen and Pranata, 1990]. Partial positive charges are given on the hydrogen atoms whereas partial negative charges are experienced on the nitrogen and oxygen atoms in the NH...O hydrogen bonds. A favorable situation is given, as in the present case, where one of the binding partners bears all the hydrogen donor groups, the other all acceptor groups. The situation is worse, if donor and acceptor sites alternate between both partners, since then secondary repulsive interactions results from complexation. This incident is given in the contact to Asn156 and clearly responded by the structure with expanded and distorted geometry (Figure 3.3B). The enhancement of adjacent hydrogen bonding contacts following the DD/AA pattern has already been described in host-guest complexes many years ago [Murray and Zimmerman, 1992].

### 3.5 Conclusion

In the present study, changes of the protonation states of ligands binding to the active site of the enzyme TGT were investigated. ITC measurements supported by  $pK_a$  calculations confirmed the uptake of one proton in case of the *lin*-benzoguanine-type ligands. The structurally related *lin*-benzohypoxanthines do not show such a pick-up of a proton. Considering the adopted binding modes, the *lin*-benzohypoxanthines do not induce a relocation of the side chain of Asp102, which binds in case of the *lin*-benzoguanines directly to the ligand via two short hydrogen bonds. This relocation of Asp102 and the presence of Asp156 in short distance to the aminopyrimidinone moiety of the ligand creates a negatively charged environment. This provokes a significant  $pK_a$  shift of the ligand's heterocyclic moiety, which therefore becomes protonated and thus positively charged. Mutational exchanges of both Asp residues by Asn underline that the cluster of negative charges is responsible for this induced  $pK_a$  shift.

However, the intrinsic  $pK_a$  values of the uncomplexed ligand under consideration must fall into a crucial window to achieve such a change in protonation upon binding. Both, the *lin*-benzoguanines and the *lin*-benzohypoxanthines exhibit a second guanidinium type motif in the five-membered imidazole portion. Our investigations reveal, however, that N(3) of the aminoimidazole moiety does not change its protonation state upon binding but the moiety experiences favorable hydrogen bond contacts to the protein avoiding secondary repulsive interactions. In summary, the *lin*-benzoguanines bear no charge in aqueous solution at pH 7.8 but become positively charged at the binding site. The charge is mainly distributed over the guanidine motif of the aminopyrimidinone and only minor effects might influence the aminoimidazole portion.

This observation explains the high potency of the *lin*-benzoguanines showing binding affinities down to the one-digit nanomolar range. Their binding is characterized by a strong enthalpic binding signature resulting from the salt bridges formed between the aminopyrimidinone moiety and the deprotonated side chains of Asp102 and Asp156 (Table 1). The contacts to both acidic residues are established in a deeply buried binding pocket that shields the formed charge-assisted interactions from solvent access. In this environment of low local dielectricity the formed charged contacts experience an enhanced affinity contribution [Bogan and Thorn, 1998]. In addition, the hydrogen bonds formed

between the aminoimidazole ring and the backbone carbonyl oxygens of Leu231 and Ala232 are expected to be energetically favorable avoiding repulsive effects.

The dramatic loss in binding affinity in case of the *lin*-benzohypoxanthines results from a loss of the interaction to Asp102, which rotates away from the ligand recognition site towards the ribose-34 binding pocket. As a consequence, the pyrimidinone ring remains uncharged, as found in aqueous solution prior to binding. Obviously, Asp156 alone does not take sufficient impact on the adjacent pyrimidinone moiety of the *lin*-benzohypoxanthines to induce a sufficiently large  $pK_a$  shift to result in protonation and consequently charge-assisted interactions.

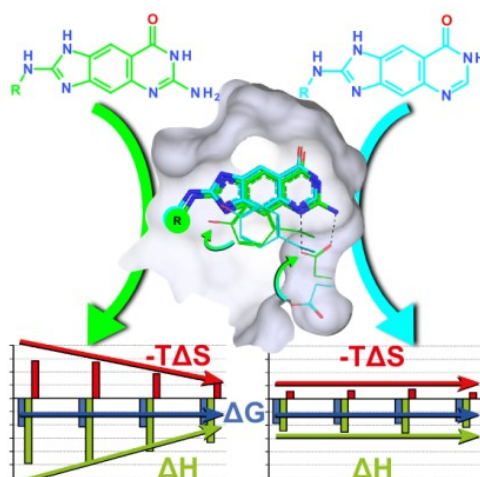
With respect to bioavailability, the aminopyrimidinone portion in the *lin*-benzoguanines exhibits the desirable feature of getting protonated upon protein binding only. Despite this advantage, overall the molecular properties of the ligands are not yet ideal regarding membrane transportation due to a large polar character of the molecules. Both the *lin*-benzoguanines and the *lin*-benzohypoxanthines show unsatisfactory membrane permeation in PAMPA measurements [Barandun *et al.*, 2012].

This study demonstrates the importance to analyze the protonation properties of structurally closely related ligands. ITC titrations, performed in different buffer conditions, allow elucidating the net protonation inventory and stoichiometry. These studies, however, have to be complemented by site-directed mutagenesis and  $pK_a$  calculations to trace where the protons go and accordingly the charges reside on the formed complex. They also help to avoid false conclusions resulting from mutually compensating effects and allow to localize the hot spots of binding and to tailor ligand properties so that transformation to the charged state only occurs upon target binding.

## 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme

### 4.1 Introductory Remarks

The following chapter was published in the scientific *Journal of Medicinal Chemistry*.<sup>3</sup> Water cluster analyses were carried out by Michael Betz (Philipps-Universität Marburg). TGT compounds were synthesized by Luzi Jakob Barandun and Christoph Hohn (ETH Zürich). All crystallographic and microcalorimetric studies have been performed by the author of this thesis along with the drafting and discussion of the paper.



### 4.2 Abstract

Lead optimization focuses on binding affinity improvement. If a flat structure–activity relationship (SAR) is detected, usually the optimization strategy will be abolished as unattractive. Nonetheless, as affinity is composed of an enthalpic and entropic contribution, factorization of both can unravel the complexity of a flat, on first sight tedious SAR. In such cases, the binding free energy of different ligands can be rather similar, its factorization into enthalpy and entropy can be distinct. We investigated the thermodynamic signature of two classes of *lin*-benzopurines binding to tRNA–guanine transglycosylase. While the differences

<sup>3</sup> Neeb *et al.* (2014). Beyond Affinity: Enthalpy-Entropy Factorization Unravels Complexity of a Flat Structure-Activity Relationship for Inhibition of a tRNA-Modifying Enzyme. *Journal of Medicinal Chemistry* 57, 5566-5578.

between the two series are hardly visible in the free energy term, they involve striking enthalpic and entropic changes. Analyzing thermodynamics along with structural features revealed that one ligand set binds to the protein without inducing significant changes compared to the apo structure, however, the second series provokes complex adaptation leading to a conformation closely similar to the substrate-bound state. In the latter state, a cross-talk between two pockets is suggested based on these observations.

### 4.3 Introduction

In a medicinal chemistry program, the optimization of a given lead structure seeks for an improvement of binding affinity. This goal is usually achieved by systematic replacement and expansion of functional groups and building blocks at the parent scaffold of the lead structure. Nowadays, the optimization process is supported by structure-based considerations, particularly if the 3D structure of the target protein is available. The value of such modeling attempts is highly appreciated, however, the experts also know that modeling might predict correct affinity correlations and structure–activity relationships based on false assumptions, which remain undiscovered as long as only affinity data are available. Without performing crystal structure determinations of every optimization candidate with the target protein, such putative misconceptions about binding modes or interaction patterns could easily remain undetected. Complex and difficult structure–activity relationships can even be reflected by featureless and flat correlations due to compensating effects. In such cases, it might be important to validate whether additional parameters beyond affinity can help to obtain more relevant insights into the properties of the studied system.

The binding affinity is a Gibbs free energy value that decomposes into an enthalpic and entropic contribution. Both properties can provide additional information about the binding event, however, as their changes are also related to the entire protein–ligand binding process, any modification of the whole system will be compressed into the measured overall  $\Delta G^0$ ,  $\Delta H^0$ , and  $T\Delta S^0$ . It is therefore extremely difficult to factorize the changes of these properties into individual contributions that can be assigned to single interactions or particular binding steps. Only within congeneric series of ligands and considering the binding to proteins classified as rigid, such decompositions have been successful [Dullweber *et al.*, 2001; Baum *et al.*, 2010; Snyder *et al.*, 2011; Biela *et al.*, 2012;

Biela *et al.*, 2012; Biela *et al.*, 2013; Martin and Clements, 2013]. In these analyses, it should not be forgotten that changes in the individual solvation pattern resulting from differences of one single water molecule can strongly perturb and shift the thermodynamic signature between ligands. Nonetheless, differences in the thermodynamic profiles of chemically closely related ligands usually indicate deviating binding features and help to validate structure–activity relationships.

In the current study, we investigated the binding of two ligand series modified at two positions of the parent *lin*-benzopurine scaffold (Scheme 4.1) [Barandun *et al.*, 2012]. The first modification involves attachment of an exocyclic amino group to a pyrimidinone headgroup, which makes the considered ligands competent to establish enhanced charge-assisted or even salt-bridge type hydrogen bonds [Neeb *et al.*, 2014]. The second modification concerns attachment of a series of substituted amino groups at the remote end of the parent scaffold [Ritschel *et al.*, 2009]. The binding affinity determined via the dissociation constant ( $K_d$  value) measured by isothermal titration calorimetry varies only slightly among individual derivatives of the two series and differs among corresponding members of both series by about one order of magnitude depending whether the amino group is present or absent at the pyrimidinone moiety [Barandun *et al.*, 2012]. It is exceptional is, however, that the attachment of the remote amino substituents of chemically rather different nature does not have much influence on the affinity. This at first glance rather unspectacular flat structure–activity relationship turns out to be very substantial if thermodynamic profiles and detailed binding modes are correlated.

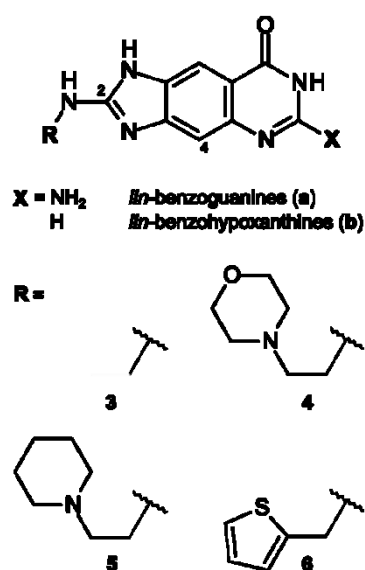
The addressed structure–activity relationship is presented for the binding of ligands to the tRNA-modifying enzyme tRNA–guanine transglycosylase (TGT). This enzyme catalyzes a complete exchange of a nucleobase at the wobble position of some tRNAs [Xie *et al.*, 2003]. Inhibition of this protein in bacteria is a promising therapeutic perspective as its function has been linked to the pathogenicity of *Shigella*, the causative agent of bacterial dysentery [Durand *et al.*, 2000; Sansonetti, 2001]. Shigellosis occurs predominantly in developing countries and is responsible for more than 100'000 lethal cases every year [Kotloff *et al.*, 1999].

In the present study, we show that two ligand series based on two slightly different parent scaffolds show quite distinct thermodynamic signatures even though the same substituents are attached. The recorded differences find an explanation in the features of

the protein conformer to which the ligands bind. Whereas one series binds to a protein conformer similar to that of the apo protein, the other series binds to the enzyme in a conformation related to the one adopted with the bound substrate. Due to considerable rearrangements of the protein, a crosstalk between binding subpockets is established, which gives rise to differences in the thermodynamic signatures. These significant differences in the binding properties would have never been detected if solely affinity data across the two ligand series would have been evaluated. Only by taking the thermodynamic profiles and the partitioning in enthalpy and entropy into account these differences do become apparent.

## 4.4 Results and Discussion

### 4.4.1 Binding Affinities and Thermodynamic Profiles of the 2-Amino-*lin*-benzopurines



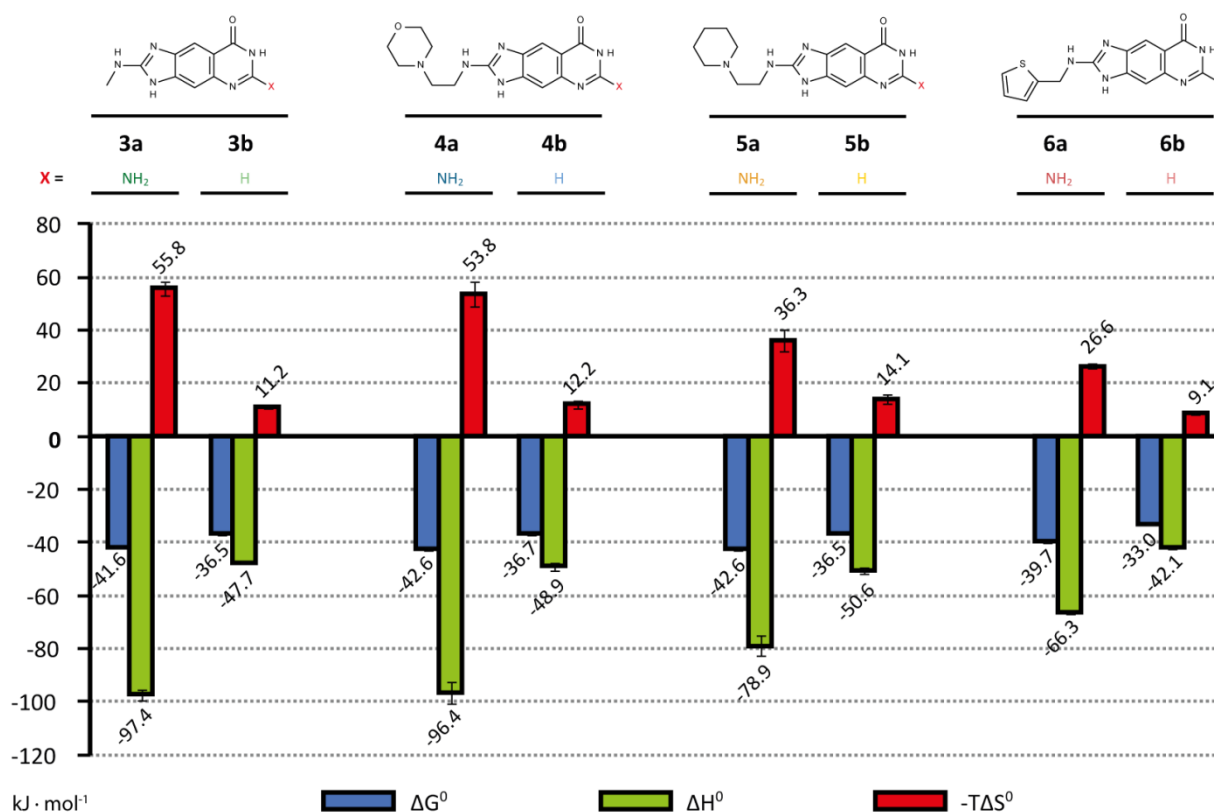
**Scheme 4.1** Chemical formulas of the investigated ligands **3a** – **6a** and **3b** – **6b**.

To establish a structure–activity relationship for the two compound series **3a** – **6a** and **3b** – **6b** (Scheme 4.1), we determined their inhibitory potency ( $K_i$  values). In previous studies, we characterized this property by a functional biochemical assay, which records by how much the incorporation of radioactively labeled guanine into the tRNA is suppressed when the catalytic properties of the target enzyme TGT are blocked [Meyer *et al.*, 2006]. This assay is a rather indirect measure, as guanine itself exhibits some inhibitory potency and a size-dependent inhibition model has to be considered for the bound ligands. Inhibitors decorated

by large substituents are competitive with tRNA binding whereas inhibitors of the size of the natural nucleobase substrates can block the enzyme, even if the tRNA is intermediately covalently attached to the protein. They compete for the binding site of the exchanged nucleobase. In the latter situation, the protein must be present in the substrate-bound conformation. As this binding assay superimposes multiple steps to the actual binding event and characterizes the inhibitory potential of the ligands rather indirectly, we used isothermal titration calorimetry (ITC) in the present study instead. This method directly reveals the dissociation constant  $K_d$  for the binary protein–ligand binding event. From this, the Gibbs free energy of binding can be accessed using the equation  $\Delta G = R \cdot T \cdot \ln K_d$ . The measured data are listed in Figure 4.1. Remarkably, the ITC binding data of **3a** – **6a** match quite well with those determined by the biochemical assay, whereas for the series **3b** – **6b** deviations up to a factor of 16 (for **3b**) are found [Ritschel *et al.*, 2009; Barandun *et al.*, 2012]. Apart from differences in the applied assay buffer conditions and the temperature used to run the experiments (ITC: 25°C, biochemical assay: 37°C), conformational differences given for the protein and discussed later in this contribution might give rise to these deviations.

The Gibbs free energy of binding shows rather similar values across the *lin*-benzoguanine **3a** – **6a** (mean:  $-41.6 \pm 1.4 \text{ kJ} \cdot \text{mol}^{-1}$ ) and *lin*-benzohypoxanthine **3b** – **6b** series (mean:  $-35.7 \pm 1.8 \text{ kJ} \cdot \text{mol}^{-1}$ ) indicating a flat and at first sight rather unspectacular structure–activity relationship. Without consulting the additional thermodynamic properties, this finding could have been the end of the current drug development study. The more it is surprising that the thermodynamic signature factorizes in both series very differently in enthalpy and entropy. All inhibitors exhibit strong exothermic binding and, as expected, this strong negative enthalpic contribution is larger for the *lin*-benzoguanine series. Remarkably, the enthalpic signature scatters for the *lin*-benzohypoxanthines series across a rather small window of  $\Delta\Delta H^0 = 8.5 \text{ kJ} \cdot \text{mol}^{-1}$  whereas the *lin*-benzoguanines spread over a nearly four times larger range of  $\Delta\Delta H^0 = 31.1 \text{ kJ} \cdot \text{mol}^{-1}$ . Among the latter compound series, **6a** exhibits with  $-66.3 \text{ kJ} \cdot \text{mol}^{-1}$  the smallest exothermic signal followed by **5a** with  $-78.9 \text{ kJ} \cdot \text{mol}^{-1}$ . Inhibitors **3a** and **4a** show the largest negative enthalpy,  $-97.4 \text{ kJ} \cdot \text{mol}^{-1}$  and  $-96.4 \text{ kJ} \cdot \text{mol}^{-1}$ . The entropic contribution is unfavorable for all ligands of both series. Remarkably, the methyl derivative **3a** and the morpholinoethyl derivative **4a** exhibit nearly identical entropic signatures. Overall, the entropic contribution is detrimental to the free energy of binding, resulting in pronounced enthalpy/entropy compensation in both series.





**Figure 4.1** Chemical formulas and thermodynamic profiles of the investigated ligands. The compounds **3** – **6** differ in their substitution pattern in 2-position. Additionally, their scaffolds vary among each other from *lin*-benzoguanines ( $X = \text{NH}_2$ ) to *lin*-benzohypoxanthines ( $X = \text{H}$ ). Values for the Gibbs free enthalpy are shown as blue bars, buffer corrected enthalpy values as green bars, and entropy values as red bars, each in units of  $\text{kJ} \cdot \text{mol}^{-1}$ . Measurements to obtain  $\Delta G^0$  comprise at least six measurements. Compounds **3a**, **4a**, **5a**, and **6a** show a buffer dependency. On average, approximately one proton is picked up by the ligand upon binding. Thus,  $\Delta H_{obs}$  was measured at least in duplicate in three different buffer systems (Hepes, Tris, Tricine). Subsequently, enthalpy values of *lin*-benzoguanines were corrected for buffer contribution.  $\Delta G^0$  and  $\Delta H_{bind}$  values for *lin*-benzohypoxanthines represent the mean of at least three independent measurements.  $-T\Delta S^0$  was calculated according to the Gibbs-Helmholtz equation.

The observed trends in the thermodynamic signatures indicate that the attached substituents in the ribose-33 pocket are not the sole determinant for the observed profiles, as the relative differences in the thermodynamic properties among ligands with the same 2-substituents do not match across the two series **a** and **b**. Also the differences between ligand pairs within the two series do not correspond to one another. Focusing on a mutual comparison of ligands from both series with identical 2-substitution reveals differences in  $\Delta\Delta H^0$  of  $24.2 \text{ kJ} \cdot \text{mol}^{-1}$  to  $49.7 \text{ kJ} \cdot \text{mol}^{-1}$ . At first glance, they should map the energy

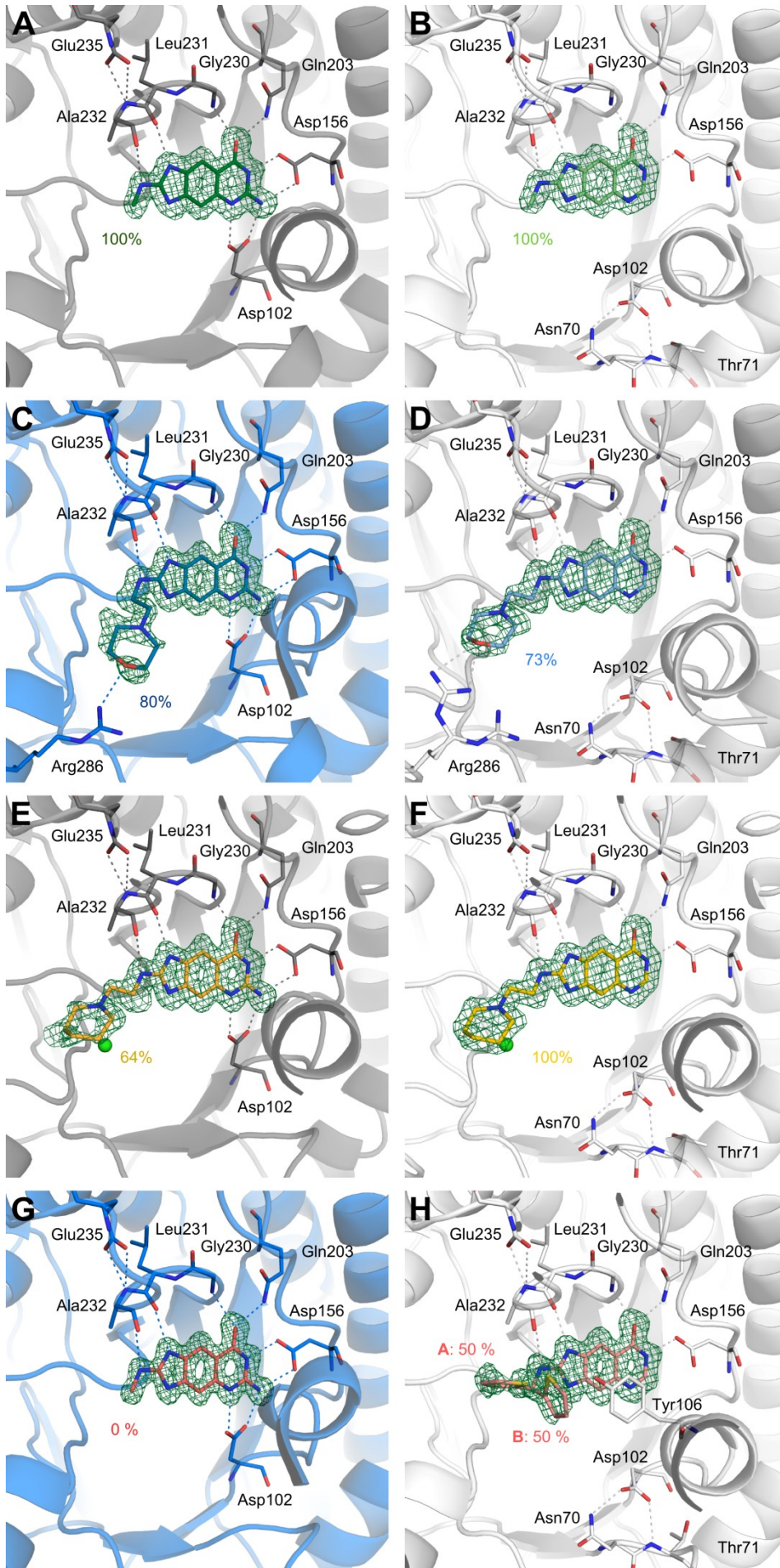
difference resulting from the exocyclic NH<sub>2</sub> group at the pyrimidinone moiety. Nonetheless, the deviating factorization of the thermodynamic data of the two compound series indicates a more complex structure–activity relationship than initially anticipated. Thus, to obtain a more detailed insight into the binding properties of the two ligand series, we resolved the crystal structures of the studied inhibitors bound to the enzyme.

#### 4.4.2 Crystal Structures and Binding Modes of 2-Amino-*lin*-benzopurines in *Z. mobilis* TGT

The 2-amino-*lin*-benzoguanines **3a** – **6a** were determined at pH 5.5 with a resolution of 1.25 – 1.49 Å, the corresponding *lin*-benzohypoxanthines **3b** – **6b** provided crystals that diffracted to 1.33 – 1.65 Å (Figure 4.2). The structures of **3a** [Neeb *et al.*, 2014], **3b** [Barandun *et al.*, 2012], and **4a** [Neeb *et al.*, 2014] had been previously determined.

In all structure determinations, the tricyclic *lin*-benzoguanine scaffold is well-defined in the electron density. It adopts a binding pose with favorable  $\pi$ -stacking interactions to the adjacent Tyr106 and Met260. As reported previously, ligands based on the *lin*-benzoguanine scaffold form hydrogen bonds to the side chain functional groups of Asp102, Asp156, Gln203, and the backbone NH group of Gly230 (Figures 4.2A, C, E, G) [Hörtner *et al.*, 2007; Stengl *et al.*, 2007; Kohler *et al.*, 2009]. In addition, the backbone carbonyl groups of Leu231 and Ala232 are hydrogen-bonded to the aminoimidazole portion of the ligand. Exposure of the two backbone C=O groups to the binding pocket in parallel fashion requires a peptide-backbone flip compared to the apo protein determined at a pH value of 5.5. The adjacent Glu235 carboxylate group is deprotonated and accepts two hydrogen bonds from the neighboring backbone NH groups of Ala232 and Val233. The interaction pattern stabilizes the backbone flip and makes a dual ladder of parallel hydrogen bonds to the bound ligand possible avoiding unfavorable secondary repulsive interactions among the closely approaching hydrogens in the hydrogen-bonding arrays [Immekus; Brenk *et al.*, 2003; Neeb *et al.*, 2014].

Detailed analyses of the protonation inventory overlaid to the binding event showed that N(5) of the *lin*-benzoguanine scaffold becomes protonated whereas N(3) remains in the neutral uncharged state [Neeb *et al.*, 2014].



#### 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme

**Figure 4.2** Binding modes of the investigated ligands to TGT (PDB IDs: **3a**, 4PUK; **3b**, 3S1G; **4a**, 4PUJ; **4b**, 4Q4R; **5a**, 4Q4O; **5b**, 4Q4P; **6a**, 4Q4S; **6b**, 4Q4Q). The protein is represented as cartoon; similar protein states with respect to helix  $\alpha$ A are colored identically (cp. Figure 7). The ligand and interacting residues are represented as sticks (nitrogen = blue, oxygen = red, sulfur = yellow). Chloride ions are shown as green spheres. For clarity, the  $\pi$ -stacking residues Tyr106 and Met260 are not shown, except when involved into the binding of the 2-substituent. Water molecules are also not displayed in the representation. The *lin*-benzopurine scaffold located in the guanine 34 binding pocket is well defined in the  $|F_o| - |F_c|$  difference electron density (green) at a sigma level between 2.5 and 3.0. The scaffold forms several interactions with Asp102, Asp156, Gln203, Gly230, Leu231, and Ala232 (dashed lines). The hydrogen bond to Leu231 is enabled by a ligand induced backbone-flip, which is stabilized by Glu235. The substituents are not equally well defined as the scaffolds suggesting residual mobility to some extent. Occupancies for the different substituents are given using the corresponding color of the ligand.

The *lin*-benzohypoxanthine scaffold adopts a very similar binding mode, and the interaction pattern to Asp156, Gln203, Gly230, Leu231, and Ala232 is analogously established [Barandun *et al.*, 2012]. The *lin*-benzohypoxanthines lack the exocyclic  $\text{NH}_2$  group at the pyrimidinone moiety. This results in different contacts with the adjacent Asp102. The latter residue orients its carboxylate group off from the binding site and remains in a geometry very similar to that found in the apo structure of the enzyme [Brenk *et al.*, 2003]. Instead of forming a direct contact to the bound ligand, Asp102 experiences hydrogen bonds to the carboxamide  $\text{NH}_2$  of Asn70 and the backbone NH group of Thr71 (Figures 4.2B, D, F, H). As a consequence of this outwards rotated pose of the Asp102 carboxylate group, a network formed by up to six adjacent water molecules is established involving the carboxylate groups of Asp102 and Asp156 [Barandun *et al.*, 2012]. As the parent ligand scaffold is no longer involved in a direct contact to Asp102, a small displacement of the ligand is observed resulting in slightly reduced distances to the remaining residues in contact with the *lin*-benzohypoxanthines. ITC measurements revealed that the *lin*-benzohypoxanthines bind to the protein without entrapping a proton at N(5) [Neeb *et al.*, 2014]. As a result, the *lin*-benzohypoxanthines bind significantly weaker than the *lin*-benzoguanines. The salt-bridge type hydrogen bonds to Asp156 formed by the *lin*-benzoguanines are replaced by more weakly charge-assisted contacts in the *lin*-benzohypoxanthines [Neeb *et al.*, 2014].

In previous studies [Ritschel *et al.*, 2009], it has been assumed that the substituents attached in 2-position to the *lin*-benzopurine scaffold of both parent structures exhibits large residual mobility in the bound state as no properly defined electron density was detected for this part of the ligands. Screening for modified crystallization conditions performed in the current study revealed for most of the studied ligands difference electron density in the crystals, which allows assignment of the 2-substituent in ordered geometry. Nonetheless, reduced population and elevated B-factors found for this part of the ligands suggest also in the recent structure determinations some residual mobility or scatter over multiple orientations.

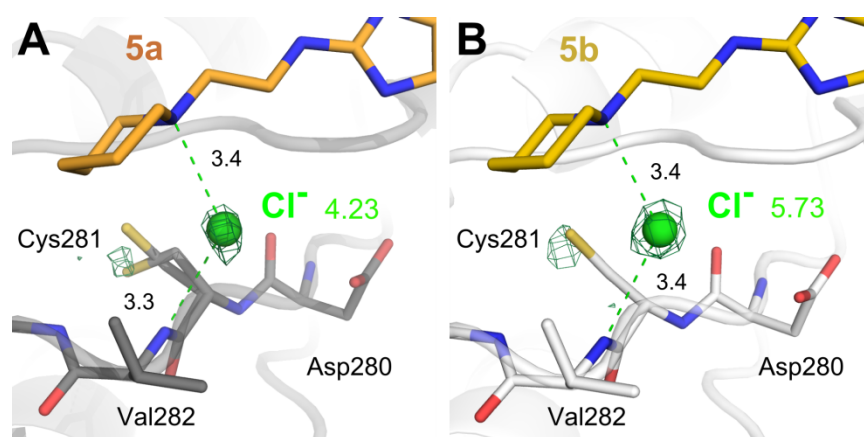
The difference electron density  $|F_o| - |F_c|$  fully defines the binding mode of **3a** and **3b** (Figure 4.2A, B). The 2-methyl group at both parent scaffolds points towards Val282 in the ribose-33 pocket (cp. Figure 4.3). The remaining part of the ribose-33 pocket is virtually unoccupied by the ligand due to the small size of the methyl group, instead several water molecules are found in the pocket.

The morpholinoethyl substituent of **4a** has been refined in the displayed *gauche* conformation to a population of 80%. The remaining tricycle converges to an occupancy of 100% (Figure 4.2C). A similarly reduced population (73%) is found for the 2-substituent in **4b**, even though the substituent adopts an *all-trans* conformation in this case (Figure 4.2D). In both examples, elevated B-factors are assigned to the morpholinoethyl substituents in comparison to the parent scaffolds (**4a**:  $B_{morpholine} = 21.5 \text{ \AA}^2$  vs.  $B_{tricycle} = 9.1 \text{ \AA}^2$ ; **4b**:  $B_{morpholine} = 26.4 \text{ \AA}^2$  vs.  $B_{tricycle} = 18.4 \text{ \AA}^2$ ), suggesting enhanced residual mobility of this part of the ligand. Although different conformers with slightly different placements in the ribose-33 pocket are experienced, the terminal ether oxygen of the morpholine ring forms expanded hydrogen bond contacts to Arg286 (**4a**: 3.2 Å; **4b**: 3.1 and 3.1 Å). In the latter case, refinement indicates the presence of two disordered conformers of the arginine side chain exhibiting equal population. The six-membered heterocycle exhibits a favorable chair conformation in **4a** and **4b**.

The nearly isostructural piperidinoethyl substituents of **5a** and **5b** are both well visible in the difference electron density with virtually identical poses matching that of the *all-trans* conformer in **4b** (Figure 4.2E, F). Whereas the substituent at the *lin*-benzoguanine scaffold refines to 64% occupancy, the one attached to the *lin*-benzohypoxanthine moiety is fully populated. Also for these ligands, the B-factors suggest enhanced residual mobility

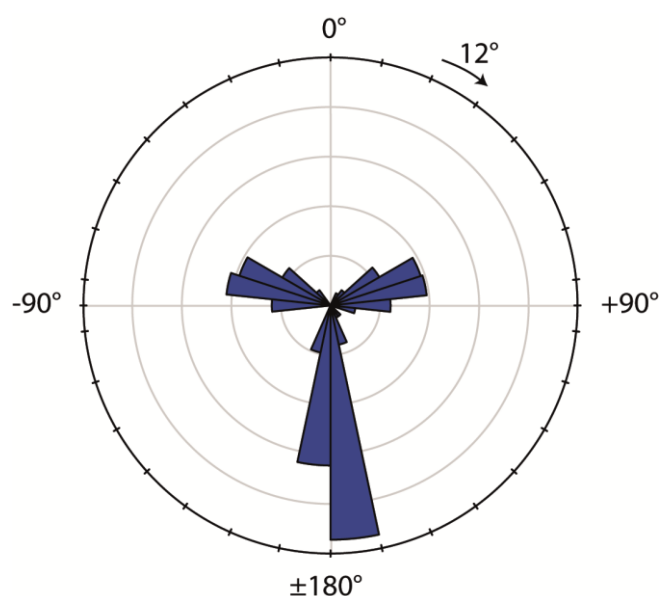
#### 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme

(**5a**:  $B_{\text{piperidine}} = 19.8 \text{ \AA}^2$  vs.  $B_{\text{tricycle}} = 13.1 \text{ \AA}^2$ ; **5b**:  $B_{\text{piperidine}} = 22.8 \text{ \AA}^2$  vs.  $B_{\text{tricycle}} = 17.2 \text{ \AA}^2$ ). In both structures, the piperidine ring adopts the energetically favored chair conformation. As a special feature, both ligands entrap a chloride ion at a position occupied in the other structures by a water molecule. Assignment of a chloride ion to this density peak was confirmed by evaluating the density maps for anomalous scattering using the program *ANODE* [Thorn and Sheldrick, 2011]. In the structure of **5b**, the anomalous density signal of  $\sigma = 5.73$  appears close to the signal of the structural zinc atom and the sulfur atoms of some methionines. In the complex TGT·**5a**, an anomalous signal of  $\sigma = 4.23$  is found. Most likely, the entrapping of the chloride ion correlates with the occupancy of the piperidine side chain, thus a reduced value for the chloride signal is reasonable in the latter structure. Refinement with Phenix [Adams *et al.*, 2010] converged to 69% chloride occupancy, well in agreement with the substituent population in **5a**. Crystal growth of TGT is performed under 1 M NaCl concentration; accordingly, the uptake of a chloride ion appears likely. This assumption is supported by the fact that the most likely protonated piperidine nitrogen ( $pK_a$  value  $\approx 11.2$  [Berrien *et al.*, 2007]) forms a hydrogen bond to the chloride ion. This ion is further stabilized through contacts to the backbone NH group of Val282 and an adjacent water molecule (Figure 4.3).



**Figure 4.3** Position of the captured chloride ion. The protein is represented as cartoon, the ligand and interacting residues as sticks (nitrogen = blue, oxygen = red, sulfur = yellow). Chloride ions are shown as green spheres. The anomalous signal calculated with the program *ANODE* is shown at a sigma level of 3.0. The chloride ion is clearly defined by the anomalous electron density interacting with the most likely protonated nitrogen of the ligand's piperidinoethyl substituent and the backbone NH of Val282. Occupancy refinement results in a similar value as found for the compound's substituent. In TGT·**5a** (PDB ID: 4Q4O), it is occupied to 100%, in TGT·**5b** (PDB ID: 4Q4P) to 69%. Peak intensities for the anomalous signal are given as green numbers.

In the complexes of **4a**, **4b**, and **5a** the substituent refines to reduced occupancy. The analysis of the difference electron density on a reduced sigma level suggests the presence of a minor populated *gauche* or *anti* conformer, respectively. Consulting the geometries found in small molecule crystal structures deposited in the CSD [Allen, 2002] reveals preferred angular distributions around  $\pm 60^\circ$  and  $180^\circ$  for an  $N^+-CH_2-CH_2-NH$  torsion fragment (Figure 4.4). The distribution suggests only minor preference for the *gauche* conformation. Accordingly, we also assume in the TGT complexes studied here that two arrangements are adopted.



**Figure 4.4** Rose plot of the CSD search results for the torsion angles of a  $N^+-CH_2-CH_2-NH$  linker as found in the substituents of **4** and **5**. The circles represent the frequency [%] of the matched torsion angle in steps of 20 percent. The position on the circumference gives the value of the torsion angle. The data set comprised 414 structures (updated CSD version 5.34; filters: not disordered, no errors, not polymeric, only organics). The carbon atoms were defined to be acyclic. Overall, *gauche* conformations are present in 39% of all analyzed structures, and the *anti* conformation is adopted in 61% of all cases.

Interestingly enough, the difference electron density does not disclose any orientation of the thiophenomethyl substituent in **6a** even though the better diffracting sulfur atom should be easier to detect. The electron density allows placement of the methyl group, however, the terminal thiophene moiety appears to be scattered over multiple orientations (Figure 4.2G). In contrast, the analogous *lin*-benzohypoxanthine derivative **6b**

discloses the thiophenomethyl substituent in two equally populated orientations (Figure 4.2H). The first conformer A orients the thiophene moiety into the ribose-33 pocket and experiences weak hydrophobic interactions with the side chain of Ala232. In the second conformation B, the thiophene ring rotates out of the ribose-33 pocket and approaches Tyr106. A weak hydrogen bond between the thiophene sulfur atom and the Tyr106 OH is established (3.1 Å). Also in this structure, the B-factors of the thiophene substituent refine to a value significantly larger as those of the parent scaffold ( $B_{thiophene A} = 27.5 \text{ \AA}^2$ ,  $B_{thiophene B} = 21.7 \text{ \AA}^2$  vs.  $B_{tricycle} = 11.2 \text{ \AA}^2$ ).

#### 4.4.3 Correlation of Thermodynamic Signatures and Binding Modes

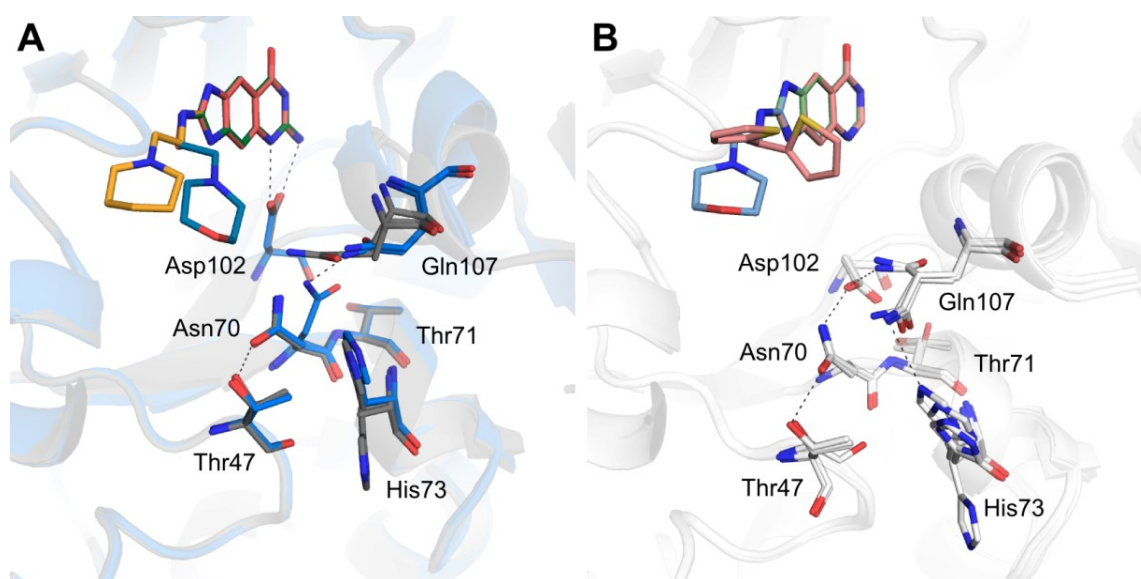
As described above, the *lin*-benzoguanines **3a** – **6a** form, upon protonation at N(5), two bidentate salt bridges to Asp102 and Asp156. In contrast, the *lin*-benzohypoxanthines **3b** – **6b** do not induce a similar reorientation of Asp102 and the carboxylate group of this residue remains in its orientation pointing away from the binding pocket. In consequence, the  $pK_a$  shift at N(5) of the *lin*-benzohypoxanthine moiety is not provoked and one of the salt bridges assigned in case of the *lin*-benzoguanines to the guanidinium portion is lost. The second one transforms into two most likely attenuated charge-assisted hydrogen bonds to Asp156 [Neeb *et al.*, 2014]. As they lack secondary repulsive interactions [Jorgensen and Pranata, 1990; Murray and Zimmerman, 1992], they can still contribute significantly. The thermodynamic and structural data recorded for thrombin-ligand binding have shown that the loss of a bidentate salt-bridge is equivalent to about  $-25 \text{ kJ} \cdot \text{mol}^{-1}$  in  $\Delta\Delta H^0$  [Baum *et al.*, 2009]. This value is approximately matched by the congeneric pairs **5a/5b** and **6a/6b** whereas **3a/3b** and **4a/4b** show significantly larger enthalpy differences. This finding indicates that the thermodynamic signatures in the current compound series are determined by additional effects superimposed on the sole loss of a salt bridge.

Comparing the structural data of the apo protein with those of the *lin*-benzohypoxanthine complexes, it becomes evident that Asp102 virtually remains in its original conformation upon ligand binding (Figure 4.5B) [Brenk *et al.*, 2003; Barandun *et al.*, 2012]. Also the network of water molecules found in the apo structure is maintained and contributes contacts between the protein and the accommodated ligands. Accordingly, the *lin*-benzohypoxanthines **3b** – **6b** bind to the enzyme without inducing major changes



compared to the apo structure, which might require a strong enthalpic price to be paid. Notwithstanding, overall, a significant enthalpy-driven binding is observed for the *lin*-benzohypoxanthine series.

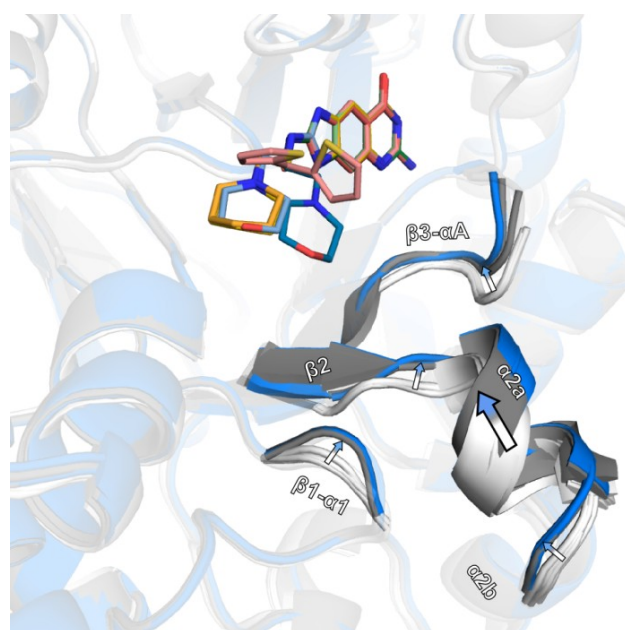
The structural situation is different for the *lin*-benzoguanines. Their binding triggers the rearrangements of Asp102 towards the binding pocket. This movement narrows the binding pocket, and the network of water molecules is displaced from the binding site in contrast to the binding of the *lin*-benzohypoxanthines. Subsequent to the reorientation of Asp102 and as a consequence of the loss of interactions with the adjacent residues such as Asn70, a cascade of structural rearrangements is induced, which results in a collapse of the interaction patterns among Asp102, Asn70, Thr47, and Gln107 previously seen in the apo protein (Figure 4.5) [Brenk *et al.*, 2003; Stengl *et al.*, 2007].



