# Structural and Functional Studies of tRNA-Guanine Transglycosylase: A putative Drug Target for Shigellosis Therapy

**Dissertation**

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Pharmazie

der PHILIPPS-UNIVERSITÄT MARBURG

vorgelegt von

**Bernhard Stengl**

aus Roth bei Nürnberg

Marburg/Lahn 2006

Vom Fachbereich Pharmazie der Philipps-Universität Marburg
als Dissertation angenommen am: 06. Juli 2006

Erstgutachter: Prof. Dr. Gerhard Klebe
Zweitgutachter: PD Dr. Klaus Reuter

Tag der mündlichen Prüfung: 06. Juli 2006

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Herrn Prof. Dr. G. KLEBE am Institut für Pharmazeutische Chemie des Fachbereichs Pharmazie der Philipps-Universität Marburg in der Zeit von Oktober 2002 bis Februar 2006 durchgeführt.

„Wirklich innovativ ist man nur dann,
wenn mal etwas daneben gegangen ist.“

WOODY ALLEN

für HANNA

# Abbreviations

| | |
|---|---|
| Å | Ångström ( 1Å = $10^{-10}$ m) |
| $A_{600}$ | absorption at 600 nm |
| Amp | Ampicillin |
| aqua bidest. | double destillated water |
| ArcTGT | TGT involved in archaeosine modification |
| CATH | Protein Structure Classification Database (Class Architecture Topology Homology) |
| Cm | Chloramphenicol |
| CMC | critical micellar concentration |
| DMSO | dimethylsulfoxid |
| dNTP | desoxynucleosidtriphosphate |
| DTT | dithiothreitol |
| *E. coli* | *Escherichia coli* |
| ECY2 | unmodified *E. coli* $tRNA^{Tyr}$ |
| EDTA | ethylendiamintetraacetate |
| FAE | follicle-associated epithelia |
| h | hour |
| HEPES | 2-[4-(2-hydroxyethyl)piperazino]ethansulfonic acid |
| IPTG | isopropylthio-β-galactosid |
| kb | kilo bases |
| kDa | kilo Dalton |
| $K_{ic}$ | competitive inhibition constant |
| $K_{iu}$ | uncompetitive inhibition constant |
| Km | kanamycin |
| LB | Luria - Bertani complex medium |
| M | molarity (mol · $L^{-1}$) |
| MES | 2-morpholinoethansulfonic acid |
| min | minute |
| NTP | nucleosidtriphosphate |
| PAGE | polyacrylamide gel electrophoresis |
| PAI | pathogenicity island |
| PEG | polyethylenglycol |

| | |
|---|---|
| PCR | polymerase chain reaction |
| PDB | PROTEIN DATA BANK |
| *P. horikoshii* | *Pyrococcus horikoshii* |
| PPase | inorganic pyrophosphatase |
| $preQ_0$ | 7-cyano-7-deazaguanine |
| $preQ_1$ | 7-aminomethyl-7-deazaguanine |
| Q | 7-(((4,5-cis-dihydroxy-2-cyclopenten-1-yl)amino) methyl)-7-deazaguanosine |
| QueA | S-adenosylmethionine:tRNA-ribosyltransferase-isomerase |
| QueTGT | TGT involved in Q modification |
| SCOP | Structural Classification of Proteins Database |
| SDS | sodiumdodecylsulfate |
| *S. flexneri* | *Shigella flexneri* |
| SPB | standard phosphate binding motif |
| SPR | surface plasmon resonance |
| TCA | trichloroacetic acid |
| TGT | tRNA-guanine transglycosylase |
| TIM-barrel | triose-phosphate isomerase (TIM) / $(\beta\alpha)_8$ barrel |
| *T. maritima* | *Thermotoga maritima* |
| Tris | tris-(hydroxymethyl)-aminomethan |
| w/v | weight per volume |
| w.t. | wild type |
| YadB | glutamyl-queuosine $tRNA^{Asp}$ synthetase |
| *Z. mobilis* | *Zymomonas mobilis* |

# Table of contents

# 1. Introduction and Motivation

## 1.1 Structure-based drug design and TGT

Structure-based drug design profits from the enormous amount of available protein crystal structures. In May 2006 36,400 structures were deposited in the protein data bank (PDB; http://www.rcsb.org). It is, however, not only the discovery of new proteins that provides valuable information. Also multiple crystal structures of the same protein in complex with different substrates or of homologous or structurally related proteins from other species are of enormous value. In many cases, the comparison of these structures gives insight into the characteristics and the properties of the protein of interest. The currently available crystal structures can be divided into 8,800 groups with a sequence identity of more than 40%, indicating similar fold and related functions (PDB-SELECT: May, 2006). Thus, on average for each deposited structure four structurally related entries are available.

TGT, the tRNA – guanine transglycosylase is a relevant target for the design of inhibitors against *Shigella*, the causative agent of bacterial dysentery. Computer-based drug design, including the method of virtual screening, enabled the discovery of compounds from structurally very different classes that were capable to inhibit TGT [Grädler *et al*., 2002; Brenk *et* al., 2003; Brenk *et* al., 2004][1-3]. Some of these compounds were suited for crystal structure analysis in complex with the TGT from *Zymomonas mobilis*. These structures form a pool of 33 available TGT crystal structures, including structures of TGT in complex with various substrate molecules, TGT mutants and TGTs from two other bacterial species (PDB; May 2006). In addition to these crystal structures, currently 131 *tgt* gene sequences have been deposited in the UniProt Knowledgebase (SwissProt & TrEMBL; http://www.expasy.org).

The careful analysis of these structures and sequences in combination with kinetic data of substrates and inhibitors is fruitful in two mutually related ways. Firstly, it provides a profound understanding of the structural prerequisites necessary for catalysis [Xie *et al.*, 2003][4]. In particular innate protein flexibility has to be considered in this context. But also water molecules can be identified to be relevant for catalysis or for the stabilization of alternative binding geometries [Brenk *et al*., 2003][2].

Secondly, the knowledge of deviating binding competent conformers can be implemented in the further development of available inhibitor series or in the computer-based search of new inhibitor classes [Brenk *et al*., 2004][1]. Such attempts might finally result in a potent antibiotic against *Shigella* TGT, that allows to minimize the symptoms of bacterial dysentery.

# 1.2 Shigellosis

## 1.2.1 Disease and treatment

Shigellosis or bacillary dysentery is caused by bacteria belonging to the genus *Shigella*. Oral uptake of contaminated drinking water or food possibly initiates an infection. The bacteria pass through the stomach, the small intestine and finally reach the colon where they invade the colon epithelium and the mucosa. This results in a destructive recto-colitis which is responsible for the dysenteric symptoms like watery diarrhoea, fever, intestinal cramps and emission of mucopurulent and blood stools [Sansonetti *et al*., 2001][5]. Due to infiltration of inflammatory cells in the colon epithelium mucosal ulcers are often observed [Mathan & Mathan, 1991][6]. In the absence of an effective treatment, secondary complications may occur like septicaemia, pneumonia or haemolytic uremic syndrome [Bennish, 1991][7].

Shigellosis is a global burden with an estimated annual number of 165.7 million episodes. Only 1.5 million cases occur in industrialized countries, many of them endemically in children day care centres, custodial institutions or as traveller's diarrhoea. Almost no fatalities are observed. The vast majority of 163.2 million episodes takes place in developing countries, often epidemically, with a death rate of 1.1 million. Most of the episodes occur in displaced populations. Civil war in areas with dense population and natural disasters (flooding, drought) often cause hunger and poverty forcing people to leave their home. Mostly, these displaced people are crowded in areas with poor sanitation and insufficient supplies of clean water. Concomitant malnutrition is widely spread. Children are affected most severely from such conditions. Thus, 69 % (112.6 million) of all episodes and 61 % (660.000) of all fatalities involve children under 5 years of age [Kotloff *et al*., 1999][8]. Particularly in Africa the interference with HIV / AIDS is a serious problem.

For effective treatment clean water, sufficient food supply and hygiene conditions avoiding the fecal – oral spreading route are inevitable. This includes washing of

hands as well as the use of separate dishes and towels for patients and other members of the household. Additionally, the control of houseflies by bait and trap strategies significantly reduces the spreading of *Shigella*. Houseflies are a key vector for the faecal contamination of human food [Cohen *et al*., 1991][9].

Antibiotics can be used to treat shigellosis and reduce the period of bacterial excretion from the patient. In developing countries they are used to stop epidemic spreading [Jennison & Verma, 2004][10]. A major problem for shigellosis treatment is the extraordinary ability of *Shigella* to acquire plasmid-encoded resistance to antimicrobial drugs that constituted the first-line therapy. Sulfonamides, tetracycline, ampicilline and trimethoprim-sulfamethoxy-azole were highly efficient drugs two decades ago, but are becoming more and more ineffective [Kotloff *et al*., 1999][8]. For trimethoprim the resistance rate rose from 3 % in 1975 to 98 % in 1988 [Heikkila *et. al*., 1990][11]. A survey from Israel reported for the period 1991 - 2000 high resistances to trimethoprim (94 %) and ampicillin (85 %), significantly increased resistances to tetracycline (23 % to 87 %) and emerging resistances to quinolones (0.5-2 %) [Ashkenazi *et al*., 2003][12]. During an epidemic outbreak in Zaire in 1994 a *Shigella* strain was identified that was resistant to all commonly used antibiotics [Goma Epidemiology Group, 1995][13]. These findings underline the need to monitor resistance and to develop new, innovative antibiotics that maintain the ability of successful treatment.

A further important goal is the development of vaccines that prevent infection. Since the 1940s this has been attempted with little success. Current research approaches using deeper insight into *Shigella* pathogenicity, however, made promising progress. But still these vaccines do not yet fulfil efficacy and safety requirements for the treatment of humans [Jennison & Verma, 2004][10].

### 1.2.2 *Shigella – Escherichia* relationship

*Shigellae* are GRAM-negative, nonsporulating, facultative anaerobic bacilli. The genus *Shigella* is divided into four different ‘species’ and a varying number of serotypes, based on biochemical differences and variations in their O-antigen [Sansonetti, 2001][14]. All four ‘species’ are spread world wide, although significant regional differences are observed [Kotloff *et al*., 1999][8].

*Shigella flexneri* (6 serotypes)

is the most abundant species worldwide (median value of 60 %). It dominates in developing countries with 55 - 85 % (industrialized countries: 16 %).

*Shigella sonnei* (1 serotype)

is the next most common species (15 %). It is the most abundant species in industrialized countries (77 %), responsible for predominantly endemic episodes. In developing countries it is limited to 5 - 30%.

*Shigella boydii* (8 serotypes)

is less common (6 %) and almost equally distributed all over the world with a slight peak in the Middle East (15 %).

*Shigella dysenteriae* (16 serotypes)

is as common as *Shigella boydii* (6 %) in average. In industrialized countries it is rarely observed (1 %). Significant peaks are found in South Asia (27 %) and sub-Saharan Africa (31 %). These peaks result from often deadly epidemic episodes caused by *S. dysenteria* subtype 1 ("Shiga bacillus"). This subtype produces Shiga toxin, a potent cytotoxin. Shiga toxin comprises two subunits and is capable to induce severe inflammations. Upon release from the bacilli it spreads all over the body of the patient damaging organs like kidneys, brain as well as red blood cells. This causes severe additional complications for the diarrhoea patient like haemorrhagic colitis and the haemolytic uremic syndrome (HUS) [O'Loughlin & Robins- Browne, 2001][15].

The genus *Shigella* belongs to the family of *Enterobacteriaceae*. Strictly speaking it does not constitute a separate genus. Genetic analysis revealed that *Shigella* belongs to the core of *Escherichia coli* strains [Escobar-Páramo *et al*., 2004][16]. Thus, it would be more appropriate to consider *Shigella* as a subtype of *E. coli* and call it 'enteroinvasive *E. coli* ' (EIEC).

### 1.2.3 Cellular and molecular pathogenicity

*Shigellae* are very infective. Only 10 – 100 bacteria can cause a disease in an adult. After passage through the stomach and the small intestine they are capable to invade the colon epithelium. The passage is not possible through the apical side of the epithelial cells (Fig. 1.1a). Invasion occurs indirectly via M cells. They are

specialized to transport antigens (bacteria, etc.) through the colon epithelium to present them to macrophages which are in tight association with them. Both are located in follicle-associated epithelia (FAE). The FAE overly the mucosa-associated lymph nodes that are responsible for intestinal immunity. *Shigella* have the unusual capacity to enter the M cell associated macrophages without being damaged. After phagocytic uptake by the macrophage, they escape from the phagosome. Inside the macrophages they induce apoptosis. Upon apoptotic death inflammatory interleukins (IL-1β and IL-18) are produced and released into the mucosa after lysis, together with the intact *Shigella*. The bacteria then enter the colon epithelial cells from the basal side. Once inside these cells they propagate and spread from cell to cell – well protected from immune cells. Infected epithelial cells release IL-8. This interleukin, together with the two above mentioned ones, attracts macrophages and polymorphonuclear leucocytes (PMN) from subepithelial tisues. The latter disrupt the integrity of the epithelial barrier and facilitate further *Shigella* invasion from the colon. Thus, in a sort of 'snowball effect' the initial inflammatory reaction is amplified and results in an uncontrolled tissue destabilization. Nevertheless, in contrast to macrophages, PMN are able to kill *Shigella* inside their phagosomes. Together with NK cells and T lymphocytes, immune cells attracted by IL-18, they eventually cope with the infection. In addition IL-18 triggers the production of the interferon IFN-γ which activates the immune system. [Sansonetti *et al.*, 2001; Sansonetti *et al.*, 2001; Jennison & Verma, 2004][5, 10, 14].

In order to enable cell entry, intracellular mobility, cell spreading and induction of apoptosis *Shigella* produces virulence factors. These virulence factors are used to reprogram the cellular machinery of epithelial as well as immune cells by activating innate transport and signalling pathways [Fernandez & Sansonetti, 2003; Van Nhieu *et al.*, 2000][17, 18].

The first step of invasion into an epithelial cell is the interaction of a tube like type III secretion apparatus with the host cell membrane (Fig 1.1b). A pore is formed on the tip of the secretion apparatus by the virulence factors IpaB and IpaC. Exposure of IpaC into the host cell cytoplasm activates a Src tyrosine kinase. A further signalling cascade results in actin dependent filopodia and lamellipodia formation. These structures form in the surrounding of the secretion apparatus tip and finally enclose *Shigella*, resulting in macropinocytotic uptake. To provide lamellipodia with enough actin and to avoid polymerization in the secretion apparatus contact region, IpaA is

secreted through the secretion tube porus into the host cell cytoplasm. Within the cell IpaA binds to vinculin and available F-actin from adjacent cytoskeleton elements of the host cell is depolymerised.

After the uptake of *Shigella* into the host cell the macropinocytotic vacuole lyses and the bacterium is released into the cytoplasm. To enable intracellular mobility the virulence factor IcsA is produced. It is located in the bacterial cell wall, exposed to the host cell cytoplasm. Binding of the host cell proteins N-WASP and ARP2/3 results in actin polymerization that pushes the bacteria through the cytoplasm.

The described virulence factors are only the most prominent ones among a vast set of further virulence factors involved in generating pathogenicity.

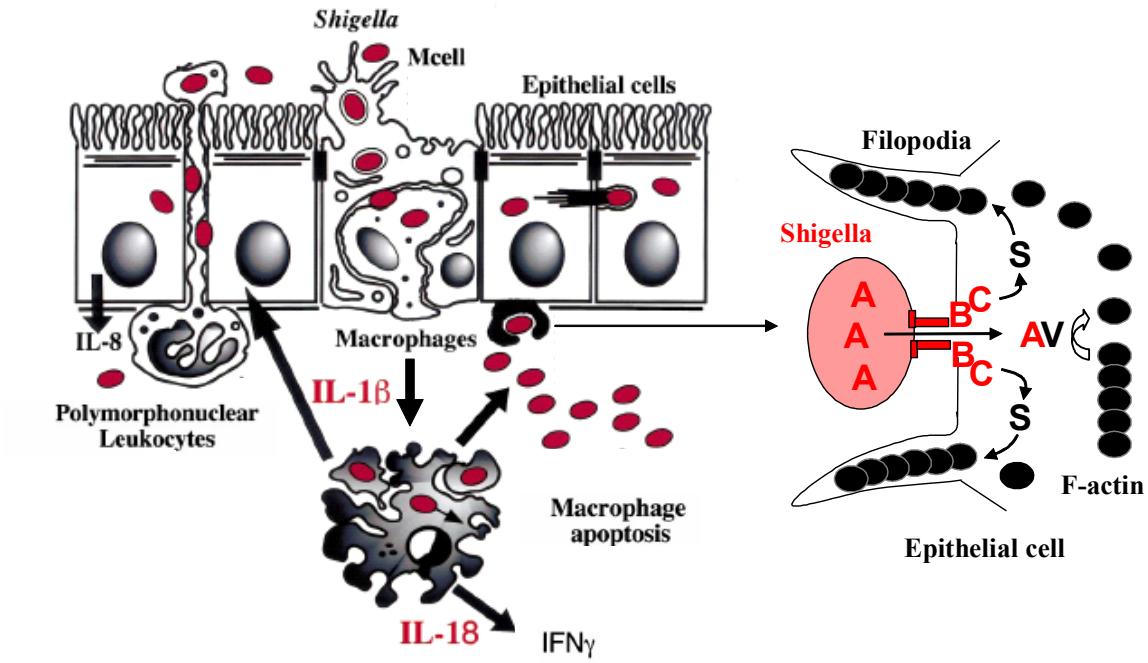



**Fig. 1.1** a) Pathway of epithelial colonization and disintegration – figure modified from Sansonetti (2001)[5]; b) macropinocytotic uptake of *Shigella*, A: ipaA, B: ipaB, C: ipaC, V: vinculin, S: Src kinase, red: *Shigella* virulence factors, black: host cell proteins

## 1.2.4 Regulation of pathogenicity

The virulence factor genes are located on a 214-kb virulence plasmid, which was isolated and sequenced from *S. flexneri*. Coding sequences are scattered all over the plasmid. One block of 30 kb shows a particular dense pattern of genes and is called pathogenicity island (PAI). The *mxi* / *spa*- and *ipa*-loci found in this region code for proteins necessary to establish the type III translocon and to allow cell entry [Sansonetti, 2001][5]. The expression of virulence genes is organized hierarchically [Dorman & Porter, 1998][19]. VirF and VirB are the key transcription activators for virulence gene expression. The *virF* and *virB* genes are plasmid encoded. Expression of virF directly activates the transcription of virulence factor genes like *icsA* as well as the transcription of the *virB*-gene. The *virB*-gene product then activates the transcription of the *mxi* / *spa*- and *ipa*-genes. Thus, VirF is in the centre of pathogenicity regulation (Fig. 1.2).

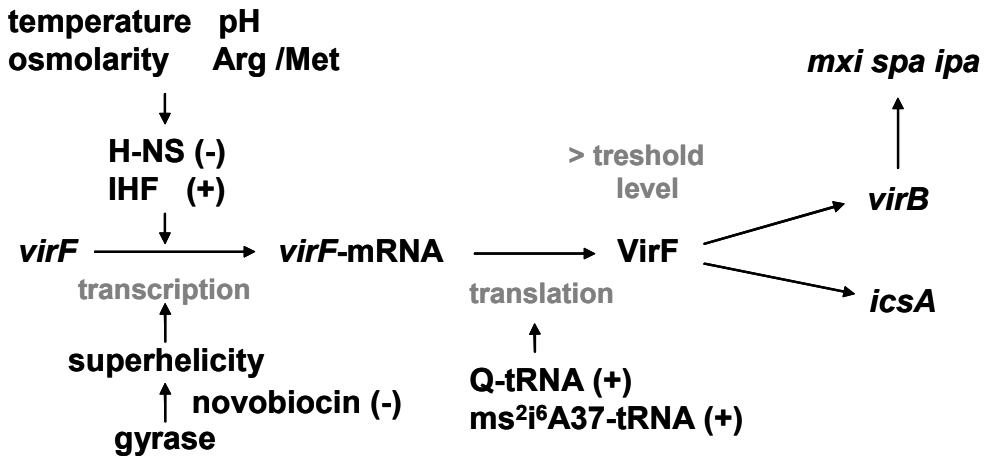



**Fig. 1.2** Regulation of VirF expression

Remarkably, it is not the absence or presence of VirF that regulates virulence. Instead, a threshold level for VirF exists above which virulence factors are produced. The amount of VirF depends on various environmental and internal factors. Such environmental factors are pH, osmolarity, temperature and nutrition factors. Below 30° C *Shigella* is not virulent. The *virF*-gene transcription is fully activated at 37° C, pH 7.4, physiological osmolarity and in the presence of free amino acids (arginine / methionine). The virF- gene promoter is positively regulated by the transcription factor IHF and negatively by H-NS. Both transcription factors are encoded by the

chromosome and their expression is sensitive to the mentioned environmental factors [Dorman & Porter, 1998; Durand *et al*., 2000; Durand & Björk, 2003][19-21].

Apart from this 'classical' regulation the virF level also depends on an accurate transcriptional and translational machinery [Durand *et al*., 2000][21]. Each intervention affecting this machinery may also influence VirF expression. The antibiotic novobiocin for instance inhibits gyrase, thus decreasing negative superhelicity in DNA. This results in decreased virF levels. On the translational level the expression of virF can be influenced as well. Efficient translation of *virF*-mRNA at the ribosome requires the presence of modified tRNA molecules. Modification of tRNA bases is very common in nature [Björk, 1996][22]. Two modifications were demonstrated to exhibit major influence on the *virF* mRNA translational speed. In specific tRNA molecules the highly modified nucleosides queuosine (Fig. 1.4) in position 34 (the anticodon wobble position) or 2-methylthio-$N^6$-isopentenyladenosine ($ms^2i^6A37$) in position 37 (adjacent to the anticodon) have to be present (Fig. 1.3). *Shigella* mutants lacking one of these modifications show significantly reduced virulence. This was tested in mutational studies where gene knock-out in the tRNA modification pathway resulted in the absence of these modifications. The first step of A37 modification is catalyzed by the *miaA* gene product. Mutation of the *miaA* gene reduces the VirF level to 10%, and the haemolytic activity to 10 - 20% compared to the wild type [Durand *et al*., 1997][23]. In tRNA position 34 the *tgt* / (*vacC*)-gene product catalyzes the incorporation of a queuine precursor into tRNA. Mutation of the *tgt*-gene reduces both, VirF level and haemolytic activity, to 50 - 60% of the wild type [Durand *et al*., 1994; Durand *et al*., 2000][21, 24]. Thus, tRNA modifying enzymes could represent promising targets for the development of antibiotics. Inhibition of such specific tRNA modification steps should result in significantly reduced virulence of *Shigella*.

For *E. coli* the *tgt*-gene product and its function has been characterized in detail (chapter 2.1). It codes for the tRNA – guanine transglycosylase (TGT). A crystal structure of this protein for the structurally very similar *Zymomonas mobilis* TGT is available [Romier *et al*., 1996][25]. Thus, this TGT can be used for structure-based drug design to develop potent inhibitors and finally to test the hypothesis of significant virulence reduction.

# 1.3 Queuosine-modification

## 1.3.1 tRNA-modification

Transfer RNA (tRNA) maturation involves a series of post-transcriptional processing steps resulting in fully functional tRNA molecules. Among these maturation steps nucleoside modification is the most remarkable one. For tRNA a wealth of structural changes of canonical nucleosides has been described. Typically ~ 10 % of all nucleosides are modified in tRNA, but as many as 25 % can be affected. More than 80 modifications have already been described so far [Björk, 1995][26]. Many of them are conserved across broad phylogenetic boundaries. These modifications range from methylations and thionylations to extensive 'hypermodifications' of canonical bases in multiple enzymatic steps. Examples for such 'hypermodifications' are queuosine, wybutosine or 2-methylthio-$N^6$-isopentenyl-adenosine (Fig. 1.3 and 1.4). The role of such modifications depends on their position. Modifications outside the anticodon region are thought to influence structural integrity or to serve as recognition determinants for the ribosome. Modifications within or around the anticodon are proposed to fine-tune translational speed and fidelity or to influence the occurence of frame-shifting events. Nevertheless, the present understanding is still rudimentary [Björk *et al*., 1999; Iwata-Reuyl, 2003][27, 28].

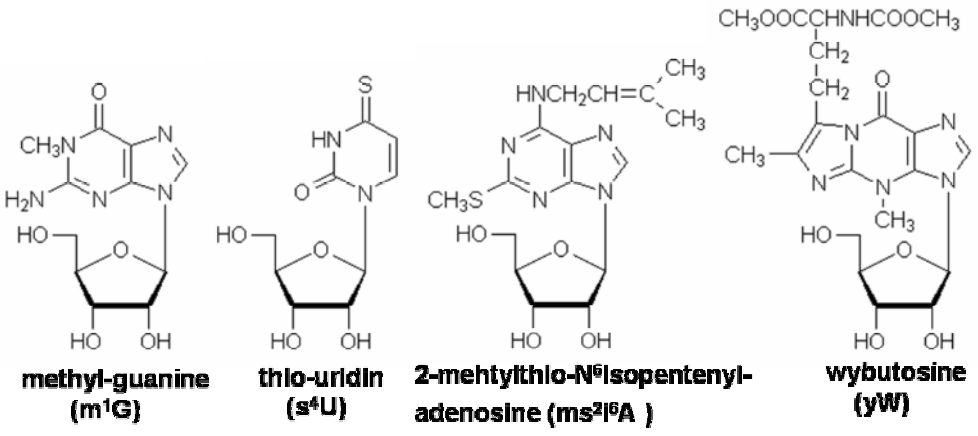



**Fig. 1.3** Examples for modified and 'hypermodified' tRNA bases (from: *The RNA Modification Database*: http://medlib.med.utah.edu/RNAmods/)

## 1.3.2 Queuosine-modification pathway

The modification of tRNA in position 34 with queuosine (7-(((4,5-cis-dihydroxy-2-cyclopenten-1-yl)amino)methyl)-7-deazaguanosine = Q ; Fig. 1.4) is one of the most pronounced modifications known to date. Queuosine is found in most eubacterial and eukaryotic species. Only few species are not capable of Q-synthesis, among them *Saccharomyces cervisiae* and the eubacterial division of Actinobacteria (*Mycobacerium, Corynebacerium, Streptomyces, Bifidobacterium*) [Reader *et al*., 2004][29].

Queuosine is found in four specific tRNAs with an anticodon constituted by G in position 34, U in position 35 and a variable base in position 36 ($G_{34}U_{35}N_{36}$, N = A,C,G,U). These four tRNAs are coding for the amino acids asparagine, aspartic acid, histidine and tyrosine [Okada & Nishimura,1979; Okada *et al*., 1979][30, 31].

The physiological relevance of the Q-modification is not yet fully understood. *E. coli* mutants defective in Q-synthesis exhibit an apparently normal phenotype and growth rate during favourable growth conditions. Upon entry into stationary growth phase viability drops significantly, but the physiological background remains elusive [Noguchi *et al*., 1982][32]. In mammals the transcription of phenylalanine hydroxylase, involved in tyrosine biosynthesis, was shown to be Q-dependent [Marks & Farkas, 1997][33]. The absence of queuine might result from an mRNA mistranslation [Iwata-Reuyl, 2003][28]. Together with the knowledge of Q-dependent VirF translation it can be assumed that the Q-modification influences anticodon base pairing. Q-modification in tRNA seems to increase efficiency of interactions of the wobble position with specific mRNAs thus resulting in increased translational speed and correct frame-shifting.

In Eubacteria Queuosine-tRNA is produced in a multi step reaction (Fig. 1.4). The initial step of tRNA modification is performed by the enzyme tRNA – guanine transglycosylase (QueTGT). It specifically recognizes the four tRNAs mentioned above by a common $U_{33}G_{34}U_{35}$ sequence. [Nakanishi *et al*., 1994; Curnow & Garcia, 1995][34, 35] QueTGT replaces guanine 34 ($G_{34}$) by the Q-precursor $preQ_1$ (7-(aminomethyl)-7-deazaguanine) [Okada & Nishimura,1979][30]. Most probably, $preQ_1$ is produced from guanosine triphosphate (GTP), [Kuchino *et al*., 1976][36], by means of the *queC*-, *queD*-, *queE*- and *queF*-gene products. These genes have been identified in a comparative genomics study [Reader *et al*., 2004][29]. The involvement of QueC in $preQ_1$ biosynthesis has been demonstrated in detail [Gaur & Varshney,

2005][37]. The roles of the *qeuD*- and *queE*-gene products are still unclear. QueF performs the reaction step previous to QueTGT. The NADPH-dependent enzyme catalyzes the reduction of $preQ_0$ (7-cyano-7-deazaguanine) to $preQ_1$ [Van Lanen *et al*., 2005][38]. Surprisingly, both bases, $preQ_0$ and $preQ_1$, are capable of binding to QueTGT, however, $preQ_1$ is preferentially incorporated into tRNA [Hoops *et al*., 1995][39]. After incorporation $preQ_1$ is transformed to the functional base queuine in two steps [Iwata-Reuyl, 2003].[28, 40]. The first one is performed by S-adenosylmethionine:tRNA ribosyltransferase-isomerase (the QueA enzyme). QueA transfers a ribosyl moiety from S-adenosylmethionine to $preQ_1$ resulting in epoxyqueuosine formation (oQ, (7-((N-(2,3-epoxy-4,5-cis-dihydroxycyclopent-1-yl)amino) methyl)-7-deazaguanosine). In the second step epoxyqueuosine is reduced by a still unknown cofactor $B_{12}$-dependent enzyme to Q In some bacterial species, among them *E. coli* and EIEC, queuosine-$tRNA^{Asp}$ is further modified to glutamyl-queuosine by YadB, a glutamyl-queuosine $tRNA^{Asp}$ synthetase [Blaise *et al*., 2004; Campanacci *et al*., 2004; Dubois *et al*., 2004; Salazar *et al*., 2004][41-44]. YadB aminoacetylates one of the hydroxyl groups of the queuosine-pentenyl moiety *via* the transfer from glutamyl-AMP. The glutamylation is prone to hydrolysis and has a rather short half-life.

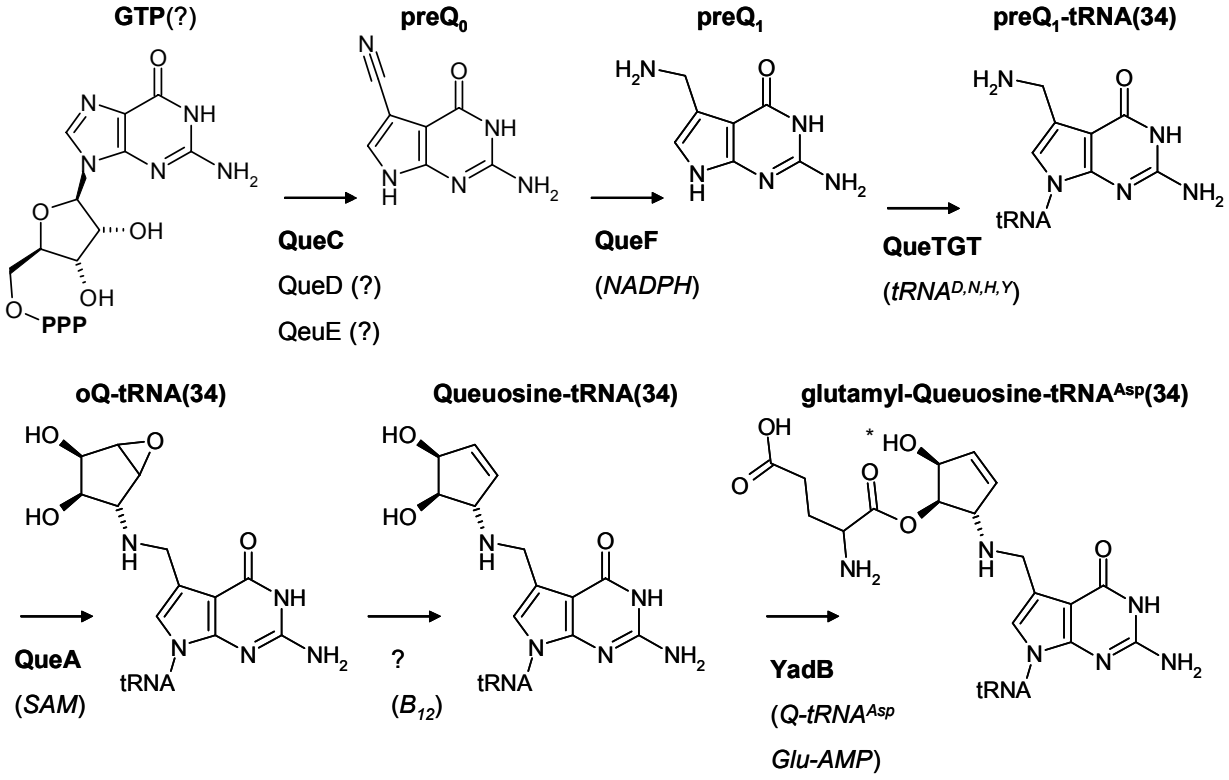



**Fig. 1.4** Queuine modification pathway (* which OH-group is glutamylated is not known)

Degradation of Q-tRNAs in *E. coli* is performed by the enzyme 'Colicin E5'. It specifically releases queuosine from tRNA by cleaving the phosphodiester bonds to U35 and U33 [Lin *et al*., 2005][45].

Crystal structures are available of the enzymes QueF from *Bacillus subtilis* [Swairjo *et al*., 2005][46] and QueTGT from *Zymomonas mobilis* [Romier *et al*., 1996][25] and *Thermotoga maritima* (PDB-code: 2ASH). Of QueA from *B. subtilis* (PDB-code: 1YY3) and *T. maritima* (PDB-code: 1VKY) coordinates are deposited in the PDB.

In Eukaryota queuosine is present in the same four tRNAs in position 34 as in Eubacteria. Nevertheless, the Q-modification process differs fundamentally. Eukaryota are not capable to synthesize queuosine *de novo*. Thus, the base queuine is a nutrition factor. It has to be acquired by food or from bacterial commensals of the digestive apparatus. This assumption was indicated by reduced queuosine levels in proliferating eukaryotic tissue cells. During growth the amount of free queuine is limited and cannot be restored quickly enough. The incorporation of queuine into tRNA is performed in a single step reaction. It is performed by the eukaryotic QueTGT accepting queuine as substrate, in contrast to eubacterial TGT which is accepting $preQ_1$. For Eukaryota further modifications of queuine are described resulting in glycosylations of the pentenyl hydroxyl groups. However, the enzymes responsible for the formation of these glycosylated β-D-mannosyl-Q-tRNAs and β-D-galactosyl-Q-tRNAs are still unknown [Iwata-Reuyl, 2003][28].

### 1.3.3 Archaeosine-modification in Archaebacteria

In Archaebacteria no queuine modified tRNAs are found. Nevertheless, central parts of the Q-modification pathway are present in Archaebacteria as well (Fig. 1.5). Deviating from Eubacteria and Eukaryota the tRNA – guanine transglycosylase from Archaebacteria (ArcTGT) incorporates the $preQ_1$ precursor $preQ_0$ into position 15 of the dihydrouridine loop (D-loop) of archaebacterial tRNAs. For ArcTGT crystal structures are available in complex with $preQ_0$, guanine and full length $tRNA^{Val}$ [Ishitani *et al*., 2002; Ishitani *et al*., 2003][47, 48].

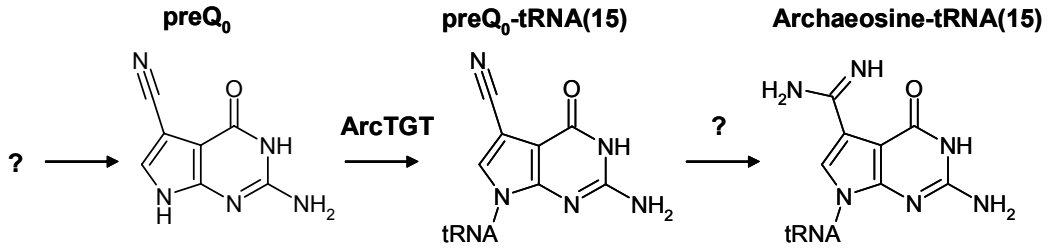



**Fig. 1.5** Archaeosine-modification pathway

The further modification steps of $preQ_0$ differ significantly from Eubacteria and Eukaryota. In yet unknown steps, $preQ_0$ is modified to archaeosine (7-formamidino-7-deazaguanine), a modified base found in virtually all of the archaebacterial tRNAs [Sprinzl *et al*., 1998][49].

$G_{15}$ or archaeosine$_{15}$ is buried in the tRNA core and involved in the formation of salt bridges between the positively charged formamidino group of archaeosine and RNA phosphate backbone groups. Archaeosine modification is thought to stabilize the canonical L-shape of tRNA under the predominantly high temperature most Archaebacteria are exposed to [Gregson *et al*., 1993; Iwata-Reuyl, 2003][28, 50].

## 1.4 Aim of the project

In two preceding PhD projects structure-based design resulted in the development of inhibitors of the QueTGT from *Z. mobilis*. A test system was established, [Grädler, PhD Thesis, 2000][51], and by the combination of classical structure-based design with computer-based methods various classes of TGT inhibitors have been discovered [Brenk, PhD Thesis, 2003][52].

QueTGT and ArcTGT crystal structures in complex with tRNA which became available in recent time, however, provide deeper insight into the molecular foundations of the base exchange reaction and substrate specificity. In particular the *Z. mobilis* TGT crystal structure in complex with a tRNA substrate unravelled the misinterpretation of previous data [Xie *et al*., 2003][4]. Asp280 instead of the initially suggested Asp102 was identified as nucleophile of the base exchange reaction. Additionally, kinetic analysis revealed that the reaction pathway follows a ping-pong mechanism [Goodenough-Lashua & Garcia, 2003][53]. Therefore, the assumed model of the base exchange mechanism required revision. However, the newly collected evidence has not been discussed in literature with respect to important mechanistic details.

In the first part of this thesis structural and functional analysis attempt to integrate the current knowledge to a comprehensive picture concerning the structural basis of the base exchange reaction in TGTs. This is relevant as the binding pocket in the surrounding of the nucleophile Asp280 is intended to be addressed in structure-based design approaches.

Detailed analysis of the available structural and sequence data should provide a new functional model concerning the molecular basis of the reaction mechanism.

TGTs from the three kingdoms of life exhibit pronounced differences in substrate specificity. In a mutant study the molecular basis for substrate promiscuity in QueTGT, discovered in the previous study, will be investigated by means of crystal structure and kinetic analysis.

The occurrence of dimers in QueTGT crystal structures and the putative functional relevance of such dimers, indicated by the crystalline complexes with bound tRNA, will be assessed by means of structure and sequence analysis.

Finally, the comparison of the available TGT crystal structures in an evolutionary context should give a more comprehensive understanding of this enzyme family.

In the second part of this thesis the consequences of the recently collected knowledge concerning the reaction pathway will be implemented in structure-based design approaches to develop more potent inhibitors.

An appropriate binding assay has to be developed taking competitive and uncompetitive inhibition contributions into consideration. They are relevant in a ping-pong reaction mechanism and were not considered in the previous assay. Additionally, the revalidation of relevant members of compound classes studied in previous investigations is required to adjust structure – activity relationship.

Based on a quinazolinone scaffold, developed in a previous study, new lead structures will be developed. This effort continues the successful collaboration with the group of Prof. Diederich (ETH Zürich) on the design and synthesis of potent inhibitors.

A virtual screening hit should be evaluated in more detail to test the relevance of a surprising binding mode suggested by docking. Systematically, structural modifications of this hit should be developed and tested. For synthesis a cooperation has been initiated with the group of Prof. Link (University of Greifswald).

A series of TGT binders has been discovered in plant extracts by the ligand fishing method developed in the group of Prof. Matusch (University of Marburg). These compounds should be characterized by means of kinetic and crystal structure analysis.

In chapter 2 results from the structural and functional analysis will be presented. In chapter 3 necessary modifications of the assay, revalidated previous inhibitor series and results from structure-based design approaches will be presented. In chapter 4 the results of this study will be summarized. The applied methods will be presented in chapter 5.

# 2. Structural and Functional Analysis

## 2.1 QueTGT – ArcTGT: base exchange reaction

Crystal structures of QueTGT and ArcTGT in complex with tRNA became available during this thesis [Xie *et al*., 2003; Ishitani *et al*., 2003][4, 48]. The key impact of these structures was the discovery of a misinterpretation of previous data. Asp280, instead of Asp102, is the nucleophile of the base exchange reaction [Xie *et al*., 2003][4]. Therefore, the model of the base exchange mechanism requires revision. However, no comprehensive interpretation is available. Analyses of TGT crystal structures and sequences provide a new functional model concerning the molecular basis of the reaction mechanism and the residues involved in this process. These results have already been published in Stengl *et al*. (2005)[54].

### 2.1.1 TGTs in the tree kingdoms of live

Although tRNA – guanine transglycosylases (TGT) are present in all three kingdoms of life they accept deviating bases as substrates incorporated at different positions into tRNA (Fig 2.1). Nevertheless, their principal architecture and the underlying reaction mechanism are highly conserved among the three kingdoms [Stengl *et al*., 2005][54].

Structurally all TGTs adopt the highly populated triose-phosphate isomerase (TIM)-type $(\beta\alpha)_8$-barrel fold with specific insertions involved in tRNA recognition and binding. These insertions are an N-terminal antiparallel $\beta$-sheet and a 'zinc binding' site close to the C-terminus (Fig. 2.1) [Romier *et al*., 1996; Romier *et al*., 1997; Ishitani *et al*., 2002][25, 47, 55]. The overall shape of TGTs is sufficiently unique to form a homologous superfamily within the TIM-barrel fold (SCOP database (version 1.65) [Andreeva *et al*., 2004][56]; CATH database (version 2.5.1) [Pearl *et al*., 2000][57]). The TGT superfamily is subdivided into two groups, QueTGT and ArcTGT. QueTGT includes the eubacterial and the eukaryotic TGTs. They are both involved in the modification of tRNA in position 34 with queuine. Structurally they are very similar and share the additional insertion $\beta$7-$\alpha$6. In contrast, the archaebacterial TGT is part of the archaeosine modification pathway in position 15 of tRNAs, thus it is referred to as ArcTGT. Apart from the deviating modification site also some remarkable structural differences can be noted. In ArcTGT three supplementary C-terminal

domains are present (C1, C2, C3 = PUA) and the insertion β7-α6 is missing [Stengl *et al*., 2005][54].

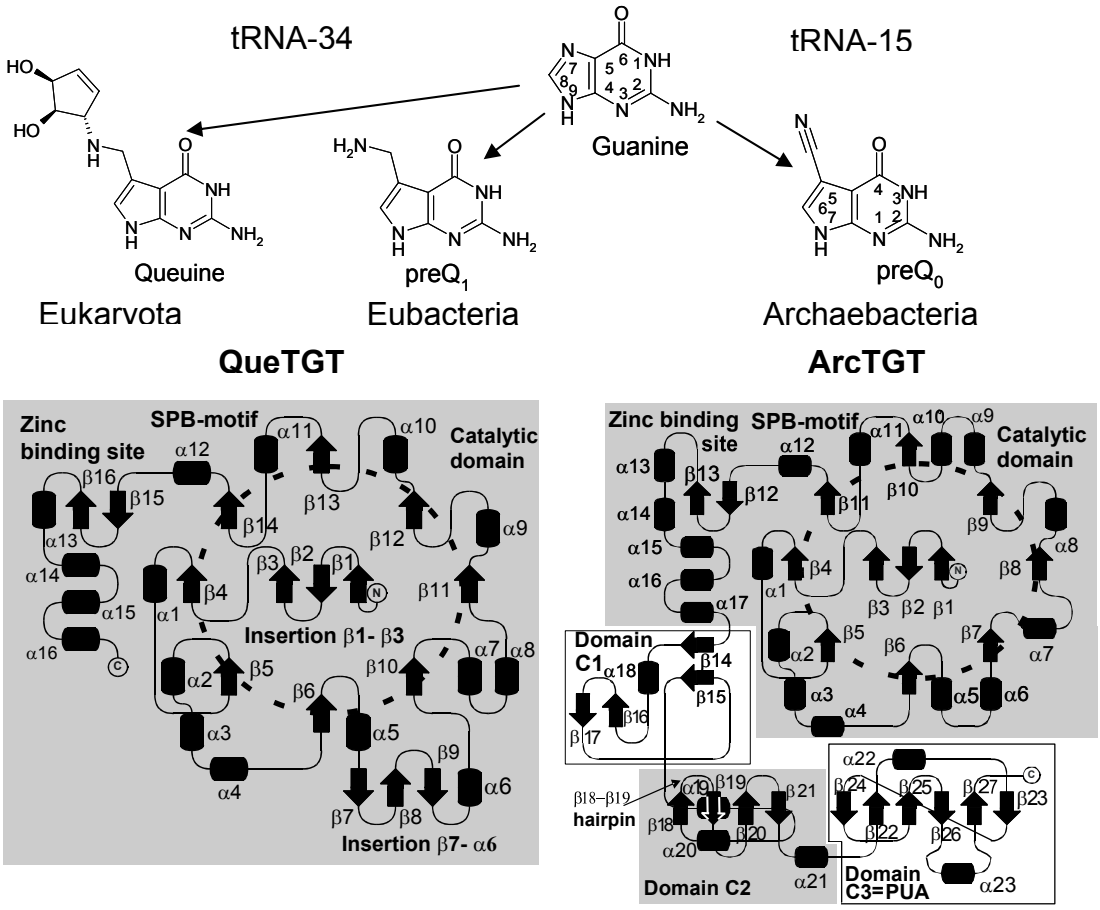



**Fig 2.1** ArcTGT/QueTGT secondary structures and substrate specificity

Crystal structure analyses of ArcTGT and QueTGT in complex with tRNA substrates allowed to extract valuable information about substrate recognition [Xie *et al*., 2003; Ishitani *et al*., 2003][4, 48]. At first glance, a direct comparison of substrate recognition by ArcTGT and QueTGT is rather difficult. The two guanine residues, 15 and 34, recognized and replaced by the two TGTs are located at completely different positions in the tRNA strand. As a consequence, the overall binding geometry of tRNA with respect to the catalytic domain is remarkably different (Fig. 2.2a). Nevertheless, the substrate binding pockets of both, QueTGT and ArcTGT, accommodate a trinucleotide sequence. It consists of the guanine nucleotide addressed by the respective enzyme as well as of the directly preceeding and the following nucleotide. Thereby, the QueTGT bound trinucleotide, the specificity regulating UGU-sequence, is present in a very similar conformation as the one bound to ArcTGT (Fig. 2.2b) [Stengl *et al*., 2005][54]. To make these bases addressable by

TGTs, in both cases tRNAs have to undergo specific conformational changes (Fig. 2.2c). In case of QueTGT the tRNA anticodon loop bends to the opposite direction compared to uncomplexed tRNA molecules. In contact with ArcTGT the D-arm of the tRNA protrudes in order to allow modification of the usually buried $G_{15}$ residue. The tRNA then adopts an unusual conformation called λ-conformation stabilized by a new structural element called 'DV'-helix (details are given in chapter 2.1.4.1).

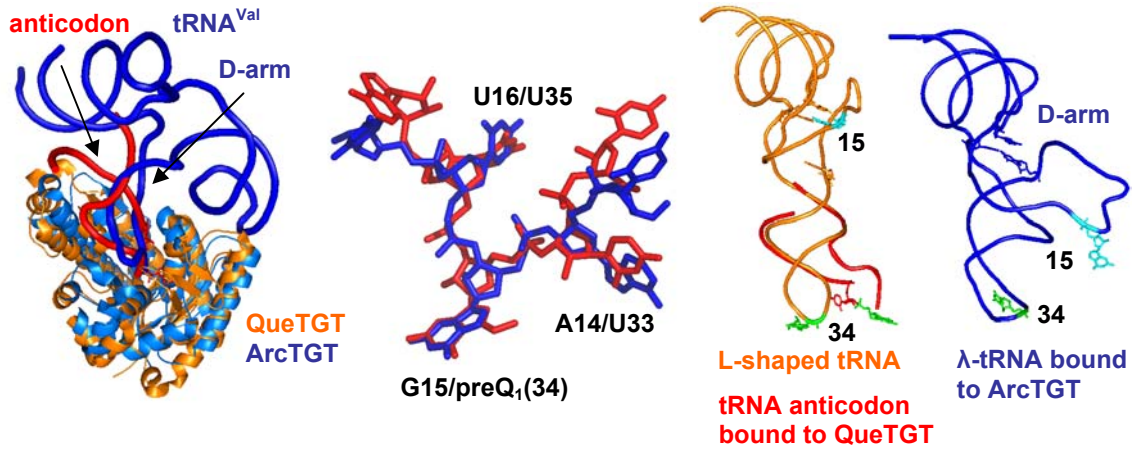



**Fig. 2.2** a) ArcTGT/QueTGT global recognition of tRNA; b) trinucleotide recognition of G34 and G15; c) conformational changes of tRNA bound to TGT

All TGTs share a common mechanism indicated by the high degree of structural conservation in the surrounding of the central guanine residue (G15 or G34 respectively) and the active site (Asp280). Most of the residues involved in QueTGT and ArcTGT catalysis are retained or at least conservatively replaced. The sequence alignment of Table 2.1 gives an impression of the degree of conservation. The following numbering refers to *Z. mobilis* TGT representative for QueTGTs (residues$^{Q}$) and *P. horikoshii* TGT representative for ArcTGTs (residues$^{A}$). The most important residues of QueTGT are Asp 280$^{Q}$, the catalytic nucleophile, and the residues contributing to the recognition of the guanine-like skeleton of the substrates: Asp102$^{Q}$, Asp156$^{Q}$, Gln203$^{Q}$ and Gly230$^{Q}$. With Asp249$^{A}$, Asp95$^{A}$, Asp130$^{A}$, Gln169$^{A}$ and Gly196$^{A}$ they have identical counterparts in ArcTGTs.

The TGT superfamily exhibits a pronounced difference in substrate specificity and promiscuity [Okada & Nishimura, 1979; Hoops *et al*., 1995; Shindo-Okada *et al*., 1980, Bai *et al*., 2000, Watanabe *et al*., 1997] .[30, 39, 58-60] . Although every TGT has a preferred substrate (Fig. 2.1) some TGTs show an extended reservoir of substrates.

While ArcTGT exclusively accepts $preQ_0$, eubacterial QueTGT accepts $preQ_0$ and $preQ_1$, however with a significant preference for $preQ_1$. Eukaryotic QueTGT, additionally to queuine also accepts $preQ_0$ and $preQ_1$ as substrates (Tab.2.2). This results from differences in the constitution of the specificity region in the binding pocket of the three TGTs (Tab. 2.1). The reason for these deviating specificities will be discussed in detail in chapter 2.2. It is important to understand these differences in order to be able to develop selective inhibitors that may address the eubacterial *Shigella* TGT but not the eukaryotic human TGT.