**Figure 4.5** Cascade of residue rotation initiated by the movement of Asp102. The same representation was applied as noted for Figure 4.2. For crystal structure analysis, all complexes containing *lin*-benzoguanines (A) and *lin*-benzohypoxanthines (B) were aligned, respectively. A) The side chains of Thr47, Asn70, His73, Gln107, and Asp102 are involved in the domino effect. All complex structures with *lin*-benzoguanine-based ligands show a high flexibility of the involved amino acid residues. Due to the missing interaction to Asp102, Asn70 is not fixed in an ordered conformation interacting either with Thr47 or Gln107. Gln107 can be found in multiple conformations, indicated by a partly missing electron density for its side chain. The previously found interaction to His73 cannot be detected in any of the complex structures under investigation. B) In binary complexes with *lin*-benzohypoxanthine-type inhibitors, residue Asp102 serves as an anchor point fixing the side chain of Asn70 in a position competent to build hydrogen bonds to Thr47 similar to the apo enzyme. The side chain of Gln107 shows different conformations capable to interact either with His73 or Asp102.

#### 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme

Even some movements of secondary structural elements are observed compared to the apo protein and the *lin*-benzohypoxanthine complexes (Figure 4.6). Accordingly, the binding of the *lin*-benzoguanines involves a fair number of adaptation processes, which all will be linked to significant changes of the thermodynamic properties. The formation of the salt bridges was already mentioned as an enthalpically driven step. The displacement of ordered water molecules from a protein binding site usually results in an entropically favorable binding signal [Baum *et al.*, 2009; Biela *et al.*, 2012].



**Figure 4.6** Influence of Asp102 on secondary structure elements of TGT. An overlay of all investigated structures of the TGT-bound *lin*-benzohypoxanthines (white) and *lin*-benzoguanines (dark gray and blue) shows substantial differences in the position of the  $\beta$ 3- $\alpha$ A-loop,  $\beta$ -strand 2, helix  $\alpha$ 2A, and the  $\beta$ 1- $\alpha$ 1-loop. The binding of members of the *lin*-benzohypoxanthine series has only a slight effect on the global protein structure, which stays closely to the state of the apo enzyme. Contrary, due to the up-rotation of Asp102 towards the ligand, members of the *lin*-benzoguanine series trigger a kind of domino effect by which the  $\beta$ 3- $\alpha$ A-loop,  $\beta$ -strand 2, helix  $\alpha$ 2A, and the  $\beta$ 1- $\alpha$ 1-loop move closer to the guanine 34 binding pocket.

In **4a**, **4b**, **5a**, **6a**, and **6b**, a residual mobility of the 2-substituent is observed that might suggest an entropically favorable contribution to binding. Most likely, these effects are beneficial for ligand binding but they will be partly compensated by the rearrangements crystallographically observed in the protein. They will cost a price in either enthalpy or

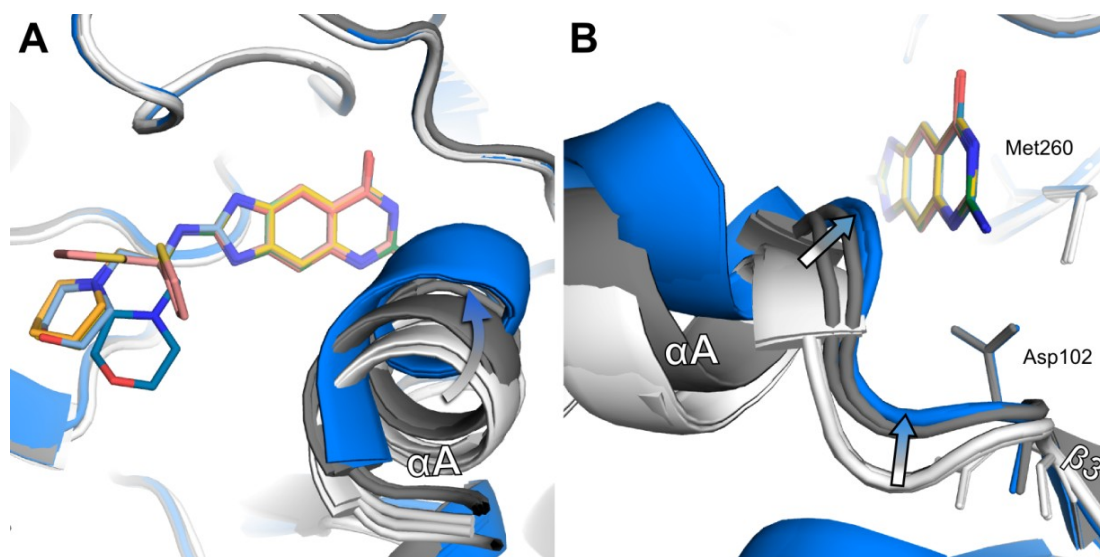
entropy. Nonetheless, overall a net improvement in the Gibbs free energy of binding of about  $-6 \text{ kJ} \cdot \text{mol}^{-1}$  is observed for the *lin*-benzoguanines over the *lin*-benzohypoxanthines. Detailed factorization appears difficult in the present case, also as the price for the desolvation will be distinct for the *lin*-benzoguanine and *lin*-benzohypoxanthine series.

The detailed comparison of the thermodynamic profiles of the *lin*-benzoguanines **3a** – **6a** and *lin*-benzohypoxanthines **3b** – **6b** unravels another striking feature (Figure 4.1). Whereas the *lin*-benzohypoxanthines factorize quite similarly in enthalpy and entropy, the members of the *lin*-benzoguanine series exhibit large differences in their profiles. This observation suggests that the *lin*-benzohypoxanthines do not experience large differences in their thermodynamic properties. Thus, hardly any dependence on the actual size and chemical properties of the substituents attached at the 2-position is observed. This picture appears different for the *lin*-benzoguanines as here the thermodynamic profile is significantly modulated by the properties of the attached 2-substituent. Obviously, in the latter compound series a crosstalk between the rearrangement of Asp102 and the binding to the ribose-33 pocket is given, whereas across the *lin*-benzohypoxanthine series this interdependence can hardly be recognized.

To find some explanations for these differences in the pocket crosstalk, we have analyzed the crystal structures of the various complexes. Interestingly enough, binding of the *lin*-benzoguanines, which recruits Asp102 for ligand recognition translates into a significant shift of helix  $\alpha\text{A}$  (Figure 4.7). This helix adopts virtually the same geometry in the apo protein and the *lin*-benzohypoxanthine-bound complexes. It is placed nearly perpendicular to the guanine-34 recognition site and extends from Tyr106 to Leu110. Once the *lin*-benzoguanines are bound to the enzyme, this helix is significantly shifted towards the ligand-binding site. This movement is triggered by the rearrangement of Asp102. A comparison of the enzyme structure with the bound substrate tRNA and the discussed *lin*-benzoguanine inhibitor complexes is quite conclusive. Obviously, also the binding of the tRNA induces this shift of helix  $\alpha\text{A}$  very similar to that observed in the *lin*-benzoguanine complexes but clearly distinct from the geometry found for the apo protein and the *lin*-benzohypoxanthine complexes. A predominant role can be assigned to Tyr106, which resides on this helix. Upon the shift of the helix, the Tyr106 side chain penetrates by  $2.5 \text{ \AA}$  deeper into the ribose-33 pocket. This replacement allows Tyr106 to experience whether any ligand portion is bound to the ribose-33 pocket. In the apo protein and the structurally very similar *lin*-benzohypoxanthine

#### 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme

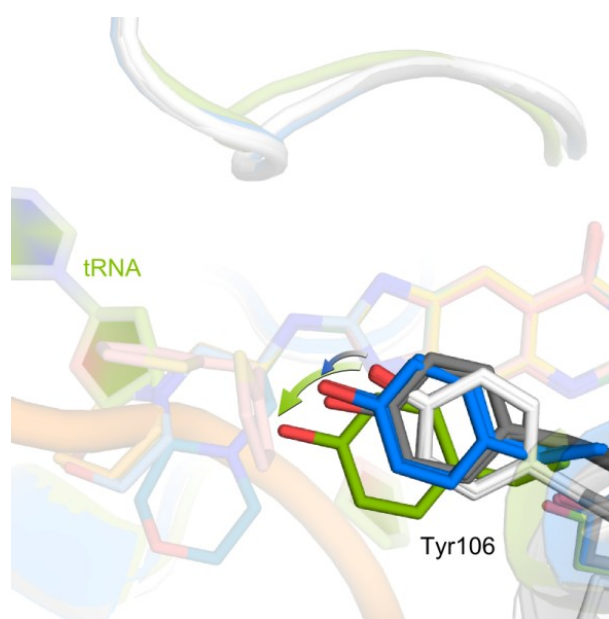
complexes, Tyr106 remains in a remote position, but for the *lin*-benzoguanine complexes, Tyr106 is pushed to varying extent (1.3 – 1.9 Å) into a proximate position bordering the ribose-33 pocket (Figure 4.8).



**Figure 4.7** Influence of the various ligands on helix  $\alpha A$ . The protein is represented as cartoon. The ligands are shown as sticks and colored according to Figure 4.2. A) All binary TGT complexes with *lin*-benzohypoxanthine-based inhibitors exhibit one fixed helix conformation (white). *lin*-Benzoguanine-TGT complexes show different conformations of the helical part: TGT-**3a** (green), TGT-**4a** (blue), and TGT-**5a** (orange) show a similar position among each other, which deviate from that in the *lin*-benzohypoxanthine-TGT complexes (gray). An even larger perturbation results from the binding of **6a** which leads to strong repositioning of Gln107, Val108, and Met109 concomitant with an up-rotation of the helix (blue). Additionally, TGT-**4a** shows a second populated binding mode with an up-rotated helix conformation as similarly found in complex TGT-**6a** (blue). B) Modulations of the backbone trace of the  $\beta 3$ - $\alpha A$  loop upon formation of the different complexes. *lin*-Benzohypoxanthine complexes are displayed in white, TGT-**3a**, TGT-**5a** in gray, and TGT-**4a**, TGT-**6a** in blue. All white structures show a conserved orientation of helix  $\alpha A$ . It is noticeable that the backbone of the  $\beta 3$ - $\alpha A$  loop adopts a conformation distal to the tricyclic ligand core. In *lin*-benzoguanine-TGT complex structures, the whole loop region is positioned closer to the ligand's scaffold. Thereby, the closest position can be found, if helix  $\alpha A$  is rotated upwards (TGT-**4a**, TGT-**6a**). TGT-**5a** adopts an intermediate orientation since Tyr106 indicates that an up-rotated helix conformation could be possible, although not observed in electron density.

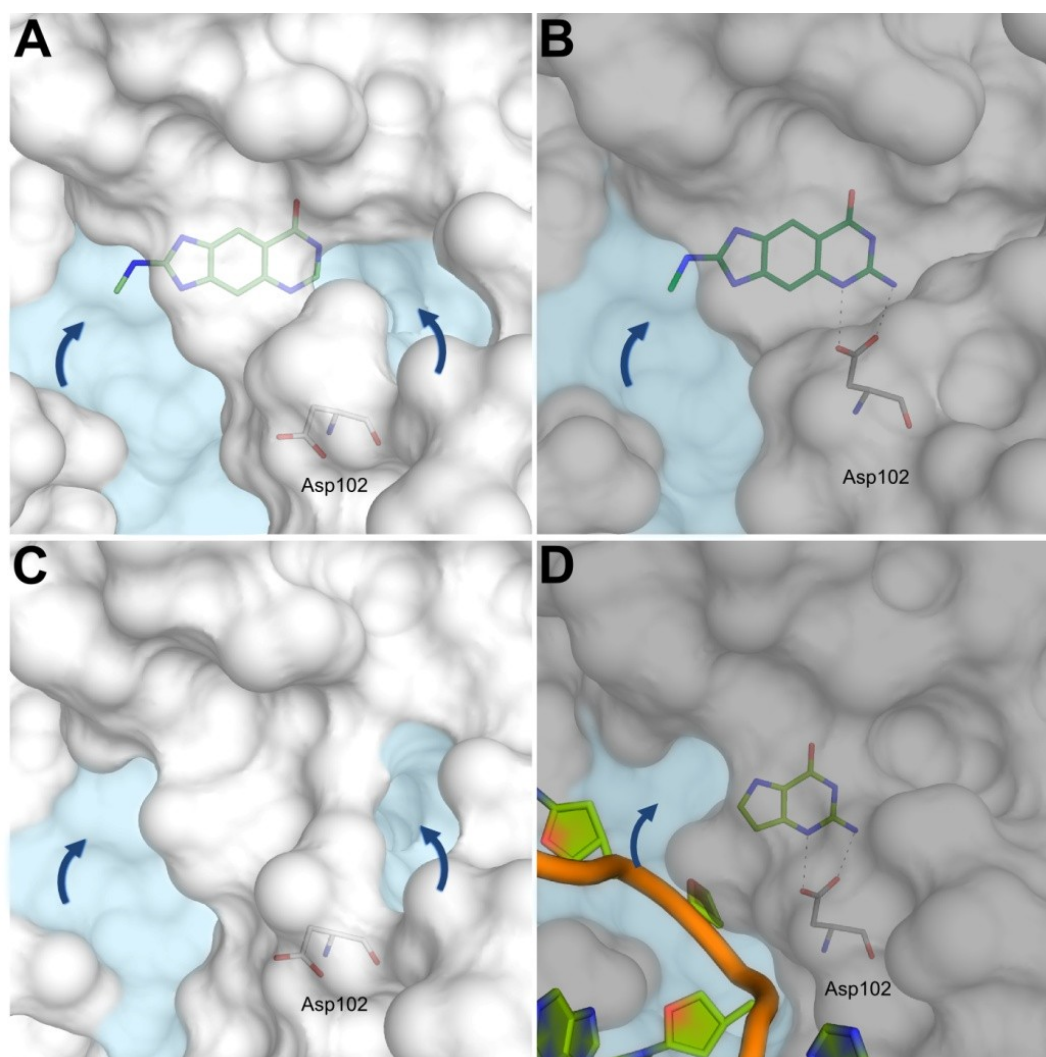
Obviously, the *lin*-benzohypoxanthines bind to TGT in a conformation that closely resembles that of the apo protein whereas the *lin*-benzoguanines provoke an arrangement that approximates the architecture of the enzyme in the substrate-bound state as observed

in the covalent tRNA intermediate (PDB ID: 1Q2R) trapped by crystallization with the substrate mimetic 9-deazaguanine (Figure 4.9). This finding explains why the thermodynamic binding signature of the latter inhibitors is more sensitive to the properties of the attached 2-substituent as the presence of such a substituent is recognized and thus transmitted via Tyr106 and the shifted helix  $\alpha$ A. With respect to the enzymatic function, such a behavior appears very reasonable as the enzyme must distinguish what is bound to the ribose-33 recognition pocket. Discrimination between correct and false substrates relies on the recognition of the nucleobase in this pocket. Accordingly, the binding properties of this part of the substrate must be transmitted and have to result in a varying thermodynamic binding profile. Exactly this feature we also observe between the *lin*-benzohypoxanthine and *lin*-benzoguanine-type inhibitor series, which gives rise to the rather complex structure–activity relationship between both series.



**Figure 4.8** Overlay of the TGT structure binding the natural substrate tRNA (PDB ID: 1Q2R) as well as the analyzed *lin*-benzoguanine and *lin*-benzohypoxanthine-bound structures. Helix  $\alpha$ A as well as the tRNA are visualized in cartoon representation (*lin*-benzohypoxanthines = white, *lin*-benzoguanines = gray/blue, 1Q2R = green). The ligands, nucleic acid bases, and Tyr106 are highlighted as sticks (nitrogen = blue, oxygen = red). The  $\alpha$ -helix A of the TGT-tRNA complex adopts an orientation similar to the *lin*-benzoguanine-bound protein. In the same way, Tyr106 of the latter complexes reaches deeper into the ribose-33 subpocket. Obviously, these positional changes induce a cooperativity between the guanine-34 recognition site and the ribose-33 subpocket leading to the observed diverse thermodynamic profiles.

#### 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme



**Figure 4.9** Influence of substrate recognition on the guanine-34 recognition site. As representatives, the complexes TGT-3b (A), TGT-3a (B), the apo enzyme at pH 7.8 (C), and the ternary complex TGT-9-deazaguanine-tRNA (D) are illustrated. The latter complexes mimic an intermediate situation during the nucleobase exchange reaction when tRNA is covalently attached and guanine is cleaved off. This situation is captured in the cocrystallized ternary complex using the geometrically identical but chemically unreactive 9-deazaguanine. The protein is shown in surface representation (white = apo state, gray = substrate bound state). The ligands are displayed as sticks, the tRNA stem loop as cartoon (nitrogen = blue, oxygen = red, sulphur = yellow, phosphorus = orange). In complex TGT-3b (PDB ID: 3S1G, A), the *lin*-benzohypoxanthine is not recognized as substrate-like by Asp102 as the exocyclic NH<sub>2</sub> group is missing. Asp102 adopts a conformation remaining within the protein similar to apo TGT (PDB: 4PUN, C). Therefore, they share a similar protein surface (white) exhibiting two entries to the active site indicated by arrows. In complexes TGT-3a (PDB ID: 4PUK, B) and TGT-9-deazaguanine-tRNA (PDB ID: 1Q2R, D), the *lin*-benzoguanine and the substrate mimetic 9-deazaguanine, respectively, are recognized as substrate surrogates by the enzyme. Asp102 rotates towards the guanine-34 recognition site and forms hydrogen bonds to the aminopyrimidinone moiety of the ligand. Provoked by the structural changes due to the up-rotation of Asp102, the second entrance to the guanine-34 recognition site is closed (gray).

We mentioned above the deviating potency profiles obtained by the biochemical assay and the ITC data, particularly for the *lin*-benzohypoxanthine series (differences in  $K_i$  values up to a factor of 16 for **3b**). ITC measures directly the binding to the protein conformer competent to accommodate a ligand, yielding  $K_d$  values. The biochemical assay requires the protein to be in the state able to recognize the tRNA substrate, thus with Asp102 oriented towards the binding pocket. Only the *lin*-benzoguanines will bind to this conformation, not the *lin*-benzohypoxanthines. Thus, a conformational change would be required to accommodate the latter ones. However, as a size-dependent inhibition model has to be assumed for the ligands binding to the enzyme, small inhibitors can show a significant inhibitory component even if the intermediately bound tRNA is attached to the protein. Supposedly, this inhibitory component can only be realized by the *lin*-benzoguanines and not by the *lin*-benzohypoxanthines. Possibly, this effect also has influence on the inhibitory properties determined by the biochemical assay or ITC and underestimates the intrinsic binding potency of the *lin*-benzohypoxanthines in the biochemical assay.

## 4.5 Conclusions

The two congeneric series of inhibitors, varied at the pyrimidinone moiety and the 2-amino substituents of a *lin*-benzopurine scaffold, show a rather flat and unattractive SAR with nearly unchanged binding affinities within the series and an offset of about one order of magnitude between both series. The loss in affinity between the *lin*-benzoguanine and the *lin*-benzohypoxanthine series can be related at first glance to the loss of a salt bridge formed to the pyrimidinone moiety with and without an attached amino group. The unmodulated affinity recorded for the attachment of a 2-amino substituent at the remote end of the parent *lin*-benzopurine scaffold can hardly be explained on the affinity level. The thermodynamic signature accessible by ITC measurements provides some surprises as in the *lin*-benzohypoxanthine series all ligands factorize similarly in enthalpy and entropy whereas the *lin*-benzoguanines show strong and complex changes in the enthalpy/entropy signature. Interestingly enough, the crystallographic analysis shows that the *lin*-benzohypoxanthines bind to the protein without inducing major conformational changes of the enzyme

#### 4. Beyond Affinity: Enthalpy–Entropy Factorization Unravels Complexity of a Flat Structure–Activity Relationship for Inhibition of a tRNA-Modifying Enzyme

compared to the apo structure. With respect to the thermodynamic properties, small or large 2-substituents of deviating chemical nature experience the same hardly discriminating profile. The *lin*-benzguanines bind to a protein conformation approximating the enzyme in the substrate-bound geometry. Major conformational changes are all triggered by the reorientation of Asp102. This residue is important for substrate recognition. Simultaneously, the protein architecture varies in a way that a crosstalk between the binding signature resulting from the ligand portion accommodating the guanine-34 and the ribose-33 pocket is established. In consequence, quite surprisingly, the attached 2-substituents at the *lin*-benzguanine scaffold receive an individually modulated thermodynamic signature across the compound series.

The example shows, even though a detailed partitioning in individual thermodynamic contributions is impossible due to the given complexity of the system, that a combination of thermodynamic signature analysis and structural information provides an insight into structure–activity relationships beyond sole affinity considerations. They are important for a rational and predictive design of novel inhibitors.



## 5. Creating a Resistance Model for TGT: the Effect of Mutations on Flexible *lin*-Benzoguanine Substituents

### 5.1 Introduction

After the successful design of high-affinity inhibitors for a distinct target, which exhibit favorable bioavailability features, the development of drug resistance due to active-site as well as non-active-site mutations often constitutes a painful backlash [Erickson and Burt, 1996]. Such mutations can either result in a decreased binding affinity of the ligand or an enhanced activity of the enzyme overcoming the inhibitory effect [Luque *et al.*, 1998].

In order to tackle this problem investigations were directed towards the analysis of the enthalpic and entropic properties of various inhibitors and how their binding is influenced by the appearance of resistance mutations [Velazquez-Campoy *et al.*, 2000; Velazquez-Campoy *et al.*, 2001; King *et al.*, 2012]. Data collected for HIV-protease inhibition suggested that ligands binding with a mainly entropic contribution to the target enzyme are more susceptible to variations in the amino acid sequence. It could be shown in case of HIV-protease that favorable entropy could mainly be achieved by rigidifying the inhibitor in a way that it optimally fits the active site of the enzyme. As a consequence, the ligand loses less conformational degrees of freedom upon binding. Simultaneously, however, its ability to escape mutational variations within the binding site is significantly reduced. Contrary to rigid molecules, flexible compounds were found to respond to mutational perturbations at critical positions of the target protein with a less pronounced decrease in binding affinity. As a strategy to overcome the loss in conformational entropy upon binding of a previously flexible ligand, an enthalpic compensation was conceived. To achieve this goal, polar groups were introduced to the molecules, which were assumed to form strong hydrogen bonds or salt bridges, respectively. As a result not only enthalpy-entropy compensation was partly avoided but also the binding specificity to the target enzyme could be increased.

As a model system to investigate the behaviour of different ligands towards mutations, we selected the tRNA modifying enzyme tRNA-guanine transglycosylase, which plays a key role in the pathogenicity of *Shigella* bacteria, the causative agent of Shigellosis [Sansone, 2001]. Disappointingly, the treatment of the disease becomes more and more

challenging due to the increasing resistance development of *Shigella* to current antibiotics [Ashkenazi *et al.*, 2003]. Therefore, the design of new selective antibiotics inhibiting TGT is of utmost importance. In the present study, we chose 2-amino-*lin*-benzoguanine type TGT inhibitors as model compounds to test their resistance tolerance in TGT binding [Hörtner *et al.*, 2007].

As reported in chapter 3 and 4 these ligands exhibit pronounced enthalpic binding due to various favorable van der Waals interactions, hydrogen bonds and salt bridges formed between the protein and the ligand [Neeb *et al.*, 2014]. Crystal structure analysis revealed that the substituents attached in 2-position of the parent *lin*-benzoguanine scaffold exhibit enhanced residual mobility indicated by increased B-factors compared to the parent scaffold and by the presence of multiple conformations of the 2-substituents [Neeb *et al.*, 2014]. This flexibility of the 2-substituents might be advantageous as ligand portions capable to bind with multiple orientations might be better suited to cope and tolerate resistance mutations of the target protein. We hypothesized that such ligands or ligand portions are potentially more appropriate for targeting mutated TGT variants since no specific interactions with any protein side chains are required and the ligand can still find an acceptable binding pose.

Overall, this study was intended to show whether a ligand exhibiting high residual mobility is better able to avoid a loss in binding affinity upon a target protein mutation than a ligand adopting only one ordered binding mode.

## 5.2 Results

### 5.2.1 Mutations of Wild Type TGT

The simulation of resistance development requires the selection of amino acids appropriate for mutagenesis. Since resistance mutations usually do not affect residues, which are directly involved in catalysis, we focused on amino acids present in the uridine-33 subpocket. The number of exposed side chains in this pocket is limited due to the fact that the area of this pocket, which is located next to the ligand's substituent, mainly contains backbone atoms and glycine residues (Figure 5.1A). Accordingly, Val262 was chosen for mutational analysis, and replaced by threonine, aspartate and cysteine, respectively. The presence of each

mutated codon was confirmed by sequencing and mass spectrometry (see appendix). Subsequently, the putative influence of the respective mutations on the binding mode of the ligand were analyzed by MD simulations (not shown), and the steric and electronic consequences of the newly introduced side chains were investigated by X-ray crystallography.

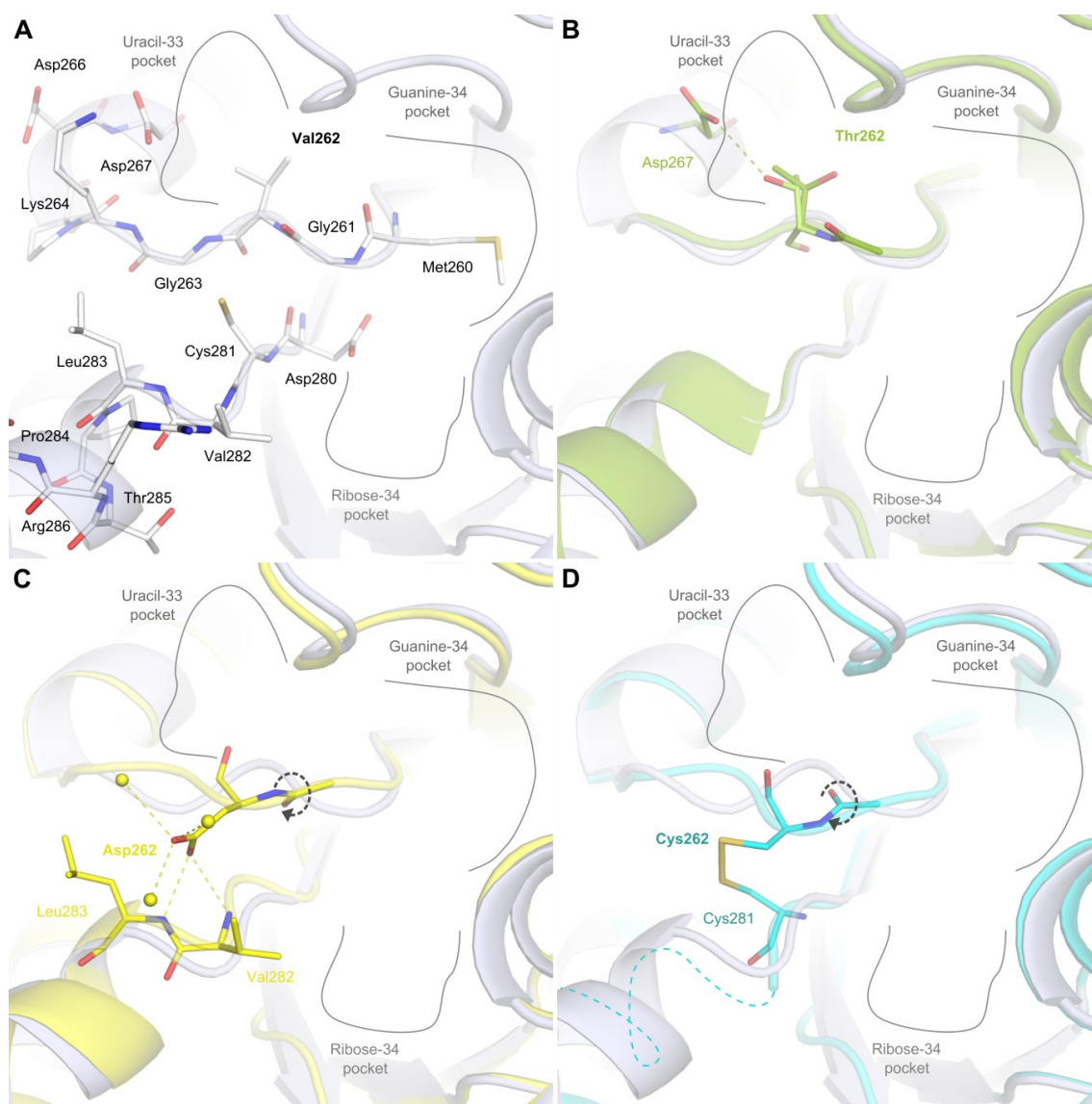
The apo structure of TGT(Val262Thr) was determined at a resolution of 1.24 Å (Figure 5.1B, green). The mutated threonine residue adopts a conformation similar to that of Val262 in the wild type enzyme. In addition, a second orientation is indicated by the difference electron density, which could, however, not satisfyingly be resolved during refinement. Yet, in this alternative position Thr262 seems to be slightly shifted into the ribose-33 subpocket and most likely forms, via its side chain hydroxyl, a hydrogen bond to the side chain carboxylate of the adjacent Asp267. The solvation pattern within the pocket is similar to the one observed in the apo structure of the wild type.

The apo structure of TGT(Val262Asp) shows a nominal resolution of 1.33 Å (Figure 5.1C, yellow). The mutated aspartate is well-defined by the electron density. Compared to the apo structure of the wild type, the peptide bond between Gly261 and Asp262 is flipped. As a consequence, the backbone of Asp262 reaches into the ribose-33 subpocket with its side chain interacting with the backbone NH groups of Val282 (3.08 Å) and Leu283 (2.79 Å) as well as three close-by water molecules (2.79 – 3.05 Å). To avoid unfavorable contacts to the side chain of Asp262 the positions of Asp280, Cys281, Val282 and Leu283 are slightly altered compared to the wild type.

The nominal resolution of the TGT(Val262Cys) structure amounts to 1.44 Å (Figure 5.1D, cyan). Surprisingly, the crystal structure shows a disulfide bridge, which is partially formed between Cys262 and Cys281 despite of the presence of DTT as reducing agent during crystallization. Similar to TGT(Val262Asp) the peptide bond between Gly261 and Cys262 is flipped. The formation of the disulfide bridge is concomitant with the rearrangement of several amino acids resulting in the dislocation of residues Leu283 to Arg289.

Kinetic characterization of the mutated variants revealed both for TGT(Val262Thr) and TGT(Val262Asp) a turn-over number which was, compared to the wild type enzyme, only slightly reduced (see Appendix 10.6). Therefore, and as the Val262Thr and the Val262Asp mutations each led to different alterations of the uridine-33 subpocket, they provide a good platform to investigate various resistance scenarios.

## 5. Creating a Resistance Model for TGT: the Effect of Mutations on Flexible lin-Benzoguanine Substituents



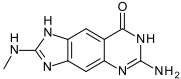
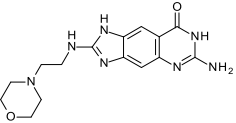
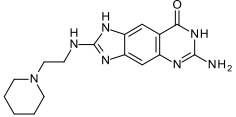
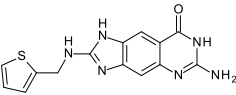
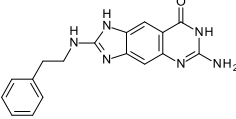
**Figure 5.1** Apo structures of the analysed mutated variants of TGT. The protein is shown in cartoon representation. The mutated residues as well as adjacent amino acids possibly affected by the respective mutation are depicted as sticks (nitrogen = blue, oxygen = red, sulphur = yellow). For clarity, water molecules, which are not in contact with the mutated residues are not shown. Interactions formed by the mutated residue with its local environment are visualized as dashes. For comparison, the structure of the wild type enzyme (transparent) is superimposed in images B – D. A) Apo structure of wild type TGT (PDB ID: 1PUD [Romier *et al.*, 1996]). The uridine-33 subpocket is mainly bordered by backbone atoms restricting possible sites appropriate for site-directed mutagenesis. B) TGT(Val262Thr). This variant shows a similar structure as the wild type enzyme. Accordingly, Thr262 adopts an orientation similar to that of the wild type, although a second conformation interacting via its hydroxyl group with the side chain of Asp267 is weakly indicated by the difference electron density. C) TGT(Val262Asp). Asp262 interacts with the backbone NH groups of Val282 (3.08 Å) and Leu283 (2.79 Å) resulting in the positional deviations of residues Gly261 and Asp262 as well as Asp280 to Leu283 from the wild type. D) TGT(Val262Cys). Cys262 forms a disulfide bond with Cys281 leading to a similar backbone orientation of Gly261 and Cys262 as observed in TGT(Val262Asp). Additionally, the formation of the disulfide bridge results in the dislocation of residues Leu283 to Arg289 (dotted cyan line).

Unfortunately, the Val262Cys mutant did not yield kinetics that could be evaluated conclusively in a Michealis-Menten plot despite of its present activity. Presumably, the partially formed disulfide bridge between Cys262 and Cys281 caused the inconsistent data. Nonetheless, the mutated variant still exhibits reasonable kinetic activity. In order to produce reliable kinetic results for this variant, TGT(Val262Cys) has to be expressed, purified and crystallographically analyzed in the absence of oxygen and under inert gas conditions that guarantee Cys262 and Cys281 to be present in their reduced state.

### 5.2.2 Binding Affinities

To address the question whether a ligand showing residual mobility constitutes a more resistance tolerant antibiotic, the  $K_d$  values of different 2-amino-*lin*-benzoguanines were measured against the three generated mutants using isothermal titration calorimetry. The obtained values were compared to the data obtained for the wild type protein (Table 5.1).

**Table 5.1** Binding affinities of *lin*-benzoguanine type inhibitors to mutated TGT variants as determined by isothermal titration calorimetry at least in triplicates.

Ligand	Chemical formulae	$K_d$ [nM]			
		WT	Val262Thr	Val262Asp	Val262Cys
<b>3a</b>		49 ± 5	44 ± 5	158 ± 5	16 ± 1
<b>4a</b>		32 ± 7	27 ± 7	181 ± 17	18 ± 4
<b>5a</b>		34 ± 7	34 ± 4	152 ± 9	---[a]
<b>6a</b>		81 ± 8	63 ± 15	557 ± 20	27 ± 2
<b>7</b>		57 ± 14	36 ± 5	359 ± 18	19 ± 2

<sup>a</sup> $K_d$  for the Val262Cys variant was measured against the oxidized protein and is thus not listed.

Thereby, **3a** served as a reference as, due to the small size of its 2-substituent, it hardly occupies the ribose-33 subpocket.

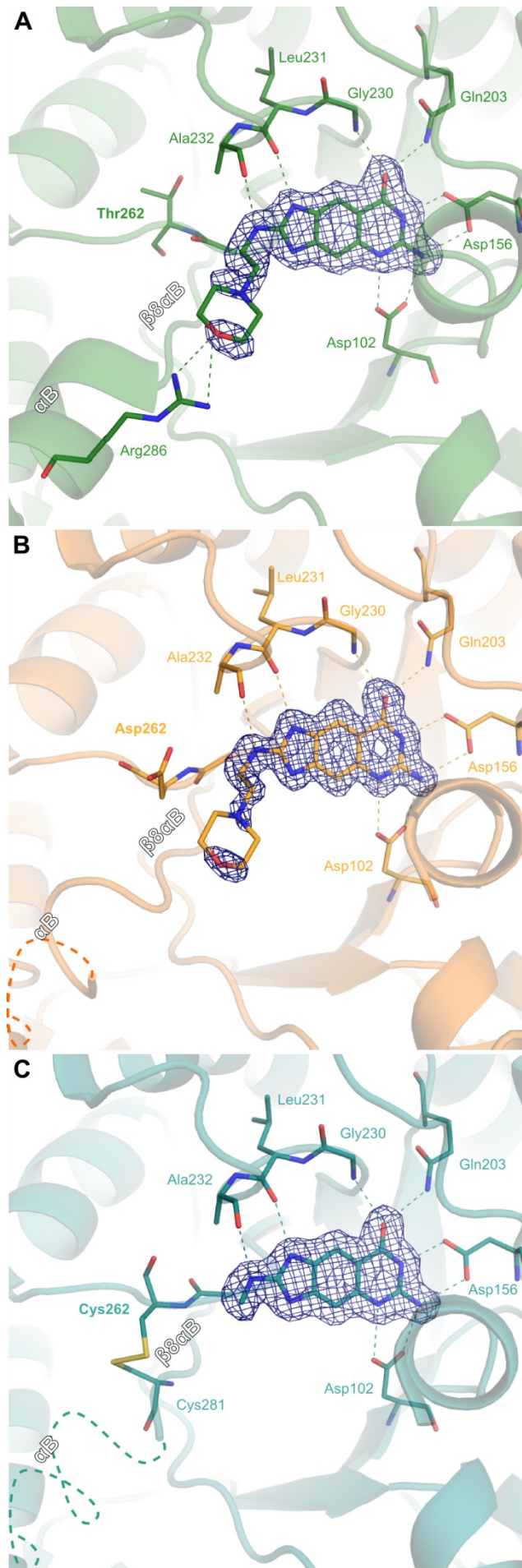
Interestingly, all ligands respond to the introduced mutations in a similar way, independently of the respective 2-substituent. As expected, the binding affinities for the TGT(Val262Thr) resemble, within the range of experimental accuracy, those obtained for the wild type enzyme. In this variant only minor changes in the binding pocket were introduced. In contrast, the binding affinities to TGT(Val262Asp) are consistently decreased by a factor of three to seven. Remarkably, all 2-amino-*lin*-benzoguanines show a two- to threefold increase in potency towards TGT(Val262Cys).

### 5.2.3 Exploring the Adaptations of **4a** Targeting the Mutated TGT Variants

As a representative for the 2-amino-*lin*-benzoguanine type ligands, **4a** was chosen and its binding mode with the mutated TGT variants was investigated by crystal structure analysis. We focused on the response of ligand binding on the modified protein environment as well as on the orientation of the 2-substituent within the altered uridine-33 subpocket.

The structure of TGT(Val262Thr) in complex with **4a** shows a maximum resolution of 1.89 Å. The parent *lin*-benzoguanine scaffold adopts the previously described binding mode involving the  $\pi$ -stacking residues Tyr106 and Met260 as well as hydrogen bonds to Asp102, Asp156, Gln203, Gly230, Leu231 and Ala232 (Figure 5.2A). As similarly observed in the corresponding complex structures of the wild type enzyme, the 2-substituent is positioned within the ribose-34 subpocket adopting a *gauche*-conformation. The morpholine ring is weakly defined most likely showing the favorable chair conformation. Its occupancy amounts to 82%. Via its morpholino oxygen the substituent interacts with the side chain of Arg286. Contrary to the apo structure of TGT(Val262Thr), no difference electron density indicating a second conformation of Thr262 is observed. The side chain of Thr262 is solely present in an orientation similarly found for the valine side chain in wild type TGT.

The binary complex of TGT(Val262Asp) with **4a** was determined at a resolution of 1.38 Å. As expected, the *lin*-benzoguanine scaffold as well as its 2-substituent show the same interaction pattern as observed in previous structures (Figure 5.2B). The morpholine ring is weakly defined by the electron density. It is refined in its chair conformation to an occupancy of 69%.



**Figure 5.2** Binding mode of **4a** as bound to the different mutated TGT variants. The protein is displayed in cartoon representation. The ligand as well as interacting protein residues are highlighted as sticks (nitrogen = blue, oxygen = red, sulphur = yellow). Hydrogen bonds are visualized as dashes. Delocalized protein residues are indicated by dotted lines. The  $2|F_o| - |F_c|$  density is shown at a *sigma* level of 1.0 in blue. For clarity, water molecules are not shown. A) Binding mode of **4a** to TGT(Val262Thr). The same interaction patterns are experienced by the *lin*-benzoguanine scaffold as in the wild type. The 2-substituent shows a *gauche* conformation forming a hydrogen bond to Arg286 at a distance of 3.0 Å. Its occupancy amounts to 82%. Thr262 adopts a conformation similar to that of Val262 in the wild type protein. B) Binding mode of **4a** to TGT(Val262Asp). The *lin*-benzoguanine scaffold forms the same interactions as previously described. Its 2-substituent binds in a *gauche* conformation and is refined to an occupancy of 69%. The side chain of Arg286, usually interacting with the terminal ether oxygen of the morpholine ring, is not resolved in the electron density. Upon binding of the *lin*-benzoguanine scaffold Ser287 and Gly288 get completely disordered (dotted orange line). The hydrogen bonds experienced by the mutated aspartate to Val282 and Leu283 in the apo protein are broken and its side chain faces the most likely protonated morpholine ring. C) In the complex of **4a** bound to TGT(Val262Cys) only the tricyclic core of the ligand is clearly defined. Contrary to the other crystal structures, the 2-substituent could not be assigned to the difference electron density. Similarly as in the apo state, Cys262 and Cys281 form a disulfide bridge. Due to minor rearrangements next to the site of mutation, helix  $\alpha$ B is completely disordered (dotted blue line).

The difference electron density for the mutated aspartate is not as properly defined as in the apo structure. The hydrogen bonds to Val282 and Leu283 are broken and the side chain of Asp262 is rotated towards **4a** being in van der Waals contact with the ligand's morpholino moiety. Furthermore, the difference electron density for the region comprising Asp280 to Arg286 indicates higher residual mobility than observed in the apo form. An increased mobility is also observed for the neighboured amino acids Ser287 and Gly288 whereby no electron density can be detected for the latter one.

A data set with a maximum resolution of 1.72 Å was collected for the TGT(Val262Cys)-**4a** complex. Also in the complex structure, the disulfide bridge between Cys262 and Cys281 is clearly visible. Contrary to the above-described inhibitor complexes, the present crystal structure shows well-defined electron density only for the parent scaffold while no defined density is observed for the 2-substituent (Figure 5.2C) indicating enhanced flexibility of this part of the ligand. Also no defined electron density can be assigned to helix  $\alpha$ B comprising the amino acids Leu283 to Asn290.



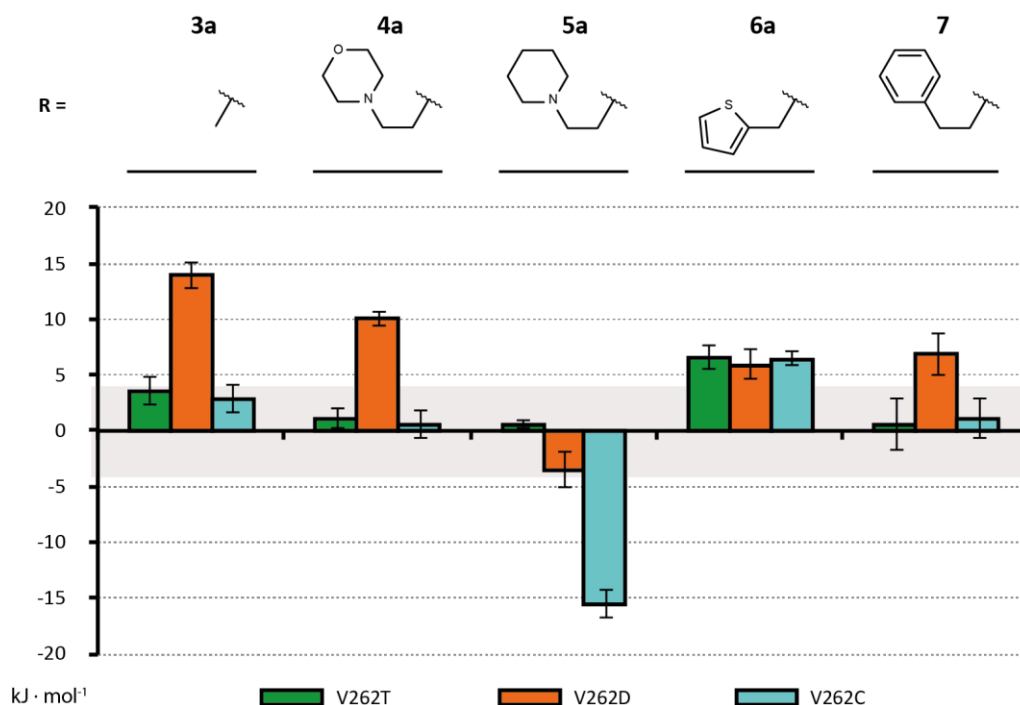
### 5.3 Discussion

To investigate the effects of the introduced mutations on the binding affinity to the 2-amino-*lin*-benzoguanine scaffold, the  $K_d$  values of **3a**, which is 2-substituted with only a methyl group, and the mutated variants were determined first. In a second step, the influences of the 2-substituents were evaluated by comparing the affinities and crystallographically determined binding modes of the remaining ligands to that of the reference ligand **3a**.

Within the range of experimental accuracy all ligands exhibit the same affinity towards TGT(Val262Thr) as towards the wild type. This result is well consistent with the crystal structures obtained from this variant: Firstly, the binding of the *lin*-benzoguanine scaffold does not take any influence on the order of the protein's secondary structure elements as shown by the comparison of the TGT(Val262Thr) apo structure and its complex structure with **4a**. Secondly, the morpholinoethyl substituent of the ligand adopts a similar *gauche*-conformation in TGT(Val262Thr) as described for the wild type. Also the enthalpies recorded for the binding of the different ligands to this mutated variant are similar to those observed for the wild type (Figure 5.3, green bars). Therefore, it is highly probable that the substituents of **5a**, **6a** and **7** behave upon binding to TGT(Val262Thr) similarly as described for wild type TGT in chapter 4.