**Tab. 2.1** Sequence alignment of important QueTGT and ArcTGT residues from *Z. mobilis*, *E. coli*, *Homo sapiens*, *Saccharomyces cervisiae*, *Pyrococcus horikoshii* and *Archaeoglobus fulgidus*

| **Species** | **Guanine binding site** | | | | | | | | **Zinc binding site** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **102** | **104** | **106** | **156** | **203** | **230** | **231** | **260** | **318** | **320** | **323** | **349** | |
| *Zymomonas* | D | G | Y | D | Q | G | L | M | C | C | C | H | **QueTGT (eub.)** |
| *Escherichia* | D | G | F | D | Q | G | L | M | C | C | C | H | |
| *Homo* | D | G | F | D | Q | G | L | M | C | C | C | H | **QueTGT (euk.)** |
| *Saccharomyc.* | D | G | F | D | Q | G | L | M | C | C | C | H | |
| *Pyrococcus* | D | S | F | D | Q | G | V | F | C | C | C | H | **ArcTGT** |
| *Archaeoglobus* | D | S | F | D | Q | G | V | F | C | C | C | H | |

| | **Ribose 15/34 binding site** | | | | | | | **Substrate specificity** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **45** | **68** | **70** | **107** | **258** | **261** | **280** | **158** | **232** | **233** | **234** | **235** | |
| *Zymomonas* | V | L | N | Q | Y | G | D | C | A | V | G | E | **QueTGT (eub.)** |
| *Escherichia* | V | L | N | Q | Y | G | D | C | A | V | G | E | |
| *Homo* | V | L | N | Q | Y | G | D | V | S | G | G | E | **QueTGT (euk.)** |
| *Saccharomyc.* | V | L | N | Q | Y | G | D | V | S | G | G | E | |
| *Pyrococcus* | V | I | N | Q | H | G | D | P | V | P | L | M | **ArcTGT** |
| *Archaeoglobus* | V | I | N | Q | H | G | D | P | V | P | L | M | |

**Tab. 2.2** Substrate specificity and promiscuity in TGTs

| **Enzyme** | **Guanine**[a] | **$preQ_0$**[a] | **$preQ_1$**[a] | **Queuine**[a] |
|---|---|---|---|---|
| Archaebacterial TGT | +[b] | + | - | - |
| Eubacterial TGT | + | + | + | - |
| Eukaryotic TGT | + | + | + | + |

[a] for chemical formulae see Figure 2.1

[b] +: is accepted as substrate – : is not accepted

## 2.1.2 Eubacterial QueTGT

### 2.1.2.1 Introduction into the tRNA – QueTGT complex

As in other enzymes with this fold the active site of TGT is located in the C-terminal centre of the TIM-barrel structure (Fig. 2.3a). Upon binding of tRNA multiple interactions are formed with residues from the 'zinc binding' site [Xie *et al*., 2003][4]. The tRNA trinucleotide sequence, $U_{33}G_{34}U_{35}$, which is strictly conserved in all Q-specific tRNAs is specifically recognized (Fig. 2.3b). $U_{33}$ and $U_{35}$ are forming polar interactions *via* functional groups of the uracil base. As $U_{33}$ is present in all tRNAs, in particular the amino acids interacting with $U_{35}$ are highly conserved to guarantee specific recognition. In contrast to the flat binding region of $U_{33}$, $U_{35}$ is located in a buried subpocket [Xie *et al*., 2003][4].

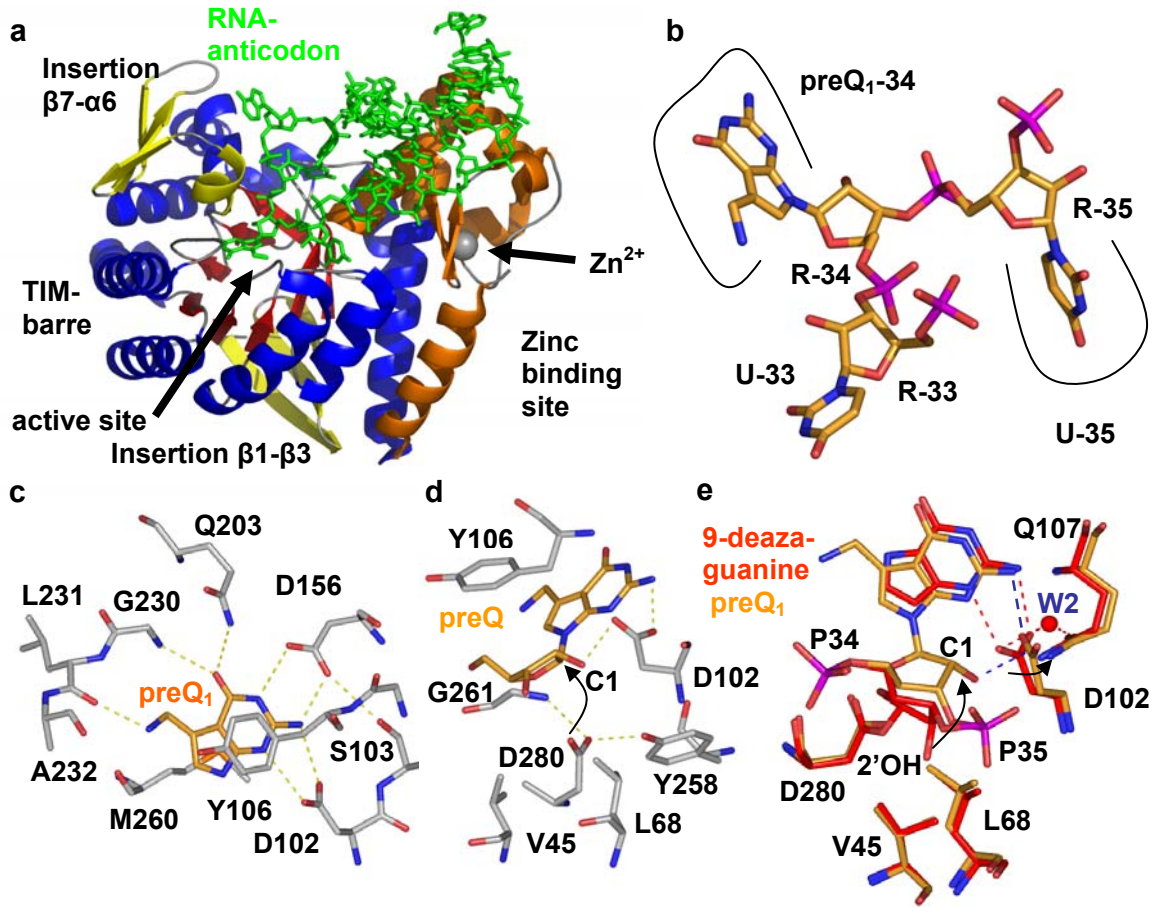



**Fig. 2.3** Substrate recognition by TGT: a) anticodon recognition; b) UGU recognition sequence; c) TGT·preQ$_1$; d) nucleophilic attack by Asp280; e) hydrophobic subpocket accepting the ribose(34) 2'OH-group

The residues involved in $G_{34}$ / $preQ_1$ binding can be subdivided into two groups. The residues contributing to the recognition of the guanine-like skeleton of QueTGT substrates are Asp102, Ser103, Asp156, Gln203 and Gly230 and Leu231. The residues Tyr106 and Met260 perform a sandwich-type hydrophobic stacking with the base of the substrate (Fig. 2.3c). Asp280 is the nucleophile of the base exchange reaction. It is well stabilized by H-bonds formed with Tyr258 and Gly261 (Fig. 2.3d). The nucleophile is located adjacent to a small hydrophobic subpocket formed by Val45 and Leu68 which is important for the stabilization of an intermediate conformation during the base exchange reaction (Fig. 2.3e).

#### 2.1.2.2 New model for the base exchange mechanism in QueTGT

The base exchange catalyzed by TGT follows a ping-pong reaction mechanism resulting in the irreversible incorporation of $preQ_1$ [Goodenough-Lashua & Garcia, 2003][53]. In a first step, tRNA binds to TGT and $G_{34}$ is cleaved off the tRNA. In the intermediate reaction state tRNA ribose 34 is covalently bound to TGT. In a second, reverse reaction step, $preQ_1$ replaces $G_{34}$ in the active site and is incorporated into the tRNA (Fig. 2.4).

The comparison of available crystal structures of *Z. mobilis* TGT in complex with tRNA substrate, $preQ_1$ and a guanine-type inhibitor allowed to deduce single steps of the catalytic reaction. They enabled to develop a detailed, new functional model for the course of the base exchange reaction. Previous assumptions, discussing Asp102 as the catalytic nucleophile are corrected and functional roles are assigned for residues that were previously unconsidered [Stengl *et al*., 2005][54].

Following the reaction pathway TGT recognizes the tRNA substrate via the ‘zinc binding’ site and $G_{34}$ is specifically recognized in the active site by Asp102, Asp156, Gln203 and Gly230 (Fig. 2.4a). A polar contact towards the peptide NH group of Ala 232 is mediated via a water molecule (W1). $G_{34}$ is buried in the active site *via* hydrophobic stacking interactions with Met260 and Tyr106.

Asp280, instead of the previously assumed Asp102, located adjacent to the ribose ring 34, acts as catalytic nucleophile. This residue is well kept in position by Tyr258 and Gly261. Tyr258 itself is arrested and kept in position through several hydrophobic interactions (Met43, Leu100, Met153, Phe199, Met260, Met278). Asp280, Gly261 and Tyr258 as well as its neighbouring hydrophobic residues are

conserved in eubacterial and eukaryotic TGTs emphasizing their particular role to guarantee accurate adjustment of the nucleophile [Xie *et al*., 2003][4].

The Asp280 carboxylic oxygen attacks the C1 carbon of ribose 34 in an $S_N2$ reaction and pulls the ribose ring towards Asp280 [Xie *et al*., 2003][4]. The ribose performs a 40° rotational movement anchored by adjacent ribose phosphate groups $P_{34}$ and $P_{35}$ (Fig. 2.3e). The rotation causing the rupture of the covalent bond between C1 and $G_{34}$ is controlled by Asp102 with its carboxy group H-bonded to the 2'OH-group of the rotating ribose. The 2'OH group is finally released towards a hydrophobic cleft formed by Val45 and Leu68 [Stengl *et al*., 2005][54]. In this orientation the polar group cannot form any H-bond to the enzyme and experiences only weak and rather unfavourable interactions (Fig. 2.3e). Supposedly, the unfavourable intermediate occupancy of the hydrophobic pocket through the 2'OH group serves as a kind of tense spring state and stores energy for the conformational movements required for the reverse reaction step during the ping-pong reaction pathway. Thus, this geometry guarantees a sterically favoured but electrostatically unfavoured intermediate state orientation.

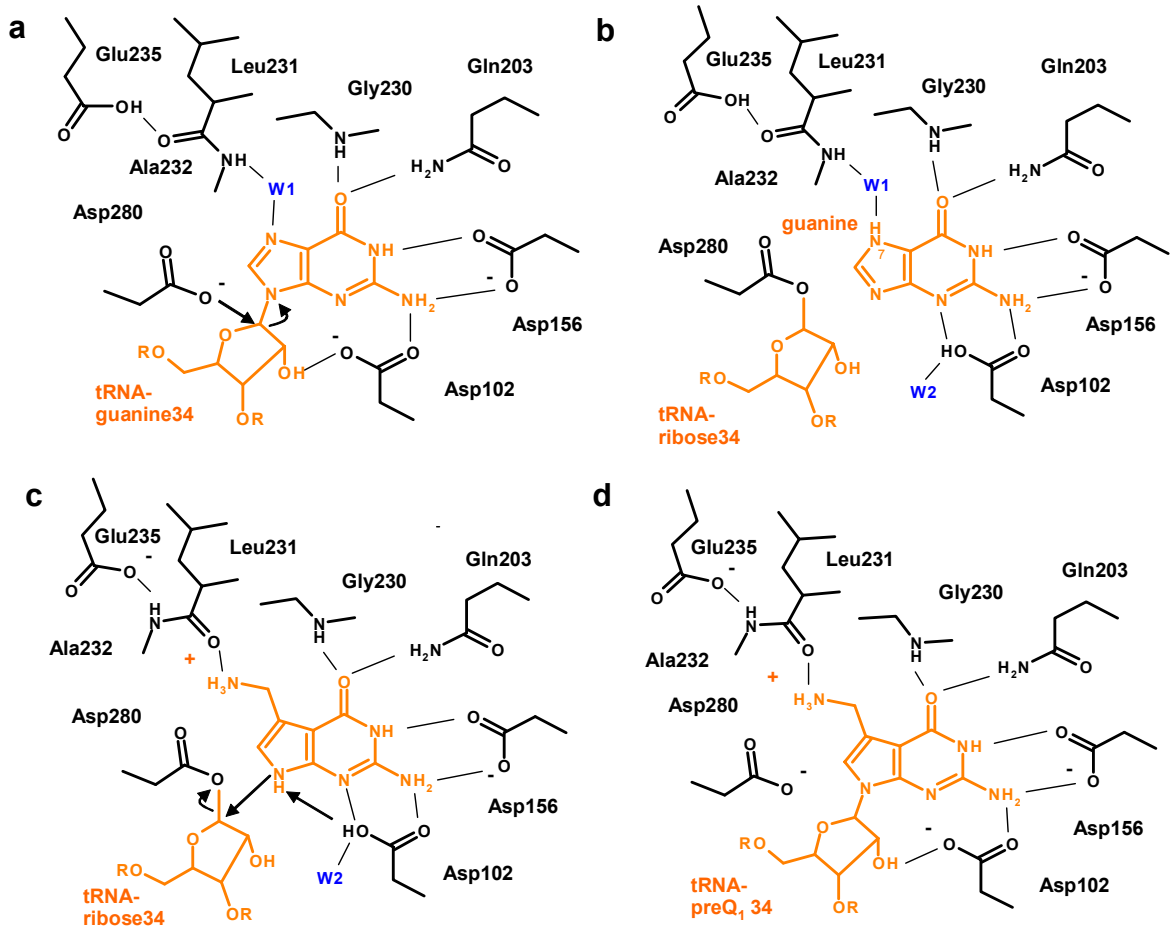



**Fig. 2.4** Base exchange mechanism in eubacterial QueTGT

Guanine is reprotonated in the binding pocket after cleavage from tRNA, either in position 7 or in position 9. Reprotonation in position 7 could be supported by water molecule W1, which is then forming a hydroxy-anion (Fig. 2.4b). The fact that the methyl group of Ala232 in the amide-exposing conformation reduces the available space of the binding pocket and perfectly shields this water molecule, speaks in favour of this position [Stengl *et al*., 2005][54]. Alternatively, reprotonation in position 9 is possible as well, then supported by Asp102 serving as general acid. Another water molecule, W2, bridging Asp102 with Gln107 at the upper rim of the binding pocket would be suited to shuffle a proton into the active site [Xie *et al*., 2003][4].

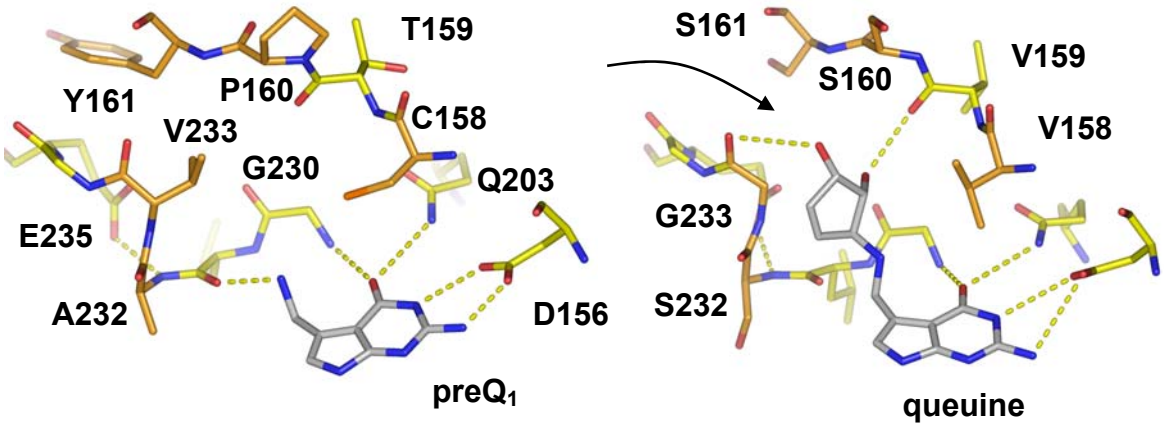
After reprotonation guanine and W1/$OH^-$ leave the binding pocket and are replaced by $preQ_1$. To be able to accommodate $preQ_1$, the Leu231-Ala232 peptide bond supposedly undergoes a flip which is controlled by Glu235 acting as acid/base system (Fig. 2.4b-c) [Stengl *et al*., 2005][54]. $PreQ_1$ is bound in a similar fashion as guanine except of the aminomethyl group that directly interacts with the flipped Leu231 carbonyl oxygen. For activation $preQ_1$ is deprotonated in position 9 by Asp102 then acting as a general base. The proton is shuffled from the active site supposedly mediated by water molecule W2. In a reverse $S_N2$ reaction step the now activated $preQ_1$ nucleophilically attacks the C1 carbon of ribose 34. Upon product formation the covalent bond towards Asp280 is cleaved and the ribose 34 2'OH group is pushed out of its unfavourable environment, stabilized via H-bond formation with Asp102 (Fig. 2.4c-d). In this process the Asp102 carboxy group rotates and expels W2 from the active site followed by the final release of the $preQ_1$-modified tRNA from TGT. Accordingly, Asp102 in addition to its function as general acid/base imposes a strong directional driving force thus controlling the structural changes upon product formation. Finally, the modified tRNA is released from the binding pocket and TGT is ready for a new base exchange cycle.

### 2.1.3 Eukaryotic QueTGT

Eukaryotic QueTGTs show a high sequence identity compared to eubacterial TGTs; e.g. 43% between *H. sapiens* and *Z. mobilis* [Deshpande & Katze, 2001][61]. Concerning the overall sequence composition they differ from eubacterial QueTGTs by a conserved 18 amino acid C-terminal extension. With respect to the active site, all residues involved in catalysis are highly conserved (Tab. 2.1). Thus, it can be

assumed that the overall tertiary structure as well as the course of the base exchange reaction is identical to eubacterial QueTGTs.

Nevertheless, eukaryotic QueTGT exhibits an extended substrate specificity (Tab. 2.2). The preferred substrate is the spatially expanded base queuine, in addition to $preQ_0$ and $preQ_1$. However, Eukaryota are not capable of queuine synthesis and acquire this base by specific uptake. Therefore, the two Q-precursors should not be present in eukaryotic cells. The further extended substrate specificity towards queuine results from a spatial extension of the binding pocket. A homology model based on the *C. elegans* sequence suggested that the replacement of Val233 from eubacterial TGT to Gly233 in eukaryotic TGT significantly enlarges the binding pocket (Tab. 2.1). This allows the binding of extended $preQ_1$-type substrates such as queuine [Romier *et al*., 1997][55].




a b

**Fig. 2.5** a) Specificity pocket in eubacterial *Z. mobilis* QueTGT with bound $preQ_1$; b) homology model of human eukaryotic QueTGT bound to queuine; Q was modelled into the binding pocket with MOLOC (Eubacteria and Eukaryota specific residues in orange)

A more recent homology model of the human TGT based on the structure of *Z. mobilis* TGT was produced using MODELLER with the help of Andreas Evers (group of Prof. Klebe, University of Marburg). It suggests a set of residues involved in the modification of the QueTGT binding pocket. Presumably, a Pro160Ser and Tyr161Ser exchange facilitate the entry of the queuine pentenyl moiety into the binding pocket, due to an extended and more flexible entry site. The exchange of Val233Gly extends the volume of the binding pocket allowing the accommodation of the pentenyl moiety. Finally, the Cys158Val / Ala232Ser modifications might fully

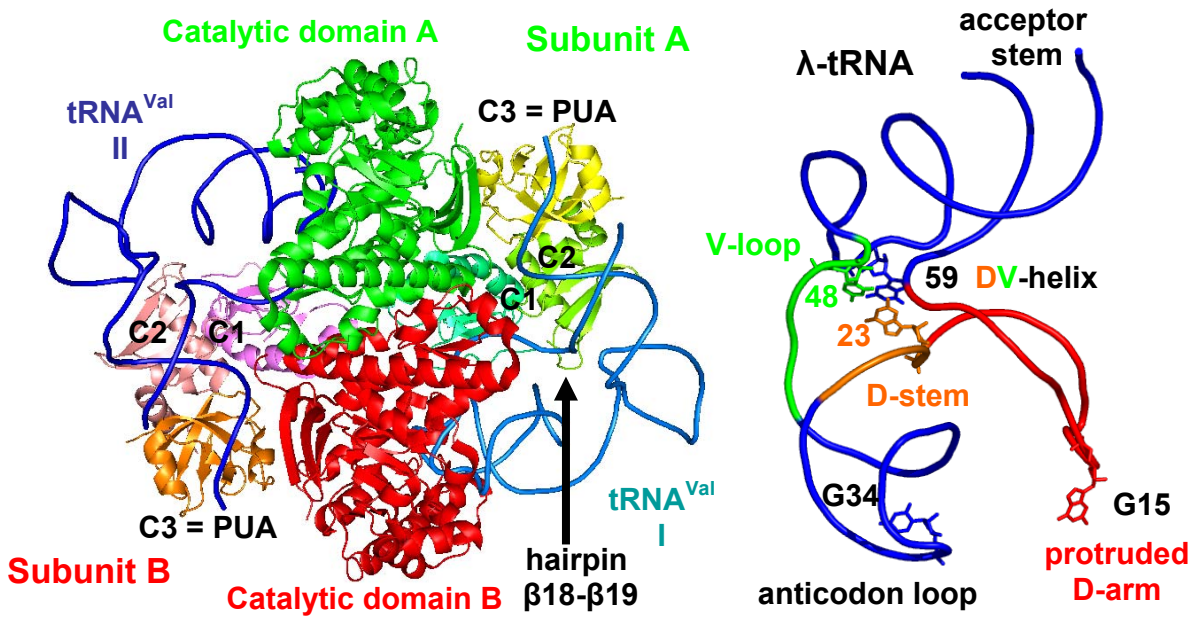
adapt the binding pocket for queuine recognition, indicated by minimization of Q into the modelled binding pocket of human TGT (Fig. 2.5 and Tab. 2.1).
In literature for the eukaryotic QueTGT hetero complex formation with a 60kDa protein is described. However, the nature of this complex remains obscure. The protein belongs to the family of ubiquitin-specific processing proteases (USB). The expression rate of this protein is elevated in leukemic and colon cancer cells. Whether there is any regulatory relationship QueTGT is involved in, still has to be elucidated [Ishiwata, 2004].

## 2.1.4 Archaebacterial ArcTGT

### 2.1.4.1 Introduction into the tRNA – ArcTGT complex

The catalytic domains of archaebacterial ArcTGTs share only about 20 – 25 % sequence identity with eubacterial QueTGTs [Romier *et al.*, 1997][55]. Nevertheless, the tertiary structure is highly conserved (Fig. 2.1). In addition to the catalytic domain, ArcTGT is characterized by three supplementary C-terminal domains required to address the deviating tRNA modification site (Fig. 2.1 and 2.6a) [Ishitani *et al.*, 2002][47]. In order to modify the buried $G_{15}$ residue the tRNA conformation has to undergo an enormous rearrangement [Ishitani *et al.*, 2003][48]. Crystal structure analysis of the archaebacterial TGT from *Pyrococcus horikoshii* in complex with $tRNA^{Val}$ showed that this rearrangement produces the so-called λ-shaped tRNA. It exhibits a conformation which has never been observed before. In canonical L-shaped tRNAs $G_{15}$ is positioned in the D-loop. As it is involved in the formation of tertiary stacking interactions with the bases $C_{48}$ and $A_{59}$, it is buried within the tRNA core. In the λ-shaped conformation the usually rigid D-arm protrudes and the D-loop becomes accessible. The conformation is stabilized via a specific helical element that is not found in canonical L-shaped tRNA. This so-called 'DV'-helix is formed by bases from the variable loop and bases normally involved in formation of the D-stem (Fig. 2.6b). As a consequence of this spatial rearrangement, the former position of $G_{15}$ is now occupied by $G_{23}$ at the end of the 'DV'-helix. The observation of this tRNA conformation gave reason to postulate an additional function of the variable loop with respect to tRNA maturation.
In order to enable these conformational changes in tRNA the supplementary C-terminal domains are required. Within the C2-domain the β18-β19 hairpin is

interacting with tRNA core bases after 'DV'-helix formation and supposedly crucial for the stabilization of the λ- conformation. The C-terminal domain C3 represents a PUA (pseudouridine synthase and archaeosine TGT) domain, widespread among RNA binding proteins (Fig. 2.6a) [Ferré-D'Amaré, 2003][62]. Nevertheless, the PUA domain seems not to be fully relevant for tRNA recognition. An ArcTGT mutant with a deletion of the PUA domains maintains reduced catalytic activity [Sabina & Söll, 2006][63].




a b

**Fig. 2.6** a) tRNA stabilization by ArcTGT in λ- conformation requires dimer formation; b) λ- tRNA is stabilized via 'DV'-helix formation

Catalytic functionality of archaebacterial TGTs requires dimer formation, resulting in a 2:2 complex of ArcTGT with tRNA (Fig. 2.6a). Both TGT subunits forming the dimer are involved in the recognition of an attached tRNA substrate molecule. While one TGT subunit recognizes the tRNA and stabilizes the λ- shape via its three *C*-terminal domains, the other subunit catalyzes the base exchange after accommodation of $G_{15}$ in the active site of its catalytic domain.

### 2.1.4.2 New model for the base exchange mechanism in ArcTGT

The catalyzed base exchange reaction in ArcTGT slightly differs form QueTGT. Here, $preQ_0$ is incorporated into tRNA in position 15. Similar to QueTGT also in ArcTGT a trinucleotide sequence is recognized in the surrounding of the active site (Fig. 2.2b).

However, $G_{15}$ addressed by ArcTGTs is not embedded within a conserved sequence motif. In the crystal structure of *P. horikoshii* TGT·tRNA$^{Val}$ complex, binding of $A_{14}$ and $U_{16}$, both flanking $G_{15}$, is mainly achieved through hydrophobic interactions between the bases and amino acid residues lining the substrate binding pocket. This feature obviously permits ArcTGTs a pronounced promiscuity in base recognition. In this respect it should be noted that not in all archaebacterial tRNAs, containing a guanine at position 15, this residue is modified. The structural prerequisite for $G_{15}$ modification in archaebacterial tRNAs is still unknown but is supposedly associated with “DV”-helix formation ability [Ishitani *et al.,* 2003][48].

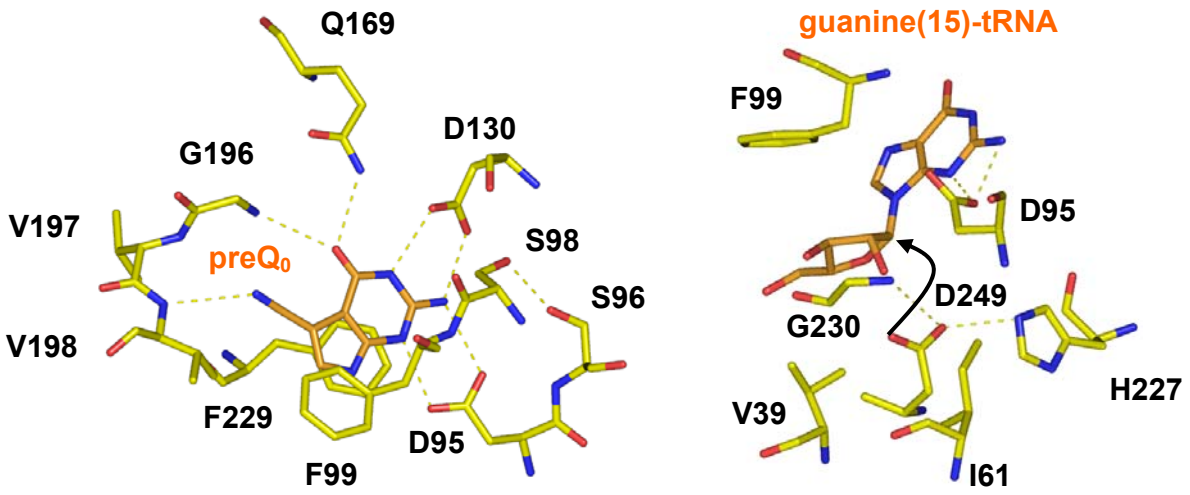
The structure of the active site and the residues responsible for the base exchange reaction in ArcTGT are largely conserved compared to QueTGT (Tab. 2.1). The detailed analysis of ArcTGT and QueTGT crystal structures revealed many similarities but also some differences in the observed substrate binding modes [Stengl *et al*., 2005][54]. Following the numbering of *P. horikoshii* TGT (residue$^{A}$) Asp95$^{A}$, Ser96$^{A}$, Asp130$^{A}$, Gln169$^{A}$ and Gly196$^{A}$ contribute to the recognition of the guanine-type skeleton of preQ$_0$ (Fig. 2.7a). Phe229$^{A}$, corresponding to Met260$^{Q}$ in QueTGT (residue$^{Q}$), hydrophobically stacks with the base of the substrate. Phe99$^{A}$ corresponding to Tyr106$^{Q}$ cannot perform a similar stacking onto the substrate base in *P. horikoshii* TGT. This is caused by the peptide backbone of Phe99$^{A}$ which is slightly shifted off from the recognition base. This difference in recognition results from the most remarkable difference between ArcTGT and QueTGT guanine binding pockets, namely a Gly105$^{Q}$ / Ser98$^{A}$ exchange (Tab. 2.1). It causes a deviating stabilization pattern of specific binding pocket residues. While in *Z. mobilis* TGT Ser103$^{Q}$ is H-bonded to Asp156$^{Q}$ (Fig. 2.3c), in *P. horikoshii* TGT the Ser96$^{A}$ side chain is H-bonded to Ser98$^{A}$ within the same loop resulting in a deviating loop geometry and subsequently in an altered substrate recognition pattern (Fig. 2.7a).

Due to the fact that no crystal structure of a covalent intermediate with an ArcTGT is available, residues likely to be important for catalysis can only be suggested taking reference to QueTGT [Stengl *et al*., 2005][54]. Asp249$^{A}$ is in an equivalent position to Asp280$^{Q}$ and supposedly performs the nucleophilic attack onto the ribose carbon. An aspartic acid in this position is conserved in the TGTs of all kingdoms (Tab 2.1). Similar to QueTGT, in the available ArcTGT structures the side chain position of Asp249 is stabilized via H-bonds with two adjacent residues (Fig. 2.7b): On one side to His227$^{A}$, a residue only conserved in archaebacteria, on the other side via the

backbone NH of Gly230$^{A}$. In QueTGT the conserved Tyr258$^{Q}$ occupies the equivalent position of His227$^{A}$. Gly230$^{A}$ is structurally conserved in all three kingdoms and corresponds to Gly261$^{Q}$. While in *Z. mobilis* TGT Asp280$^{Q}$ is firmly clamped by Tyr258$^{Q}$ and Gly261$^{Q}$ (Fig. 2.3d), in *P. horikoshii* TGT the Asp249$^{A}$ side chain distances of the carboxy group towards the two residues His227$^{A}$ and Gly230$^{A}$ fall into a range of 2.5 - 3.5Å depending on the presence of the bound tRNA substrate.

The tRNA-ribose 15 moiety is able to place its 2'OH into a hydrophobic pocket of ArcTGTs in a similar way as observed for tRNA-ribose 34 when bound to QueTGT. In ArcTGTs, this pocket is composed by the conserved residues Val39$^{A}$ and Ile61$^{A}$, corresponding to Val45$^{Q}$ and Leu68$^{Q}$ in *Z. mobilis* TGT (Fig. 2.7b and 2.3d).

The binding mode of Asp95$^{A}$ in ArcTGT, a further residue conserved across all kingdoms, is equivalent to Asp102$^{Q}$ in QueTGT, but adopts a slightly different geometry. It does not form an H-bond to the 2'OH group in the complex with bound tRNA.




a b

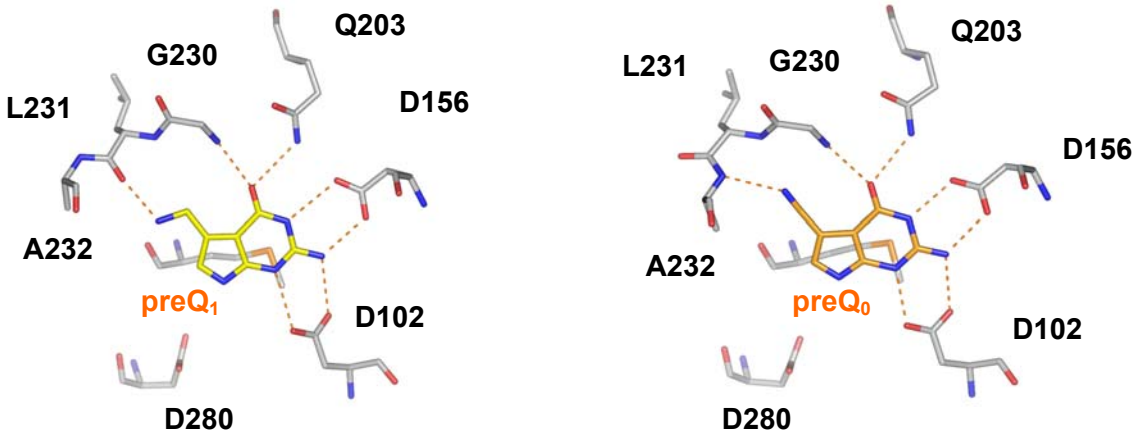
**Fig. 2.7** a) ArcTGT·preQ$_0$ binding site; b) residues assumed to be involved in the base exchange mechanism

Summarizing the observations all features necessary to perform the base exchange reaction are similarly exhibited in both, QueTGT and ArcTGT. Hence, an evolutionarily highly conserved mechanism must be assumed originating from an ancient ancestor already existing before the separation of the three kingdoms (see chapter 2.4).

# 2.2 QueTGT – ArcTGT: substrate specificity

## 2.2.1 QueTGT – ArcTGT: regulation of substrate specificity

Eubacterial QueTGT exhibits to some degree substrate promiscuity (Tab. 2.2). In addition to the natural substrate $preQ_1$ it also accepts its biosynthetic precursor $preQ_0$. *Z. mobilis* TGT crystal structures in complex with $preQ_0$ and $preQ_1$ clearly demonstrate the capability of both bases to bind to this TGT [Brenk *et al.*, 2003][64]. Substrate promiscuity results from a functional and sterical adaptation of the binding pocket caused by the reorientation of the Leu231/Ala232 peptide bond (Fig. 2.8).




a b

**Fig. 2.8** a) *Z. mobilis* TGT binding to $preQ_1$; b) *Z. mobilis* TGT binding to $preQ_0$

From the analyses of several *Z. mobilis* TGT crystal structures in complex with different substrates a structural explanation for the observed extended substrate specificity can be provided [Stengl *et al.*, 2005][54]. During the base exchange reaction replacement of guanine by $preQ_1$ requires an adjustment of the binding pocket geometry (chapter 2.1.2.2 and Fig. 2.4). To achieve this, the Leu231/Ala232 peptide bond undergoes a flipping movement. Binding of guanine(34)-tRNA results in the W1 mediated contact to the amide group of Ala232. As the available tRNA-bound TGT crystal structure has no sufficient resolution this structural features has been extracted from the pyridazindione-type inhibitor (**H6**) bound crystal structure (see chapter 3.2.1 and 3.2.2) [Brenk *et al.*, 2003][2]. This inhibitor includes the key structural elements of guanine. Upon binding of $preQ_1$ the properties of the binding pocket are

altered. Proper recognition of the aminomethyl group of $preQ_1$ requires the NH donor functionality to be exchanged by the CO acceptor functionality via rearrangement of the Leu231/Ala232 peptide bond switch. Furthermore, the new orientation of the methyl group of Ala232 extends the size of the binding pocket allowing to accommodate the kinked aminomethyl group of $preQ_1$.

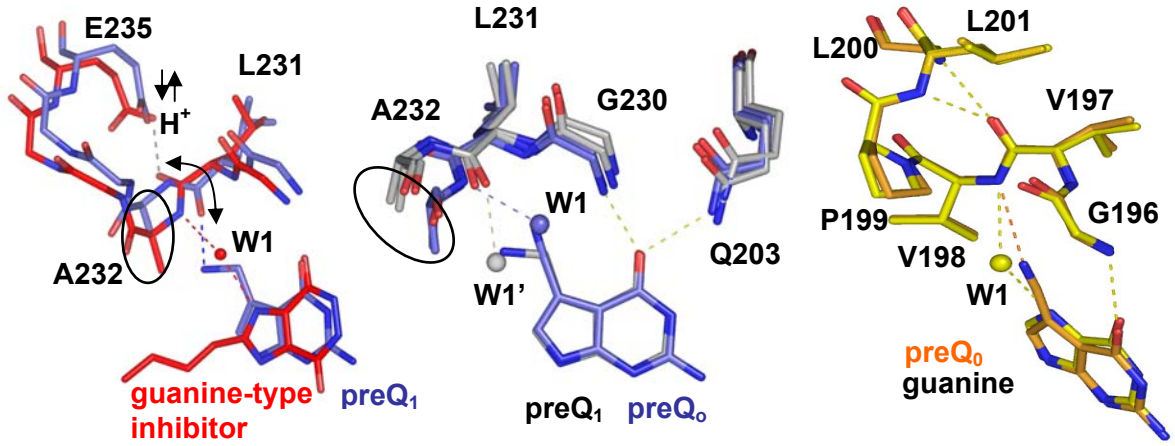



**Fig. 2.9** a) Leu231/Ala232 peptide switch, allowing subsequent binding of guanine and $preQ_1$ is controlled by Glu235; b) water molecules in uncomplexed TGT structures at pH 5.5 (blue) and pH 8.5 (grey) indicate two spatially deviating interaction sites favourable to accompdate polar groups of $preQ_0$ and $preQ_1$; c) substrate recognition in ArcTGT·guanine and ArcTGT·$preQ_0$

The peptide switch is stabilized from the side opponent to the binding pocket by the side chain carboxy group of Glu235, which is strictly conserved in all QueTGTs [Stengl *et al*,. 2005][54]. Depending on the protonation state of this carboxy group, Glu235 either donates an H-bond towards the backwards exposed carbonyl group or accepts an H-bond from to the amide group of the reoriented peptide bond thus serving as a general acid/base mediating the peptide switch (Fig. 2.9a). As expected for such a general acid/base, the peptide switch gated by Glu235, can already be triggered in uncomplexed QueTGT depending on the applied pH conditions (Fig. 2.9b). In *Z. mobilis* TGT, crystallized at pH 5.5, Glu235 is protonated and hydrogen-bonds the CO group of the peptide bond. In consequence, the NH group is exposed towards the binding pocket. At pH 8.5, Glu235 experiences deprotonation which triggers the switch of the peptide bond now binding the Glu235 carboxylate *via* its amide group. Accordingly, the CO group is exposed towards the binding pocket.

Associated with the peptide flip in the uncomplexed crystal structures two water molecules (W1 and W1') are bound to either the carbonyl- or the amide-group exposed to the binding pocket. They are found at two deviating, clearly distinct positions and indicate within a narrow spatial range the favourable, putative binding sites of $preQ_1$, $preQ_0$ or the W1 assisted binding mode of the pyridazindione-inhibitor **H6** (Fig. 2.9b). From these findings it can be concluded that the binding pocket conformation suited for guanine binding additionally allows the binding or $preQ_0$ by replacing W1, thus explaining the extended substrate specificity of QueTGT.

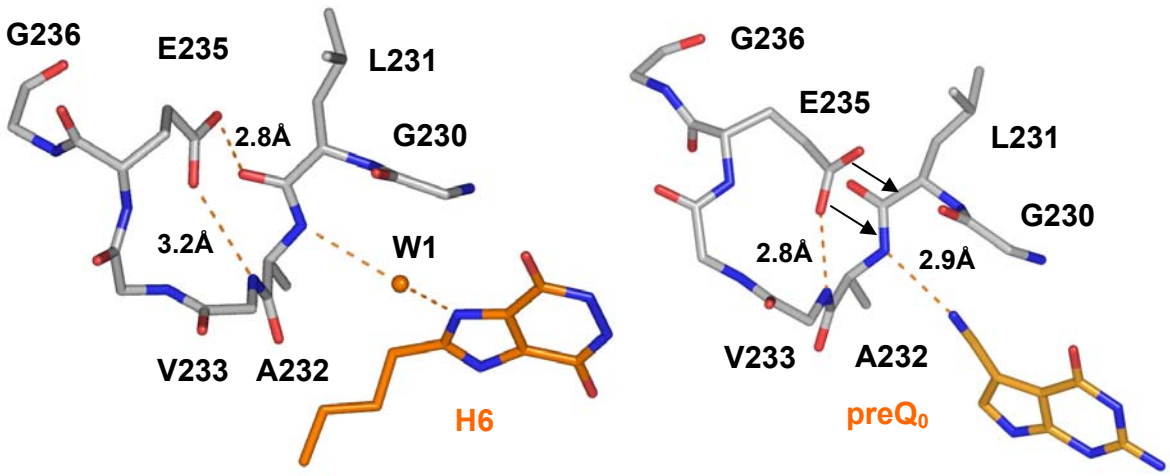
In contrast to QueTGT, ArcTGT features an exclusive substrate specificity towards guanine and $preQ_0$, (Tab. 2.2). This results from a significantly reduced adaptability of the binding pocket [Brenk *et al*., 2003][64]. The peptide bond corresponding to $Leu231^Q$/$Ala232^Q$ which performs the peptide flip in QueTGTs is represented by $Val197^A$/$Val198^A$ in ArcTGT and presents an amide donor group towards the substrate (Tab. 2.1 and Fig. 2.9c). At its back side, this peptide bond lacks a group gating a peptide switch, as observed in QueTGTs, due to an entirely different stabilization geometry. In *P. horikoshii* TGT, the two invariant backbone NH-bonds of $Leu200^A$ and $Leu201^A$, which will not support a peptide switch as mediated by $Glu235^Q$, are present in the second sphere of amino acids around the active site (Fig. 2.9c). Additionally, $Pro199^A$ increases the rigidity of this loop geometry. For ArcTGTs the peptide switching functionality is not required. The interstitial water molecule which in the binary *P. horikoshii* TGT·guanine complex bridges guanine with the $Val197^A$/$Val198^A$ peptidic NH group, can directly be replaced by the acceptor nitrile group of $preQ_0$. This binding mode is similar to that of $preQ_0$ in QueTGT (Fig. 2.9b/c).

Apart from the fact that Archaebacteria are not capable to synthesize $preQ_1$, incorporation of this base into tRNA by ArcTGT is also not observed under experimental conditions. From the analysis of the ArcTGT crystal structures two explanations can be provided. Firstly, with the invariant NH donor group no functionality suited for binding is exposed to the binding pocket. Secondly, the hydrophobic side chain of $Val198^A$ is limiting the available space of the binding pocket and disables the binding of the kinked $preQ_1$ amino-methyl group. This provides a conclusive explanation why ArcTGT exhibits no substrate promiscuity similarly to QueTGT.

## 2.2.2 QueTGT substrate selectivity – TGT(E235Q) mutant

### 2.2.2.1 Introduction

The eubacterial QueTGT exhibits the above-described substrate promiscuity. QueTGT accepts, additionally to the natural substrate $preQ_1$, also its biosynthetic precursor $preQ_0$. Promiscuity results from the Leu231/Ala232 peptide switch gated from the opponent side of the binding pocket by the general acid/base Glu235 (see chapter 2.2.1). $PreQ_0$ and $preQ_1$ are simultaneously present in the cell. Therefore, the eubacterial QueTGT has to distinguish between the two substrates to guarantee the selective incorporation of $preQ_1$ into tRNA. In *E. coli TGT* $K_m$ values for $preQ_1$ and $preQ_0$ have been determined [Hoops *et al*., 1995][39]. The $K_m$ of $preQ_1$ (0.4 µM) is six fold lower than that of $preQ_0$ (2.4 µM). For *Z. mobilis* TGT no kinetic data have yet been determined.




a b

**Fig. 2.10** a) *Z. mobilis* TGT binding to **H6**; Glu235 is H-bonded to the peptide switch; b) *Z. mobilis* TGT binding to $preQ_0$; Glu235 stacks on top of the hydrophobic face of the peptide switch (for Val233 only the backbone is shown)

The $preQ_0$ and $preQ_1$ bound crystal structures of *Z. mobilis* TGT provide no clear explanation for the selective incorporation of $preQ_1$ [Brenk *et al*., 2003][64]. For binding of $preQ_1$ charge assistance might be relevant. The positively charged amino-methyl group of $preQ_1$, binding to the carbonyl group of the Leu231/Ala232 peptide switch, could compensate the negative charge of Glu235. In the latter TGT·$preQ_1$ complex, the carbonic acid functional group of Glu235 is in close contact to the amide group of

Ala232 (2.8 Å) suggesting this functional group to be deprotonated (Fig. 2.9a). In the crystal structure of TGT·preQ$_0$ the acceptor nitrile group of preQ$_0$ is H-bonded to the amide group of Ala232. However, an unexpected and unique orientation of the terminal acid functional group of Glu235 is observed. The Glu235 side chain stacks with its hydrophobic surface on top of the Leu231/Ala232 peptide bond. Additionally, the side chain shift provokes the Glu235/Gly236 peptide bond to flip [Brenk *et al*., 2003][64]. This new geometry is stabilized by an H-bond to the amide of Val233 (Fig. 2.10b). The short distance of 2.8 Å for this H-bond suggests deprotonation and charged state of the carboxylate group of the Glu235 side chain in this orientation. In all other available TGT crystal structures with the amide group of Ala232 exposed to the binding pocket, the side chain of Glu235 is binding to the carbonyl group of Leu231 and it occurs most presumably in its protonated state. This is suggested by the interaction geometry observed in the TGT·**H6** complex [Brenk *et* al., 2003][2]. The Glu235 side chain is in short contact with the carbonyl group of Leu231 (2.8 Å) and forms a longer interaction (3.2 Å) with the amide group of Val233 (Fig. 2.10a). This binding mode supports the assumption that the carbonic acid functionality is present in its protonated state.

Due to the missing kinetic data for *Z. mobilis* TGT the influence of these structural modifications on the binding of preQ$_0$ and preQ$_1$ remains unresolved. Therefore, the kinetic parameters for both substrates were determined to estimate the influence of the observed binding pocket conformations on selectivity in this species. Additionally, a TGT(E235Q) mutated enzyme was constructed and crystallized in complex with preQ$_0$ and preQ$_1$. As the terminal amide functionality of Gln235 will be permanently present in a protonated state it should stabilize and firmly fix the peptide switch by forming a hydrogen bond to the carbonyl group of Leu231. Thus, it should arrest the peptide switch it the amide-exposing conformation, independent of applied pH conditions or substrates. Selectivity is expected to be modulated in favour of preQ$_0$ binding.

#### 2.2.2.2 Results

The E235Q mutated *Z. mobilis* TGT was constructed by site-directed mutagenesis. Details for the construction are given in chapter 5.2.4. The purification of TGT(E235Q) followed the standard protocol for wild type TGT (chapter 5.2.5). The mutated TGT(E235Q) enzyme crystallized under similar conditions and in the same

space group as the wild type. For TGT(w.t.) structures of crystals grown at pH 5.5 and 8.5 are available. Therefore, crystallization attempts were performed for TGT(E235) at different pH conditions as well (for conditions see chapter 5.5.1). Crystals of the uncomplexed TGT(E235Q), suited for X-ray analysis, grew at two pH values (pH 6 and pH 8.5). In contrast to TGT(w.t.) crystals at pH 5.5 could not be obtained. At pH 8.5 the crystal structure of uncomplexed TGT(E235Q) was determined to a maximum resolution of 1.55Å. The structure of crystals grown at pH 6 was resolved to a maximum resolution of 1.57Å. Crystallographic data are given in Table 6.2.1. Details of data collection and structure solution are given in chapter 5.5.2 and 5.5.3. In both crystal structures the residues in the vicinity of the active site, including the mutated Gln235, are well defined and adopt virtually identical orientations. The amide group of Ala232 is exposed to the binding pocket. The terminal amide group of Gln235 stacks with its hydrophobic surface on top of the Leu231/Ala232 peptide bond. The Glu235 side chain is H-bonded to the amide function of Val233 and the carbonyl function of Gly230. A detailed discussion of the interaction pattern, particularly in comparison with other structures, follows in the next chapter. To investigate the binding modes of $preQ_0$ and $preQ_1$, co-crystallisation experiments of TGT(E235Q) with both bases were performed at pH 6 (for conditions see chapter 5.5.1). Applying lower pH values in cocrystallization attempts with both bases was already successful for the wild type enzyme [Brenk *et al*., 2003]. At this pH, Asp102 rotates into the binding pocket and forms a double hydrogen bond with the substrate. Only $preQ_0$ could be successfully cocrystallized with TGT(E235Q) resulting in a crystal structure with a maximum resolution of 1.7 Å. $PreQ_0$ is well defined in the binding pocket. The crystal structure of the respective cocrystallization experiment of TGT(E235Q) with $preQ_1$ was refined to a maximum resolution of 1.6 Å. Although all binding pocket residues are well defined, no electron density of a bound ligand could be identified. The observation of a split conformation for Asp102, resulting from the rotation of the side chain into the binding pocket indicates, however, the limited ability of $preQ_1$ binding. The Asp102 rotation is only observed upon ligand binding and has been described in similar fashion for inhibitor-bound crystal structures [Brenk *et al*., 2004][1]. Supposedly some parts of the crystal have been populated by bound $preQ_1$, however, the population is not sufficient to assign a distinct binding geometry.

Kinetic parameters for the wild type and the mutated TGT(E235Q) have been determined with respect to $preQ_0$, $preQ_1$, guanine and tRNA. For the latter two substrates the determination is based on the monitoring of tritium labelled [8-$^3$H]-guanine which is incorporated in tRNA at position 34 *via* TGT catalysis. After labelling tRNA is precipitated and monitored *via* liquid scintillation counting. The applied procedure is described in chapter 5.3.1/5.3.2 and has recently been published by us in Meyer *et al*. (2006)[65]. For $preQ_0$ and $preQ_1$ this method can not be applied. Therefore, an appropriate enzyme kinetic assay had to be developed. In this assay [8-$^3$H]-guanine-labelled $tRNA^{Tyr}$ was previously produced and then used as substrate for TGT to monitor the base exchange against $preQ_0$ or $preQ_1$. The decreasing level of labelled tRNA can be monitored and from these data kinetic parameters can be derived. Details about the procedure and the production of labelled tRNA are given in chapter 5.3.2.

In Tab. 2.3 the kinetic data for *Z. mobilis* TGT with respect to the four substrates is given. TGT(E235Q) as well as the wild type enzyme incorporate $preQ_0$ and $preQ_1$ into tRNA. For the wild type enzyme the main difference between the two substrates is the turn-over rate. Values for $k_{cat}$ differ by a factor of 10. For TGT(E235Q) a 45 fold decrease in $K_m$ for $preQ_1$ has been determined with respect to the wild type. $k_{cat}$ values, however, remain virtually identical to those of the wild type.

**Tab. 2.3** Kinetic paramteters for TGT(w.t.) and TGT(E235Q)

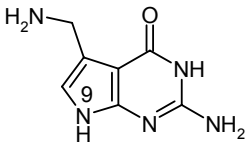
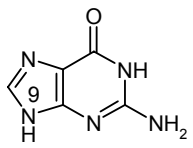
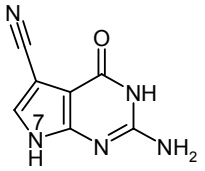
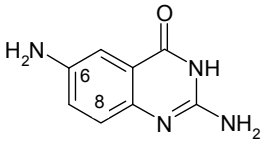
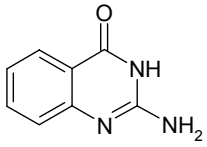
| **TGT (w.t.)** | **$tRNA^{Tyr}$** | **[$^3$H]-guanine** | **$preQ_1$** | **$preQ_0$** |
|---|---|---|---|---|
| $K_m$ [µM] | 0.9 ± 0.2 | 1.2 ± 0.2 | 0.7 ± 0.2 | 0.9 ± 0.2 |
| $k_{cat}$ [$s^{-1}$] | $2.7 \cdot 10^{-2}$ | $2.8 \cdot 10^{-2}$ | $5.1 \cdot 10^{-2}$ | $0.6 \cdot 10^{-2}$ |
| $k_{cat}/K_m$ [$\mu M^{-1}s^{-1}$] | $3.0 \cdot 10^{-2}$ | $2.3 \cdot 10^{-2}$ | $7.3 \cdot 10^{-2}$ | $0.6 \cdot 10^{-2}$ |

| **TGT (E235Q)** | **$tRNA^{Tyr}$** | **[$^3$H]-guanine** | **$preQ_1$** | **$preQ_0$** |
|---|---|---|---|---|
| $K_m$ [µM] | 1.0 ± 0.1 | 3.3 ± 0.3 | 32 ± 7 | < 0.5 |
| $k_{cat}$ [$s^{-1}$] | $3.5 \cdot 10^{-2}$ | $3.8 \cdot 10^{-2}$ | $5.0 \cdot 10^{-2}$ | $0.3 \cdot 10^{-2}$ |
| $k_{cat}/K_m$ [$\mu M^{-1}s^{-1}$] | $3.5 \cdot 10^{-2}$ | $1.2 \cdot 10^{-2}$ | $0.2 \cdot 10^{-2}$ | $> 0.5 \cdot 10^{-2}$ |

### 2.2.2.3 Discussion of the kinetic data

The $K_m$ values for $preQ_0$ and $preQ_1$ show no significant difference (Tab. 2.3). Within the error bars $preQ_0$ ($K_m$ = 0.9 µM) binds equally well as the natural substrate $preQ_1$ ($K_m$ = 0.7 µM). However, a significant difference results for the speed of their incorporation into tRNA. $PreQ_1$ is incorporated almost ten times faster than $preQ_0$. Obviously *Z. mobilis* TGT achieves substrate selectivity not *via* binding affinity but *via* the turn-over rate. The $K_m$ of guanine (1.2 µM) is slightly reduced compared to that of $preQ_1$ and also $k_{cat}$ is only half the value observed for $preQ_1$. However, the latter value refers to the exchange of guanine by tritium-labelled guanine, which is reversible. As the incorporation of $preQ_1$ into tRNA is reported to be irreversible, the backwards base exchange reaction can be neglected for this case [Okada *et al*., 1979][31].

**Tab. 2.4** Comparison of natural substrates and quinazolinone-based inhibitors

| **conformation** | **carbonly Leu231 H-bonded Glu235** | **amide Ala232 H-bonded Glu235** | **amide Ala232 stacking Glu235** |
|---|---|---|---|
| **natural substrates** |  |  |  |
| | **$preQ_1$** | **guanine** | **$preQ_0$** |
| **Km [µM]** | **0.7** | **1.2** | **0.9** |
| **quinazolinone inhibitors** |  |  | |
| | **Q1** | **Q2** | |
| **$K_{iu}$ [µM]** | **0.6** | **1.7** | |

Results from inhibition experiments of quinazolinone-based inhibitors to *Z. mobilis* TGT support the assumption of an approximately two-fold affinity difference for guanine and $preQ_1$ (Tab. 2.4). In the crystal structures of TGT with 6-amino-quinazolinone (**Q1**) the binding pocket conformation is similar to that in the complex with $preQ_1$ [Brenk *et al*., 2004][1]. TGT bound to quinazolinone (**Q2**), however, exhibits a binding pocket similar to the one expected for guanine [Meyer *et al*., 2004][66]. For **Q1** ($K_{iu}$ = 0.6 µM) and **Q2** ($K_{iu}$ = 1.7 µM ) the binding affinity differs by a factor of three [Meyer *et al*., 2006][65]. For details concerning crystal structures and affinity data see chapters 3.2.1 and 3.2.3.

The similar $K_m$ values for $preQ_0$ and $preQ_1$ in *Z. mobilis* TGT indicate virtually identical energy contents for the two observed binding pocket conformations. For *E. coli* TGT a six-fold difference in $K_m$ values for $preQ_1$ (0.4µM) and $preQ_0$ (2.4µM) was reported [Hoops *et al*., 1995][39]. No clear explanation for the affinity differences between both species could be identified. The kinetic data were determined using a method which is comparable to this study.

In the TGT(E235Q) mutated enzyme $K_m$ for $preQ_0$ is virtually identical to that of the wild type (0.9 µM vs. < 0.5 µM). Due to the limited accuracy of the assay this value could not be determined more precisely. In contrast to $preQ_0$, the $K_m$ of $preQ_1$ significantly increases from 0.7 to 32 µM, whereas the values for $tRNA^{Tyr}$ and guanine are not significantly altered. Even though $K_m$ is strongly affected, the $k_{cat}$ values for $preQ_0$ and $preQ_1$ remain virtually identical. Still $preQ_1$ is significantly faster incorporated than $preQ_0$. Both rates differ by approximately a factor of 20. This suggests that substrate recognition is obviously gated by Glu235. However, once accommodated in the active site, the reaction rate of the base exchange is determined by other factors. Regarding $k_{cat}/K_m$ as indicator for the catalytic efficiency, the overall selectivity is inverted in favour of $preQ_0$ for TGT(E235Q).

$k_{cat}$ values for $preQ_1$, guanine and $preQ_0$ are similarly decreasing in TGT(w.t.) and TGT(E235Q). From this finding it can be concluded that the deviating binding modes of the three bases exhibit only minor influence on the turn-over rates. Most presumably the innate ability of these bases to perform a nucleophilic attack is responsible for the observed trend in reactivity. Nitrogen N9 of the respective bases nucleophilically attacks C1 of the covalently bound tRNA ribose 34 (for chemical formulae see Fig. 2.4; for mechanism see chapter 2.1.2.2). N9 in $preQ_1$ is the most potent nucleophile, as it is in no conjugation with the exocylic amino methyl group.

N9 in guanine is less potent due to an electron withdrawing effect of N7 being in conjugation with N9. $PreQ_0$ is the least potent base in this series. Here the exocylcic cyano group is in conjugation with the endocyclic nitrogen, resulting in an even more pronounced electron withdrawing effect.

Surprisingly, in *E. coli* TGT $V_{max}$ for $preQ_0$ (4.2 µM $s^{-1}$ $mg^{-1}$) is virtually identical to that of $preQ_1$ (2.6 µM $s^{-1}$ $mg^{-1}$), although the latter base is the preferred substrate for this TGT [Hoops *et al.*, 1995][39]. However, the comparison with *Z. mobilis* TGT provides no satisfactory explanation for this observation.

#### 2.2.2.4 Discussion of TGT(E235Q) crystal structures

The crystal structures of uncomplexed TGT(E235Q) revealed an unexpected binding mode for the mutated glutamine residue. All residues in the vicinity of this residue are well defined in the crystal structure (Fig. 2.11a). However, Gln235 is not H-bonded to the carbonyl group of Leu231 as expected. Instead, the Gln235 side chain stacks with its hydrophobic surface on top of the Leu231/Ala232 peptide bond. Together with the side chain of Tyr161 it is involved in a sandwich-like stacking interaction (Fig. 2.11b). The Glu235 side chain is additionally stabilized via two H-bonds formed with the amide function of Val233 (3.0 Å) and the carbonyl function of Gly230 (2.9 Å). Concerning the peptide switch, the Leu231/Ala232 peptide bond exposes the Ala232 amide function to the binding pocket stabilizing water molecule (W1).

The superposition with the uncomplexed TGT(w.t) crystal structure derived at pH 5.5 illustrates the differences with respect to the glutamate / glutamine side chain geometry. These differences are likely supported by a Gln235/Gly236 peptide flip. (Fig. 2.11d). While the orientation of Ala232 including W1 is nearly identical in both structures, the 235 side chains adopt significantly different geometries. Although Glu235 is supposed to be protonated at this pH and thus structurally highly isosteric to glutamine, deviating geometries are found in both structures.

The crystal structures of TGT(E235Q) reveal another significant difference to the wild type enzyme. In the TGT(w.t.) the orientation of the Leu231/Ala232 peptide switch can be triggered in a pH dependent manner, regulated by the Glu235 protonation state. The E235Q mutated enzyme, however, permanently seems to expose the amide function of Ala232 towards the binding pocket, at least under the conditions screened in the experiments. The two structures found in crystals grown at pH 6 and

pH 8.5 are virtually identical with respect to the orientation of Leu231 - Gln235 (Fig 2.11c). Thus, conformational rearrangements seem not to be relevant for TGT(E235Q) compared to the wild type (Fig. 2.9a).

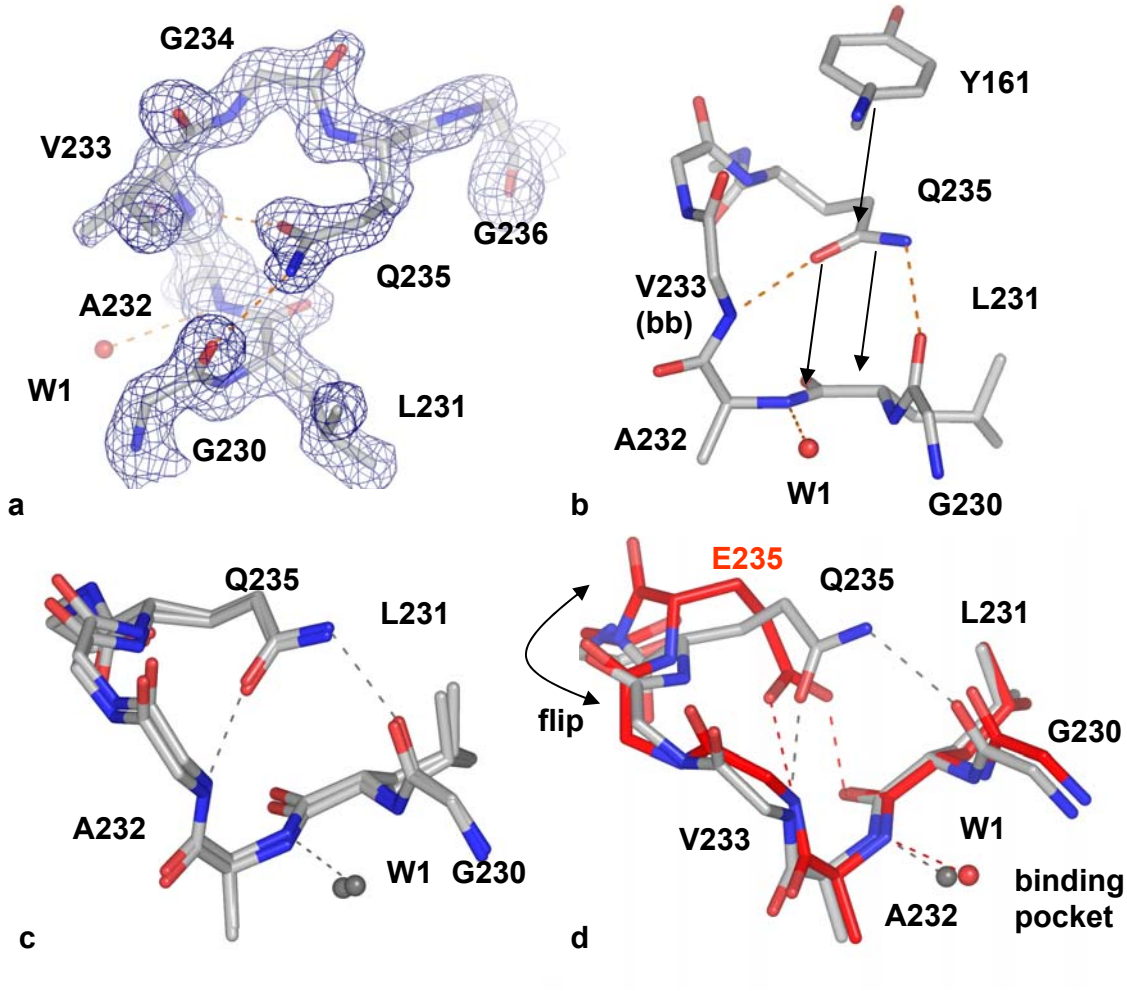



**Fig 2.11** a) Electron density of the G230-G236 loop in apo TGT(E235Q) crystallized at pH 8.5 and contoured at 1.3σ in a 2|Fo| - |Fc| map; b) hydrophobic stacking stabilizing the geometry of Q235; c) superposition of TGT(E235Q) crystallized at pH 8.5 and pH 6; d) superposition of G230 - G236 in apo TGT(E235Q) crystallized at pH 8.5 and apo TGT(w.t.) crystallized at pH 5.5

Cocrystallization of TGT(E235Q) at pH 6 with $preQ_0$ was successful, in contrast to the attempts with $preQ_1$. $PreQ_0$ is well defined in the TGT(E235Q) binding pocket displaying distinct interactions with Asp102, Asp156, Gln203, Gly230 and Ala232 (Fig. 2.12a). The superposition with the uncomplexed crystal structure at pH 6 provides an explanation for this result (Fig. 2.12b). Both binding pocket geometries are virtually identical. Independent of pH, TGT(E235Q) provides a binding pocket suited to accommodate $preQ_0$. Upon binding of this base, the water molecule W1 is

expelled from the binding pocket and Asp102 is rotated towards the ligand. Thus, almost no binding pocket rearrangement is required for binding. This finding is supported by the kinetic analysis. $K_m$ of $preQ_0$ in TGT(E235Q) is low, with some care, supposedly slightly lower than for the wild type.

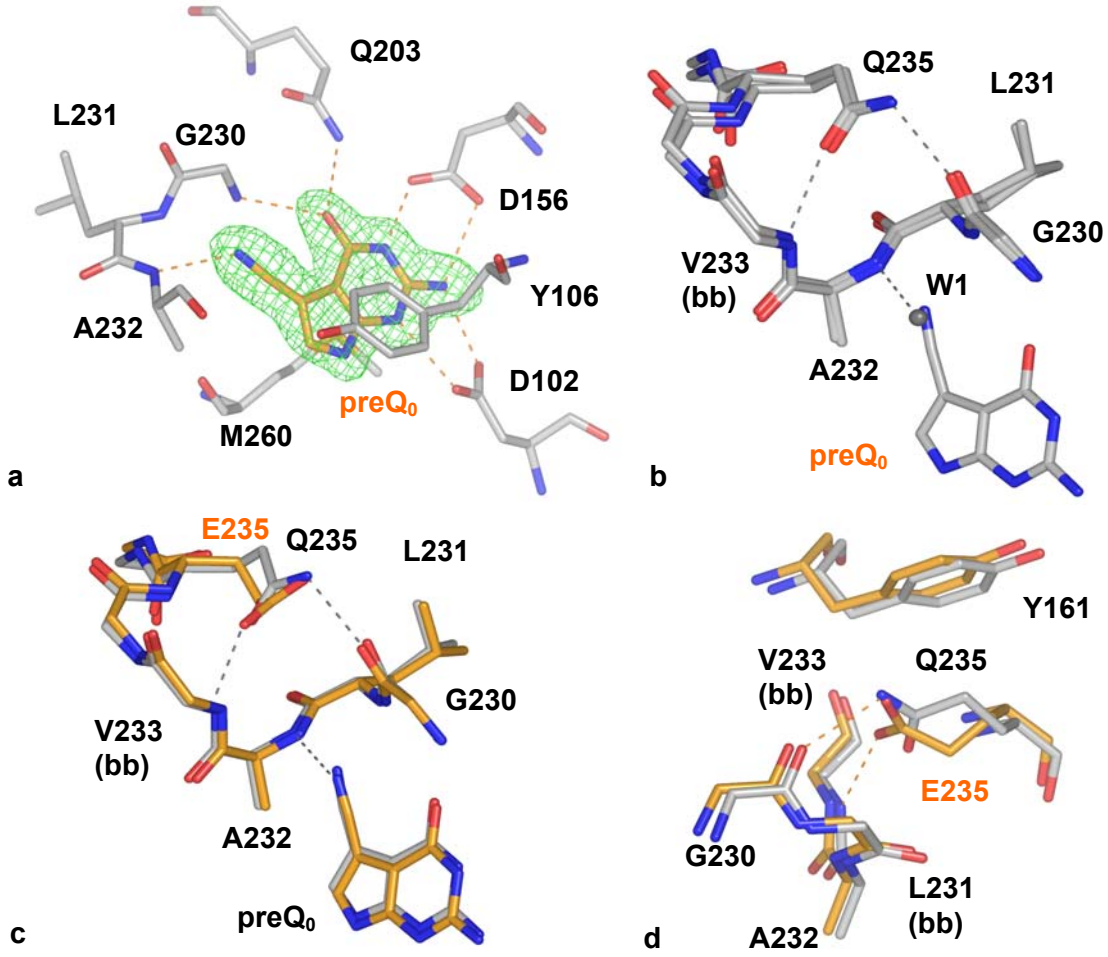



**Fig 2.12** a) Binding mode of $preQ_0$ in TGT(E235Q) contoured at 3σ in the |Fo| - |Fc| map refined in the last cycle excluding ligand coordinates; b) structural alignment of apo TGT(E235Q) crystallized at pH 8.5 and TGT(E235Q)·$preQ_0$; c) structural alignment of TGT(w.t.)·$preQ_0$ and TGT(E235Q)·$preQ_0$; d) the same alignment viewed from another direction

In case of $preQ_1$, however, cocrystallization was not successful. No electron density of a bound ligand could be identified. Due to a split conformation of Asp102, it can be assumed that $preQ_1$ has partly populated the crystal. The increased $K_m$ of 32µM for $preQ_1$ in TGT(E235Q) reflects its limited binding competence. Nevertheless, the binding mode of $preQ_1$ remains elusive. Whether $preQ_1$ is accommodated in the binding pocket facing the exposed amide nitrogen or whether the peptide switch is still in operation to expose the Leu231 carbonyl group remains unresolved.

The superposition of the crystal structures of the wild type and TGT(E235Q) in complex with $preQ_0$ show almost identical conformations for Glu235 and Gln235. Even the 235/236 peptide flip is found with almost identical geometries (Fig. 2.12c). Nevertheless, the detailed comparison reveals some significant differences. They result from the deviating orientations of the ethylene linker of glutamate and glutamine (Fig. 2.12d). In case of the wild type TGT, complexed with $preQ_0$, the carbonic acid function adopts a geometry that disables H-bond formation with the carbonyl group of Gly230. Instead a short contact is formed with the amide of Val233 (2.8 Å), suggesting Glu235 to be negatively charged. This might explain why Glu235 avoids short contact to the carbonyl group of Gly230. In this geometry even the stacking with Tyr161 is less perfect than in the crystal structure of the mutated enzyme.

Based on the available crystal structures, the following conclusions can be drawn. In uncomplexed TGT(w.t.) Glu235 binds to the peptide switch. It changes its protonation state whether it is faced by a hydrogen-donor or -acceptor functionality of the peptide switch. Supposedly, this situation is also given if a guanine or $preQ_1$ are accommodated in the binding pocket. However, upon binding of $preQ_0$ this interaction mode is not preserved. Most presumably the negatively polarized cyano group of $preQ_0$ that interacts with the amide group of Ala232 modifies the local dielectric conditions. This prevents Glu235 to pick up a proton and to form a hydrogen bond to the carbonyl of the peptide switch. Therefore, the Glu235 side chain has to find an alternative geometry. The negative charge on the Glu235 carboxylate function disables favourable interactions with the carbonyl group of Gly230. Accordingly, only a hydrophobic stacking with the neighbouring peptide bond of the switch remains as binding geometry. In contrast the terminal amide group of Gln235 in the mutated enzyme is capable to form a hydrogen bond via its carboxamide protons to Gly230. Additionally, also the geometry for stacking with Tyr161 is more favourable. Obviously, the binding mode of Gln235, clearly distinct from the binding geometry of Glu235 in its protonated state, is rather favourable and is conserved across different pH conditions. One may speculate, as no direct contact to the peptide switch is formed, whether the switch is still operational in the mutant. At least, the kinetic experiments indicate, that $preQ_1$ is still recognized as substrate, even though with much less efficiency.

In summary these considerations provide an explanation for the significantly modulated selectivity in favour of $preQ_0$ in the mutated enzyme. The crystal structures of both isoforms recognize $preQ_0$ comparably. However, the mutated enzyme adopts already in the uncomplexed state a geometry suited for $preQ_0$ binding. In contrast, the wild type adopts this conformation only once the nucleobase enters the active site.

#### 2.2.2.5 Summary and outlook

The crystal structure analysis of the TGT(E235Q) clearly demonstrated that kinetic data from mutational studies have to be discussed carefully and cautiously. By means of protein engineering a more conservative exchange than glutamic acid / glutamine can hardly be performed. The present example shows that kinetic data is reflected as expected, but structural biology demonstrates that nature finds for this exchange an unexpected solution.

Nevertheless, this study also shows that a single mutation enables the inversion of selectivity in TGT. The fact that turn-over rates are hardly affected by the mutation clearly indicates that selectivity regulation in TGT is a multifactorial process.

The successful modulation of selectivity is an important step towards the understanding of selectivity and specificity determining features in TGTs. Mutational studies with *E. coli* TGT replacing Asp156 by various other amino acids altered the specificity towards xanthine in case of the TGT(D156N) mutated enzyme. However the catalytic activity was reduced [Todorov & Garcia, 2006][67]. In this context it might be interesting to study mutational exchanges in the close neighbourhood of the attacking nucleophile Asp280. In QueTGT this residue is tightly kept in position by H-bonds formed with Gly261 and Tyr258 (Fig. 2.3d). In the archaebacterial ArcTGT the tyrosine is replaced by a strictly conserved histidine. Likely, this exchange will have pronounced influence on the catalytic properties.

These considerations provide a perspective towards an ambitious goal: the modification of substrate specificity towards bases other than $preQ_0$ or $preQ_1$ and their efficient incorporation into tRNA. This would allow to study the translational process in more detail *via* well designed modification of the accuracy of the wobble base pairing.

# 2.3 Homodimer formation in QueTGT

The crystal structures of TGT from two species (*Z. mobilis*, *Thermotoga maritima*) exhibit homodimers. This prompted us to investigate the structural properties of these dimers, in particular their occurrence among TGTs of other species and their possible functional relevance.

### 2.3.1 Dimer formation in solution and in crystals

During the initial biochemical characterization of TGT enzymes from different species dimer and multimer formation in pure TGT solutions has been observed. However, their occurrence is controversially discussed [Reuter & Ficner, 1995][68]. *E. coli* TGT seems to form dimers at low protein concentrations and multimers at higher concentrations. For *Z. mobilis* TGT a molecular mass of 55 kDa was determined by gel filtration. This value falls between monomer (43 kDa) and dimer (86 kDa). Whether the presence of TGT as dimer has any functional relevance or whether it is only formed upon tRNA binding, possibly in a species dependent manner, has not yet been examined.

Uncomplexed TGT crystallizes in the space group *C*2 with one TGT molecule per asymmetric unit. In this space group two-fold symmetry imposes homodimer formation with a buried solvent accessible surface of 1,667Å$^2$ [Romier *et al*., 1996][25]. Within the unit cell, two symmetry related pairs of TGT dimers are present due to *C* centering (Fig. 2.13a). The crystal structure of tRNA-bound *Z. mobilis* TGT also shows a homodimer, however not imposed by symmetry. It binds a tRNA anticodon stem loop within the asymmetric unit (Fig. 2.13b – $TGT_A$, $TGT_B$) [Xie *et al*., 2003][4]. Superposition of this tRNA-bound homodimer with the uncomplexed C2 symmetrical TGT dimer reveals an identical contact interface geometry (Fig. 2.13d).

The recently determined crystal structure of *Thermotoga maritima* TGT (PDB-code: 2ASH) also exhibits a homodimer in the asymmetric unit. The enzyme of the thermophilic eubacterium shares a sequence identity of 47% (63% homology) with TGT from *Z. mobilis* (retrieved from SWISS-PROT, [Boekmann *et al*., 2003])[69]. All essential residues in the active site are conserved which indicates an unchanged enzyme function (Tab. 2.5). The superposition of this structure with the C2 symmetrical dimer reveals an identical orientation of all structural elements (Fig. 2.13c/d).

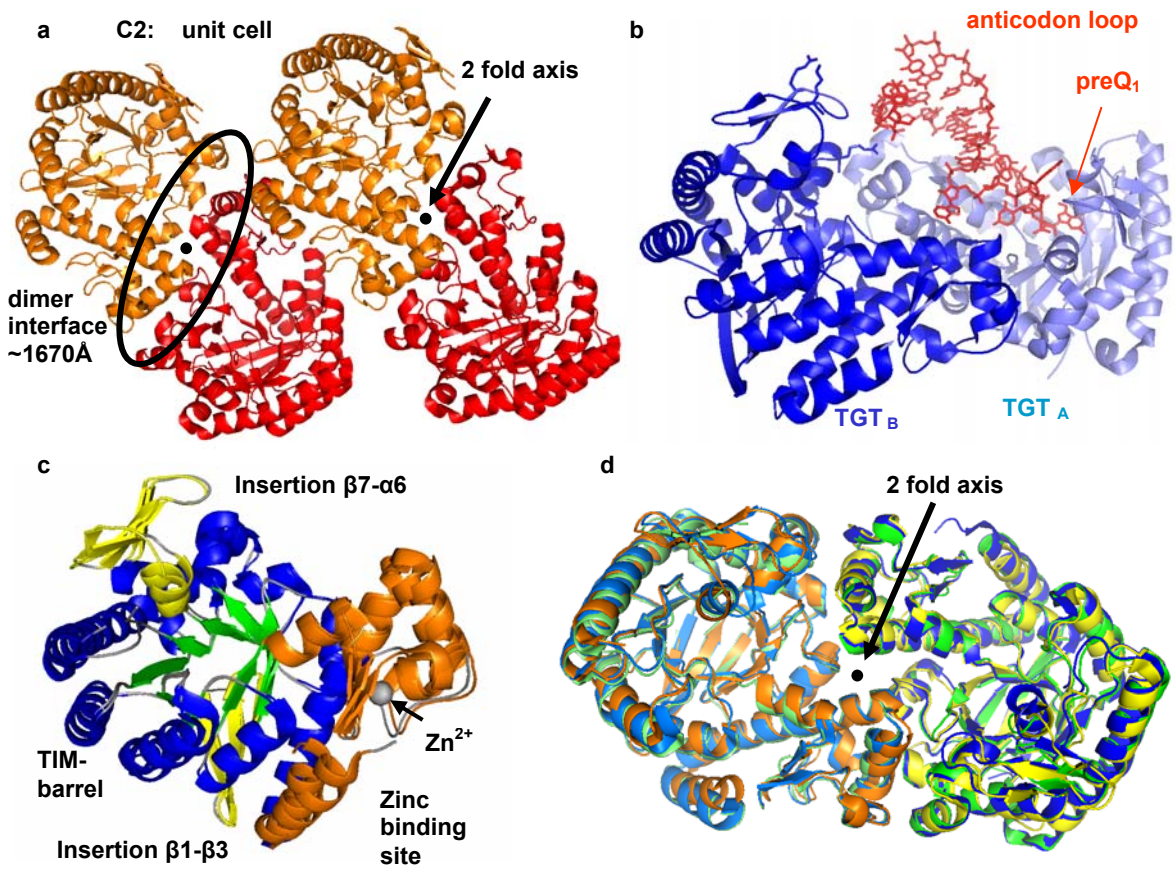



**Fig. 2.13** a) Uncomplexed *Z. mobilis* TGT crystal structure: unit cell content in space group *C*2 formed by four symmetry equivalents; b) crystal structure of *Z. mobilis* TGT in complex with tRNA anticodon stem loop: two TGT molecules binding on stemloop are found in the asymmetric unit; c) superposition of crystal structures of *Z. mobilis* and *T. maritima* TGT coloured by structural elements; d) superposition of uncomplexed *Z. mobilis* TGT (dimer formed by two fold symmetry – green), *Z. mobilis* TGT in complex with anticodon stem loop (dimer from asymmetric unit – blue) and uncomplexed *T. maritima* TGT (dimer from asymmetric unit – yellow/orange)

### 2.3.2 Sequence comparison of 21 TGTs from different species

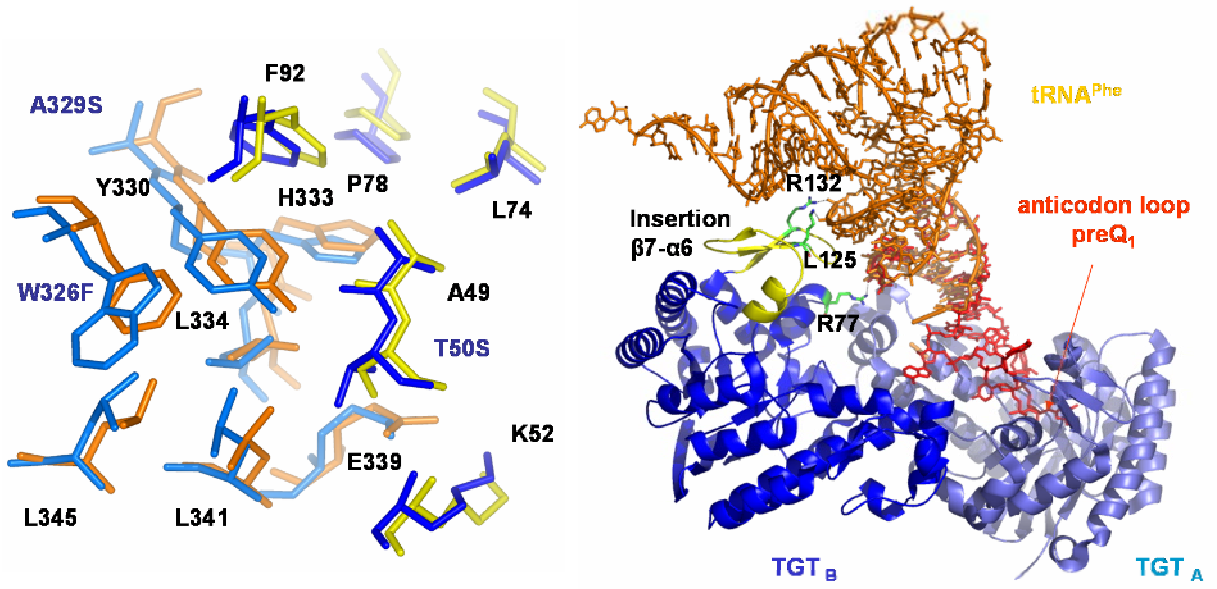
The structural superposition of the dimer interfaces from *Z. mobilis* and *T. maritima* TGT indicated that the involved residues are highly conserved (Fig. 2.14a). To collect further evidence for the presence of a conserved interface, a sequence alignment was performed based on 21 TGT sequences retrieved from SWISS-PROT and aligned using CLUSTAL_W [Boekmann *et al.*, 2003] (Tab. 6.1)[69]. The sequences originate from 21 species and were selected as representatives of major eubacterial and eukaryotic subdivisions. The selection followed a recently presented global

phylogeny using knowledge from 191 sequenced genomes [Ciccarelli *et al*., 2006] [70]. For reasons of comparison, at first the residues of the active site were aligned (Tab. 2.5 – following *Z. mobilis* numbering). The composition of the active site is highly conserved among all TGTs, except for *Chlamydia* species. Here, the active site exhibits some notable differences. Furthermore, a *Chlamydia*-specific insertion of 18 residues between position 114 and 115 deviates from all other TGTs and suggests a slightly different function (Tab. 6.1). Thus, the *Chlamydia* TGT was excluded from any further considerations. Across eubacterial TGTs the core of residues, forming the dimer interface, is highly conserved (Tab. 2.6). Lys52, Leu74, Pro78, Phe92 from the TIM-barrel site of the interface and Tyr330, His333, Leu334, Glu339, Leu345 from the zinc-binding subdomain have identical interaction partners in most eubacterial divisions. At position 49 (Gly/Ala), 50 (Thr/Ser/Cys), 326 (Phe/Tyr/Trp), 329 (Ala/Ser) and 341 (Leu/Phe) homologous residues are found. These results suggest that in Eubacteria dimer formation as general property of TGTs seems likely.
In eukaryotic TGTs Lys52, Leu74 and Pro78 are conserved and at positions 50 (Thr/Ser), 92 (Phe/Met), 329 (Ala/Ser), 330 (Tyr/Phe), 339 (Glu/Asn) and 345 (Leu/His) residues similar to those in Eubacteria are found (Tab. 2.6). Although, the overall degree of conservation is smaller, compared to Eubacteria, dimer formation is possibly also given in Eukaryotes.

### 2.3.3 Functional model for the QueTGT dimer

In contrast to QueTGT, the archaebacterial TGT is known to form functional dimers upon catalysis. This has been described in detail in chapter 2.1.4.1. However, the structural superposition of *P. horikoshii* TGT with *Z. mobilis* TGT revealed a completely different composition of the dimer interface. This may result from significantly different tRNA binding geometries and modification sites resulting in a 2:2 complex of TGT with tRNA. Although a direct analogy between the archaebacterial and eubacterial enzyme is not given, the fact of a functional dimer for Archaebacteria together with the presence of a highly conserved dimer interface in Eubacteria suggests that dimer formation is required for function also for the latter class. In the 2:1 complex of *Z. mobilis* TGT with tRNA, a TGT homodimer ($TGT_A$, $TGT_B$) is bound to an elongated anticodon stem loop which served as substitute for full-length tRNA [Xie *et al*., 2003][4]. In the crystal structure, the $U_{33}G_{34}U_{35}$ recognition

sequence of the anticodon forms strong interactions to $TGT_A$. As the two active sites of $TGT_A$ and $TGT_B$ pack closely together only one tRNA anticodon loop is bound in this crystal form. Although $TGT_B$ is not involved in the recognition of the tRNA substrate, it forms an interaction with the anticodon stem loop. Arg77 is H-bonded to the phosphate groups of $C_{27}$ and is found in close contact to the phosphate group of $G_{26}$. This arginine is highly conserved among all eubacterial and eukaryotic TGTs (few exceptions: Lys/Gln – Tab. 2.6).




a b

**Fig. 2.14** a) Structural alignment of the dimer interfaces in *Z. mobilis* TGT (blue/slate blue) and *T. maritima* TGT (yellow/orange); b) structural superposition of $tRNA^{Phe}$ from yeast onto the tRNA anticodon stem loop from the tRNA-complexed *Z. mobilis* TGT crystal structure; residues from $TGT_B$ assumed to interact with tRNA are coloured in green

To analyze further putative interaction sites of $TGT_B$ with respect to full-length tRNA, the geometry of $tRNA^{Phe}$ (PDB code: 1EHZ) was modelled onto the anticodon stem loop (Fig. 2.14b). This superposition shows two positively charged residues, Lys125 and Arg132, of $TGT_B$ to be well placed to interact with phosphate groups of the tRNA D-stem bases in position 11 and 12. The two residue side chains were manually adjusted towards tRNA to demonstrate their ability to form H-bonds with the tRNA. Lys125 and Arg132 are located in a TIM-barrel insertion ($\beta$7-$\alpha$6) characterized by three antiparallel $\beta$-strands. This insertion is conserved among eubacterial and eukaryotic TGTs [Stengl *et al.*, 2005][54]. Even though Lys125 and Arg132 are not

strictly conserved among Eubacteria and Eukaryota, at both positions predominantly Arg or Lys are found (Tab. 2.6).

As these putative interaction sites with the bound tRNA seem to be widely conserved, it can be concluded that the eubacterial (and most presumably also eukaryotic) TGTs form functional dimers upon catalysis. While only one molecule performs the base exchange reaction, the other serves to stabilize the bound tRNA in its geometry via interactions with Arg77 and insertion β7-α6. This hypothesis would explain why the residues in the dimer interface are remarkably conserved across various species.

### 2.3.4 Outlook

The performed sequence and structural alignments provide strong indications for functional dimer formation in TGT. However, this hypothesis still has to be tested experimentally. Gelfiltration experiments are a good starting point for such investigations. Initial experiments performed with a Superdex 200 10/30 prepgrade column indicate that *Z. mobilis* TGT is present as monomer in solutions of uncomplexed TGT. Further experiments performed with TGT in the presence of full length $tRNA^{Tyr}$ indicate the formation of dimers and multimers. However, the results from these experiments are not yet conclusive enough, as the applied conditions require further validation. Dynamic light scattering could offer an alternative to assess complex formation. Additionally, mutational studies modifying the structure and integrity of the dimer interface could verify the above-discussed assumptions.

**Tab. 2.5** Alignment of residues involved in substrate recognition, base exchange and complexation of zinc from 21 TGT sequences retrieved from SWISS-PROT

| species | Guanine 34 / preQ$_1$ binding site | | | | | | | | | Ribose 34 binding site / catalytic site | | | | | | | | | zinc binding site | | | | classification |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 102 | 106 | 107 | 156 | 203 | 230 | 231 | 235 | 260 | 45 | 47 | 68 | 70 | 73 | 258 | 261 | 280 | 282 | 318 | 320 | 323 | 349 | **Eubacteria** |
| *Escherichia* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | γ-Proteobacteria |
| *Vibrio* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| *Pseudomonas* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| *Neisseria* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | β-Proteobacteria |
| *Agrobacterium* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | S | C | S | H | α-Proteobacteria |
| *Rickettsia* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| *Zymomonas* | D | Y | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| *Helicobacter* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | ε-Proteobacteria |
| *Gloeobacter* | D | F | Q | D | Q | G | V | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Cyanobacteria |
| *Thermotoga* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Thermotogae |
| *Chlamydia p.* | D | F | Q | D | H | G | S | R | L | V | T | F | N | H | H | G | D | S | C | C | C | H | Chlamydiae |
| *Chlamydia t.* | D | F | Q | D | H | G | S | K | L | V | T | F | N | H | H | G | D | S | C | C | C | H | |
| *Treponema* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Spirochaetes |
| *Bacillus* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Firmicutes |
| *Lactobacillus* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| *Clostridium* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| | | | | | | | | | | | | | | | | | | | | | | | **Eukaryotes** |
| *Schizosacch.* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Fungi |
| *Caenorhabditis* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Ecdysozoa |
| *Drosophila* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |
| *Mus* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | Vertebrata |
| *Homo* | D | F | Q | D | Q | G | L | E | M | V | T | L | N | H | Y | G | D | V | C | C | C | H | |

**Tab. 2.6** Alignment of residues suggested to be involved in dimer formation and tRNA stabilization from 19 TGT sequences retrieved from SWISS-PROT

| species | TIM-barrel side | | | | | | zinc binding subdomaine side | | | | | | | | $TGT_B$ - tRNA | | | classification |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 49 | 50 | 52 | 74 | 78 | 92 | 326 | 329 | 330 | 333 | 334 | 339 | 341 | 345 | 77 | 125 | 132 | **Eubacteria** |
| *Escherichia* | G | T | K | L | P | F | Y | A | Y | H | L | E | L | L | R | R | P | γ-Proteobacteria |
| *Vibrio* | G | T | K | L | P | F | Y | S | Y | H | L | E | L | L | R | R | K | |
| *Pseudomonas* | G | T | K | L | P | F | F | A | Y | H | L | E | L | L | R | A | K | |
| *Neisseria* | G | S | K | L | P | F | F | A | Y | H | L | E | L | L | R | K | K | β-Proteobacteria |
| *Agrobacterium* | G | T | K | L | P | L | Y | A | Y | H | L | E | L | L | R | K | L | α-Proteobacteria |
| *Rickettsia* | G | T | K | L | P | F | Y | A | Y | H | L | E | L | L | Q | S | K | |
| *Zymomonas* | A | T | K | L | P | F | W | A | Y | H | L | E | L | L | R | K | R | |
| *Helicobacter* | G | C | K | M | P | F | Y | A | Y | H | L | E | T | L | R | K | K | ε-Proteobacteria |
| *Gloeobacter* | A | T | K | L | P | F | F | A | Y | H | L | E | L | L | Q | R | L | Cyanobacteria |
| *Thermotoga* | A | S | K | L | P | F | F | S | Y | H | L | E | L | L | K | R | K | Thermotogae |
| *Treponema* | A | T | K | L | P | F | Y | A | Y | H | L | E | L | L | R | Q | R | Spirochaetes |
| *Bacillus* | A | T | K | L | P | F | Y | A | Y | H | L | E | F | L | R | R | K | Firmicutes |
| *Lactobacillus* | A | S | K | L | P | F | F | A | Y | H | L | E | F | L | R | K | K | |
| *Clostridium* | G | V | K | L | P | F | Y | A | Y | H | L | E | L | L | R | N | R | |
| | | | | | | | | | | | | | | | | | | **Eukaryotes** |
| *Schizosacch.* | A | S | K | L | P | F | E | A | Y | N | S | E | V | L | K | L | P | Fungi |
| *Caenorhabditis* | G | T | K | L | P | M | Y | A | Y | H | S | E | V | L | R | E | M | Ecdysozoa |
| *Drosophila* | G | T | K | L | P | F | Y | S | Y | H | I | E | V | L | R | R | Q | |
| *Mus* | A | T | K | L | P | F | H | A | F | A | L | N | T | H | R | R | E | Vertebrates |
| *Homo* | A | T | K | L | P | F | H | A | F | A | L | N | A | H | R | R | E | |

# 2.4 Classification of the TGT superfamily

The tRNA bound crystal structures of ArcTGT and QueTGT allow, due to their substrate binding modes, the extraction of valuable information about the evolutionary origin of TGTs within the TIM-barrel fold. Taking the available amino acid sequence data of TGTs and further enzymes involved in the tRNA modification pathway into consideration, also the relationship within the superfamily can be elucidated. The presented results have already been summarized in Stengl *et al*. (2005)[54].

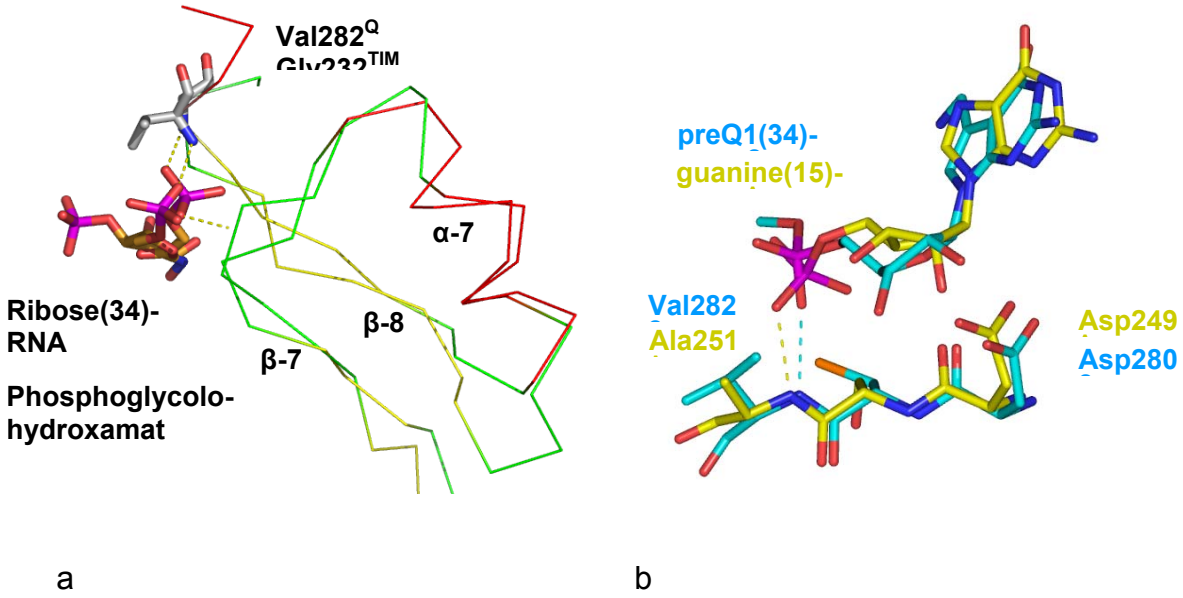
## 2.4.1 Evolutionary origin of the TGT superfamily

The TGT superfamily is grouped together with enzymes adopting a triose-phosphate isomerase (TIM)/$(\beta\alpha)_8$ barrel-like fold. This fold is very abundant and approximately 10 % of all enzymes share this geometry [Gerlt, 2000][71]. The biological roles of 85 % of the known reaction types performed by TIM barrel enzymes are associated with metabolism [Nagano *et al*., 2002][72].

During the last two decades, several evolutionary classification models for TIM-barrel enzymes have been proposed [Farber & Petsko, 1990; Lesk *et al*., 1989][73, 74]. Recent considerations favour the model of divergent evolution starting from a common ancestor, rather than convergent evolution. For about a dozen of the assumed 26-29 homologous superfamilies, sequential, structural and functional evidence has been presented to support this assumption [Copley & Bork, 2000; Nagano *et al*., 2002][72, 75]. Among these, presumably closer related superfamilies, a standard phosphate binding (SPB) motif, involved in the recognition of substrate phosphate groups, is widely spread. Thus, it was used as one criterion to achieve higher order classification. This structural element ranges from the 7$^{th}$ $\beta$-strand to the 8$^{th}$ $\alpha$-helix of the TIM barrel motif, characterized by a high structural homology. Additionally conserved phosphate binding positions at the ends of the adjacent 7$^{th}$ and 8$^{th}$ TIM barrel strands are present.

A straight-forward assignment of TGTs to the TIM barrel fold is difficult. As TGTs catalyze reactions attributed to the information pathway, they constitute one of the few examples for TIM barrel enzymes not involved in metabolism. Furthermore, due

to unusual insertions into the TIM barrel responsible for RNA binding, TGTs form a separated superfamily (see chapter 2.1.1). Furthermore, almost no global structural relationship to other superfamilies could be detected. However, detailed analysis of this fold with specific structural alignment methods revealed that TGTs share the standard phosphate binding motif with some other mutually related superfamilies [Nagano *et al*., 2000]. For *Z. mobilis* TGT the overall $C_\alpha$-RMSD with respect to the SPB motif is only 1.7Å-2.0Å. $Gly261^Q$ and $Val262^Q$ at the end of strand β-13 (≡ $7^{th}$ TIM barrel strand), as well as $Val282^Q$ at the end of strand β-14 (≡ $8^{th}$ TIM barrel strand) were predicted to be in conserved positions for substrate phosphate binding *via* backbone interactions.




a b

**Fig 2.15** a) Conserved SPB- motif in QueTGT and triose-phosphate isomerase; b) conserved SPB- motiv in QueTGT and ArcTGT

The tRNA complexed QueTGT structure revealed that the predicted $Val282^Q$ is indeed involved in binding of the substrate's phosphate group of $G_{34}$. The second residue, $Gly261^Q$, is involved in binding of a substrate ribose hydroxy group [Stengl *et al*., 2005][54]. A structural alignment of the SPB motif of QueTGT and triose-phosphate isomerase (TIM) from *Gallus gallus*, with their respective ligands shows equivalently positioned phosphate groups, binding residues and orientation towards $Val282^Q$ and $Gly232^{TIM}$ (Fig. 2.15a). Furthermore, in QueTGTs $Val282^Q$ is one of the well conserved residues next to the active site. In the available ArcTGT structure from *P. horikoshii* $Ala251^A$, corresponding to $Val282^Q$, displays a similar phosphate binding mode, however, to the phosphate group of $G_{15}$ (Fig. 2.15b). This finding supports the

allocation of the TGT superfamily to other presumably homologous SPB containing superfamilies.

The SPB motif is also a key structural element to understand why TGT is not involved in metabolism as most of the other TIM barrel enzymes. Comparison of evolutionarily related TIM barrel enzymes involved in metabolism shows no strict correlation with the metabolic pathway they are involved in. Often related enzymes perform tasks in different metabolic pathways. An evolutionary model suggests a process called 'enzymatic recruitment' between different metabolic pathways as one possible driving force to develop new protein functions [Copley & Bork, 2000][75]. This model requires an ancient enzyme possessing a broader substrate specificity as starting point. New functions can evolve in case of different compounds sharing a common structure and being accepted as substrates by this enzyme. Specific groups (e.g. phosphate groups) represent such common substructures. Duplication and diversification of the respective gene finally result in two distinct protein functions involved in different metabolic pathways.

Applying this evolutionary model to TGT results in the following consideration. The ancient predecessor of nowadays existing TGTs was possibly involved in central metabolism accepting and binding phosphorylated substrates via its SPB motif. Due to the presence of phosphate moieties in tRNA molecules they could have served as some sort of 'accessory'- substrate for this ancestor. After gene duplication and evolutionary modification, one such copy could have evolved with specific tRNA recognition and modification properties as primordial TGT. This could explain why the TGT ancestor has possibly departed from pure metabolism and developed as a catalyst with a function rather unusual for an enzyme with a TIM barrel fold (Fig. 2.16).

### 2.4.2 Classification within the TGT superfamily

The presence of TGTs across broad phylogenetic boundaries indicates the evolutionary origin of the primordial TGT in a prokaryotic predecessor before the separation of the three kingdoms of life. The further evolution of TGT within the three kingdoms, resulting in deviating substrate specificities, however, is not certain at all. The following considerations try to fit the current knowledge to a consistent picture (Fig. 2.16).

Due to the fact that QueTGT and ArcTGT are involved in different tRNA modification pathways it can be speculated that the initial split between the eubacterial and the archaebacterial lineage took place before the two pathways have been fully established. Thus, for the last common ancestor it only seems clear that its TGT was involved in tRNA modification. It is uncertain at which tRNA position the modification was performed. It is also uncertain whether $preQ_0$ or $preQ_1$ was accepted as substrate. The lower number of required steps for biosynthesis of $preQ_0$ argues in favour of this base. The less complex architecture of QueTGT speaks in favour of $preQ_1$, although this assumption is challenged by the fact that QueTGT accepts $preQ_0$ as substrate as well. However, in Eubacteria QueTGT was finally located in the Queuine-pathway influencing transcription and in Archaebacteria ArcTGT was integrated into the Archaeosine-pathway increasing tRNA stability.