The  $K_d$  values of the different 2-amino-*lin*-benzoguanines measured for TGT(Val262Asp) show, compared to the wild type TGT, a significant loss in binding affinity. Taking a closer look at the crystal structure of the TGT(Val262Asp)·**4a** complex, the position and orientation of the morpholinoethyl substituent exhibits no major difference to the WT TGT·**4a** complex. However, it can be clearly noticed that the  $\beta 8\alpha B$ -loop as well as helix  $\alpha B$  show a significantly increased flexibility upon ligand binding, which leads to the complete dislocation of two amino acids. In contrast to the structure in complex with **4a**, this part of TGT(Val262Asp) adopts a more rigid geometry in the apo structure. Possibly, the observed increase in flexibility is responsible for the drop in binding affinity. Moreover, this loss is more pronounced for **6a** and **7**, which are endowed with hydrophobic substituents, while **4a** and **5a**, endowed with polar most likely charged substituents, experience no changes in their  $K_d$ s compared to the reference ligand **3a**. Additionally, a significant loss in binding enthalpy and simultaneous gain in entropy is experienced by the ligands **3a**, **4a**, **6a** and **7** upon mutation of Val262 to aspartate (Figure 5.3, orange bars). An exception represents **5a**.

Certainly, electrostatic interactions either attractive (**4a** and **5a**) or repulsive (**6a** and **7**) will influence the resulting binding constants and thermodynamic profiles.



**Figure 5.3** Enthalpy differences of the investigated ligands **3a** – **7** compared to wild type TGT. The  $\Delta\Delta H$  values are shown as bars in units of  $\text{kJ} \cdot \text{mol}^{-1}$ .  $\Delta H_{obs}$  was measured in triplicates. The errors were calculated according to error propagation. The light gray box represent the  $\pm 4 \text{ kJ} \cdot \text{mol}^{-1}$  area, within that data are considered not significantly different.

The binding potency of the investigated ligands towards the protein slightly increases upon mutation of Val262 to cysteine. Although parts of helix  $\alpha B$  are delocalized in the structure of this variant in complex with **4a** no impact on the  $K_d$  value is noticeable in comparison to Val262Asp·**4a**. Inspecting the crystal structure, however, shows already a high flexibility of this part of the protein in the apo state of the mutant variant. The observed affinity gain and enthalpy differences (Figure 5.3, blue bars) can hardly be explained by means of the crystal structures since the disulfide bridge is presumably not present in solution and only formed during crystallization in consequence of an exposure of the sample to oxygen over several days. Also the fact that stored and subsequently thawed protein shows a drastically decreased affinity in ITC measurements supports the assumption that the freshly expressed protein is initially present in its reduced state.

## 5.4 Conclusions

The present chapter describes how various *lin*-benzoguanine type inhibitors with different flexible 2-substituents respond, with respect to their binding affinities, to mutations changing the properties of the flat, solvent-exposed uridine-33 subpocket of the target protein TGT.

The 2-amino-*lin*-benzoguanines included in this study exhibit varying behaviours of their 2-substituents: While **3a** serves as a reference ligand lacking any substituent to be deeply placed into the ribose-33 subpocket, **4a** and **5a** are endowed with flexible substituents attached to the tricyclic parent scaffold. Conformationally, the attached substituents might swap between *gauche*- and *anti*-conformation giving rise to enhanced residual mobility. For **6a** this flexibility is particularly pronounced, which is why no properly defined difference electron density for the thiophenomethyl substituent was detected in multiply recorded data sets. In contrast, **7** binds to TGT with its phenylethyl substituent in one single fully occupied orientation.

Our studies confirmed that gaining reliable information concerning the development of potential resistances requires the presence of an appropriate resistance model. As a prerequisite for the reasonable interpretation of crystal structures, these have to reflect the behaviour of the protein in solution best possible, which, in the present study, only applied to TGT(Val262Thr) and TGT(Val262Asp). Considering the potency of a compound we were able to show that not only the substituent, which binds into the pocket of interest, has to be taken into account but also the binding of the parent scaffold. In case of TGT(Val262Asp) the mere binding of the tricyclic core provoked disorder within the protein structure next to the site of mutation, which drastically decreased the binding affinity of the investigated inhibitors. Moreover, our studies indicate that the various 2-substituents do not show a large difference in their potency compared to the reference ligand **3a**. If at all, the investigated 2-amino-*lin*-benzoguanines only show any significant variance with respect to the 2-substituent when binding to TGT(Val262Asp) (factor three to seven). However, these could rather be explained by their chemical properties than by differences in their occupancies. Obviously, those ligands, which are endowed with a positively charged substituent, are able to attract the Asp262 side chain carboxylate, which leads, compared to the uncharged hydrophobic derivatives, to a less pronounced loss of affinity. The more

electron-rich aromatic substituents, in contrast, experience a distinct loss in potency most likely due to unfavourable electrostatic interactions.

In summary, the responses of the tested ligands upon mutations in the uridine-33 binding pocket of TGT proved diverse results. Although the crystal structures determined in this study suggest closely related binding modes of the ligands to the mutated variants, no clear correlation with respect to their 2-substituents was found. This leads to the conclusion that their potency predominantly results from the binding of their tricyclic scaffold.

## 6. Addressing a New Subpocket of TGT by Elongated 2-Amino-*lin*-benzoguanines

### 6.1 Introduction

The rational design and synthesis of potent ligands which do not only occupy the tunnel-shaped guanine-34 recognition site but also target the neighboring subpockets has been one of the major goals in the development of new TGT inhibitors. To this end various substituents were attached to the parent tricyclic scaffold of *lin*-benzoguanine. As appropriate substitution sites the 2-position and/or the 4-position were used resulting in potent ligands binding in the low nanomolar range [Stengl *et al.*, 2007; Kohler *et al.*, 2009; Ritschel *et al.*, 2009; Ritschel *et al.*, 2009; Barandun *et al.*, 2013]. While in the crystal structures of these compounds in complex with TGT the position of the 4-substituent could be clearly assigned to the difference electron density in nearly all cases, in no case any reasonable difference electron density could be spotted for the 2-substituent [Kohler *et al.*, 2009; Ritschel *et al.*, 2009]. This observation led to the assumption that the attached 2-substituents may be highly flexible even in the enzyme bound state.

Thus, further substituents were synthesized with the goal to fix this ligand portion within the uracil-33 subpocket by implementing basic ring systems that were meant to interact with backbone groups within the pocket. Similarly, polar terminal functions were attached to the ring systems that were meant to experience hydrogen bonds to Asp267 and Lys264 located at the far end of the uracil-33 subpocket. Furthermore, combined ring systems were added to mimic the uracil portion of the natural substrate and to achieve high shape-complementary. Finally, for crystallographic reasons, various substituents containing the stronger diffracting sulphur were synthesized [Ritschel, 2009].

Unfortunately, none of the chemical modifications resulted in an ordered binding mode of the 2-substituent visible in the crystal structure and, consequently, all previous studies had to consider binding geometries suggested by docking. However, also docking yielded multiple orientations of the attached 2-substituents, which made it impossible to detect the actual binding mode reliably [Ritschel, 2009; Ritschel *et al.*, 2009].

In the present study we performed an excessive screening for modified crystallization conditions, which ultimately enabled us to trace the positions of the various 2-substituents [Neeb *et al.*, 2014]. Although new interactions between the substituents and the protein could be identified within the ribose-34 subpocket, the originally addressed uracil-33 subpocket remained unoccupied. Therefore, we decided to extend the substitution pattern of the ligand in order to occupy additional chemical space and gain affinity. Thus, we designed modified ligands based on a successfully crystallized ligand exhibiting a phenylethyl substituent in 2-position.

Since docking results had not been able to indicate the binding poses of the 2-substituent in previous studies, we used MD simulations as an alternative computational method to predict binding poses, particularly of ligands exhibiting residual mobility. Subsequently, we determined the crystal structures of the designed ligands in complex with TGT to check for the actual binding modes, and quantified their binding affinities via a biochemical assay.

## 6.2 Results and Discussion

### 6.2.1 Ligand Design

As a starting point for ligand design we selected the 2-amino-*lin*-benzoguanine **7** decorated with a phenylethyl substituent (Table 6.1).

The first criterion for this choice had been bioavailability since the class of *lin*-benzoguanine-based inhibitors shows only poor permeability in PAMPA measurements [Barandun *et al.*, 2012]. From a bioavailability-point-of-view the phenylethyl substitution bears no additional protonation site that potentially lowers the permeability through cell membranes at a physiological pH value as it is the case, e.g., for the basic morpholinoethyl or piperidinoethyl substituent.

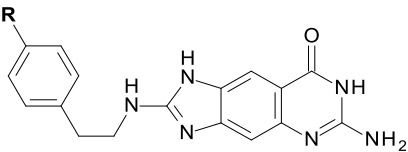
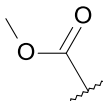
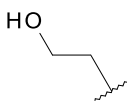
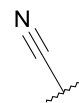
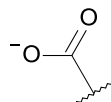
Secondly, we had focused on the binding affinity ( $K_i$ ) of the ligands. Remarkably, attachment of the phenylethyl substituent increases the potency of the parent inhibitor scaffold by a factor of about ten in the biochemical assay although, compared to the parent scaffold, no further polar interactions are provided [Ritschel *et al.*, 2009].

Finally, we had selected this substitution pattern from crystallographic considerations. Solely the phenylethyl substitution had resulted in a ligand that could be

refined to an occupancy of 100% while all other members of the 2-amino-*lin*-benzoguanine compound class under investigation had exhibited a reduced occupancy due to scatter over multiple conformations.

In the present study, we designed and synthesized, based on the parent scaffold **7**, four different elongation motifs exhibiting diverse chemical properties (Table 6.1).

**Table 6.1** Chemical formulae and binding affinities of the extended 2-amino-*lin*-benzoguanines.

Compound		$K_i$ [nM]
<b>7</b>	H	$10 \pm 3$ <sup>[a]</sup>
<b>7a</b>		$2 \pm 2$
<b>7b</b>		$7 \pm 3$
<b>7c</b>		$21 \pm 8$
<b>7d</b>		$30 \pm 7$

<sup>a</sup> $K_i$  value according to Ritschel *et al.*, 2009.

### 6.2.2 MD Simulations

MD simulations were performed using the program *AMBER* under NTP conditions ( $p = 1\text{bar}$ ,  $T = 300\text{K}$ ). **7a** and **7d** were simulated over 10 ns. **7b** and **7c** were trapped in a local minimum over a 10 ns period. Thus, for these ligands simulation time was extended to 100 ns. Subsequently, the ligand conformations taken up along the trajectory were hierarchically clustered into families with maximal RMSDs of 2 Å using the program *ptraj* and representative arche types were visually inspected. The results are listed in Table 6.2.

**Table 6.2** Results of the hierarchical clustering after the performed MD simulations. Given is the total simulation time, the number of observed clusters, the most populated binding poses, their percental occurrence, and the RMSD compared to the binding mode observed in the subsequently determined crystal structure.

	time	# clusters	Binding poses <sup>[a]</sup>	Occurrence	RMSD <sup>[b]</sup>
<b>7a</b>	10 ns	3		63.6%	0.4 Å
				63.6%	0.4 Å
<b>7b</b>	100 ns	8		46.8%	4.9 Å
				4.3%	0.4 Å
<b>7c</b>	100 ns	6		76.1%	3.5 Å
				2.5%	1.6 Å
<b>7d</b>	10 ns	8		34.2%	1.3 Å
				8.8%	1.1 Å

<sup>a</sup>Color code: green = most represented, blue = most similar to crystal structure, yellow = crystal structure.

<sup>b</sup>RMSD values were calculated using *fconv* after alignment on the tricyclic scaffold without hydrogens [Neudert and Klebe, 2011].

All simulations were able to find a binding pose of the 2-substituent similar to the one observed in the crystal structure within an RMSD between 0.4 Å and 1.6 Å. Thereby, the pose of **7a** matched particularly well (RMSD = 0.4 Å) and, moreover, was the most representative in terms of percental occurrence, closely similar to the occupancy found by



X-ray crystallography (63.6% vs. 74%). For the remaining ligands **7b**, **7c** and **7d** different binding poses were suggested as the most populated orientations deviating between 1.3 Å to 4.9 Å compared to those found in the subsequently determined crystal structures.

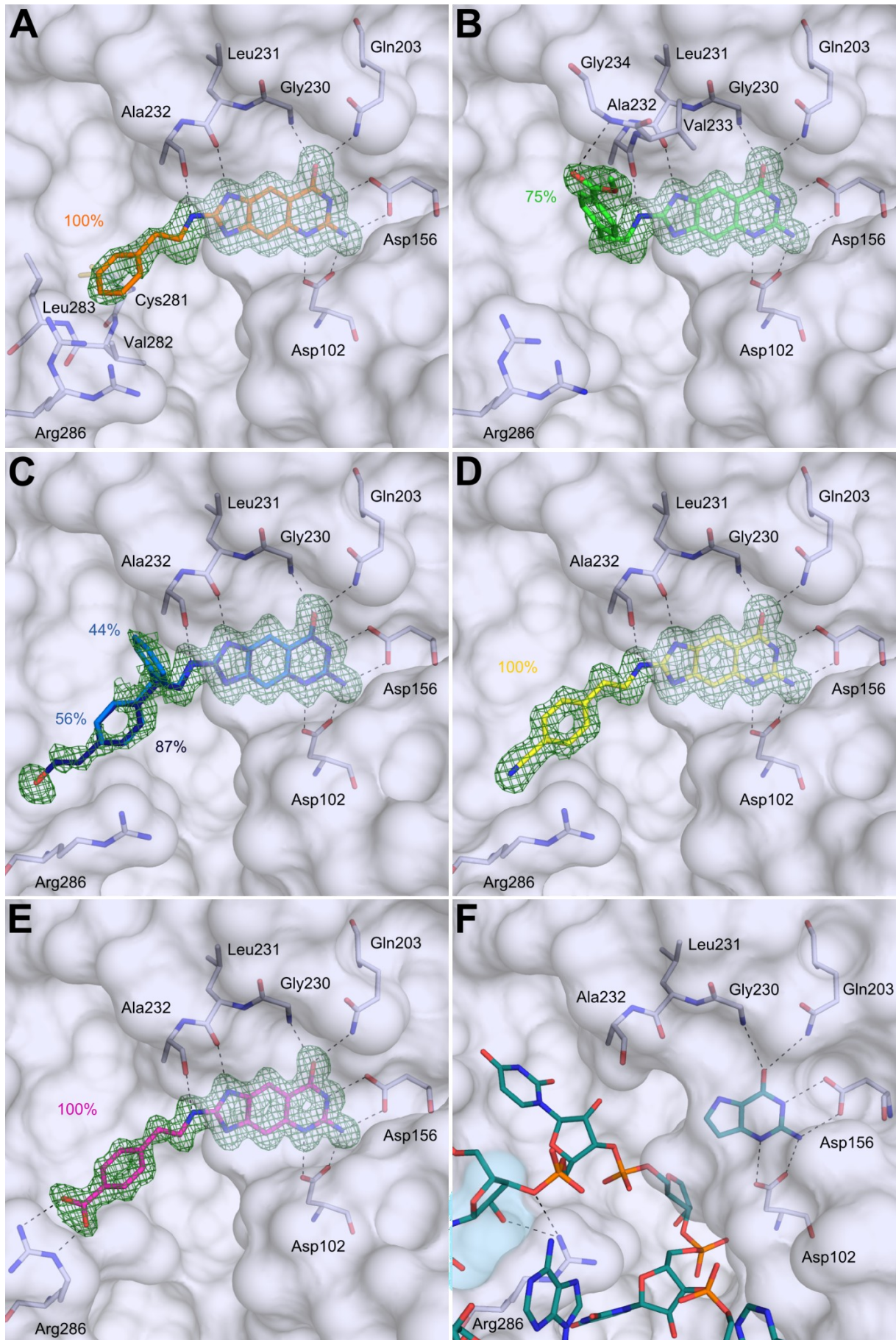
### 6.2.3 Affinity Data

Binding affinities were obtained via a radioactive enzyme assay. Thereby, the incorporation of [8-<sup>3</sup>H]guanine into tRNA<sup>Tyr</sup> (ECY2) at position G34 at pH 7.3 and 37°C was measured by liquid scintillation counting. Inhibition constants were calculated by the comparison of the initial velocities of the base-exchange reaction in absence and presence of the ligand.

All analyzed ligands show a binding affinity in the low one- to two-digit nanomolar range (Table 6.1). Reference ligand **7** exhibits a binding affinity of about 10 nM [Ritschel *et al.*, 2009]. Interestingly, the attachment of the methylester in **7a** is able to achieve a significant enhancement in binding affinity by a factor of five. For **7b** further decoration of the phenyl moiety does not yield a significant change in the  $K_i$  value, although its ligand efficiency is decreased by the attachment of the hydroxyethyl group. In contrast, **7c** and **7d** show a slightly decreased potency. Compared to the unsubstituted phenylethyl lead structure the value for the derivative **7c** with an attached nitrile group is reduced about twofold, while the decoration with a carboxylate group in **7d** results in a loss by a factor of three. Obviously, the desolvation costs especially of the attached polar functional group of **7d** are larger than the actual gain in affinity due to newly formed interactions.

### 6.2.4 Crystal Structure Analysis

For the investigated ligands crystal structures were determined with a resolution of 1.14 – 1.40 Å. In all structures the fully occupied *lin*-benzoguanine scaffold is well-defined in the difference electron density (Figure 6.1). It is placed into the guanine-34 recognition pocket forming a  $\pi$ -stacking between the side chains of Tyr106 and Met260 and establishing the same interactions to neighbored amino acids as previously described [Stengl *et al.*, 2007; Kohler *et al.*, 2009; Ritschel *et al.*, 2009; Ritschel *et al.*, 2009; Barandun *et al.*, 2013].



**Figure 6.1** Binding modes of the analyzed *lin*-benzoguanines **7** (A), **7a** (B), **7b** (C), **7c** (D) and **7d** (E) as well as a tRNA analogue (F). The protein surface is shown in white. The ligands and interacting residues are displayed as sticks (nitrogen = blue, oxygen = red, phosphorus = orange). Hydrogen bonds are visualized as dashes. The  $|F_o| - |F_c|$  difference electron density is illustrated at a *sigma*-level of  $\sigma = 2.5$  as green meshes. For clarity the  $\pi$ -stacking residues Tyr106 and Met260 as well as water molecules are not shown. A – E) The tricyclic parent scaffold is well defined in the difference electron density experiencing hydrogen bonds to residues Asp102, Asp156, Gln203, Gly230, Leu231 and Ala232 within the guanine-34 recognition site. Depending on its substitution pattern, the phenylethyl substituent in 2-position adopts either an orientation within the ribose-33 subpocket or is rotated out of it facing the backbone of Val233 and Gly234. Only **7d** (E) forms additional polar interactions to the guanidinium head group of Arg286. While in the complex structures TGT·**7**, TGT·**7a** and TGT·**7d** the side chain of Arg286 closes the ribose-32 subpocket, **7b** and **7c** are able to bind to that site turning the side chain of Arg286 apart. F) Binding of tRNA induces the same conformation of Arg286 as found in the complexes TGT·**7b** and TGT·**7c** opening the ribose-32 subpocket (blue). Ribose-32 interacts with the side chain of Arg286 via two charge-assisted hydrogen bonds.

The  $|F_o| - |F_c|$  difference electron density fully defines the binding mode of **7** (Figure 6.1A). Yet, increased B factors of the 2-substituent compared to the tricyclic parent scaffold are observed ( $B_{phenyl} = 31.5 \text{ \AA}^2$  vs.  $B_{tricycle} = 17.2 \text{ \AA}^2$ ). The phenylethyl substituent occupies the ribose-33 subpocket adopting an all-*trans* conformation. Weak hydrophobic interactions are experienced between the side chains of Ala232, Cys281, Val282, Leu283 and the 2-substituent, which covers in lid-like fashion the hydrophobic residues in this pocket.

The 2-substituent of **7a** refines to a reduced occupancy of 75% indicating higher flexibility of this moiety compared to that observed for **7** ( $B_{phenyl} = 28.9 \text{ \AA}^2$  vs.  $B_{tricycle} = 14.6 \text{ \AA}^2$ ). The C-CH<sub>2</sub>-CH<sub>2</sub>-NH linker adopts a *gauche* conformation with a torsion angle of  $-61.3^\circ$ , which, most likely, is energetically less favourable (Figure 6.1B). Contrary to the *gauche*-isomer of the substituent in **4a** (Chapter 4), the one in **7a** is oriented out of the ribose-33 subpocket facing the backbone of Val233 and Gly234. Obviously, an extended all-*trans* conformation of this substituent would lead to a clash with Arg286 and, therefore, is not adopted. Instead, the substituent is forced to fold back into the described conformation moving the ligand out of the ribose-33 pocket. Weak van-der-Waals interactions are formed to the side chain of Ala232. Additionally, an expanded hydrogen-bond contact between the backbone NH group of Gly234 and the carbonyl oxygen of the ester group in **7a** is formed (3.51 Å). Remarkably, such a folded conformation is observed for an attached 2-substituent for the first time.

For **7b** two data sets showing different results with respect to the substituent's binding mode were obtained (Figure 6.1C). The first structure displays the 2-substituent in two conformations with a summed overall occupancy of 100%. In both conformations the substituent is not sufficiently resolved in the difference electron density due to the residual mobility of that part of the molecule resulting in an ill-defined difference electron density for the hydroxyethyl moiety. In the first conformer, the C-CH<sub>2</sub>-CH<sub>2</sub>-NH linker adopts an all-*trans* conformation with an occupancy of 56% and enhanced residual mobility ( $B_{phenyl} = 20.3 \text{ \AA}^2$  vs.  $B_{tricycle} = 11.9 \text{ \AA}^2$ ). The adopted orientation is similar to that observed for **7**. In the second conformer, the linker exhibits *gauche* conformation resulting in an analogous binding mode as observed for **7a** ( $B_{phenyl} = 19.8 \text{ \AA}^2$  vs.  $B_{tricycle} = 11.9 \text{ \AA}^2$ ). In this orientation the substituent refines to 44%. In a second independently collected data set using another crystal the whole ligand is visible in the difference electron density in the all-*trans* conformation only, although a slightly reduced occupancy for the 2-substituent of 87% is refined. With increasing distance from the parent scaffold the difference electron density becomes more smeared out and, accordingly, the assigned temperature factors increase ( $B_{substituent} = 26.4 \text{ \AA}^2$  vs.  $B_{tricycle} = 12.4 \text{ \AA}^2$ ). The position of the hydroxyethyl portion can hardly be assigned to the density. Interestingly, the side chain of Arg286 is shifted out of the ribose-33 subpocket presumably to create sufficient space to accommodate the hydroxyethyl moiety of the ligand now extending its binding towards the ribose-32 subpocket. This region of the protein has never been occupied by synthetic ligands so far.

Comparably, also the 2-substituent of **7c** adopts an all-*trans* conformation penetrating into the ribose-32 subpocket (Figure 6.1D). The whole ligand is fully occupied, although its 2-substituent shows increased temperature factors compared to the parent scaffold ( $B_{substituent} = 17.4 \text{ \AA}^2$  vs.  $B_{tricycle} = 9.8 \text{ \AA}^2$ ). The terminal nitrile group does not experience specific interactions with any of the amino acid residues found in the ribose-32 subpocket. Only a single water molecule is located in close distance (3.0 Å) to the nitrile functional group. The side chain of Arg286 has to change its orientation and is shifted out of the ribose-33 subpocket adopting the previously found conformation.

The binding mode of **7d** is fully defined in the  $|F_o| - |F_c|$  difference electron density (Figure 6.1E). The C-CH<sub>2</sub>-CH<sub>2</sub>-NH linker adopts an all-*trans* conformation reaching into the ribose-33 subpocket. Under the applied pH conditions the terminal carboxylate group is most likely deprotonated. In this state, it forms a bidentate salt-bridge to the guanidinium

moiety of Arg286. In order to do so, the polar group of Arg286 bends towards the ribose-32 subpocket. Despite of the strong electrostatic interaction between the carboxylate group of the ligand and the side chain of Arg286, the 2-substituent exhibits B factors, which are similarly increased as those of the remaining derivatives investigated in this study ( $B_{\text{substituent}} = 21.1 \text{ \AA}^2$  vs.  $B_{\text{tricycle}} = 11.1 \text{ \AA}^2$ ).

### 6.3 Conclusions

In the present chapter the binding modes of five extended 2-amino-*lin*-benzoguanines were predicted using MD simulations as *in silico* design technique. Subsequent to synthesis their binding modes were determined experimentally by crystal structure analyses whereby modified crystallisation conditions were developed. Previous studies based on docking predictions and crystal structure analyses using soaking protocols had failed to actually characterize the binding poses of the attached 2-substituents [Ritschel, 2009; Ritschel *et al.*, 2009]. In all cases enhanced disorder and scatter over multiple orientations of the substituent was suggested. Therefore, we applied molecular dynamic simulations as this method is better suited to explore the accessible configuration space of the 2-substituent. In one example the MD simulations predict the binding mode, which was subsequently found via X-ray crystallography, well. Also in the other three cases binding modes, which are similar to those found in the crystal structures, are predicted by MD. However, the attributed populations of the different configurations do not match with the refined occupancies in the respective crystal structures.

The subsequently determined crystal structures of the four TGT-ligand complexes, which were based on a co-crystallization protocol, provide a more detailed insight into the binding mode of the extended 2-substituents. While in previous studies our design hypotheses were based on *in silico* docking performed to interpret the ill-defined difference electron density in the area where the 2-substituent would be expected, reasonable  $|F_o| - |F_c|$  density is observed for nearly all ligands in the analyses presented in this contribution. The crystal structures indicate different conformations of the attached 2-substituents, which will either reach into the ribose-33 subpocket, if the C-CH<sub>2</sub>-CH<sub>2</sub>-NH linker adopts an all-*trans* conformation, or out of the ribose-33 pocket facing the backbone of Val233 and Gly234, if the linker folds into a *gauche* orientation. Surprisingly, the terminal

groups of **7b** and **7c** are able to force Arg286 into a conformation pointing off the ribose-34 subpocket. The latter residue acts as a kind of flexible gatekeeper, which controls access to the ribose-32 subpocket. In a conformation oriented out of the ribose-33 pocket, it adopts a position similarly found in the structure with the natural RNA substrate (Figure 6.1F). In this geometry the pocket is opened towards the uridine/cytosine-32 recognition site, a pocket, which previously had never been addressed by any of our synthetic ligands. Remarkably, the displacement of the flexible side chain of Arg286 does not result in a significant affinity loss.

The deduced information concerning the binding modes of the extended 2-substituent opens the perspective for a new strategy to design more potent ligands forming new interactions within the nucleoside-32 subpocket. This region offers polar side chains to be addressed similarly as by the ribose-32 portion of the natural substrate tRNA.

Two messages can be learnt from this contribution. First of all, the visibility of substituents in the difference electron density may depend on the applied crystallization protocol. In consequence, not every ill-defined electron density inevitably indicates the residual mobility or the scatter of a substituent in the protein-bound state. It may instead result from the particular protein conformation preformed during crystallization. This preformed conformation may prevent the mutual conformational adaptation of the protein and the ligand and thus, prevent the ligand from adopting a defined and, from an enthalpic point of view, favorable conformation when bound to the crystallized protein. Accordingly, it is of utmost importance to distinguish if soaking or co-crystallization protocols were applied when discussing the results of structural studies including protein-bound ligands. As a consequence, the direct comparison of protein-ligand complex structures must be deemed problematic, if the considered crystals were not produced via the same crystallization protocol.

The second message concerns the attachment of polar groups at the terminal end of ligands. In the present compound series no dramatic change in binding affinity was experienced in comparison to the original phenylethyl reference substituent, although in the case of **7d** even a novel salt bridge contact to Arg286 is established. Nonetheless, this ligand, which is endowed with a carboxylate group, loses affinity. This may, at least in part, be due to the high desolvation costs for such a charged functional group. Definitely, this example underlines that a solvent-exposed salt-bridge as the one formed to Arg286 at the rim of the ribose-33 pocket has hardly any enhancing effect on binding affinity. Such an interaction will

only improve binding affinity strongly, if it is formed in a deeply buried environment well-shielded from the surrounding water. Only there a charge-assisted contact is strongly enhanced by the local environment of low dielectricity. This has to be well regarded if putative interaction sites in flat solvent-exposed binding pockets are thought of as potential hot-spots to be addressed in a drug design study.

## 7. 5-Azacytosine as a Novel Scaffold to Inhibit *Z. mobilis* TGT with Expected Improved Bioavailability and Synthetic Accessibility

### 7.1 Introduction

Fragment-based lead discovery (FBLD) has been well established as an alternative to high-throughput screening (HTS) in the field of lead finding to develop novel drug-like candidates. Prerequisite for the success of such approaches was the introduction of sensitive biophysical methods such as surface plasmon resonance (SPR) [Perspicace *et al.*, 2009; Navratilova and Hopkins, 2010] or NMR [Fejzo *et al.*, 1999; Dalvit *et al.*, 2002], which are able to detect weak ligand-protein binding. FBLD comprises several steps starting with the design of an appropriate fragment library, followed by *in-vitro* testing of fragments detected in biophysical prescreening by X-ray crystallography and reliable affinity measurements. In a subsequent phase promising fragments have to be evolved to putative lead structures either by fragment growing, linking or merging. This strategy is accompanied by multiple steps of optimization consisting of modelling, structure determination and assay measurements.

Although HTS focuses on the testing of a large magnitude of compounds in the range of  $10^6$ , this method faces a number of problems with respect to subsequent optimization, such as too high lipophilicity or poor solubility of the initial hits leading to false positives in the primary HTS assays [Scott *et al.*, 2012]. Also discovered hits do not necessarily fill binding pockets of the target protein in an optimal manner as the initial test compounds are rather large with respect to their molecular weights ranging from 300 to 500 Da [Lipinski *et al.*, 2001]. Even though discovered hits might show low-micromolar binding, their suitability as starting point for the optimization is difficult to estimate since the contribution of individual interactions to binding affinity remains virtually undefined.

Optimal compound selection for FBLD was suggested to follow the Astex Rule of Three, which means that the molecular weight should be below 300 Da, not more than three hydrogen bond acceptor or donor functionalities should be present and the calculated logP value has to be lower than three [Congreve *et al.*, 2003]. Attributed to the smaller size of fragments, test libraries usually comprise only  $10^3$  members and still are assumed to cover the chemical space sufficiently well, much better than HTS collections can actually achieve



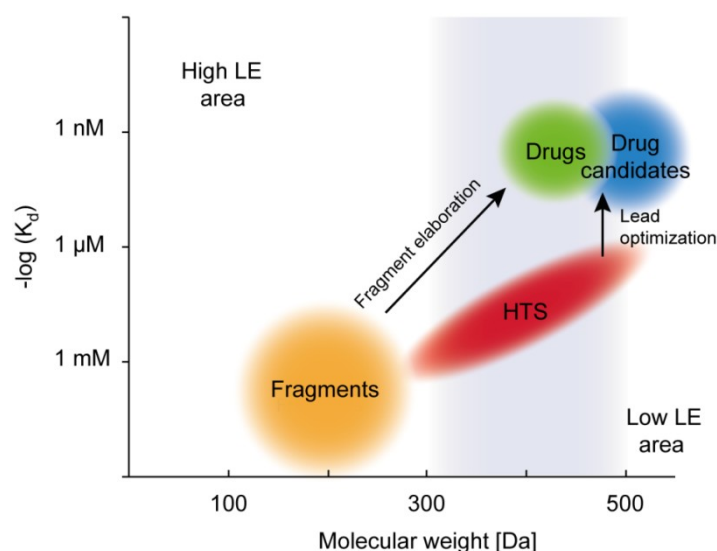
[Makara, 2007]. Moreover, the quality of the formed interactions between the small and often optimally fitting fragments and the target's amino acids is better and more efficient than with HTS compounds since a weak binding fragment has to establish binding via fewer interactions.

To better relate the potency of a compound to its actual size ligand efficiency (LE) measures have been suggested, which determine the ratio between Gibbs free energy of binding ( $-\Delta G$  ( $kcal\ mol^{-1}$ )) and the number of non-hydrogen atoms of the compound [Kuntz *et al.*, 1999]. Thus, efficient ligands will only be attributed to high affinity binding, if they simultaneously exhibit low molecular weight. More recently, the concept of group efficiency (GE) has been proposed to estimate the contribution of an added functional group to the free energy during optimization [Verdonk and Rees, 2008]. Group efficiency defines the gain in affinity ( $kcal\ mol^{-1}$ ) compared to the molecule lacking the functional group under investigation related to the number of added non-hydrogen atoms.

The optimization of fragment and HTS hits is illustrated in Figure 7.1. The goal to hit by putative drug candidates is an area of high LE with a molecular weight not exceeding 500 Da. HTS hits start at lower LE than fragments as they exhibit already a molecular weight approaching the 500 Da limit. Therefore, LE of HTS hits has to be increased while there is only little room to simultaneously introduce new functional groups, which augment molecular weight. Thus, interactions experienced by the HTS hits have to be replaced by alternative ones leading to stronger interactions whereby even a reduction in ligand size might be necessary. In contrast, fragment hits exhibit low LE but they offer multiple options to attach additional groups to achieve more contacts with the protein and improve LE.

Recently, the 6-amino-1,7-dihydro-8*H*-imidazo[4,5-*g*]quinazolin-8-ones (*lin*-benzoguanines) and 1,7-dihydro-8*H*-imidazo[4,5-*g*]quinazolin-8-ones (*lin*-benzohypoxanthines) were introduced as potent scaffolds inhibiting the tRNA-modifying enzyme tRNA-guanine transglycosylase (TGT) from *Zymomonas mobilis* [Stengl *et al.*, 2007; Barandun *et al.*, 2012]. Substituents attached at position 2 and 4 of the parent scaffolds lead to binding affinities in the nanomolar range [Hörtner *et al.*, 2007; Kohler *et al.*, 2009]. Despite their high potency, *lin*-benzopurines do not exhibit ideal properties for good *in vivo* activity as parallel artificial membrane permeability assay (PAMPA) measurements suggest that they may hardly be transported over biomembranes [Barandun *et al.*, 2012]. Unfortunately, the most promising compounds only accumulate within the membrane but do not succeed in transportation.

## 7. 5-Azacytosine as a Novel Scaffold to Inhibit *Z. mobilis* TGT with Expected Improved Bioavailability and Synthetic Accessibility



**Figure 7.1** Comparison of fragment based lead discovery (FBLD) and high throughput screening (HTS) approaches modified according to Scott *et al.*, 2012. The molecular weight of putative hits is plotted against their potency. The right corner represents the low LE area, the left corner the high LE area. Fragment hits (orange) exhibit low molecular weight as well as low binding affinity. Upon optimization fragments are iteratively grown, linked or merged to a putative drug candidate (blue) of higher molecular weight settled in the Lipinski area between 300 and 500 Da. By this procedure potency is gained by attaching additional functional groups and high LE is reached. Contrary, HTS hits (red) show already a high molecular weight, which makes the introduction of new functional groups difficult. Therefore, existing functionalities have to be replaced by ones of higher quality. Even a reduction of the molecule's size might be necessary.

Therefore, the search for novel scaffolds with improved bioavailability is of utmost importance. In the present study we embarked onto a fragment-based approach aimed at the reduction of the tricyclic ring system of the *lin*-benzopurines. Starting point for the design was a fragment suggested by modelling, which consists of a 5-azacytosine core. In its structural and physicochemical properties it exhibits similarity with the natural substrate guanine and bears a directly linked phenyl ring. We focused on the optimization of this initial fragment by attaching additional groups in order to improve its LE.

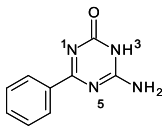
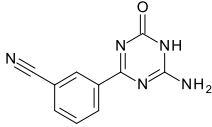
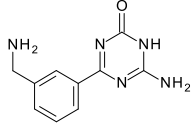
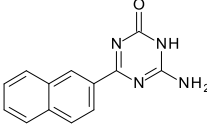
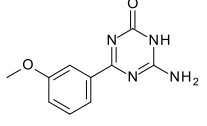
## 7.2 Results and Discussion

### 7.2.1 Ligand Design

*Z. mobilis* TGT exerts a nucleobase exchange of guanine by a modified base in the wobble position G34 of the tRNA-anticodon loop. Thereby, the molecular recognition of the

substrate is established by the amino acid side chains of Asp102, Asp156, Gln203 and the backbone NH of Gly230. Additionally, the nucleobase is sandwiched in  $\pi$ -stacking fashion between Tyr106 and Met260 located above and below the binding site [Xie *et al.*, 2003; Tidten *et al.*, 2007].

**Table 7.1** Chemical formulae, calculated cLog P values, corresponding  $K_i$  values to *Z. mobilis* TGT determined by the radioactive enzyme assay and ligand efficiencies of the 5-azacytosine analogues.

		cLog P <sup>[a]</sup>	$K_i$ [ $\mu$ M]	LE
	<b>8</b>	0.40	465 $\pm$ 139	0.32
	<b>9</b>	0.25	346 $\pm$ 210	0.30
	<b>10</b>	-0.74	107 $\pm$ 31	0.34
	<b>11</b>	1.39	4 $\pm$ 2	0.41
	<b>12</b>	0.24	1476 $\pm$ 648	0.24

<sup>a</sup>cLog P values were calculated using the Properties Viewer online software on chemicalize.org.

The 5-azacytosine core is deduced from the natural substrate guanine. During design it was assumed to establish similar interactions to Asp102, Asp156, Gln203 and Gly230. The attached phenyl ring is meant to enhance binding affinity by  $\pi$ -stacking interactions to Tyr106 and Met260. In addition, it serves as an anchor point to decorate the novel scaffold by easily accessible chemistry with substituents in 2- and 4-position. These substituents were planned to reach into the hydrophobic ribose-34 subpocket flanked by Val45, Leu68, Leu100 and Val282 and the solvent-exposed uridine-33 cavity made up of various backbone atoms and sealed at the far end by residues Lys264 and Asp267.

In addition to the described scaffold, also ligands exhibiting small decorations at the phenyl moiety were synthesized in order to search for promising substituents as starting points for fragment growing. These included a nitrile, an aminomethyl and a methoxy group in meta position of the phenyl ring as well as an additional fused aromatic ring leading to a naphthyl substituent. Details concerning the physicochemical properties of ligands **8** – **12** are listed in Table 7.1.

### 7.2.2 Inhibition Mode and Affinity Data

The blocking of the function of the target protein TGT can follow, depending on the size of the active-site ligand, an inhibition mechanism which is competitive or non-competitive with simultaneous binding of the tRNA. Therefore, prior to affinity measurements the inhibition mode was determined by a trapping experiment [Xie *et al.*, 2003]. The results are visualized in Figure 7.2A. All investigated ligands show only one band attributed to the mass of TGT with the bound ligand at approximately 43 kDa. This clearly indicates that a competitive binding mode is given not allowing the tRNA molecule to bind to the protein-inhibitor-complex. Obviously, the tRNA is sterically blocked from simultaneous binding to the protein likely due to the attached phenyl ring at the basic scaffold (Figure 7.2B).

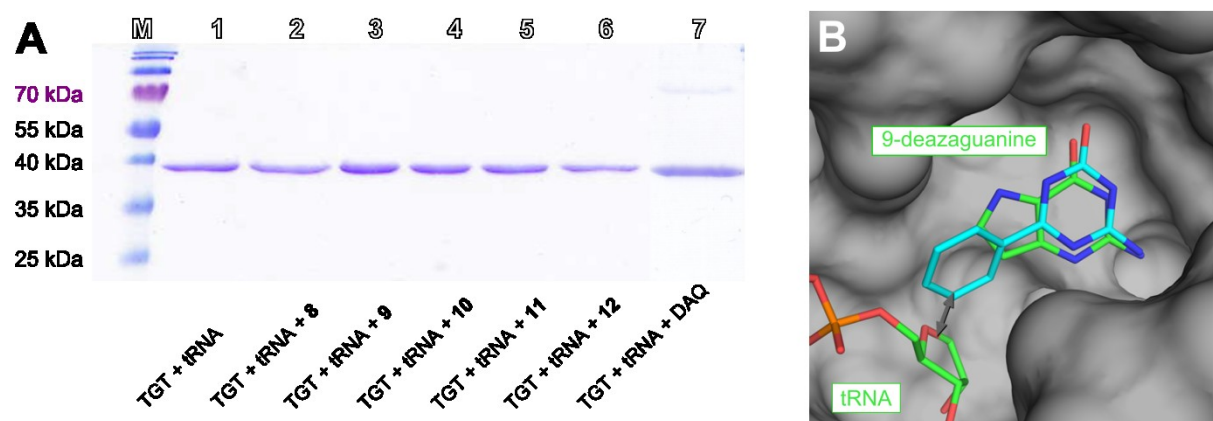
Binding affinities were determined via a radioactive enzyme assay. Thereby, the incorporation of [ $^3\text{H}$ ]guanine into tRNA<sup>Tyr</sup> (ECY2) at position G34 at pH 7.3 and 37°C is measured by liquid scintillation counting [Stengl *et al.*, 2005]. Inhibition constants were calculated by the comparison of the initial velocities of the base-exchange reaction in the absence and presence of the inhibitor.

The phenyl derivative **8** shows an inhibitory constant of 465  $\mu\text{M}$ . Compared to the *lin*-benzoguanine scaffold affinity drops by a factor of approximately 100 [Hörtner *et al.*, 2007]. The substitution of the phenyl ring in meta position by a nitrile group in **9** does not result in a noticeable change of the  $K_i$  value within the range of error, yielding a value of 346  $\mu\text{M}$ . By introducing an aminomethyl group (**10**) an approximately 4.3-fold increase in binding strength could be measured, resulting in a  $K_i$  value of 107  $\mu\text{M}$ .

Unexpectedly, the naphthyl substituent of **11** improves affinity significantly to the one-digit micromolar range to 4  $\mu\text{M}$ , which corresponds to a more than 100-fold affinity

increase. In case of GABA<sub>A</sub> antagonists a similar affinity gain has been observed by increasing the hydrophobicity of the ligand from a phenyl to a naphthyl moiety [Frølund *et al.*, 2002].

Contrary to the other ligands, compound **12** changes the potency to the worse by a factor of 3.2 compared to the parent scaffold **8**. Its  $K_i$  amounts to 1476  $\mu\text{M}$ .



**Figure 7.2** Inhibition mode of 5-azacytosine type ligands. A) Trapping experiment performed by SDS-PAGE using a mixture of TGT, tRNA and the different 5-azacytosines. M, size marker; DAQ, 2,6-diamino-3H-quinazolin-4-one. The reference band 1 as well as the samples 2 – 6 containing 5-azacytosine based ligands show only one band indicating a competitive inhibition mode. The reference sample containing DAQ as an uncompetitive inhibitor, which stabilizes the covalently linked TGT-tRNA complex, exhibits an additional retarded band (lane 7). B) Superimposition of the proposed binding mode of **8** (cyan) and the tRNA stem loop (green, PDB ID: 1Q2R). The ligands are shown in stick representation. Oxygen atoms are colored in red, nitrogen atoms in blue and phosphate atoms in orange. The surface of the protein is represented in gray. For clarity, the surface of Tyr106 and Cys158 is not shown. The inhibition mode is confirmed by the proposed binding mode of **8**. The phenyl ring in 6-position and ribose-34 of the tRNA are in too close proximity (2 Å) to allow the binding of **8** to the guanine-34 recognition site of the enzyme-tRNA intermediate.