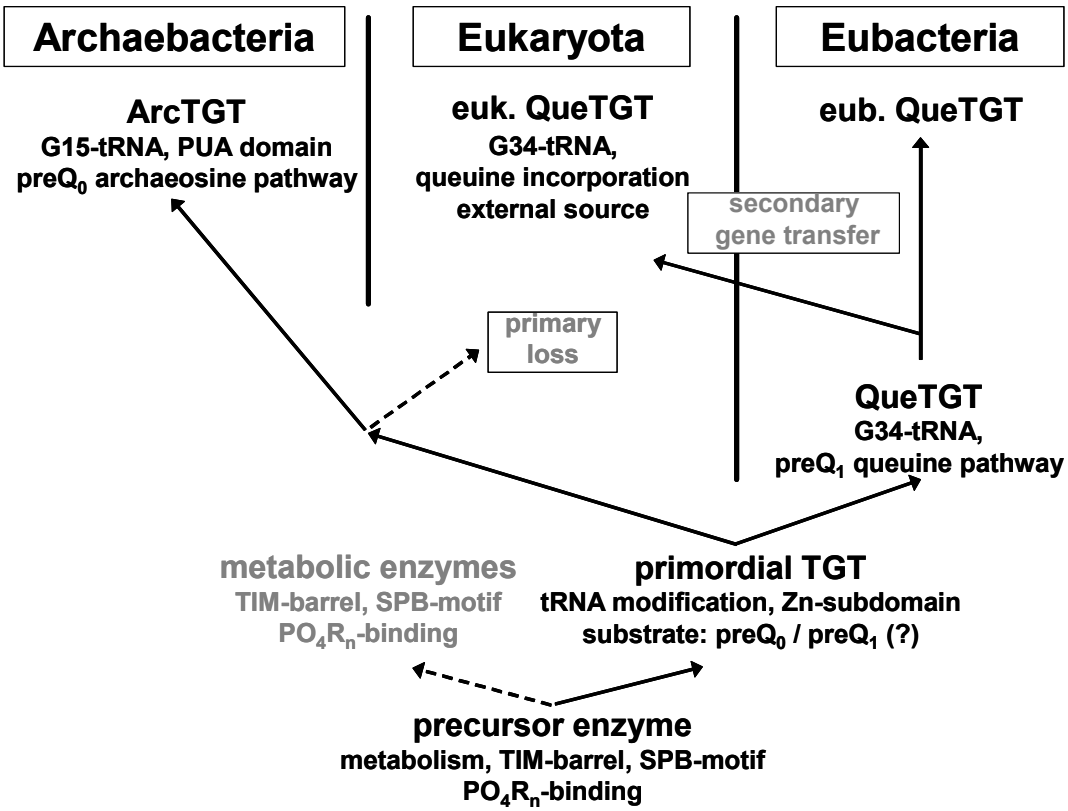



**Fig. 2.16** Evolutionary model of the TGT superfamily

The following lineage split took place between Archaebacteria and Eukaryota [Ciccarelli *et al*., 2006] [70]. As in Eukaryota neither orthologs of *arctgt* nor of *queC–E* genes involved in $preQ_0$ synthesis are found, a primary loss of the enzymes involved in tRNA modification in the initial Eukaryota has to be assumed. Secondarily however, in a later evolutionary stage the Eukaryota reacquired a single *tgt*-gene, most presumably by horizontal gene transfer. The gene was of eubacterial origin, indicated by the high degree of structural and functional identity of eukaryotic and eubacterial TGT. Most presumably the *quetgt*-gene of the primordial Eukaryota originated from an eubacterial endosymbiont, which further evolved towards the mitochondria of the present Eukaryota. This evolutionary process was possibly assisted by multiple gene translocation events from the mitochondria chromosome towards the eukaryotic chromosome. This hypothesis has been investigated in detail for proteins of the nuclear pore complex and the splicosome [Martin & Koonin, 2006][76]. As solely the *quetgt*-gene was transferred, *de novo* Q-synthesis was disabled for Eukaryota. Thus, Eukaryota were still dependent on external queuine sources and substrate specificity had to be modified towards queuine. Due to the fact that the initial Eukaryota were highly mobile predators of bacterial cells, queuine could be provided by nutrition. Even for the Metazoa with eubacterial commensals in their digestive apparatus queuine could be provided in sufficient amounts. The only present eukaryotic organisms that are known to lack queuine are green plants, obviously not possessing an external queuine source, and *Saccharomyces cerevisae*, a yeast living predominantly from vegetal sources.

These considerations corroborate the evolutionary history within the TGT superfamily and form a complex but nevertheless consistent picture.

# 3. Structure-based Inhibitor Design

The new knowledge from the functional analysis was implemented in structure-based inhibitor design. As a consequence of the detailed insight into the base exchange reaction an appropriate binding assay had to be developed taking competitive and uncompetitive inhibition contributions into consideration, which were not considered in the previous assay (chapter 3.1). Therefore, inhibition constants of relevant inhibitor classes detected and developed in previous studies were revalidated to adjust structure – activity relationship (chapter 3.2). Based on these results, a quinazolinone-based inhibitor series was further developed resulting in several new lead compounds (chapter 3.3). In chapter 3.4 a virtual screening hit detected earlier was further investigated by means of structure-based design. Finally, a compound class, isolated from plant extracts by an HPLC assisted ligand-fishing method using TGT as trap, was tested with respect to their inhibition potency (chapter 3.5).

## 3.1 Modifications of the binding assay

The determination of inhibition constants in TGT is based on monitoring of tritium labelled [8-$^{3}$H]-guanine which is incorporated in tRNA catalyzed by TGT. Labelled guanine is used as alternative substrate instead of natural preQ$_1$, for which no labelled material is available. The inhibition of TGT reduces the amount of incorporated tritium-labelled guanine in tRNA which can be monitored by liquid scintillation counting. From the reduction of the initial velocity of the base exchange reaction inhibition constants can be derived [Grädler *et al*., 2001][3]. However, the determination of inhibition constants is complicated by several factors. Firstly, low solubility might result in non-specific inhibition. Such effects can be avoided by the addition of detergents. As a secondary effect the protein solubility is improved (chapter 3.1.1). Secondly, as described above, the inhibition mechanism in TGT is more complex than previously assumed. As the base exchange follows a ping-pong mechanism competitive and uncompetitive inhibition contributions have to be considered in dependence on the ligand properties (chapter 3.1.2 - 3.1.3). To face these complications the previously used assay was modified (chapter 3.1.4). All necessary modifications of the assay setup recently have been published by us in

Meyer *et al*. (2006)[65]. In a short outlook an alternative approach will be presented that might be suited to replace the radioactive assay (chapter 3.1.5).

## 3.1.1 Detergents effect ligand and protein solubility

### 3.1.1.1 Detergents and non-specific inhibition

Many of the tested compounds exhibit low solubility under the applied assay conditions, especially those coming from the class of pteridines and quinazolinones (chapter 3.2). Effects of low solubility can be manifold. As a consequence of possible precipitation, the resulting inhibition potency could be underestimated. However, also the opposite could be observed, as reported recently. Compound with low solubility might also exhibit unspecific inhibition. The risk of such inhibition exists for a large variety of compounds and has been referred to as “promiscuous inhibition” [McGovern *et al*., 2002][77]. Such compounds show up repeatedly as inhibitors of various enzymes in biological assays and do not follow a target-specific mode of action. Instead, such behaviour is thought to result from compound agglomeration that adsorbs the enzyme onto its aggregated surface. In consequence, the enzyme is inactivated. Adding detergents, such as *Tween 20* or *Triton X-100*, to the respective assay solutions obviously reduces aggregate formation and thus allows to distinguish “promiscuous” from specific inhibitors [McGovern *et al*., 2002][77]. To avoid possible effects of “promiscuous” or non-specific inhibition as well as other undesired precipitation effects in the aqueous TGT assay, *Tween 20* was added in 5% of its critical micellar concentration (CMC) of 59 µM (chapter 5.3.1). The amount of *Tween 20* used is based on the results reported by Ryan *et al*. (2003)[78]. An example for the influence of *Tween* 20 on affinity data in TGT is presented in chapter 3.5.

### 3.1.1.2 Detergents and TGT solubility

As a supplementary effect, the use of detergents influences not only ligand but also protein solubility [Ryan *et al*., 2003][78]. In particular *Tween 20* was described to reduce protein absorption onto plastic surfaces of cups usually used for kinetic measurements. Such artificial immobilization results in a reduced effective enzyme concentration in the assay solution. Therefore, the kinetic parameters for TGT were

re-examined to quantify the possible impact of *Tween 20* (Tab. 3.1 and chapter 5.3.2). Furthermore, accurately determined kinetic parameters are required to derive inhibition constants. $K_m$ values for TGT with respect to both substrates $tRNA^{Tyr}$ and [8-$^3$H]-guanine remain within the range already published for *Z. mobilis* TGT ($K_m(tRNA^{Tyr})$: 0.2-1.0 $\mu$M; $K_m$([8-$^3$H]-guanine): 0.38-1.3 $\mu$M) [Reuter & Ficner, 1995; Brenk *et al.*, 2003][64, 68]. In contrast $k_{cat}$ obtains higher values than reported previously. Most likely, this observation results from a higher effective TGT concentration in the *Tween 20*-containing assay solution. Accordingly, it can be assumed that the principal catalytic properties are not affected by added *Tween 20*.

**Tab. 3.1** Kinetic parameters for TGT(w.t.) in the presence and absence of *Tween 20* (previous values from Brenk *et al*. (2003)[64])

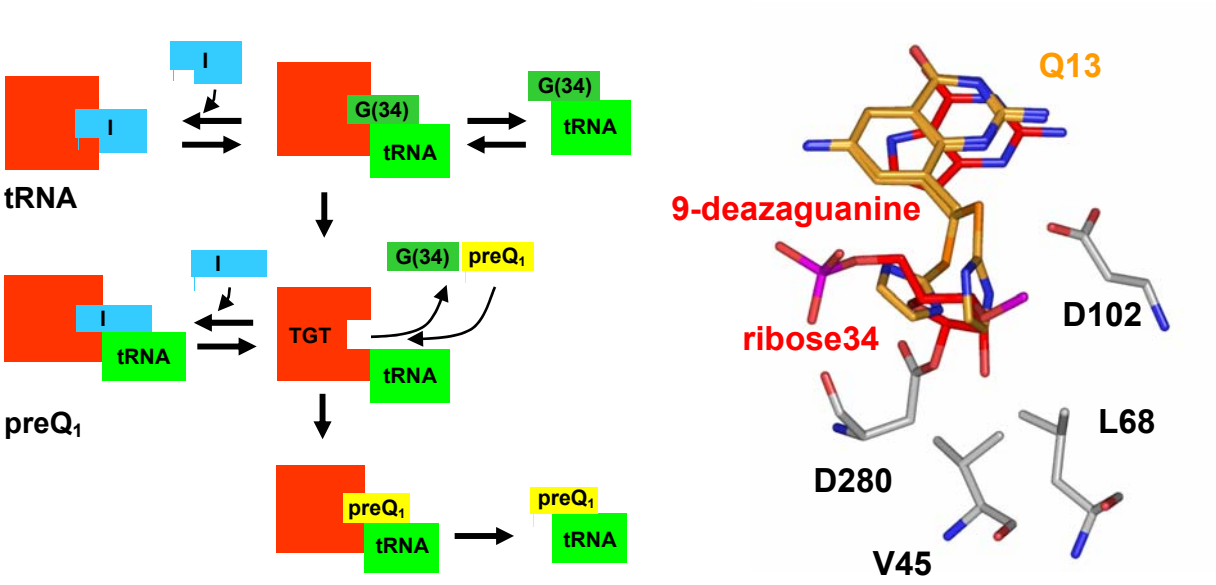
| **TGT (w.t.)** | **$tRNA^{Tyr}$** | previous values | **[8-$^3$H]-guanine** | previous values |
|---|---|---|---|---|
| $K_m$ [µM] | 0.9 ± 0.2 | 1.0 ± 0.4 | 1.2 ± 0.2 | 0.4 ± 0.1 |
| $k_{cat}$ [$s^{-1}$] | $2.7*10^{-2}$ | $1.4*10^{-2}$ | $2.8*10^{-2}$ | $1.1*10^{-2}$ |
| $k_{cat}/K_m$ [$\mu M^{-1}s^{-1}$] | $3.0*10^{-2}$ | $1.4*10^{-2}$ | $2.3*10^{-2}$ | $2.9*10^{-2}$ |

### 3.1.1.3 Detergents and inhibitor preincubation

Preincubation of *Z. mobilis* TGT with a putative inhibitor was reported to affect the initial velocity of the enzyme-catalyzed reaction [Brenk *et al.*, 2003][64]. A 10 minute preincubation of TGT with the inhibitor to be tested prior to the addition of substrate showed reduced initial velocity by about 50%. In contrast to this finding, the addition of *Tween 20* (*c* = 5% CMC) rather resulted in a slight increase of the initial velocity after a 10 minute preincubation and increased the initial velocity in the absence of *Tween 20* by a factor of three. It can be assumed that the previously observed preincubation effect resulted from a superimposed slow adsorption/desorption process of the enzyme on the plastic cup surface which is unmasked by the use of *Tween 20*.

## 3.1.2 Inhibition of a ping-pong reaction

Detailed analyses of the TGT base exchange reaction by means of kinetic and crystallographic studies provided a better insight into the molecular details of the reaction pathway (chapter 2.1.2.2). The base exchange reaction follows a ping-pong mechanism [Goodenough-Lashua & Garcia. 2003][53]. Subsequently, TGT accepts two substrates, tRNA and $preQ_1$. For each substrate the enzyme provides a distinct binding pocket arrangement. While tRNA interacts with uncomplexed TGT, $preQ_1$ binds into a modified binding pocket partially formed upon the covalent binding of tRNA to Asp280. In a ping-pong reaction two modes of inhibition are possible (Fig. 3.1a). Either the binding of tRNA is prevented (inhibition of the first reaction step), or the binding of $preQ_1$ is prevented (inhibition of the second reaction step).




a b

**Fig. 3.1** a) Inhibition modes in the TGT ping-pong reaction pathway; b) structural superposition of covalent TGT-tRNA intermediate stabilized by 9-deazaguanine and **Q13** binding to uncomplexed TGT; the **Q13** substituent would interfere with the ribose binding site (for chemical formula in Fig. 3.3)

The molecular bases for the inhibition of the second reaction step can be extracted from the crystal structure of TGT in complex with a tRNA anticodon stem loop stabilized by 9-deazaguanine. The structure of the "ternary" complex revealed that this base served as non-reactive substrate mimic for $preQ_1$. 9-Deazaguanine accommodates the binding pocket *via* non-covalent interactions and stabilizes the covalent TGT-tRNA intermediate state (Fig. 3.1b). Accordingly, this site can either

accommodate active substrates (e.g. guanine or $preQ_1$) or other small non-reactive ligands. In the series of already studied inhibitors, some compounds exhibit molecular dimensions similar to the natural substrates, thus competing with $preQ_1$ for binding. These considerations appear in particular relevant for a quinazolinone-based inhibitor series of TGT. Therefore, this inhibitor class was used by us as a test case [Meyer *et al*., 2006][65]. Competition with $preQ_1$ binding is expected for the quinazolinone lead structures **Q1** and **Q2** (for chemical formulae see Fig. 3.3). However, once decorated by additional substituents, as in **Q13**, steric clashes with the covalently bound tRNA ribose moiety at position 34 might occur. Accordingly, it can be assumed that these extended compounds compete only with tRNA binding to uncomplexed TGT (inhibition of the first reaction step) (Fig. 3.1b).

### 3.1.3 Detection of the inhibition modes

To verify the assumption of two independent modes of inhibition, a trapping experiment was performed (chapter 5.3.3). It is based on the study of *Xie et al.* (2003)[4] where the covalent TGT-RNA intermediate was stabilized by 9-deazaguanine. In this analysis 9-deazaguanine was replaced by **Q1** and **Q2**. In different reaction mixtures, TGT was first incubated with an excess of tRNA together with two small sized (**Q1** and **Q2**) and substituted inhibitors (**Q5** and **Q13**). Then, an SDS-PAGE analysis of these mixtures was performed (Fig. 3.2). Bands are observed for uncomplexed TGT at 43.5 kDa (lane 2). Incubation of TGT with tRNA results in two additional, faint, retarded bands at about 70 kDa (lane 3). These bands correspond to covalent TGT-tRNA complexes [Nonekowski & Garcia, 2001][79]. The detection of two bands for the covalent complex results from different tRNA conformers present under the applied PAGE conditions as already reported elsewhere [Kung & Garcia, 1998][80]. The occurrence of well-defined bands for TGT-tRNA in the presence of small inhibitors such as **Q1** and **Q2** (lane 6 and 7) demonstrates the potential of these ligands to stabilize the covalent intermediate similarly to 9-deazaguanine. Once exposed to inhibitors exhibiting a large substituent at their basic scaffolds such as **Q5** and **Q13** (lane 4 and 5), the respective bands are only very faint similar to the situation when uncomplexed TGT is mixed with tRNA in absence of a small molecule inhibitor (lane 3). This finding indicates that inhibitors of the size of **Q5** and **Q13** are not capable to stabilize the covalent TGT-tRNA complex.

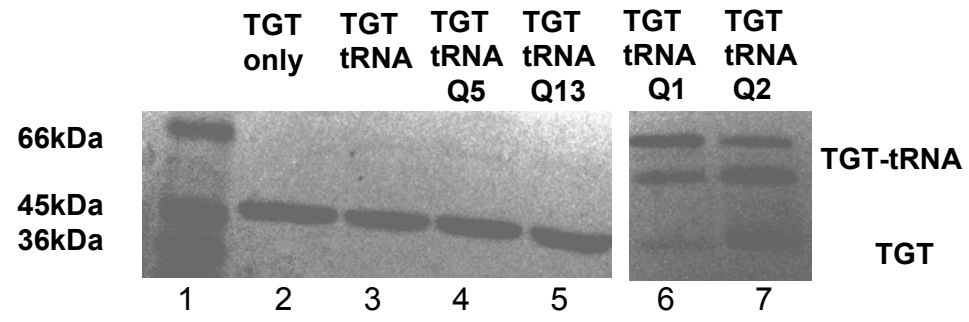



**Fig. 3.2** Trapping experiments with quinazolinone-based inhibitors

## 3.1.4 Modification of tRNA-based tritium labelling assay

In consequence of the size-dependent inhibition model with respect to the two substrates tRNA and [8-$^3$H]-guanine, replacing preQ$_1$ in the assay, the procedure to determine inhibition constant had to be modified (chapter 5.3.3 and 5.3.4). Inhibitors with large side chains are expected to be competitive only with tRNA substrate binding to TGT (competitive inhibition with respect to tRNA). Small inhibitors of comparable size to the natural nucleobase substrates, however, exhibit a more complex inhibition profile (Fig. 3.3). They either occupy the base exchange site prior to tRNA binding to TGT (competitive inhibition with respect to tRNA). It can be concluded hat this type of binding is possible considering the formation of binary TGT·inhibitor complexes observed in many crystal structures [Brenk *et al*., 2004; Meyer *et al*., 2004][1, 66]. Alternatively, these ligands compete with preQ$_1$ or guanine after tRNA is covalently attached to TGT (uncompetitive inhibition with respect to tRNA; competitive inhibition with respect to preQ$_1$ / [8-$^3$H]-guanine). Accordingly, to compare small guanine-sized and larger, extended inhibitors, their binding constants determined with respect to tRNA binding are required to provide directly comparable inhibitory values. Thus, for inhibitors of the guanine-type size two inhibition constants (competitive inhibition constant: $K_{ic}$ ; uncompetitive inhibition constant: $K_{iu}$) have to be defined that fully characterize their TGT inhibition potency with respect to tRNA binding.

Following the modified protocol, inhibition constants were determined with respect to tRNA. Thus, the tRNA$^{Tyr}$ concentration was now limiting, while [8-$^3$H]-guanine was used in excess. Due to the fact that in the previously applied assay protocol, inhibition constants have been determined with respect to [8-$^3$H]-guanine as limiting substrate (competitive inhibition with respect to preQ$_1$ / [8-$^3$H]-guanine), also

inhibition constants for relevant compounds acting as sole competitive inhibitors from previous design approaches had to be re-determined (chapter 3.2) [Meyer et al., 2006][65].

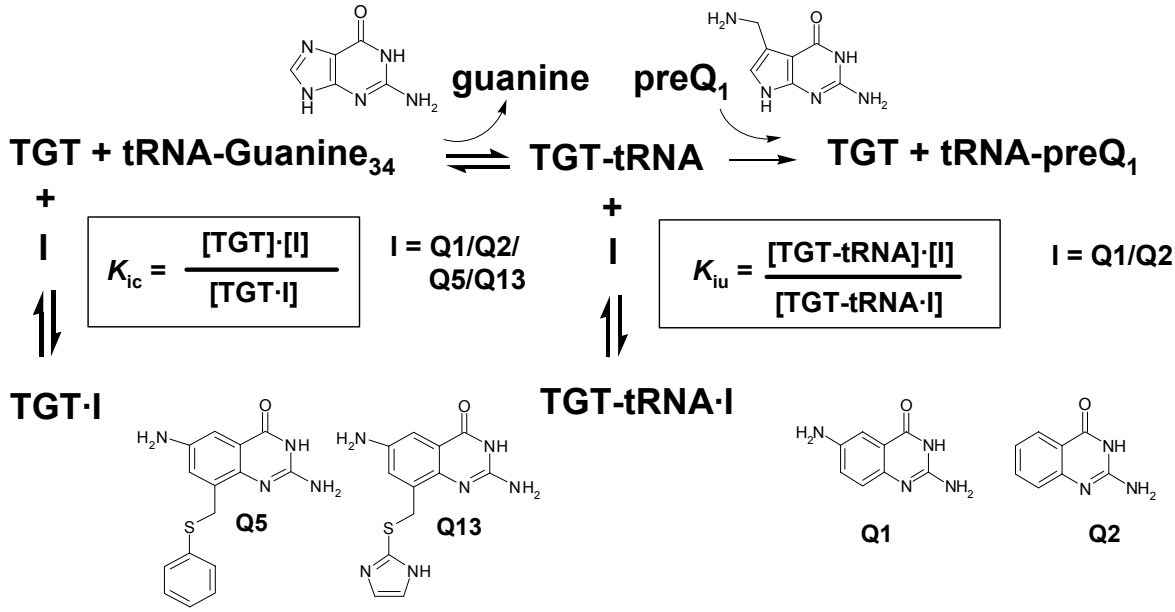



**Fig. 3.3** Size-dependent competitive ($K_{ic}$) and uncompetitive ($K_{iu}$) inhibition modes with respect to tRNA binding

### 3.1.4.1 Determination of non-competitive inhibition

Small-sized inhibitors, such as **Q1 and Q2**, act as competitive and uncompetitive inhibitors with respect to tRNA binding. They can be identified by trapping experiments (chapter 3.1.3). The competitive/uncompetitive inhibition process can be described by a double reciprocal *Lineweaver-Burk* plot according to equation (1).

$$\frac{1}{v_0} = \left[ \frac{K_m}{V_{max}} \cdot \left(1 + \frac{[I]}{K_{ic}}\right) \right] \cdot \frac{1}{[S]} + \frac{1}{V_{max}} \cdot \left(1 + \frac{[I]}{K_{iu}}\right) \qquad (1)$$

$v_o$: initial velocity of [8-$^3$H]-guanine incorporation, $K_m$ of tRNA$^{Tyr}$, $V_{max}$: maximum velocity of the uninhibited [8-$^3$H]-guanine incorporation, [I]: inhibitor concentration, $K_{ic}$: competitive inhibition constant, [S]: tRNA concentration, $K_{iu}$: uncompetitive inhibition constant

This relationship exhibits two inhibition constants, $K_{ic}$ and $K_{iu}$. While competitive inhibition results in a raise of $K_m$, uncompetitive inhibition decelerates $V_{max}$. Both inhibition constants contribute to the total inhibition of TGT with respect to tRNA binding. To describe the suppression of the readout of the assay, namely the velocity of the incorporation of [8-$^3$H]-guanine in tRNA replacing unlabeled guanine ($v_o$), both

have to be considered. In the literature, such a composite type of inhibition is referred to as mixed or non-competitive inhibition [Segel, 1993; Bisswanger, 2000][81, 82].

The two inhibition constants can be calculated using the tritium-labelling assay by monitoring the initial velocity $v_o$ of the base exchange reaction at various inhibitor and tRNA substrate concentrations. $K_{ic}$ rises $K_m$ towards the apparent value $K_m^{app}$ and $K_{iu}$ reduces $V_{max}$ towards the apparent value $V_{max}^{app}$. By applying formula (2) and (3) the inhibition constants $K_{ic}$ and $K_{iu}$ can be calculated from the kinetic parameters of the uninhibited reaction together with the apparent kinetic parameters determined in the presence of inhibitor.

$$K_m^{app} = K_m \cdot \left(1 + \frac{[I]}{K_{ic}}\right) \tag{2}$$

$$V_{max}^{app} = \frac{V_{max}}{\left(1 + \frac{[I]}{K_{iu}}\right)} \tag{3}$$

### 3.1.4.2 Determination of competitive inhibition

For inhibitors with a large side chain such as **Q5** or **Q13** the competitive inhibition constant $K_{ic}$ is supposedly sufficient to fully describe their inhibition properties as these compounds are too large to be significantly accommodated by the binding pocket as long as the covalent TGT-tRNA adduct is formed. $K_{iu}$ is hardly relevant in this case, and the respective inhibition type can be described as predominantly competitive. Equation (4) describing the latter situation results from a simplification of equation (1). The competitive inhibition constant $K_{ic}$ either for guanine-sized or substituted inhibitors describes in both cases the potential to impede tRNA binding to TGT. Thus, the $K_{ic}$ values for both types of inhibitors are directly comparable.

$$\textbf{(4)} \quad \frac{1}{v_0} = \left[ \frac{K_m}{V_{max}} \cdot \left(1 + \frac{[I]}{K_{ic}}\right) \right] \cdot \frac{1}{[S]} + \frac{1}{V_{max}} \tag{4}$$

The procedure to determine the inhibition constants follows the protocol developed by Grädler *et al*. (2001)[3], however, with minor modifications. As limiting substrate now $tRNA^{Tyr}$ is used and for the calculation of the inhibition constants the $K_m$ value of $tRNA^{Tyr}$ in the presence of *Tween 20* is applied (Tab. 3.1).

A detailed discussion of the newly determined inhibition constants of **Q1**, **Q2**, **Q5** and **Q13** will be presented in chapter 3.2.3.

### 3.1.5 SPR-based binding assay

The tritium-labelling assay follows a sophisticated and time-consuming procedure. The number of compounds for which inhibition constants can be determined is limited to approximately two per day. Furthermore, the amounts of consumed material and its disposal (radioactive waste) makes the procedure quite expensive. Therefore, it is an important goal to search for alternative binding assays. Initial attempts to establish a fluorescence based assay were not successful [Brenk, PhD Thesis, 2003][52].

Another alternative is the development of a surface plasmon resonance (SPR)-based binding assay [Brenk, PhD Thesis, 2003][52]. The surface plasmon resonance spectroscopy allows to follow the interactions of molecules. One of the interaction partners is immobilized on a chip with a gold surface. Interactions of the mobile partner with the immobilized partner change the refraction index of total refraction in the solution next to of the gold surface. This change can be detected and used for the calculation of binding constants [Rich & Myszka, 2000][83].

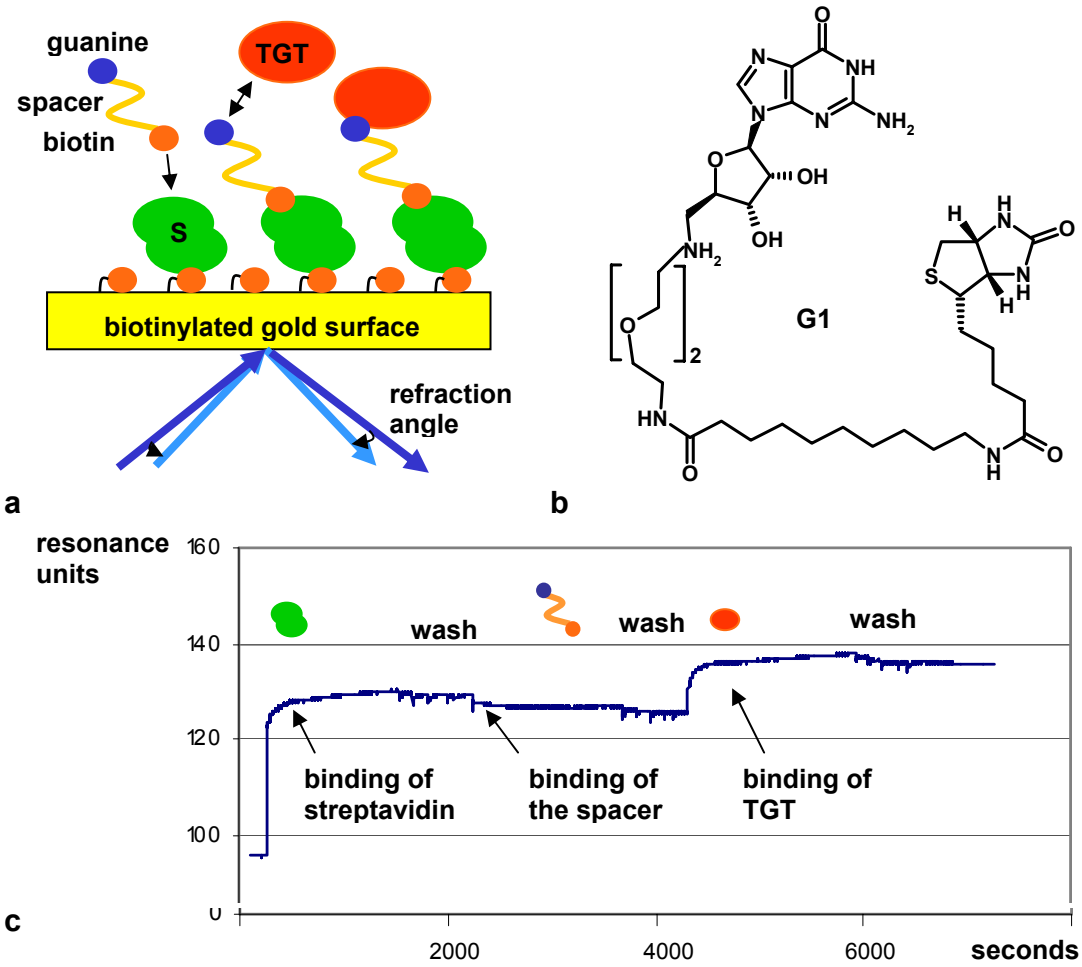



**Fig 3.4** a) SPR-based assay design using TGT and streptavidin (S); b) chemical structure of the linker L1; c) results from an SPR-based binding experiment using linker G1.

To adjust this method for TGT, an assay system was developed (Fig. 3.4a). The selected interaction partners for TGT is streptavidin. This protein is a strong binder of biotin and usually forms homodimers. To enable binding of TGT to streptavidin a set of spacer molecules was synthesized by C. Herforth (group of Prof. Link, University of Greifswald). The terminus of the spacer molecule **G1**, labelled by a guanosine moiety ('G' for guanosine), is suited to bind to TGT. The other end bears a biotin moiety, suited to bind to streptavidin (Fig. 3.4b). The spacers bridging between both ends had a length that should allow binding to both active sites. In an initial test using the tritium-labelling assay **G1** was shown to bind to TGT.

The assay procedure was designed as following (Fig. 3.4c). In a first step, streptavidin is immobilized on a biotinylated gold surface. In a second step, the linker is added. With its biotinylated end it binds to the remaining free binding pocket of streptavidin, then TGT is added. Due to the binding of TGT to the guanosine terminus of the linker, the refraction of the gold surface is altered, which can be monitored. In the presence of inhibitor less TGT should be available for binding to **G1**, thus a smaller change of refraction should be monitored. Unfortunately, in initial attempts no binding of TGT to the linker could be observed [Brenk, PhD Thesis, 2003][52].

In collaboration with M. Hartmann (group of Prof. M. Keusgen, University of Marburg) the experiment with **G1**, the compound with the longest available spacer, was repeated (chapter 5.4). In these experiments binding of 10pg TGT / $mm^2$ to the gold surface could be monitored. This finding demonstrates, that this assay is principally suited to replace the radioactive assay. However, one major problem remains. The SPR-assay is only suitable for aqueous solutions and does not tolerate DMSO. Due to the low solubility of many of the present TGT inhibitors, DMSO has to be added to the assay solutions. Thus, before inhibition can be recorded either DMSO has to be substituted or more soluble inhibitors have to be developed. Nevertheless, this technique provides a promising starting point for a new assay setup.

# 3.2 Revalidation of structure – affinity data

Due to the necessary modifications concerning the assay protocol, inhibition constants for already crystallized and characterized compounds had to be re-determined and their structure – activity relationship had to be revised. Competitive and uncompetitive contributions towards inhibition with respect to tRNA binding have to be considered separately (see chapter 3.1 and 5.3.1 - 5.3.4). The previously published values refer to competitive inhibition with respect to guanine binding. Therefore, they are not comparable to the values resulting from the modified method. Due to the elaborate assay procedure it was not yet feasible to redetermine inhibition constants for all previously reported compounds. Although the previously determined values are not discriminative with respect to the inhibition mode, they roughly reflect the trends in the binding properties correctly. In the following tables, values referring to the former method are marked with a star (*), values referring to the modified method are given in bold.

In the first section of this chapter, binding competent conformations of the TGT binding pocket will be systematized. Due to structural adaptations of the binding pocket upon inhibitor binding, identified in previous studies, four relevant conformations have to be considered (chapter 3.2.1). A quinazolinone based inhibitor series, developed in a previous study, served as a starting point for the development of new TGT inhibitor series (chapter 3.3). Accordingly, reevaluation of inhibition constants and structure – activity relationship of virtually all compounds of this series is relevant for further design approaches (chapter 3.2.3). Of other series only selected representatives were tested (chapter 3.2.2).

## 3.2.1 TGT binding pocket conformations

The TGT binding pocket performs some pronounced adaptations upon substrate or inhibitor binding. Deviating binding pocket geometries are relevant for different inhibitor classes. The combinations of two adaptive sites result in four binding pocket geometries relevant for inhibitor binding which are systematically presented in Figure 3.5. These conformations have been observed in crystal structures of pyridazindione- and quinazolinone-based inhibitor classes in complex with TGT [Grädler *et al*., 2001; Brenk *et al*., 2003; Brenk *et al*., 2004; Meyer *et al*., 2004][1-3, 66]. The first adaptation

results from the Leu231/Ala232 peptide switch, which has already been described in chapter 2.2. Either a carbonyl donor or an amide accepter functionality is exposed towards the binding pocket and modifies the size of the pocket. The second adaptation refers to Asp102. In uncomplexed structures, this residue points out of the binding pocket and is involved in an H-bonding network with Asn70 and Thr47. Upon substrate binding of e.g. $preQ_1$ the side chain of Asp102 rotates into the binding pocket and forms two hydrogen bond with the substrate. As a consequence, the size of the guanine recognition pocket is significantly decreased, distinct active site water molecules are expelled and the properties exposed to the binding site are fundamentally altered. Furthermore, the rotational movement of Asp102 provokes a reorientation / disordering of the side chains of Asn70 and Thr47 [Brenk *et al*., 2003; Brenk *et al*., 2004][1, 64].

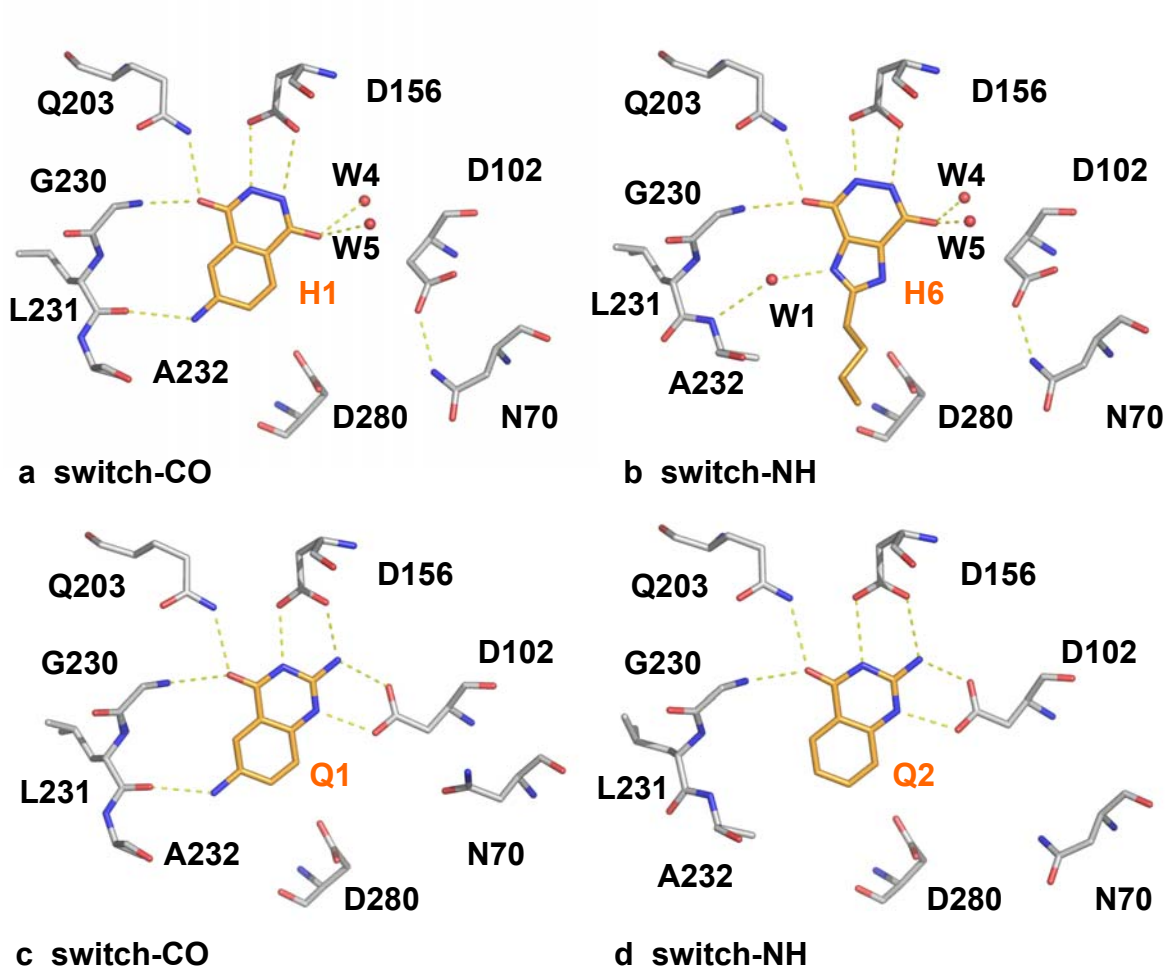



**Fig. 3.5** Binding pockets of TGT relevant for inhibitor binding differ in the Leu231/Ala232 peptide switch and in the orientation of Asp102 (in/out)

## 3.2.2 Revalidation relevant compound classes

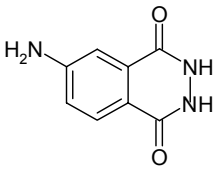
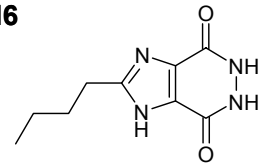
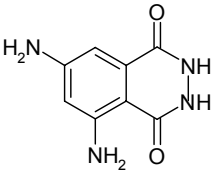
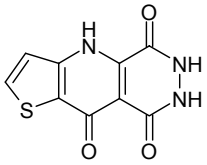
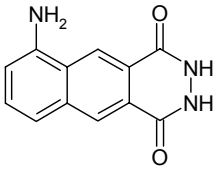
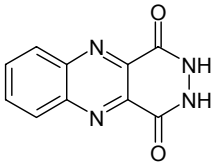
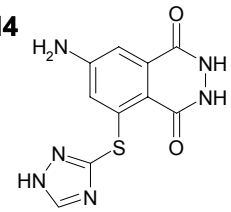
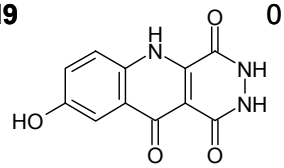
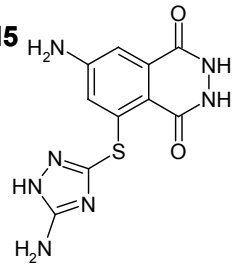
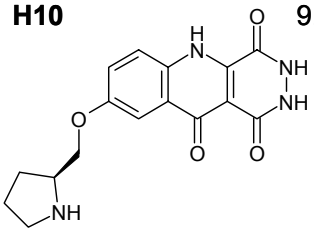
### 3.2.2.1 Pyridazindione-based inhibitor series

In the first rational design approach performed with *Z. mobilis* TGT, the pyridazindione **H1** ('H' for hydrazide) has been discovered as a promising starting point for the development of TGT inhibitors [Grädler *et al*., 2001][3]. This compound showed an inhibition constant of 8.3 µM, following the former assay, and crystal structure analysis in complex with TGT was successful (Fig. 3.5a). Due to its lead-like scaffold the structure was used for further design. Modifications resulted in compound **H2** and **H3** with similar binding properties as **H1** (Tab. 3.2).

All compounds bind to the Leu231 carbonyl exposing / Asp102out conformation with the waters W4 and W5 present in the binding pocket (Fig. 3.5a). Attempts, to address the catalytic nucleophile Asp280 by substituting the scaffold **H1** had only limited success. Crystal structures of **H4** and **H5** in complex with TGT revealed that the triazole-based substituents were not well oriented in this pocket. Both compounds were weaker binders of TGT [Grädler *et al*., 2001][3].

Searches for alternative pyridazindione-type leads in the NNC-database revealed **H6** and **H7** as the most promising hits [Brenk *et al*., 2003][2]. With a inhibition constant of 83µM, following the former assay, **H6** was a ten-fold weaker binder than **H1** (Tab. 3.2). Nevertheless, a well defined crystal structure in complex with TGT could be obtained of compound **H6** [Brenk *et al*., 2003][2]. Surprisingly, it shows a modified binding mode compared to **H1**. The binding pocket adopts the Ala232 amide exposing / Asp102out conformation and an interstitial water molecule (W1) is bridging towards Ala 232 (Fig. 3.5b). This crystal structure was very important for the investigation of the base exchange mechanism and substrate promiscuity (chapters 2.1 and 2.2). Of **H7** no crystal structure in complex with TGT could be obtained, probably due to the low solubility of the compound. Nevertheless, **H7** was used as lead for a structure-based design approach [Brenk *et al*., 2003][84]. Modifications of the scaffold resulted in **H8** and **H9** with similar or slightly improved affinity. Substitution of **H8** resulted in **H10**, but no stronger binding was observed. Due to the low solubility in water, of **H8** – **H10** no crystal structures in complex with TGT are available.

**Tab. 3.2** Pyridazindione-based inhibitors

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **H1** |  | **75 ± 9**<br>8.3 ± 0.4* | 1.95Å<br>1ENU | **H6** |  | **62 ± 40**<br>83 ± 18* | 2.1Å<br>1N2V |
| **H2** |  | **73 ± 13**<br>0.3 ± 0.1* | 1.95Å<br>1F3E | **H7** |  | 5.0 ± 1.2* | unpub.<br>(ln01) |
| **H3** |  | 0.3 ± 0.1* | 1.95Å<br>unpub.<br>(ug01) | **H8** |  | 5.6 ± 0.4* | --- |
| **H4** |  | 54 ± 14* | 2.1Å<br>unpub.<br>(hd03) | **H9** |  | 0.7 ± 0.1* | --- |
| **H5** |  | 38 ± 1* | 1.4Å<br>unpub.<br>(hd04) | **H10** |  | 9.4 ± 0.1* | --- |

$K_{ic}$ in [µM] with average error; (*): former assay; bold: modified assay;
crystal structures: PDB-code and maximum resolution

Inhibition constants for **H1**, **H2** and **H6** were revalidated following the new assay procedure. Trapping experiments revealed that uncompetitive inhibition (stabilization of the covalent TGT-tRNA complex) is not relevant for this compound class, although similar molecular dimensions as $preQ_1$ are given. This may result from the fact that in the tRNA-bound complex Asp102 is rotated into the active site (Fig. 3.1b). The pyridazindiones, however, require Asp102 to point out of the binding pocket (Fig. 3.5b/d). The redetermination of competitive inhibition constants revealed that the pyridazindiones were overestimated in their binding properties by the previously used assay. The affinity differences, observed for **H1** and **H6**, resulting from the Leu231/Ala232 peptide switch could not be confirmed. Instead, the new data

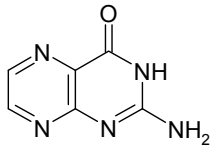
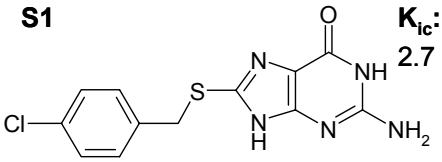
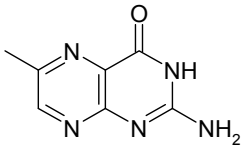
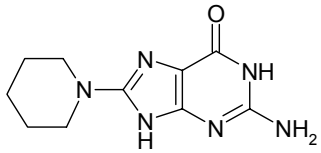
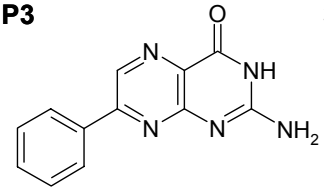
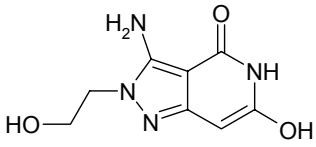
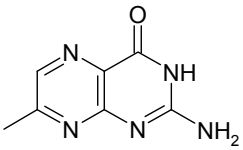
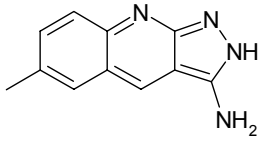
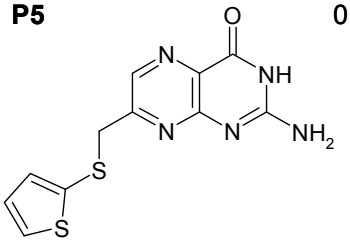
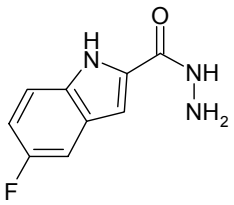
suggests that the orientation of the Leu231/Ala232 peptide switch has only minor influence on binding affinity.

#### 3.2.2.2 Pteridines and virtual screening hits

A virtual screening campaign was performed on *Z. mobilis* TGT based on the crystal structure of TGT·**H6** and TGT·**H1** to create a composite pharmacophore hypothesis [Brenk *et al*., 2003][2]. It included the binding pocket information for the Leu231/Ala232 peptide switch in both orientations. Nine compounds were finally selected and tested for TGT inhibition (Tab.3.3). Pteridine-based ligands were retrieved which showed micromolar inhibition (Tab. 2.4; **P1** - **P3**; 'P' for pteridine). Further groups of compounds had a guanine like scaffold (**S1** – **S3** ; 'S' for screening) or only low similarity to the natural substrates (**S4**; **S5**). They all were significantly weaker binders compared to the pteridines. For all compounds no crystal structure in complex with TGT could be obtained, most probably due to their low solubility.
Pteridine **P1** was used as scaffold for a structure-based design approach [Brenk *et al*., 2003][84]. First **P4** was synthesized and tested. Although a slight affinity decrease was observed, this compound provided a starting point for further substitutions in position 7 to address the hydrophobic ribose 34 binding pocket formed by Val45 and Leu68. Of all tested compounds **P5**, substituted with a hydrophobic thiophene substituent showed the best affinity. Also for these compounds crystal structure analysis in complex with TGT was not successful.

Due to the structural similarity of **P1** with guanine, this compound was tested for non-competitive inhibition. In a trapping experiment **P1** was identified to stabilize the covalent TGT-tRNA intermediate. With an uncompetitive inhibition constant of 0.7 µM is one of the most potent uncompetitive inhibitors known to date. The competitive inhibition constant of 2.2 µM indicates that a guanine-type scaffold structure is more favourable for binding than the pyridazindione scaffold. Compared to the pyridazindione scaffold **H1** a gain of affinity by a factor of 40 with respect to the competitive inhibition constant is observed for **P1**. For compound **S1** the trapping experiment revealed that uncompetitive inhibition is not relevant. With respect to competitive inhibition (51 µM) this guanine derivative experiences a significant loss in affinity compared to **P1**.

**Tab. 3.3** Pteridine-based inhibitors and virtual screening hits

| | | | | |
|---|---|---|---|---|
| **P1** |  | **$K_{ic}$: 2.2 ± 0.6**<br>**$K_{iu}$: 0.7 ± 0.2**<br>0.6 ± 0.2* | **S1** |  | **$K_{ic}$: 51 ± 13**<br>2.7 ± 0.3* |
| **P2** |  | 0.25 ± 0.1* | **S2** |  | 37 ± 7* |
| **P3** |  | 3.8 ± 0.1* | **S3** |  | 8.1 ± 1.0* |
| **P4** |  | 5.6 ± 1.7* | **S4** |  | 156 ± 36* |
| **P5** |  | 0.45 ± 0.05* | **S5** |  | 72 ± 5* |

$K_{ic}$ in [µM] with average error; (*): former assay; bold: modified assay;

## 3.2.3 Revalidation of quinazolinones

### 3.2.3.1 Development of the quinazolinone-based inhibitor series

The most potent TGT inhibitors were derived from the class of quinazolinones (3*H*-chinazolin-4-on). This class was designed in a collaborative project with E. Meyer (group of Prof. F. Diederich, ETH Zürich). The compounds were synthesized in Zürich, structural and initial kinetic analysis were performed in Marburg by R. Brenk (group of Prof. G. Klebe, University of Marburg).

Quinazolinones are characterized by inhibition constants in the lower micromolar range and suited for crystal structure analysis in complex with TGT [Meyer *et al.*,

2002; Brenk *et al*., 2004, Meyer *et al*., 2004][1, 66, 85]. The lead structure **Q1** was developed *via* the combination of structural elements of **H1** and $preQ_1$ (Tab. 3.4). With an inhibition constant of 0.35 µM, following the former assay, **Q1** appeared as a potent inhibitor. Crystal structure analysis of TGT·**Q1** revealed a binding mode similar to **TGT·H1** (Fig. 3.5a/c). However, one major structural difference could be observed. Asp102 is rotated into the binding pocket and engages in a bifurcated hydrogen bond to **Q1**. This binding mode resembles that of $preQ_1$. Upon rotation of Asp102 the water molecules W4 and W5 are expelled from the binding site.

**Q2** lacks the 6-amino group of **Q1**. It was synthesized to probe for the relevance of the exocyclic amino group for binding [Meyer *et al*., 2004][66]. With an inhibition constant of 20 - 50 nM, following the previous assay, it was a surprisingly tight binder, although in structural terms this could not be explained. The crystal structure of TGT·**Q2** revealed a deviating binding mode compared to **Q1** (Fig. 3.5d). This compound binds to the peptide switch with the amide oriented to the binding pocket. In contrast to the crystal structure of TGT·**H6** the interstitial water molecule W1 is not found in this structure.

Compounds **Q1** and **Q2** were used as starting points to further explore the binding site properties and to develop more potent inhibitors [Meyer *et al*., 2002; Brenk *et al*., 2004, Meyer *et al*., 2004][1, 66, 85]. Decorations at position 8 were intended to address the hydrophobic pocket formed by Val45 and Leu68. **Q3** and **Q4** served as starting points for a series of compounds modified in this position (Tab. 3.4). Repulsive interactions of the 8-methyl group of **Q4** with Asp102 were identified in the TGT·**Q4** crystal structure. With an inhibition constant of 7µM, following the previous assay, **Q4** exhibited a significant loss of affinity [Meyer *et al*., 2004][66].

Further substituted compounds in position 8 can be classified in two series with respect to the spacer and side chain properties. In the first series compounds with different spacers have been synthesized. Ethyl, sulfanylmethyl and phenoxymethyl spacers have been tested to link the scaffold with a phenyl substituent (Tab. 3.5). The second series is based on 6-amino-quinazolinone with a sulfanylmethyl spacer. In this series side chains with aliphatic, aromatic and polar substituents have been investigated (Tab. 3.6).

Modifications at position 6 of the quinazolinone scaffold were synthesized to study the relevance of the peptide switch for inhibitor binding (Tab. 3.7). For synthesis quinazolinone with a phenylsulfanylmethyl substituent was used as core fragment.

TGT was identified to accept a broad variety of functionalities in this position (-CN, -$NH_2$, -OH, -H, -Br, $CH_2NH_2$).

Crystal structure analyses of substituted inhibitors in complex with TGT was successful only in case of **Q12**, **Q13**, **Q14** and **Q15** as a result of enhanced solubility [Brenk *et al*., 2004][1]. Although the quinazolinone scaffold is well defined in the binding pocket in all structures, the ligand side chains are not properly ordered. While for the spacer split conformations could be observed in some cases, additional side chains are disordered in all cases (Fig. 3.6).

To estimate the relevance of uncompetitve inhibition for small sized inhibitors and the impact of ther detergent *Tween 20* for substituted inhibitors, the revalidation of most compounds appeared essential. The revalidated affinity data was published by us in Meyer *et al*. (2006)[65].

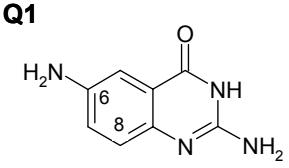
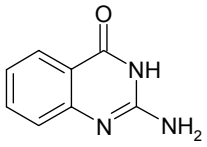
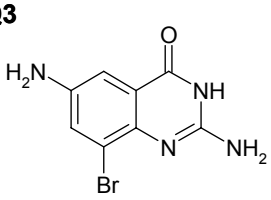
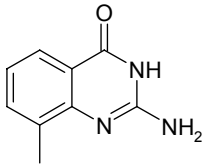
#### 3.2.3.2 Non-competitive inhibition by small-sized quinazolinones

Ligands **Q1** and **Q2** were identified by the trapping experiment as uncompetitive inhibitors (see chapter 3.1.3). The $K_{iu}$ constants indicate decreasing power to stabilize the covalent complex from **Q1** (0.6 µM) to **Q2** (1.7 µM) (Tab. 3.4). This descending order is also reflected by the intensities of the bands at 70 kDa in the SDS-PAGE of the trapping experiment (Fig. 3.2). Facing the competitive inhibition constants ($K_{ic}$), **Q2** (2.1 µM) shows a similar binding potency compared to **Q1** (1.5 µM). From these findings it can be concluded that the previously reported, very low inhibition constants of **Q1** (350 nM) and **Q2** (20 - 50 nM), [Meyer et al. 2004], resulted from neglecting the uncompetitive inhibition contribution in the former assay.

**Q1** and **Q2** exhibit a similar uncompetitive inhibition potency as **P1** ($K_{iu}$: 0.7 µM; Tab. 3.3). Obviously, this results from their similar chemical structures allowing to stabilize the covalent TGT-tRNA intermediate state. Also the competitive inhibition constants for **P1**, **Q1** and **Q2** are in a similar range. Compared to the pyridazindiones **H1** and **H6** ($K_{ic}$: 60 - 80 µM) they gain affinity by a factor of 40 - 60 (Fig. 3.2). The comparison of the respective crystal structures in Figure 3.5 indicates that a guanine-type scaffold structure is more favourable for binding by allowing Asp102 to rotate into the binding pocket and providing a better binding geometry for Asp156 .

To evaluate the impact of 8-substitution at the quinazolinone scaffold, **Q4** was reevaluated as well (Tab. 3.4). It also shows non-competitive inhibition, but the competitive contribution dominates ($K_{ic}$ = 3.7 µM vs. $K_{iu}$ = 19.1 µM). Consistently, the 70 kDa covalent bands in the respective trapping experiment are less pronounced than for **Q2** (data not shown). In structural terms this observation can be explained by spatial conflicts with the attached methyl group in **Q4** most likely interfering with the ribose ring 34 of the covalently bound tRNA (c.f. Fig. 3.1b). This effect is even more pronounced for inhibitors bearing larger side chains at this position, which was confirmed by trapping experiments (Fig. 3.2). Therefore, the impact of uncompetitive inhibition is further reduced and can be neglected in the assay determination procedure (for more details see chapter 3.3.2.1).

**Tab. 3.4** Competitive and uncompetitive inhibition constants of small-sized inhibitors

| | | | | | |
|---|---|---|---|---|---|
| **Q1** |  | **$K_{ic}$: 1.5 ± 0.4**<br>**$K_{iu}$: 0.6 ± 0.2**<br>0.35 ± 0.12* | **Q2** |  | **$K_{ic}$: 2.1 ± 0.5**<br>**$K_{iu}$: 1.7 ± 0.4**<br>0.02-0.05* |
| **Q3** |  | 9.1 ± 1.0* | **Q4** |  | **$K_{ic}$: 3.7 ± 0.9**<br>**$K_{iu}$: 19.1 ± 4.8**<br>7.0 ± 1.8* |

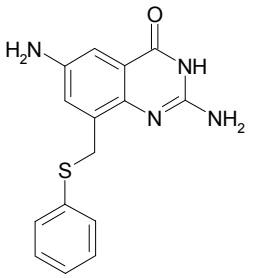
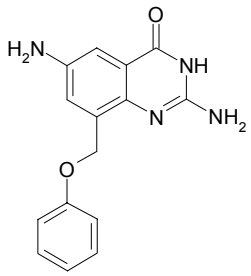
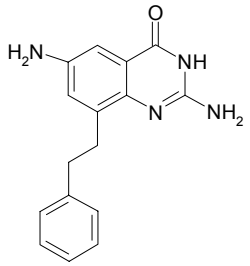
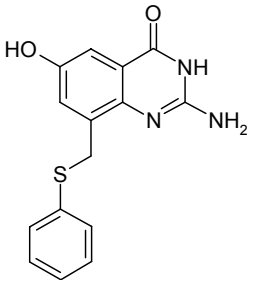
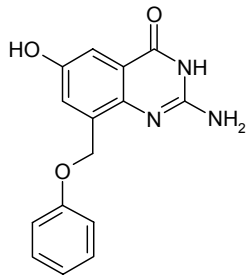
$K_{ic}$ in [µM] with average error; (*): previous assay; bold: modified assay

### 3.2.3.3 Revalidation of substituted quinazolinones

All substituted quinazolinones were identified as primarily competitive inhibitors *via* trapping experiments. The revalidation of inhibition constants revealed that they were modulated in a non-linear fashion, most probably due to contributions of non-specific inhibition resulting from low solubility [Meyer *et al*., 2006][65]. In consequence, the previously made assumptions about structure – activity relationship had to be revised and corrected also for substituted inhibitor series [Brenk *et al*., 2004; Meyer *et al*., 2004; Meyer *et al*., 2002][1, 66, 85].

Especially the previously reported large difference in binding affinity between S-, O- and C-atoms in the linker, that prompted to postulate a sulphur effect, is no longer evident [Meyer *et al*., 2002][85]. Instead **Q7**, the compound with an ethylene linker, turned out to be the most potent inhibitor of this series, however with only a slight affinity advantage (Tab. 3.5).

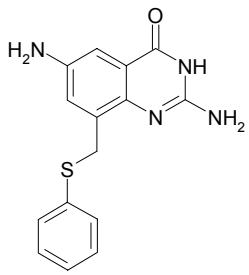
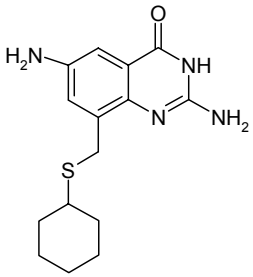
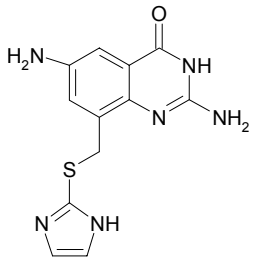
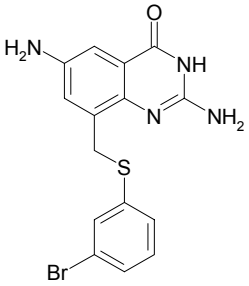
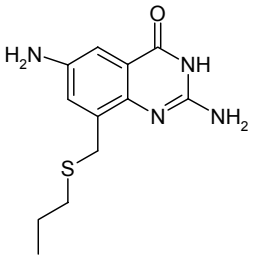
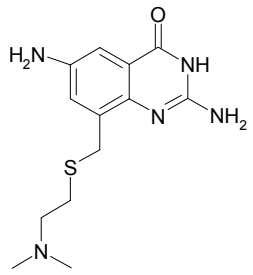
**Tab. 3.5** Revalidation of 6-aminoquinazolinones – series with different spacers

| | | | | | |
|---|---|---|---|---|---|
| **Q5** | **3.8 ± 0.4**<br>0.1 ± 0.02* |  | **Q6** | **5.7 ± 1.4**<br>5.6 ± 0.4 * |  |
| **Q7** | **2.6 ± 0.9**<br>3.6 ± 1.2* |  | | | |
| **Q8** | **5.1 ± 1.3**<br>0.25 ± 0.05* |  | **Q9** | **8.7 ± 3.0**<br>4.6 ± 1.4* |  |

$K_{ic}$ in [µM] with average error; (*): previous assay; bold: modified assay

Additionally, the revalidation showed that compounds with substituents other than unsubstituted phenyl rings experience reduced affinity (Fig. 3.6). Whereas in the previous assay no preference for aromatic, aliphatic and polar substituents could be identified, the redetermined values indicates a preference for the phenyl-substituted compounds. Compared to **Q5**, aliphatic substituents (**Q12**) and polar substituents (**Q13**, **Q14**) loose affinity by factors of four to nine. But also further substitution of the phenyl ring of **Q5** resulted in significantly weaker binding (**Q10**).

**Tab. 3.6** Revalidation of series with aromatic, aliphatic and polar side chains

| | | | | | |
|---|---|---|---|---|---|
| **Q5** | **3.8 ± 0.4**<br>0.1 ± 0.02* |  | **Q11** | 5.4 ± 0.6 * |  |
| **Q13** | **34 ± 9**<br>1.4 ± 0.6* |  | | | |
| **Q10** | **23 ± 25**<br>0.6 ± 0.1* |  | **Q12** | **14 ± 4**<br>7.7 ± 1.6* |  |
| **Q14** | **22 ± 6**<br>3.5 ± 0.9* |  | | | |

$K_{ic}$ in [µM] with average error; (*): previous assay; bold: modified assay

With respect to modifications in position 6 the enzyme accepts a diverse set of functionalities without showing dramatic changes in binding (Tab. 3.7). The structural basis for this promiscuity, associated with the Leu231/Ala232 peptide switch, has already been discussed in detail in chapter 2.2 and supports the results form the pyridazindione-based inhibitor series (chapter 3.2.2). The only exception in the quinazolinone series is compound **Q15** with a four-fold affinity loss compared to **Q5**. However, due to the enhanced solubility, most likely resulting from the charged amino-methyl group, crystal structure analysis of the compound in complex with TGT could be successfully performed [Brenk *et al.*, 2004][1]. It indicates repulsive van der Waals contacts of the 6-methyl group with the carbonyl group of Leu231 to be responsible for the affinity loss (Fig. 3.6) .

**Tab. 3.7** Quinazolinones – modifications in position 6 to address the peptide switch by Leu231/Ala232

| | | | | | |
|---|---|---|---|---|---|
| **Q5** | **3.8 ± 0.4**<br>0.1 ± 0.02* | **Q8** | **5.1 ± 1.3**<br>0.25 ± 0.05 * | **Q15** | **16 ± 4**<br>1.7 ± 0.4* |
| **Q16** | **4.0 ± 1.4**<br>1.1 ± 0.3* | **Q17** | **6.4 ± 1.6**<br>4.1 ± 1.2* | **Q18** | 1.1 ± 0.05* |

$K_{ic}$ in [µM] with average error; (*): previous assay; bold: modified assay;

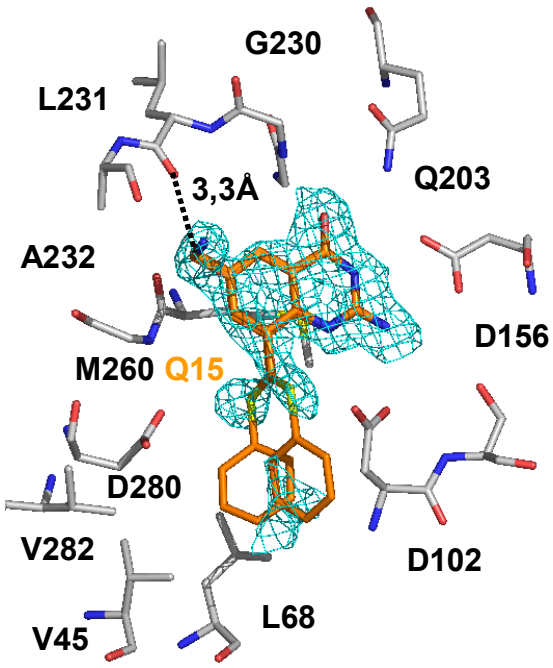



**Fig. 3.6** Crystal structure of **Q15** in complex with TGT contoured at 1.0σ in the 2 $|F_o|$ - $|F_c|$ density map (adapted from Brenk *et al*. (2004)[1]

# 3.3 Quinazolinone-based inhibitors

The revalidation of structure – affinity data for a quinazolinone-based inhibitors series, as a consequence of the modified assay, was presented in chapter 3.2.3. Based on the quinazolinone scaffold three series of inhibitors have been developed and tested:

7-amino-quinazolinones (3.3.1),
*lin*-benzoguanines (3.3.2) and
2-amino-*lin*-benzoguanines (3.3.3).

All compounds presented in this section were developed in cooperation with Emmanuel A. Meyer and Simone Hörtner (group of Prof. François Diederich, ETH Zürich, Switzerland).

## 3.3.1 7-Amino-quinazolinones

The knowledge from the revalidated quinazolinones served as a starting point for a new design approach. Comparison of **Q1** and **Q2** clearly revealed that the presence of the 6-amino group is not associated with a significant gain in affinity (Tab. 3.7). Possibly, this is due to the structural adaptability of the binding pocket exemplified by the Leu231/Ala232 peptide switch (chapter 2.2). Additionally, substitution of the quinazolinone scaffold in position 8 clearly resulted in an affinity loss, due to repulsive interactions of the substituent with Asp102 [Meyer *et al*., 2004][66]. To circumvent direct repulsion, a set of compounds was developed based on the quinazolinone scaffold **Q2**, substituted in position 7 (Fig. 3.7). An amine function was thought to serve as anchor fragment to address the hydrophobic ribose 34 pocket formed by Val45 and Leu68, avoiding the repulsive interactions observed for substitution in position 8. This compound series was synthesized by Simone Hörtner (group of Prof. F. Diederich, ETH Zürich).

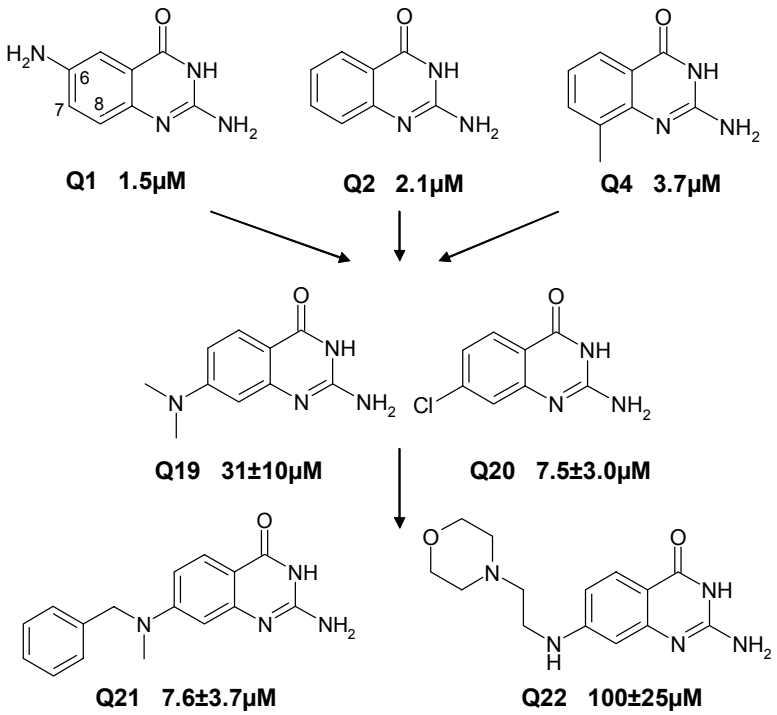


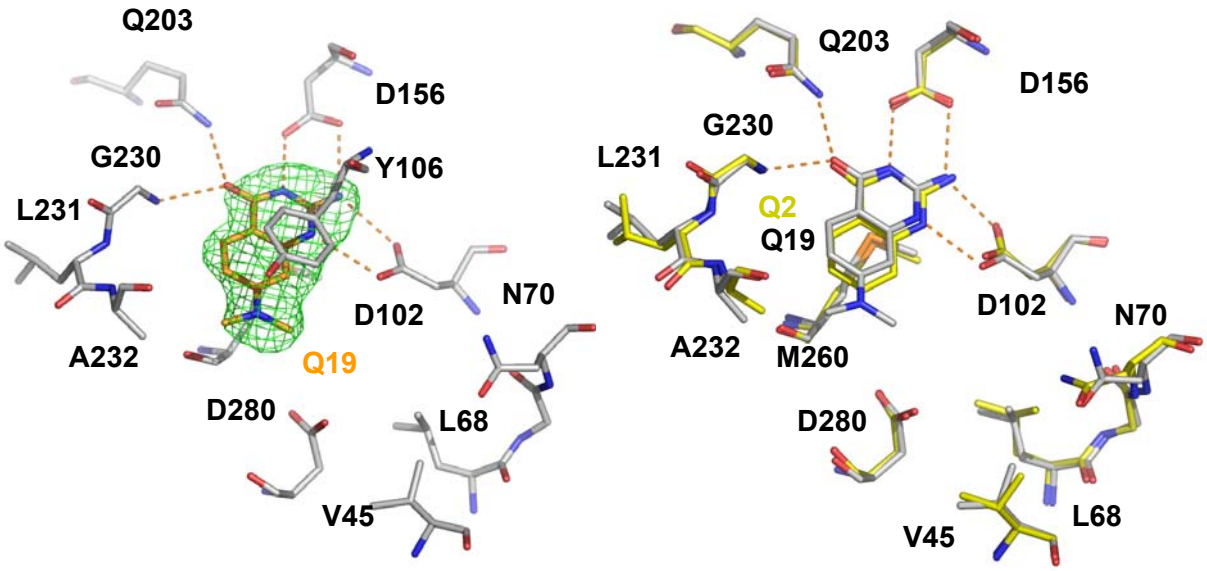

**Fig. 3.7** Design strategy and inhibition constants ($K_{ic}$) for 7-substituted quinazolinones

All newly designed compounds **Q19** - **Q22** were tested for uncompetitive inhibition in trapping experiments. No stabilization of the covalent complex could be observed. Thus, this series represents a set of predominantly competitive inhibitors. Substitution in 7-position seems to interfere significantly with the covalently bound ribose 34 in the TGT-tRNA complex, impeding uncompetitive inhibition. Concerning competitive inhibition all 7-substituted inhibitors showed significantly reduced affinity compared to the parent structure **Q2**. In case of **Q19** affinity decreased by a factor of 15, however phenyl substitution in case of **Q21** could restore affinity by a factor of 4.

Crystal structure analyses have been successfully performed with **Q19** and **Q21** as binary complexes with *Z. mobilis* TGT, resulting in crystals diffracting up to a maximum resolution of 2.0 – 2.15 Å (for methods see chapter 5.5; for crystallographic data see Tab. 6.2.2).

In the crystal structure of TGT·**Q19,** the ligand is sufficiently well defined in the $G_{34}$ binding pocket (Fig. 3.8a). The most pronounced difference electron density is observed for the dimethylamino substituent next to Asp280, while the quinazolinone scaffold is less well defined. The resolution (2.15 Å) and the quality of the crystal structure (R-factor: 23.8 % and $R_{free}$-factor: 35.0 %) is limited with high average B-

values for both protein (58.4 Å$^2$) and ligand (60.6 Å$^2$). Such a reduced quality does not allow an unambiguous placement of crystal water molecules next to the binding pocket.




a b

**Fig 3.8** a) Crystal structure of TGT·**Q19** with the ligand contoured at 2.5σ in the |Fo| - |Fc| difference electron density map refined in the last step excluding ligand coordinates; b) structural superposition of the active sites of TGT·**Q19** and TGT·**Q2**

In the crystal structure of TGT·**Q21** only the quinazolinone scaffold is well defined in the difference electron density. The phenyl substituent is likely scattered over multiple conformations resulting in a rather diffuse difference electron density. In the last refinement cycle, an average B-value of 93.1 Å$^2$ is assigned to its orientation presented in Figure 3.9a. Due to the higher resolution (2.0 Å) and the slightly better quality of the diffraction data (R-factor: 21.5 % and $R_{free}$-factor: 30.3 %) the resulting average B-values for the protein (43.5 Å$^2$) and the 7-amino-quinazolinone scaffold of the ligand (55.3 Å$^2$) adopt more reasonable values compared to TGT·**Q19**. Additionally, it was possible to locate two crystal water molecules in the binding pocket, exhibiting average B-values of 45.0 Å$^2$. The presence of these two water molecules adjacent to the hydrophobic ribose 34 pocket formed by Val45 and Leu68 clearly indicates that the phenyl substituent is exposed to the solvent sticking towards the ribose 33 binding site.

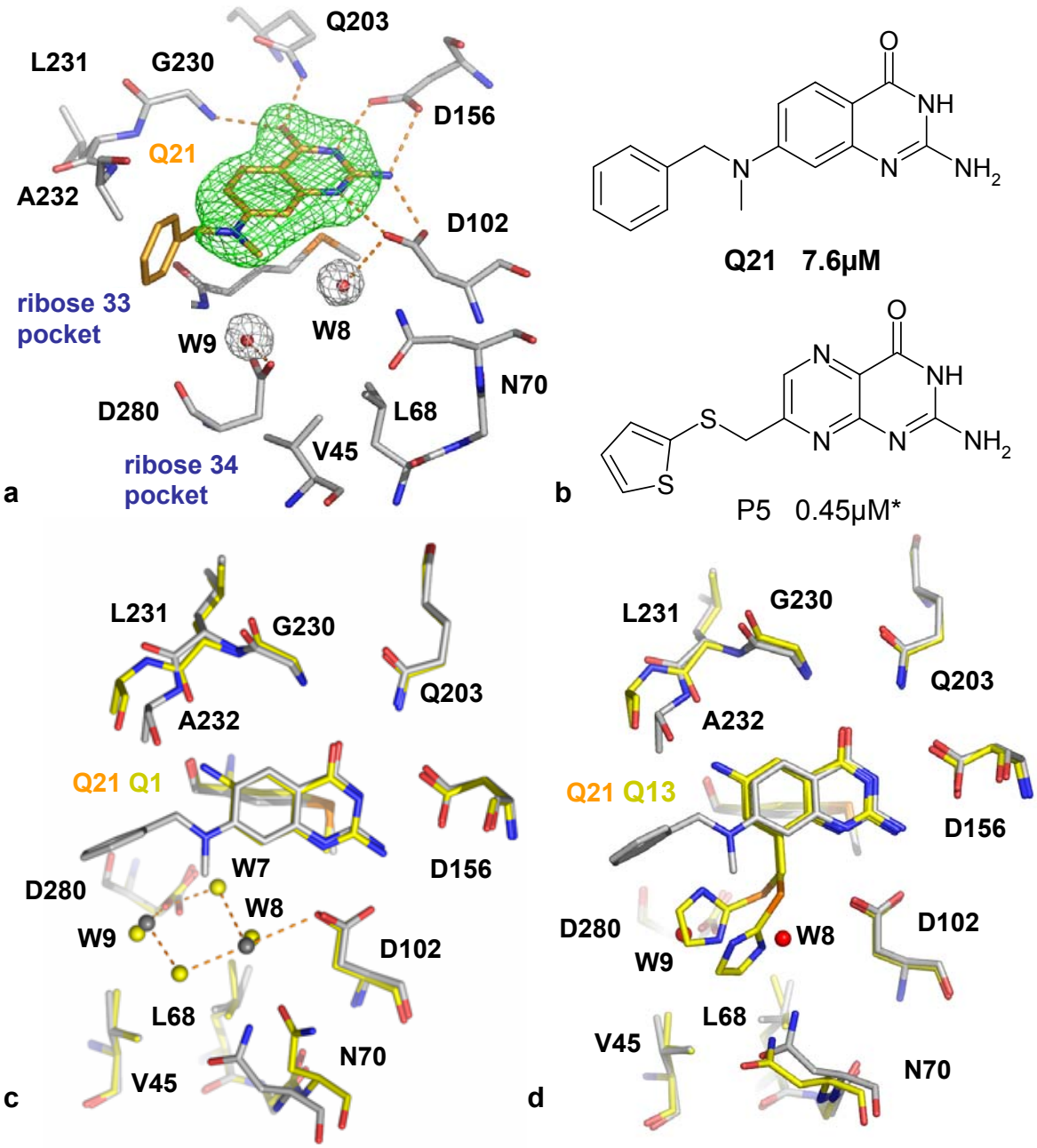



**Fig. 3.9** a) Crystal structure of TGT·**Q21** with the ligand contoured at 2.0 σ in the |Fo| - |Fc| difference electron density map; b) structural similarity of the quinazolinone inhibitor **Q21** and the pteridine inhibitor **P5** (* $K_i$ referring to former assay); c) structural superposition of the active sites of TGT·**Q19** and TGT·**Q1**; d) structural superposition of the active sites of TGT·**Q19** and TGT·**Q13**

The comparison of the binding mode of **Q19** with the binding mode of **Q2** reveals almost identical orientations for the ligands and the active site residues (Fig. 3.8b). Both compounds are bound to a protein conformer with the Leu231/Ala232 peptide switch in its amide exposing geometry. Thus, from the TGT·**Q19** structure it is difficult to extract obvious reasons for the observed affinity loss.

The comparison of TGT·**Q21** with the high resolution structure of TGT·**Q1** (1.2 Å maximum resolution), however, provides an explanation for the affinity loss of the 7-substituted quinazolinones (Fig 3.9c). The obvious difference in TGT·**Q1** is that the Leu231/Ala232 peptide switch is present in the carbonyl exposing conformation. The important difference, however, seems that binding of **Q21** expels a cluster of water molecules mediating interactions between Asp102 and Asp280 from the active site. In TGT·**Q1** these polarizable waters, particularly W7 and W8, are thought to buffer and compensate for the negative charge of the nucleophile Asp280 and Asp102 which recognizes the substrate and catalyzes the proton abstraction. In TGT·**Q21** W7 is expelled from the active site as its distance to the ligand's methyl group would only amount 2.6Å. In consequence, Asp280 looses its H-bonding contact to Asp102 *via* W8.

Crystal structure analyses of TGT·**Q2** and TGT·**Q4** had already shown that the 8-methyl-substituted **Q4** produces unfavourable contacts to the carboxy group of Asp102 and additionally disrupts the water network [Meyer *et al*., 2004][66]. Subtracting the uncompetitive inhibitory contributions reveals an almost two-fold affinity decrease of **Q4** (3.7µM) compared to **Q2** (2.1 µM). Therefore, it can be concluded for TGT that besides unfavourable ligand – protein contacts perturbation of the crystal water network in between Asp280 and Asp102 is detrimental to binding. Both factors have to be compensated by favourable contacts experienced in other regions of the binding pocket.

This assumption is underlined by the superposition of TGT·**Q21** with TGT·**Q13** (Fig. 3.9d). In the latter structure the imidazole substituent penetrates into the hydrophobic ribose 34 pocket and all crystal water molecules are expelled from this site. The binding mode of the most potent substituted quinazolinone **Q7** (2.6 µM) should be similar to that of TGT·**Q13** (34 µM), however the former exposes a phenyl substituent into this pocket. Unfortunately, no crystal structure of **Q7** could be determined. Its competitive inhibition constant is similar to that of **Q1.** Supposedly, the free enthalpy price paid for the replacement of the water cluster next to Asp280 and Asp102 is virtually compensated by additional favourable contacts formed by the added side chain. Obviously, in case of **Q21** replacement of the water cluster is avoided and its phenyl substituent orients towards the ribose 33 binding site which opens to the solvent. Hydrophobic stacking of the **Q21** phenyl substituent with the methyl group of

Ala231 is thought to stabilize this orientation and explains the four-fold affinity gain compared to **Q19**.

The binding mode observed for **Q21** also provides an explanation for the surprisingly low binding constant of **P5** (Fig. 3.9b). With 0.45 µM it is the strongest binder from the pteridine series, although this inhibition constant still refers to the former protocol (Tab. 3.3) [Brenk *et al.,* 2003][84]. Due to the low solubility of **P5** no crystal structure in complex with TGT could be determined. **P5** was designed to bind in the hydrophobic ribose 34 pocket, but the structural similarity of **P5** and **Q21** suggests alike binding modes. Both substituents (thiophen and phenyl ring) are expected to orient towards the ribose 33 binding site minimizing the interference with the water cluster.

In summary, the design concept to substitute the quinazolinone skeleton at position 7 was not as successful as expected with respect to inhibitory potency. However, from the affinity data and the crystal structures in complex with TGT valuable information for further design can be extracted. Firstly, substitution in position 7 hampers uncompetitive binding. All 7-substituted derivatives predominantly compete with tRNA binding. Secondly, the water network in between Asp280 and Asp102 is an important structural element and should only be replaced by polar interactions. Thirdly, addressing the ribose 33 binding pocket instead of the ribose 34 pocket might result in stronger binding due to favourable interactions with Ala232.

### 3.3.2 *lin*-Benzoguanines

The 6-amino-quinazolinone scaffold **Q1** was used as a starting point for the development of a new class of TGT inhibitors. Extended by an imidazole moiety *lin*-benzoguanine (6-aminoimidazol[4,5-*g*]quinazolin-8(7*H*)-one) **L1** was developed as putative scaffold (Tab. 3.8). In case of the quinazolinones, the most potent binders were decorated with aromatic side chains (Tab. 3.5 and 3.6). Therefore, systematic substitutions in 3- and 4-position by aromatic substituents attempted to address the hydrophobic ribose 34 binding pocket formed by Val45, Leu68 and Val282 adjacent to the $G_{34}$ binding pocket (Tab. 3.6). For the 7-amino-quinazolinones it has already been demonstrated, that addressing this pocket is not favourable, due to the perturbation of a water cluster (chapter 3.3.1). But *lin*-benzoguanines and 7-amino-quinazolines were developed and analyzed at the same time. Therefore, the considerations concerning the water cluster are equally relevant also for the *lin*-

benzoguanines. All compounds presented in Table 3.8 were synthesized by E. Meyer (group of Prof. F. Diederich, ETH Zürich). *Lin*-benzoguanines represent potent TGT inhibitors (chapter 3.3.2.1) and allowed successful crystal structure analyses (chapter 3.3.2.2 – 3.3.3.5). The binding of one inhibitor significantly interferes with the integrity of the TGT dimer interface (chapter 3.3.2.6). The inhibitor complexed crystal structures gave insight into the adaptability of the binding pocket and allowed to investigate the structural basis of induced-fit adaptations observed upon tRNA binding (chapter 3.3.2.7 and 3.3.2.8).

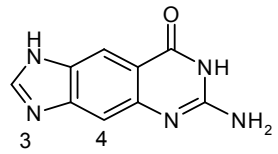
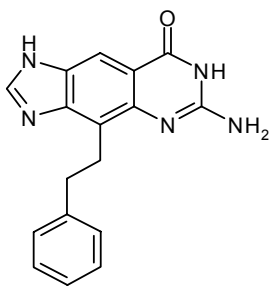
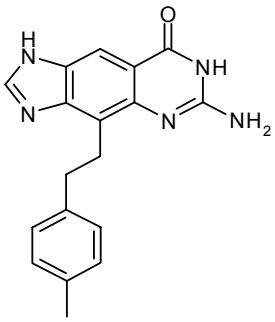
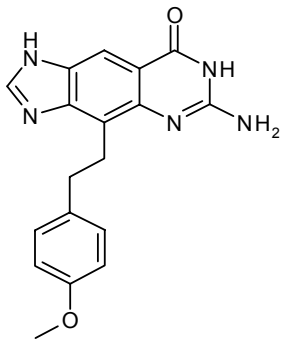
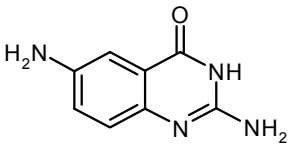
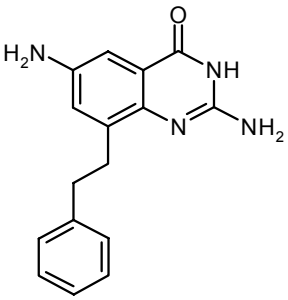
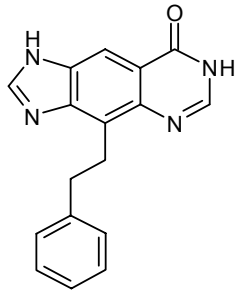
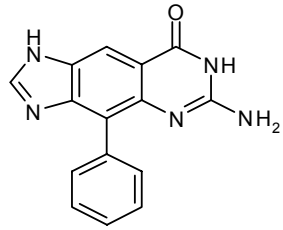
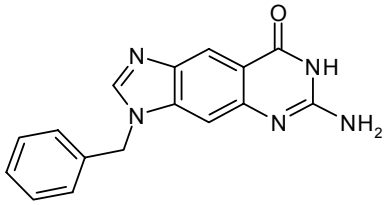
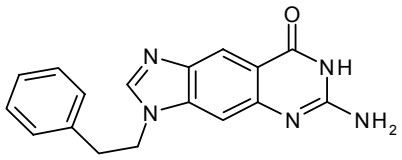
### 3.3.2.1 Inhibition constants

All *lin*-benzoguanine based inhibitors were tested in trapping experiments for uncompetitive inhibition. Uncompetitive inhibition is only relevant for the basic scaffold **L1** ($K_{iu}$: 7.9 µM) although this value is already significantly reduced compared to **Q1** ($K_{iu}$: 0.6 µM) (Tab. 3.8). Obviously, the extended scaffold interferes to some degree with ribose 34 covalently bound in the tRNA-TGT intermediate complex. To assess a possible uncompetitive contribution of substituted *lin*-benzoguanine inhibitors, the $K_{iu}$ of **L2** has been determined. It is separated from competitive tRNA binding by a factor of 50 (Tab. 3.8). Obviously, for substituted compounds uncompetitive inhibition is still possible but on a very reduced level. In consequence, as this contribution appears only marginal, which was also indicated by the trapping experiments, only $K_{ic}$ values were determined for the remaining derivatives.

Comparison of competitive inhibition constants ($K_{ic}$) for *lin*-benzoguanine **L1** (4.1 µM) with respect to 6-aminoqinazolinone **Q1** (1.5 µM) shows a slight decrease of inhibitory potency with respect to the former. Substitution of **L1** in position 4 as realized in **L2**, results in a slightly improved affinity (Tab. 3.8). With a $K_{ic}$ of 1.0 µM **L2** is the most potent TGT inhibitor of this series and slightly stronger than the equally substituted 6-amino-quinazolinone **Q7** (2.6 µM). Compound **L5** lacks the C(6)-$NH_2$ group compared to ligand **L2.** Its three-fold loss in affinity demonstrates the beneficial contribution of this function for binding as it mediates polar contacts to Asp156 and Asp102. Additional substitutions of the phenyl ring in **L2** by *p*-Me (**L3**, 6.9 µM) or *p*-OMe (**L4**, 3.7 µM) appears detrimental to binding. Immediate attachment of the phenyl group realized in **L6** results in a substantial loss in activity (29 µM). Similarly, the N(3)-substituted *lin*-benzoguanines experience a affinity loss (**L7**, 15.4 µM; **L8**, 7.3 µM). Structural evidence for these trends will be discussed in

the following chapters. Details for synthesis and inhibition constants recently have been published by us in Meyer *et al*. (2006)[65].

**Tab. 3.8** Competitive inhibition constants of *lin*-benzoguanine based inhibitors

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **L1** | **4.1 ± 1.0**<br>$K_{iu}$: 7.9 ± 2.0 |  | **L2** | **1.0 ± 0.3**<br>$K_{iu}$: ~ 50 |  | **L3** | **6.9 ± 1.7** |
|  | **L4** | **3.7 ± 0.9** |  | | | | |
| **Q1** | **1.5 ± 0.4**<br>$K_{iu}$: 0.6 ± 0.2 |  | **Q7** | **2.6 ± 0.9** |  | **L5** | **2.8 ± 1.3** |
|  | **L6** | **29 ± 7** |  | | | | |
| **L7** | **15 ± 4** |  | **L8** | **7.3 ± 3.8** |  | | |

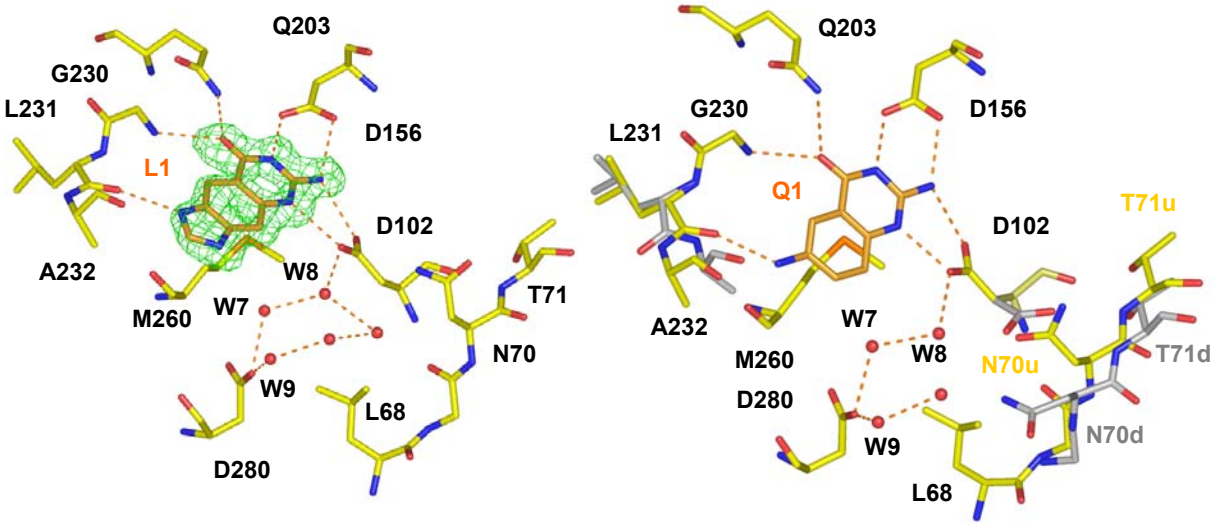
### 3.3.2.2 Crystallization experiments

Crystallization experiments have been performed with all new inhibitors as binary complexes with *Z. mobilis* TGT. For details see chapter 5.5. The improved solubility of the *lin*-benzoguanine derivatives with respect to the previously investigated 6-amino-quinazolinone enhanced the chances for successful soaking. For **L1**, **L2**, **L3** and **L4** soaking at a pH of 5.5 resulted in crystals diffracting to a maximum resolutions of 1.58 – 2.1Å. Details of the data collection and refinement statistics are given in Table 6.2.3. In case of **L5** and **L6** no bound ligand could be identified in the

active site after soaking. In case of **L7** and **L8** the substitution at N(3) reduced solubility. For **L7**, no bound ligand could be detected after soaking, for **L8** the binding pocket was partly occupied by the ligand. Nevertheless, most likely due to reduced population and/or structural disorder, the diffraction data was not sufficient for structural refinement.

### 3.3.2.3 Crystal structure of *lin*-benzoguanine

The assumed binding mode of *lin*-benzoguanines could be confirmed by crystal structure analysis of **L1** in complex with TGT at a resolution of 1.7Å. The structure was solved and refined by Ruth Brenk (PhD Thesis, 2003)[52]. **L1** is well defined in the $G_{34}$ binding pocket and in tight contact with the protein forming 7 H-bonds and 60 van der Waals interactions. 94 % of its solvent accessible surface is buried upon complexation (Fig. 3.10a and Tab. 3.9).




a b

**Fig. 3.10** a) Crystal structure of **L1** in the binding pocket of TGT determined at 1.7Å resolution and contoured at 2.4σ in the |Fo| - |Fc| density map refined excluding ligand coordinates in the last refinement cycle; b) binding mode of **Q1** in the binding pocket of TGT; coloured in grey: apo conformation, coloured in yellow: inhibitor bound conformation.

The observed binding mode is similar to that of the natural substrate $preQ_1$ or the initial lead **Q1** (Fig. 3.10b). The residual electron density of **L1** is better defined compared to **Q1**. Both, **L1** and the residues forming the $G_{34}$ binding pocket are properly ordered whereas for **Q1** several split side chain conformations were observed [Brenk *et al.*, 2004][1]. It had been concluded that **Q1** occupied not all

binding pockets in the crystal upon soaking due to reduced ligand solubility. Thus, the diffraction pattern, recorded as average for the entire crystal, comprises a superposition of the ligand-bound and the ligand-free binding pockets. The former was refined to 60 % occupancy. It shows, superimposed with the unoccupied pocket, for Asp102, Asn70 (Gly69-His73), Leu231 and Ala232 two alternative conformations. This gives rise to the interpretation that upon ligand binding Asp102 has to rotate towards the ligand and Asn70 moves concertedly together with the Val45-His73 backbone strand by 2Å. Furthermore, the Leu231/Ala232 peptide bond is flipped to expose its carbonyl oxygen to H-bond the 6-amino group of **Q1**. In TGT·**L1** the ligand is fully occupied. Binding pocket residues adopt only one conformation similar to TGT·**Q1** in the ligand-bound state. Asp102 recognizes the ligand via a double hydrogen bond. The movement of Asn70 significantly extends the volume of the adjacent hydrophobic cavity formed by Val45 and Leu68. A similar cluster of crystal waters, particularly W7, W8 and W9, stabilize Asp102 and Asp280 in both structures (Fig. 3.10a/b).

**Tab. 3.9** $K_{ic}$ values for **L1-4**, average B-values for scaffold and substituent, number of formed H-bonds, van der Waals contacts in the binding pocket and percentage of buried solvent accessible surface for the ligands.

| **Compound** | **L1** | **L2 u** [a] | **L2 d** [a] | **L3 u** | **L4 d (TGT$_1$)** [b] | **L4 d (TGT$_2$)** [b] |
|---|---|---|---|---|---|---|
| $K_{ic}$ [µM] | 4.1 | 1.0 | | 6.9 | 3.7 | |
| av. B-value scaffold | 20.2 | 15.2 | 15.0 | 21.3 | 39.0 | 29.3 |
| av. B-value substit. | – | 27.8 | 23.1 | 45.6 | 54.0 | 32.9 |
| Nr. of H-bonds | 7 | 7 | 7 | 7 | 7 | 7 |
| Nr. of v. d. Waals contacts | 60 | 96 | 86 | 92 | 88 | 102 |
| buried solvent accessible surface [%] | 94 | 93 | 93 | 90 | 95 | 93 |

a: **L2** was separately calculated for u- and d-conformation

b: **L4** refers to the two symmetry independent dimers in TGT$_1$·**L4** and TGT$_2$·**L4**

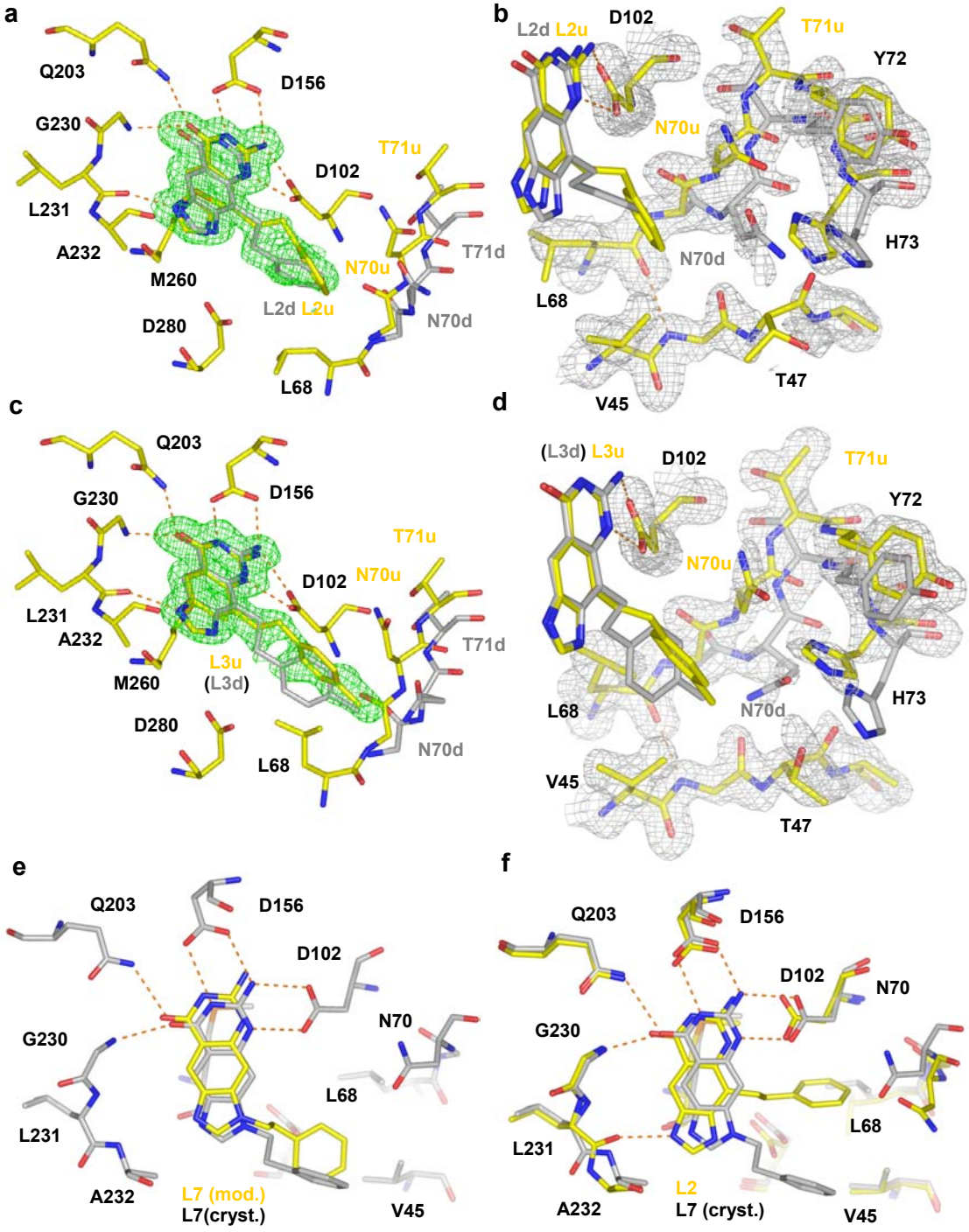
#### 3.3.2.4 Substituted *lin*-benzoguanine crystal structures

In consequence, further design concentrated on filling the enlarged hydrophobic cavity, simultaneously expelling the water cluster. Crystal structure analyses of TGT·**L2** and TGT·**L3** showed that this goal has been successfully achieved and the aromatic substituents occupy the cavity (Fig. 3.11a/c). In both structures, the basic *lin*-benzoguanine scaffold is located equally as in TGT·**L1**. Surprisingly, no significant gain in affinity was observed for the substituted compounds **L2** - **L4** (Tab. 3.8). Obviously, the expected entropic gain does not pay entirely for the enthalpic price of the replacement of the water cluster next to Asp280 and Asp102 (Fig. 3.10a), as already described in detail for the 7-amino-quinazolinones in chapter 3.3.2. The remaining cost virtually compensates for the additional contacts formed by the introduced ligand side chain (60 van der Waals interactions in **L1** vs. 86 – 96 in **L2** and **L3** – see Tab. 3.9). All substituted derivatives cross this region of the binding pocket with hydrophobic, hardly polarizable portions unable to form any polar interactions.

For **L7** and **L8** the chemical structure provides a further explanation for the slight affinity loss compared to **L1** (Tab. 3.8). In **L1** – **L4** N(1) is protonated and the Leu231/Ala232 peptide switch is present in the conformation allowing H-bonding *via* the Leu231 carbonyl group (Fig. 3.10a/c). Substitution at N(3) of the imidazole moiety of **L7** and **L8** would require N(1) to be present in deprotonated state, at least as long as this position of the ligand remains uncharged and N(1) does not pick-up a proton. In this unprotonated state, N(1) is not expected to induce a flip of the peptide switch and the amide nitrogen of Ala232 would remain exposed to the binding pocket. But in this conformation the methyl group of Ala232 limits space at the upper rim of the binding pocket for ligand accommodation (see TGT·**Q2** in Fig. 3.5). The *lin*-benzoguanine scaffold, however, requires extended volume in this region compared to **Q2**, thus proper hosting of **L7** and **L8** appears sterically unfavourable.