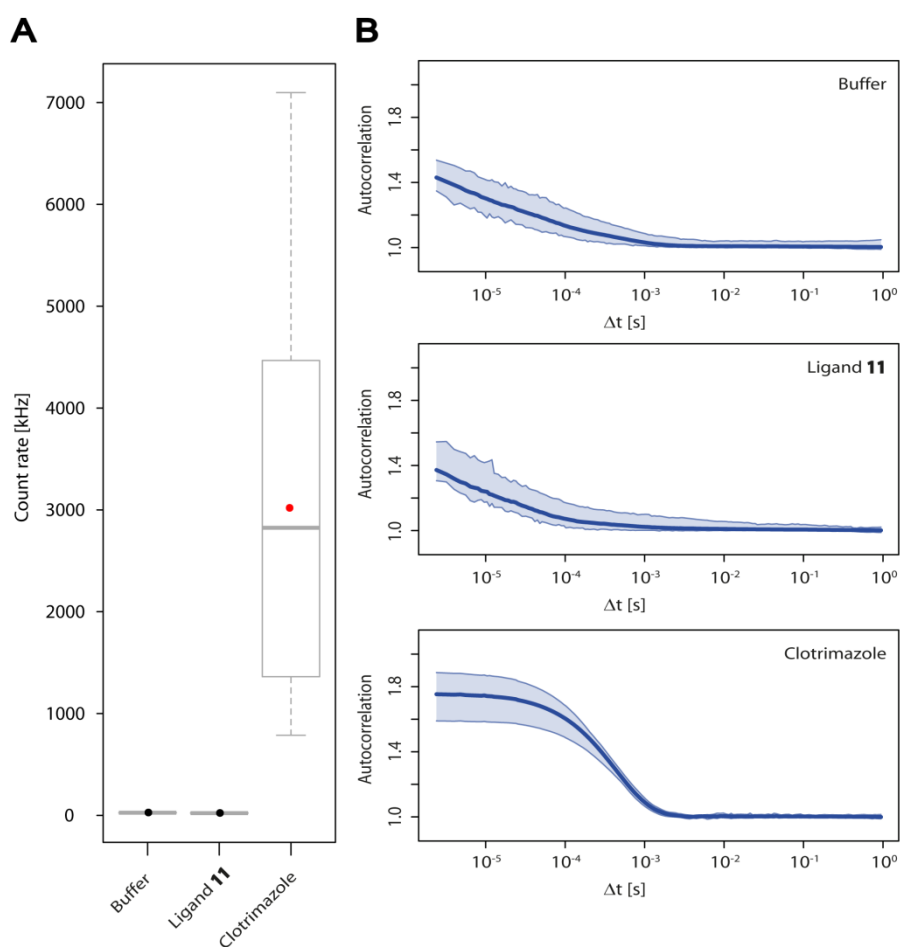
### 7.2.3 Putative Agglomeration Behaviour of the Naphthyl Derivative **11**

Hydrophobic ligands tend to agglomerate above a critical agglomeration concentration (CAC), which is typically found in the micromolar range [McGovern and Shoichet, 2003; Coan and Shoichet, 2008; Doak *et al.*, 2010]. The agglomerate binds to the protein surface in unspecific manner and suggests false positive signals [Feng and Shoichet, 2006; Feng *et al.*, 2007; Thorne *et al.*, 2010]. According to Table 7.1 **11** exhibits the lowest solubility in aqueous solution in agreement with the cLog P value of 1.39. Therefore, the question arises, whether

## 7. 5-Azacytosine as a Novel Scaffold to Inhibit *Z. mobilis* TGT with Expected Improved Bioavailability and Synthetic Accessibility

the high potency of **11** is due to an unspecific binding event. To rule out this eventuality the following experiments were performed:

Dynamic light scattering (DLS) measurements were carried out to identify potential aggregate formation [Feng *et al.*, 2005]. A ligand solution was analyzed under the same conditions as applied in the assay measurements (200 mM Hepes pH 7.3, 20 mM MgCl<sub>2</sub> and 0.037% (v/v) Tween 20 at 37°C).



**Figure 7.3** DLS measurements. A) Boxplot showing the count rates of the analyzed samples in kHz. The gray box represents the interquartile range (IQR) with its borders giving the first (Q1) and third quartile (Q3) of the measurements. The median is represented by a gray bar within the box. The average count rate is visualized by a circle; the red color indicates that the signal exceeds the 10-fold of the blank value. The borders of the dashed lines give Q1 – 1.5 IQR and Q3 – 1.5 IQR, respectively. While **11** shows a count rate in the range of the blank measurement (24 kHz vs. 27 kHz), clotrimazole exhibits an over 100-fold increased count rate of 3020 kHz. B) Autocorrelation functions of the samples plotted against time. **11** shows a similar autocorrelation function as the measurement of the sole buffer. Clotrimazole exhibits a sigmoidal autocorrelation function indicating its behaviour as strong aggregate former.

The ligand concentration was adjusted to 20  $\mu\text{M}$  corresponding to the highest concentration of the dilution series in the assay. Samples were measured iteratively every ten minutes over a period of one hour and compared to the data observed by the strong aggregate former clotrimazole [Seidler *et al.*, 2003].

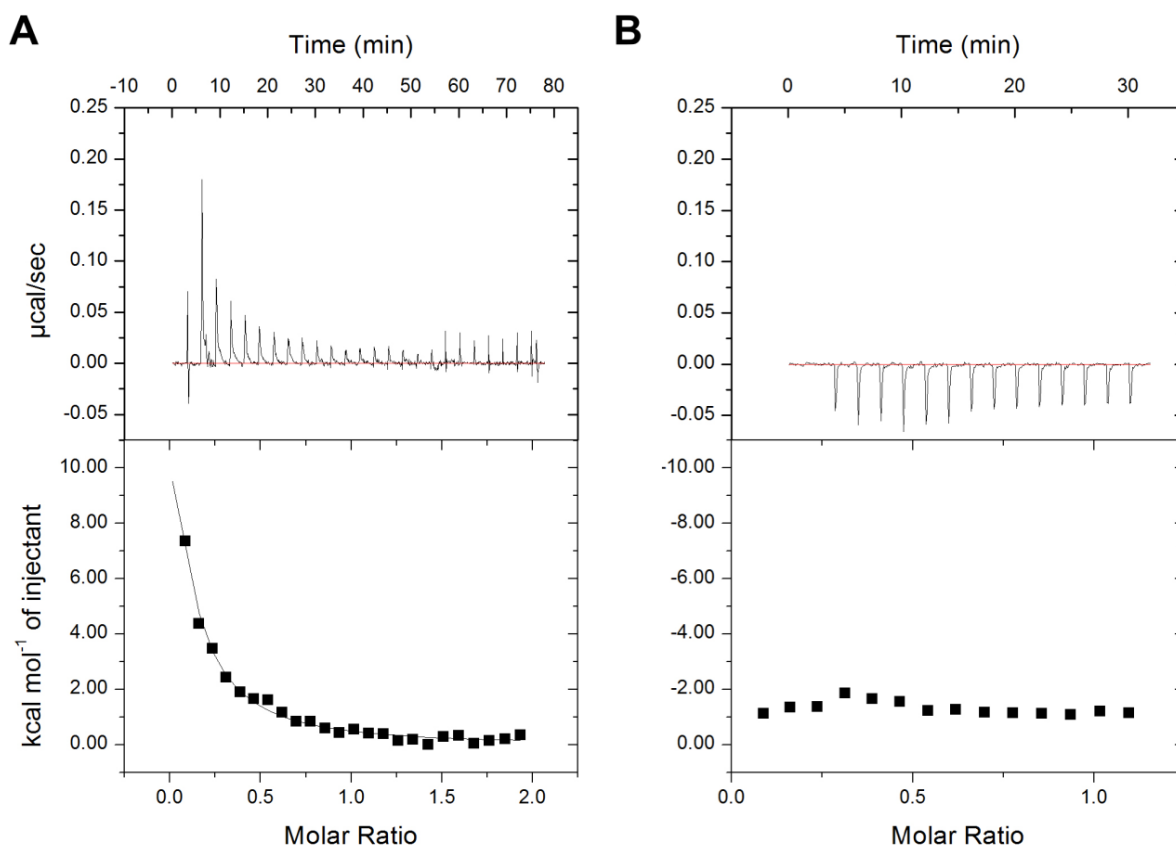
According to theory, the autocorrelation function of a DLS measurement shows a sigmoidal curve in case of an agglomeration event accompanied by the significant exceeding of the count rate compared to blank measurements, which displays the number of photons detected per second and therefore, serves as an indicator for signal strength and sample quality.

Figure 7.3 summarizes the results of the analyzed compounds. The average count rates are shown in Figure 7.3A. While **11** exhibits an average count rate in the range of the blank measurement (24 kHz vs. 27 kHz), clotrimazole exhibits a signal height which is drastically increased by a factor about 100 (3020 kHz). Furthermore, only the autocorrelation function of clotrimazole exhibits a sigmoidal shape (Figure 7.3B). Therefore, DLS experiments suggest that **11** does not form aggregates in solution over the period of an assay measurement at the chosen concentration.

Furthermore, the binding of **11** to an alternative protein was tested in ITC measurements according to an application note suggested by the manufacturer (28-9815-96). Thereby, the ligand was titrated either to the target protein TGT or bovine serum albumin (BSA), respectively.

BSA is commonly used in biochemical assays to reduce the tendency for agglomeration [McGovern *et al.*, 2002]. The mechanism of this interference is not yet fully understood. Titration of **11** to BSA should only yield a binding isotherm if the ligand agglomerates in solution and thus, unspecifically mounts to the BSA surface.

The obtained thermograms are shown in Figure 7.4. As already indicated by DLS also the ITC measurements support the assumption that no aggregate formation has warped the assay results: Only in case of the titration of **11** to TGT a binding isotherm is observed (Figure 7.4A), while the addition to BSA yields only heat signals of dilution (Figure 7.4B). The  $K_d$  value of **11** determined by ITC averages to 7  $\mu\text{M}$ , which is in good accordance with the results of the [8-<sup>3</sup>H]guanine assay.



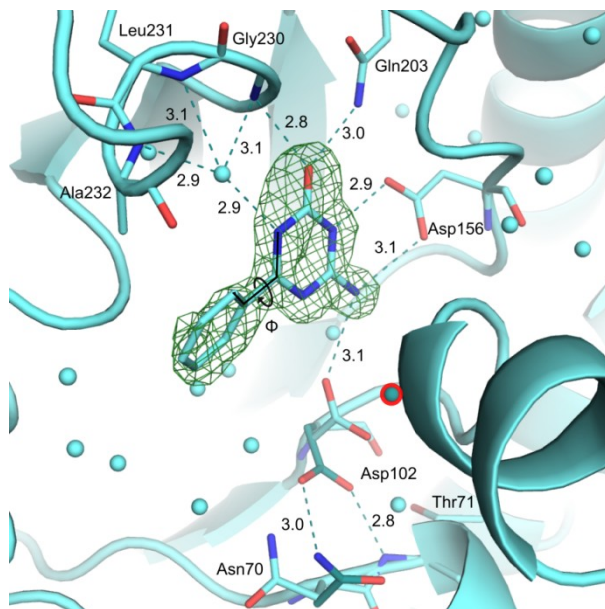
**Figure 7.4** Thermograms for the binding of **11** to TGT (A) and BSA (B). Titrating **11** to its target protein TGT yields an endothermic binding isotherm. Due to the low affinity of the ligand it does not exhibit a sigmoidal shape. The resulting  $K_d$  value is comparable to that observed in the enzyme assay. However, the reaction shows an unreasonable binding stoichiometry. Therefore, a second experiment revealing potentially unspecific binding to the protein surfaces was performed by titrating **11** in a BSA solution. Contrary to A only small peaks of dilution could be observed during this titration. Details concerning the setup of the measurements are listed in the experimental section.

#### 7.2.4 Crystal Structure Analysis

A crystal structure of TGT in complex with **8** was determined at a resolution of 1.62 Å. The 6-phenyl-5-azacytosine scaffold is well defined in the  $|F_o| - |F_c|$  difference electron density (Figure 7.5). Most likely, due to its low solubility, the ligand is not fully populated in the guanine recognition site (84% occupancy). The ligand skeleton intercalates similar to the natural substrate guanine between Tyr106 and Met260 (Figure 7.6). Additionally, it forms hydrogen bonds to the side chains of Asp102, Asp156, Gln203 and the backbone NH of Gly230. A further interaction is mediated via a water molecule between N1 of the 5-azacytosine moiety and the backbone NH of Ala232. As a donor function is missing in this



area no backbone flip is induced as in case of preQ<sub>1</sub> or *lin*-benzopurines [Stengl *et al.*, 2007; Barandun *et al.*, 2012].

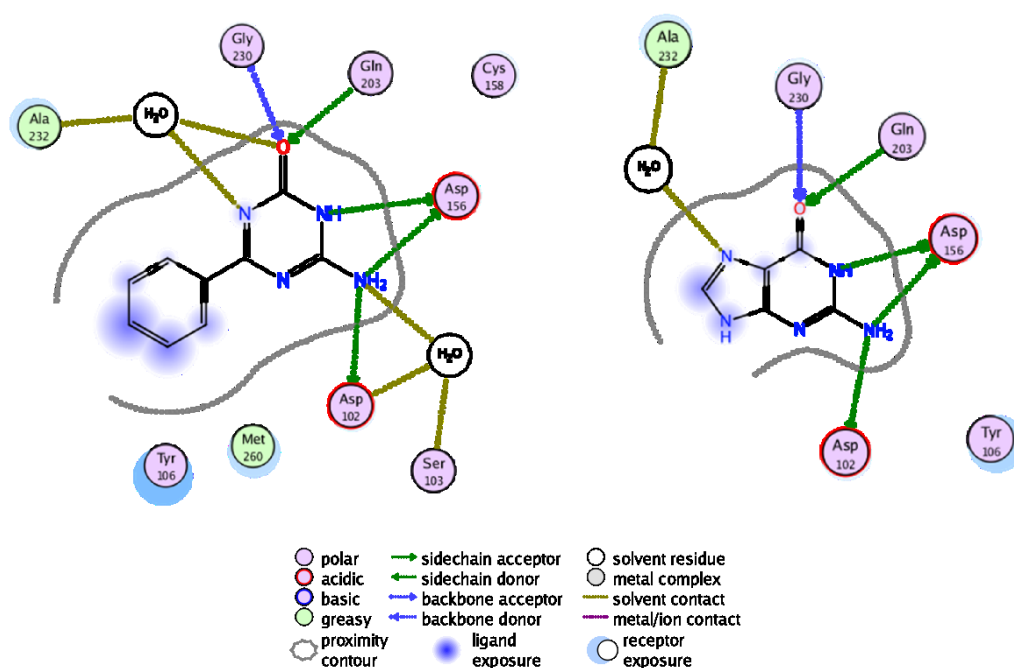


**Figure 7.5** Binding mode of **8** within the guanine-34 recognition site (PDB ID: 4Q4M). The protein is shown as a cartoon, interacting residues as well as **8** in stick representation (carbon = cyan, oxygen = red, nitrogen = blue). Water molecules are represented as spheres. Hydrogen bonds are visualized as dashed lines with their distances given in Å. The 6-phenyl-5-azacytosine scaffold is contoured at  $3\sigma$  in the  $|F_o| - |F_c|$  density map (green) of the structural model refined after removing the ligand. Compound **8** adopts a similar binding mode as the natural substrate guanine with an occupancy of 84%. Several hydrogen bonds are formed to the side chains of Asp102, Asp156, Gln203 and the backbone NH group of Gly230. In addition, the ligand interacts with the backbone of Gly230, Leu231 and Ala232 via a close-by water molecule. Contrary to guanine, **8** does not show a fully planar binding geometry and the phenyl moiety exhibits a torsion angle  $\Phi$  of about  $35^\circ$  with respect to a plane through the heterocycle. Furthermore, the side chain of Asp102 adopts two different positions: In its ligand-facing orientation it is able to interact with the exocyclic amino function of the 5-azacytosine ring via one hydrogen bond (cyan, 42%). The second orientation is similar to that in TGT apo structures forming hydrogen bonds to the side chain of Asn70 and the backbone of Thr71 (teal, 58%). The water molecule encircled in red will only be present, if Asp102 is rotated out of the pocket.

Unlike previously known substrates and ligands, the ring system of **8** does not adopt a planar geometry at the binding site. The phenyl ring is rotated approximately  $35^\circ$  out of plane of the heterocycle (Figure 7.5). Furthermore, the hydrogen bond formed to Asp102 suggests suboptimal interaction geometry. In the binary complex of TGT and guanine (PDB

## 7. 5-Azacytosine as a Novel Scaffold to Inhibit Z. mobilis TGT with Expected Improved Bioavailability and Synthetic Accessibility

ID: 2PWU) the substrate shows an occupancy of 75% [Tidten *et al.*, 2007]. Asp102 is present in two conformations, one facing the ligand by establishing two parallel hydrogen bonds to N3 and the exocyclic amino function of guanine (75%). The second conformer (25%) is rotated out of the pocket interacting with the side chain of Asn70 and the backbone NH of Thr71, an orientation also found in the apo enzyme [Brenk *et al.*, 2003]. The structure indicates that the rotation of Asp102 into the binding pocket is ligand-induced and correlates with the observed occupancy of the compound.

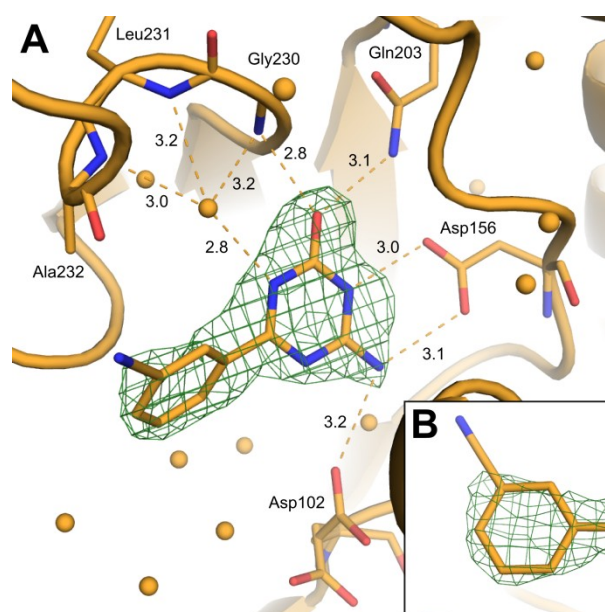


**Figure 7.6** 2D ligand interaction diagram of **8** and guanine generated with MOE. The two ligands exhibit similar interaction patterns within the guanine-34 recognition site forming hydrogen bonds to the side chains of Asp102, Asp156, Gln203 and the backbone NH group of Gly230. The water mediated interaction between the backbone NH group of Ala232 and N1 of the 5-azacytosine **8** is performed by N7 in case of guanine as the binding partner.

Although **8** is occupied to 84%, the higher populated conformation A of Asp102 is rotated out of the guanine recognition site (58%). In that case, a close-by water molecule is observed adjacent to the aspartate side chain (Figure 7.5). The other conformer (42%) of Asp102 orients towards the ligand and the above-mentioned water molecule is displaced. However, the side chain of Asp102 adopts a conformation orthogonal to a best plane

through **8**. Only one H-bond is formed to the carboxylate oxygens via the exocyclic amino group of the 5-azacytosine. Due to the one-sided fixation of the its scaffold, **8** exhibits a rather loose binding with an average B-factor of  $26.4 \text{ \AA}^2$  showing its highest flexibility at the phenyl portion.

For complex TGT·**9** a maximum resolution of  $1.79 \text{ \AA}$  was achieved. The binding mode observed for **8** is also found for **9** (Figure 7.7A). The occupancy of the common 6-phenyl-5-azacytosine core amounts to 84%. Unfortunately, sufficient difference electron density for the nitrile substituent in meta position is lacking. A small bulge next to the phenyl ring suggests that the nitrile moiety points towards the ribose-33 subpocket (Figure 7.7B).



**Figure 7.7** Binding mode of **9**. The same options for visualization have been applied as in Figure 7.5. A) The binding mode of the 6-phenyl-5-azacytosine scaffold of **8** is confirmed by **9**. The heterocycle is clearly defined by the  $|F_o| - |F_c|$  density map (green) showing an occupancy of 84%. However, adequate difference electron density is missing for the attached nitrile substituent at the phenyl ring. B) Close-up of the phenyl moiety and its nitrile substituent in meta position. A small bulge of the difference electron density suggests the nitrile functionality to point towards the adjacent ribose-33 subpocket.

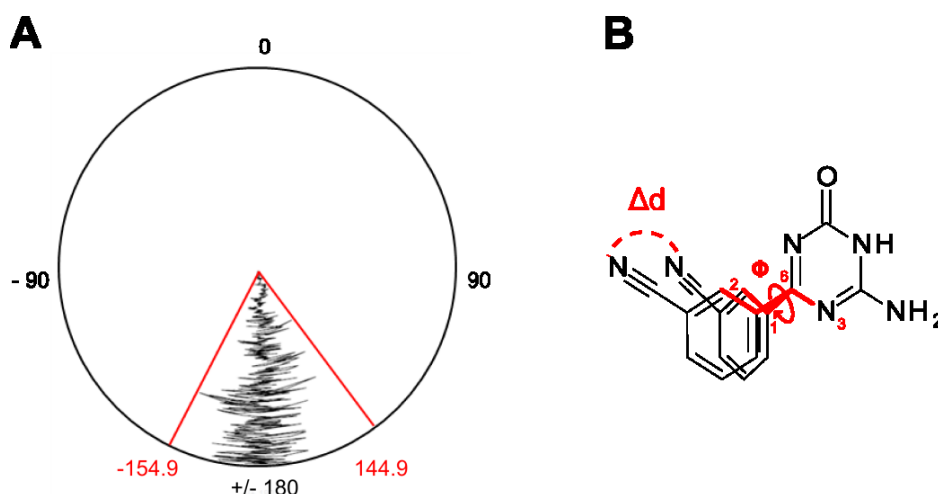
Again, Asp102 adopts the previously described conformations. The occupancy of the side chain conformer facing the ligand amounts to 46%. Thus, the orientation of Asp102 resembling the one, which is found in the apo enzyme, also represents the predominant conformation in this complex. Similarly, the ligand shows elevated B-factors of  $36.3 \text{ \AA}^2$ .

The binary complex of TGT·**10** shows a resolution of 1.76 Å. The 6-phenyl-5-azacytosine core refines to 88% occupancy. The position of the aminomethyl substitution cannot be clearly assigned to the  $|F_o| - |F_c|$  electron density due to an enhanced residual flexibility of this ligand portion with an average B-factor of 48.3 Å<sup>2</sup>. The conformer of Asp102 oriented by a monodentate H-bond towards the 5-azacytosine moiety represents 46% population.

Data sets for TGT·**11** and TGT·**12** were collected with a resolution of 1.35 Å and 1.74 Å, respectively, including crystals grown under varying co-crystallization conditions. Unexpectedly, difference electron density neither for the 5-azacytosine core nor for the substituted phenyl ring could be observed, probably due to low solubility (compound **11**) or rather reduced potency (compound **12**).

### 7.2.5 MD Simulation

To investigate the properties of the 6-phenyl-5-azacytosine scaffold within the guanine recognition site and to get an idea about its flexibility considering the meta substituent, a short molecular dynamic (MD) simulation of 10 ns was carried out for the complex TGT·**9**.



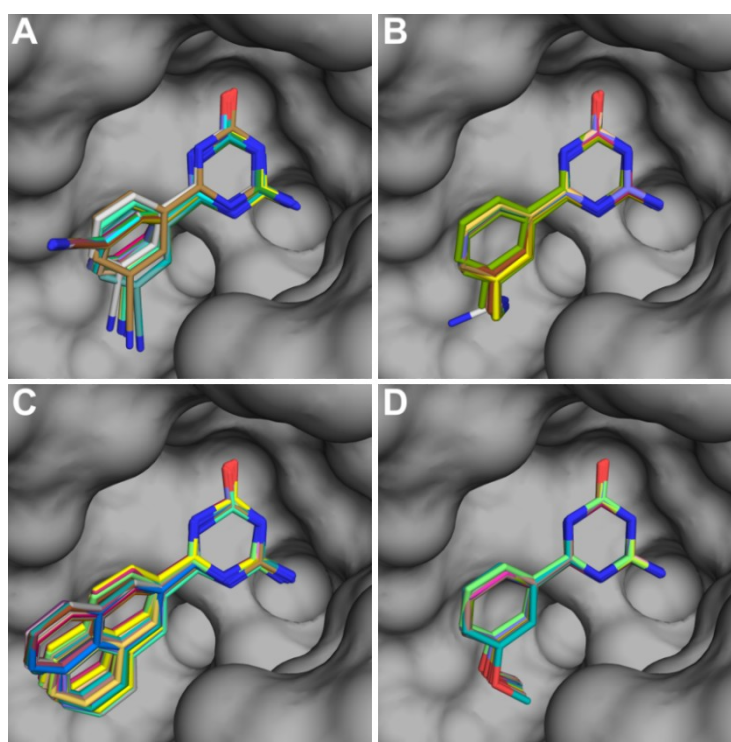
**Figure 7.8** Results of the MD simulation performed for **9**. A) Dialplot tracing the movement of the phenyl moiety in 6-position along the trajectory. The radius of the plot represents the time scale of the simulation, the circumference the value for the calculated dihedral. The plotted dihedral ranges between -154.9° and 144.9°. B) Visualisation of the distance, which is covered by the nitrogen of the nitrile group due to the variance in the torsion angle marked in red.  $\Delta d$  is calculated as segment along the circular arc.

The trajectory was analyzed by means of the torsion angle formed between N3 and C6 of the 5-azacytosine moiety and C1 and C2 of the attached phenyl ring (Figure 7.8B, red). The results are visualized in terms of a dialplot in Figure 7.8A.

The torsion angle oscillates by about  $60^\circ$  between extrema of  $144.9^\circ$  and  $-154.9^\circ$  with an average value of  $171.2^\circ$ . This torsion oscillation translates into a wiggling motion of the nitrile group swinging along a circular arc (Figure 7.8B). Overall, a maximum scatter of  $3.8 \text{ \AA}$  is covered by the carbon and  $4.9 \text{ \AA}$  by the nitrogen atom of the nitrile group explaining why no properly defined density is visible for this part of the molecule.

### 7.2.6 Docking Solutions

The crystal structures of the various TGT-ligand complexes do not provide a satisfactory explanation for the affinity differences of the inhibitors **8** – **12**. To obtain additional insight into possible binding modes, **9** – **12** were docked into the guanine recognition site. The suggested binding modes are summarized in Figure 7.9.

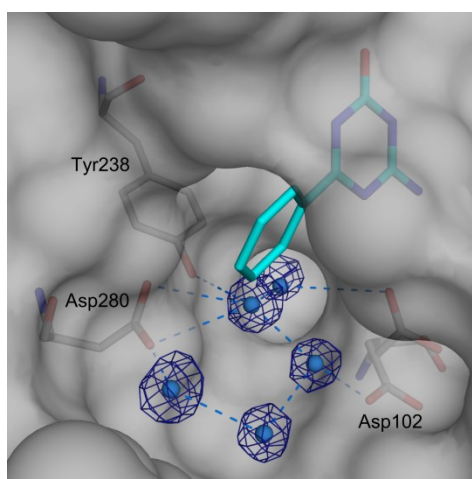


**Figure 7.9** Docking solutions of ligands **9** – **12** generated by *GOLD*. The protein surface is shown in gray except for Tyr106 and Cys158, which are omitted from the plot for clarity. The ligands are visualized in stick representation (oxygen = red, nitrogen = blue).

## 7. 5-Azacytosine as a Novel Scaffold to Inhibit *Z. mobilis* TGT with Expected Improved Bioavailability and Synthetic Accessibility

The scaffold intercalates between Tyr106 and Met260 and hydrogen bonds are formed to the side chains of Asp102, Asp156, Gln203 as well as the backbone NH of Gly230. As expected, the docked binding mode of **9** suggests the nitrile substitution to be oriented towards the uracil-33-subpocket without forming any directional interaction. The torsion angle between the 5-azacytosine and the phenyl ring averages to 50°. However, also the alternative orientation of the nitrile group towards the ribose-34 pocket is observed on the less well ranked docking solutions.

In contrast to **9**, docking suggests one consistent binding mode for **10** placing the aminomethyl substituent into the ribose-34 subpocket next to the side chains of Asp102 and Asp280. Thereby, the substituent displaces a water molecule and participates in interactions with a water cluster placed between the two aspartates (Figure 7.10).

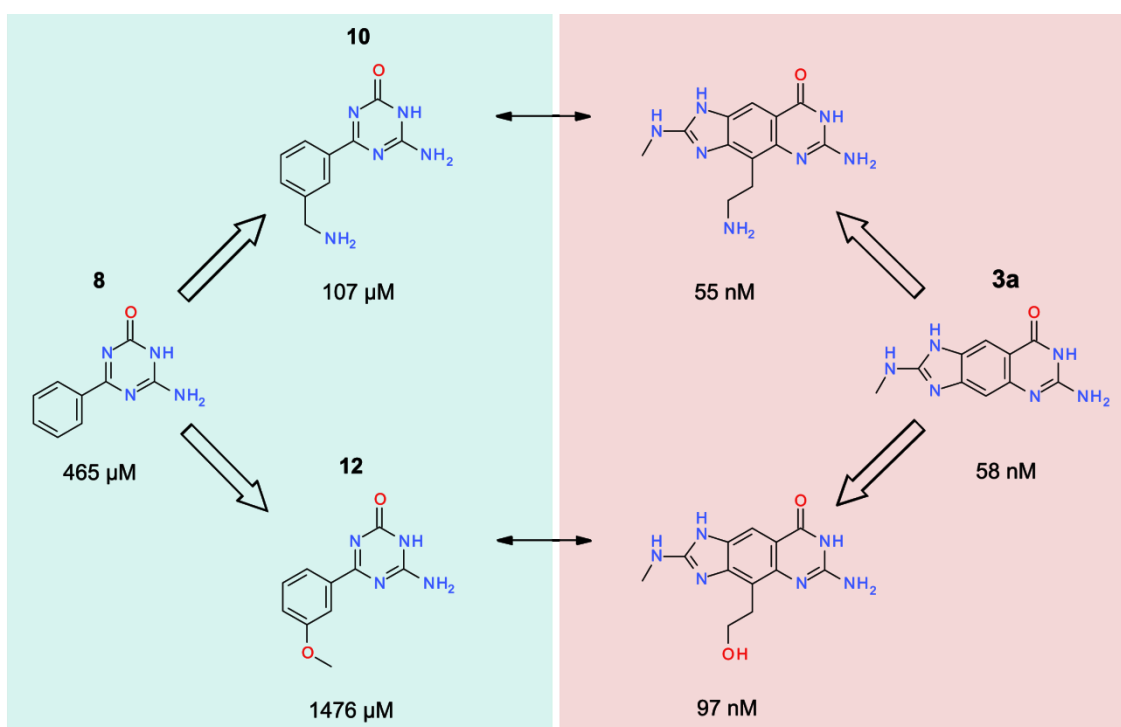


**Figure 7.10** Five-membered water cluster located between Asp102 and Asp280 in TGT-**8**. The protein surface is shown in gray. Water molecules are represented as blue spheres and contoured at  $1\sigma$  in the  $2|F_o| - |F_c|$  density map (dark blue). Compound **8** and residues forming hydrogen bond contacts to the water cluster are visualized in stick representation (carbon, gray/cyan; oxygen, red; nitrogen, blue). Hydrogen bonds are indicated by blue dashes up to a maximum distance of 3.6 Å.

In these docking solutions the torsion angle between the 5-azacytosine and the phenyl ring amounts to 10° to 25°. The enhanced binding affinity of **10** is in agreement with this binding mode and mirrors the properties of similarly substituted *lin*-benzoguanine inhibitors. In this series, an attached terminal amino group also replaces a water molecule and participates in a cluster of water molecules located between Asp102 and Asp280

without a loss in affinity compared to the unsubstituted parent structure (Figure 7.11) [Ritschel *et al.*, 2009]. In case of **10** even a gain in affinity by a factor of 4.3 is noticed.

The loss in binding affinity of **12** by a factor of 3.2 can be explained by similar considerations. The meta methoxy substituent at the phenyl ring is placed into the ribose-34-subpocket upon docking. Thereby, it disturbs the mentioned water cluster and experiences unfavorable contacts to the most likely charged side chains of Asp102 and Asp280, which explains its loss in binding affinity.



**Figure 7.11** Influence of various substituents placed into the ribose-34 subpocket on the binding affinity to *Z. mobilis* TGT. On the left hand side, ligands **8**, **10** and **12** representing 5-azacytosine type inhibitors are displayed while the right hand side shows members of the *lin*-benzoguanines with similar substituents at the 4-position of the scaffold. In each case the compound in the center of the scheme serves as a reference lacking a substitution, which binds to the ribose-34 subpocket. In both ligand series a substituent exhibiting a primary amino function is related to affinity gain. This functionality displaces a water molecule of the network located between Asp102 and Asp280 and is able to actively interfere with this network.

For the placement of the naphthyl substituent **11** docking suggests two different binding modes. For the better ranked docking solutions a coplanar orientation of the naphthyl ring with the heterocycle is found with deviations from a common plane of 25° and

40°. The less well ranked solutions show the naphthyl substituent in perpendicular orientation to the cytosine moiety. This orientation of the naphthyl ring appears less favourable as it penetrates between Asp102 and Asp280 and disturbs the five-membered water cluster resulting in a loss of binding affinity.

### 7.3 Conclusions

In the present chapter a new class of fragment-like TGT inhibitors with potentially increased bioavailability has been studied. The 5-azacytosine scaffold was selected as it should exhibit similar features as the closely related natural substrate guanine. Considering the key interactions to Asp102 and Asp156, which are prerequisite to induce a protonation change of the ligand and therefore provoke the potent binding of a charged ligand was assumed to be transferable to the 5-azacytosines. Another advantage of the compound class is its easy synthetic accessibility allowing for the production of multiple derivatives.

Unfortunately, the newly introduced unsubstituted 6-phenyl-5-azacytosine scaffold showed only binding affinity in the three-digit micromolar range. Further decorations in meta position did not yield a satisfactory increase in potency. Successfully accomplished crystal structures of **8**, **9**, and **10** indicate that the interaction to Asp102 is not optimally formed to the 5-azacytosine scaffold. The latter amino acid has been shown to exhibit a key role in substrate recognition, ligand protonation and pocket crosstalk (Chapter 4). Especially the protonation provoked by a strong  $pK_a$  shift at N(5) in case of the *lin*-benzguanines resulted in a substantial increase in potency due to the formation of two bidentate salt bridges between Asp102, Asp156 and the ligand [Neeb *et al.*, 2014]. Additionally, the  $pK_a$  value assigned to the change in protonation at N(5) is essential for the protonation behaviour of the ligand and the generation of a charged species. In contrast to the *lin*-benzguanines, the  $pK_a$  value of the 5-azacytosine moiety is lowered by almost one logarithmic unit from approximately 4.4 to 3.5 [Tyagi *et al.*, 2003; Barandun *et al.*, 2012].

This difference makes a  $pK_a$  shift induced by the two adjacent aspartate residues rather unlikely. In consequence of the absent protonation of the 5-azacytosine moiety no charged species seems to be created and the salt bridges to the protein are supposedly not formed.



In summary, as shown by crystal structure analyses and assay measurements, the low  $pK_a$  value of the novel parent scaffold paralleled by the deviation from planarity of the attached phenyl ring most likely results in a perturbation of the key interaction to Asp102. In consequence, only unsatisfying LEs are observed for this scaffold. Obviously, the formation of salt bridges to Asp102 and Asp156 are essential for potent binding.

Surprisingly, the naphthyl derivative **11** shows significantly improved binding with a remarkable LE of  $0.41 \text{ kcal mol}^{-1}$  similar to those of *lin*-benzoguanine type inhibitors. Unfortunately, a crystal structure of TGT in complex with this ligand could not be determined. Accordingly, it remains purely speculative to assume that this ligand exhibits properties to induce a sufficient  $pK_a$  shift to bind in protonated state. At least such behavior could explain the substantial increase in binding affinity. Despite good LE of **11**, the compound appears less suited for further optimization due to its low solubility and the danger of aggregation at higher concentrations even though at lower concentrations this behaviour could be ruled out by DLS experiments. As general strategy, polar groups are introduced to a fragment to allow for the formation of hydrogen bonds to the protein and to achieve sufficient solubility. In later design cycles hydrophobic groups can be attached to the lead fragment. In order to follow this basic rule, the design of alternative fragment-like TGT inhibitors should focus on an ammeline scaffold, which exhibits an elevated  $pK_a$  of 4.5 [Schmitt *et al.*, 1997], and an attached second amino group suitable to avoid unfavourable contacts to Asp102.

## 8. Summary and Outlook

The present thesis deals with structure-based design, in particular the characterization and improvement of selective antibiotics targeting the enzyme tRNA–guanine transglycosylase. TGT plays a key role in the pathogenicity of *Shigella*, the causative agent of Shigellosis, performing a base exchange reaction that yields modified tRNA. This modification is essential for *Shigella* to invade the epithelial cells of the colon. The worldwide increasing resistance of *Shigella* species against currently administered antibiotics requires the urgent development of novel drugs for the treatment of Shigellosis especially to cure patients in developing countries.

Recently, the *lin*-benzopurine scaffold was introduced as promising starting point for the structure-based design of TGT inhibitors. Two classes, namely the *lin*-benzoguanines and *lin*-benzohypoxanthines varying in an exocyclic amino functionality, were found to inhibit the target enzyme in the nanomolar range. However, the analyzed molecules did not yet show ideal drug metabolism and pharmacokinetic features since firstly they showed poor permeation through cell membranes in PAMPA measurements and secondly their attached 2-substituents binding to the ribose/uracil-33 subpocket were poorly defined in the difference electron density in crystal structures complicating the establishment of a structure–activity relationship.

In a comprehensive study the first issue was addressed by focusing on the protonation inventory of *lin*-benzopurines prior to and after complex formation. Different *in vitro* and *in silico* methods comprising isothermal titration calorimetry, site-directed mutagenesis and Poisson-Boltzmann calculations were applied to identify and verify the site of protonation.

Prerequisite for this assignment is the set-up of ITC measurements appropriate to analyze the buffer dependency of the *lin*-benzopurines, which in return allows concluding the proton uptake or release by the system under investigation (Chapter 2). Subsequently, the robustness of the settings was tested by investigating the influence of protein and ligand impurities on the derived thermodynamic profiles. The latter comprised organic and inorganic impurities. Due to the low solubility of *lin*-benzopurines combined with unfavorable polar properties such impurities are complicating the preparation of pure compounds. Accompanied by measurements including the second model system trypsin, our

investigations showed that protein impurities are not necessarily influencing the thermodynamic properties of high affinity ligands binding to the target protein. Their presence is rather reflected in the binding stoichiometry. However, caution has to be paid especially in case of analyzing unknown binding reactions since incorrect binding stoichiometry of the interacting species can be suggested. Contrary to protein impurities, impure ligand preparations were found to influence the partitioning of the enthalpic and entropic term significantly while the binding free energy remains largely unaffected. In case of an impurity exhibiting similar binding properties to the intended ligand the actual thermodynamic profile cannot be deduced from ITC measurements and further purification is inevitable. Even organic and inorganic impurities that are not able to interact with the target protein can influence the thermodynamic properties drastically. In the fortunate case of knowledge about the exact composition of the impurity, elementary analysis provides a helpful tool to support re-fitting the derived thermodynamic data yielding a good approximation to the actual enthalpic and entropic terms. This procedure, however, should be applied with caution since only a small deviation is accepted until changes in the thermodynamic signature are considered as significant. Especially in congeneric series, which might show only marginal differences in their binding enthalpy and entropy, sole re-fitting may lead to a misinterpretation of data. Instead, more effort should be spent into the purification protocol of the ligands. If pure ligands cannot be obtained at all, reverse titrations might be consulted. However, this set-up also bears several problems like the requirement of large amounts of pure protein, which must be stable in the applied buffer system over the period of titration. For this purpose, dynamic light scattering has been shown to provide valuable insights to access whether protein or ligand aggregation influences the experiments.

The established ITC technique has then been used to analyze the protonation inventory of the complex formation between *lin*-benzopurines and TGT (Chapter 3). Initial ITC measurements performed with *lin*-benzoguanine-type ligands suggested the uptake of one proton by either the protein or the ligand. Focusing on the ligand side, subsequent  $pK_a$  calculations identified two basic nitrogens within the guanidine moiety of either the aminoimidazole or the aminopyrimidinone portion of the parent scaffold exhibiting significant changes in their  $pK_a$  properties upon complexation. Even though being the less basic group, the protonation of the aminopyrimidinone moiety was suggested to be more

likely; however, the results obtained for the aminoimidazole did not allow to fully exclude protonation at this site. Thus, we analyzed different ligands in further ITC experiments, each of them excluding one of the basic sites by chemical modifications. Surprisingly, the *lin*-benzoguanines exhibited a distinct buffer dependency while the *lin*-benzohypoxanthines did not show this behavior. To explain these results crystal structures of both inhibitor classes were consulted. A major difference in their binding modes was observed for the aminopyrimidinone portion: While the *lin*-benzohypoxanthines bind to a TGT conformation closely similar to the apo enzyme interacting with only one aspartate within the guanine-34 recognition site, addition of the exocyclic amino functional group in case of the *lin*-benzoguanines induces the rotation of a second aspartate towards the binding pocket. Obviously, the negatively charged environment of both aspartates in short distance provokes a  $pK_a$  shift in case of the *lin*-benzoguanines strong enough to induce the uptake of a proton. If only one aspartate residue is oriented towards the binding pocket this shift is not sufficient for protonation as shown for the *lin*-benzohypoxanthines. This hypothesis could be confirmed by generating mutants with the aspartates replaced by asparagines, respectively. Performing similar ITC experiments titrating the *lin*-benzoguanines to the Asn-mutants did not show the previously observed buffer dependency anymore. Previous *in silico* studies concerning the same issue proposed that the aminoimidazole portion of the parent scaffold would also change protonation state in the protein-bound state. According to our ITC measurements, this moiety would only become protonated upon binding if a protein residue simultaneously released a proton. The sole candidate to feature this property is a proximal glutamate residue, which is known to change its protonation state depending on the bound ligand or the applied environmental pH value. pH-Soaking experiments unravelled that this residue is already deprotonated at the applied pH value and thus, the aminoimidazole portion must remain in uncharged state. The gain in binding affinity previously ascribed to the protonation event could be attributed to an array of two parallel H-bonds directly formed with two adjacent C=O groups of the protein backbone avoiding secondary repulsive interactions.

PAMPA measurements unravelled that the parent scaffold of the *lin*-benzoguanines exhibits a too polar character to permeate through the membrane of cells. Taking the present studies into account, it does not come as a surprise that the *lin*-benzohypoxanthines do not show an improvement of the PAMPA scores. Future design strategies must address

the polarity of the tricyclic scaffold focusing on the aminoimidazole portion. An approach, in which this part of the molecule is masked by appropriate functional groups turning it into a pro-drug, could help to achieve the desired partitioning.

Considering the binding affinities across the series of *lin*-benzopurines, a rather flat structure–activity relationship is observed. Therefore, additional insight into the driving forces of binding was gained by factorizing the free binding energy into enthalpy–entropy contribution (Chapter 4). A modified crystallization protocol enabled to spot the flexible 2-substituent of the analyzed 2-amino-*lin*-benzopurines in most cases yielding the basis to link structural features to the derived thermodynamic data. Crystal structures unravelled that the *lin*-benzohypoxanthines bind to TGT in a state closely similar to the apo enzyme showing Asp102 in an orientation oriented off from the binding pocket. In contrast to that, the *lin*-benzoguanines induce the rotation of Asp102 towards the binding pocket. Thus, a proton is picked up by the aminopyrimidinone moiety of the parent scaffold then forming an additional bidentate salt-bridge to Asp102. Triggered by the movement of Asp102 several amino acid residues and secondary structure elements rearrange and transform the protein to a conformation as similarly found in the tRNA-substrate bound state. These changes were found to result in major differences determining the binding affinity in a biochemical assay, which yields an inhibitory constant ( $K_i$ ), and ITC measurements, from which a dissociation constant ( $K_d$ ) can be derived. In the biochemical assay the incorporation of radioactively-labeled guanine into tRNA is measured. As a prerequisite, TGT is required in its substrate-bound state. Thus, *lin*-benzohypoxanthines appear less potent since the protein has to be transformed into the substrate-bound state first. Contrary, ITC measurements record the change in heat released or absorbed during the binding reaction and a pure dissociation constant of the binding equilibrium to the required protein conformer is derived. Thereby, TGT binds the *lin*-benzohypoxanthines in its apo state showing an increased binding potency.

As expected, bindings of both series, *lin*-benzoguanines and *lin*-benzohypoxanthines, were found to be enthalpy-driven. Thereby, the *lin*-benzohypoxanthines exhibit a less pronounced enthalpic term due to their missing interaction to Asp102, which can be partly compensated by a crystallographically conserved water cluster located at the bottom of the guanine-34 recognition site. While the thermodynamic profiles of the *lin*-benzohypoxanthine series remains nearly unchanged, data for the *lin*-benzoguanine series are found to be quite diverse. Thus, assigning the contribution of the attached 2-substituent was hardly possible.

Obviously, the structural changes triggered by Asp102 in case of the *lin*-benzoguanines enable a cross-talk between ribose/uracil-33 subpocket addressed by the 2-substituent and the guanine-34 recognition site occupied by the parent scaffold. Helix  $\alpha$ A located nearly perpendicular to the bound ligands was shown to play a crucial role in mediating the interactions between both pockets. After the *lin*-benzoguanines are recognized by the protein via Asp102 as substrate-like this helix is, amongst others, changing its position as observed similarly in tRNA·TGT complexes. Accompanied by this movement, Tyr106 penetrates deeper in the ribose/uracil-33 subpocket to experience whether this part of the protein is occupied. This procedure guarantees to differentiate correct from false substrates.

Further design cycles should consider that ligands addressing the ribose/uracil-33 subpocket have to be recognized as substrate-like triggering the rotation of Asp102 into the guanine-34 binding pocket in order to provoke the described cross-talk. Additionally, only the thus discriminated ligands are able to modulate binding affinity by forming bidentate salt-bridges to both aspartates, if their  $pK_a$  values fall into a range appropriate to induce a protonation change.

Based on the observations that 2-amino-*lin*-benzoguanines show a reduced occupancy combined with elevated temperature factors of the 2-substituent, a high flexibility of that part of the molecule was assumed. Therefore, we tested whether a binder with high residual mobility can avoid a loss in binding affinity compared to a binder adopting one ordered binding mode and therefore achieve an advantage as a more competent antibiotic less affected by resistant mutations (Chapter 5). After selecting Val262 within the ribose-33 subpocket as mutation site, three different mutants were expressed (Val262Thr, Val262Asp, Val262Cys) and crystallized. Subsequently, the binding affinities ( $K_d$ ) of various 2-amino-*lin*-benzoguanines were determined and discussed with respect to structural information derived from crystal structures of the TGT mutants in complex with reference **4a**.