To obtain further evidence for these assumptions, **L8** was placed to the difference electron density of the partly occupied binding pocket of TGT·**L8** and minimized in this orientation using MOLOC to test whether the assumed orientation could be correct. Both scaffold orientations are virtually identical (Fig. 3.10e). The superposition of TGT·**L8** with the binding mode of TGT·**L2** supports the assumptions made above. As a result of the Leu231/Ala232 peptide switch present in the NH-

exposing orientation, the **L8** scaffold is shifted and rotated inside the binding pocket compared to **L2** (Fig. 3.10f).




Legend on the following page

**Fig. 3.11** a) Crystal structure of TGT·**L2** in the binding pocket of TGT contoured at 2.4σ in the |Fo| - |Fc| density map. The phenyl substituent adopts two conformations (grey=d, yellow=u) paralleled by two conformations for N70; b) electron density of the TGT·**L2** N70 loop contoured at 1.0σ in the 2 |Fo| - |Fc| density map. Associated with the split conformations of **L2** the residues G69-H73 adopt two arrangements; c) crystal structure of TGT·**L3** contoured at 2.4σ in the |Fo| - |Fc| density map. The phenyl substituent was only refined in up-conformation (yellow) although the binding pocket conformation also suggests presence of the down-conformation (grey); d) electron density of the TGT·**L3** N70 loop contoured at 1.0σ in the 2|Fo| - |Fc| density map. Residues G69–H73 adopt two almost equally distributed conformations; e) structural superposition of the assumed binding mode of TGT·**L8** indicated by the partly occupied binding pocket (grey) and minimization with MOLOC (yellow); f) structural superposition of the assumed binding mode of TGT·**L8** and the crystal structure of TGT·**L2**

#### 3.3.2.5 Split conformations in TGT·L2 and TGT·L3

Analysis of TGT·**L2** and TGT·**L3** reveals an unexpected conformational splitting for both, ligand and active site residues. Both structures were resolved to a resolution of 1.58 Å, thus allowing the interpretation of partial occupancy. In Figure 3.11a/c the ligand coordinates were excluded from the last refinement to test the quality of the model. For reasons of visualization and comparability identical σ levels were applied for respective figures.

With respect to a best plane defined by the *lin*-benzoguanine skeleton, the 4-substituents orient either above (up-conformation = u) or below this plane (down-conformation = d). Both orientations are observed side-by-side in the same crystal suggesting virtually equivalent energy content. Both orientations have already been reported for substituted quinazolinone-type inhibitors (e.g. **Q15** – Fig. 3.6) [Brenk *et al*., 2004][1]. However, in these cases structural disorder was too pronounced and the difference electron densities of ligand substituents and active site residue side chains were not well enough defined to unambiguously relate these split conformations to specific structural adaptations of the binding pocket.

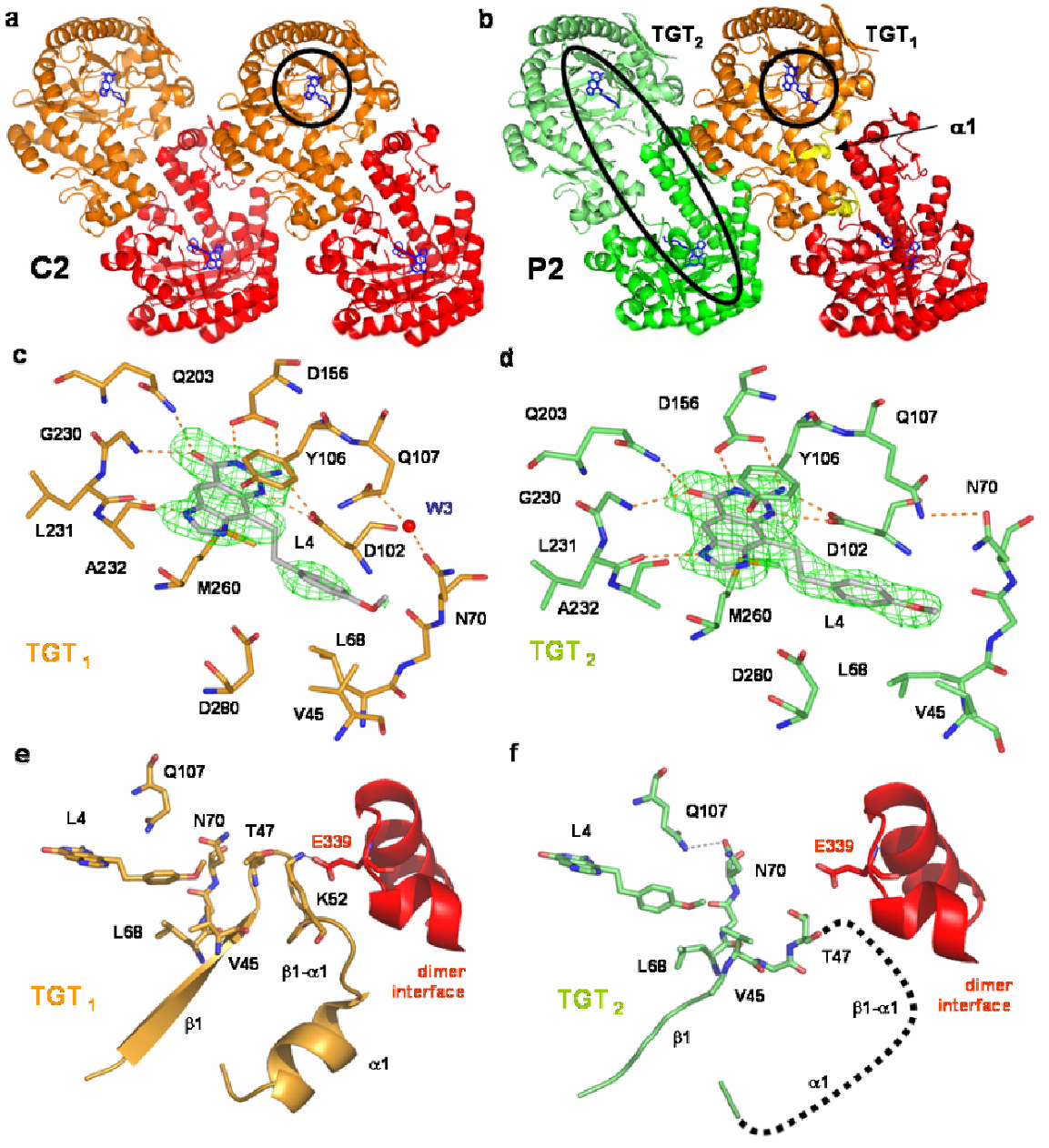
In TGT·**L2** the d-conformation is slightly higher populated (occupancy of 56 %) and a low average B-value of 23.1 Å$^2$ is found for the phenyl substituent. The u-conformation was refined to an occupancy of 44 % with a slightly increased average B-value of 27.8 Å$^2$ (Fig. 3.11a and Tab. 3.9). While in both conformations the

buried solvent accessible surface (93 %) and the number of H-bonds formed with active site residues (7) is identical, the amount of established van der Waals contacts in u-conformation is lower (86 vs. 96). This may provide an explanation for the lower occupancy of the u-conformation and possibly indicates a slight energetic disadvantage. The two observed conformations of the substituents correlate with two different binding site conformers involving the Asn70 backbone stretch from Gly69 to His73. They differ by 1.4 Å with respect to averaged $C_\alpha$ rmsd. In Figure 3.11b the shift of the Gly69 – His73 backbone is visualized with the most pronounced shift of 1.9 Å experienced by Thr71. For the d-conformation, Asn70 (N70d) is found in virtually the same position as in uncomplexed TGT. In the u-conformation Asn70 (N70u) adopts a position similar to that in TGT·**L1** and its carboxyamine group performs a parallel stacking with the phenyl ring of the substituent. Accordingly, both conformations are largely identical to either the ligand-bound or ligand-free conformations observed in TGT·**Q1** (Fig. 3.10b). But in contrast to this structure, in TGT·**L3** both protein conformers are accommodating the ligand.

In the crystal structure of TGT·**L3**, u- and d-conformation split differently. Electron density suggests that the aromatic substituent favours u-conformation, however with reduced accuracy, as indicated by the rather diffuse side chain electron density (Fig. 3.11c). With an average B-value of 45.6 $Å^2$ the substituent exhibits a significantly higher value compared to the remaining scaffold atoms (21.3 $Å^2$). The ligand is to 90 % buried from solvent accessibility, forming 7 H-bonds and 92 van der Waals contacts (Tab.3.9). Surprisingly Asn70 is found almost equally distributed in u- and d-conformation (Fig. 3.11d). Consulting TGT·**L1** and TGT·**L2** it can be concluded that the Asn70d conformer is only observed once the *lin*-benzoguanine scaffold is substituted. Therefore, we assume that also for **L3**, the d-conformation of the side chain is present, however distributed over multiple states, thus resulting in a rather blurred and inaccurately defined electron density. In consequence, this conformation was not included in the refinement. Nevertheless, we added its possible geometry in Figure 3.11c/d to show its putative contacts to Asn70d. Supposedly unfavourable contacts experienced by the additional para-methyl group in the d-conformation with Asn70d result in the less-well defined binding mode. Directly facing the orientation of Asn70d in TGT·**L2** and TGT·**L3** shows that this residue is further pushed out of space as a consequence of accommodating the extended para-methyl-phenyl substituent in **L3** (Fig. 3.11a/c).

### 3.3.2.6 Destabilization of crystal contacts in TGT·L4

A different crystal structure is observed for TGT·**L4**. This results from a space group transition observed upon soaking. TGT crystallizes in space group *C*2 with one TGT molecule per asymmetric unit and usually this crystal symmetry is conserved upon soaking. In crystals with this space group the TGT molecule is supplemented by a symmetry equivalent in the crystal packing to form a homodimer with a buried solvent accessible surface of 1,667Å$^2$, as already described by Romier *et al*. (1996)[25]. Two dimers finally compose the *C*2 unit cell symmetry related by *C* centering (Fig. 3.12a). The putative relevance of dimer formation for catalysis has already been described in chapter 2.3.




Legend on the following page

**Fig. 3.12** a) TGT·**L2** in space group C2 with one TGT molecule per asymmetric unit. Two TGT molecules form a symmetric dimer via a two-fold axis. In the unit cell two identical dimers are present symmetry related by *C* centering; b) In TGT·**L4** *C* centering degenerates upon ligand soaking, reducing crystal symmetry to the space group *P*2. In consequence, two independent TGT dimers are found in the unit cell. In $TGT_2$ helix $\alpha$1 and loop $\beta$1-$\alpha$1, involved in dimer interface formation, are disordered. In $TGT_1$ helix $\alpha$1 and the dimer interface are ordered, similarly to the dimers in *C*2; c) The $TGT_1$·**L4** binding pocket is structurally similar to *C*2 complexes. **L4** is contoured at 2.4$\sigma$ in the |Fo| - |Fc| density map. W3 is found in an equal position to N70 as in TGT·**L1**; d) The $TGT_2$·**L4** binding pocket (dimer with modified interface). **L4** is contoured at 2.4$\sigma$ in the |Fo| - |Fc| density map. Due to the disordered helix $\alpha$1 V45-T47, L68-N70 and Q107 adopt conformations that significantly extend the binding pocket size; e) Binding pocket and dimer interface of $TGT_1$·**L4**. Helix $\alpha$1 and loop $\beta$1$\alpha$1 are fully ordered; f) Binding pocket and dimer interface of $TGT_2$·**L4**. Helix $\alpha$1 and loop $\beta$1$\alpha$1 are disordered and in consequence, **L4** penetrates deeper into the pocket and becomes properly ordered.

Surprisingly, TGT·**L4** displays two molecules, $TGT_1$ and $TGT_2$, per asymmetric unit in space group *P*2 (Fig. 3.12b). Binding of **L4** interferes with dimer formation at least in one of the two dimers present in the *C*2 unit cell, thus reducing crystal symmetry by loss of the *C* centering. The cell dimensions in *P*2 remain virtually identical as in *C*2 (cell parameters are given in Tab. 6.2.3). This provides an explanation for the surprising fact that the crystal did not crack upon soaking. In $TGT_2$·**L4** ligand binding extensively modifies the binding pocket as well as the dimer interface. Interestingly enough in $TGT_1$·**L4** neither the active site nor the interface are affected by ligand binding. Instead, they show a very similar geometry to all previously determined structures (e.g. TGT·**L2** and TGT·**L3**) and will be discussed with respect to these, first. No split conformations are observed in $TGT_1$·**L4** as detected for **L2** and **L3**. Only the d-conformation is present (Fig. 3.12c). The electron density of the phenyl-methoxy substituent is defined with only limited accuracy and the average B-value of 54.0 $Å^2$ for the substituent is significantly higher than that of the remaining scaffold (39.0 $Å^2$). Nevertheless, the ligand is buried to 95 % from solvent access and forms 7 H-bonds and 88 van der Waals contacts with the protein (Tab. 3.9). Also Asn70 adopts only the d-conformation and forms close contacts with the methoxy group of the substituent. Obviously this further size-expanded substituent (cf. **L2** and **L3**) avoids accommodation in u-conformation. Thus, the substituent of **L4** inverts the shift

of Asn70 towards u-conformation caused by the unsubstituted *lin*-benzoguanine scaffold in TGT·**L1** (Fig. 3.10a). The position occupied by the carboxamine group of Asn70 in u-conformation in other structures is occupied in the present case by a water molecule (W3) mediating a contact towards Gln107 (Fig. 3.12c). The side chain of this residue is fully ordered in the present structure, unlike in TGT·**L2** and **L3** where it adopts no uniquely defined conformation.

The $TGT_2$ dimer, however, exhibits one major structural distortion. Half of the loop $\beta$1-$\alpha$1 and the whole helix $\alpha$1 are not defined in the electron density (Fig. 3.12d/f). In $TGT_1$ and other *C*2 crystallized structures Ala49, Thr50 and Lys52 residing in the $\beta$1-$\alpha$1 loop form interactions with Tyr330, His333, Leu334, Glu339 and Leu341 from the zinc binding-subdomain of the tightly bound dimer molecule. Especially notable is the charged interaction between Lys52 and Glu339 (Fig. 3.12e). Although the residues Ala49, Thr50 and Lys52 within loop $\beta$1-$\alpha$1 are not ordered in $TGT_2$, further interactions within the dimer interface maintain the structural integrity of the $TGT_2$-dimer arrangement (see chapter 2.3). Adjacent to Ala49-Lys52 the residues Val45-Thr47 in loop $\beta$1-$\alpha$1 are located at the outer rim of the active site, being part of the above discussed small hydrophobic ribose 34 binding site. In $TGT_1$ the substituent of **L4** is distributed over several states in the binding pocket, with high average B-values (scaffold: 39.0 $Å^2$, side chain 54.0 $Å^2$), whereas in $TGT_2$ the ligand is well defined within the binding pocket (scaffold: 29.3 $Å^2$, side chain 32.9 $Å^2$ – see Tab. 3.9). Figures 3.12c/d illustrate this difference in binding accuracy. In all symmetric *C*2-dimers the loop residues Val45–Thr47 stabilize the geometry of Leu68–Asn70 and obviously restrict the spatial extension of the binding pocket (Fig. 3.12e). In $TGT_2$ Val45–Thr47, the last ordered residues of loop $\beta$1-$\alpha$1, point away from the binding pocket thus extending its size. This allows Leu68-Asn70 to shift away from the ligand. The Leu68 side chain rotates away from the ligand and H-bonding between the Asn70 and Gln107 side chain further stabilizes and extends the size of the binding pocket (Fig. 3.12f). As a result **L4** penetrates deeper into the binding pocket of $TGT_2$ than in $TGT_1$ and adopts, most likely, a more favourable geometry than in the latter. This is also suggested by a increase of van der Waals contacts from 88 to 102 (Tab. 3.9).

From these observations it can be deduced that the binding of **L4** interferes in one of the symmetry independent molecules with the geometry of the TGT-dimer interface. As both contact geometries are present side-by side in the same crystal it is possible

to directly compare both arrangements involving Val45-Lys52 in the β1-α1 loop. In $TGT_2$ these residues adopt a geometry that allow a properly ordered geometry of **L4** in the binding pocket. Nevertheless, the extended binding pocket geometry takes influence on the TGT-dimer interface resulting in a structural disorder involving Ala48 to the end of helix α1. In $TGT_1$ this helix remains ordered and the dimer interface is found as in all previously determined structures, but **L4** is not properly accommodated in the obviously too narrow binding pocket.

To estimate whether both, dimer formation and destabilization, is a crystallographic artifact or whether it is of functional relevance that might be used in structure-based design, requires further investigations as already suggested in chapter 2.3. If homodimer formation turns out to be a prerequisite for catalysis in TGT, **L4** provides an outlook to a new class of inhibitors, that interferes with TGT dimer formation.

#### 3.3.2.7 Induced fit adaptations and water molecules

The induced-fit adaptations observed in TGT·**L1** - **L3** and $TGT_1$·**L4** can similarly be observed when TGT is bound to tRNA. Therefore, conformers of TGT observed upon inhibitor binding will be compared to the adaptations registered in due course of substrate binding. The structures of uncomplexed TGT, [Romier *et* al.,1996][25], and the one bound to a tRNA anticodon stem loop, [Xie et al., 2003)[4], reveal a series of adaptations in the vicinity of the $G_{34}$/$preQ_1$ binding pocket (Fig. 3.13a). Asp102 involved in H-bonding to Asn70 and Thr47 in the uncomplexed situation, rotates towards $preQ_1$. Simultaneously, Asn70 and Thr47 engage in the stabilisation of the U35 phosphate group. During this transition Thr47 shifts by about 1.8 Å and the backbone strand from Gly69 to His73 is translocated by 1.0 Å ($C_\alpha$-rmsd 69 - 73). Additionally, the Thr47/Ala48 peptide bond is flipped. The carbonyl group of Thr47 is now binding to the main chain amide nitrogen of Asn70, thus stabilizing the shifted geometry of the Gly69-His73 backbone stretch.

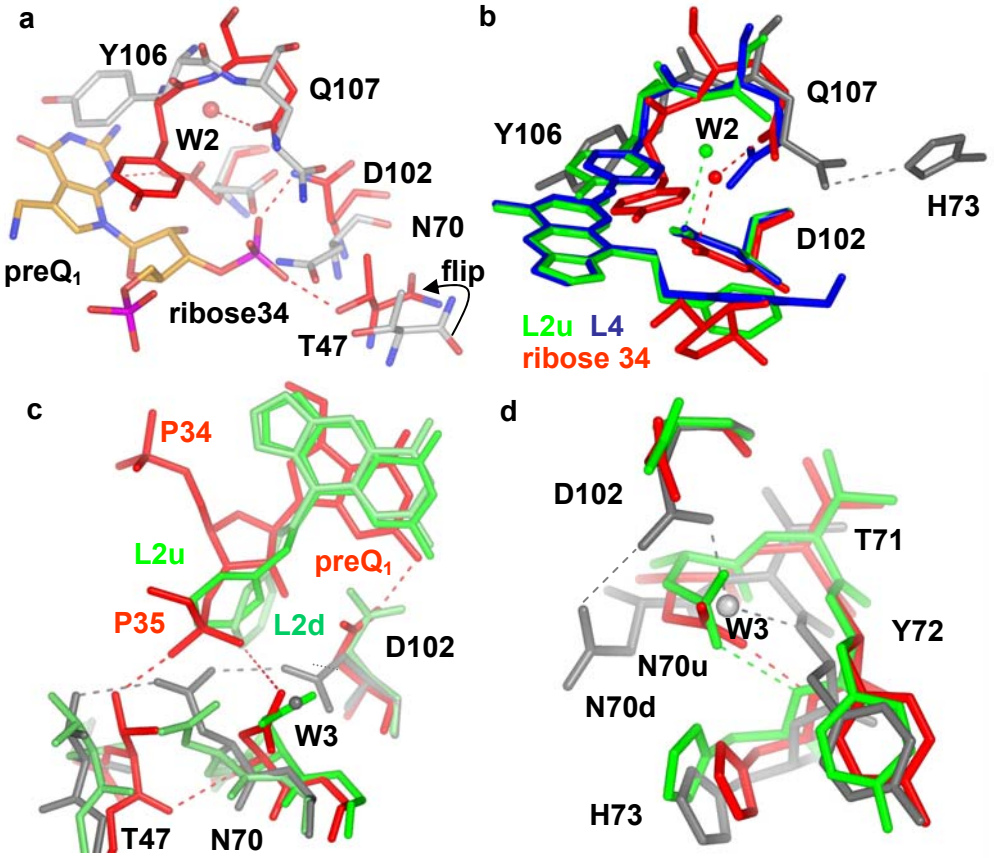



**Fig. 3.13** a) Structural superposition of uncomplexed TGT (grey) and preQ$_1$-tRNA-bound TGT (orange/red) reveals multiple conformational adaptations; b) structural superposition of uncomplexed TGT (grey), preQ$_1$-bound tRNA (red), TGT·**L2** in d-conformation (green) and TGT$_1$·**L4** (blue). **L2** and **L4** do not allow the shift of Q107 as in the tRNA bound state, due to stacking of the side chain to *lin*-benzoguanine. The Q107 side chain is either disordered (green) or W2 is not found in the binding pocket (blue); c) structural superposition of uncomplexed TGT (grey), preQ$_1$-tRNA bound TGT (red) and TGT·**L2** in u- and d-conformation (green). Upon tRNA binding the H-bond network of D102-N70-T47 (grey) is disrupted. All residues are involved in substrate binding. In TGT·**L2** N70u is similar to the tRNA-bound, N70d to the uncomplexed and T47 in an intermediate conformation; d) structural superposition of uncomplexed TGT (grey), preQ$_1$-tRNA bound TGT (red) and TGT·**L2** in u-conformation (green). W3 is stabilized by the amide group of Y72. In the tRNA bound state and in TGT·**L2** N70 replaces W3 forming an H-bonding to the amide group of H73.

Tyr106, closing up the $G_{34}$/$preQ_1$ recognition pocket in the uncomplexed state performs a hydrophobic stacking with the ribose ring adjacent to $preQ_1$ once tRNA is bound (Fig. 3.13a). In this state, the Tyr106 backbone is shifted by 1.9 Å compared to uncomplexed TGT. In consequence the neighbouring Gln107 side chain is no longer hydrogen-bonded to His73 which is located at the outer rim of the binding pocket. Instead Gln107 orients towards the $G_{34}$/$preQ_1$ recognition site to subsequently involve Asp102 in a water (W2) mediated interaction. Presumably, this water molecule is important in the catalytic mechanism assisting Asp102 and Gln107 to shuffle a proton out of the binding pocket (Fig. 2.4).

Inhibitor binding of TGT·**L2**, TGT·**L3** and $TGT_1$·**L4** partially reflect these structural rearrangements, but obviously they do not proceed beyond these substrate-induced adaptations. The discussion will focus on Asn70/Thr47 and Tyr106/Gln107 as they monitor the underlying processes. The adaptations close to Asn70/Thr47 are associated with the rotation of Asp102 towards the base recognition site in a domino-type rearrangement. All ligands interact with Asp102 similarly to $preQ_1$ in its tRNA-bound state (Fig. 3.13c). Asn70, after having lost Asp102 as binding partner, seeks for a new stable geometry. In the crystal structure of TGT·**L2**, the two almost equally populated conformations of Gly69 - His73 correspond closely ($C_\alpha$-rmsd Gly69 - His73: $\leq$ 0.5Å) to the conformers found in either uncomplexed or tRNA-bound TGT which mutually differ by $\geq$ 1Å (Fig. 3.13c). In uncomplexed TGT and $TGT_1$·**L4** a water molecule (W3) is found in exactly the same position that the Asn70 side chain occupies in the tRNA-bound state (Fig. 3.20c/d). In uncomplexed TGT this water molecule is stabilized by contacts to the backbone amide of Tyr72 and the side chain of Asp102. In $TGT_1$·**L4** the latter contact is replaced by a contact to Gln107 as Asp102 is rotated towards the ligand. In the tRNA-bound TGT the shifted Asn70 displaces this water and forms a backbone amide interaction to His73 (Fig. 3.13d). This interaction allows the Asn70 side chain to properly accommodate the binding of the U35 phosphate group that serves as one of the hinges controlling the rotation of ribose 34 (Fig. 2.3e). A similar backbone amide interaction to His73 is realized by Asn70u in TGT·**L2**, although here the carboxamide terminus is rotated by about 45° with respect to the tRNA-bound situation as the side chain stacks on top of the ligand's phenyl ring (Fig. 3.13d). These findings underline the relevance of W3 in

stabilizing the uncomplexed binding pocket conformation and allowing by its substitution the modification of the binding pocket geometry.

Also Thr47 has to reorient after losing close contact to the Asn70 side chain upon tRNA binding (Fig. 3.13a). In the crystal structure of TGT·**L2**, Thr47 is lacking a proper interaction partner (Fig. 3.13c). The residue remains rather mobile indicated by a poorly defined electron density and an augmented average B-value of 57.3 $Å^2$. The backbone flip is observed similar as in the tRNA-bound state but no backbone shift can be observed. Thus, no H-bond towards the backbone of Asn70 is formed. Obviously this allows Asn70 to be present in two alternative conformations. However, within the domino-type rearrangement process the presence of a rather disordered residue points to the relevance of interaction partners to stabilize alternative geometries.

The conformational changes observed for Tyr106/Gln107 in the inhibitor-bound states are less pronounced than those observed upon tRNA binding. This results from the fact that Tyr106 performs a hydrophobic stacking onto the *lin*-benzoguanine scaffold in inhibitor-bound states whereas it is stacking onto ribose 34 in the tRNA-bound state (Fig. 3.13b). The latter requires a pronounced backbone shift with respect to the uncomplexed conformation (1.6 Å compared to uncomplexed TGT). The smaller shift of Tyr106 in inhibitor bound crystal structures (0.9 Å compared to uncomplexed TGT) results in deviating conformations of Gln107 that impede the water mediated contact by W2 towards Asp102 observed in the tRNA-bound state. In TGT·**L2**, the Gln107 side chain is not visible in the electron density but water molecule W2 can still be observed H-bonded to Asp102 (Fig. 3.13b – green). In $TGT_1$·**L4**, the Gln107 side chain points directly towards Asp102 and no mediating water molecule can be detected (Fig. 3.13b – blue). Instead, H-bonding to W3 can be observed (Fig. 3.12c). Thus, the rotation and shift of Tyr106 out of the base recognition site triggers the rearrangement of Gln107 along with an appropriate orientation of the catalytically important water molecule W2.

### 3.3.2.8 Discussion and outlook

In the presented crystal structures pronounced induced-fit adaptations could be observed resembling those experienced by TGT upon tRNA binding. In this context, the question arose whether the adaptations result as specific responses to the properties of each individual inhibitor or whether they point to structural adaptations potentially required by the protein in order to accomplish its functional task? If the latter assumption is correct, it can be assumed that these adaptations correspond to low energy transformations required for efficient functioning of the enzyme. Most of the studied inhibitors place their substituent in two alternative conformations into the hydrophobic ribose 34 binding pocket. With respect to the strand hosting Asn70 the protein responds differently to both placements: One geometry relates closely to the uncomplexed, the other to the tRNA-bound situation. Considering Tyr106/Gln107 and Thr47, these residues adopt conformations half-way in-between the uncomplexed and tRNA-bound situation. From this finding we can conclude for inhibitor binding: TGT is highly adaptive, however this multiplicity is obviously required to fulfill its functional role. A ligand to be bound selects one of the binding competent conformers of the protein and stabilizes this state in the complex. Very similar considerations have been suggested to rationalize the binding properties of different inhibitors to aldose reductase, [Urzhumzev *et al.,*1997; Howard *et al*., 2004; Sotriffer *et al*., 2004][86-88], or trypsin mutants [Rauh *et al*., 2002; Rauh *et al*., 2003][89, 90]. Nevertheless, the picture might be more complicated as indicated by inhibitor **L4**. The destabilization of a whole structural element might point to adaptations beyond the naturally observed flexibility of the enzyme.

A further conformational adaptation within the TGT binding pocket leads to another important aspect of ligand binding and catalytic activity: water molecules tightly associated with the binding pocket. The Leu231/Ala232 peptide flip controls substrate recognition and promiscuity [Stengl *et al*., 2005][54]. Depending on the orientation of the peptide switch either the carbonyl- or the amide-group is exposed to the binding pocket. In two uncomplexed TGT crystal structures exhibiting both orientations, due to deviating pH conditions used for crystallization, these functional groups are H-bonded to two distinct water molecules in deviating positions. Upon ligand binding, they either get replaced or mediate an interaction between the bound ligand and the exposed peptide bond functionality (W1 – Fig. 2.9a/b). The present study indicates further water molecules as important for the substrate or ligand

binding process. Water molecule W3 serves as a surrogate for an alternative protein side chain conformer that allows tRNA binding in the correct orientation (Fig. 3.13d). Water molecule W2 is only found in the active site when the substrate tRNA is bound, as its accommodation requires a conformational adaptation of the binding pocket (Fig. 3.13b). As W2 is likely to receive a proton from $preQ_1$ during catalysis its presence is highly relevant for the enzymatic reaction (Fig. 2.4). In the ligand bound crystal structure of TGT·**L1** two waters, W7 and W8 are bridging between Asp102 and Asp280, the catalytic nucleophile. Attempts to replace them by hydrophobic ligand portions were geometrically successful, however, they did not parallel the expected increase in binding affinity. Despite the assumed entropic benefit for water release obviously the price for their replacement by a hydrophobic substituent appears detrimental to binding. Possibly, the waters compensate for the negative charge of Asp280 and should be replaced by polarizable ligand functionalities that are able to form polar interactions.

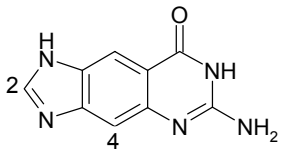
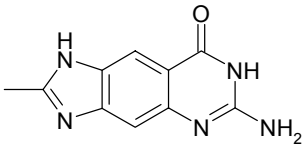
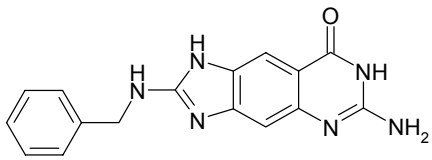
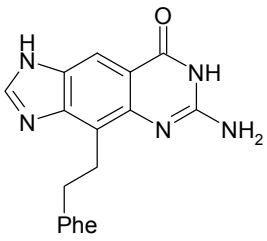
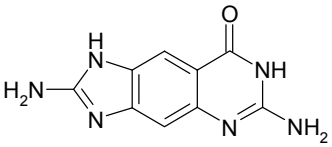
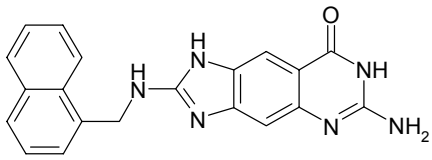
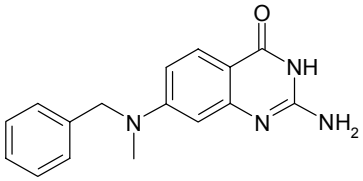
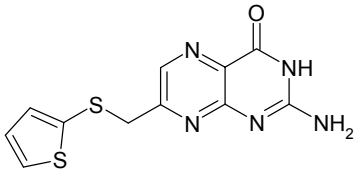
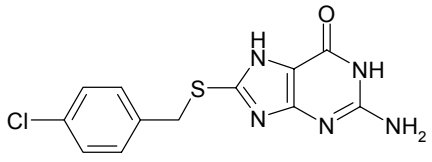
In summary, these findings underline the structural and functional relevance of water molecules associated with the structural adaptability of the binding pocket. The knowledge from this inhibitor series should be considered in further design approaches, particularly in new attempts trying to address the hydrophobic ribose 34 binding pocket.

### 3.3.3 2-Amino-*lin*-benzoguanines

Substitution of *lin*-benzoguanine **L1** in position 2 resulted in the discovery of 2-amino-*lin*-benzoguanine **L10** as a potent new lead structure. The initially synthesized sample of **L1** showed impurities due to a small amount of unexpected synthesis byproduct. Surprisingly, this impurity was detected by crystal structure analysis. Soaking of this compound sample into TGT crystals resulted in a difference electron density with additional density in position 2 with respect to the *lin*-benzoguanine scaffold (Tab. 3.14d). Thus, a new sample of better purified **L1** was synthesized. Soaking resulted in a crystal structure with a difference electron density corresponding to the chemical structure of **L1** (Fig. 3.10a). Nevertheless, with a $K_i$ of 150 nM the inhibition constant for the previous impure sample was significantly lower than that of pure **L1** ( 250 nM – values referring to the former assay setup) [Brenk, PhD Thesis, 2003][52]. Unfortunately, the chemical structure of the byproduct could not

be identified from the sample, as it was present in a too small amount. However, from the observations it could be concluded that the 2-substituted byproduct of *lin*-benzoguanine displayed a potent TGT inhibitor. Although present only in traces, it was capable of replacing **L4** in the TGT crystals and resulted in significantly stronger inhibition of the base exchange reaction.

**Tab. 3.10** Competitive inhibition constants of 2-amino-*lin*-benzoguanines in comparison with other relevant compounds

| | | | | | |
|---|---|---|---|---|---|
| **L1** | **4,100 ± 1,000**<br>impure: 150* / pure: 250* |  | **L9** | **300 ± 75** |  |
| **L11** | **100 ± 25** |  | | | |
| **L2** | **1,000 ± 300** |  | **L10** | **125 ± 50** |  |
| **L12** | **75 ± 15** |  | | | |
| **Q21** | **7,600 ± 3,700** |  | **P5** | 450± 50* |  |
| **S1** | **51,000 ± 13,000** |  | | | |

$K_{iu}$ in [nM]; (*): previous assay

To test 2-substituted *lin*-benzoguanine compounds **L9** and **L10** have been synthesized. In further steps **L10** was extended by a phenyl (**L11**) and a naphtyl (**L12**) moiety. All synthesis were performed by Simone Hörtner (group of Prof. F. Diederich, ETH Zürich). To identify the inhibition modes for the new inhibitor series, trapping experiments were performed. In contrast to **L1**, no stabilization of the covalent TGT-tRNA intermediate could be observed. Thus, 2-substituted *lin*-benzoguanines represent competitive inhibitors of tRNA, binding to uncomplexed

TGT. The competitive inhibition constant ($K_{ic}$) for all compounds of the new series are in the nanomolar range (Tab. 3.10). Compared to the scaffold **L1** (4,100 nM) the methyl-substituted **L9** (300 nM) showed an affinity enhancement of more than a factor of 10. For the amino-substituted **L10** (125 nM) the gain in affinity was even more pronounced. Further substitution of the amino group, intending to address the ribose 33 binding site, resulted only in minor affinity gain (**L11**: 100 nM; **L12**: 75 nM). The results of the trapping experiments and affinity data for **L9** and **L10** were kindly provided by Tina Ritschel (group of Prof. Klebe, University of Marburg).

Crystal structure analysis was successfully performed with **L12** as binary complex with *Z. mobilis* TGT (chapter 5.5). Soaking resulted in crystals diffracting up to a maximum resolution of 1.95 Å (for crystallographic data see Fig. 6.2.4). In TGT·**L12** only the *lin*-benzoguanine scaffold is well defined (Fig. 3.14a). The scaffold forms seven H-bonds with residues of the binding pocket and one long H-bonding interaction to the carbonyl-group of Ala232 *via* the endocyclic 2-amino group (distance: 3.6 Å). The scaffold is well positioned in the binding pocket and exhibits an average B-value of 40.6 $Å^2$. The average B-value of the protein is 28.2 $Å^2$. The naphtyl substituent, however, is scattered over multiple conformations, sticking into the direction of the solvent exposed ribose 33 binding pocket. This is indicated by the rather diffuse difference electron density and makes it difficult to locate the substituent. In the last refinement step it exhibited an average B-value of 86.9 $Å^2$ for the orientation shown in Figure 3.14a. Additionally, it was possible to locate crystal water molecules (W7 and W8) in the binding pocket (average B-value of 40.0 $Å^2$).

To find an explanation for the significant affinity gain, the crystal structure of TGT·**L12** was compared with other available TGT structures. The superposition with TGT·**L1** indicates similar orientations of the scaffolds (Fig. 3.14b). Also the binding pocket conformation is fairly identical. Similar to TGT·**L1** the Gly69-His73 loop is shifted with respect to uncomplexed TGT, thus extending the size of the ribose 34 binding pocket. In both structures the water molecules W7 and W8 are present mediating an H-bonding contact between Asp280 and Asp102. In the 4-substituted *lin*-benzoguanine inhibitor series the presence of these waters was identified to be highly relevant for binding affinity, as it might assist to compensate the charge on the nucleophile Asp280 (see chapter 3.3.2.4). However, in TGT·**L12** only W7 and W8 could be identified, most presumably resulting from the deviating orientation of the Asn70 side chain and the disordered naphtyl side chain of the ligand impeding proper

accommodation of W9 and further water molecules. Thus, the comparison of these two crystal structures does not provide an explicit explanation for the pronounced affinity gain for the 2-substituted *lin*-benzoguanines.

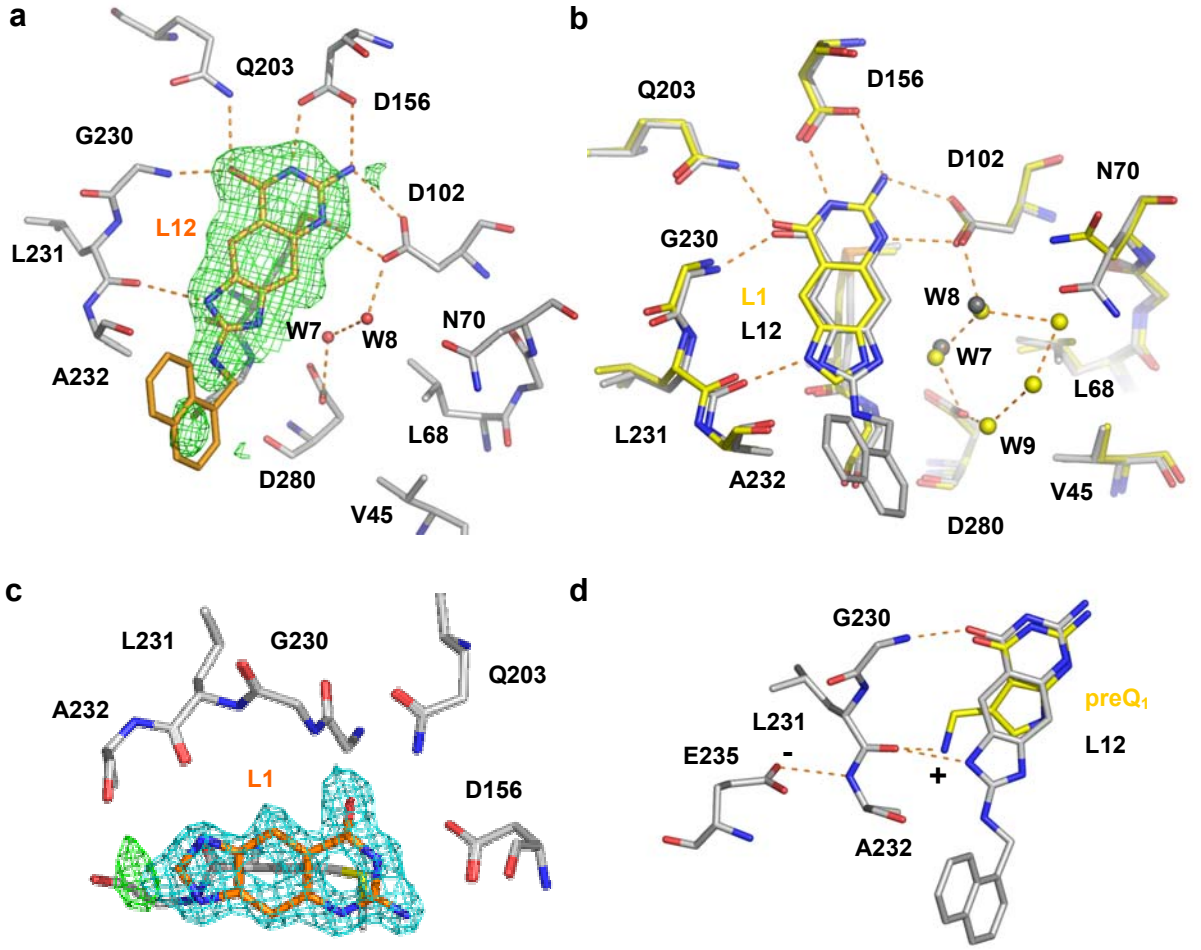



**Fig. 3.14** a) Crystal structure of TGT·**L12** with the ligand contoured at 2.0σ in the |Fo| - |Fc| difference electron density map refined in the last cycle excluding ligand coordinates to test the model; b) structural superposition of the active sites of TGT·**L1** and TGT·**L12**; c) crystal structure of TGT·**L1** (impure sample) with the ligand contoured at 1.0σ in the 2 |Fo| - |Fc| density map; additional difference electron density contoured at 2.5σ in the |Fo| - |Fc| difference electron density map indicates the presence of a small 2-substituted derivative of **L1** in the binding pocket; d) structural superposition of the Leu231/Ala232 peptide switch of TGT·preQ$_1$ and TGT·**L12** indicates charge assisted binding.

An explanation for the affinity gain might be given by considering charge-assisted hydrogen bonds in case of the 2-substituted *lin*-benzoguanines. This is suggested taking possible p$K_a$ shifts of relevant functional groups into account. For preQ$_1$ the positive charge on the exocyclic amino group supposedly assists binding (Fig. 3.14d). The carbonyl group of Leu231 is stabilized in its orientation by the negatively charged side chain of Glu235 (chapter 2.2.1 - 2.2.2). Thus, the Leu231/Ala232 peptide bond projects to some extend its negative charge into the binding pocket. The p*K*a value for an exocyclic amino-methyl group falls into the range of 10 (cf. p$K_a$ of 1-amino-1-phenyl-methane: 10.5; Fig. 3.15). Thus, preQ$_1$ is likely protonated under physiological pH and capable of compensating the negative charge of Glu235 mediated *via* the peptide bond. In general charge assistance is thought to contribute significantly to binding affinity [Böhm & Klebe, 1996][91]. In contrast, for **L1** no charge assisted H-bond is expected. The p$K_a$ of the imidazole moiety in *lin*-benzoguanine is 4.5 [Brenk, PhD Thesis, 2003][52]. The assay, however, is applied at a pH of 7.3. Thus, **L1** binds most likely deprotonated. Similar considerations hold for the anilin-type 6-amino group of **Q1** (p$K_a$ of 2-amino-naphthaline: 4.1; Fig. 3.15). Substitution of the *lin*-benzoguanine scaffold in position 2 is expected to significantly affect the p$K_a$ value of the imidazole moiety. In case of **L9**, a shift towards higher p$K_a$ is expected as suggested by the p$K_a$ differences of benzimidazole / 2-methyl-benzimidazole (5.4 vs. 6.1) or imidazole / 2-methyl-imidazole (7.0 vs. 7.8). The p$K_a$ value of the imidazole moiety will be further shifted by the adjacent negatively charged Glu235. This might result in protonation of the imidazole moiety at a pH of 7.3. The resulting charge assisted H-bond could explain the ten-fold affinity gain going from **L1** (4100 nM) to **L9** (300 nM).

For 2-amino-*lin*-benzoguanine an even more pronounced p$K_a$ shift can be expected, as indicated by the comparison of 2-amino-benzimidazole / benzimidazole (5.4 vs. 7.5). The additional amino group generates a guanidinium functionality in the *lin*-benzoguanine scaffold (p$K_a$ of guanidinium: 13.7; Fig. 3.15). Thus, 2-amino-*lin*-benzoguanine should bind in its protonated form and mimic the charge assisted binding mode of preQ$_1$. Compared to **L9**, for **L10** an additional H-bond can be formed to the carbonyl-group of Ala232. This provides an explanation for additional, more than two fold affinity gain of **L10** (125 nM).

However, to test this hypothesis, determination of the exact p$K_a$ values for the inhibitors as well as *in silico* p$K_a$ calculations of the compounds in complex with TGT

are required. Additionally, **L9** binding should be assayed at different pH conditions to assess a putative pH-dependence of the binding constants.

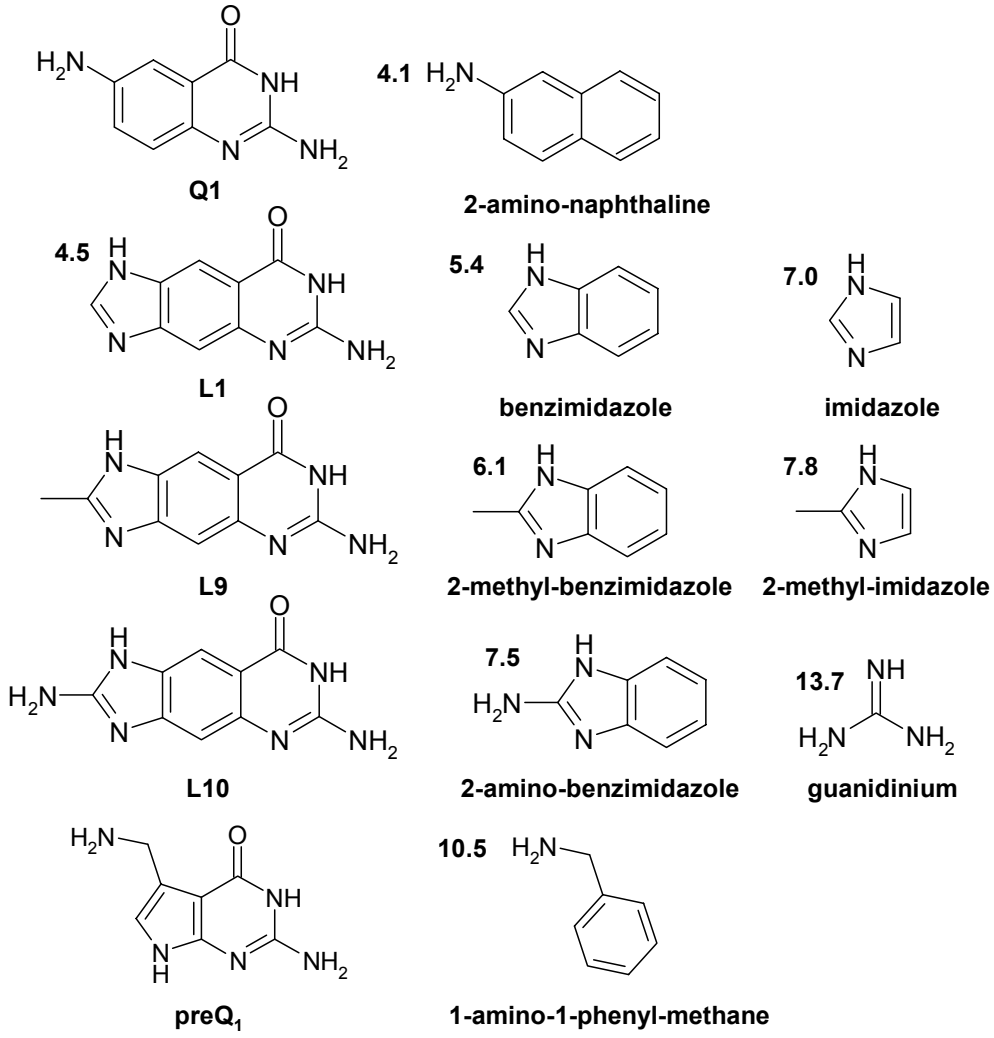
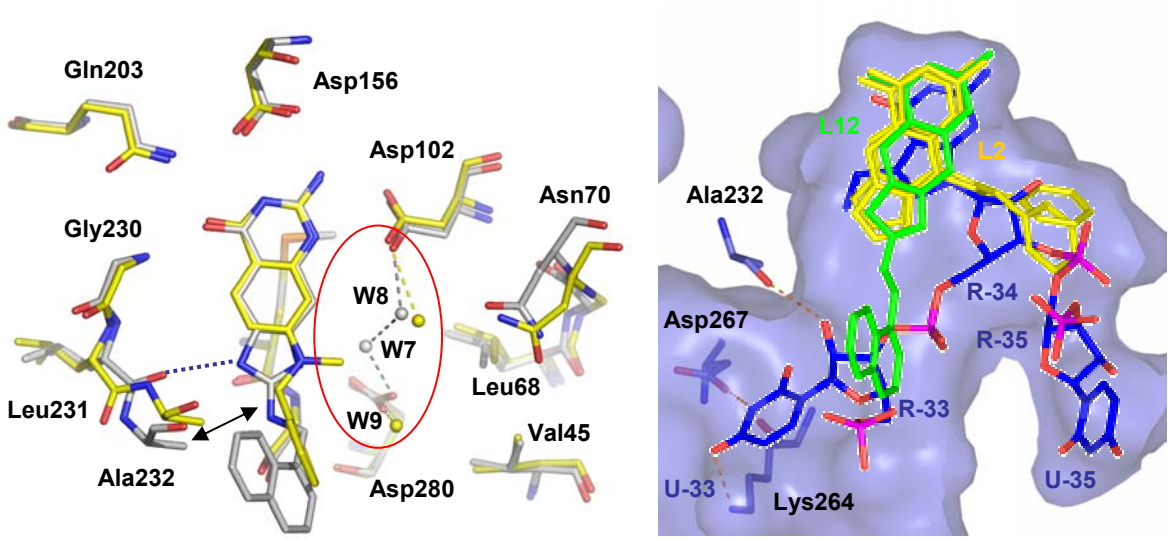



**Fig 3.15** p$K_a$ values for nitrogen groups in relevant scaffold structures; values for fragments taken from [Brown et al., 1955; Albert et al., 1948; Bruice & Schmir, 1958][92-94]

The distinct orientation of the substituent towards the ribose 33 binding site marks the significant difference of this inhibitor series, resulting from the substitution of the *lin*-benzoguanine scaffold in position 2. All previously designed TGT inhibitors were substituted to address the hydrophobic ribose 34 binding pocket formed by Val45 and Leu68 [Grädler *et al*., 2001; Brenk *et al*., 2003; Meyer *et al*., 2002][3, 84, 85]. Also the *lin*-benzoguanines, substituted at position 3 and 4 bind into this pocket (chapter 3.3.2.4). In the present approach it was intended to direct the substituents towards the ribose 33 binding pocket, flanked by Ala232 and Gly261, to provide a scaffold that will allow further extension towards the uracil 33 binding pocket. However,

substitution at the 2-amino group, represented by **L11** and **L12**, had only minor influence on affinity. From the TGT·**L12** crystal structure in can be concluded that these substituents experience too little contact with the protein and remain solvent exposed. No significant contribution to binding can be expected.




a b

**Fig. 3.16** a) Structural superposition of the active sites of TGT·**L12** and TGT·**Q21**; b) structural superposition of the binding modes of 4-substituted **L2** and 2-substitued **L12** in the binding pocket of TGT bound to preQ$_1$-tRNA

Compounds that might adopt binding modes similar to **L12** have been identified in previous inhibitor series [Brenk *et al*., 2003, Brenk *et al*., 2003][2, 84]. E.g. **P5**, **S1** and **Q21** possess substituents to be placed into the ribose 33 binding site (Tab. 3.10). For **Q21** this binding mode was confirmed crystallographically, although unexpected, as the design intended to address the ribose 34 pocket (chapter 3.3.1). In TGT·**Q21** the scaffold is well accommodated, whereas the phenyl substituent is not properly ordered. The structural superposition of TGT·**Q21** and TGT·**L12** suggests criteria to correlate affinity with structure (Fig. 3.16a). **Q21** binds 100-fold weaker than **L12** (75 nM). One explanation for this drop is the missing charge assistance for **Q21**. The Leu231/Ala232 peptide switch adopts the amide exposing conformation and Glu325 will be uncharged. Furthermore, **Q21** places its N-methyl group in a region that interferes with the water cluster. The phenyl substituent of **Q21** contributes significantly to binding. Compared to the unsubstituted scaffold **Q19** (31,000 nM) a

four-fold stronger binding is experienced. Structurally this can be explained by hydrophobic stacking of the phenyl ring with the methyl group of Ala232. In TGT·**L12** a similar contact is given, however, it involves to some degree also the 2-amino group. This might explain the only marginal gain in affinity compared to the unsubstituted **L10** (Tab. 3.10). In summary, the charge assistance and the unperturbed water network between Asp102 and Asp280 seem to be the key factors for the superior binding of **L12**.