Compared to the wild type, the analyzed 2-amino-*lin*-benzoguanines showed the same binding affinity towards variant Val262Thr. This result was expected since the properties of the pocket were only slightly changed. Contrary, potencies towards variant Val262Asp decreased significantly while the  $K_d$  values towards variant Val262Cys increased. The observed changes could be related to the binding of the parent scaffold rather than to the different 2-substituents inducing disorder of the protein in proximal distance to the

mutation site. Only marginal differences could be ascribed to the properties of the substitution pattern in case of Val262Asp, most likely due to electrostatic attractions and repulsions, respectively. Nevertheless, this fact indicates that flexible portions of a drug are able to modulate binding affinity in the unfortunate situation of resistance mutations.

In chapter 6 molecular dynamic simulations were used to predict the binding modes of extended 2-amino-*lin*-benzoguanines, which could not be sufficiently described by docking results in previous studies. Unfortunately, MD simulations succeeded only in one out of four cases to reliably describe the binding mode of the ligand in a way to match with the determined crystal structure. Nevertheless, subsequent crystal structure analyses using modified crystallization conditions resulted in reasonable difference electron density for all extended 2-amino-*lin*-benzoguanines. The refined data sets unravelled novel aspects important for the further design and characterization of TGT inhibitors:

Firstly, using a modified crystallization protocol we were able to spot the 2-substituent of the extended 2-amino-*lin*-benzoguanines, which bind to the ribose-32 subpocket that has never been occupied before. The collected data sets provide the basis for new synthesis strategies forming novel interactions within this part of the protein that is better shielded than its neighboring solvent-exposed environment. In that context, another ligand binding to the ribose-33 subpocket via a bidentate salt bridge showed a reduced binding potency. This example unravelled that solvent-exposed interactions formed between ligand and protein in the flat ribose-33 subpocket, if at all, hardly contribute to binding affinity since high desolvation costs have to be paid in this environment of high dielectric conditions. Therefore, further drug design cycles addressing this part of the protein have to account for an optimization of rather hydrophobic 2-substituents.

Secondly, the results of this chapter emphasize the importance of the applied crystallization conditions. Poorly defined electron density does not always indicate ligands or substituents exhibiting high mobility in the protein-bound state. The applied crystallization protocol takes a major impact on the derived difference electron density and has to be considered in the discussion of the obtained structures.

The class of 5-azacytosines was investigated as a novel scaffold to inhibit TGT with potentially increased bioavailability (Chapter 7). Similarly as the *lin*-benzopurines, also the 5-azacytosines are deduced from the natural substrate guanine and were expected to

establish similar binding features. An attached phenyl ring serves as an anchor point for substituents reaching into the neighbored ribose-33 and ribose-34 subpockets. Compared to the *lin*-benzoguanines, this new class of inhibitors offers the advantage of better synthetic accessibility. The  $pK_a$  value of its aminopyrimidinone moiety (N(5)), however, is found between the one of *lin*-benzohypoxanthines and *lin*-benzoguanines and represented a crucial issue to be analyzed in respect to its protonation behavior.

Unfortunately, the ligands show only binding affinities in the low micromolar range except of one compound decorated with a naphthyl moiety, which exhibits a binding potency in the one-digit micromolar range similar as the closely related guanine. Crystal growth proved to be difficult due to the low solubility of all compounds. Nevertheless, crystal structures could be solved for three out of five ligands. As expected, the 5-azacytosine scaffold binds to the guanine-34 recognition site similarly as guanine, however, the key interaction to Asp102 that was proven to be of utmost importance for substrate recognition, ligand protonation, and pocket crosstalk is poorly established. The reason for the disturbance of this interaction can be deduced from the binding mode of the 6-phenyl-5-azacytosines: Different from all known TGT inhibitors, they do not bind to the protein in a planar fashion since the attached phenyl ring rotates out of the plane defined by the heterocycle. In consequence, the – compared to *lin*-benzoguanines – low  $pK_a$  cannot be shifted into a window appropriate for ligand protonation by the interactions formed to Asp102 and Asp156 and the drop in affinity can be explained. For the more potent naphthyl derivative a crystal structure with bound ligand could not be determined. Up to this point it remains unclear, whether the naphthyl substituent induces the conformational transition of Asp102 to form the desired bidentate salt bridge to the parent scaffold and thereby, induces the protonation step necessary to increase binding affinity.

Despite its promising potency, the naphthyl derivative is not an appropriate starting point for further optimization since it tends to form aggregates at higher concentrations providing significant problems in its handling during synthesis and subsequent analysis. Instead, further design cycles should concentrate on the closely related ammelines. An additional exocyclic amino functionality attached to the heterocycle does not only represent a linker region creating more space towards the side chain of Asp102 but also increases the  $pK_a$  value about one logarithmic unit thus likely falling into a window that allows protonation of the parent scaffold.



## 8. Zusammenfassung und Ausblick

Die vorliegende Arbeit behandelt das strukturbasierte Wirkstoffdesign, insbesondere die Charakterisierung und Optimierung selektiver Antibiotika gegen das Enzym tRNA-Guanin Transglykosylase. Die TGT spielt eine Schlüsselrolle in der Pathogenität von *Shigellen*, den Erregern der Shigellenruhr, indem sie eine Basenaustauschreaktion katalysiert, die modifizierte tRNA liefert. Diese Modifizierung ist essenziell für *Shigellen*, um die Epithelzellen des Darms zu befallen. Weltweit steigen die Resistenzen von *Shigellen* gegen gängige Antibiotika an, sodass dringend neue Wirkstoffe zur Behandlung der Shigellenruhr insbesondere von Patienten in Entwicklungsländern benötigt werden.

Kürzlich wurde das *lin*-Benzopurin-Grundgerüst als vielversprechender Ausgangspunkt für das strukturbasierte Design von TGT-Inhibitoren eingeführt. Zwei Klassen wurden als potente Inhibitoren im nanomolaren Bereich identifiziert – die *lin*-Benzoguanine und *lin*-Benzohypoxanthine, die sich in einer exozyklischen Aminofunktion voneinander unterscheiden. Allerdings zeigen die untersuchten Moleküle noch keine idealen Eigenschaften in Bezug auf Metabolismus und Pharmakokinetik: Zum einen weisen sie eine geringe Permeation durch Zellmembranen in PAMPA-Messungen auf. Zum anderen ist ihr angefügter Substituent in 2-Position, der in der Ribose/Uracil-33 Tasche bindet, nur schwach durch Differenzelektronendichte in entsprechenden Kristallstrukturen definiert. Dies erschwert es eine Struktur–Wirkungsbeziehung aufzustellen.

In einer umfassenden Studie wurde der erstgenannte Sachverhalt adressiert, indem der Protonierungszustand der *lin*-Benzopurine vor und nach der Komplexbildung betrachtet wurde. Verschiedene *in vitro* und *in silico* Methoden einschließlich Isothermaler Titrationskalorimetrie, ortsgerichteter Mutagenese und Poisson-Boltzmann-Berechnungen wurden angewendet, um die Protonierungsstelle zu identifizieren und zu verifizieren.

Voraussetzung dafür ist die Etablierung von ITC-Messungen, die sich eignen die Pufferabhängigkeit der *lin*-Benzopurine zu analysieren, die wiederum Rückschlüsse auf die Protonenaufnahme bzw. –abgabe des betrachteten Systems erlaubt (Kapitel 2). Anschließend wurde die Robustheit dieser Versuchsbedingungen getestet, indem der Einfluss von Protein- und Ligandverunreinigungen auf die erhaltenen thermodynamischen

Daten untersucht wurde. Diese enthielten sowohl organische als auch anorganische Verunreinigungen. Verschuldet durch die geringe Löslichkeit und ungünstige Polarität der *lin*-Benzopurine erschweren solche Verunreinigungen die Darstellung reiner Präparate. Zusammen mit Messungen an dem zweiten Modellsystem Trypsin wurde deutlich, dass Verunreinigungen des Proteins nicht zwangsläufig die thermodynamischen Eigenschaften von hochaffinen Liganden beeinflussen. Ihr Vorhandensein wird vielmehr in der Bindungsstöchiometrie widerspiegelt. Dieses Vorgehen muss jedoch mit Vorsicht betrachtet werden, insbesondere wenn unbekannte Bindungsreaktionen untersucht werden, da Rückschlüsse auf eine falsche Bindungsstöchiometrie der interagierenden Partner gezogen werden können. Im Gegensatz zu Proteinverunreinigungen beeinflussen verunreinigte Ligandpräparate die Aufspaltung in Enthalpie und Entropie signifikant, während die Freie Bindungsenergie größtenteils unberührt bleibt. Weist eine Verunreinigung ähnliche Bindungseigenschaften wie der beabsichtigte Ligand auf, kann das tatsächliche thermodynamische Profil nicht aus den ITC-Messungen abgeleitet werden und eine weitere Aufreinigung ist unumgänglich. Aber auch organische und anorganische Verunreinigungen, die nicht in der Lage sind mit dem Zielprotein zu interagieren, können sich drastisch auf die thermodynamischen Eigenschaften auswirken. Eine Elementaranalyse kann eine hilfreiche Methode darstellen, um die erhaltenen Daten so anzupassen, dass man eine gute Annäherung an die tatsächlichen enthalpischen und entropischen Beiträge erhält. Für dieses Vorgehen ist allerdings die genaue Kenntnis über die Zusammensetzung der Verunreinigung nötig. Jedoch sollte bei diesem Verfahren bedacht werden, dass nur geringfügige Abweichungen akzeptiert werden bis ein Unterschied der Daten als signifikant angesehen wird. Vor allem in Ligandserien, die unter Umständen nur marginale Änderungen in ihrer Bindungsenthalpie und –entropie zeigen, kann eine blinde Anpassung der Daten zu Missinterpretationen führen. Stattdessen sollte das Hauptaugenmerk auf geeigneten Aufreinigungsprotokollen liegen. Wenn eine vollständige Aufreinigung unmöglich sein sollte, sollten reverse Titrations in Betracht gezogen werden. Jedoch birgt dieses Verfahren wiederum neue Problemstellungen wie den Einsatz von großen Mengen reinem Protein, das in dem angewendeten Puffersystem über die Dauer der Titration stabil in Lösung vorliegen muss. Zu diesem Zweck konnte gezeigt werden, dass die Methode der Dynamischen Lichtstreuung wertvolle Einblicke gibt, ob Protein- oder Ligandaggregationen die Experimente beeinflusst.

Die etablierte ITC-Methode wurde anschließend genutzt, um das Protonierungsverhalten während der Komplexbildung zwischen *lin*-Benzopurinen und TGT zu analysieren (Kapitel 3). Zunächst mit *lin*-Benzoguaninen durchgeführte Messungen suggerierten die Aufnahme von einem Proton, entweder durch das Protein oder den Liganden. Mit Fokus auf letzteren konnten mittels  $pK_a$ -Berechnungen zwei basische Stickstoffe der Guanidinstruktur identifiziert werden, die sowohl in dem Aminoimidazol- als auch in dem Aminopyrimidinanteil des Grundgerüsts enthalten ist. Beide Teilstrukturen zeigten signifikante Änderungen ihres  $pK_a$ -Wertes während der Komplexbildung. Obwohl weniger basisch, war eine Protonierung der Aminopyrimidinonstruktur wahrscheinlicher. Die erhaltenen Ergebnisse für die Aminoimidazolstruktur erlaubten es jedoch nicht, diese als Protonierungsort vollständig auszuschließen. Daher wurden verschiedene Liganden in weitere ITC-Experimenten einbezogen, die jeweils eines der basischen Strukturelemente aufgrund chemischer Modifikationen ausschließen. Überraschenderweise zeigen alle *lin*-Benzoguanine eine deutliche Pufferabhängigkeit, während die *lin*-Benzohypoxanthine dieses Verhalten nicht aufweisen. Um dieses Result zu erklären, wurden Kristallstrukturen beider Inhibitor Klassen herangezogen. Ein Hauptunterschied im Bindungsmodus der Aminopyrimidinonpartialstruktur konnte beobachtet werden: Während die *lin*-Benzohypoxanthine an eine dem Apo-Enzym ähnelnde Konformation der TGT binden, bei der die Liganden mit nur einem Aspartat in der Guanin-34-Tasche interagieren, induziert die zusätzliche Aminogruppe im Falle der *lin*-Benzoguanine die Rotation einer zweiten Aspartatseitenkette in Richtung Bindetasche. Offensichtlich verursacht die negativ geladene Umgebung beider Aspartate in kurzer Distanz die Änderung des  $pK_a$ -Wertes im Falle der *lin*-Benzoguanine, die stark genug ist, um eine Protonenaufnahme zu begünstigen. Ist nur ein Aspartatrest zur Bindetasche orientiert, ist diese Änderung für eine Protonierung nicht ausreichend, wie für die *lin*-Benzohypoxanthine gezeigt. Diese Hypothese konnte durch Mutanten abgesichert werden, bei denen jeweils ein Aspartat- gegen einen Asparaginrest ausgetauscht wurde. In analogen ITC-Experimenten, in denen die *lin*-Benzoguanine zu der jeweiligen Asn-Mutante titriert wurden, konnte die zuvor festgestellte Pufferabhängigkeit nicht weiter beobachtet werden. In früheren *in silico* Studien, die dasselbe Thema adressierten, wurde die These aufgestellt, dass der Aminoimidazolteil des Grundgerüsts ebenfalls seine Protonierung im proteingebundenen Zustand ändern würde. Unseren ITC-Messungen zufolge würde dieser Teil nur dann bei der Bindung protoniert, wenn eine

Aminosäureseitenkette gleichzeitig ein Proton abgibt. Der einzige Kandidat, der diese Eigenschaft aufweist, ist ein nahegelegener Glutamatrest, von dem bekannt ist, dass er seinen Protonierungszustand abhängig von dem gebundenen Liganden oder dem vorherrschenden pH-Wert der Umgebung ändern kann. pH-Soaking-Experimente ergaben, dass diese Seitenkette schon vor der Ligandbindung bei dem in den Messungen verwendeten pH-Wert deprotoniert vorliegt und daher die Aminoimidazolgruppe im ungeladenen Zustand verbleiben muss. Der Affinitätsgewinn, der zuvor dem Protonierungsereignis zugeschrieben wurde, konnte zwei parallel ausgebildeten H-Brücken zugeordnet werden, die unmittelbar zu zwei benachbarten C=O Gruppen des Proteinrückgrates gebildet werden. Auf diese Weise werden sekundär repulsive Interaktionen vermieden.

PAMPA Messungen deckten auf, dass das Grundgerüst der *lin*-Benzoguanine einen zu polaren Charakter aufweist, um durch Zellmembranen zu wandern. Unter Berücksichtigung der vorliegenden Studien ist es nicht überraschend, dass die *lin*-Benzohypoxanthine keine Verbesserung des PAMPA-Scores zeigen. Zukünftige Design-Strategien müssen sich mit der Polarität des trizyklischen Grundgerüsts mit besonderem Fokus auf der Aminoimidazolstruktur beschäftigen. Durch einen Ansatz, in dem dieser Molekülteil durch geeignete funktionelle Gruppen maskiert ist und den Liganden so in ein Pro-Pharmakon überführt, könnte die gewünschte Partitionierung erreicht werden.

Unter Berücksichtigung der Bindungsaffinitäten ist eine eher flache Struktur–Wirkungsbeziehung über die Serien der *lin*-Benzopurine zu beobachten. Daher wurden zusätzliche Einblicke in die treibenden Kräfte der Bindung über die Aufspaltung der Freien Bindungsenergie in enthalpische und entropische Beiträge gewonnen (Kapitel 4). Ein modifiziertes Kristallisationsprotokoll ermöglichte es, den flexiblen 2-Substituenten der analysierten 2-Amino-*lin*-Benzoguanine in den meisten Fällen zu erkennen. Dies legte die Grundlage, um strukturelle Eigenschaften mit den erhaltenen thermodynamischen Daten in Verbindung zu setzen. Die Kristallstrukturen deckten auf, dass die *lin*-Benzohypoxanthine an einen Zustand der TGT binden, der dem Apo-Enzym ähnelt und Asp102 in einer der Bindetasche abgewandten Orientierung zeigt. Im Gegensatz dazu induzieren die *lin*-Benzoguanine die Rotation von Asp102 zur Bindetasche. Dadurch wird ein Proton von der Aminopyrimidinonstruktur des Grundgerüsts aufgenommen und die Ausbildung einer

zusätzlichen, zweizähligen Salzbrücke zu Asp102 ermöglicht. Ausgelöst durch die Bewegung von Asp102 ordnen sich verschiedene Aminosäurereste und sekundäre Strukturelemente um, sodass das Protein in eine Konformation überführt wird, die dem tRNA gebundenen Zustand ähnelt. Diese Änderungen resultieren in wesentlichen Unterschieden zwischen der Bestimmung der Bindungsaffinitäten in einem biochemischen Assay, der eine Inhibitionskonstante ( $K_i$ ) liefert, und ITC-Messungen, aus denen man eine Dissoziationskonstante ( $K_d$ ) erhalten kann. Im biochemischen Assay wird der Einbau von radioaktiv markiertem Guanin in die tRNA gemessen. Als Voraussetzung wird die TGT in ihrem substratgebundenen Zustand benötigt. Dadurch erscheinen die *lin*-Benzohypoxanthine weniger potent, da das Protein zunächst in die substratgebundene Form überführt werden muss. Im Gegensatz dazu erfassen ITC-Messungen die Änderungen der Wärme, die während der Bindung abgegeben oder aufgenommen wird und man erhält eine reine Dissoziationskonstante des Bindungsgleichgewichtes zwischen Ligand und benötigtem Proteinkonformer. Dabei bindet die TGT die *lin*-Benzohypoxanthine in ihrem Apo-Zustand, wodurch eine verbesserte Bindungsstärke verzeichnet wird.

Wie erwartet, binden beide Ligandserien, *lin*-Benzohypoxanthine und *lin*-Benzoguanine, enthalpisch. Dabei weisen die *lin*-Benzohypoxanthine einen weniger ausgeprägten enthalpischen Beitrag auf, da die Interaktion zu Asp102 fehlt. Diese kann teilweise durch ein kristallographisch konserviertes Wassernetzwerk am Grunde der Guanine-34-Tasche kompensiert werden. Während die thermodynamischen Profile der *lin*-Benzohypoxanthin-Serie fast unverändert bleiben, sind die Daten für die *lin*-Benzoguanine recht divers. Dadurch war es nur schwer möglich, den Beitrag des angefügten 2-Substituenten zu bestimmen. Offensichtlich ermöglichen die strukturellen Änderungen induziert durch Asp102 im Falle der *lin*-Benzoguanine eine Kommunikation zwischen Ribose/Uracil-33-Tasche, die vom 2-Substituenten adressiert wird, und der Guanin-34-Tasche, die das Ligandgrundgerüst besetzt. Helix  $\alpha A$ , die fast orthogonal zu dem gebundenen Liganden steht, spielt eine entscheidende Rolle dabei, die Interaktionen zwischen den beiden Taschen zu vermitteln. Nachdem die *lin*-Benzoguanine über Asp102 als substratähnlich vom Protein erkannt werden, ändert unter anderem diese Helix ihre Position wie es auch in ähnlicher Art und Weise in tRNA-TGT Komplexen beobachtet wird. Begleitet von dieser Bewegung dringt Tyr106 tiefer in die Ribose/Uracil-33-Tasche vor und nimmt

wahr, ob dieser Teil des Proteins besetzt ist oder nicht. Dieses Vorgehen garantiert richtige von falschen Substraten zu unterscheiden.

Zukünftige Designzyklen sollten berücksichtigen, dass Liganden, die die Ribose/Uracil-33-Tasche adressieren, als substratähnlich erkannt werden müssen und damit die Rotation von Asp102 in die Guanin-34-Tasche bewirken, sodass die beschriebene Kommunikation zwischen den Bindetaschen erfolgen kann. Zusätzlich sind nur die so unterschiedenen Liganden in der Lage die Bindungsaffinität durch zweizählige Salzbrücken zu beiden Aspartaten zu modulieren, vorausgesetzt ihr  $pK_a$ -Wert fällt in einen Bereich, der geeignet ist eine Protonierungsänderung zu bewirken.

Basierend auf der Beobachtung, dass die 2-Amino-*lin*-Benzoguanine eine erniedrigte Besetzung in Verbindung mit erhöhten Temperaturfaktoren des 2-Substituenten zeigen, wurde eine hohe Flexibilität dieses Molekülteils angenommen. Darauf aufbauend wurde getestet, ob ein Ligand mit hoher residualer Beweglichkeit einen Bindungsverlust eher vermeiden kann als ein Ligand, der einen geordneten Bindungsmodus einnimmt, und so einen Vorteil als ein wirksameres Antibiotikum erzielen kann, das weniger von Resistenzentwicklungen betroffen ist (Kapitel 5). Nachdem Val262 inmitten der Ribose-33-Tasche als Mutationsstelle ausgewählt wurde, folgte die Expression und Kristallisation drei verschiedener Mutanten (Val262Thr, Val262Asp, Val262Cys). Anschließend wurde die Bindungsaffinität ( $K_d$ ) verschiedener 2-Amino-*lin*-Benzoguanine bestimmt und in Bezug auf strukturellen Informationen diskutiert, die aus Kristallstrukturen der TGT-Mutanten im Komplex mit Ligand **4a** erhalten wurden.

Verglichen mit dem Wildtyp zeigten die analysierten 2-Amino-*lin*-Benzoguanine die gleiche Bindungsaffinität zu Mutante Val262Thr. Dieses Ergebnis war wie erwartet, da die Eigenschaften der Tasche nur leicht verändert wurden. Im Gegensatz dazu verringerte sich die Potenz gegenüber Mutante Val262Asp signifikant während die  $K_d$ -Werte zu Mutante Val262Cys anstiegen. Die beobachteten Änderungen konnten eher in Zusammenhang mit dem Grundgerüst als mit den verschiedenen 2-Substituenten gebracht werden, das Unordnung seitens des Proteins in naher Umgebung zur Mutationsstelle verursacht. Nur marginale Unterschiede konnten den Eigenschaften des Substitutionsmusters im Falle von Val262Asp zugeschrieben werden, die höchstwahrscheinlich auf elektrostatische Anziehungs- und Abstoßungskräfte zurückzuführen sind. Nichtsdestotrotz deutet dieser Fakt

an, dass flexible Teile eines Arzneistoffmoleküls dazu in der Lage sind, die Bindungsaffinität in der ungünstigen Situation einer Resistenzentwicklung durch Mutation zu modulieren.

In Kapitel 6 wurden Molekulardynamiksimulationen genutzt, um den Bindungsmodus verlängerter 2-Amino-*lin*-Benzoguanine vorauszusagen, der nicht zufriedenstellend durch vorherige Dockingstudien beschrieben werden konnte. Unglücklicherweise konnte nur eine von vier MD-Simulationen den Bindungsmodus des Liganden zuverlässig voraussagen, sodass dieser mit dem der gelösten Kristallstruktur übereinstimmte. Nichtsdestotrotz resultierte die anschließende Kristallstrukturanalyse unter modifizierten Kristallisationsbedingungen in einer angemessenen Differenzelektronendichte für alle verlängerten 2-Amino-*lin*-Benzoguanine. Die verfeinerten Datensätze deckten neuartige Aspekte auf, die entscheidend für das weitere Design und die Charakterisierung von TGT Inhibitoren sind:

Erstens waren wir dazu in der Lage mit einem modifizierten Kristallisationsprotokoll den 2-Substituenten der verlängerten 2-Amino-*lin*-Benzoguanine zu erkennen, der in der Ribose-32-Tasche bindet, die zuvor noch nie besetzt wurde. Die gesammelten Datensätze liefern die Grundlage für neue Synthesestrategien, durch die neuartige Interaktion in diesem Proteinteil geschlossen werden können, der besser abgeschirmt ist als seine benachbarte, solventexponierte Umgebung. In diesem Zusammenhang zeigte ein anderer Ligand, der in der Ribose-33-Tasche eine zweizählige Salzbrücke zum Protein ausbildet, eine erniedrigte Bindungspotenz. Dieses Beispiel verdeutlicht, dass solventexponierte Interaktionen, die zwischen Ligand und Protein innerhalb der flachen Ribose-33-Tasche gebildet werden, wenn überhaupt nur wenig zur Bindungsaffinität beitragen, da hohe Desolvatisierungskosten in der Umgebung hoher Dielektizitätsverhältnisse gezahlt werden müssen. Daher müssen zukünftige Wirkstoffdesignzyklen, die diesen Teil des Proteins adressieren, die Optimierung eher hydrophober 2-Substituenten berücksichtigen.

Zweitens unterstreichen die Ergebnisse dieses Kapitels die Wichtigkeit der verwendeten Kristallisationsbedingungen. Schlecht definierte Elektronendichte deutet nicht notwendigerweise auf Liganden oder Substituenten hin, die eine hohe Beweglichkeit im proteingebundenen Zustand besitzen. Das verwendete Kristallisationsprotokoll beeinflusst die erhaltene Differenzelektronendichte maßgeblich und muss in der Diskussion der erhaltenen Strukturen berücksichtigt werden.

Die Substanzklasse der 5-Azacytosine wurde als neuartiges Grundgerüst untersucht, das die TGT mit potentiell erhöhter Bioverfügbarkeit inhibiert (Kapitel 7). Ähnlich den *lin*-Benzopurinen sind auch die 5-Azacytosine von dem natürlichen Substrat Guanin abgeleitet und vergleichbare Bindungseigenschaften wurden erwartet. Ein angefügter Phenylring dient als Ankerpunkt für Substituenten, die in die benachbarten Ribose-33 und Ribose-34-Taschen reichen. Im Vergleich zu den *lin*-Benzoguaninen bietet diese neue Inhibitorklasse den Vorteil der verbesserten synthetischen Zugänglichkeit. Der  $pK_a$ -Wert ihrer Aminopyrimidinoneinheit (N(5)) liegt jedoch zwischen dem der *lin*-Benzohypoxanthine und *lin*-Benzoguanine und stellt somit einen kritischen Sachverhalt dar, der in Bezug auf das Protonierungsverhalten der 5-Azacytosine genauer betrachtet werden muss.

Unglücklicherweise zeigen die Liganden nur Bindungsaffinitäten im niedrig mikromolaren Bereich. Ausnahme bildet ein Präparat, das mit einer Naphthyleinheit dekoriert ist und eine Bindungspotenz im einstellig mikromolaren Bereich aufweist, ähnlich dem nahe verwandten Guanin. Das Kristallwachstum stellte sich als schwierig heraus, da alle Präparate durch eine geringe Löslichkeit gekennzeichnet waren. Nichtsdestotrotz konnten Kristallstrukturen für drei der fünf Liganden gelöst werden. Wie erwartet bindet das 5-Azacytosin-Grundgerüst in der Guanin-34-Bindungstasche ähnlich wie Guanin, jedoch ist die Interaktion zu Asp102 schwach ausgebildet, die sich als äußerst wichtig für die Substraterkennung, Ligandprotonierung und Kommunikation zwischen den Bindetaschen herausgestellt hat. Der Grund für die Störung dieser Interaktion kann von dem Bindungsmodus der 6-Phenyl-5-Azacytosine abgeleitet werden: Anders als alle bekannten TGT-Inhibitoren binden sie nicht in einer planaren Art und Weise an das Protein, da der angefügte Phenylring aus der Ebene rotiert, die durch den Heterozyklus definiert wird. In Folge wird der im Vergleich zu den *lin*-Benzoguaninen niedrigere  $pK_a$ -Wert nicht durch die Interaktionen zu Asp102 und Asp156 in ein Fenster geeignet für die Ligandprotonierung verschoben und der Einbruch der Affinität kann dadurch erklärt werden. Für das potentere Naphthylderivat konnte keine Kristallstruktur mit dem gebundenen Liganden bestimmt werden. Bis zu diesem Zeitpunkt bleibt unklar, ob der Naphthylsubstituent einen konformationellen Wechsel von Asp102 veranlasst, um die gewünschte zweizählige Salzbrücke zum Grundgerüst auszubilden und damit den Protonierungsschritt induziert, der für den Affinitätsgewinn nötig ist.



Trotz der vielversprechenden Bindungsstärke ist das Naphthylderivat kein geeigneter Startpunkt für weitere Optimierungen, da es dazu neigt bei höheren Konzentrationen Aggregate zu bilden. Dies bereitet erhebliche Probleme im Umgang während der Synthese und nachgeschalteter Analyse. Stattdessen sollten sich zukünftige Designzyklen auf die nahe verwandten Ammeline konzentrieren. Eine zusätzliche Aminogruppe außerhalb des Heterozyklus stellt nicht nur eine Verbindungsregion dar, die einen größeren Abstand zu der Seitenkette von Asp102 schafft, sondern erhöht auch den  $pK_a$ -Wert um eine logarithmische Einheit und fällt damit wahrscheinlich in einen Bereich, der die Protonierung des Grundgerüsts erlaubt.

## 9. Materials and Methods

### 9.1 *In-silico* Methods

#### 9.1.1 pK<sub>a</sub> Calculations

A consistent charge model was produced by a modified version of the charge distribution algorithm that initially was developed by Gasteiger and Marsili, named “partial equalization of orbital electronegativities” (PEOE) [Barken and Gasteiger, 1980]. According to previous studies on the target, a dielectric constant of  $\epsilon = 20$  was chosen to describe the properties of the binding pocket considering an implicit solvent model [Ritschel *et al.*, 2009].

For the calculation of the pK<sub>a</sub> values all titratable groups within a radius of 12Å around the active site were considered for the site-site interaction portion of the pK<sub>a</sub> calculations (C $\gamma$  of Tyr106 was taken as center of the selection). Thereby, the following residues were identified as titratable groups: Lys52, Tyr72, His73, Asp/Asn102, Tyr106, Asp/Asn156, Glu157, Cys158, Tyr161, Glu173, Tyr226, Glu235, Asp238, Glu239, Asp245, Tyr258, Lys264, Asp266, Asp267, Asp280, Cys281, Tyr354, and Tyr381. With the program *REDUCE*, all hydrogens were added to the protein in which all acidic residues were deprotonated and basic ones were protonated [Word *et al.*, 1999]. For the uncharged state of the ligand, *SYBYL* atom types were assigned and an explicit charge of 0 was set for the hydrogen atom of the titratable group prior to the partial charge calculation.

After this preparation, the Poisson–Boltzmann calculation was started using the program *MEAD* [Bashford, 1997]. The resulting pK<sub>a</sub> shifts of the crucial amino acid residues are listed in Tables 3.3 – 3.4 at pH 7.8. A detailed list of all analyzed residues is shown in the appendix.

#### 9.1.2 Molecular Dynamic Simulation

Molecular dynamic (MD) simulations were performed with the program *AMBER11* [Case *et al.*, 2005]. In chapter 6 the docking results (8.1.3) of the corresponding ligands served as starting coordinates, in chapter 7 the complex structure TGT-9. Prior to the simulation, the pdb-file had to be modified: All crystallographically determined water molecules were extracted. Missing amino acids as well as the most probable conformer of missing side chains were added. In case of more than one visible conformation, the highest occupied one was kept during the simulation. Protonation states of histidines were inspected visually and

set as HID (hydrogen at  $\delta$ -position, HIS#: 90, 257, 319, 332) or HIE (hydrogen at  $\epsilon$ -position, HIS#: 73, 145, 333). The zinc ion was mimicked by four massless dummy atoms, each with a charge of +0.5 using the CaDA approach by Pang [Pang *et al.*, 2000; Pang, 2001]. Cysteine residues involved in the binding of the zinc ion were set as CYM (CYS#: 308, 310, 313), histidines as HIN (HIS#:349).

Due to the required computational time the calculations were based on only one TGT monomer. Parameters for the ligands were generated with the program *antechamber* using the general amber force field (gaff) [Wang *et al.*, 2004], its charges were calculated via bond charge correction (bcc) [Jakalian *et al.*, 2000; Jakalian *et al.*, 2002]. Addition of hydrogen atoms to the protein, neutralization of the system by adding two sodium ions and solvation of the complex in a TIP3P water box was done with the *tleap* [Jorgensen *et al.*, 1983].

After a minimization of the water box comprising 100 steps and the whole system comprising 500 steps performed with a generalized Born solvent model, all following simulations included periodic boundary conditions, the Particle Mesh Ewald procedure (PME) [Darden *et al.*, 1993] and the SHAKE algorithm [Ryckaert *et al.*, 1977] using the ff99SB force field [Hornak *et al.*, 2006] with a cut-off of 10 Å. Thereby, the system is heated up to 300 K stepwise (0 K ... 150 K ... 225 K ... 300 K) over a period of 150 ps fixing the TGT monomer with weak restraints ( $25 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$ ). Subsequently, the pressure is adjusted to 1 bar over a time scale of 50 ps followed by a productive simulation for 10 to 100 ns using 2 fs time steps under NPT conditions. The trajectory derived under these conditions was further analyzed with the program *ptraj*, whereby every second frame was included into the analyses. The dialplot was generated using *R*.

### 9.1.3 Docking

Docking was performed using *GOLD* Suite v5.1 [The Cambridge Crystallographic Data Centre: Cambridge, U.K.], the default values for the genetic algorithm (GA) [Jones *et al.*, 1997] and the scoring function *ChemScore* [Eldridge *et al.*, 1997]. For each ligand 30 – 100 GA runs were performed with 100'000 operations. As a protein model the coordinates of the crystal structure 2QZR (chapter 6) and TGT-8 (chapter 7) were used. Since Asp102 is present in two different conformations in TGT-8, its side chain was kept flexible during docking. The binding site included all atoms within a distance of 10 Å to the sulfur of Met260. Ligands were drawn using the *MOE* builder.

## 9.2 Experimental Methods

### 9.2.1 Chemicals

Chemicals used in the context of this work are listed in table 9.1.

**Table 9.1** Used chemicals as well as their manufacturer in alphabetical order.

<b>Name</b>	<b>Manufacturer</b>
Acetic Acid 100%	ROTH
Acrylamide/bisacrylamide (30% / 0.8%)	ROTH
Agar-agar	ROTH
Agarose	Fluka
Ammonium peroxydisulfate (APS)	ROTH
Ammonium sulfate	ROTH
Ampicillin	ROTH
Anhydrotetracycline hydrochloride	Fluka
Antarctic Phosphatase	New England BioLabs
Biotinylated thrombin	Novagen
Bovine serum albumin (BSA)	Sigma Aldrich
Bromphenol blue	Merck
Chloramphenicol	ROTH
Chloroform:Isoamyl alcohol 24:1	ROTH
Complete™ mini protease inhibitor, EDTA free	Roche
Coomassie Brilliant Blue R250	ROTH
d-Desthiobiotin	IBA
Dimethylformamide	Merck
Dimethylsulfoxide (DMSO) p. a.	ROTH
1,4-Dithiothreitol (DTT)	ROTH
EcoRV-HF®	New England BioLabs
Ethanol 96%	ROTH
Ethylenediaminetetraacetate (EDTA)	Merck
Glucose	ROTH
Glycerol	Acros Organics
[8- <sup>3</sup> H]-Guanine	Hartmann Analytics
Guanine hydrochloride	Sigma Aldrich
Hydrochloric acid	Fisher Scientific
2-(4-Hydroxyphenylazo)-benzoic acid (HABA)	Fluka
2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethanesulfonic acid (Hepes)	ROTH
Isopropyl-β-D-thiogalactopyranoside (IPTG)	ROTH
Kanamycin sulfate	AppliChem
Magnesium chloride hexahydrate	Merck
6x Mass Ruler Loading Dye (R0621)	Fermentas

β-Mercaptoethanol	Merck
2-(N-Morpholino)ethanesulfonic acid (MES)	ROTH
Methanol p. a.	Fluka
NheI	New England BioLabs
PageRuler™ Prestaint Protein Ladder (SM0671/2)	Fermentas
Phenol:Chloroform:Isoamyl alcohol 25:24:1	ROTH
Polyethylene glycol MW 400 (PEG 400)	Sigma Aldrich
Polyethylene glycol MW 8000 (PEG 8000)	Fluka
Protein assay dye reagent concentrate	BioRad
Rotiphorese® Gel 30 (Acrylamide:N,N-Methylenbisacrylamide 37.5:1)	ROTH
Rotiphorese® 10x SDS-PAGE	ROTH
Rotiszint®eco plus	ROTH
Sodium chloride	ROTH
Sodium dodecyl sulfate	ROTH
Sodium hydroxide	Grüssing
SYBR® Safe DNA gel stain	invitrogen
T4 DNA Ligase	New England BioLabs
Tetramethylenediamine (TEMED)	ROTH
Trichloroacetic acid (TCA)	Fluka
N-(Tris(hydroxymethyl)methyl)glycine (Tricine)	Merck
2-Amino-2-(hydroxymethyl)-propane-1,3-diol (Tris)	ROTH
Tryptone	ROTH
Tween 20	Sigma Aldrich
Yeast extract	ROTH

### 9.2.2 Equipment

Equipment used in this work is given in table 9.2.

**Table 9.2** Employed equipment as well as its manufacturer in alphabetical order.

Device	Manufacturer
ÄKTA FPLC	GE Healthcare
Autoklave type FVA2	Fedegari
Centrifuge Avanti J-10 / J-25	Beckman Coulter
Centrifuge Biofuge fresco	Heraeus
Centrifuge Biofuge pico	Heraeus
Centrifuge Multifuge 3	Heraeus
Dialysis membrane Zellu Trans 4,000 – 6,000 MWCO	ROTH
Elektrophoresis cell OWL EASYCAST™ B1A	Thermo Scientific
FiveEasy pH meter	Mettler Toledo

HERAsafe biological safety cabinet	Thermo Scientific
Incubator shaker Innova 4200	New Brunswick Scientific
Incubator shaker Innova 4230	New Brunswick Scientific
Incubator type B5060 EC-CO2	Heraeus
iTC200 microcalorimeter	GE Healthcare
Magnetic stirrer IKA-COMBIMAG REO	Janke & Kunkel
Microscope Typ SZ 60	Olympus
Micro weighing scale type CP2P	Sartorius
Mini Cyclor™ PTC-150	MJ Research
Nanodrop 2000c	Thermo Scientific
NAP-25 Sephadex column	GE Healthcare
Phenyl Sepharose column XK16	Amersham Bioscience
Pipettes	Eppendorf
Precision weighing scale type 404/13	Sauter
Q Sepharose column XK26	Amersham Bioscience
Safe Imager™ Blue Light-Transilluminator	invitrogen
Shaker Polymax 1040	Heidolph
SDS-PAGE Mini-PROTEAN® Tetra Electrophoresis System	BioRad
SpectroSize™ 300	Molecular Dimensions
Speed-Vac DNA mini	Jouan Nordic
Strep-Tactin Superflow Sepharose column	IBA
Scintillation Counter TRI-CARB® 1900CA	Packard
Thermomixer Comfort 2mL	Eppendorf
Ultrasonic Sonifier 250	Branson
Ultrasonic bath Elmasonic S 10/(H)	Elma
UV-Vis Spectrometer Smart Spec 3000	BioRad
Vakuum pump type N811KN.18	KNF
Vivaspin® 20 centrifugal concentrator 30,000 MWCO	Sartorius
Vortex mixer VF1	Janke & Kunkel
Weighing machine type 572/45	Kern
Zetasizer Nano ZS	Malvern Instruments

### 9.2.3 Buffer, Solutions and Media

Table 9.3 comprises a list of the buffers, solutions and media that were used in the context of this work. If not explicitly described, chemicals were dissolved in demineralized water. The desired pH value was achieved by the addition of sodium chloride solution (10 M) or hydrochloric acid (32%) using a FiveEasy pH meter. Buffers were freshly prepared and sterile filtered through a Millipore Steritop™ bottle top filter with a pore size of 0.22 µm before usage.

**Table 9.3** Used buffer, solutions and media in alphabetical order.

<b>Name</b>	<b>Composition</b>
Ampicillin solution	100 mg/ml ampicillin, sterile filtered
Anhydrotetracycline solution	2 mg/ml anhydrotetracycline in DMF, sterile filtered
APS solution	10% (w/v) ammonium peroxydisulfate
Antarctic Phosphatase Reaction	
Buffer (1x)	50 mM Bis-Tris-propane hydrochloride, 1 mM MgCl <sub>2</sub> , 0.1 mM ZnCl <sub>2</sub>
Assay buffer	200 mM HEPES, 20 mM MgCl <sub>2</sub> , 0.037% (v/v) Tween 20, pH 7.3
Buffer A	10 mM Tris, 1 mM EDTA, 1 mM DTT, pH 7.8
Buffer B	10 mM Tris, 1 mM EDTA, 1 mM DTT, 1 M NaCl, pH 7.8
Buffer C	10 mM Tris, 1 mM EDTA, 1 mM DTT, 1 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> , pH 7.8
Buffer E	100mM Tris, 150mM NaCl, 1mM EDTA, 2.5mM d-desthiobiotin, pH 7.8
Buffer R	100mM Tris, 150mM NaCl, 1mM EDTA, 1 mM HABA, pH 7.8
Buffer W	100mM Tris, 150mM NaCl, 1mM EDTA, pH 7.8
Chloramphenicol solution	34 mg/ml chloramphenicol in ethanol, sterile filtered
Cryo buffer A	50 mM MES, 0.3 M NaCl, 2 % (v/v) DMSO, 4 % (w/v) PEG8000, 30% (v/v) glycerole, pH 5.5
Cryo buffer B	50 mM Tris, 0.3 M NaCl, 2% (v/v) DMSO, 4% (w/v) PEG8000, 30% (v/v) glycerole, pH 8.5
Cryo buffer C	80 mM MES, 8% (v/v) DMSO, 10.4% (w/v) PEG8000, 20% (v/v) PEG 400, pH 5.5
Cryo buffer D	80 mM Tris, 8% (v/v) DMSO, 6.4% (w/v) PEG8000, 20% (v/v) PEG400, pH 7.8
Crystallisation buffer A	100 mM MES, 10% (v/v) DMSO, 11 – 13% (w/v) PEG8000, pH 5.5
Crystallisation buffer B	100 mM Tris, 10% (v/v) DMSO, 6 – 8% (w/v) PEG8000, pH 8.5
CutSmart™ Buffer	50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 µg/ml BSA
Destaining solution	10% (v/v) acetic acid, 40% (v/v) methanol
Dialysis buffer	10 mM Tris, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.8
Guanine solution	188 µM guanine hydrochloride
IPTG stock solution	1 M IPTG, steril filtered
ITC buffer A	50 mM HEPES, 200 mM NaCl, 0.037% (v/v) Tween 20, pH 7.8
ITC buffer B	50 mM Tris, 200 mM NaCl, 0.037% (v/v) Tween 20, pH 7.8
ITC buffer C	50 mM Tricine, 200 mM NaCl, 0.037% (v/v) Tween 20, pH 7.8
Kanamycin solution	30 mg/ml kanamycin sulfate
Laemmli buffer	0.25 M Tris, 2 M glycerole, 1% (w/v) SDS, pH 8.3
Luria-Bertani agar	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, 1.5% (w/v) agar-agar, autoclaved
Luria-Bertani medium	1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, autoclaved
Makro seeding buffer	100 mM Tris, 10% (v/v) DMSO, 5% (w/v) PEG 8000, pH 7.8
High salt buffer	10 mM Tris, 1 mM EDTA, 2 M NaCl, pH 7.8
Lysis buffer	20 mM Tris, 10 mM EDTA, 1 mM DTT, 1 tablet/50 ml Complete™, pH 7.8
Running gel buffer	1 M Tris, pH 8.8

SDS solution	10% (w/v) SDS
SDS sample buffer (4x)	250 mM Tris, 8% (w/v) SDS, 40% (v/v) glycerole, 0.04% (w/v) bromphenol blue, 8% (w/v) $\beta$ -mercaptoethanol, pH 6.8
SOC medium	20% (w/v) tryptone, 5% (w/v) yeast extract, 0.5% (w/v) NaCl, 10 mM $MgCl_2$ , 0.4% (w/v) glucose
Stacking gel buffer	1 M Tris, pH 6.8
Staining solution	10% (v/v) acetic acid, 40% (v/v) methanol, 1 g/l Coomassie Brilliant Blue R250
T4 DNA Ligase Reaction Buffer	50 mM Tris-HCl, 10 mM $MgCl_2$ , 1 mM ATP, 10 mM DTT
TAE buffer (50x)	2 M Tris, 50 mM EDTA, pH 7.8
TCA solution A	5% (w/v) trichloroacetic acid
TCA solution B	10% (w/v) trichloroacetic acid
Tween solution	1% (w/v) Tween 20
2xYT medium	16% (w/v) tryptone, 10% (w/v) yeast extract, 5% (w/v) NaCl, autoclaved

#### 9.2.4 Bacterial Strains and Plasmids

Plasmids and bacterial strains are given in Tables 9.4 and 9.5.