Similar arguments hold for the comparison of 2-substituted and 4-substituted *lin*-benzoguanines (Fig. 3.16b). In case of the 4-substitution the water network between Asp280 and Asp102 is expelled from the binding pocket without providing appropriate structural compensation. Most presumably also charge assisted binding is not given for these derivatives.

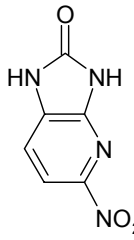
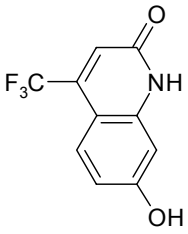
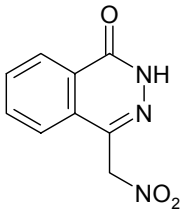
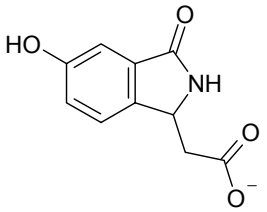
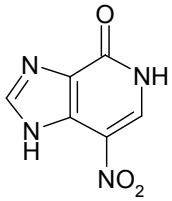
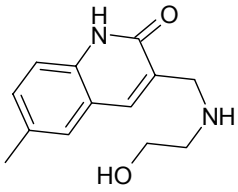
In summary, the 2-amino-*lin*-benzoguanine scaffold provides a promising starting point for further design. However, the phenyl and naphtyl substituents should be replaced by better suited substituents specifically addressing the ribose 33 binding site and to reach out the uracil 33 binding pocket.

# 3.4 Benzimidazolin-2-one-based inhibitors

## 3.4.1 Nitro-substituted virtual screening hits

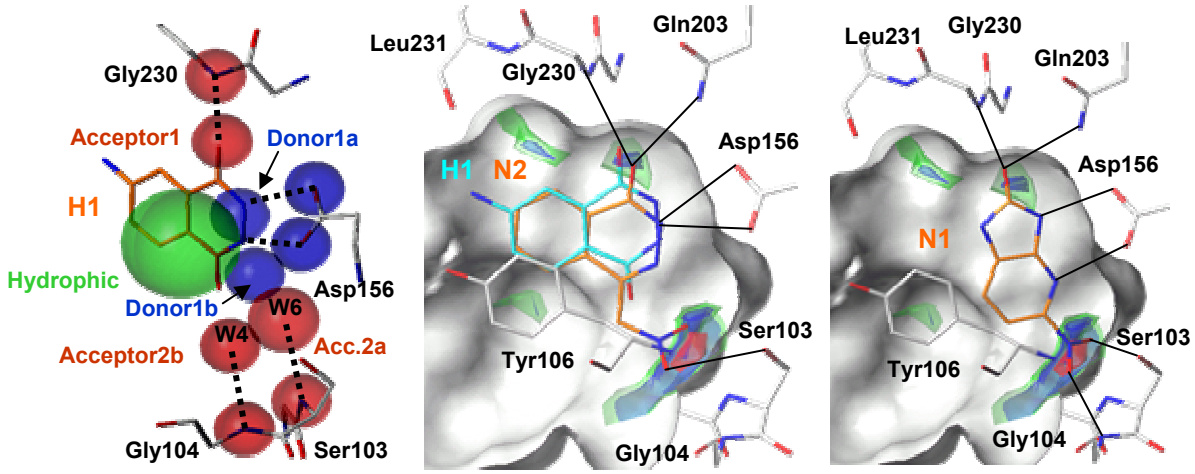
In a virtual screening based on *Z. mobilis* TGT interesting hits substituted with a nitro group were discovered (Tab. 3.11) [Brenk *et al*., 2004][1]. Inhibition in the two-digit micromolar range made them suited to follow-up with an optimization program.

**Tab. 3.11** Nitro-substituted screening hits as potent TGT inhibitors

| | | | | | |
|---|---|---|---|---|---|
| **N1** |  | **36 ± 24**<br>27 ± 3* | **S6** |  | 158 ± 17* |
| **N2** |  | 31 ± 5* | **S7** |  | 403 ± 33* |
| **N3** |  | 15.1 ± 0.1* | **S8** |  | 58 ± 15* |

$K_{ic}$ in [µM] with average error; (*): previous assay; bold: modified assay;

In this virtual screening scenario the positions of several crystal waters next to Asp102 and Asp156, observed in **TGT·H1**, were considered in a pharmacophore hypothesis (Fig. 3.17a). The water positions W4 and W6 were translated as acceptor groups into the pharmacophore, potentially interacting with Ser103 and Gly104.
Six screening hits were selected for testing (Tab. 3.11). The most potent compounds exhibit nitro groups (**N1** - **N3**). The modelled binding modes of **N1** and **N2** are given in Figure 3.17b/c. The nitro groups of both compound are assumed to form interactions with Ser103 and Gly104.




a b c

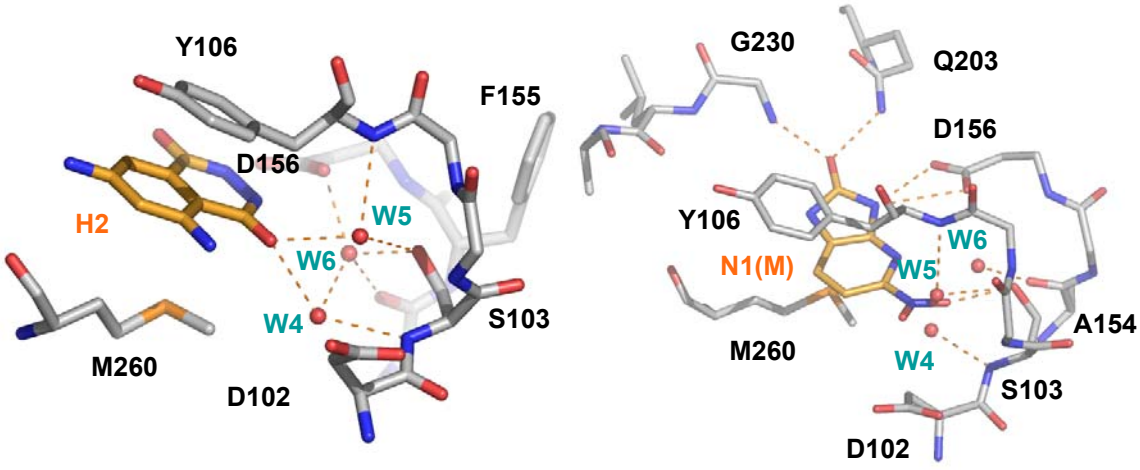
**Fig. 3.17** a) Protein-based pharmacophore hypothesis considering the hydrophobic centre from **H1** and the water molecules W4 and W6; the size of the spheres corresponds to the allowed matching tolerance; b) superposition of the docking mode of **N2** with **H1**; assumed H-bonds are depicted as lines; coloured isopleths indicate favourable interaction sites of donor groups as suggested by DrugScore; c) docking mode of **N1** oriented perpendicularly with respect to the principle molecular axis of **N2** and **H1** with rectangular orientation of the scaffold structure N1 compared to H1; (figures adapted from Brenk et al., 2004[1]).

Unfortunately no crystal structure could be obtained of any of these screening hits in complex with TGT. Docking simulations suggested a binding mode for **N1** with the principle molecular axis orientated perpendicular to that of **H1** or **N2** (Fig. 3.17b/c). This observation, together with the good synthetic accessibility of **N1**, prompted us to select **N1** as novel lead for further optimization.

## 3.4.2 N1 – scaffold evaluation

**N1** (5-nitro-3*H*-imidazo[4,5-*b*]-pyridin-2-(1*H*)one), initially determined to 27 µM [Brenk *et al*., 2004][1], showed under the modified assay conditions no uncompetitive inhibition (stabilization of the covalent TGT-tRNA complex). Likely this results from the fact that **N1** blocks Asp102 from rotation into the active site, a conformational prerequisite for covalent tRNA binding (Fig. 3.1). This finding speaks in favour of the assumed binding mode of **N1** (Fig. 3.17c). The competitive inhibition constant was determined to 36µM which agrees within the assumed error range with the former evaluation (Tab. 3.11).

Prior to modifications of **N1** the suggested binding mode was revised and compared with the crystal structure of TGT·**H2** to better estimate the role of water molecules found in the binding pocket. In TGT·**H2** they mediate contacts between ligand and protein (Fig. 3.18a). W4 is hydrogen-bonded to the amide nitrogen of Ser103 and the ligand, W5 to the side chain of Ser103 and the ligand. A third water molecule, W6, is buried deeply in the $G_{34}$ binding pocket stabilizing the side chain geometries of Asp156, Ser103 and the position of W4. The docked binding mode of **N1** suggests that the nitro group of the ligand replaces W4 and W5. The urea-type scaffold is stabilized *via* H-bonds to Gly230, Gln203 and Asp156. W6 appears to remain in the binding pocket (Fig. 3.18b).

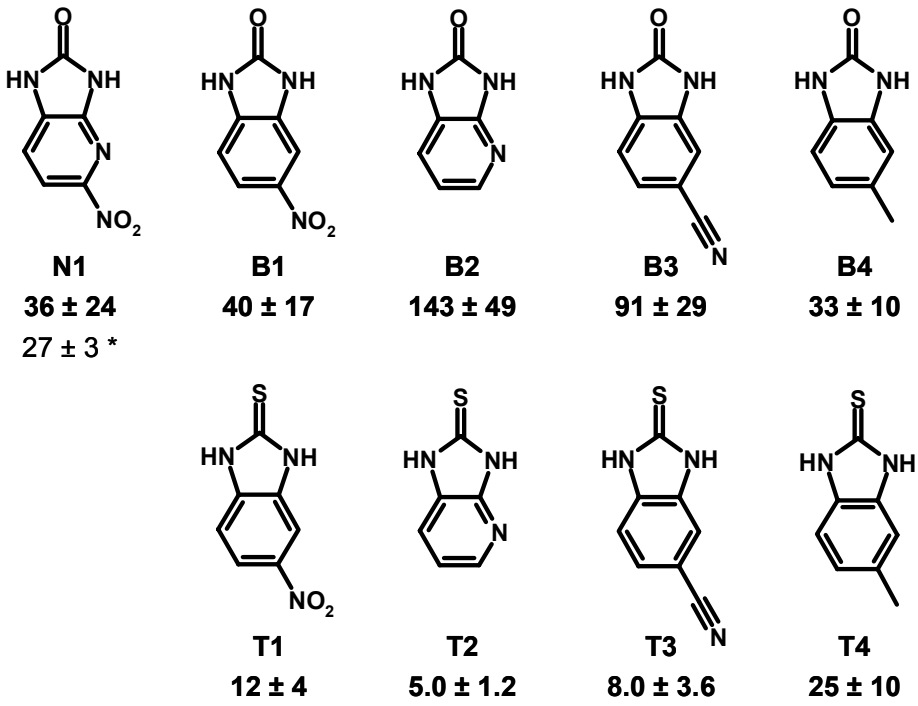



a b

**Fig. 3.18** a) Stabilization of a cluster of crystal waters in the TGT·**H2** binding pocket; b) assumed binding mode of **N1** (docking and MOLOC minimization); W4 and W5 are thought to be replaced upon inhibitor binding.

To investigate the influence of the nitro group on binding derivatives of **N1** were synthesized by Tim Larsen (Prof. A. Link, University of Greifswald) (Tab. 3.12). Compound **B2** (143 µM) clearly demonstrates the relevance of the nitro group for binding. As it was intended to replace the nitro group also by other functional groups the scaffold had to be modified. To simplify synthesis a phenyl ring was introduced resulting in **B1** (5-nitro-benzimidazolin-2-one; 'B' for benzimidazole) exhibiting very similar affinity. Obviously the pyridine nitrogen is not significantly contributing to binding. Subsequently, the nitro-group was replaced by a cyano group (**B3**) or a methyl group (**B4**). The assumed binding modes for **B1**-**B4** (minimized with MOLOC) are all very similar to that of **N1** (Fig. 3.19).

**Tab. 3.12** benzimidazolin-2one-based scaffold series ( $K_{ic}$ in [µM])




(*) inhibition constant refers to previous assay

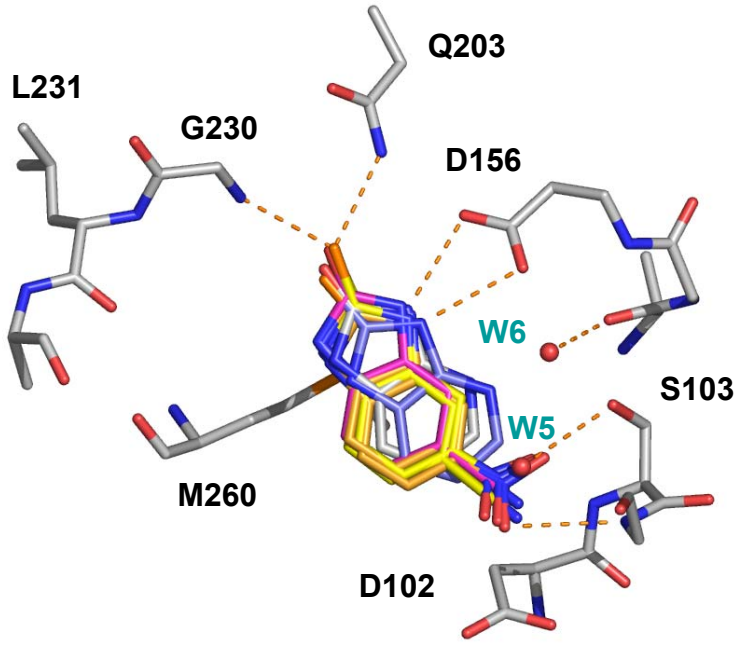


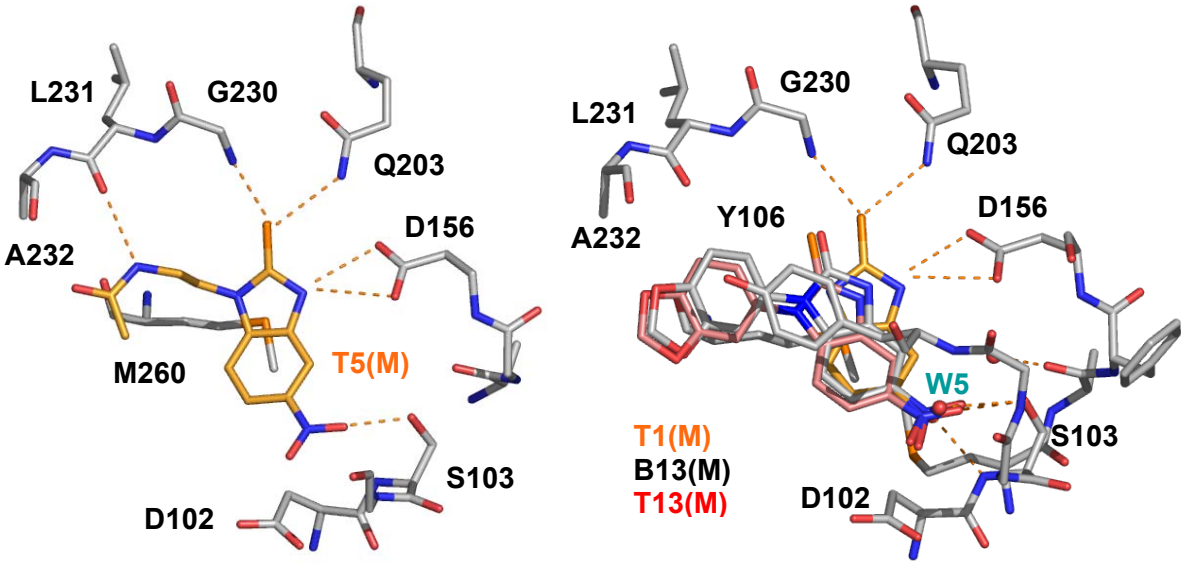

**Fig. 3.19** Superposition of assumed binding modes of **N1**, **B1** – **B4** and **T1** - **T4** (MOLOC minimization)

In a second series the urea carbonyl function of **B1** was replaced by a thiourea unit (benzimidazolin-2-thione-based inhibitor series **T1** - **T4**; 'T' for -thione). Initial minimizations of **T1** - **T4** indicated similar binding modes to **B1**-**B4**, despite of the deviating size and binding properties of sulphur (Fig. 3.19). Surprisingly, all thioureas bind significantly better than the respective oxygen analogues (Tab. 3.12). **T1** (12µM) gains three-fold affinity compared to **B1**. The most pronounced difference is observed for the pyridine analogue **T2** which departs from **B2** by a factor of 30. With an inhibition constant of 5.0 µM it is one of the most potent inhibitor of the series.
To verify the assumed binding modes, crystallization trials for **T1**, **T2** and compounds presented in the following (**B7**, **B15**, **T5**, **T17**) were performed. Neither soaking nor cocrystallization was yet successful. No difference electron density in the $G_{34}$ binding pockets of the studied crystal structures, indicating the presence of a ligand, could be observed. Accordingly, any further design has to be based on the assumed docking modes.

### 3.4.3 Substituted inhibitor series

To address further parts of the binding pocket and to stabilize the binding geometry of the novel scaffold, substituted derivatives of **B1** and **T1** have been synthesized. In a first series polar and aliphatic substituents have been attached (Tab. 3.14). In all cases the inhibition constants determined for the thiourea derivatives are lower than of the equivalent oxy analogues. All inhibition constants are similar to the parent structures or increased by factors of two to four. Only for **T5** a significant gain in affinity compared to the unsubstituted scaffold could be observed. With an inhibition constant of 2.0µM, it is the best inhibitor of this series.
Unfortunately, also in this case soaking into TGT crystals was not successful. The minimized binding mode of **T5** suggests an additional H-bond to be formed by the acetamide substituent with the carbonyl group of Leu231 which might contribute to some degree to binding (Fig. 3.20a).
In a second series aromatic substituents have been attached to the scaffold. Using phenyl substituted side chains, binding affinity increases significantly (Tab. 3.13). **B14** showed limited solubility at higher concentrations under the assay solution. Even presence of detergent *Tween 20* could not compensate the low solubility (chapter 3.1.1.1). **B13** and **T13**, substituted with piperonyl moieties, are more promising. The

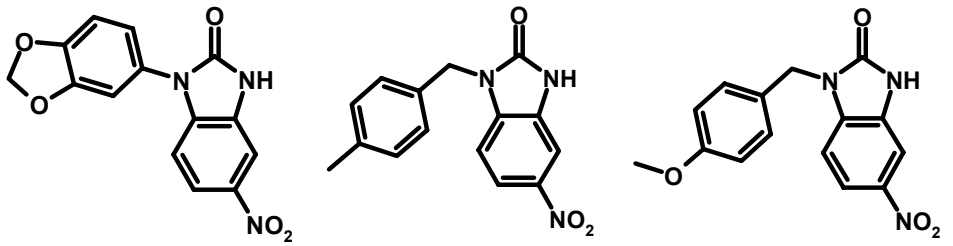
inhibition constant of **B13** is similar to that of **B1**. **T13** gains affinity by a factor of two (5.1 µM) compared to the scaffold **T1**. The assumed binding modes of the latter compounds indicate stacking interactions with Tyr106 (Fig. 3.20b).



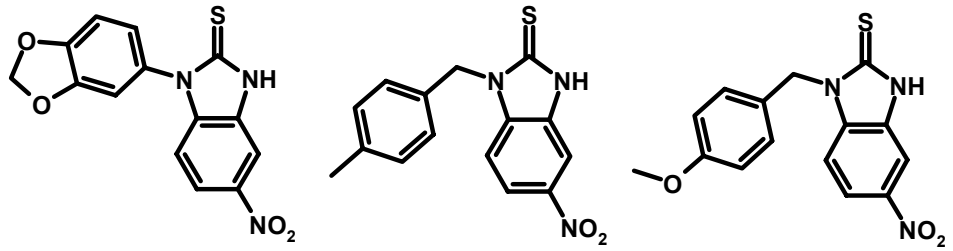

a b

**Fig. 3.20** a) Assumed binding mode of **T5** (MOLOC minimization); b) comparison of assumed binding modes of **T1**, **T13** and **B13** (MOLOC minimization)

**Tab. 3.13** Aromatic derivatives of **B1** and **T1** ( $K_{ic}$ in [µM])

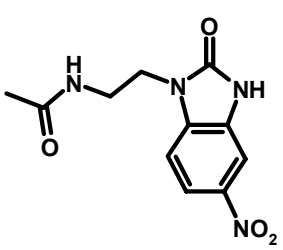


**B13** 39 ± 19 **B14** >50 (limited solub.) **B15** 100 ± 18

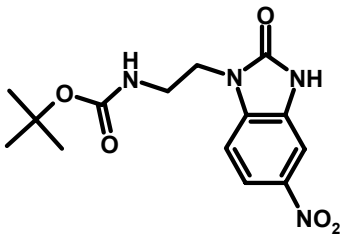


**T13** 5.1 ± 3.1 **T14** 83 ± 49 **T15** 35 ± 9

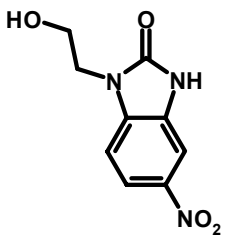
**Tab. 3.14** Polar and aliphatic derivatives of **B1** and **T1** ( $K_{ic}$ in [µM])



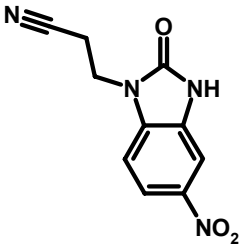
**B5**
**in synthesis**



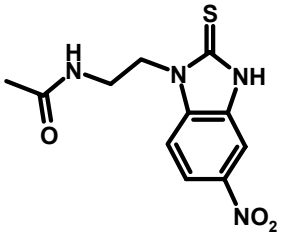
**B6**
**150 ± 40**



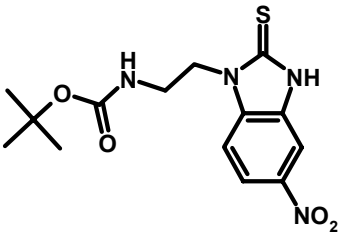
**B7**
**58 ± 35**



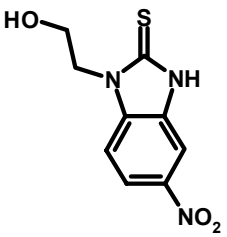
**B8**
**33 ± 6**



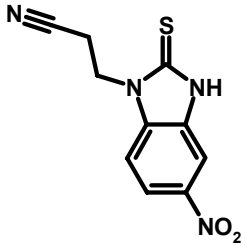
**T5**
**2.0 ± 0.3**



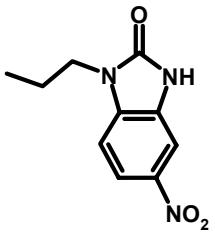
**T6**
**29 ± 13**



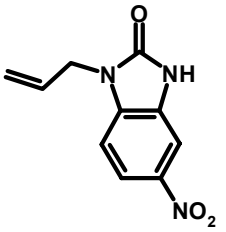
**T7**
**18 ± 4**



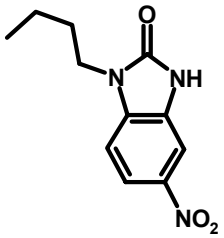
**T8**
**31 ± 11**



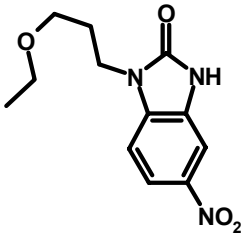
**B9**
**56 ± 26**



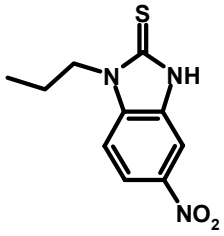
**B10**
**130 ± 33**



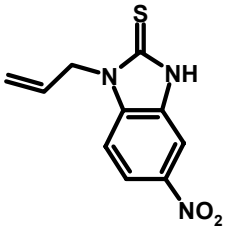
**B11**
**64 ± 7**



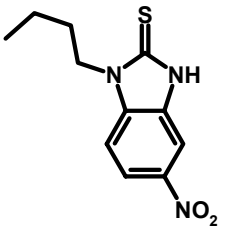
**B12**
**100 ± 40**



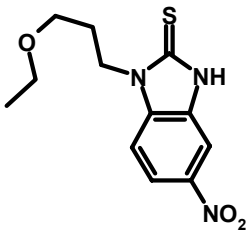
**T9**
**13 ± 6**



**T10**
**57 ± 9**



**T11**
**37 ± 13**

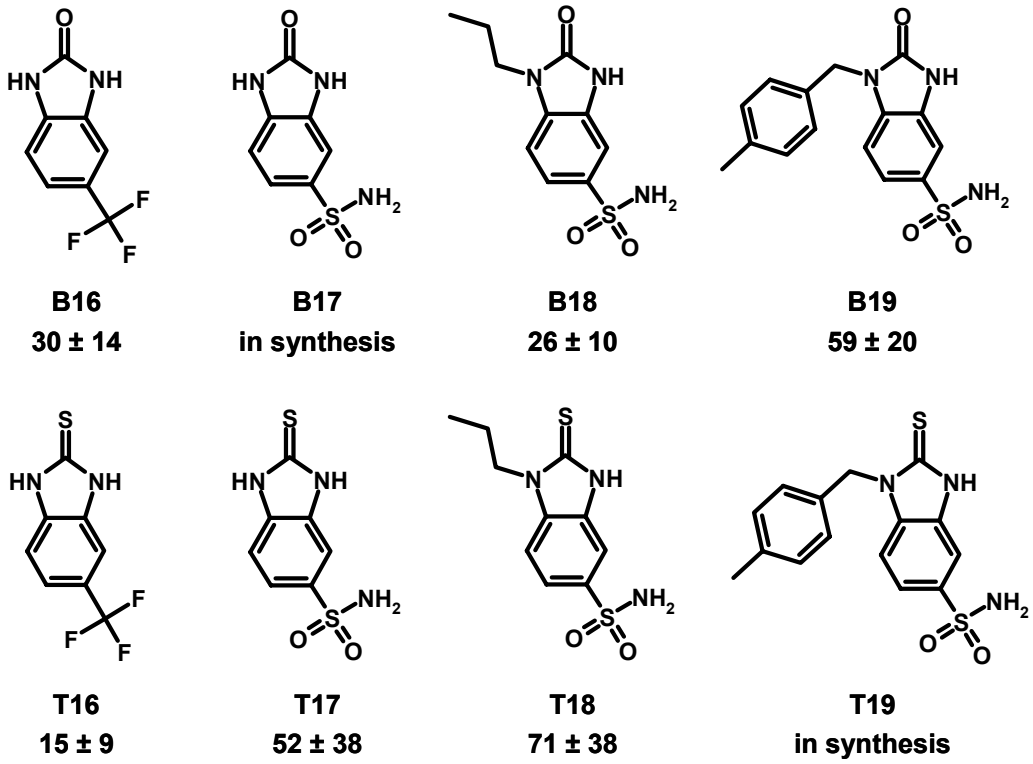


**T12**
**21 ± 5**

## 3.4.4 Sulfonamide-substituted scaffolds

Another aspect of the synthesis program attempted the replacement of the nitro group by functional groups with tetrahedral geometry (Tab. 3.15). In particular substitution with a sulfonamide moiety appeared favourable. The structural superposition suggests the two sulfonamide oxygens in an appropriate geometry to replace W4 and W5.

**Tab. 3.15** Benzimidazolin-2-one /-thione series with trifluor and sulfonamide substituted scaffold ( $K_{ic}$ in [µM])




In case of **B16** and **T16** a $CF_3$ group was attached to the scaffold. The inhibition constants are similar to those observed for scaffolds with a nitro group. In case of the sulfonamides reverse trends in affinity might favour urea-series instead of the thioureas. **T18** is a three fold weaker binder than **B18**. For the corresponding pair of nitro derivatives the thiourea derivative is favoured (**T9**: 13 µM) / **B9**: 56 µM). Other pairs from both series still have to be synthesized. Possibly they will confirm the indicated trends.

### 3.4.5 Outlook for benzimidazolin-2-ones

To confirm the assumed binding modes for the benzimidazolin-2-one inhibitor series crystal structure analysis is definitely required. In particular **B13** / **T13** might be likely candidates due to their good solubility and the assumed stacking interaction with Tyr106. Of the sulphonamide series, **B18** seems to be suited best for crystallization trials.

Furthermore, synthesis of other substituted derivatives appears mandatory. Particularly the free amines of **B5** / **T5** might contribute with a charged side chain, similar to preQ$_1$, to binding affinity and stabilize the urea-scaffold in the binding pocket.

## 3.5 Apigenin-based inhibitors

### 3.5.1 Ligand fishing

In collaboration with Daniela Heller (group of Prof. R. Matusch, University of Marburg) apigenin-based inhibitors of TGT have been investigated. Putative inhibitors of TGT were isolated from plant extracts applying the method of ligand fishing. This method was developed in the group of Prof. Matusch [Lenz, PhD Thesis, 1999][95]. Daniela Heller adjusted the method of ligand fishing to TGT [Heller, PhD Thesis, 2005][96]. In a first step TGT is incubated with compounds from a plant extract to allow specific binding to the enzyme (Fig. 3.21). In following centrifugation and washing steps using microcon centrifugal filter units all compounds that are not specifically bound to TGT are washed off from the enzyme. In a last step the enzyme is denatured releasing specific binders which can be identified by CapLC. *Via* preparative HPLC these substances are then isolated directly from the plant extract.

Twelve plant species were selected that were described in literature to be used in dysentery treatment by indigenous peoples or showed antibacterial activity. Plant extracts of the respective species were produced or purchased. Out of the plant extracts from *Passiflora incarnata* L., *Centella aquatica* (L.) URBAN and *Solanum dulcamara* L. compounds binding to TGT could be identified. The chemical structure of the compound isolated from *P. incarnata* was determined as isovitexin-2''-*O*-β-glucosid which belongs to the flavonoids of the apigenin type. Flavonoids are present

in amounts of 0.5 to 2.5 % in the extract of 'Passiflorae herba', among them predominantly C-glycosids of apigenin and luteolin (Tab. 3.16) [Heller, PhD Thesis, 2005][96].

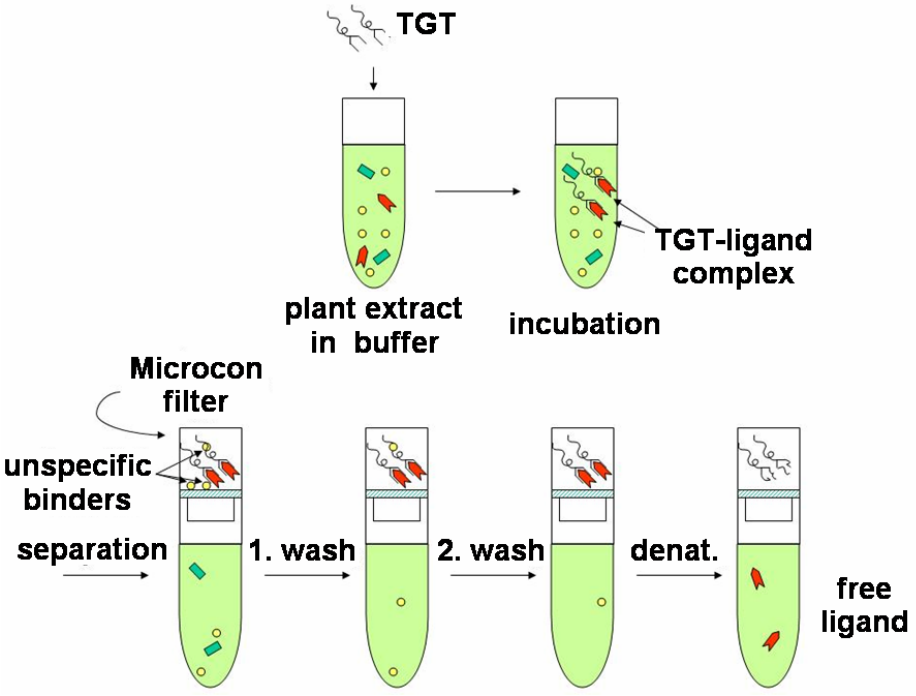
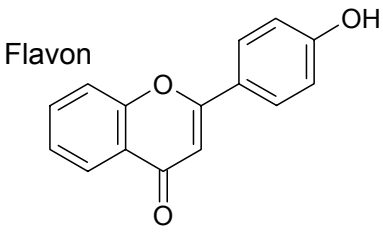
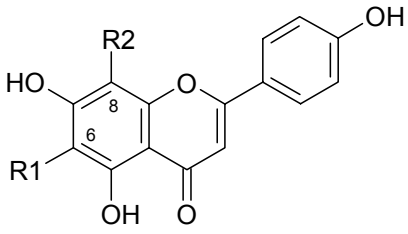



**Fig. 3.21** Work flow for ligand fishing (figure adapted from Heller (PhD Thesis, 2005)[96])

**Tab. 3.16** Chemical structure of flavonoids of the apigenin type
(C-Gluc: C-glycosidic β-D-glucose; C-Arab: C-glycosidic α-L-arabinose)

| Flavon  |  | |
|---|---|---|
| **Flavonoid** | **R1** | **R2** |
| Apigenin | H | H |
| Vitexin | H | C-Gluc |
| Schaftosid | C-Gluc | C-Arab |
| Isoschaftosid | C-Arab | C-Gluc |
| Isovitexin | C-Gluc | H |
| Isovitexin-2''-*O*-β-glucosid | C-Gluc-O-β-Gluc | H |

## 3.5.2 Determination of inhibition constants

The inhibition constant of the isolated isovitexin-2''-*O*-β-glucosid was determined together with those of the parent structures isovitexin and apigenin. To identify specificity and selectivity determining features the inhibition constants of vitexin, schaftosid and isoschaftosid, structurally related C-glycosids of apigenin, were determined as well. For apigenin, vitexin, isovitexin, schaftosid and isoschaftosid the commercially available pure compounds were used in the assay.

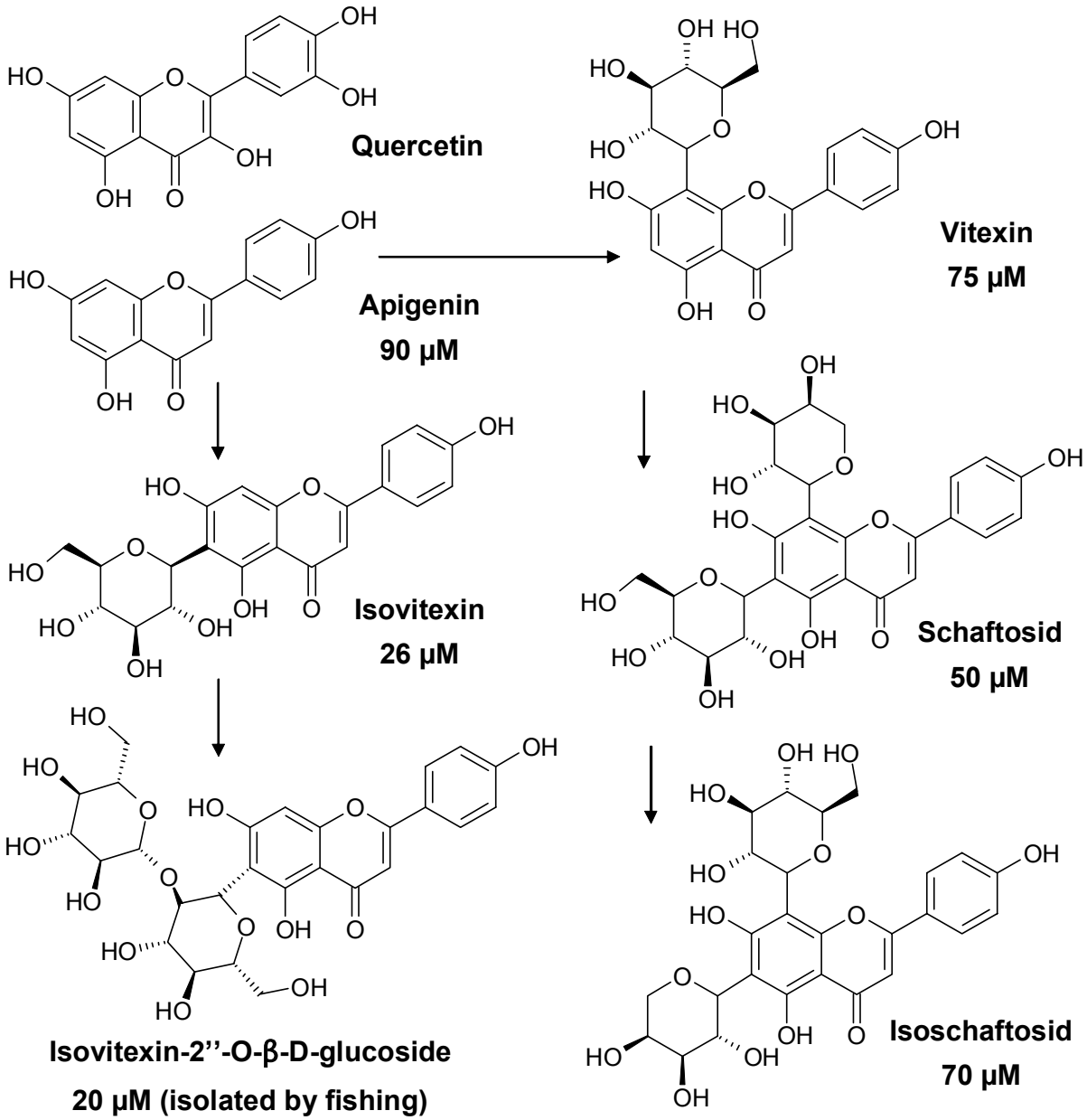



**Fig. 3.22** Competitive inhibition constants of flavonoids of the apigenin type (values refer to the modified assay protocol)

In an initial experiment, it was discovered that apigenin is a non-specific inhibitor of the TGT base exchange reaction (see chapter 3.1.1). Due to a chemical structure similar to quercetin, (Fig. 3.22), which was identified as non-specific inhibitor of relevant pharmaceutical targets, this experiment appeared essential [McGovern *et al*., 2002][77]. In the absence of the detergent *Tween 20*, the inhibition constant of apigenin was determined to 7.4µM, compared to 78µM in the presence of *Tween 20* (values refer to the former assay protocol – see chapter 3.1.4). This finding clearly demonstrates the ability of apigenin to inhibit TGT which is, however, overestimated in the absence of a detergent.

Isovitexin-2''-*O*-β-glucosid is the most potent inhibitor of all tested compounds with a $K_{ic}$ of 20µM. Isovitexin exhibits a similar binding affinity as isovitexin-2''-*O*-β-glucosid, whereas the parent structure apigenin looses affinity by a factor of four. The 8-substituted apigenins (vitexin, schaftosid, isoschaftosid) are capable to inhibit TGT as well, but the inhibitory potency remains similar to that of the apigenin scaffold. All values given in Figure 3.22 refer to the modified assay protocol (chapter 3.1.4). Facing the only small affinity differences, it remains not understood why Isovitexin-2''-*O*-β-glucosid can be isolated by ligand fishing, although all other compounds are known to be present in the *Passiflora* plant extract as well.

## 3.5.3 Docking experiments

To identify the binding modes of isovitexin-2''-*O*-β-glucosid and isovitexin, crystal structure analysis in complex with TGT have been performed. In particular isovitexin-2''-*O*-β-glucosid seemed to be suited for the experiments due to its good solubility in water. However, neither soaking nor cocrystallization were successful. No difference electron density in the binding pocket, indicating the presence of a ligand, could be identified.

To get some ideas about putative binding modes, docking experiments were performed with GOLD using apigenin and isovitexin-2''-*O*-β-glucosid as ligands.

The crystal structure of TGT·**H2** (Fig. 3.18a), after removing the ligand and all water molecules, was used for docking. Applying standard settings in GOLD 3.0.1 twenty solutions were created for apigenin in a region of 20Å surrounding the $G_{34}$ binding pocket. For isovitexin-2''-*O*-β-glucosid the number of accepted solutions was

increased to 25 to take the higher flexibility of the compound into consideration (chapter 5.6.2).

For apigenin 19 of 20 solutions cluster together. The only exception is found on rank 19. Apigenin is rotated by 180° with respect to the orientation found in the main cluster (Fig. 3.23b). All solutions are located in the $G_{34}$ binding pocket of TGT stabilized *via* hydrophobic stacking interactions with Tyr106 and Met260. In all solutions the 2-hydroxyphenyl ring points out of the binding pocket (Fig. 3.23d). The best solution, according to the GOLDScore scoring function, forms the following interactions with the active site (Fig. 3.23c). The 4-carbonyl group is H-bonded to Gly230, the 5-hydroxy group interacts with Gln203 and Asp156, and the 7-hydroxy group is forming an H-bond with Ser103.

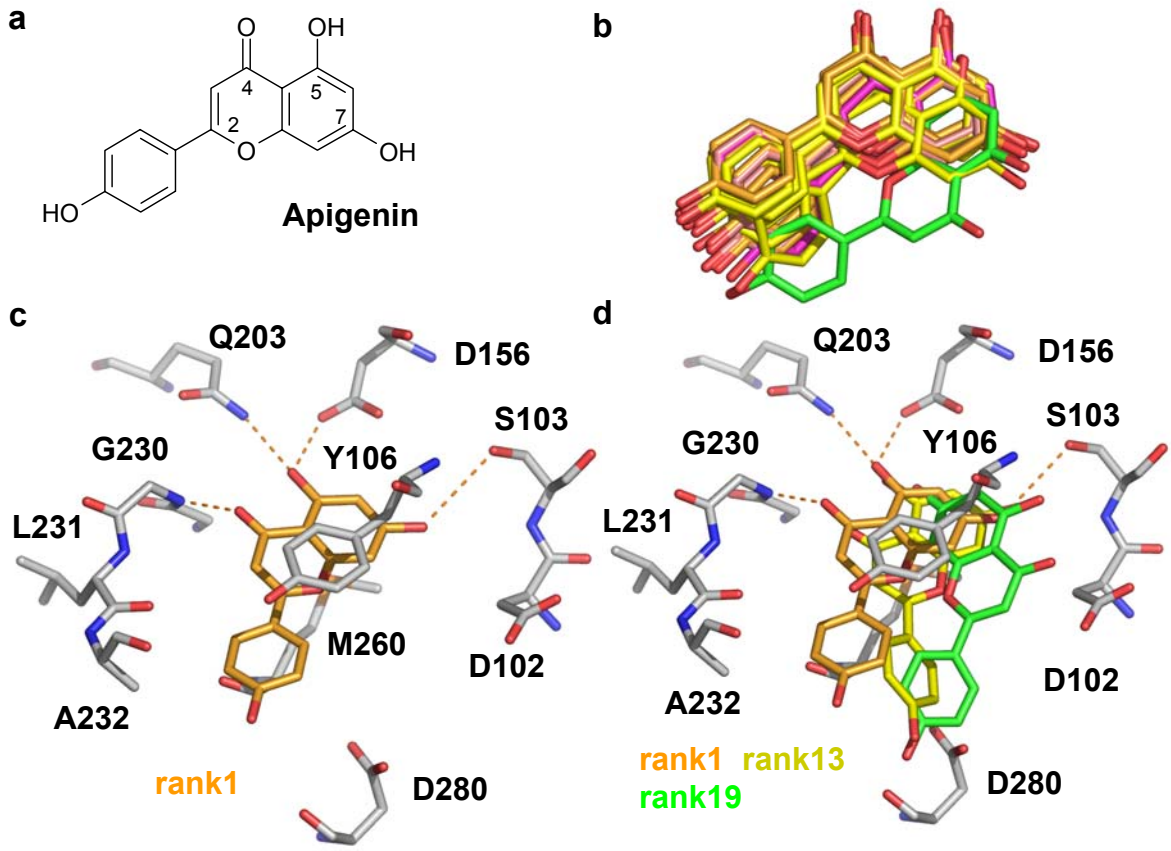



**Fig. 3.23** a) Chemical structure of apigenin; b) superposition of 20 docking solutions for apigenin from GOLD; the solution with reversed orientation is shown in green; c) binding mode of the apigenin docking solution ranked on position 1; d) comparison of docking solutions ranked on the positions 1, 13 and 19;

For isovitexin-2''-*O*-β-glucosid the docking solutions form two clusters with respect to the orientation of the 2-hydroxyphenyl group. Two solutions from rank 18 and 25 form

a small cluster. The main cluster comprises all other solutions (Fig. 3.24a). The superposition of the best ranked solutions from the two clusters with the TGT-bound tRNA shows that in the main cluster the 2-hydroxyphenyl substituent points into the $G_{34}$/preQ$_1$ binding pocket (Fig. 3.24b). In the second cluster this substituent is located in the $U_{33}$ binding pocket. For the binding geometry found on rank 1 only a few specific interactions with the $G_{34}$-binding site are encountered (Fig. 3.24c). The 2-hydroxyphenyl substituent is H-bonded to Gln203 and Asp156. Two H-bonds are formed by the apigenin scaffold with Ala232 and Asp280. The second glucose moiety, however, fits well into the subpocket formed by Val45 and Val282.

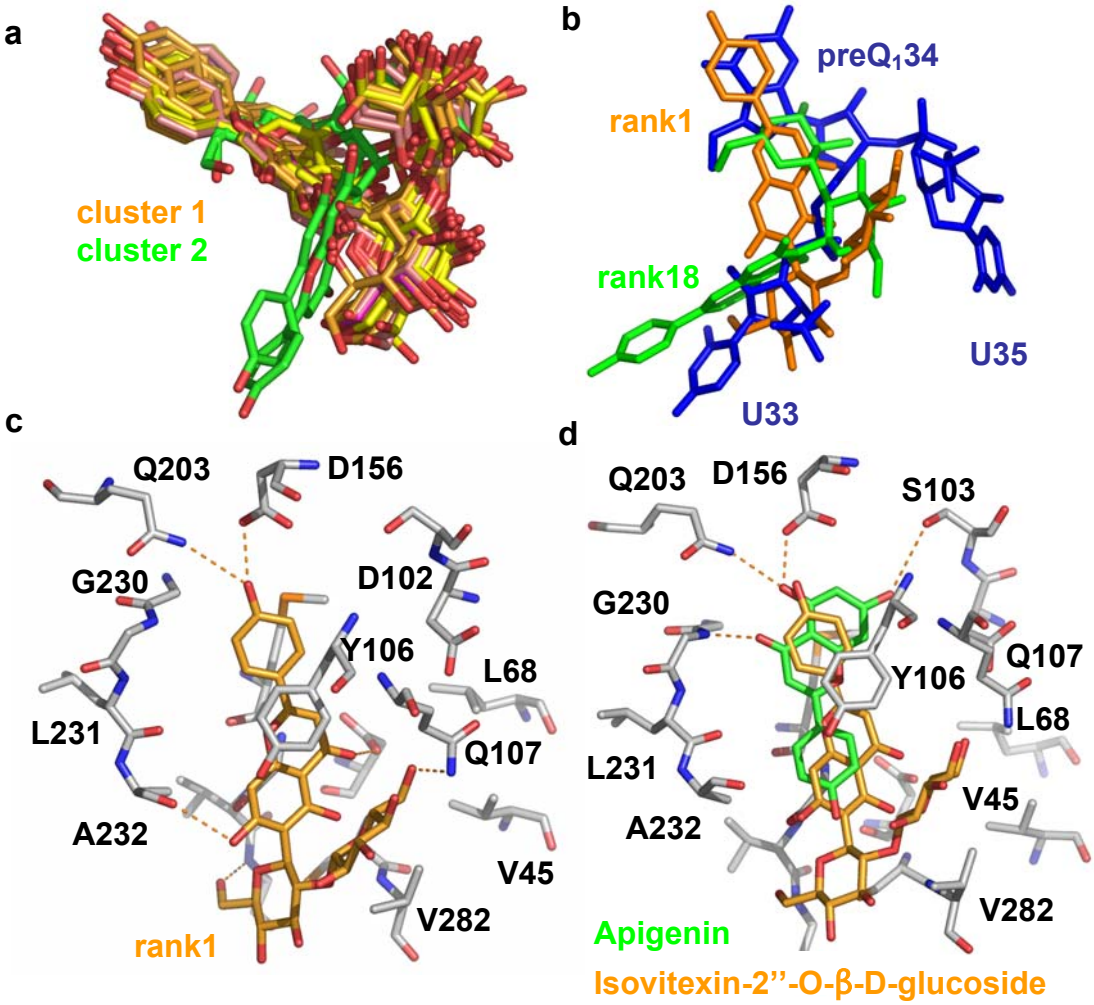



**Fig. 3.24** a) Superposition of 25 docking solutions for isovitexin-2''-*O*-β-glucosid from GOLD; b) superposition of representatives from the two clusters (orange and green) on tRNA bound to TGT (blue); c) binding geometry of isovitexin-2''-*O*-β-glucosid (orange) found on rank 1; d) superposition of binding geometries of apigenin (green) and isovitexin-2''-*O*-β-glucosid (orange) found on rank 1

The comparison of the first rank docking solutions of apigenin and isovitexin-2''-*O*-β-glucosid reveals a surprising inversion of the apigenin scaffold orientation (Fig. 3.24d). While the apigenin scaffold addresses most of the residues involved in the recognition of $G_{34}$, in case of isovitexin-2''-*O*-β-glucosid the interactions with this pocket are reduced. Due to the scaffold inversion interactions are dominated by hydrophobic stacking with Tyr106 and Met260.

The suggested structure - affinity model could be tested using C-glycosides structurally similar to isovitexin-2''-*O*-β-glucosid, but based on alternative flavonoid scaffolds. In case of quercetin (Fig. 3. 22), for instance, the additional hydroxy group of the 2-phenyl ring would be suited to form additional interactions with the $G_{34}$ binding site.