**Table 9.4** Used plasmid vectors as well as their characteristics and manufacturers in alphabetical order.

Name	Characteristics	Manufacturer
pASK-IBA13plus-ZM10	$Amp^r$ , $Cm^r$ ; Strep-tag II <sup>®</sup> ; ColE1-origin, <i>tet</i> -promotor; inserted <i>tgt</i> -gene as 1.2 kb <i>BsaI</i> fragment	[Jakobi, 2013]
pET9d-ZM4	$Cm^r$ , $Km^r$ ; ColE1-origin, <i>tac</i> -promotor, coding for <i>laqI<sup>q</sup>malE lacZ<math>\alpha</math></i> ; inserted <i>tgt</i> -gene as 1.3 kb <i>BamHI/NcoI</i> fragment	[Reuter and Ficner, 1995]
pET9d-ZM4-D102N	$Cm^r$ , $Km^r$ ; TGT-D102N	[Romier <i>et al.</i> , 1997]
pET9d-ZM4-E235Q	$Cm^r$ , $Km^r$ ; TGT-E235Q	[Tidten <i>et al.</i> , 2007]
pPR-IBA2-ZM10	$Amp^r$ , $Cm^r$ ; Strep-tag II <sup>®</sup> ; ColE1-origin, T7-promotor based expression; inserted <i>tgt</i> -gene as 1.2 kb <i>NheI/EcoRV</i> fragment	This work
pPR-IBA2-ZM10-D156N	$Amp^r$ , $Cm^r$ ; TGT-D156N	This work
pPR-IBA2-ZM10-V262C	$Amp^r$ , $Cm^r$ ; TGT-V262C	This work
pPR-IBA2-ZM10-V262D	$Amp^r$ , $Cm^r$ ; TGT-V262D	This work
pPR-IBA2-ZM10-V262T	$Amp^r$ , $Cm^r$ ; TGT-V262T	This work
ptRNA2	$Amp^r$ ; <i>E. coli tRNA<sup>Tyr</sup></i> as <i>BstNI</i> fragment under control of T7-promotor in pTZ18U	[Curnow <i>et al.</i> , 1993]



**Table 9.5** Used *E. coli* host cells including their genotype and vendor in alphabetical order.

Host cell	Genotype	Vendor
BL21(DE3)pLysS	F <sup>-</sup> <i>dcm ompT hsdS</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>gal λ</i> (DE3) [pLysS Cam <sup>r</sup> ]	Stratagene
BL21-CodonPlus(DE3)-RIPL	F <sup>-</sup> <i>ompT hsdS</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>dcm</i> <sup>+</sup> Tet <sup>r</sup> <i>gal λ</i> (DE3) <i>endA Hte</i> [argU proL Cam <sup>r</sup> ] [argU ileY leuW Strep/Spec <sup>r</sup> ]	Stratagene
Rosetta 2(DE3)	F <sup>-</sup> <i>ompT hsdSB</i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ) <i>gal dcm</i> (DE3) pRARE2 <sup>3</sup> (Cam <sup>r</sup> )	Stratagene
XL10-Gold	Tet <sup>r</sup> Δ( <i>mcrA</i> )183 Δ( <i>mcrCB-hsdSMR-mrr</i> )173 <i>endA1 supE44 thi-1 recA1 gyrA96 relA1 lac Hte</i> [F <sup>-</sup> <i>proAB lacI</i> <sup>q</sup> ZΔM15 Tn10 (Tet <sup>r</sup> ) Amy Cam <sup>r</sup> ]	Stratagene

## 9.2.5 Molecular biological methods

### 9.2.5.1 Cloning

In order to achieve a higher protein yield the sequence encoding the strep-tagged wild type TGT of the pASK-IBA13plus-ZM10 plasmid [Jakobi *et al.*, 2014] was cloned into the T7-promotor based expression vector pPR-IBA2 (Table 9.6). Therefore, the *tgt*-gene as well as the new vector were separately cut with the restriction enzymes *NheI*-HF<sup>®</sup> and *EcoRV*-HF<sup>®</sup>. For dephosphorylation the cut vectors were treated with Antarctic Phosphatase.

**Table 9.6** Composition of the digestion, dephosphorylation and ligation mixtures.

Digestion	Dephosphorylation	Ligation
pPR-IBA2 vector /		pPR-IBA2 vector 0.5 μl
pASK-IBA13plus-ZM10 1 μg		<i>tgt</i> -gene 9.5 μl
<i>NheI</i> -HF <sup>®</sup> 1 μl	Antarctic Phosphatase	1 μl T4 DNA Ligase 1 μl
<i>EcoRV</i> -HF <sup>®</sup> 1 μl		
CutSmart™ Buffer (10x) 5 μl	Antarctic Phosphatase	T4 DNA Ligase
ddH <sub>2</sub> O 42 μl	Reaction Buffer (1x)	6 μl Reaction Buffer (10x) 2 μl
	ddH <sub>2</sub> O	3 μl ddH <sub>2</sub> O 7 μl

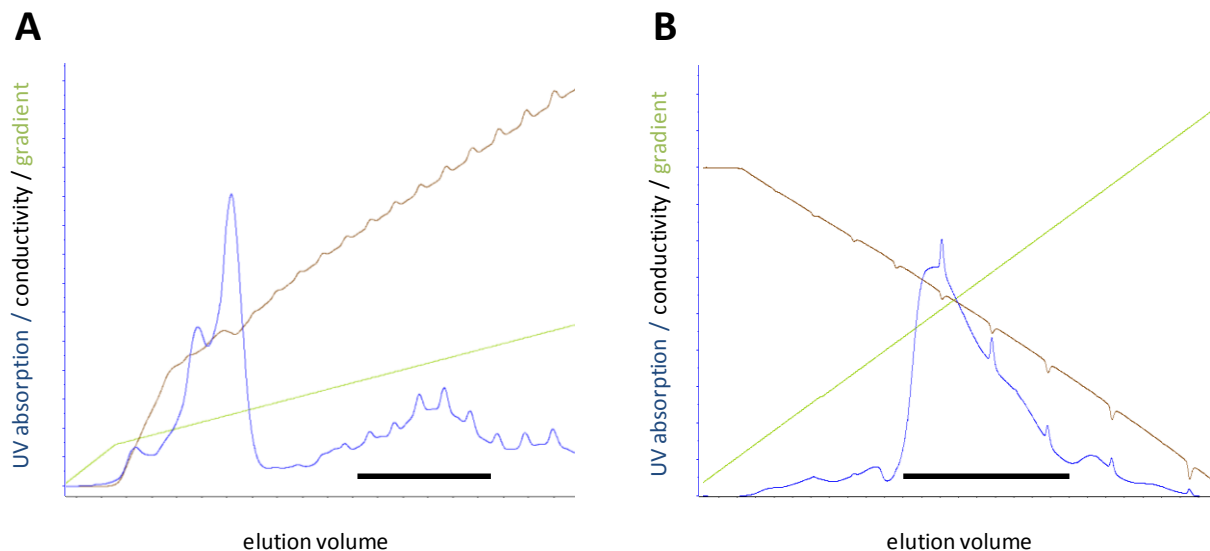
The *tgt*-gene (373 kDa) was separated from the cut pASK-IBA13plus (840 kDa) by agarose gel electrophoresis and the corresponding band cut out of the gel. For the extraction of the DNA out of the gel the QIAquick Gel Extraction Kit (Qiagen) was used. Subsequently, the *tgt*-gene and the pPR-IBA2 vector were ligated for one hour at 37°C. For transformation 10 μl of the sample were used. The resulting plasmid was named pPR-IBA2-ZM10.

### 9.2.5.2 Protein Expression and Purification

#### 9.2.5.2.1 Expression of the *Z. mobilis tgt* gene via the vector pET9d

Bacterial cells BL21(DE3)pLysS transformed with pETZM4 [Reuter and Ficner, 1995] were grown in a pre-culture of 100 ml LB medium containing 30 mg/l kanamycin and 34 mg/ml chloramphenicol at 37°C and 220 rpm shaking overnight and transferred into two liters LB medium. The main culture was raised to an optical density at 600 nm ( $OD_{600}$ ) of 0.8 under the same conditions as listed for the pre-culture. Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was added in a concentration of 1 mM to induce overproduction for 20 hours at 14°C. Cells were harvested by centrifugation (10000 rpm at 4°C) and resuspended in 50 ml lysis buffer. The suspension was sonicated six times for 90 seconds on ice and subsequently centrifuged for 1 hour at 20000 rpm (Beckman Coulter, J-25.50).

The purification via FPLC comprised two steps: Firstly, the clear supernatant was applied onto a Q Sepharose XK26 (anion exchange column) previously conditioned on buffer A. After an initial washing step with buffer A the protein was eluted by a linear gradient reaching from 0 to 100% of buffer B at a flow rate of 4 ml/min (Figure 9.1A). Fractions containing TGT were identified by SDS gel electrophoresis [Laemmli, 1970].



**Figure 9.1** Chromatograms recorded during the purification of TGT wild type using the anion exchange column Q Sepharose XK26 (A) and the hydrophobic interaction column Phenyl Sepharose XK16 (B). The UV absorption (blue line; proportional to the protein amount), the conductivity (black line; indicating the NaCl concentration) and the adjusted gradient (green line) are shown in dependence of the elution volume. The fractions containing the desired protein as confirmed by SDS PAGE are marked by a black bar.

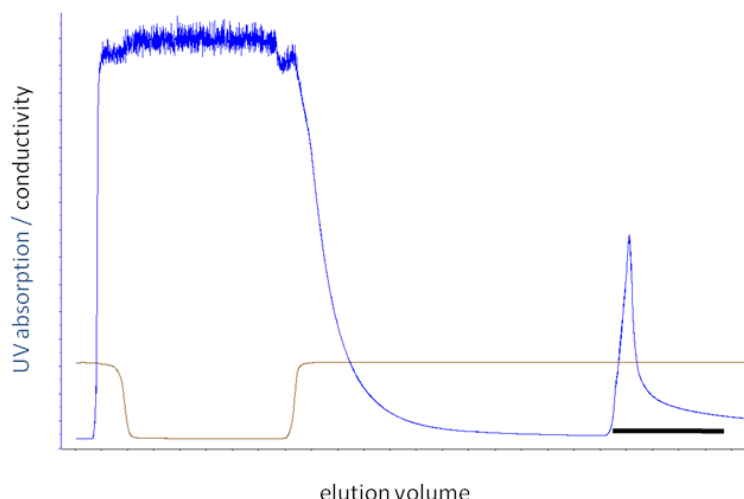
Secondly, 1 M ammonium sulphate was added to the chosen TGT fractions, which were then loaded onto a Phenyl Sepharose XK16 (hydrophobic interaction column) previously conditioned on buffer C. The column was washed with buffer C and the protein was eluted by a linear gradient reaching from 0 to 100% of buffer A at a flow rate of 2 ml/min (Figure 9.1B). Again, fractions containing the desired protein were identified by SDS gel electrophoresis and subsequently concentrated to a final concentration of at least 3 mg/ml using VIVASPIN®20 centrifugal concentrators (Sartorius) with a cutoff of 30,000 Da.

The final purification step was reached by micro-crystallization. Therefore, the protein solution was dialyzed against 5 l of buffer A. The precipitated protein was harvested by centrifugation and resolved with high salt buffer to the required concentration. The derived protein solution was stored at -20°C.

#### 9.2.5.2.2 Expression of the *Z. mobilis tgt* gene via the vector pPR-IBA2

Bacterial cells BL21-CodonPlus(DE3)-RIPL (V262T, V262D, V262C) and Rosetta 2(DE3) (D156N) transformed with pPR-IBA2-ZM10 were grown in a pre-culture of 100 ml 2×YT medium containing 100 mg/l ampicillin and 34 mg/ml chloramphenicol at 37°C and 220 rpm shaking overnight and transferred into two liters 2×YT medium. The main culture was raised to an optical density at 600 nm (OD<sub>600</sub>) of 0.8 under the same conditions as listed for the pre-culture. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at a concentration of 1 mM to induce overproduction for 16 hours at 14°C. Cells were harvested by centrifugation (10000 rpm at 4°C) and resuspended in 50 ml lysis buffer. The suspension was sonicated six times for 90 seconds on ice and subsequently centrifuged for 1 hour at 20000 rpm (Beckman Coulter, J-25.50).

In the course of this work it was observed that also nucleic acids bind to the Strep-Tactin Superflow column the clear supernatant was firstly applied onto a Q Sepharose™ XK26 column previously conditioned on buffer A. After an initial washing step with buffer A the protein was eluted by a linear gradient reaching from 0 to 100% of buffer B at a flow rate of 4 ml/min (Figure 9.1A). Fractions containing TGT were identified by SDS gel electrophoresis [Laemmli, 1970] and loaded onto a Strep-Tactin Superflow column. After a stable UV signal was reached by a washing step with buffer W the protein was eluted with buffer E at a flow rate of 2 ml/min (Figure 9.2).



**Figure 9.2** Chromatogram recorded during the purification of strep-tagged wild type TGT using the Strep-Tactin Superflow column. The UV absorption (blue line; proportional to the protein amount) is plotted in dependence of the elution volume. The fractions containing the desired protein as confirmed by SDS PAGE are marked by a black bar.

Subsequently, the sample was dialyzed (cutoff 4,000 – 6,000) to higher salt conditions against dialysis-buffer for at least 4 hours.

The strep-tag® II was removed via the Thrombin Cleavage Capture Kit (Novagen®) for 16 hours at 20°C following the manufacturer's instructions. Afterwards, the cleaved tag as well as the biotinylated Thrombin was captured using streptavidin agarose via filtration.

For the stability of the protein it was of utmost importance to cleave the tag before concentrating the sample via VIVASPIN®20 centrifugal concentrators (Sartorius; 30,000 MWCO) and exchanging the buffer to high salt conditions. The protein solution was stored at -20°C.

### 9.2.5.3 Mutagenesis

In order to generate mutated plasmids the QuikChange® Lightning Site-Directed Mutagenesis Kit (Stratagene) based on the principle of the polymerase chain reaction (PCR) was used. Oligonucleotide primers were designed following the vendor's instructions and synthesized, purified and analysed by MWG Operon (Ebersberg). For the PCR reaction the protocol suggested by the manufacturer (Table 9.7) was applied using pPR-IBA2-ZM10 as a template. The oligonucleotide primers are listed in Table 9.8.

**Table 9.7** Composition and applied protocol of the PCR reaction.

Composition of the reaction:		Applied protocol:		
dsDNA template	10 – 100 ng	<b>Starting phase</b>	120 s	95°C
Primer forward	125 ng			
Primer backward	125 ng	<b>Denaturation</b>	20 s	95°C
10x buffer	5 µl	<b>Hybridisation</b>	10 s	60°C
dNTP mix	1 µl	<b>Synthesis</b>	150 s	68°C
ddH <sub>2</sub> O	ad 50 µl			RPT 18x
DNA polymerase	1 µl	<b>End phase</b>	300 s	68°C

After the reaction parental methylated and hemimethylated DNA was digested by *Dpn* I. Subsequently, 2 µl of the PCR product was transformed into XL10-GOLD cells during an incubation period of 30 minutes followed by a heat shock of 30 seconds at 42°C. To select only mutated plasmids the host cells were raised in SOC medium for one hour at 37°C, plated on LB agar containing 100 mg/l ampicillin and incubated over night at 37°C. Finally, a single clone was picked and transferred into 10 ml of LB medium containing 100 mg/l ampicillin. The resulting culture was shaken overnight at 220 rpm and 37°C.

700 µl of the bacterial cell cultures were mixed with 300 µl glycerol and stored at -80°C. The remaining culture was harvested for plasmid extraction using the peqGOLD Plasmid Miniprep KitII (PEQLAB Biotechnologie GmbH) following the manufacturer's instructions. For the extracted and purified plasmid the presence of the desired mutation as well as the absence of any further mutations were confirmed by sequence analysis (Eurofins MWG Operon, Ebersberg). Subsequently, the plasmid was transformed into BL21-CodonPlus(DE3)-RIPL or Rosetta 2(DE3) cells by heat shock for 20 seconds at 42°C followed by the selection routine as previously described.

**Tabelle 9.8** Used oligonucleotide primers in site-directed mutagenesis.

Oligonucleotide primer	Sequence 5'- ... -3' <sup>[a]</sup>
D156N_forward	5'-ATT GTT ATG GCA TTT <u>AAT</u> GAA TGT ACC CCG TAT-3'
D156N_backward	5'-ATA CGG GGT ACA TTC <u>ATT</u> AAA TGC CAT AAC AAT-3'
V262C_forward	5'-CAT TAT CTG ATG GGT <u>TGT</u> GGT AAA CCG GAT GAT-3'
V262C_backward	5'-ATC ATC CGG TTT ACC <u>ACA</u> ACC CAT CAG ATA ATG-3'
V262D_forward	5'-CAT TAT CTG ATG GGT <u>GAT</u> GGT AAA CCG GAT GAT-3'
V262D_backward	5'-ATC ATC CGG TTT ACC <u>ATC</u> ACC CAT CAG ATA ATG-3'
V262T_forward	5'-CAT TAT CTG ATG GGT <u>ACT</u> GGT AAA CCG GAT GAT-3'
V262T_backward	5'-ATC ATC CGG TTT ACC <u>AGT</u> ACC CAT CAG ATA ATG-3'

<sup>a</sup>The mutated codon is underlined.

#### 9.2.5.4 *In-vitro* Transcription and Purification of *E. coli* tRNA<sup>Tyr</sup>

The preparation of *E. coli* tRNA<sup>Tyr</sup> was carried out using the T7 RiboMAX™ Express Large Scale RNA Production System (Promega). Thereby, the synthesis reaction was performed using the ptRNA2 plasmid (Table 9.4), which was subsequently digested by RQ1 DNase.

The purification comprised two extraction steps using a 2.1 ml mixture of phenol:chloroform:isoamyl alcohol 125:24:1 (v/v/v) followed by another extraction using a 1 ml mixture of chloroform:isoamyl alcohol 24:1 (v/v). Subsequently, the sample was applied to a NAP-25 column and the eluate of approximately 1.5 ml evaporated overnight at 1 mbar and 34°C. The tRNA was resolved in HEPES buffer and monomerized by applying a temperature gradient reaching from 20°C to 70°C in 10°C steps each for 10 minutes. A final monomerization step at 70°C was performed for two hours after the addition of MgCl<sub>2</sub> to a final concentration of 1 mM. Again, the sample was purified via a NAP-25 column and afterwards evaporated overnight under the previously mentioned conditions. The tRNA was resolved in assay buffer and the concentration adjusted to 200 μM using UV-Vis spectroscopy. 1 μM tRNA<sup>Tyr</sup> corresponds to an absorption of 0.703  $A_{260nm}$ .

#### 9.2.5.5 Agarose Gel Electrophoresis

For gel electrophoresis a 1% (w/v) agarose gel was used. Thus, 0.5 g agarose were dissolved in 49.5 ml 1×TAE buffer by heating the suspension up in a microwave. 5 μl SYBR® Safe DNA gel stain solution were added directly before pouring the gel. 30 μl of the DNA sample were mixed with 5 μl 6×Mass Ruler Loading Dye and loaded on the gel. Fastruler™ DNA Ladder Middle Range and Fastruler™ DNA Ladder Low Range served as references. The gel was run for one hour at 100 V in 1×TAE buffer. Subsequently, bands became visible under UV light.

#### 9.2.5.6 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

The identification and purity of TGT was analyzed via SDS-PAGE. For that purpose, gels were freshly prepared before usage consisting of an upper stacking and a lower running gel (Table 9.9).

**Table 9.9** Stacking and running gel composition.

	Stacking gel (2x)	Running gel (2x)
Acrylamide/bisacrylamide (30% / 0.8%)	1 ml	3 ml
APS solution	60 $\mu$ l	60 $\mu$ l
TEMED	6 $\mu$ l	6 $\mu$ l
Stacking gel/running gel buffer	1.2 ml	2 ml
SDS solution	60 $\mu$ l	60 $\mu$ l
ddH <sub>2</sub> O	3.75 ml	0.95 ml

30  $\mu$ l of the sample were mixed with 10  $\mu$ l 4 $\times$ SDS sample buffer and incubated at 95°C for 15 minutes. After centrifugation of the sample at 13000 rpm for one minute 15  $\mu$ l were loaded to an SDS gel. The PageRuler™ Prestain Protein Ladder served as a reference.

SDS-PAGE was performed in Laemmli buffer at 130 V. Afterwards protein bands were uncovered by incubation the gel in staining solution overnight and exposing it to destaining solution for one hour, whereby each step was accompanied by gentle shaking.

### 9.2.5.7 Determination of Protein and DNA Concentrations

Concentrations of protein, DNA and tRNA solutions were determined by UV-Vis spectroscopy at 280 nm and 260 nm. Previous to the measurement a blank value was generated using the buffer free of any sample to be analyzed.

After calculating the extinction coefficient with ProtParam the concentrations of protein solutions were calculated using Equation 9.1. For *Z. mobilis* TGT expressed without the strep-tag II an absorption of 0.778  $A_{280nm}$  correlates to a concentration of 1 mg/ml (23.4  $\mu$ M). For the tagged variant an absorption of 0.875  $A_{280nm}$  corresponds to 1 mg/ml before and an absorption of 0.789  $A_{280nm}$  after cleavage.

$$C_{Protein} = \frac{A_{280nm}}{\epsilon_{\lambda} \cdot d}$$

(Equation 9.1)

The purity of protein samples could be concluded from the  $A_{260nm}/A_{280nm}$  ratios listed in Table 9.10 [Sambrook and Russell, 2001].

**Table 9.10** Purity of protein samples derived by their  $A_{260nm}/A_{280nm}$  ratio.

Protein [%]	Nucleic acid [%]	$A_{260nm}/A_{280nm}$ ratio
100	0	0.57
95	5	1.06
90	10	1.32
70	30	1.73

Additionally, a Bradford assay was applied on the basis of protein-dye binding using bovine serum albumin in a concentration range of 0.2 to 0.9 mg/ml dissolved in the corresponding buffer as a standard. Both methods resulted in a closely similar protein concentration.

The concentration of extracted and purified DNA was calculated according to Equation 9.2.

$$c_{DNA} = A_{260nm} \cdot 50 \mu\text{g/ml} \cdot \text{dilution factor}$$

(Equation 9.2)

## 9.2.6 Enzyme Kinetics

### 9.2.6.1 Trapping Experiment

The base exchange mechanism allows two putative inhibition modes. Firstly, competitive binding is possible, which either allows tRNA or the small molecule ligand to bind to the active site. Secondly, a mixed binding mode can be observed in the case of small guanine-like inhibitors, allowing the formation of a tertiary complex after the tRNA is bound to the protein and guanine present in position 34 is cleaved and expelled from the binding pocket. To analyze the inhibition mode of the 5-azacytosine type inhibitors a trapping experiment was conducted as described by Xie *et al.*, 2003.

5  $\mu\text{M}$  *Z. mobilis* TGT, 100  $\mu\text{M}$  *E. coli* tRNA<sup>Tyr</sup> (ECY2) and 200 – 500  $\mu\text{M}$  ligand depending on its solubility in 10  $\mu\text{L}$  of 200 mM Hepes buffer pH 7.3 containing 20 mM MgCl<sub>2</sub> were incubated for 1 h at room temperature. After that, 10  $\mu\text{L}$  SDS buffer was added to the mixture and incubated for another hour to guarantee sufficient unfolding. 10  $\mu\text{L}$  of each mixture were loaded onto a 15% SDS gel. The same conditions as described under 8.2.5.6 were applied. After electrophoresis, protein bands were stained with Coomassie™ brilliant



blue R-250 (Bio-RAD). In case of a competitive inhibition mode only one band is visible at 43 kDa. An uncompetitive binding mode is indicated by two bands. The additional band corresponds to the covalent tRNA-TGT complex. As a reference for uncompetitive binding 2,6-diamino-3H-quinazolin-4-one (DAQ) was used [Brenk *et al.*, 2004].

### 9.2.6.2 Enzyme Assay

The method used for the kinetic characterization of TGT is based on the work of Grädler *et al.*, 2001 and Stengl *et al.*, 2005. Due to their low solubility the ligands were first dissolved in 100% DMSO and subsequently diluted to the desired concentration containing 5% DMSO with assay buffer. The protein dissolved in the same buffer was added to the various ligand samples with a final concentration of 150 nM and incubated for 10 min at 37°C. Additionally, a reference without the ligand only containing DMSO and buffer was incubated under the same conditions. Subsequently, base exchange was started by the 1:1 addition of a similarly prepared solution containing *E. coli* tRNA<sup>Tyr</sup> (ECY2; 3 μM) and a mixture of guanine and radioactively labelled [8-<sup>3</sup>H]guanine (20 μM). 15 μL aliquots were removed from this mixture (76 μL) every two minutes and pipetted on Whatman GC-F glass microfiber filters. The reaction was immediately quenched in 10% (w/v) trichloroacetic acid solution at 0°C. In order to separate the tRNA from excess [8-<sup>3</sup>H]guanine not incorporated into tRNA, the glass microfiber filters were washed twice in a 5% (w/v) TCA solution for 10 min followed by an additional washing step using technical grade ethanol over 20 min. The labelled tRNA was captured in the filters during the described steps. Afterwards the filters were dried at 60°C for at least 30 min. The resulting count rate was obtained after the addition of 4 mL Rotiszent™ to each filter by liquid scintillation counting.  $K_i$  values were determined using the method described by Dixon at least in duplicate [Dixon, 1953]. Thereby, the slope of the plotted count rates (GraFit 4.09™, Erithacus Software) yields the initial velocities at given inhibitor concentrations:

$$v_i = \frac{v_{max} \cdot [S]}{K_m \cdot \left(1 + \frac{[I]}{K_i}\right) + [S]}$$

(Equation 9.3)

The reproducibility of the maximal velocity  $v_0$  defined by equation 9.3 [Michaelis and Menten, 1913] is low and therefore, has to be determined again for every series of measurements. The combination of equation 9.3 and 9.4 yields the inhibition constant  $K_i$  corresponding to equation 9.5. Thereby, a Michaelis-Menten constant of  $0.9 \mu\text{mol} \cdot \text{L}^{-1}$  was used for data evaluation [Biela *et al.*, 2013].

$$v_0 = \frac{v_{max} \cdot [S]}{K_m + [S]} \quad (\text{Equation 9.4})$$

$$v_i = \frac{v_{max} \cdot [S]}{K_M \cdot \left(1 + \frac{[I]}{K_i}\right) + [S]} \quad \cup \quad v_0 = \frac{v_{max} \cdot [S]}{K_M + [S]} \Rightarrow \frac{v_0 \cdot K_M + [S]}{v_i \cdot M} = \frac{1}{K_i} \cdot [I] + \left(\frac{[S]}{K_M} + 1\right)$$

(Equation 9.5)

$v_0$	Initial velocity in absence of an inhibitor
$v_i$	Initial velocity in presence of an inhibitor
$K_M$	Michaelis-Menten constant of tRNA <sup>Tyr</sup>
$[S]$	Concentration of tRNA <sup>Tyr</sup>
$[I]$	Inhibitor concentration
$K_i$	Competitive inhibition constant

### 9.2.6.3 Kinetic characterization

Deviating from the described enzyme assay various *E.coli* tRNA<sup>Tyr</sup> concentrations (0.5  $\mu\text{M}$ , 1.0  $\mu\text{M}$ , 1.5  $\mu\text{M}$ , 3.0  $\mu\text{M}$ , 6.0  $\mu\text{M}$  15.0  $\mu\text{M}$ ) were used for the kinetic characterization of the TGT mutants while keeping the guanine (15% radioactively labelled) and protein concentration (75 nM) fixed. This time the guanine / radioactively labelled guanine substrate (20  $\mu\text{M}$ ) was pre-incubated together with tRNA (150 nM) in assay buffer for 10 minutes at 37°C. Subsequently, the protein dissolved in assay buffer was added in a 1:2 manner. 15  $\mu\text{l}$  aliquots were removed from this mixture (76  $\mu\text{l}$ ) every four minutes over a total period of 16 minutes and pipetted on Whatman GC-F glass microfiber filters. The reaction was immediately quenched in 10% (w/v) TCA solution at 0°C followed by the previously described washing steps. Afterwards the filters were dried at 60°C for at least 30 minutes. The resulting count rate was obtained after the addition of 4 ml Rotiszent™ to each filter by liquid scintillation counting. After converting the derived cpm signal into the amount of

incorporated radioactive guanine via a calibration line the initial velocities  $v_0$  for each tRNA concentration were plotted against the reaction time (GraFit 4.09™, Erithacus Software). Based on this plot a Michaelis-Menten curve was generated yielding the Michaelis-Menten constant  $K_M$  as well as  $k_{cat}$  according to the Michaelis-Menten Equation 9.6. Measurements were at least carried out in duplicate.

$$v_0 = \frac{\frac{k_{cat}}{[E]} \cdot [S]}{K_M + [S]} \quad (\text{Equation 9.6})$$

$v_0$	Initial velocity
$k_{cat}$	Catalytic constant of the base exchange
$K_M$	Michaelis-Menten constant of tRNA <sup>Tyr</sup>
[S]	Concentration of tRNA <sup>Tyr</sup>
[E]	Concentration of <i>Z. mobilis</i> TGT

### 9.2.7 Isothermal Titration Calorimetry

ITC measurements were performed using a Microcal iTC<sub>200</sub> microcalorimeter system (GE Healthcare). The protein was dissolved in the experimental buffer to a final concentration of 10  $\mu\text{M}$  for the titrations of *lin*-benzoguanines, 20  $\mu\text{M}$  for the titrations of *lin*-benzohypoxanthines, and 30  $\mu\text{M}$  for the titrations of the 5-azacytosine derivative **11**, each containing 3% DMSO. Due to their low solubility the ligands were first dissolved in 100% DMSO and diluted with the buffer solution to a final DMSO concentration of 3%. The ligand concentration in the syringe was adjusted to 200 – 300  $\mu\text{M}$  with the experimental buffer. To examine a potential buffer dependency of the binding process, experiments were carried out in three different buffer systems containing either 50 mM HEPES, Tris or Tricine, 200 mM NaCl and 0.037% Tween 20, pH 7.8 at least in duplicate. The experimentally determined enthalpy  $\Delta H_{obs}$  was plotted against the heat of ionization  $\Delta H_{ion}$  as reported in literature and fitted by linear regression [Christensen, 1976; Fukada and Takahashi, 1998]. The intercept of the y axis represents the enthalpy corrected for buffer contributions and the slope discloses the number of protons released or taken up per mole formed complex upon binding according to Equation 9.7. The positive slope for the measurements involving the *lin*-

benzoguanine series indicated the uptake of one proton by the ligand. Therefore, enthalpy values were corrected for buffer dependency as described above. The *lin*-benzohypoxanthine series did not show a significant proton uptake. Thus, the mean of three independent measurements was calculated.

$$\Delta H_{obs} = \Delta H_{bind} + n_{H^+} \cdot \Delta H_{ion} \quad (\text{Equation 9.7})$$

All ITC experiments were run at 25°C after a stable baseline had been achieved. The reference cell contained filtered demineralized water. The initial delay before the injections were started and the spacing between each injection were adjusted to 180 s. The first injection contained 0.3 – 0.5 µL of the ligand solution followed by 14 – 24 injections of 1.0 – 2.0 µL. A stirring speed of 1000 rpm was chosen. Raw data were collected as released heat per time.

To analyze the raw data using the Origin 7.0 software, the baseline and integration limits were adjusted manually. After integrating the area under the peaks, the first data point was removed due to its reduced accuracy [Mizoue and Tellinghuisen, 2004]. The influence of the heat of dilution was corrected considering the heat contributions collected after saturation of the protein.  $K_d$  as well as  $\Delta H^0$  were extracted applying a single-site binding model as provided by the manufacturer. Subsequently,  $-T\Delta S^0$  was calculated according to the Gibbs-Helmholtz equation.

### 9.2.8 Dynamic Light Scattering

The potential aggregation behavior of the 5-azacytosine derivative **11** was determined by Dynamic Light Scattering (DLS) under assay conditions (200 mM, HEPES, 20 mM MgCl<sub>2</sub>, 0.037% (v/v) Tween 20, 5% (v/v) DMSO, pH 7.3) at a final concentration of 20 µM using a SpectroSize™ 300 (Molecular Dimensions Limited) equipped with an optical power of 100 mW at a wavelength of 660 nm at 37°C. The scattered light was detected at a 90° angle. The given data comprised at least two measurements of 30 runs over 15 s. Clotrimazole served as a reference as a strong aggregator at a concentration of 100 µM in assay buffer containing 5% (v/v) DMSO [Seidler *et al.*, 2003]. Count rate and autocorrelation function were plotted using the program *R*.

The calculation of the particle size was based on the Brownian motion. The bigger a particle, the slower is its movement through a medium and the slower is the change in its scattered light. This relationship is expressed by the Stokes-Einstein equation:

$$d = \frac{k_B T}{3\pi\eta D} \quad (\text{Equation 9.8})$$

$d$	Hydrodynamic diameter
$k_B$	Boltzmann constant
$T$	Temperature in Kelvin
$\eta$	Viscosity of the medium
$D$	Diffusion coefficient

## 9.2.9 X-ray Crystallography

### 9.2.9.1 *Z. mobilis* TGT Crystallization

Crystals were grown in the presence of the inhibitor using the sitting drop vapor diffusion method at 291 K. The protein solution was adjusted to  $12 \text{ mg} \cdot \text{mL}^{-1}$  by dilution with high salt buffer (10 mM Tris, 2 M NaCl, 1 mM EDTA, pH 7.8) and incubated with the inhibitor previously dissolved in 100% DMSO at a final concentration up to 1.5 mM depending on the solubility of the corresponding ligand. This solution was mixed with 1.5  $\mu\text{L}$  reservoir solution (100 mM MES, pH 5.5, 10% (v/v) DMSO, 11 – 13% (w/v) PEG 8000) to a 3  $\mu\text{L}$  droplet. The reservoir contained 1.0 mL of the above-mentioned solution. Within one week crystals showing an appropriate size for data collection were obtained.

For crystals grown in the presence of a ligand which did not show an appropriate size or form sufficient for diffraction experiments a soaking protocol was applied (TGT-**3a**, TGT-**5b**, TGT Val262Thr-**4a**, TGT Val262Asp-**4a**, TGT Val262Cys-**4a**, TGT-**7**). For that purpose, crystals were grown in the absence of the ligand according to the previously described protocol. Instead of mixing the protein solution with the inhibitor before crystallization, pre-grown apo crystals were transferred to a 3  $\mu\text{L}$  droplet of reservoir solution mixed with the stock solution containing the desired ligand to a final concentration of 1 mM. The droplet was sealed against 1.0 mL reservoir solution and soaked into the crystal overnight.

### 9.2.9.2 pH Soaking

For pH soaking a single crystal grown at pH 5.5 was transferred and incubated over a period of 5 min into droplets consisting of reservoir solution and buffer containing 100 mM Tris, pH 7.8, 10% (v/v) DMSO, 8% (w/v) PEG 8000 mixed in a 3:2, 1:2, and 2:3 manner, respectively. Subsequently, the crystal was equilibrated in a droplet of sole buffer pH 7.8 overnight to guarantee a homogeneous pH over the whole crystal. As a cryoprotectant for the apo crystal 20% (v/v) PEG 400 (cryo buffer D) was used.

### 9.2.9.3 Data Collection

For data collection the crystals were transferred into the corresponding cryoprotectant solution (cryo buffers A – C, Table 8.3) for 20 seconds followed by immediate flash-freezing in liquid nitrogen. Due to their low affinity, the corresponding 5-azacytosine based ligands were added to the cryoprotectant solution in the same concentration as applied to the crystallization conditions to prevent diffusion out of the crystals.

Data sets were collected at the BESSY II (Helmholtz-Zentrum, Berlin, Germany) beamline 14.2 at a wavelength of  $\lambda = 0.91841 \text{ \AA}$  and 14.3 at a wavelength of  $\lambda = 0.89460 \text{ \AA}$  both using a Rayonix MX225 CCD detector. Additionally, complex structures were collected at the PETRA III (EMBL, Hamburg, Germany) beamline P14 at a wavelength of  $\lambda = 1.23953 \text{ \AA}$  and  $\lambda = 0.97627 \text{ \AA}$  using a PILATUS 6M-F detector. To minimize radiation damage, all data sets were collected at cryo-conditions (100 K).

All TGT crystals showed the monoclinic space group C2 containing one monomer in the asymmetric unit. Data sets TGT-**3a**, TGT-**4a**, TGT-**4b**, TGT-**5a**, TGT-**5b**, TGT-**6a**, TGT-**6b**, TGT Val262Thr, TGT Val262Asp, TGT Val262Cys, TGT Val262Thr-**4a**, TGT Val262Cys-**4a**, TGT-**7**, TGT-**7a**, TGT-**8**, TGT-**9**, and TGT-**10** were processed and scaled with the *HKL2000* package [Otwinowski and Minor, 1997]. Data processing and scaling for TGT Asp102Asn-**3a**, Asp156Asn-**3a**, TGT-apo pH 7.8, TGT-**7b**, TGT-**7c**, and TGT-**7d** were performed with *XDS* and *XSCALE* [Kabsch, 2010], respectively. Data set TGT Val262Asp-**4a** was processed with *iMOSFLM 1.0.6* [Battye *et al.*, 2011] and subsequently scaled with the program *SCALA* of the ccp4 program suite [Winn *et al.*, 2011]. Cell dimensions, data collection and processing statistics are given in the appendix (Table 10.8).

#### 9.2.9.4 Calculation of the Anomalous Map

The anomalous electron density was generated using the program *ANODE* [Thorn and Sheldrick, 2011]. Unit-cell parameters, space-group notation, and atom coordinates were extracted from the pdb-files. Reflection indices,  $|F_A|$  and  $\alpha$  values originated from the reflection file derived with *XPREP* [Sheldrick, 2008].

#### 9.2.9.5 Structure Determination and Refinement

The coordinates of the TGT apo-structure 1PUD served as a starting model for molecular replacement using the program *Phaser MR* of the ccp4 program suite [McCoy, 2007]. Structures were refined using the program *Phenix* [Adams *et al.*, 2010] starting with a first cycle of simulated annealing using default parameters. Further refinement cycles comprised the coordinate xyz, occupancy and individual B-factor refinement as well as applying metal restraints for the zinc ion. In case of the structures TGT-**3a**, TGT-**6b**, and TGT-apo pH 7.8 the weights between X-ray target and stereochemistry restraints were optimized in addition to individual atomic displacement parameter (ADP) weights refined for TGT-**3a**, TGT-**4a**, TGT-**4b**, TGT-**5a**, TGT-**5b**, TGT-**6a**, TGT-**6b**, TGT-**7**, TGT-**7a**, TGT-**7b**, TGT-**7c**, TGT-**7d**, and TGT-apo pH 7.8. The temperature factors of structures TGT-**3a**, TGT-**4a**, TGT-apo pH 7.8, TGT-**4b**, TGT-**5a**, TGT-**6a**, TGT-**6b**, TGT Val262Thr, TGT Val262Asp, TGT Val262Cys, TGT Val262Asp-**4a**, TGT-**7**, TGT-**7a**, TGT-**7b**, TGT-**7c**, and TGT-**7d** were refined anisotropically, while for structures TGT Asp102Asn-**3a**, Asp156Asn-**3a**, TGT-**5b**, TGT Val262Thr-**4a**, TGT Val262Cys-**4a**, TGT-**8**, TGT-**9**, and TGT-**10** a TLS refinement was performed after selecting appropriate TLS groups with the phenix.find\_tls\_groups option [Painter and Merritt, 2006]. The calculation of the  $R_{\text{free}}$  value comprised a 2 – 5% fraction of the data.

For all structures amino acid side chains were fitted according to their  $\sigma_A$ -weighted  $2|F_o| - |F_c|$  and  $|F_o| - |F_c|$  electron density obtained in the program *Coot* [Emsley and Cowtan, 2004]. After the initial refinement cycles, the zinc ion as well as water and glycerol molecules were implemented in the model. For adding water molecules, the option “update waters” included in Phenix was used after increasing the hydrogen bond length threshold for the solvent-model and solvent-solvent contacts to 2.3 Å. The inserted molecules were visually reviewed afterwards. Ligand restraints were generated by the CSD based

gradeWebServer [<http://grade.globalphasing.org>] in case of **3a**, **4a**, **4b**, **5a**, **5b**, **6a**, **6b**, **7b**, **7c** and **7d**. Ligands were built and minimized using the program *MOE* [*MOE 2012.10*] and geometric restraints calculated subsequently by the program *Monomer Library Sketcher* [Winn *et al.*, 2011] in case of **7**, **7a**, **8**, **9**, and **10**. Multiple protein residue conformations were assigned in case a reasonable electron density was observed and were kept during refinement if the side chain with the lowest occupancy showed a value of at least 20%. In case of **4a**, **4b**, **5a**, **5b**, **6a**, **6a**, and **7b** the occupancies of the 2-substituent were refined due to elevated B-factors observed for this portion. In the same manner the occupancies of **8**, **9**, and **10** were refined for the whole molecule. Ramachandran plots were generated with the program *PROCHECK* [Laskowski *et al.*, 1993]. For the analysis of temperature factors the program *Moleman* was used [Kleywegt, 2001]. The burial of molecular ligand portions in the protein binding pocket was computed using the program *MS* [Connolly, 1983].

### 9.3 Synthesis

Ligands were synthesized and purified as described in detail elsewhere [Hörtner *et al.*, 2007; Barandun *et al.*, 2012; Neeb *et al.*, 2014; Neeb *et al.*, 2014].



## 10. Appendix

### 10.1 Calculation of $pK_a$ Values of Titratable Groups of the TGT Active Site

#### 10.1.1 TGT Wild Type

**Table 10.1** Calculated  $pK_a$  values of the amino acid residues within a 12 Å sphere around C $\gamma$  of Tyr106 of TGT wild type and **3a** before and after complex formation.

Residue	$pK_a$ prior to complex formation <sup>[a]</sup>	$pK_a$ in complexed state	
		N3	N5
CYS-281	10.40	10.32	10.25
CYS-158	11.87	11.96	12.23
ASP-156	5.44	4.40	3.02
ASP-280	2.18	0.59	1.44
ASP-266	3.24	3.21	3.24
ASP-267	0.66	0.48	0.66
ASP-245	3.80	3.79	3.80
ASP-238	3.98	3.98	3.98
ASP-102	1.63	1.64	-2.12
LYS-52	8.75	8.74	8.71
LYS-264	12.39	12.35	12.37
HIS-73	4.08	3.97	3.98
<b>3a</b> <sup>[b]</sup>		6.65	7.00
GLU-235	4.20	3.55	4.30
GLU-239	3.24	3.24	3.23
GLU-173	4.57	4.58	4.60
GLU-157	3.13	3.06	3.16
TYR-72	12.02	11.96	11.93
TYR-106	10.77	11.60	11.53
TYR-226	11.66	11.67	11.67
TYR-381	10.83	10.81	10.82
TYR-354	15.55	15.53	15.52
TYR-161	10.70	10.67	10.70
TYR-258	16.51	16.76	16.65

<sup>a</sup>For the Poisson-Boltzmann calculations a dielectric constant of  $\epsilon = 20$  was applied. <sup>b</sup> $pK_a$  values in aqueous solution according to Barandun *et al.*, 2012.

### 10.1.2 TGT Asp102Asn

**Table 10.2** Calculated  $pK_a$  values of the amino acid residues within a 12 Å sphere around Cy of Tyr106 of TGT Asp102Asn and **3a** before and after complex formation.

Residue	$pK_a$ prior to complex formation <sup>[a]</sup>	$pK_a$ in complexed state N5
CYS-281	8.73	8.67
CYS-158	10.40	11.00
ASP-156	2.80	2.57
ASP-280	1.37	1.19
ASP-266	3.70	3.70
ASP-267	0.47	0.52
ASP-245	3.71	3.72
ASP-238	3.90	3.89
LYS-52	8.40	8.36
LYS-264	12.97	12.97
<b>3a</b> <sup>[b]</sup>		-1.14
HIS-73	3.94	3.84
GLU-235	3.67	4.00
GLU-239	3.27	3.26
GLU-173	4.40	4.41
GLU-157	2.47	2.52
TYR-72	11.17	11.06
TYR-106	10.69	11.37
TYR-226	10.83	10.83
TYR-381	11.75	11.73
TYR-354	14.59	14.59
TYR-161	10.67	10.67
TYR-258	15.58	15.55

<sup>a, b</sup>The same conditions are applied as denoted in Table 10.1.

### 10.1.3 TGT Asp156Asn

**Table 10.3** Calculated  $pK_a$  values of the amino acid residues within a 12 Å sphere around Cy of Tyr106 of TGT Asp156Asn and **3a** before and after complex formation.

Residue	$pK_a$ prior to complex formation <sup>[a]</sup>	$pK_a$ in complexed state N5
CYS-281	9.31	9.28
CYS-158	10.46	10.76
ASP-280	2.29	1.36
ASP-266	3.42	3.40
ASP-267	1.70	1.70
ASP-245	3.97	3.95
ASP-238	4.09	4.08
ASP-102	1.29	-2.31
LYS-52	8.66	8.64
LYS-264	11.32	11.30
<b>3a</b> <sup>[b]</sup>		5.32
HIS-73	3.84	3.50
GLU-235	4.37	4.31
GLU-239	3.50	3.47
GLU-173	4.55	4.52
GLU-157	3.72	3.46
TYR-72	10.83	10.77
TYR-106	11.62	12.34
TYR-226	11.36	11.36
TYR-381	9.18	9.19
TYR-354	14.44	14.45
TYR-161	10.85	10.85
TYR-258	16.34	16.52

<sup>a, b</sup>The same conditions are applied as denoted in Table 10.1.

## 10.2 Anomalous Density Data

Table 10.4 Anomalous density data for TGT·5a.

Peak	Height	Nearest atom <sup>[a]</sup>	[Å]	Peak	Height	Nearest atom <sup>[a]</sup>	[Å]
<b>1</b>	53.36	ZN_B:ZN1	0.04	<b>33</b>	4.39	C_A:TYR381	1.37
<b>2</b>	8.69	SD_A:MET176	0.09	<b>34</b>	4.37	OD1_A:ASN304	1.34
<b>3</b>	7.46	SD_A:MET153	0.34	<b>35</b>	4.36	CA_A:VAL322	1.36
<b>4</b>	7.43	SD_A:MET278	0.32	<b>36</b>	4.35	OE2_A:GLU273	1.00
<b>5</b>	7.27	SD_A:MET75	0.16	<b>37</b>	4.30	NH2_A:ARG211	0.65
<b>6</b>	6.94	SD_A:MET250	0.16	<b>38</b>	4.27	C_A:SER205	0.42
<b>7</b>	6.73	SG_A:CYS318	0.22	<b>39</b>	4.27	OG_A:SER371	1.66
<b>8</b>	6.63	SG_A:CYS158	0.12	<b>40</b>	4.24	CA_A:GLY87	2.13
<b>9</b>	6.53	O_S:HOH1	0.11	<b>41</b>	4.23	O_S:HOH14 <sup>[b]</sup>	0.15
<b>10</b>	6.43	SD_A:MET43	0.22	<b>42</b>	4.22	O_A:SER170	0.46
<b>11</b>	6.23	SD_A:MET260	0.13	<b>43</b>	4.20	CA_A:LEU54	2.16
<b>12</b>	6.00	SD_A:MET172	0.37	<b>44</b>	4.20	N_A:ASP96	1.09
<b>13</b>	5.78	SD_A:MET240	0.28	<b>45</b>	4.18	NE_A:ARG34	2.24
<b>14</b>	5.29	SG_A:CYS318	1.18	<b>46</b>	4.17	O_A:LEU357	1.72
<b>15</b>	5.06	O_S:HOH2	0.61	<b>47</b>	4.17	CB_A:ALA352	1.58
<b>16</b>	4.94	SG_A:CYS320	0.83	<b>48</b>	4.15	OE1_A:GLN372	4.30
<b>17</b>	4.92	NZ_A:LYS312	1.40	<b>49</b>	4.13	O_S:HOH62	0.26
<b>18</b>	4.90	N_A:CYS323	1.90	<b>50</b>	4.12	C_A:VAL269	0.73
<b>19</b>	4.83	SD_A:MET358	0.11	<b>51</b>	4.12	O_A:GLU317	1.46
<b>20</b>	4.70	SD_A:MET240	1.38	<b>52</b>	4.11	CG2_A:VAL206	1.43
<b>21</b>	4.61	O_S:HOH369	1.36	<b>53</b>	4.11	O_S:HOH349	1.25
<b>22</b>	4.60	SD_A:MET344	0.46	<b>54</b>	4.11	O_A:SER287	1.16
<b>23</b>	4.58	CE1_A:HIS349	1.15	<b>55</b>	4.10	O_S:HOH359	1.55
<b>24</b>	4.55	O_A:ALA383	5.09	<b>56</b>	4.09	O_S:HOH101	0.90
<b>25</b>	4.54	CA_A:GLU235	1.14	<b>57</b>	4.08	O_S:HOH269	0.48
<b>26</b>	4.50	O_S:HOH123	0.28	<b>58</b>	4.07	N_A:ARG132	3.42
<b>27</b>	4.50	O_A:GLY230	0.31	<b>59</b>	4.06	N_A:GLU173	0.37
<b>28</b>	4.44	O_S:HOH428	3.37	<b>60</b>	4.06	CD2_A:LEU89	2.06
<b>29</b>	4.44	C_A:TRP95	1.08	<b>61</b>	4.03	NE1_A:TRP178	0.88
<b>30</b>	4.42	O_S:HOH196	1.15	<b>62</b>	4.02	O_S:HOH203	0.91
<b>31</b>	4.41	CB_A:ASP363	1.51	<b>63</b>	4.01	O_S:HOH269	3.96
<b>32</b>	4.40	SD_A:MET109	0.30				

<sup>a</sup>The nearest atom is listed in the form atom\_chain:residue. <sup>b</sup>Anomalous data for the identified chloride ion.

**Table 10.6** Anomalous density data for TGT-5b.

Peak	Height	Nearest atom <sup>[a]</sup>	[Å]	Peak	Height	Nearest atom <sup>[a]</sup>	[Å]
<b>1</b>	60.59	ZN_B:ZN1	0.03	<b>20</b>	4.37	O_F:HOH1	1.13
<b>2</b>	7.01	SD_A:MET153	0.21	<b>21</b>	4.37	SD_A:MET346	0.19
<b>3</b>	6.50	SD_A:MET43	0.42	<b>22</b>	4.31	O_A:ARG34	1.67
<b>4</b>	6.50	SD_A:MET358	0.21	<b>23</b>	4.31	SD_A:MET172	0.19
<b>5</b>	6.03	SD_A:MET344	0.38	<b>24</b>	4.30	O_A:SER366	1.59
<b>6</b>	5.98	SD_A:MET278	0.14	<b>25</b>	4.30	SG_A:CYS281	0.48
<b>7</b>	5.73	O_G:HOH1 <sup>[b]</sup>	0.11	<b>26</b>	4.29	CA_A:GLY234	4.95
<b>8</b>	5.24	SD_A:MET176	0.25	<b>27</b>	4.29	O_S:HOH292	3.36
<b>9</b>	5.01	SD_A:MET250	0.34	<b>28</b>	4.25	CB_A:THR295	1.21
<b>10</b>	4.92	O_S:HOH241	1.23	<b>29</b>	4.25	O_A:MET109	4.66
<b>11</b>	4.77	O_S:HOH290	3.05	<b>30</b>	4.22	OD2_A:ASP238	7.99
<b>12</b>	4.74	SG_A:CYS158	0.14	<b>31</b>	4.19	SG_A:CYS323	0.16
<b>13</b>	4.66	SD_A:MET93	0.50	<b>32</b>	4.17	CA_A:SER287	1.17
<b>14</b>	4.47	NH1_A:ARG34	4.05	<b>33</b>	4.16	CA_A:SER112	1.89
<b>15</b>	4.42	CA_A:ARG303	0.92	<b>34</b>	4.15	N_A:SER17	0.83
<b>16</b>	4.41	SD_A:MET260	0.43	<b>35</b>	4.10	O_S:HOH184	3.08
<b>17</b>	4.41	N_A:LEU68	1.05	<b>36</b>	4.09	C_A:LYS125	1.76
<b>18</b>	4.39	OD1_A:ASN290	2.21	<b>37</b>	4.09	CZ_A:PHE42	1.14
<b>19</b>	4.37	O_F:HOH1	1.13	<b>38</b>	4.08	CB_A:SER205	2.01

<sup>a</sup>The nearest atom is listed in the form atom\_chain:residue. <sup>b</sup>Anomalous data for the identified chloride ion.

### 10.3 Thermodynamic Data of *lin*-Benzopurines

**Table 10.6** Raw data as well as thermodynamic profiles after correction for buffer distribution of the investigated ligands measured at 25 °C.

Ligand	$K_d$ (nM)	$\Delta G^0$ (kJ · mol <sup>-1</sup> )	buffer	$\Delta H_{obs}$ (kJ · mol <sup>-1</sup> )	$-T\Delta S^0$ (kJ · mol <sup>-1</sup> ) <sup>[a]</sup>
<b>3a</b>	52.4 ± 6.9	-41.6 ± 0.3	Hepes	-74.8 ± 1.8	33.2 ± 1.9 <sup>[b]</sup>
			Tricine	-66.7 ± 0.1	25.1 ± 0.3
			Tris	-50.9 ± 0.9	9.3 ± 0.9
				<b>-97.4 ± 2.2</b>	<b>55.8 ± 2.2<sup>[c]</sup></b>
<b>3b</b>	411.0 ± 63.8	-36.5 ± 0.3	Hepes	-48.0 ± 0.2	11.5 ± 0.4
			Tricine	-47.7 ± 0.2	11.2 ± 0.5
			Tris	-47.6 ± 0.3	11.1 ± 0.5
				<b>-47.7 ± 0.3</b>	<b>11.2 ± 0.5</b>
<b>4a</b>	35.0 ± 6.9	-42.6 ± 0.5	Hepes	-78.3 ± 1.2	35.7 ± 1.3
			Tricine	-65.8 ± 0.3	23.2 ± 0.6
			Tris	-51.9 ± 1.1	9.3 ± 1.2
				<b>-96.4 ± 4.0</b>	<b>53.8 ± 4.0</b>
<b>4b</b>	369.2 ± 25.3	-36.7 ± 0.2	Hepes	-50.1 ± 0.4	13.4 ± 0.4
			Tricine	-49.3 ± 0.4	12.6 ± 0.4
			Tris	-47.4 ± 0.4	10.7 ± 0.4
				<b>-48.9 ± 1.4</b>	<b>12.2 ± 1.4</b>
<b>5a</b>	34.3 ± 6.5	-42.6 ± 0.5	Hepes	-66.7 ± 0.5	24.1 ± 0.7
			Tricine	-56.9 ± 0.5	14.3 ± 0.7
			Tris	-48.8 ± 0.4	6.2 ± 0.6
				<b>-78.9 ± 3.7</b>	<b>36.3 ± 3.7</b>
<b>5b</b>	407.4 ± 46.0	-36.5 ± 0.3	Hepes	n.d.	n.d. <sup>[d]</sup>
			Tricine	-50.6 ± 0.7	14.1 ± 0.8
			Tris	-50.6 ± 1.7	14.1 ± 1.8
				<b>-50.6 ± 1.4</b>	<b>14.1 ± 1.4</b>
<b>6a</b>	111.6 ± 11.8	-39.7 ± 0.3	Hepes	-50.9 ± 0.6	11.2 ± 0.6
			Tricine	-44.0 ± 1.5	4.3 ± 1.5
			Tris	-31.2 ± 0.8	-8.5 ± 0.8
				<b>-66.3 ± 0.8</b>	<b>26.6 ± 0.9</b>
<b>6b</b>	1638.9 ± 204.7	-33.0 ± 0.3	Hepes	-42.1 ± 0.4	9.1 ± 0.5
			Tricine	-41.8 ± 0.4	8.8 ± 0.5
			Tris	-42.3 ± 0.5	9.3 ± 0.6
				<b>-42.1 ± 0.3</b>	<b>9.1 ± 0.4</b>

<sup>a</sup> $-T\Delta S^0$  was calculated according to the Gibbs-Helmholtz equation. <sup>b</sup>Errors were estimated by means of standard deviation for  $K_d$  and  $\Delta G^0$  comprising at least six measurements and for  $\Delta H_{obs}$  at least two. The error for  $-T\Delta S^0$  was calculated according to error propagation. <sup>c</sup>Buffer corrected data are displayed in bold. <sup>d</sup>Not determined.

## 10.4 Sequencing Results of the Strep-tagged TGT Variants after Cleavage

**D156N** GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60  
**WT** GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60  
 \*\*\*\*\*

**D156N** VRATGADIILGNTYHMLLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120  
**WT** VRATGADIILGNTYHMLLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120  
 \*\*\*\*\*

**D156N** EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFNECTPYPATPSRAASSMERSMRW 180  
**WT** EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180  
 \*\*\*\*\*:\*\*\*\*\*

**D156N** AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240  
**WT** AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240  
 \*\*\*\*\*

**D156N** EMFRVLDFSVPMPLPDDKPHYLMGVGKPDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300  
**WT** EMFRVLDFSVPMPLPDDKPHYLMGVGKPDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300  
 \*\*\*\*\*

**D156N** PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIIFYQQLM 360  
**WT** PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIIFYQQLM 360  
 \*\*\*\*\*

**D156N** QKIRDSISEGRFSQFAQDFRARYFARNS 388  
**WT** QKIRDSISEGRFSQFAQDFRARYFARNS 388  
 \*\*\*\*\*

### Scheme 10.1 Sequence alignment of mutant Asp156Asn.

**V262C** GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60  
**WT** GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60  
 \*\*\*\*\*

**V262C** VRATGADIILGNTYHMLLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120  
**WT** VRATGADIILGNTYHMLLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120  
 \*\*\*\*\*

**V262C** EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180  
**WT** EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180  
 \*\*\*\*\*

**V262C** AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240  
**WT** AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240  
 \*\*\*\*\*

**V262C** EMFRVLDFSVPMPLPDDKPHYLMGCGKPDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300  
**WT** EMFRVLDFSVPMPLPDDKPHYLMGVGKPDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300  
 \*\*\*\*\*

**V262C** PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIIFYQQLM 360  
**WT** PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIIFYQQLM 360  
 \*\*\*\*\*

**V262C** QKIRDSISEGRFSQFAQDFRARYFARNS 388  
**WT** QKIRDSISEGRFSQFAQDFRARYFARNS 388  
 \*\*\*\*\*

### Scheme 10.2 Sequence alignment of mutant Val262Cys.

```

V262D      GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60
WT         GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60
          *****

V262D      VRATGADIILGNTYHLMRLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120
WT         VRATGADIILGNTYHLMRLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120
          *****

V262D      EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180
WT         EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180
          *****

V262D      AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240
WT         AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240
          *****

V262D      EMFRVLDFSVPMPLPDDKPHYLMGDKGPKDDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300
WT         EMFRVLDFSVPMPLPDDKPHYLMGDKGPKDDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300
          *****

V262D      PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIAFYQQLM 360
WT         PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIAFYQQLM 360
          *****

V262D      QKIRDSISEGRFSQFAQDFRARYFARNS 388
WT         QKIRDSISEGRFSQFAQDFRARYFARNS 388
          *****

```

Scheme 10.3 Sequence alignment of mutant Val262Asp.

```

V262T      GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60
WT         GSMVEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAAATVKALKPET 60
          *****

V262T      VRATGADIILGNTYHLMRLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120
WT         VRATGADIILGNTYHLMRLRPGAERIAKLGGLHSFMGWDRPILTDSSGGYQVMSLSSLTKQS 120
          *****

V262T      EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180
WT         EEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYPATPSRAASSMERSMRW 180
          *****

V262T      AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240
WT         AKRSRDAFDSRKEQAENAALFGIQQGSVFENLRQQSADALAEIGFDGYAVGGLAVGEGQD 240
          *****

V262T      EMFRVLDFSVPMPLPDDKPHYLMGTGPKDDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300
WT         EMFRVLDFSVPMPLPDDKPHYLMGTGPKDDIVGAVERGIDMFDCVLPTRSGRNGQAFTWDG 300
          *****

V262T      PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIAFYQQLM 360
WT         PINIRNARFSEDLKPLDSECHCAVCQKWSRAYIHHLIRAGEILGAMLMTEHNIAFYQQLM 360
          *****

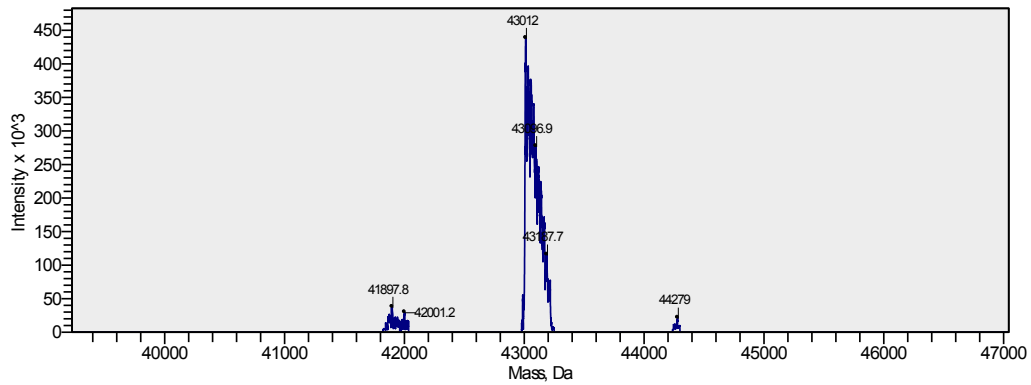
V262T      QKIRDSISEGRFSQFAQDFRARYFARNS 388
WT         QKIRDSISEGRFSQFAQDFRARYFARNS 388
          *****

```

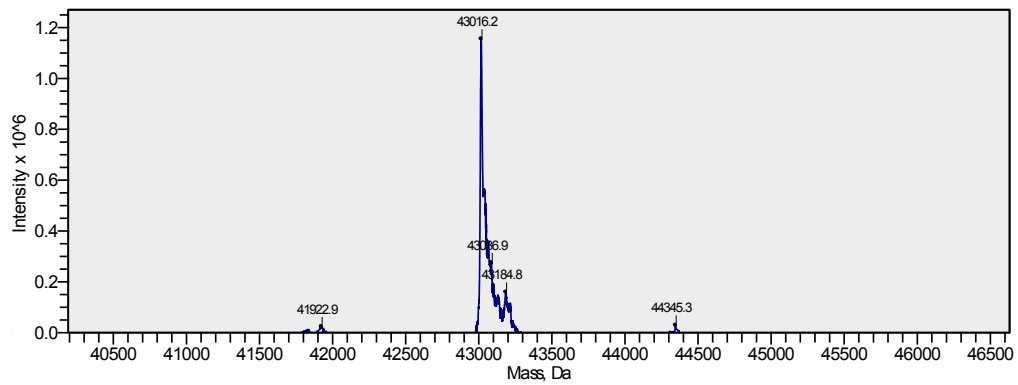
Scheme 10.4 Sequence alignment of mutant Val262Thr.



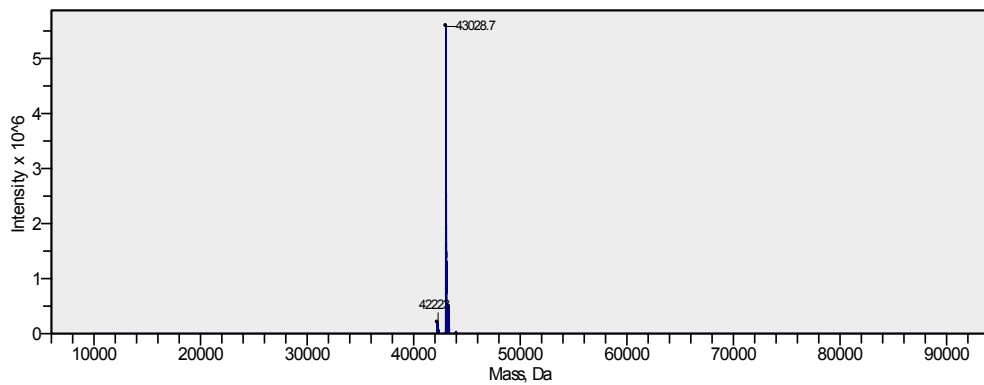
## 10.5 Mass Spectra of the Strep-tagged TGT Variants after Cleavage



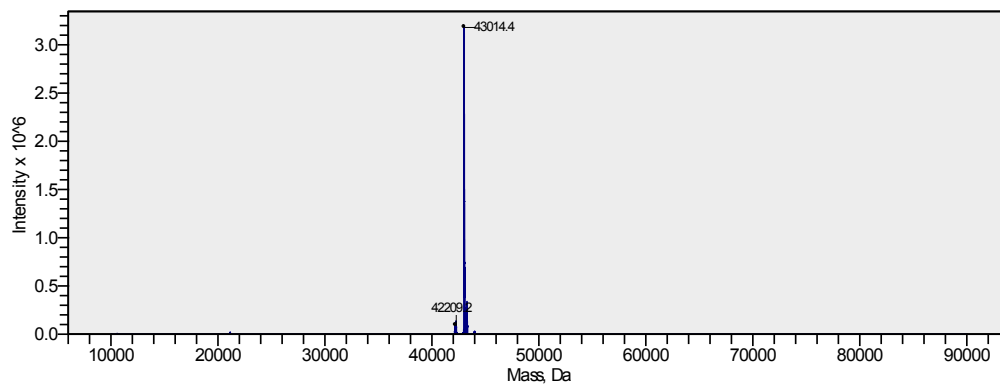
**Figure 10.1** Mass spectrum of TGT Asp156Asn (theoretical mass: 43012.8 Da).



**Figure 10.2** Mass spectrum of TGT Val262Cys (theoretical mass: 43017.8 Da).



**Figure 10.3** Mass spectrum of TGT Val262Asp (theoretical mass: 43029.7 Da).



**Figure 10.4** Mass spectrum of TGT Val262Thr (theoretical mass: 43015.8 Da).

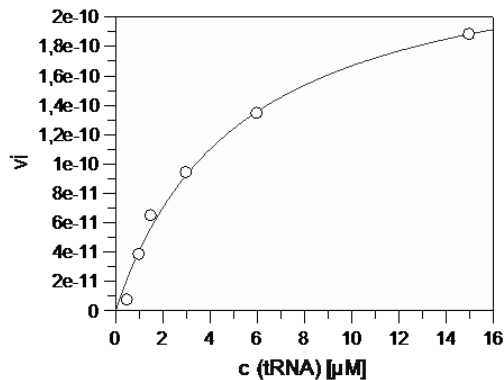
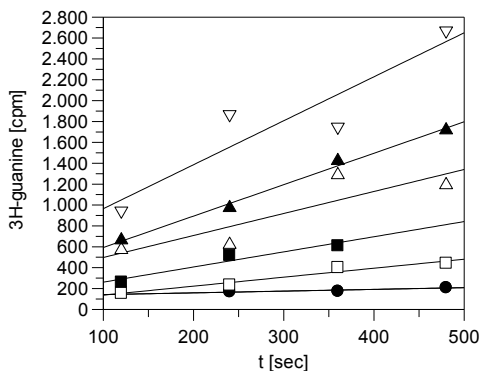
## 10.6 Enzyme Kinetics of the Strep-tagged TGT Variants after Cleavage

**Table 10.7** Initial velocities, Michaelis-Menten plot and kinetic data of the different TGT variants.

**TGT Wild Type**  
[Jakobi, 2013]

$K_M$  (tRNA) 4.8  $\mu\text{M}$   
 $k_{\text{cat}}$   $1.2 \cdot 10^{-2} \text{ s}^{-1}$

**TGT D156N**

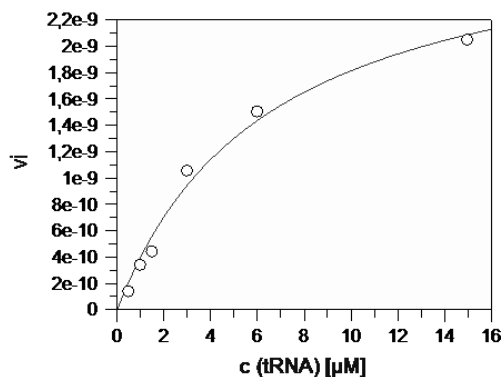
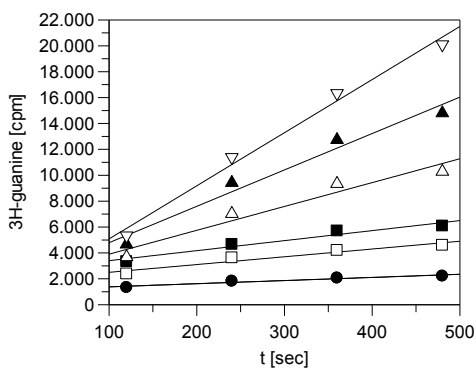


$K_M$  (tRNA) 5.25  $\mu\text{M}$   
 $k_{\text{cat}}$   $3.4 \cdot 10^{-3} \text{ s}^{-1}$

**TGT V262C**

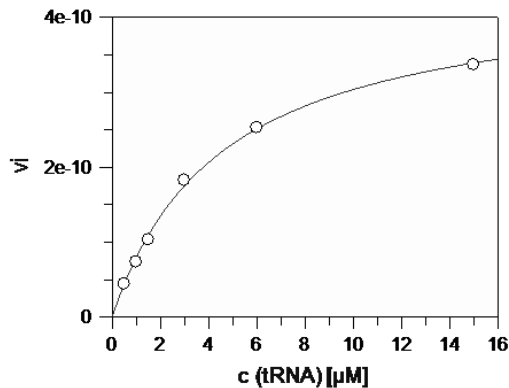
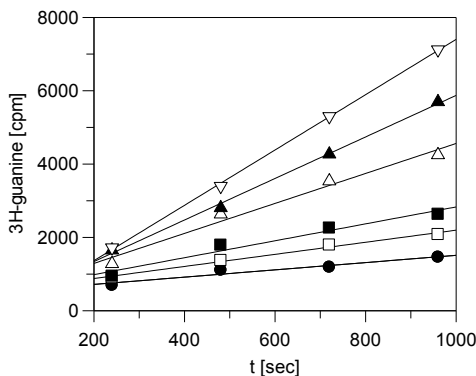
Active (multiple measurements did not yield an evaluable result)

**TGT V262D**



$K_M$  (tRNA) 6.49  $\mu\text{M}$   
 $k_{\text{cat}}$   $4.0 \cdot 10^{-3} \text{ s}^{-1}$

**TGT V262T**



$K_M$  (tRNA) 3.52  $\mu\text{M}$   
 $k_{\text{cat}}$   $5.3 \cdot 10^{-3} \text{ s}^{-1}$

## 10.7 Data Collection and Refinement Statistics

**Table 10.8** Data collection, processing and refinement statistics for the investigated TGT-ligand-complexes.

<b>Crystal data</b>	<b>TGT-3a</b>	<b>Asp102Asn-3a</b>	<b>Asp156Asn-3a</b>	<b>TGT-4a</b>	<b>Apo pH 7.8</b>
<b>PDB ID</b>	<b>4PUK</b>	<b>4PUL</b>	<b>4PUM</b>	<b>4PUJ</b>	<b>4PUN</b>
<i>A) Data collection and processing</i>					
Collection site	BESSY 14.2	PETRA P14	PETRA P14	BESSY 14.2	PETRA P14
No. crystals used	1	1	1	1	1
$\lambda$ [Å]	0.91841	1.23953	1.23953	0.91841	0.97627
Space group	C2	C2	C2	C2	C2
<i>Unit cell parameters</i>					
a [Å]	90.8	89.7	89.7	90.6	90.5
b [Å]	65.0	64.2	64.7	64.8	64.7
c [Å]	71.0	70.4	70.8	71.0	69.9
$\beta$ [°]	96.3	92.9	93.1	96.2	96.0
<i>B) Diffraction data<sup>[a]</sup></i>					
Resolution range [Å]	30 - 1.49 (1.52 - 1.49)	80 - 1.65 (1.85 - 1.65)	80 - 1.93 (1.98 - 1.93)	30 - 1.42 (1.44 - 1.42)	50 - 1.25 (1.30 - 1.25)
Unique reflections	66573 (3339)	43729 (9984)	30236 (2238)	76897 (3785)	107127 (11652)
R(I)sym [%] <sup>[b]</sup>	5.1 (48.1)	2.9 (27.5)	4.7 (50.5)	5.7 (50.5)	3.0 (50.0)
Completeness [%]	99.4 (98.8)	91.6 (73.8)	99.0 (99.4)	99.3 (98.4)	96.8 (95.3)
Redundancy	5.1 (4.9)	3.1 (2.6)	4.7 (4.6)	5.2 (5.1)	3.3 (3.3)
I/ $\sigma$ (I)	29.3 (3.4)	22.3 (3.8)	21.6 (3.4)	28.6 (3.9)	18.0 (2.7)
Wilson B-factor [Å <sup>2</sup> ]	15.7	22.4	25.8	12.8	14.5
Matthews Coefficient [Å <sup>3</sup> /Da]	2.4	2.4	2.4	2.4	2.4
<i>C) Refinement</i>					
PHENIX version	1.8.4_1496	1.8.1_1168	1.8.1_1168	1.8.4_1496	1.8.4_1496
Resolution range [Å]	27.3 - 1.49	70.3 - 1.65	70.7 - 1.93	29.5 - 1.42	45.0 - 1.25
Reflections used for R <sub>free</sub>	2000	1312	1512	2000	5357
Reflections used for R <sub>work</sub>	64506	42417	28724	74896	101770
<i>Final R values<sup>[a]</sup></i>					
R <sub>free</sub> [%] <sup>[c]</sup>	17.2	18.9	20.2	16.0	16.2
R <sub>work</sub> [%] <sup>[d]</sup>	13.9	16.7	15.9	12.9	13.9
<i>No. of atoms (non-hydrogen)</i>					
Protein atoms	2922	2829	2854	2911	2920
Water molecules	365	261	219	388	373
Ligand atoms	17	17	17	24	---
RMSD, angle [°]	1.3	1.1	1.1	1.0	1.4
RMSD, bond [Å]	0.014	0.008	0.009	0.007	0.013
<i>Ramachandran plot<sup>[e]</sup></i>					
Most favoured regions [%]	95.3	95.9	95.9	94.2	94.0
Additionally allowed regions [%]	4.4	3.8	3.8	5.4	5.6
Generously allowed regions [%]	0.3	0.3	0.3	0.3	0.3
<i>Mean B-factors [Å<sup>2</sup>]</i>					
Protein atoms	19.4	24.6	25.1	14.3	18.5
Water molecules	33.3	33.7	31.0	32.5	29.8
Ligand atoms	16.7	22.6	24.9	13.8	---

<b>Crystal data</b>	<b>TGT-4b</b>	<b>TGT-5a</b>	<b>TGT-5b</b>	<b>TGT-6a</b>	<b>TGT-6b</b>
<b>PDB ID</b>	<b>4Q4R</b>	<b>4Q4O</b>	<b>4Q4P</b>	<b>4Q4S</b>	<b>4Q4Q</b>
<i>A) Data collection and processing</i>					
Collection site	BESSY 14.2	BESSY 14.2	BESSY 14.3	BESSY 14.2	BESSY 14.2
No. crystals used	1	1	1	1	1
$\lambda$ [Å]	0.91841	0.91841	0.89460	0.91841	0.91841
Space group	C2	C2	C2	C2	C2
<i>Unit cell parameters</i>					
a [Å]	89.9	89.8	90.0	90.9	91.3
b [Å]	64.8	64.7	64.9	64.9	64.9
c [Å]	70.8	70.7	70.7	70.5	70.4
$\beta$ [°]	93.1	93.4	93.3	95.9	96.1
<i>B) Diffraction data<sup>[a]</sup></i>					
Resolution range [Å]	30 - 1.45 (1.48 - 1.45)	30 - 1.35 (1.37 - 1.35)	30 - 1.54 (1.57 - 1.54)	30 - 1.25 (1.27 - 1.25)	30 - 1.41 (1.43 - 1.41)
Unique reflections	71426 (3566)	87189 (4254)	60063 (2924)	109841 (5550)	78370 (3931)
R(I)sym [%] <sup>[b]</sup>	8.6 (46.5)	4.6 (38.8)	6.7 (30.2)	4.7 (45.5)	7.1 (49.1)
Completeness [%]	99.6 (99.9)	98.8 (98.3)	99.8 (99.5)	97.5 (98.8)	99.8 (99.9)
Redundancy	3.1 (3.0)	2.1 (2.0)	3.7 (3.0)	2.4 (2.3)	2.9 (2.9)
I/ $\sigma$ (I)	12.2 (2.2)	15.4 (2.0)	18.9 (3.6)	19.0 (2.1)	15.6 (3.1)
Wilson B-factor [Å <sup>2</sup> ]	17.3	14.9	17.5	11.9	10.7
Matthews Coefficient [Å <sup>3</sup> /Da]	2.4	2.4	2.4	2.4	2.4
<i>C) Refinement</i>					
PHENIX version	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.4_1496
Resolution range [Å]	29.8 - 1.45	27.1 - 1.35	27.2 - 1.54	22.6 - 1.25	25.2 - 1.41
Reflections used for R <sub>free</sub>	3607	4379	3033	1991	1675
Reflections used for R <sub>work</sub>	67805	82794	57026	107849	76695
<i>Final R values<sup>[a]</sup></i>					
R <sub>free</sub> [%] <sup>[c]</sup>	17.2	16.2	17.0	17.2	16.1
R <sub>work</sub> [%] <sup>[d]</sup>	13.6	13.7	15.2	13.2	12.5
<i>No. of atoms (non-hydrogen)</i>					
Protein atoms	2865	2923	2851	2892	2914
Water molecules	282	389	330	310	351
Ligand atoms	23	24	23	17	21
RMSD, angle [°]	1.0	1.1	1.1	1.0	0.9
RMSD, bond [Å]	0.007	0.007	0.008	0.006	0.006
<i>Ramachandran plot<sup>[e]</sup></i>					
Most favoured regions [%]	95.1	95.2	96.1	94.2	96.5
Additionally allowed regions [%]	4.5	4.4	3.6	5.1	3.2
Generously allowed regions [%]	0.3	0.3	0.3	0.6	0.3
<i>Mean B-factors [Å<sup>2</sup>]</i>					
Protein atoms	21.8	17.1	21.1	17.9	16.3
Water molecules	35.3	31.9	33.0	32.8	31.3
Ligand atoms	21.5	15.6	19.4	12.5	15.7 <sup>[f]</sup>

<b>Crystal data</b>	<b>Val262Thr</b>	<b>Val262Asp</b>	<b>Val262Cys</b>	<b>Val262Thr-4a</b>	<b>Val262Asp-4a</b>
<b>PDB ID</b>	<b>4Q8M</b>	<b>4IPP</b>	<b>4Q8N</b>	<b>4Q8O</b>	<b>4Q4P</b>
<i>A) Data collection and processing</i>					
Collection site	BESSY 14.2	BESSY 14.2	BESSY 14.2	BESSY 14.2	BESSY 14.2
No. crystals used	1	1	1	1	1
$\lambda$ [Å]	0.91841	0.91841	0.91841	0.91841	0.91841
Space group	C2	C2	C2	C2	C2
<i>Unit cell parameters</i>					
a [Å]	90.4	91.0	91.9	91.3	91.4
b [Å]	64.8	65.1	65.3	64.9	65.0
c [Å]	70.3	70.3	70.2	71.4	70.4
$\beta$ [°]	95.6	96.2	96.5	96.6	96.1
<i>B) Diffraction data<sup>[a]</sup></i>					
Resolution range [Å]	30 - 1.24 (1.26 - 1.24)	30 - 1.33 (1.35 - 1.33)	30 - 1.45 (1.48 - 1.45)	30 - 1.89 (1.92 - 1.89)	25.3 - 1.45 (1.53 - 1.45)
Unique reflections	113077 (5575)	93075 (4631)	72428 (3585)	32679 (1625)	72559 (10548)
R(I)sym [%] <sup>[b]</sup>	4.8 (48.0)	5.2 (24.4)	5.8 (47.1)	7.9 (43.4)	10.0 (49.5)
Completeness [%]	99.3 (98.5)	99.4 (99.9)	99.0 (99.7)	97.9 (96.3)	99.9 (99.9)
Redundancy	3.3 (3.3)	3.0 (3.0)	3.2 (3.0)	3.2 (3.2)	3.7 (3.7)
I/ $\sigma$ (I)	23.2 (2.5)	20.4 (4.4)	18.3 (2.7)	15.1 (2.9)	8.3 (2.9)
Wilson B-factor [Å <sup>2</sup> ]	12.1	11.6	14.4	21.1	12.5
Matthews Coefficient [Å <sup>3</sup> /Da]	2.4	2.4	2.4	2.4	2.4
<i>C) Refinement</i>					
PHENIX version	1.8.4_1496	1.8.1_1168	1.8.4_1496	1.8.4_1496	1.8.4_1496
Resolution range [Å]	17.5 - 1.24	19.0 - 1.33	20.8 - 1.45	23.6 - 1.89	23.3 - 1.45
Reflections used for R <sub>free</sub>	2000	2000	1983	1652	3657
Reflections used for R <sub>work</sub>	111076	91075	70443	31018	68901
<i>Final R values<sup>[a]</sup></i>					
R <sub>free</sub> [%] <sup>[c]</sup>	16.0	15.0	15.5	19.1	16.5
R <sub>work</sub> [%] <sup>[d]</sup>	14.2	13.3	13.4	15.6	13.8
<i>No. of atoms (non-hydrogen)</i>					
Protein atoms	2903	2926	2830	2838	2875
Water molecules	301	408	343	252	319
Ligand atoms	---	---	---	24	24
RMSD, angle [°]	1.1	1.1	1.0	1.1	1.0
RMSD, bond [Å]	0.007	0.008	0.007	0.009	0.007
<i>Ramachandran plot<sup>[e]</sup></i>					
Most favoured regions [%]	95.3	94.4	94.3	94.6	94.6
Additionally allowed regions [%]	4.4	5.3	5.4	5.1	5.1
Generously allowed regions [%]	0.3	0.3	0.3	0.3	0.3
<i>Mean B-factors [Å<sup>2</sup>]</i>					
Protein atoms	17.1	13.8	19.8	24.1	17.1
Water molecules	32.7	28.2	35.0	32.9	32.0
Ligand atoms	---	---	---	27.9	21.0

<b>Crystal data</b>	<b>Val262Cys-4a</b>	<b>TGT-7</b>	<b>TGT-7a</b>	<b>TGT-7b (1)</b>	<b>TGT-7b (2)</b>
<b>PDB ID</b>	<b>4Q8Q</b>	<b>4Q8T</b>	<b>4Q8U</b>	<b>---</b>	<b>---</b>
<i>A) Data collection and processing</i>					
Collection site	BESSY 14.2	BESSY 14.2	BESSY 14.2	PETRA P14	PETRA P14
No. crystals used	1	1	1	1	1
$\lambda$ [Å]	0.91841	0.91841	0.91841	0.97627	0.97627
Space group	C2	C2	C2	C2	C2
<i>Unit cell parameters</i>					
a [Å]	91.4	90.2	89.8	89.8	90.2
b [Å]	65.0	64.6	64.5	64.7	64.8
c [Å]	70.3	71.0	70.9	70.6	70.3
$\beta$ [°]	96.0	93.2	93.2	96.1	96.2
<i>B) Diffraction data<sup>[a]</sup></i>					
Resolution range [Å]	30 - 1.72 (1.75 - 1.72)	30 - 1.40 (1.42 - 1.40)	30 - 1.31 (1.33 - 1.31)	50 - 1.20 (1.24 - 1.20)	80 - 1.20 (1.24 - 1.20)
Unique reflections	37958 (1991)	80321 (3968)	96126 (4340)	122015 (11647)	123477 (11058)
R(I)sym [%] <sup>[b]</sup>	9.9 (37.4)	5.8 (48.9)	4.8 (47.4)	3.0 (50.5)	2.9 (49.2)
Completeness [%]	87.5 (91.0)	99.9 (98.3)	99.1 (89.9)	97.4 (90.5)	98.3 (94.4)
Redundancy	2.6 (2.5)	4.2 (3.8)	3.3 (2.9)	3.3 (2.8)	3.3 (2.8)
I/ $\sigma$ (I)	9.4 (2.4)	22.7 (2.4)	21.8 (2.0)	20.3 (2.4)	19.8 (2.2)
Wilson B-factor [Å <sup>2</sup> ]	15.4	17.1	15.8	12.7	12.8
Matthews Coefficient [Å <sup>3</sup> /Da]	2.4	2.4	2.4	2.4	2.4
<i>C) Refinement</i>					
PHENIX version	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.4_1496
Resolution range [Å]	29.5 - 1.72	29.8 - 1.40	18.1 - 1.31	44.6 - 1.20	69.9 - 1.20
Reflections used for R <sub>free</sub>	1898	1949	4815	6101	6174
Reflections used for R <sub>work</sub>	36060	78372	91306	115914	117303
<i>Final R values<sup>[a]</sup></i>					
R <sub>free</sub> [%] <sup>[c]</sup>	20.5	15.4	16.2	16.8	15.9
R <sub>work</sub> [%] <sup>[d]</sup>	17.3	13.8	14.0	14.4	13.8
<i>No. of atoms (non-hydrogen)</i>					
Protein atoms	2762	2899	2892	2936	2937
Water molecules	293	311	330	361	354
Ligand atoms	17	24	28	24	27
RMSD, angle [°]	1.0	1.1	1.1	1.1	1.1
RMSD, bond [Å]	0.008	0.007	0.007	0.007	0.007
<i>Ramachandran plot<sup>[e]</sup></i>					
Most favoured regions [%]	94.1	95.3	94.9	95.3	94.9
Additionally allowed regions [%]	5.6	4.4	4.7	4.4	4.8
Generously allowed regions [%]	0.3	0.3	0.3	0.3	0.3
<i>Mean B-factors [Å<sup>2</sup>]</i>					
Protein atoms	19.4	20.7	18.8	16.4	16.4
Water molecules	29.3	33.8	32.8	29.2	29.6
Ligand atoms	16.1	22.6	21.2	14.9 <sup>[f]</sup>	18.6

Crystal data	TGT-7c	TGT-7d	TGT-8	TGT-9	TGT-10
PDB ID	4Q8V	4Q8W	4Q4M	---	---
<i>A) Data collection and processing</i>					
Collection site	BESSY 14.2	PETRA P14	BESSY 14.3	BESSY 14.3	BESSY 14.2
No. crystals used	1	1	1	1	1
$\lambda$ [Å]	0.91841	0.97627	0.89460	0.89460	0.91841
Space group	C2	C2	C2	C2	C2
<i>Unit cell parameters</i>					
a [Å]	91.2	89.8	88.5	88.8	89.8
b [Å]	65.1	64.7	63.9	64.6	64.2
c [Å]	70.8	70.7	70.4	70.5	70.6
$\beta$ [°]	96.3	96.3	93.0	93.5	93.1
<i>B) Diffraction data<sup>[a]</sup></i>					
Resolution range [Å]	50 - 1.40 (1.48 - 1.40)	80 - 1.14 (1.19 - 1.14)	30 - 1.62 (1.65 - 1.62)	30 - 1.84 (1.87 - 1.84)	30 - 1.76 (1.79 - 1.76)
Unique reflections	79794 (12619)	134549 (15089)	48799 (2059)	35510 (1778)	38423 (1860)
R(I)sym [%] <sup>[b]</sup>	6.2 (48.3)	2.8 (30.9)	5.0 (30.0)	15.9 (50.0)	12.9 (46.7)
Completeness [%]	97.8 (95.8)	92.0 (80.0)	98.0 (82.6)	99.9 (100.0)	97.7 (96.3)
Redundancy	2.9 (2.9)	5.7 (5.0)	3.1 (2.0)	3.7 (3.6)	2.3 (2.3)
I/ $\sigma$ (I)	11.0 (2.0)	27.8 (4.7)	21.5 (2.8)	8.3 (2.9)	7.4 (2.2)
Wilson B-factor [Å <sup>2</sup> ]	12.1	11.6	15.6	20.9	21.4
Matthews Coefficient [Å <sup>3</sup> /Da]	2.4	2.4	2.3	2.4	2.4
<i>C) Refinement</i>					
PHENIX version	1.8.4_1496	1.8.4_1496	1.8.4_1496	1.8.1_1168	1.8.1_1168
Resolution range [Å]	21.8 - 1.40	70.3 - 1.14	26.8 - 1.62	28.7 - 1.84	29.7 - 1.76
Reflections used for R <sub>free</sub>	3990	6728	2467	1725	1922
Reflections used for R <sub>work</sub>	75804	127821	46332	32928	36501
<i>Final R values<sup>[a]</sup></i>					
R <sub>free</sub> [%] <sup>[c]</sup>	16.3	15.0	18.9	19.3	19.1
R <sub>work</sub> [%] <sup>[d]</sup>	13.1	13.4	15.7	15.8	16.2
<i>No. of atoms (non-hydrogen)</i>					
Protein atoms	2913	2927	2847	2828	2839
Water molecules	384	376	339	323	273
Ligand atoms	26	27	14	16	14
RMSD, angle [°]	1.1	1.1	1.1	1.1	1.1
RMSD, bond [Å]	0.007	0.007	0.008	0.008	0.008
<i>Ramachandran plot<sup>[e]</sup></i>					
Most favoured regions [%]	95.0	94.9	94.5	95.5	94.6
Additionally allowed regions [%]	4.7	4.8	5.2	4.2	5.1
Generously allowed regions [%]	0.3	0.3	0.3	0.3	0.3
<i>Mean B-factors [Å<sup>2</sup>]</i>					
Protein atoms	14.6	15.1	19.6	21.5	27.0
Water molecules	29.4	28.0	31.9	31.4	36.1
Ligand atoms	13.0	15.5	26.4	36.3	48.3

<sup>a</sup>Values in parentheses are statistics for the highest resolution shell. <sup>b</sup> $R(I)_{sym} = [\sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)] \times 100$ , in which  $\langle I(h) \rangle$  is the mean of the  $I(h)$  observation of reflection  $h$ . <sup>c</sup> $R_{work} = \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|$ . <sup>d</sup> $R_{free}$  was calculated as shown for  $R_{work}$  but on refinement-excluded 2 - 5 % of data. <sup>e</sup>Statistics from PROCHECK [Laskowski *et al.*, 1993]. <sup>f</sup>Averaged value of the two observed conformations.

## References

- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., Zwart, P. H. (2010). *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D* **66**, 213-221.
- Allen, F. H. (2002). The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Cryst. B* **58**, 380-388.
- Ashkenazi, S., Levy, I., Kazaronovski, V., Samra, Z. (2003). Growing antimicrobial resistance of *Shigella* isolates. *J. Antimicrob. Chemother.* **51**, 427-429.
- Baker, B. M., Murphy, K. P. (1996). Evaluation of Linked Protonation Effects in Protein Binding Reactions Using Isothermal Titration Calorimetry. *Biophys. J.* **71**, 2049-2055.
- Baranauskienė, L., Petrikaitė, V., Matulienė, J., Matulis, D. (2009). Titration Calorimetry Standards and the Precision of Isothermal Titration Calorimetry Data. *Int. J. Mol. Sci.* **10**, 2752-2762.
- Barandun, L. J., Immekus, F., Kohler, P. C., Ritschel, T., Heine, A., Orlando, P., Klebe, G., Diederich, F. (2013). High-affinity inhibitors of *Zymomonas mobilis* tRNA-guanine transglycosylase through convergent optimization. *Acta Cryst. D* **69**, 1798-1807.
- Barandun, L. J., Immekus, F., Kohler, P. C., Tonazzi, S., Wagner, B., Wendelspiess, S., Ritschel, T., Heine, A., Kansy, M., Klebe, G., Diederich, F. (2012). 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.
- Barken, F. M., Gasteiger, E. L. (1980). Excitability of a penicillin-induced cortical epileptic focus. *Exp. Neurol.* **70**, 539-547.
- Bashford, D. (1997). An Object-Oriented Programming Suite for Electrostatic Effects in Biological Molecules. An Experience Report on the MEAD Project. In *Scientific Computing in Object-Oriented Parallel Environments*; Ishikawa, Y., Oldehoeft, R., Reynders, J., Tholburn, M., Eds.; Springer: Berlin / Heidelberg.; **1343**: 233-240.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., Leslie, A. G. W. (2011). *iMOSFLM*: a new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Cryst. D* **67**, 271-281.
- Baum, B., Mohamed, M., Zayed, M., Gerlach, C., Heine, A., Hangauer, D., Klebe, G. (2009). More than a Simple Lipophilic Contact: A Detailed Thermodynamic Analysis of Nonbasic Residues in the S1 Pocket of Thrombin. *J. Mol. Biol.* **390**, 56-69.
- Baum, B., Muley, L., Heine, A., Smolinski, M., Hangauer, D., Klebe, G. (2009). Think Twice: Understanding the High Potency of Bis(phenyl)methane Inhibitors of Thrombin. *J. Mol. Biol.* **391**, 552-564.



- Baum, B., Muley, L., Smolinski, M., Heine, A., Hangauer, D., Klebe, G. (2010). Non-Additivity of Functional Group Contributions in Protein-Ligand Binding: A Comprehensive Study by Crystallography and Isothermal Titration Calorimetry. *J. Mol. Biol.* **397**, 1042-1054.
- Bensadoun, A., Weinstein, D. (1976). Assay of Proteins in the Presence of Interfering Materials. *Anal. Biochem.* **70**, 241–250.
- Berrien, J.-F., Ourévitch, M., Morgant, G., Ghermani, N. E., Crousse, B., Bonnet-Delpon, D. (2007). A crystalline H-bond cluster of hexafluoroisopropanol (HFIP) and piperidine - Structure determination by X ray diffraction. *J. Fluor. Chem.* **128**, 839-843.
- Biela, A., Khayat, M., Tan, H., Kong, J., Heine, A., Hangauer, D., Klebe, G. (2012). Impact of Ligand and Protein Desolvation on Ligand Binding to the S1 Pocket of Thrombin. *J. Mol. Biol.* **418**, 350-366.
- Biela, A., Nasief, N. N., Betz, M., Heine, A., Hangauer, D., Klebe, G. (2013). Dissecting the Hydrophobic Effect on the Molecular Level: The Role of Water, Enthalpy, and Entropy in Ligand Binding to Thermolysin. *Angew. Chem. Int. Ed.* **52**, 1822–1828.
- Biela, A., Sielaff, F., Terwesten, F., Heine, A., Steinmetzer, T., Klebe, G. (2012). Ligand Binding Stepwise Disrupts Water Network in Thrombin: Enthalpic and Entropic Changes Reveal Classical Hydrophobic Effect. *J. Med. Chem.* **55**, 6094-6110.
- Biela, I., Tidten-Luksch, N., Immekus, F., Glinca, S., Nguyen, T. X. P., Gerber, H.-D., Heine, A., Klebe, G., Reuter, K. (2013). Investigation of Specificity Determinants in Bacterial tRNA-Guanine Transglycosylase Reveals Queuine, the Substrate of Its Eucaryotic Counterpart, as Inhibitor. *PLoS One* **8**, e64240.
- Bogan, A. A., Thorn, K. S. (1998). Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **280**, 1-9.
- Bradford, M. M. (1976). Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Anal. Biochem.* **72**, 248-254.
- Brenk, R., Meyer, E. A., Reuter, K., Stubbs, M. T., Garcia, G. A., Diederich, F., Klebe, G. (2004). Crystallographic Study of Inhibitors of tRNA-guanine Transglycosylase Suggests a New Structure-Based Pharmacophore for Virtual Screening. *J. Mol. Biol.* **338**, 55-75.
- Brenk, R., Stubbs, M. T., Heine, A., Reuter, K., 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.
- Brown, A. (2009). Analysis of Cooperativity by Isothermal Titration Calorimetry. *Int. J. Mol. Sci.* **10**, 3457–3477.
- Burnouf, D., Ennifar, E., Guedich, S., Puffer, B., Hoffmann, G., Bec, G., Disdier, F., Baltzinger, M., Dumas, P. (2012). kinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *J. Am. Chem. Soc.* **134**, 559-565.

- Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., Woods, R. J. (2005). The Amber Biomolecular Simulation Programs. J. Comput. Chem. *26*, 1668-1688.
- Chaires, J. B. (2008). Calorimetry and Thermodynamics in Drug Design. Annu. Rev. Biophys. *37*, 135–151.
- Chen, Y. C., Brooks, A. F., Goodenough-Lashua, D. M., Kittendorf, J. D., Showalter, H. D., Garcia, G. A. (2011). Evolution of eukaryal tRNA-guanine transglycosylase: insight gained from the heterocyclic substrate recognition by the wild-type and mutant human and *Escherichia coli* tRNA-guanine transglycosylases. Nucleic Acids Res. *39*, 2834-2844.
- Christensen, J. J., Hansen, L. D., Izatt, R. M. (1976). Handbook of Proton Ionization Heats and Related Thermodynamic Quantities; Wiley-Interscience: New York.
- Coan, K. E. D., Shoichet, B. K. (2008). Stoichiometry and Physical Chemistry of Promiscuous Aggregate-Based Inhibitors. J. Am. Chem. Soc. *130*, 9606-9612.
- Congreve, M., Carr, R., Murray, C., Jhoti, H. (2003). A 'Rule of Three' for fragment-based lead discovery? Drug Discov. Today *8*, 876-877.
- Connolly, M. L. (1983). Analytical Molecular Surface Calculation. J. Appl. Crystallogr. *16*, 548-558.
- Conradi, R. A., Hilgers, A. R., Ho, N. F., Burton, P. S. (1991). The Influence of Peptide Structure on Transport Across Caco-2 Cells. Pharm. Res. *8*, 1453-1460.
- Curnow, A. W., Garcia, G. A. (1995). tRNA-guanine transglycosylase from *Escherichia coli*. Minimal tRNA structure and sequence requirements for recognition. J. Biol. Chem. *270*, 17264-17267.
- Czodrowski, P., Dramburg, I., Sotriffer, C. A., Klebe, G. (2006). Development, validation, and application of adapted PEOE charges to estimate pK<sub>a</sub> values of functional groups in protein-ligand complexes. Proteins *65*, 424-437.
- Czodrowski, P., Sotriffer, C. A., Klebe, G. (2007). Protonation Changes upon Ligand Binding to Trypsin and Thrombin: Structural Interpretation Based on pK<sub>a</sub> Calculations and ITC Experiments. J. Mol. Biol. *367*, 1347-1356.
- Dalvit, C., Fasolini, M., Flocco, M., Knapp, S., Pevarello, P., Veronesi, M. (2002). NMR-Based Screening with Competition Water-Ligand Observed via Gradient Spectroscopy Experiments: Detection of High-Affinity Ligands. J. Med. Chem. *45*, 2610-2614.
- Darden, T., York, D., Pedersen, L. (1993). Particle mesh Ewald - an N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089-10092.
- Daviter, T., Fronzes, R. (2013). Protein Sample Characterization. Methods Mol. Biol. *1008*, 35-62.
- Dixon, M. (1953). The Determination of Enzyme Inhibitor Constants. Biochem. J. *55*, 170-171.
- Doak, A. K., Wille, H., Prusiner, S. B., Shoichet, B. K. (2010). Colloid Formation by Drugs in Simulated Intestinal Fluid. J. Med. Chem. *53*, 4259-4265.

- Dorman, C. J., Porter, M. E. (1998). The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. Mol. Microbiol. *29*, 677-684.
- Dullweber, F., Stubbs, M. T., Musil, D., Stürzebecher, J., Klebe, G. (2001). Factorising Ligand Affinity: A Combined Thermodynamic and Crystallographic Study of Trypsin and Thrombin Inhibition. J. Mol. Biol. *313*, 593-614.
- DuPont, H. L., Levine, M. M., Hornick, R. B., Formal, S. B. (1989). Inoculum Size in Shigellosis and Implications for Expected Mode of Transmission. J. Infect. Dis. *159*, 1126-1128.
- Durand, J. M., Björk, G. R., Kuwae, A., Yoshikawa, M., Sasakawa, C. (1997). The Modified Nucleoside 2-Methylthio-N<sup>6</sup>-Isopentenyladenosine in tRNA of *Shigella flexneri* Is Required for Expression of Virulence Genes. J. Bacteriol. *179*, 5777-5782.
- Durand, J. M., Okada, N., Tobe, T., Watarai, M., Fukuda, I., Suzuki, T., Nakata, N., Komatsu, K., Yoshikawa, M., 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.
- Durand, J. M. B., Dagberg, B., Uhlin, B. E., 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.
- El Tayar, N., Tsai, R.-S., Testa, B., Carrupt, P.-A., Leo, A. (1991). Partitioning of solutes in different solvent systems: The contribution of hydrogen-bonding capacity and polarity. J. Pharm. Sci. *80*, 590-598.
- Eldridge, M. D., Murray, C. W., Auton, T. R., Paolini, G. V., Mee, R. P. (1997). Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J. Comput. Aided Mol. Des. *11*, 425-445.
- Emsley, P., Cowtan, K. (2004). *Coot*: model-building tools for molecular graphics. Acta Cryst. D *60*, 2126-2132.
- Erickson, J. W., Burt, S. K. (1996). Structural Mechanisms of HIV Drug Resistance. Annu. Rev. Pharmacol. Toxicol. *36*, 545-571.
- Escobar-Páramo, P., Clermont, O., Blanc-Potard, A. B., Bui, H., Le Bouguéneq, C., Denamur, E. (2004). A Specific Genetic Background Is Required for Acquisition and Expression of Virulence Factors in *Escherichia coli*. Mol. Biol. Evol. *21*, 1085-1094.
- Fejzo, J., Lepre, C. A., Peng, J. W., Bemis, G. W., Ajay, Murcko, M. A., Moore, J. M. (1999). The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. Chem. Biol. *6*, 755-769.
- Feng, B. Y., Shelat, A., Doman, T. N., Guy, R. K., Shoichet, B. K. (2005). High-throughput assays for promiscuous inhibitors. Nat. Chem. Biol. *1*, 146-148.
- Feng, B. Y., Shoichet, B. K. (2006). A detergent-based assay for the detection of promiscuous inhibitors. Nat. Protoc. *1*, 550-553.

- Feng, B. Y., Simeonov, A., Jadhav, A., Babaoglu, K., Inglese, J., Shoichet, B. K., Austin, C. P. (2007). A High-Throughput screen for Aggregation-Based Inhibition in a Large Compound Library. J. Med. Chem. *50*, 2385-2390.
- Fernandez, M. I., Sansonetti, P. J. (2003). *Shigella* interaction with intestinal epithelial cells determines the innate immune response in shigellosis. Int. J. Med. Microbiol. *293*, 55-67.
- Frey, B., McCloskey, J., Kersten, W., Kersten, H. (1988). New Function of Vitamin B<sub>12</sub>: Cobamide-Dependent Reduction of Epoxyqueuosine to Queuosine in tRNAs of *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. *170*, 2078-2082.
- Freyer, M. W., Lewis, E. A. (2008). Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. Methods Cell Biol. *84*, 79-113.
- Frølund, B., Jørgensen, A. T., Tagmose, L., Stensbøl, T. B., Vestergaard, H. T., Engblom, C., Kristiansen, U., Sanchez, C., Krogsgaard-Larsen, P., Liljefors, T. (2002). Novel Class of Potent 4-Arylalkyl Substituted 3-Isoxazolol GABA<sub>A</sub> Antagonists: Synthesis, Pharmacology, and Molecular Modeling. J. Med. Chem. *45*, 2454-2468.
- Fukada, H., Takahashi, K. (1998). Enthalpy and Heat Capacity Changes for the Proton Dissociation of Various Buffer Components in 0.1 M Potassium Chloride. Proteins *33*, 159-166.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocols Handbook; Walker, J. M., Ed.; Humana Press: Totowa, N. J.; 571-607.
- Gill, S. C., von Hippel, P. H. (1989). Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. Anal. Biochem. *182*, 319-326.
- Goa, J. (1953). A Micro Biuret Method for Protein Determination; Determination of Total Protein in Cerebrospinal Fluid. Scand. J. Clin. Lab. Invest. *5*, 218-222.
- Grädler, U., Gerber, H.-D., Goodenough-Lashua, D. M., Garcia, G. A., Ficner, R., Reuter, K., Stubbs, M. T., 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.
- Hall, H. K. (1957). Correlation of the Base Strengths of Amines. J. Am. Chem. Soc. *79*, 5441-5444.
- Henry, C. M. (2001). Structure-Based Drug Design. Chem. Eng. News *79*, 69-74.
- Holdgate, G. A. (2001). Making Cool Drugs Hot: Isothermal Titration Calorimetry as a Tool to Study Binding Energetics. BioTechniques *31*, 164-184.
- Hoops, G. C., Townsend, L. B., Garcia, G. A. (1995). tRNA-Guanine Transglycosylase from *Escherichia coli*: Structure-Activity Studies Investigating the Role of the Aminomethyl Substituent of the Heterocyclic Substrate PreQ1. Biochemistry *34*, 15381-15387.
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C. (2006). Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. Proteins *65*, 712-725.

- Hörtner, S. R., Ritschel, T., Stengl, B., Kramer, C., Schweizer, W. B., Wagner, B., Kansy, M., Klebe, G., 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.
- Immekus, F. P. P. (2013) *lin*-Benzopurines as Inhibitor of tRNA-Guanine Transglycosylase: Perturbance of Homodimer Formation, Import of Water Clusters and Determinants of Crystallographical Disorder. Dissertation, Philipps-Universität Marburg.
- Iwata-Reuyl, D. (2003). Biosynthesis of the 7-deazaguanosine hypermodified nucleosides of transfer RNA. Bioorg. Chem. *31*, 24-43.
- Jakalian, A., Bush, B. L., Jack, D. B., Bayly, C. I. (2000). Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. J. Comput. Chem. *21*, 132-146.
- Jakalian, A., Jack, D. B., Bayly, C. I. (2002). Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. J. Comput. Chem. *23*, 1623-1641.
- Jakobi, S. (2013). Nichtkompetitive Inhibition der tRNA-Guanin Transglycosylase durch Störung der essentiellen Protein-Protein-Interaktion. Dissertation, Philipps-Universität Marburg.
- Jakobi, S., Nguyen, T. X., Debaene, F., Metz, A., Sanglier-Cianferani, S., Reuter, K., Klebe, G. (2014). Hot-spot analysis to dissect the functional protein-protein interface of a tRNA-modifying enzyme. Proteins, in press.
- Jennison, A. V., Verma, N. K. (2004). *Shigella flexneri* infection: pathogenesis and vaccine development. FEMS Microbiol. Rev. *28*, 43-58.
- Jones, G., Willett, P., Glen, R. C., Leach, A. R., Taylor, R. (1997). Development and Validation of a Genetic Algorithm for Flexible Docking. J. Mol. Biol. *267*, 727-748.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926-935.
- Jorgensen, W. L., Pranata, J. (1990). Importance of Secondary Interactions in Triply Hydrogen-Bonded Complexes - Guanine-Cytosine vs Uracil-2,6-Diaminopyridine. J. Am. Chem. Soc. *112*, 2008-2010.
- Kabsch, W. (2010). XDS. Acta Cryst. D *66*, 125-132.
- King, N. M., Prabu-Jeyabalan, M., Bandaranayake, R. M., Nalam, M. N., Nalivaika, E. A., Özen, A., Haliloğlu, T., Yilmaz, N. K., Schiffer, C. A. (2012). Extreme Entropy-Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease. ACS Chem. Biol. *7*, 1536-1546.
- Klebe, G. (2009). Wirkstoffdesign - Entwurf und Wirkung von Arzneistoffen. 2. Auflage; Spektrum Akademischer Verlag: Heidelberg.
- Kleywegt, G. J., Zou J. Y., Kjeldgaard M. & Jones T. A. (2001). Around O. In International Tables for Crystallography; M. G. Rossmann, Arnold, E., Eds.; Kluwer Academic Publishers: Dordrecht; Vol. F, 353-356.

- Koeppe, R. E. II, Stroud, R. M. (1976). Mechanism of hydrolysis by serine proteases: direct determination of the  $pK_a$ 's of aspartyl-102 and aspartyl-194 in bovine trypsin using difference infrared spectroscopy. Biochemistry 15, 3450–3458.
- Kohler, P. C., Ritschel, T., Schweizer, W. B., Klebe, G., Diederich, F. (2009). High-Affinity Inhibitors of tRNA-Guanine Transglycosylase Replacing the Function of a Structural Water Cluster. Chem. Eur. J. 15, 10809-10817.
- Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K., Levine, M. M. (1999). Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. Bull. World Health Organ. 77, 651-666.
- Kuntz, I. D., Chen, K., Sharp, K. A., Kollman, P. A. (1999). The maximal affinity of ligands. Proc. Natl. Acad. Sci. USA 96, 9997-10002.
- Kunz, W., Henle, J., Ninham, B. W. (2004). 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. Curr. Opin. Colloid. In. 9, 19-37.
- Kupke, D. W., Dorrier, T. E. (1978). Protein concentration measurements: the dry weight. Methods Enzymol. 48, 155–162.
- Ladbury, J. E., Klebe, G., Freire, E. (2010). Adding calorimetric data to decision making in lead discovery: a hot tip. Nat. Rev. Drug Discov. 9, 23-27.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283-291.
- Lee, B. W., Van Lanen, S. G., Iwata-Reuyl, D. (2007). Mechanistic Studies of *Bacillus subtilis* QueF, the Nitrile Oxidoreductase Involved in Queuosine Biosynthesis. Biochemistry 46, 12844-12854.
- Levine, O. S., Levine, M. M. (1991). Houseflies (*Musca domestica*) as Mechanical Vectors of Shigellosis. Rev. Infect. Dis. 13, 688-696.
- Liang, Y. (2008). Applications of isothermal titration calorimetry in protein science. Acta Biochim. Biophys. Sin. 40, 565–576.
- Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 46, 3-26.
- Luque, I., Todd, M. J., Gómez, J., Semo, N., Freire, E. (1998). Molecular Basis of Resistance to HIV-1 Protease Inhibition: A Plausible Hypothesis. Biochemistry 37, 5791-5797.
- Makara, G. M. (2007). On Sampling of Fragment Space. J. Med. Chem. 50, 3214-3221.
- Mandal, J., Ganesh, V., Emelda, J., Mahadevan, S., Parija, S. C. (2012). The Recent Trends of Shigellosis: A JIPMER Perspective. J. Clin. Diagn. Res. 6, 1474-1477.

- Martin, S. F., Clements, J. H. (2013). Correlating Structure and Energetics in Protein-Ligand Interactions: Paradigms and Paradoxes. Annu. Rev. Biochem. *82*, 267-293.
- McCarty, R. M., Somogyi, A., Bandarian, V. (2009). *Escherichia coli* QueD Is a 6-Carboxy-5,6,7,8-tetrahydropterin Synthase. Biochemistry *48*, 2301-2303.
- McCarty, R. M., Somogyi, A., Lin, G., Jacobsen, N. E., Bandarian, V. (2009). The Deazapurine Biosynthetic Pathway Revealed: In Vitro Enzymatic Synthesis of PreQ<sub>0</sub> from Guanosine 5'-Triphosphate in Four Steps. Biochemistry *48*, 3847-3852.
- McCoy, A. J. (2007). Solving structures of protein complexes by molecular replacement with *Phaser*. Acta Cryst. D *63*, 32-41.
- McGovern, S. L., Caselli, E., Grigorieff, N., Shoichet, B. K. (2002). A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening. J. Med. Chem. *45*, 1712-1722.
- McGovern, S. L., Shoichet, B. K. (2003). Kinase Inhibitors: Not Just for Kinases Anymore. J. Med. Chem. *46*, 1478-1483.
- Meyer, E. A., Donati, N., Guillot, M., Schweizer, W. B., Diederich, F., Stengl, B., Brenk, R., Reuter, K., 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.
- Michaelis, L., Menten, M. L. (1913). Die Kinetik der Invertinwirkung. Biochem. Z. *49*, 333-369.
- Mizoue, L. S., Tellinghuisen, J. (2004). The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. Anal. Biochem. *326*, 125-127.
- Murray, T. J., Zimmerman, S. C. (1992). New Triply Hydrogen-Bonded Complexes with Highly Variable Stabilities. J. Am. Chem. Soc. *114*, 4010-4011.
- Myszka, D. G., Abdiche, Y. N., Arisaka, F., Byron, O., Eisenstein, E., Hensley, P., Thomson, J. A., Lombardo, C. R., Schwarz, F., Stafford, W., Doyle, M. L. (2003). The ABRF-MIRG'02 Study: Assembly State, Thermodynamic, and Kinetic Analysis of an Enzyme/Inhibitor Interaction. J. Biomol. Tech. *14*, 247-269.
- Navratilova, I., Hopkins, A. L. (2010). Fragment Screening by Surface Plasmon Resonance. ACS Med. Chem. Lett. *1*, 44-48.
- Neeb, M., Betz, M., Heine, A., Barandun, L. J., Hohn, C., Diederich, F., Klebe, G. (2014). Beyond Affinity: Enthalpy-Entropy Factorization Unravels Complexity of a Flat Structure-Activity Relationship for Inhibition of a tRNA-Modifying Enzyme. J. Med. Chem. *57*, 5566-5578.
- Neeb, M., Czodrowski, P., Heine, A., Barandun, L. J., Hohn, C., Diederich, F., Klebe, G. (2014). Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and pK<sub>a</sub> Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme. J. Med. Chem. *57*, 5554-5565.
- Neudert, G., Klebe, G. (2011). fconv: format conversion, manipulation and feature computation of molecular data. Bioinformatics *27*, 1021-1022.

- Nhieu, G. T. V., Bourdet-Sicard, R., Duménil, G., Blocker, A., Sansonetti, P. J. (2000). Bacterial signals and cell responses during *Shigella* entry into epithelial cells. Cell. Microbiol. *2*, 187-193.
- Niyogi, S. K. (2005). Shigellosis. J. Microbiol. *43*, 133-143.
- Okada, N., Nishimura, S. (1979). Isolation and Characterization of a Guanine Insertion Enzyme, a Specific tRNA Transglycosylase, from *Escherichia coli*. J. Biol. Chem. *254*, 3061-3066.
- Otwinowski, Z., Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.
- Painter, J., Merritt, E. A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Cryst. D *62*, 439-450.
- Pang, Y. P. (2001). Successful Molecular Dynamics Simulation of Two Zinc Complexes Bridged by a Hydroxide in Phosphotriesterase Using the Cationic Dummy Atom Method. Proteins *45*, 183-189.
- Pang, Y. P., Xu, K., El Yazal, J., Prendergast, F. G. (2000). Successful molecular dynamics simulation of the zinc-bound farnesyltransferase using the cationic dummy atom approach. Protein Sci. *9*, 1857-1865.
- Perspicace, S., Banner, D., Benz, J., Müller, F., Schlatter, D., Huber, W. (2009). Fragment-Based Screening Using Surface Plasmon Resonance Technology. J. Biomol. Screen. *14*, 337-349.
- Phillips, G., El Yacoubi, B., Lyons, B., Alvarez, S., Iwata-Reuyl, D., de Crécy-Lagard, V. (2008). Biosynthesis of 7-Deazaguanosine-Modified tRNA Nucleosides: a New Role for GTP Cyclohydrolase I. J. Bacteriol. *190*, 7876-7884.
- Raffa, R. B., Stagliano, G. W., Spencer, S. D. (2004). Protonation effect on drug affinity. Eur. J. Pharmacol. *483*, 323-324.
- Rauh, D., Klebe, G., Stubbs, M. T. (2004). Understanding Protein-Ligand Interactions: The Price of Protein Flexibility. J. Mol. Biol. *335*, 1325-1341.
- Rauh, D., Klebe, G., Stürzebecher, J., Stubbs, M. T. (2003). ZZ Made EZ: Influence of Inhibitor Configuration on Enzyme Selectivity. J. Mol. Biol. *330*, 761-770.
- Rauh, D., Reyda, S., Klebe, G., Stubbs, M. T. (2002). Trypsin Mutants for Structure-Based Drug Design: Expression, Refolding and Crystallisation. Biol. Chem. *383*, 1309-1314.
- Reader, J. S., Metzgar, D., Schimmel, P., de Crécy-Lagard, V. (2004). Identification of Four Genes Necessary for Biosynthesis of the Modified Nucleoside Queuosine. J. Biol. Chem. *279*, 6280-6285.
- Reuter, K., 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.
- Ritschel, T. (2009). TGT a Drug Target to Study pKa Shifts, Residual Solvation & Protein - Protein Interface Formation. Dissertation, Philipps-Universität Marburg.



- Ritschel, T., Atmanene, C., Reuter, K., Van Dorsselaer, A., Sanglier-Cianferani, S., 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.
- Ritschel, T., Hoertner, S., Heine, A., Diederich, F., 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.
- Ritschel, T., Kohler, P. C., Neudert, G., Heine, A., Diederich, F., 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.
- Robinson, N. C., Tye, R. W., Neurath, H., Walsh, K. A. (1971). Isolation of Trypsins by Affinity Chromatography. Biochemistry *10*, 2743-2747.
- Romier, C., Ficner, R., Reuter, K., Suck, D. (1996). Purification, Crystallization, and Preliminary X-Ray Diffraction Studies of tRNA-Guanine Transglycosylase from *Zymomonas mobilis*. Proteins *24*, 516-519.
- Romier, C., Reuter, K., Suck, D., Ficner, R. (1996). Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. EMBO J. *15*, 2850-2857.
- Romier, C., Reuter, K., Suck, D., Ficner, R. (1996). Mutagenesis and Crystallographic Studies of *Zymomonas mobilis* tRNA-Guanine Transglycosylase Reveal Aspartate 102 as the Active Site Nucleophile. Biochemistry *35*, 15734-15739.
- Ryckaert, J. P., Ciccotti, G., Berendsen, H. J. C. (1977). Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of n-Alkanes. J. Comput. Phys. *23*, 327-341.
- Sack, R. B., Rahman, M., Yunus, M., Khan, E. H. (1997). Antimicrobial Resistance in Organisms Causing Diarrheal Disease. Clin. Infect. Dis. *24 Suppl. 1*, S102-105.
- Sambrook, J. F., Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual (3<sup>rd</sup> Ed.); Cold Spring Harbor Laboratory Press: New York.
- Sansonetti, P. J. (2001). Microbes and Microbial Toxins: Paradigms for Microbial-Mucosal Interactions III. Shigellosis: from symptoms to molecular pathogenesis. Am. J. Physiol. Gastrointest. Liver Physiol. *280*, G319-323.
- Sansonetti, P. J. (2001). Rupture, invasion and inflammatory destruction of the intestinal barrier by *Shigella*, making sense of prokaryote-eukaryote cross-talks. FEMS Microbiol. Rev. *25*, 3-14.
- Sansonetti, P. J., Bergounioux, J. (2008). Shigellosis. In Harrison's Principles of Internal Medicine (17<sup>th</sup> Ed.); Fauci, A. S., Braunwald, E., Kasper, D. L., Hauser, S. L., Longo, D. L., Jameson, J. L., Loscalzo, J., Eds.; McGraw-Hill: New York; 962-964.
- Schaner, L. S. (1962). Passage of Drugs Across Body Membranes. Pharmacol. Rev. *14*, 501-530.

- Schmitt, P., Poiger, T., Simon, R., Freitag, D., Kettrup, A., Garrison, A. W. (1997). Simultaneous Determination of Ionization Constants and Isoelectric Points of 12 Hydroxy-s-Triazines by Capillary Zone Electrophoresis and Capillary Isoelectric Focusing. *Anal. Chem.* **69**, 2559-2566.
- Scott, D. E., Coyne, A. G., Hudson, S. A., Abell, C. (2012). Fragment-Based Approaches in Drug Discovery and Chemical Biology. *Biochemistry* **51**, 4990-5003.
- Seidler, J., McGovern, S. L., Doman, T. N., Shoichet, B. K. (2003). Identification and Prediction of Promiscuous Aggregating Inhibitors among Known Drugs. *J. Med. Chem.* **46**, 4477-4486.
- Sheldrick, G. M. (2008). A short history of SHELX. *Acta Cryst. A* **64**, 112-122.
- Snyder, P. W., Mecinović, J., Moustakas, D. T., Thomas, S. W., 3rd, Harder, M., Mack, E. T., Lockett, M. R., Héroux, A., Sherman, W., Whitesides, G. M. (2011). Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proc. Natl. Acad. Sci. USA* **108**, 17889-17894.
- Spencer, J. N., Gleim, J. E., Blevins, C. H., Garrett, R. C., Mayer, F. J. (1979). Enthalpies of Solution and Transfer Enthalpies - Analysis of the Pure Base Calorimetric Method for the Determination of Hydrogen-Bond Enthalpies. *J. Phys. Chem.* **83**, 1249-1255.
- Stengl, B., Meyer, E. A., Heine, A., Brenk, R., Diederich, F., 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.
- Steuber, H., Czodrowski, P., Sotriffer, C. A., Klebe, G. (2007). Tracing Changes in Protonation: A Prerequisite to Factorize Thermodynamic Data of Inhibitor Binding to Aldose Reductase. *J. Mol. Biol.* **373**, 1305-1320.
- Tellinghuisen, J. (2003). A study of statistical error in isothermal titration calorimetry. *Anal. Biochem.* **321**, 79-88.
- Tellinghuisen, J. (2004). Statistical error in isothermal titration calorimetry. *Methods Enzymol.* **383**, 245-282.
- Tellinghuisen, J. (2005). Optimizing Experimental Parameters in Isothermal Titration Calorimetry: Variable Volume Procedures. *J. Phys. Chem. B* **109**, 20027-20035.
- Tellinghuisen, J. (2007). Calibration in isothermal titration calorimetry: Heat and cell volume from heat of dilution of NaCl(aq). *Anal. Biochem.* **360**, 47-55.
- Tellinghuisen, J., Chodera, J. D. (2011). Systematic errors in isothermal titration calorimetry: Concentrations and baselines. *Anal. Biochem.* **414**, 297-299.
- Testa, B., Carrupt, P. A., Gaillard, P., Billois, F., Weber, P. (1996). Lipophilicity in Molecular Modeling. *Pharm. Res.* **13**, 335-343.
- Thorn, A., Sheldrick, G. M. (2011). ANODE: anomalous and heavy-atom density calculation. *J. Appl. Crystallogr.* **44**, 1285-1287.

- Thorne, N., Auld, D. S., Inglese, J. (2010). Apparent activity in high-throughput screening: origins of compound-dependent assay interference. Curr. Opin. Chem. Biol. *14*, 315-324.
- Tidten, N., Stengl, B., Heine, A., Garcia, G. A., Klebe, G., Reuter, K. (2007). Glutamate *versus* Glutamine Exchange Swaps Substrate Selectivity in tRNA-Guanine Transglycosylase: Insight into the Regulation of Substrate Selectivity by Kinetic and Crystallographic Studies. J. Mol. Biol. *374*, 764-776.
- Tsamaloukas, A. D., Keller, S., Heerklotz, H. (2007). Uptake and release protocol for assessing membrane binding and permeation by way of isothermal titration calorimetry. Nat. Protoc. *2*, 695-704.
- Turnbull, W. B., Daranas, A. H. (2003). On the Value of  $c$ : Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? J. Am. Chem. Soc. *125*, 14859–14866.
- Van Lanen, S. G., Kinzie, S. D., Matthieu, S., Link, T., Culp, J., Iwata-Reuyl, D. (2003). tRNA Modification by S-Adenosylmethionine:tRNA Ribosyltransferase-Isomerase. Assay Development and Characterization of the Recombinant Enzyme. J. Biol. Chem. *278*, 10491-10499.
- Velazquez-Campoy, A., Freire, E. (2006). Isothermal titration calorimetry to determine association constants for high-affinity ligands. Nat. Protoc. *1*, 186-191.
- Velazquez-Campoy, A., Kiso, Y., Freire, E. (2001). The Binding Energetics of First- and Second-Generation HIV-1 Protease Inhibitors: Implications for Drug Design. Arch. Biochem. Biophys. *390*, 169-175.
- Velazquez-Campoy, A., Todd, M. J., Freire, E. (2000). HIV-1 Protease Inhibitors: Enthalpic versus Entropic Optimization of the Binding Affinity. Biochemistry. *39*, 2201-2207.
- Verdonk, M. L., Rees, D. C. (2008). Group Efficiency: A Guideline for Hits-to-Leads Chemistry. ChemMedChem *3*, 1179-1180.
- Wadsö, I. (2000). Needs for standards in isothermal microcalorimetry. Thermochim. Acta *347*, 73-77.
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., Case, D. A. (2004). Development and Testing of a General Amber Force Field. J. Comput. Chem. *25*, 1157-1174.
- Wang, W. (2000). Lyophilization and development of solid protein pharmaceuticals. Int. J. Pharm. *203*, 1–60.
- World Health Organization (2001). Antimicrobial resistance in shigellosis, cholera and campylobacteriosis. WHO/CDS/CSR/DRS/2001.8. Department of Communicable Disease Surveillance and Response, Geneva.
- World Health Organization (2005). Guidelines for the control of shigellosis, including the epidemics which were caused by *Shigella dysenteriae* type 1. Department of Child and Adolescent Health and Development, Geneva.
- World Health Organization (2009). Diarrhoeal Diseases. Retrieved from [http://www.who.int/vaccine\\_research/diseases/diarrhoeal/en/index6.html](http://www.who.int/vaccine_research/diseases/diarrhoeal/en/index6.html).

- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Cryst. D* **67**, 235-242.
- Wiseman, T., Williston, S., Brandts, J. F., Lin, L. N. (1989). Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. *Anal. Biochem.* **179**, 131–137.
- Word, J. M., Lovell, S. C., Richardson, J. S., Richardson, D. C. (1999). Asparagine and Glutamine: Using Hydrogen Atom Contacts in the Choice of Side-Chain Amide Orientation. *J. Mol. Biol.* **285**, 1735-1747.
- Xie, W., Liu, X., Huang, R. H. (2003). Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. *Nat. Struct. Biol.* **10**, 781-788.
- Yonemura, H., Imamura, T., Soejima, K., Nakahara, Y., Morikawa, W., Ushio, Y., Kamachi, Y., Nakatake, H., Sugawara, K., Nakagaki, T., Nozaki, C. (2004). Preparation of Recombinant  $\alpha$ -Thrombin: High-Level Expression of Recombinant Human Prethrombin-2 and Its Activation by Recombinant Ecarin. *J. Biochem.* **135**, 577–582.
- Zhang, Y. L., Zhang, Z. Y. (1998). Low-Affinity Binding Determined by Titration Calorimetry Using a High-Affinity Coupling Ligand: A Thermodynamic Study of Ligand Binding to Protein Tyrosine Phosphatase 1B. *Anal. Biochem.* **261**, 139-148.

## Acknowledgement

Die vorliegende Arbeit wäre nicht ohne die Hilfe zahlreicher Menschen entstanden, die mich während meiner Zeit in der AG Klebe unterstützt, ermutigt und mit neuen Ideen bereichert haben. Nachfolgend möchte ich diesen danken:

*Prof. Dr. Gerhard Klebe* danke ich herzlich für die Aufnahme in seine Arbeitsgruppe, das in mich gesetzte Vertrauen bei der Bearbeitung meines Dissertationsthemas, sowie sein stets offenes Ohr und die daraus hervorgegangenen konstruktiven Ansätze, wenn ein Experiment nicht wie erwartet verlief oder das gewünschte Ergebnis zeigte. Darüber hinaus bedanke ich mich für seine Korrekturen der Paperdrafts und des Manuskripts der vorliegenden Dissertationsschrift, sowie die finanzielle Unterstützung zahlreicher Tagungsreisen.

*Prof. Dr. Klaus Reuter* gebührt mein Dank für die umfassenden Tipps im Bezug auf das praktische Arbeiten im S1- und Isotopenlabor, die konstruktiven Verbesserungsvorschläge dieses Manuskripts sowie das Erstellen des Zweitgutachtens.

*Dr. Luzi Jakob Barandun, Christoph Hohn und Prof. Dr. François Diederich* (ETH Zürich) danke ich für die erfolgreiche Zusammenarbeit bei dem Design neuer TGT-Inhibitoren, der unermüdlichen Synthesebereitschaft und den informativen Video-Konferenzen. Weiterhin gebührt ihnen mein Dank für einen schönen Aufenthalt in Zürich.

*Prof. (apl.) Dr. Andreas Heine* danke ich für die Einführung in das Fachgebiet der Röntgenstrukturanalyse, sowie seine unermüdliche Geduld und Hilfestellung bei der Bearbeitung zahlreicher Datensätze.

*Dr. Holger Steuber* gebührt mein Dank für das Sammeln von Datensätzen am Synchrotron in Berlin und Hamburg, sowie für seine Hilfe bei der anschließenden Prozessierung.

*Stefan Grüner, Dr. Paul Czodrowski und Michael Betz* danke ich für die bereichernden Ideen und die Zusammenarbeit im Zusammenhang mit den eingereichten Veröffentlichungen.

Bei *Hans-Dieter Gerber* bedanke ich mich für die kompetente Beantwortung zahlreicher Fragen zur Ligandsynthese und –analytik, bei *Steffi Dörr* für die hilfreichen Tipps und Tricks zum Arbeiten im S1-Labor.

*Lydia Hartleben und Christian Sohn* danke ich herzlich für ihre administrative und technische Unterstützung, ohne die das reibungslose Arbeiten in der Gruppe nicht möglich wäre.

*Dr. Peter Kolb* danke ich für die gute Büronachbarschaft.

Den Administratoren der Arbeitsgruppe (*Sven Siebler, Gerd Neudert, Michael Betz, Timo Krotzky, Denis Schmidt, Felix Gut, Felix Terwesten, Phong Nguyen*) danke ich für ihre hohe Einsatzbereitschaft bei der Wartung des Computernetzwerkes.

Dem Team der Massenspektrometrie Lahnberge rund um *Dr. Uwe Linne* danke ich für das Vermessen von Masseproben zur Analytik der exprimierten Proteine.

Bedanken möchte ich mich bei dem *Beamline-Support* des *BESSY II* in Berlin, sowie des *EMBL* in Hamburg. Dem *Helmholtz-Zentrum für Materialien und Energie* in Berlin gebührt mein Dank für die Übernahme der Reisekosten nach und Übernachtungskosten in Berlin.

Mein Dank gebührt ebenfalls der AG Keller, im Besonderen *Jun. Prof. Dr. Sandro Keller* und *Dr. Jana Bröcker*, die mich während meines Besuchs an der TU Kaiserslautern mit offenen Armen empfangen und meine Arbeit auf dem Gebiet der ITC bereichert haben.

*Dr. Jana Brüßler* danke ich für ihre Geduld bei der Einarbeitung in die DLS Messtechnik.

Der gesamten AG Klebe gebührt mein Dank für die angenehme Arbeitsatmosphäre. Dem TGT Team bestehend aus *Dr. Inna Biela, Dr. Stephan Jakobi, Frederik Ehrmann, Phong Nguyen* und *Dr. Florian Immekus* danke ich für die gute Zusammenarbeit und zahlreiche Diskussionen, die in festgefahrenen Situationen neue Anreize geschaffen haben.

Darüber hinaus bedanke ich mich bei *Neli Radeva* für die inspirierenden Frozen Yogurt-Pausen, sowie bei *Dres. Adam* und *Inna Biela, Dr. Ina Lindemann, Dr. Helene Köster, Barbara Wienen, Denis Schmidt, Dr. Stephan Jakobi, Maren Jakobi, Stefan Krimmer* und *Nicole Bertoletti* für eine tolle Zeit in Marburg. Ihr seid zu unverzichtbaren Freunden geworden!

Weiterhin danke ich meinen Freunden aus Schul- und Studienzeiten, die mich auch während meiner Promotionszeit begleitet haben, insbesondere *Dominik, Almuth, Christina, Kristina, Désirée, Isabel, Philip, Carolin, Alexandra* und *Thomas*.

Der tiefste Dank gebührt meinen Eltern *Reiner* und *Martina Neeb*, meinem Bruder *Robin*, sowie meinen Großeltern *Ruth* und *Heinz*, die mich in den vergangenen Jahren zu jeder Zeit und in jeder Situation mit ihrem Interesse an meiner Arbeit, immerwährendem Verständnis und aufbauenden Worten unterstützt haben. Ohne euren Rückhalt hätte diese Arbeit nicht gelingen können!

## Erklärung

(gemäß §10 der Promotionsordnung)

Ich versichere, dass ich meine Dissertation

**Investigations on *lin*-Benzopurines With Respect to Dissociation Behavior, Pocket Cross-Talk, Targeting Resistance Mutants, Residual Mobility, and Scaffold Optimization**

selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den

.....  
(Manuel Neeb)

## Publications

Grüner S., Neeb M., Barandun L. J., Sielaff F., Shun K., Hohn C., Steinmetzer T., Diederich F., Klebe G.:

*Impact of Protein and Ligand Impurities on ITC-derived Protein-Ligand Thermodynamics*, *Biochim. Biophys. Acta* **2014**, 1840: 2843-2850.

Neeb M., Czodrowski P., Heine A., Barandun L. J., Hohn C., Diederich F., Klebe G.:

*Chasing Protons: How Isothermal Titration Calorimetry, Mutagenesis and  $pK_a$  Calculations Trace the Locus of Charge in Ligand Binding to a tRNA-Binding Enzyme*, *J. Med. Chem.* **2014**, 57: 5554-65.

Neeb M., Betz M., Heine A., Barandun L. J., Hohn C., Diederich F., Klebe G.:

*Beyond Affinity: Enthalpy-Entropy Factorization Unravels Complexity of a Flat Structure-Activity Relationship for Inhibition of a tRNA-modifying Enzyme*, *J. Med. Chem.* **2014**, 57: 5566-78.

Wolters S., Neeb M., Berim A., Schulze Wischeler J., Petersen M., Heine A.:

*Structural analysis of coniferyl alcohol 9-O-methyltransferase from *Linum nodiflorum* reveals a novel active-site environment*, *Acta Cryst. D* **2013 May**; 69(Pt 5): 888-900.

Zheng M., Pavan G.M., Neeb M., Schaper A.K., Danani A., Klebe G., Merkel O.M., Kissel T.:

*Targeting the blind spot of polycationic nanocarrier-based siRNA delivery*, *ACS Nano.* **2012 Nov 27**; 6(11): 9447-54.

## Poster Presentations

- 11/2011      Developments in Protein Interaction Analysis (DiPIA), Boston, Massachusetts, USA  
*Thermodynamic Characterization of Putative Inhibitors against Shigellosis*  
 (Manuel Neeb, Luzi Jakob Barandun, François Diederich, Gerhard Klebe)
- 12/2012      4th joint BER II and BESSY II User Meeting, Berlin, Germany  
*From Model Proteins to More Complex Systems – Studying Thermodynamics of TGT-Ligand-Complexes*  
 (Manuel Neeb, Luzi Jakob Barandun, François Diederich, Gerhard Klebe)
- 03/2014      New Approaches in Drug Design & Discovery, Rauschholzhausen, Germany  
*Isothermal Titration Calorimetry of tRNA-Guanine Transglycosylase Inhibitors*  
 (Christoph Hohn, Luzi Jakob Barandun, Manuel Neeb, Florian Immekus, Philipp Kohler, Sandro Tonazzi, Tina Ritschel, Gerhard Klebe, François Diederich)



Aus Gründen des Persönlichkeitsschutzes wird von der elektronischen Veröffentlichung des Lebenslaufes abgesehen.