# 4. Summary and Outlook

## 4.1 Summary

In this study structure-based drug design approaches for *Zymomonas mobilis* tRNA - guanine transglycosylase (TGT) are presented. Functional analysis and a mutational study provided deeper insight into the catalyzed reaction to support design strategies. Structurally, the *Z. mobilis* enzyme is highly homologous to the TGT from *Shigella sp.*, the causative agent of bacterial shigellosis, for which no crystal structures are available. The development of TGT inhibitors, one of the key pathogenicity factors of *Shigella sp.*, could result in a potent new antibiotic.

The detailed understanding of functional properties of TGT is highly relevant for structure-based design as it provides valuable information about the binding pocket. TGTs are evolutionary ancient enzymes, present in all kingdoms of life, that catalye guanine exchange within their cognate tRNAs by modified 7-deazaguanine bases. The eubacterial TGT catalyzes the exchange of guanine 34 in four specific tRNA molecules by the modified base $preQ_1$. Although distinct bases are incorporated into tRNA at different positions in a kingdom-specific manner, the catalytic subunits of all TGTs are structurally conserved (chaper 2.1).

Crystal structures of QueTGT and ArcTGT in complex with tRNA became available during this thesis and revealed a misinterpretation of previous data [Xie *et al*., 2003][4]. Asp280, instead Asp102, is the nucleophile of the base exchange reaction. Therefore, the model of the base exchange mechanism required revision. However, no comprehensive interpretation was available. Analysis of TGT crystal structures and sequences were performed and provided a new functional model concerning the molecular basis of the reaction mechanism and the residues involved in this process (chapter 2.1). Substrate binding modes indicate an evolutionarily conserved base exchange mechanism in all kingdoms of life with the conserved aspartate 280 (following *Z. mobilis* numbering) serving as nucleophile *via* covalent binding to C1' of the guanosine ribose in an intermediate state.

Eubacterial TGT exhibits substrate promiscuity compared to archaebacterial TGT. In addition to the natural substrate $preQ_1$, also the biochemical precursor $preQ_0$ can be incorporated into tRNA. The structural prerequisite for promiscuity was identified

(chapter 2.2.1). A peptide switching functionality is present in the active site, gated by the general acid/base Glu235. The flip of the Leu231/Ala232 peptide bond modifies the properties of the binding pocket to allow accommodation of tRNA-bound guanine and $preQ_1$. The binding pocket that accepts guanine, however, is also capable to recognize $preQ_0$. As no kinetic data for $preQ_0$ and $preQ_1$ was available, selectivity regulation remained unresolved. Therefore, the missing data for the wild type were determined. To invert selectivity a TGT(E235Q) mutated enzyme was constructed (chapter 2.2.2) This mutant was intended to stabilize the peptide switch in a geometry favouring the binding of $preQ_0$. The kinetic characterization and the crystallization of the TGT(E235Q) in complex with $preQ_0$ allowed a detailed comparison with the wild type enzyme. In TGT(w.t.), $preQ_1$ and $preQ_0$ exhibit similar $K_m$ values. Selectivity in favour of $preQ_1$ is guaranteed via turn-over rates. In the mutated enzyme the $K_m$ value for $preQ_1$ was significantly higher while the $K_m$ of $preQ_0$ remained virtually identical. In the TGT(E235Q) crystal structures, the peptide switch is stabilized by the glutamine side chain and permanently arrested in a geometry suited for $preQ_0$ binding. The observed geometry, however, differes significantly from the expectation. Concerning $k_{cat}$, $preQ_1$ is still the preferred substrate in the mutated enzyme. This suggests that substrate recognition is obviously regulated by Glu235. However, once accommodated in the active site the reaction rate of the base exchange is determined by other factors. Regarding $k_{cat}/K_m$, an indicator for the catalytic efficiency, the overall selectivity is inverted in favour of $preQ_0$ for TGT(E235Q).

From the detailed analysis of TGT crystal structures and sequences it can be hypothesized that for eubacterial and the eukaryotic TGTs, dimer formation is relevant to perform the catalytic reaction (chapter 2.3). In all available crystal structures of *Z. mobils* TGT homodimers are present in the crystals. Also in the crystal structure of *Thermotoga maritima* TGT homodimers are present exhibiting an almost identical interface geometry. Detailed analyses of these contacts and consulting available TGT sequences from Eubacteria and Eukaryota revealed conservation across the contacting residues. This suggests that eubacterial and eukaryotic TGT most probably act as homodimers in catalysis. It is hypothesized that one unit of the dimer performs the catalytic reaction. The second is required to recognize and properly orient the bound tRNA for the catalytic reaction likely *via*

other conserved residues that were identified in docking studies. These assumptions, however, need further experimental validation.

The tRNA-bound crystal structures of ArcTGT and QueTGT allow, due to their substrate binding modes, the extraction of valuable information about the evolutionary origin of the TGT superfamily (chapter 2.4). Although the TGT superfamily is not involved in metabolism, as almost all other enzymes with TIM barrel fold, it shares a standard phosphate binding (SPB) motif with other homologous superfamilies [Nagano *et al*., 2005][72]. The structural comparison of the tRNA-bound TGT structures with other TIM barrel enzymes reveals conserved phosphate binding modes that confirm the previously made assumptions. Additionally, a model for the evolution within the TGT superfamily has been developed.

Results from structural and functional analysis are highly relevant for structure-based design approaches and inhibitor evaluation. In case of TGT, the new insight into the base exchange mechanism required the modification of the previously applied protocol to determine inhibition constants (chapter 3.1). The base exchange reaction was identified to follow a ping-pong mechanism [Goodenough-Lashua & Garcia, 2003][53]. Therefore, competitive and uncompetitive inhibition contributions have to be considered depending on the ligand's size and properties. Uncompetitive inhibition is predominantly relevant for small-sized compounds with molecular size and properties similar to guanine or $preQ_1$. To face these complications the previously applied assay was modified. In trapping experiments uncompetitive binders can be identified. The new assay setup allows the separate evaluation of competitive and uncompetitive inhibition contributions with respect to tRNA binding. Additionally, to avoid the effects of non-specific inhibition, the detergent *Tween 20* is used in the assay. Such a type of inhibition is often observed for compounds with low solubility – and most of the tested compound series, presented in this study, exhibit low solubility. As a secondary effect, however, the protein solubility is significantly increased by the detergent. Therefore, the kinetic parameters were redetermined and corrected values were applied in the assay. Concerning $K_m$ values no significant differences could be identified, however, concerning $V_{max}$ a dramatic increase was detected.

Due to the required modifications of the assay protocol, inhibition constants for already crystallized and characterized compounds were redetermined and their structure – activity relationship was revalidated (chapter 3.2). In particular small-sized

inhibitors have been overestimated concerning their inhibitory potency due to the previously neglected uncompetitive contribution. However, uncompetitive inhibition is only relevant for inhibitor classes with scaffold structures similar to guanine. In particular inhibitors from the class of quinazolinones have been revalidated as this scaffold served as template for further structure-based design approaches described in this study. For substituted compound series the revalidation revealed that inhibition constants are modulated in non-linear fashion. Most presumably contributions of non-specific inhibition resulting from low solubility were compensated by the detergent *Tween 20*. The above described only small affinity differences for the Leu231/Ala232 peptide switch were confirmed in this series. However, the previously reported sulphur effect, [Meyer *et al.*, 2002][85], could not be confirmed. Instead, the accommodation of the ribose 34 binding pocket has only minor influence on binding affinity.

Based on the quinazolinone scaffold three new TGT inhibitor series were developed (chapter 3.3). The compounds were synthesized by E. Meyer and S. Hörtner (group of Prof. F. Diederich, ETH Zürich).

A series of 7-substituted-quinazolinones (chapter 3.3.1) exhibits a decreasing inhibitory potency. Nevertheless, some valuable information for further design cycles is available from this series. Firstly, substitution in position 7 disables uncompetitive inhibition of the base exchange reaction. Secondly, successful crystal structure analysis of one compound in complex with TGT revealed that a water network inside the binding pocket in the neighbourhood of the nucleophile Asp280 is highly relevant for binding affinity and should only be replaced by means of polar interactions. Unexpectedly, the substituent of one crystallographically analyzed derivative points into the direction of the ribose 33 binding site and not, as planned, towards the ribose 34 pocket.

A series of *lin*-benzoguanines (chapter 3.3.2) exhibits a similar inhibitory potency as derivatives of 6-amino-quinazolione scaffold compounds. Inhibition constants are in the single digit micromolar range. Due to an improved solubility of the *lin*-benzoguanines four crystal structures in complexes with TGT have been determined up to a resolution of 1.58 Å – 2.1 Å. These structures give insight into the structural flexibility of TGT necessary to perform catalysis. In three of the structures molecular rearrangements are observed that match with conformational changes also noticed upon tRNA substrate binding. Several water molecules are involved in

these rearrangement processes. Two of them, W2 and W3, demonstrate the structural and the catalytic importance of water molecules during the TGT base exchange reaction. In the fourth crystal structure the inhibitor influences interactions of TGT dimers present in the crystals, due to the rearrangement of a protein loop. This loop is involved in inhibitor binding as well as in dimer interface formation. In due course of this rearrangement a space group transition compared to uncomplexed crystals can be observed upon soaking.

Substitution of *lin*-benzoguanine in position 2 resulted in an inhibitor series with significantly enhanced binding affinity (chapter 3.3.3). With inhibition constants up to the two digit nanomolar range it is the most potent TGT inhibitor series yet known. The successful crystal structure analysis of one compound in complex with TGT provides some explanations for the significant affinity gain. Substitution in position 2 is not interfering with the above mentioned crucial water cluster in the neighbourhood of Asp280 and the ribose 34 binding pocket. Instead, the substituents orient towards the U33 binding channel, although no specific interactions are formed. The assumption of a charge assisted binding mode might be the key element to explain the significant affinity gain. *Lin*-benzoguanines are binding via their imidazole moiety to the Leu231/Ala232 peptide switch in an orientation that requires Glu235 to be negatively charged. In case of *lin*-benzoguanine the p$K_a$ of this moiety is 4.5. Most presumably this moiety is uncharged at the assay condition of pH of 7.3. Substitution of the scaffold in 2-position supposedly causes a raise of the p$K_a$ which results in protonation and charge-assisted binding.

A structurally very different series of inhibitors was realized based on a scaffold structure identified by virtual screening (chapter 3.4). Modification of this scaffold resulted in a benzimidazolin-2-one based inhibitor series. 1- and 5-substituted derivatives of benzimidazolin-2-one and benzimiazolin-2-thione were synthesized by T. Larsen (Prof. Link, University of Greifswald). The inhibition constants of these compounds are in a range of 2 – 150µM. Crystallization attempts of selected compounds in complex with TGT were not successful. Thus, any structure – activity relationships have to be based on docking results. These do not yet provide a fully conclusive picture.

Applying an alternative screening method, described as “ligand fishing”, a compound binding to TGT could be discovered and identified from a plant extract (chapter 3.5). Isolation and characterization were performed by D. Heller (group of Prof. R.

Matusch, University of Marburg). This compound, isovitexin-2''-*O*-β-glucosid, exhibits an inhibition constant of 20 µM. Unfortunately, crystallization in complex with TGT was not successful. Docking studies do not suggest a fully convincing specific recognition of this residue in the active site of TGT. Therefore, the exact binding mode still remains to be elucidated.

## 4.2 Outlook

Concerning the reaction rate of the base exchange, $preQ_1$ is incorporated ten fold faster into tRNA than $preQ_0$ (chapter 2.2.2). Tyr258 which is stabilizing the binding geometry of the nucleophile Asp280 in QueTGT could be involved in the regulation of turn-over rates. This is indicated by structural comparisons with ArcTGT where this tyrosine is conservatively replaced by histidine. To test this hypothesis TGT(Y258F) and TGT(Y258H) mutants could be constructed.

Eubacterial TGTs are supposed to form functional dimers (chapter 2.3). Whether dimers are only present in crystals or whether they can also be observed in solutions in the presence of tRNA could be tested by gel filtration or dynamic light scattering.

2-Amino-*lin*-benzoguanine is a very promising scaffold for further inhibitor design cycles that might result in subnanomolar inhibitors. To validate the assumption of a charge assisted binding mode for this compound determination of the exact $pK_a$ values as well as *in silico* $pK_a$ calculations of the compound in complex with TGT are required. Additionally, inhibition constants should be assayed at different pH conditions to assess a putative pH-dependence of the binding constants.
The substituents added to the 2-amino-*lin*-benzoguanine scaffold orient into the uracil 33 binding site without forming specific interactions. Introduction of substituents that specifically address residues of the ribose 33 and uracil 33 binding pocket might result in inhibitors with inhibition constants in the lower nanomolar range.
Potent inhibitors of this series should be tested for inhibition of *S. flexneri* TGT or alternatively the almost identical *E. coli* TGT.

For the benzimidazolin-2-one based inhibitor series no distinct binding mode could be identified. To confirm the assumed binding modes further crystal structure analyses are required. In particular **B13** / **T13** might be likely candidates due to their good solubility and the assumed stacking interaction with Tyr106. Of the sulphonamide series, **B18** seems to be suited best for crystallization trials.

Furthermore, synthesis of other substituted derivatives appears mandatory. Particularly the free amines of **B5** / **T5** might contribute with a charged side chain, similar to $preQ_1$, to binding affinity and stabilize the urea-scaffold in the binding pocket.

For isovitexin-2''-*O*-β-glucosid binding modes can only be assumed from docking. The suggested structure - affinity model could be tested using C-glycosides structurally similar to isovitexin-2''-*O*-β-glucosid, but based on alternative flavonoid scaffolds. In case of quercetin, for instance, the additional hydroxy group of the 2-phenyl ring would be suited to form additional interactions with the $G_{34}$ binding site.

# 5. Materials and Methods

For additional information concerning the applied methods see also Grädler (PhD Thesis, 2000)[51] and Brenk (PhD Thesis, 2003)[52].

## 5.1 Chemicals and materials

Chemicals used in this study were purchased at Sigma-Aldrich, Merck and Roth.
Plastic materials (pipette tips and reaction cups) were autoclaved before use (30 min, 121 °C, 1.3 bar).

## 5.2 Biochemical methods

### 5.2.1 Media and stock solutions

**LB-Medium**: 1.0 % (w/v) bacto-trypton
0.5 % (w/v) yeast extract
0.5 % (w/v) NaCl

All media were autoclaved immediately after preparation (30 min, 121°C, 1.3 bar).
For the preparation of agar-media 1.5 % (w/v) agar was added to the solution.

**Ampicillin (Amp)**: 100 µg/mL in aqua bidest.
**Chloramphenicol (Cm)**: 34 µg/mL in ethanol
**Kanamycin (Km)**: 30 µg/mL in aqua bidest.

The antibiotic stock solutions were filtered with a 0.2 µm filter.
The antibiotic stock solutions were added at 0.1% (v/v) to the respective media solutions.

## 5.2.2 Determination of concentrations

Absorptions of biological polymers in aqueous solutions were determined *via* photometry to determine their concentrations. The absorption maximum for DNA is at 260 nm, for proteins at 280 nm.

**DNA solution** (empirical formula for double stranded DNA):

**µg DNA / mL = $A_{260}$ × 50 µg/mL × factor of dilution**

The purity was estimated from the $A_{260}/A_{280}$ ratio.
At a ratio of 1.8 – 2.0 the purity is in the range of 70 - 95 %.
[Sambrook *et al*., 1989][97]

***E. coli* tRNA$^{Tyr}$ solution** [Curnow *et al*., 1993][98]:

**1 µM tRNA$^{Tyr}$ ≡ 0,703 $A_{260}$**

The purity was estimated from the $A_{260}/A_{280}$ ratio.
At a ratio of 1.8 – 2.0 the purity is in the range of 90 %.
[Sambrook *et al*., 1989][97]

***Z. mobilis* TGT solution**

**1 mg/mL (23,4 µM) *Z. mobilis* TGT (w.t.) ≡ 0,778 $A_{280}$**

The absorption coefficient was calculated from the TGT amino acid sequence. [Cantor & Schimmel, 1980][99]

### 5.2.3 Strains and plasmids

**Tab. 5.1** *E. coli* strains used in this study.

| Strain | Description | Origin |
|---|---|---|
| XL2-blue | *(mcrA)183(mcrB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac*[*F'proAB laqI*$^q$ ZΔM15 Tn5(Km$^r$)] | Stratagene |
| BL21(DE3) pLysS | F$^-$ *dcm ompT hsdS*($r_B^-m_B^-$)*galλ*(DE3) [pLys Km$^r$] | Stratagene |
| TG2 | [*supE hsdΔ5 thi Δ(lac-proAB) Δ(sre-recA) 306::Tn10* (Tet$^r$) F' (*traD36pro AB*$^+$ *lacI*$^q$ lacZΔM15)] | Stratagene |

**Tab. 5.2** Plasmids used in this study

| Plasmid | Description | Origin |
|---|---|---|
| pET9d-ZM4 | Cm$^r$, Km$^r$; ColEI-origin, *tac*-promotor, coding for *laqI*$^q$*malE lacZ*$\alpha$; inserted *tgt* gene as 1.3 kb *Bam*HI/*Nco*I-fragment in pET9d | [Reuter & Ficner, 1995][68] |
| pET9d-ZM4-E235Q | Cm$^r$, Km$^r$; TGT-E235Q | this study |
| ptRNA2 | Amp$^r$; *E. coli* *tRNA*$^{Tyr}$ as *Bst*NI-fragment under control of T7-promotor in pTZ18U | [Curnow et al., 1993][98] |

.

Glycerol stocks were prepared by mixing 200 µL glycerol with 800 µL bacteria culture and stored at -80 °C.

## 5.2.4 Cloning techniques

QuikChange™ site-directed mutagenesis (Stratagene) was used to introduce the desired glutamine mutation of glutamate 235 into the wild-type TGT expression plasmid pET9d-ZM4 (Tab. 5.3). The template plasmid was expressed in 15 mL overnight culture of *E. coli* BL21(DE3) pLysS / pET9d-ZM4 (chapter 5.2.3) in LB-medium containing Cm and Km (chapter 5.2.1) and harvested with QIAprep Spin Miniprep Kit (Qiagen) following protocols of the vendor. For site-directed mutagenesis 10 ng of template plasmid were mixed with 1 µL dNTP-mix supplied by the vendor, 125 ng forward- and 125 ng reverse primer (Tab. 5.3), 5 µL 10x-Pfu-buffer, 1 µL PfuTurbo™ DNA-polymerase and 37µL aqua bidest, following the Stratagene protocol. The solution was subjected to 20 PCR cycles of the following temperature sequence: 93 °C for 0.5 min, 45 °C for 1 min, 72 °C for 14 min. To digest the template plasmid 0.1 unit of *Dpn*I was added. The mixture was kept for 1 h at 37°C. 1 µL of the PCR product was transformed into 100 µL of *E. coli* XL2blue cells. The mixed solutions were stored for 30 min on ice, subjected to a 42 °C heat shock for 0.5 min and cooled down for 2 min on ice. 2 mL of autoclaved LB-medium was added for cell recovery and the solution was stored for 1 h at 37 °C. The transformed cells were spread on agar plates containing Km and Cm (5.2.1) and incubated at 37 °C for 1-2 days. Individual colonies were transferred into 15 mL LB-medium containing Cm and Km and grown over night. The plasmid was isolated from 10mL over night culture (see above) and glycerol stocks were prepared and stored at -80 °C. Sequencing of the entire *tgt* gene (MWG, Ebersberg) confirmed the presence of the desired mutation as well as the absence of any further unwanted mutation. Subsequently, the mutated plasmid pET9d-ZM4-E235Q was transformed into *E. coli* BL21(DE3) pLysS cells. These cells were used for the preparation of the TGT enzyme (chapter 5.2.5).

**Tab. 5.3** PCR-primers for site-directed mutagenesis (mutation underlined)

| **primer** | **sequence** |
|---|---|
| E235Q-s | 5’-GGGGGATTGGCTGTGGGTCAAGGACAGGATGAAATG-3’ |
| E235Q-a | 5’-CATTTCATCCTGTCCTTGACCCACAGCCAATCCCCC-3’ |

## 5.2.5 Preparation of TGT

Preparation of TGT was performed following the method described by Romier et al. (1996)[100].

The following buffers were used:

| | |
|---|---|
| **extraction buffer** | 20 mM Tris-HCl pH 7.8, 10 mM EDTA, 1 mM DTT<br>per 50 mL 1 tablet of Complete™ (Roche) protease inhibitor cocktail |
| **TGT-buffer A** | 10 mM Tris-HCl pH 7.8, 1 mM EDTA, 1 mM DTT |
| **TGT-buffer B** | TGT-buffer A + 1.0 M NaCl |
| **TGT-buffer C** | TGT-buffer A + 1.0 M $(NH_4)_2SO_4$ |
| **high-salt-buffer** | TGT-buffer A + 2.0 M NaCl, 0.01% (w/v) $NaN_3$ |

A 6-L culture of *E. coli* BL21(DE3) pLysS / pET9d-ZM4 (5.2.3) was grown in LB-medium containing Km and Cm at 37°C with vigorous shaking to an $A_{600}$ of 0.8. Subsequently, the temperature was lowered to 14 °C and overexpression of the *Z. mobilis tgt* gene was induced by the addition of IPTG at a final concentration of 1 mM. After another 24 h at 14 °C, the cells were harvested by centrifugation (Beckman Coulter, JA10, 4.000 rpm, 10 min, 4°C).

The cell pellet was resuspended in 50 mL of extraction buffer and the cells were disrupted with a Branson sonifier during storage on ice. After two centrifugation steps (Beckman Coulter, JA25.50, 20.000 rpm, 45 min, 4 °C) the supernatant was drop wise diluted with a saturated $(NH_4)_2SO_4$ solution to a final concentration of 80%. After centrifugation (JA25.50, 25.000 rpm, 30 min, 4 °C) the protein pellet was resuspended with extraction buffer (without Complete™) and dialyzed over night against 5 L TGT-buffer A + 0.01% (w/v) $NaN_3$ (cutoff 10,000 kDa).

The enzyme was purified in two subsequent chromatographical steps using ÄKTA FPLC device (Amersham Pharmacia Biosciences). In the first step the dialyzed protein solution was loaded onto a Q-sepharose column (XK 26, bed volume: 53 mL, Amersham) and washed with TGT-buffer A. The enzyme was eluted with a linear gradient of 0 to 100 % TGT-buffer B with a flow rate of 4 mL/min. The TGT enzyme eluted at 300 mM NaCl. Fractions containing TGT were identified by SDS-PAGE, combined and $(NH_4)_2SO_4$ was added to a concentration of 1.0 M. In the second column chromatography step a phenyl-sepharose column (XK 16, bed volume:

16mL, Amersham) was equilibrated with TGT-buffer C. Subsequently, the TGT-containing fractions were loaded onto the column. The enzyme was eluted with a linear gradient of 0 to 100 % TGT-buffer A with a flow rate of 2 mL/min. The enzyme eluted at 500 mM $(NH_4)_2SO_4$.

The fractions containing TGT were identified by SDS-PAGE, combined and concentrated to ~ 3 mg/mL using Centriplus YM30 (Milipore). The concentrated fractions were dialyzed over night in 15 mL Slide-A-Lyzer™ cassettes (Pierce) against 5 L of TGT-buffer A + 0.01 % (w/v) $NaN_3$. Upon dialysis microcrystals of TGT appeared. The microcrystals were centrifuged (JA25.50, 4.000 rpm, 20 min, 4 °C) and redissolved in high-salt-buffer to the desired TGT concentrations. The enzyme was stored at -20 °C.

| | | | |
|---|---|---|---|
| TGT for crystallization: | 20 mg/mL | 30 µL | (chapter 5.5) |
| TGT for labelling-assay: | 4 mg/mL | 5 µL | (chapter 5.3) |
| TGT for SPR-assay: | 5 mg/mL | 5 µL | (chapter 5.4) |

## 5.2.6 Preparation of $tRNA^{Tyr}$

Preparation of *E. coli* $tRNA^{Tyr}$ (ECY2) *via* in vitro transcription was done following the method described by Curnow *et al*. (1993)[98].

The following buffers were used:

| | |
|---|---|
| **NaAc-buffer** | 3.0 M NaAc pH 5.3 |
| **tRNA-buffer A** | 10 mM HEPES pH 7.3 |
| **tRNA-buffer B** | tRNA-buffer A + 2 M NaCl |
| **tRNA-buffer C** | tRNA-buffer A + 1 mM $MgCl_2$ |
| **10x Tx-buffer** | 400 mM Tris-HCl pH 8.0, 200 mM $MgCl_2$, 50 mM DTT, 10 mM spermidine |

To inactivate nucleases glass materials used during the preparation were sterilized for 4h at 180 °C.

Frequently applied precipitation of DNA and tRNA solutions was performed as following: 2.5 volumes of ethanol p.a. and 0.1 volumes of NaAc-buffer were added. For precipitation this mixture was stored at -20°C. The precipitated nucleobases were centrifuged (Beckman Coulter, JA-25.50, 15.000 rpm, 30 min, 4 °C) and after decantation the pellet was dried at room temperature for 15 min.

To produce the template plasmid required for in vitro transcription a 2-L culture of *E. coli* TG2 / ptRNA2 (5.2.3) was grown in LB-medium containing Amp (chapter 5.2.1) at 37 °C for 19 h. The cells were harvested by centrifugation (Beckman Coulter, JA-10, 4.000 rpm, 10 min, 4 °C).

The plasmid was extracted using a QIAGEN Plasmid Mega Kit (QIAGEN, Hilden), following the protocol of the vendor. The plasmid containing solution was precipitated as described above, stored over night, centrifuged (Beckman Coulter, JA-25.50, 10.000 rpm, 30min, 4°C) and redissolved in 10 mL aqua bidest. After another precipitation step, following the procedure described at the beginning of this chapter the plasmid pellet was dissolved in water to a final concentration of 1 mg/mL (chapter 5.2.2).

Restriction digestion was performed using the endonuclease *Bst*NI. Aliquots of 250 µL plasmid solution (1 mg/ml) were incubated 5 µL *Bst*N I (10.000 U/ml ; New England Biolabs), 50 µL buffer #2 (New England Biolabs), 2.5 µL BSA (100 µg/ml ; New England Biolabs), 250 µL aqua bidest. and stored over night at 60°C. The digestion mixture was then extracted with equal volumes of Roti-Phenol (Roth) and choloroform:isoamylalcohole (24:1). The aqueous supernatant was extracted again with an equal volume of choloroform:isoamylalcohole (24:1). The aqueous supernatant of this extraction was precipitated and stored over night.

After centrifugation the dried pellet was dissolved in water to a final concentration of 0.6 mg/mL. In vitro transcription of ECY2 was carried out in 15 mL falcon tubes. Aliquots of 1 mL restricted template were incubated together with 600 µL 10x Tx-buffer, 240 µL ATP (100 mM), 240 µL CTP (100 mM), 240 µL UTP (100 mM), 240 µL GTP (100 mM), 25 µL inorganic PPase (Roche Diagnostics), 5 µL RNAse-inhibitor (10,000 U/mL, Roche Diagnostics), 6000 U T7 RNA Polymerase (Amersham Bioscience) ad 6mL aqua bidest. The mixture was allowed to incubate at 37 °C for approximately 7 h. Successful transcription was confirmed with agarose gel electrophoresis and staining with ethidium-bromide. The mixture was extracted in two steps following the method described above and precipitated at -20 °C over night.

The transcribed ECY2 was purified by column chromatography. Therefore, the centrifuged and dried pellet was dissolved in 10 mL tRNA-buffer A and loaded onto an anion-exchange column (SOURCE 15Q, HR 10/10, Pharmacia). After washing with tRNA-buffer A ECY2 was eluted with a linear gradient of 0 to 100% tRNA-buffer

B with a flow rate of 1.5 mL/min. The tRNA eluted at 450 mM NaCl. Fractions containing ECY2 were identified by gel electrophoresis, combined and precipitated.
After centrifugation (JA-25.50, 10.000rpm, 1 h, 4 °C) the pellet was dissolved in tRNA-buffer A to a final concentration of < 20 µM. For the following monomerization step ECY2 was transferred into a water bath and within 45 min the temperature was risen to 70 °C. For another 30 min the tRNA was stored at this temperature. After addition of $MgCl_2$ to a final concentration of 1 mM the tRNA was stored for 2 h on ice. The solution then was precipitated and stored for 1 h at -20 °C.
After centrifugation the pellet was dissolved in tRNA-buffer 3 to a final concentration of ~ 200 µM (chapter 5.2.2). Purity of the sample was confirmed photometrically (chapter 5.2.2) and aliquots were stored at -20 °C.

## 5.3 Kinetic parameters and inhibition constants

The method has been modified according to the initial procedures described by Reuter *et al*. (1994)[101] and Grädler *et al*. (2001)[3].

The following buffer was used:

**TGT assay buffer** 200 mM HEPES pH 7.3, 20 mM $MgCl_2$,
2.95 µM = 5 % CMC *Tween 20* (*Roth*)

### 5.3.1 Workflow of the labelling assay

The determination of TGT activity was carried out in 75 *µ*L mixtures containing 150 nM *Z. mobilis* TGT and variable concentrations of *E. coli* $tRNA^{Tyr}$, guanine (7.5% radio-labelled [8-$^3$H]-guanine (*Hartmann Analytik*)) and inhibitor in TGT assay buffer. Inhibitors that were added to the assay solution were, due to limited solubility in aqueous solutions, dissolved in DMSO. The final assay solution contained up to 5 % DMSO. Reactions were started by adding tRNA and guanine/[8-$^3$H]-guanine to the protein solution. Prior to addition of the substrate, protein and protein/inhibitor mixtures were preincubated at 37 °C for 10 min to adjust the solution to the assay temperature. The reacting mixture was kept at 37 °C, and 15 *µ*L aliquots were taken at intervals of 1 to 4 min. Aliquots were immediately transferred to glass fiber (GC-F) filters (*Whatman*) and quenched with 10 % (w/v) trichloroacetic acid at 0 °C for

15 min. Unbound guanine was washed from the filters in 7 min intervals twice with 5% (w/v) trichloroacetic acid (TCA) and twice with ethanol. Filters were dried at 60° for 45 min and then transferred in 4 mL of scintillation cocktail for lipophilic samples (Roth). Tritium incorporated into tRNA was quantified using liquid scintillation counting (Results from the aliquots were used to calculate initial velocity using *GraFit*[102].

## 5.3.2 Kinetic parameters

Michaelis-Menten parameters for tRNA and guanine were determined separately in triplicate and average values were calculated. Kinetic parameters for guanine were measured using 150 nM TGT, 15 µM tRNA, and variable concentrations of guanine (7.5 % [8-$^3$H]-guanine) in the range of 0.5 - 20 µM. Kinetic parameters for tRNA were measured using 150 nM TGT, 20 µM guanine/[8-$^3$H]-guanine, and variable concentrations of tRNA (0.25-15 µM). Initial velocities in *counts per minute* were transferred to [µM/min] using a calibration constant derived from liquid scintillation counting of guanine/[8-$^3$H]-guanine solutions with variable concentrations. Kinetic parameters were determined via double-reciprocal linearization using the method of Edie-Hofstee and linear regression using *GraFit*[102].

Michaelis-Menten parameters for $preQ_0$ and $preQ_1$ were calculated *via* monitoring of [8-$^3$H]-guanine excorporation from $tRNA^{Tyr}$ radio-labelled in position 34. To produce radioactively labelled tRNA 50 µM unmodified $tRNA^{Tyr}$ was incubated with 500 nM TGT and 10 µM [8-$^3$H]-guanine in TGT assay buffer for 1 h. TGT is extracted from the reaction mixture by the addition of equal volumes of Roti-Phenol (Roth) and choloroform:isoamylalcohole (24:1). The aqueous supernatant is once again extracted with an equal volume of choloroform:isoamylalcohole (24:1). The aqueous supernatant containing the radioactively labelled tRNA is separated from guanine / [8-$^3$H]-guanine *via* gel filtration using NAP-columns (GE Healthcare, Life Science) and TGT assay buffer. The labelled tRNA was dried by vacuum centrifugation and dissolved in tRNA-buffer C (chapter 5.2.6) to a final concentration of 200 µM. Kinetic parameters for $preQ_0$ and $preQ_1$ were measured using 150 nM TGT(w.t) / TGT(E235Q) and 15 µM labelled $tRNA^{Tyr}$. For $preQ_0$ the concentration was varied in a range of 0.5 - 15 µM. For $preQ_1$ variable concentrations in a range of 0.5 - 15 µM for TGT(w.t.) and 2 - 80 µM for TGT(E235Q) were applied. Initial velocities in *counts per minute* were calculated from the decreasing tritium labelling level of tRNA due to

the incorporation of the respective substrate bases. Initial velocities were transferred to [$\mu$M/min] using a calibration constant derived from liquid scintillation counting of guanine/[8-$^3$H]-guanine solutions with variable concentrations. Kinetic parameters were determined via double-reciprocal linearization using the method of Eadie-Hofstee and linear regression using *GraFit*[102].

## 5.3.3 Trapping experiment

To distinguish pure competitive from non-competitive inhibitors an initial trapping experiment followed by SDS-PAGE was performed with each inhibitor. 5 µM *Z. mobilis* TGT, 100 µM *E. coli* tRNA$^{Tyr}$, and 1 mM of the respective inhibitor (dissolved in DMSO) in 10 µL of 100 mM HEPES buffer, pH 7.3, 20 mM $MgCl_2$, and 5 mM dithiothreit were incubated for 1 h at 25°. A total of 10 µL SDS loading buffer was added and incubated for another 1 h at 25°. 10 µL of each sample were loaded onto a 15% SDS gel. After electrophoresis gels were stained with Coomassie blue.

## 5.3.4 Inhibition constants

### 5.3.4.1 Inhibition constants for pure competitive inhibition

The inhibition assay was performed using 150 nM TGT, 20 $\mu$M guanine/[8-$^3$H]-guanine, and tRNA at two concentrations (1 $\mu$M and 1.5 $\mu$M). Six reaction mixtures for each tRNA concentration were prepared. To five of them, inhibitor dissolved in DMSO (5% final volume) at variable concentrations was added. Initial velocities for the reaction mixtures were determined, and $K_{ic}$ determination was performed using *Dixon* plots. As $V_{max}$ determination is only possible with limited accuracy for independent measurements, the modified equation (5) published by Grädler *et al.* (2001)[3] was used to calculate $K_{ic}$.

Linear regression of data points derived from this equation with *GraFit*[102] resulted in a straight line with the slope $1/K_{ic}$.

$$\frac{v_0}{v_i} \cdot \frac{K_m + [S]}{K_m} = \frac{1}{K_{ic}} \cdot [I] + \left( \frac{[S]}{K_m} + 1 \right) \quad (5)$$

$v_0$ initial velocity at given [S] concentration in the absence of inhibitor, $v_i$ initial velocity at given [S] concentration in the presence of inhibitor, [S] tRNA concentration, [I] inhibitor concentration, $K_m$ Michaelis-Menten constant of tRNA, $K_{ic}$ competitive inhibition constant

#### 5.3.4.2 Inhibition constants for mixed inhibition

The inhibition assay was performed using 75 nM TGT, 20 $\mu$M guanine/[8-$^3$H]-guanine, and variable tRNA concentrations (0.25 - 15 $\mu$M). Kinetic parameters were determined once in the absence of inhibitor ($K_m$, $V_{max}$) and twice in the presence of specific inhibitor concentrations [I] to calculate ($K_m^{app}$, $V_{max}^{app}$) via double-reciprocal linearization and linear regression using *GraFit*[102]. Contributions of $K_{ic}$ and $K_{iu}$ towards non-competitive inhibition can be calculated from the following equations (2) and (3):

$$K_m^{app} = K_m \cdot \left( 1 + \frac{[I]}{K_{ic}} \right) \quad (2)$$

$$V_{max}^{app} = \frac{V_{max}}{\left( 1 + \frac{[I]}{K_{iu}} \right)} \quad (3)$$

$V_{max}$ maximum velocity of uninhibited reaction, $V_{max}^{app}$ apparent maximum velocity of inhibited reaction, $K_m$ Michaelis-Menten constant of tRNA, $K_m^{app}$ apparent Michaelis-Menten constant of tRNA in the presence of inhibitor, [I] inhibitor concentration, $K_{ic}$ competitive inhibition constant, $K_{iu}$ uncompetitive inhibition constant

Due to the elaborate $K_i$ determination procedure inhibition constants were determined only in duplicate or triplicate. Estimated standard deviations $\sigma$ were calculated from the range R of the resulting inhibition constants [Fischer & Hannappel, 1996][103].

$\sigma = R \cdot 0.896$ (for double measurements)

$\sigma = R \cdot 0.596$ (for triple measurements)

## 5.4 SPR-based assay

The SPR-spectroscopic experiments were performed using Jandratek device at 22 °C.

The following buffers were used:

| | |
|---|---|
| **Phosphate-buffer** | 120 mM NaCl, 30 mM $NaH_2PO_4$ / $(Na)_2HPO_4$ pH 7.3 |
| **SPR-buffer** | 98 % Phosphate buffer + 2 % high-salt-buffer (chapter 5.2.5) |

TGT for SPR-assay (chapter 5.2.5) was mixed with 245 µL of Phosphate-buffer to a final concentration of 100 µg/mL. Linker G1 (3.2.5) was dissolved in SPR-buffer to a final concentration of 20 µg/mL. Streptavidin was dissolved in SPR-buffer to a final concentration of 500 µg/mL.
A biotinylated gold chip was mounted into the test chamber and washed with 10 µL of SPR-buffer for 10 min. Then the buffer was removed. 10 µL of streptavidin solution were filled into the test chamber and incubated for 15 min. Unbound streptaviden was washed from the chip in five washing steps using 10 µL of SPR-buffer. 10 µL of linker G1 solution was filled into the test chamber and incubated for 15 min. Unbound linker was washed from the chip in 5 washing steps using 10 µL of SPR-buffer. 10 µL of TGT solution was filled into the test chamber. The solution in the chamber was stirred for 6 min and then incubated for another 10 minutes. Unbound TGT was washed from the chip in 5 washing steps using 10 µL of SPR-buffer. The resonance signal was permanently monitored and used to calculate the amount of bound protein.

**Protein binding** (empirical formula for Jandratek device):

**1 Resonance Unit ≡ 1 pg protein · $mm^{-2}$**

# 5.5 Crystal structure analysis

## 5.5.1 Growing of crystals

TGT crystals suitable for ligand soaking were produced in a two step procedure. Droplets were prepared by mixing 2 µL of concentrated protein solution (14 mg/mL TGT in high-salt-buffer (chapter 5.2.5)) with 2 µL reservoir solution of the respective seeding buffer (S-buffer). Micro-crystals were grown at 273 K using the hanging-drop, vapour diffusion method in the presence of 1 mL of reservoir solution of the respective seeding buffer. Micro-crystals of 0.05 $mm^3$ grew within two weeks.

| | |
|---|---|
| **S-buffer pH 5.5** | 100 mM morpholino ethylsulfonate (MES), pH 5.5, 1 mM DTT, 8 % (w/v) PEG 8.000, 10 % (v/v) DMSO |
| **S-buffer pH 6** | 100 mM MES, pH 6, 1 mM DTT, 8 % (w/v) PEG 8.000, 10 % (v/v) DMSO |
| **S-buffer pH 8.5** | 100 mM Tris, pH 8.5, 1 mM DTT, 13 % (w/v) PEG 8.000, 10 % (v/v) DMSO |

Subsequently macro-seeding was performed under similar conditions. Again droplets were prepared by mixing 2 µL of concentrated protein solution with 2 µL reservoir solution of the respective macro-seeding buffer (MS/CC-buffer). One micro-crystal was transferred into this solution. Single crystals with a size of approximately 0.7 x 0.7 x 0.2 $mm^3$ grow within two to four weeks per droplet. Small sized compounds were dissolved in DMSO and added to the droplet to a final concentration of 2 mM to allow soaking. Crystals were soaked at 293 K for one day.

| | |
|---|---|
| **MS/CC-buffer pH 5.5** | 100 mM MES, pH 5.5, 1 mM DTT, 5 % (w/v) PEG 8.000, 10 % (v/v) DMSO |
| **MS/CC-buffer pH 6** | 100 mM MES, pH 6, 1 mM DTT, 5 % (w/v) PEG 8.000, 10 % (v/v) DMSO |
| **MS/CC-buffer pH 8.5** | 100 mM Tris, pH 8.5, 1 mM DTT, 5 % (w/v) PEG 8.000, 10 % (v/v)DMSO |

Cocrystallization was performed under similar conditions. Droplets were prepared by mixing 2µL of concentrated protein solution with 2µL reservoir solution of the respective cocrystallization buffer (MS/CC-buffer). Additionally, small sized compounds were dissolved in DMSO and immediately added to the droplet to a final concentration of 2 mM. One micro-crystal was transferred into this solution. Single crystals with a size of approximately 0.7 x 0.7 x 0.2 $mm^3$ grow within two to four weeks per droplet.

## 5.5.2 Data collection

For data collection, crystals were cryoprotected using glycerol; 4 µL of the crystallization droplet were well mixed with 2.2 µL of glycerol resulting in a 35 % glycerol solution. The soaked crystal was transferred for 10 seconds into this solution and subsequently flash-frozen in liquid $N_2$. Data sets were collected at cryo conditions (100 K) with $CuK_\alpha$ radiation. (λ = 1.5418Å) using a Rigaku RU-300 rotating-anode generator at 50 kV and 90 mA equipped with either focusing mirrors (MSC, USA) and a R-AXIS IV + + image-plate system or with Xenocs focussing optics and a R-AXIS IV detector. All tested crystals exhibit monoclinic symmetry in space group *C*2 containing one monomer per asymmetric unit with Matthews coefficients of 2.3 - 2.4. One crystal soaked with **L4** showed the monoclinic space group *P*2 containing two symmetry independent monomers per asymmetric unit with a Matthews coefficients of 2.3. All data processing and scaling were performed using the HKL2000 package [Ottinowski & Minor, 1997][104]. For all refined structures unit cell dimensions for the crystals, data collection and processing statistics are given in chapter 6.2.

## 5.5.3 Structure determination and refinement

For TGT·**L4** (space group *P*2) molecular replacement with the coordinates of apo TGT crystallized at pH 5.5 (PDB-code: 1P0D) was performed using AMoRe [Navaza, 1994][105]. For *C*2 crystals grown at pH 5.5 or pH 6 coordinates of the apo TGT crystal structure grown at a pH of 5.5, (PDB-code: 1P0D) were directly applied for initial rigid-body refinement of the protein molecule followed by repeated cycles of

conjugate gradient energy minimization, simulated annealing and *B*-factor refinement using the CNS program package [Brunger *et al*., 1998][106]. For *C*2 crystals grown at pH 8.5 the coordinates of the apo TGT crystal structure grown at a pH of 8.5 (PDB-code: 1PUD) were applied. For TGT·**L1** and TGT·**L4** this program was also used in further refinement cycles to include the ligand and the water molecules to generate the final model. Refinement at the later stages for all other structures was performed with SHELXL [Sheldrick & Schneider, 1997][107]. Here, up to 20 cycles of conjugate gradient minimization were performed with default restraints on bonding geometry and *B*-values: 5 % of all data were used for $R_{free}$ calculation. Amino acid side-chains were fit to σA-weighted 2 $|F_o|$ - $|F_c|$ and $|F_o|$ - $|F_c|$ electron density maps using O [Jones *et al*., 1991][108]. Water and glycerol molecules as well as the ligand were located in the difference electron density and added to the model for further refinement cylces. During the last refinement cycles, riding H-atoms were introduced for the protein residues (not for ligand) without using additional parameters. All final models were validated using PROCHECK [Laskowski *et al*., 1993)[109]. Data refinement statistics are given in chapter 6.2.

## 5.6 Computational methods

### 5.6.1 Minimization

For some compounds crystal structure are not available. To suggest binding geometries the compounds were placed in an available TGT binding pocket and minimized with the MAB force field as implemented in MOLOC [Gerber & Müller, 1995][110]. For minimization the binding pocket was kept rigid. Only the side chains of Asn70, Asp102, Ser103, Tyr106, Asp156 and Gln203 were kept flexible to allow limited induced fit events.

The following binding pockets after removing ligand and water molecules were used for minimization: TGT·**H2** (PDB: 1F3E), TGT·preQ$_1$ (PDB: 1P0E ), TGT·**H7** (PDB: 1N2V), TGT· tRNA (PDB: 1Q2S) and TGT·**L1** (PDB: 2BBF).

### 5.6.2 Docking

Docking experiments were performed with GOLD 3.0.1. The crystal structure of TGT·**H2**, after removing the ligand and all water molecules, was used for docking. Applying standard settings in GOLD, 10 – 25 solutions were created in a region of 20Å in the vicinity of Arg286. Solutions were ranked according to the GOLDScore scoring function.

### 5.6.3 Alignment of structures

Structural alignments of dissimilar sequences (ArcTGT and QueTGT) were performed by means of the program suite SYBYL 7.0 (Tripos Inc., St. Louis, MO.). Alignment of structures with similar or identical sequences was performed with the alignment function implemented in Pymol (http://www.pymol.org).

# 6. Appendix

## 6.1 Structural alignment of 21 TGT sequences

21 TGT sequences were selected as representatives of major eubacterial and eukaryotic subdivisions. The sequences were retrieved from SWISS-PROT and aligned using CLUSTAL_W [69].

```
Escherichia             ------------MKFELDTTDGRARRGRLVFDRGVVETPCFMPVGTYGTVK-GMTPEEVE
Vibrio                  ----------MKLKFELKKKNGNARRGQLIFERGTVQTPAFMPVGTYGTVK-GMTPEEVK
Pseudomonas             ------------MNFELLATDGKARRGRLTFPRGVVETPAFMPVGTYGTVK-GMLPRDIE
Neisseria               -----------MLKFTLHKKDGYARRGTLELNHGKIETPVFMPVGTYGSVK-AMNPQNLH
Agrobacterium           --------MHEKFTFTLKSTSGGARLGEVAMPRGVIRTPAFMPVGTVGTVK-AMYLDQVR
Rickettsia              ---------MSKFSFNIHHQHKKARSGIIVTAHGEMRTPAFMPVGTRGTVK-AMLPESVA
Zymomonas               VEATAQETDRPRFSFSIAAREGKARTGTIEMKRGVIRTPAFMPVGTAATVK-ALKPETVR
Helicobacter            ------------MDFQLQATDNNARAGLLNLAHSQVATPVFMPVGTQGCIKSLDATDAQE
Gloeobacter             --------MTASFAFTIEHRDGEARAGTFATPHGPVYTPCFMPVGTQATVK-TLTPAQLA
Thermotoga              ------------MEFEVKKTFGKARLGVMKLHHGAVETPVFMPVGTNASVK-LLTPRDLE
Chlamydia_pneu.         --------MALKFHLIHQSKKSQARVGQIETSHGVIDTPAFVPVATHGALK-----GVID
Chlamydia_trac.         --------MALRFEILHQSKKSRARVGRIETAHGYIDTPAFVPVATNGALK-----GVLD
Treponema               -----MKEKKEIFTLLHQDAASPARTGVLELPHGKVLTPAFMPVGTAATVKAMTKDDLDE
Bacillus                -----MAEQPIRYEFIKECKQTGARLGKVHTPHGSFETPVFMPVGTLATVK-TMSPEELK
Lactobacillus           ------MEPAIKYRLIKKEKHTGARLGELITPHGTFPTPMFMPVGTQASVK-SLAPEELD
Clostridium             -----------MYTLIKKCGN--AKRGRFETPHGTIETPVFMNVGTLGVIKGAVSSMDLK
Schizosaccharomyces     ---MASSFPALQFKVVARCSTTRARVTDIQLPHGLVESPVFMPVGTQASLK-GVLPEQLD
Caenorhabditis          ------------MRYDVLARAGFARRGNLHLPHSIVETPVFMPVGTQGTMK-GIVPEQLV
Drosophila              --MGPSHIPPLTYKVVAECSVSKARAGLMTLRHSEVNTPVFMPVGTQGTLK-GIVPDQLI
Mus                     ------------MRLVAECSRSGARAGELRLPHGTVATPVFMPVGTQATMK-GITTEQLD
Homo                    ------------MRLVAECSRSRARAGELWLPHGTVATPVFMPVGTQATMK-GITTEQLD
                                           *:   .   :. . :* *: *.* . :*

                                    70     77             92       102  107
Escherichia             ATGAQIILGNTFHLWLRPGQEIMKLHGDLHDFMQWKGPILTDSGGFQVFSLGD-------
Vibrio                  ETGAQILLGNTFHLWLRPGQEVMKMHGDLHDFMNWQGPILTDSGGFQVFSLGD-------
Pseudomonas             DIGAQIILGNTFHLWLRPGTEVIQRHGDLHDFMQWKGPILTDSGGFQVFSLGA-------
Neisseria               DIKAQIILGNTYHLWLRPGLEVVEQFGGLHGFIGWDKPILTDSGGFQVFSLSD-------
Agrobacterium           ELGADIILGNTYHLMLRPGPERVARLGGLHELIRWPHPILTDSGGFQVMSLSG-------
Rickettsia              ETGADILLGNTYHLMLQPTAERIVQLGGLHKFMNWDKPILTDSGGFQVMSLSK-------
Zymomonas               ATGADIILGNTYHLMLRPGAERIAKLGGLHSFMGWDRPILTDSGGYQVMSLSS-------
Helicobacter            ILGAKLILANTYHMYLRPGEKVVEELGGLHRFAQFYGSFLTDSGGFQAFSLSD-------
Gloeobacter             ETGAQMILANTYHLSLQPGADIVAGAGGLHGFMQWPGPILTDSGGFQVFSLSS-------
Thermotoga              EAGAEIILSNTFHLMLKPGVEIIKLHRGLHNFMGWKRPILTDSGGFQVFSLPK-------
Chlamydia_pneu.         HSDIPLLFCNTYHLLLHPGPEAVAKLGGLHQFMGRQAPIITDSGGFQIFSLAYGSVAEEI
Chlamydia_trac.         HSNIPLMFCNTYHLIVHPGAEAIAAMGGLHQFIGRNAPIITDSGGFQIFSLAYGSVAEEI
Treponema               IG-FEIILANTYHLFLRPGIEVIKAAGGLHGFSDWKKNFLTDSGGFQVFSLSQ-------
Bacillus                AMDAGIILSNTYHLWLRPGQDIVKEAGGLHKFMNWDRAILTDSGGFQVFSLSK-------
Lactobacillus           AMGAGVILSNTYHLWLRPGEQIVKEAGGLHQFMNWKKGILTDSGGFQVFSLAK-------
Clostridium             EIGCQVELSNTYHLHLRPGDEVIKKMGGLHKFMNWDRPILTDSGGFQVFSLAK-------
Schizosaccharomyces     ALGCKIMLNNTYHLGLKPGQEVLDTVGGAHRFQSWNKNILTDSGGFQMVSLLK-------
Caenorhabditis          SMDCRILLCNTYHLGHRPGHERVKAAGGLHKMMNWNRSILTDSGGFQMVSLSK-------
Drosophila              ELNCQILLGNTYHLGLRPGIETLKKAGGLHKFMGWPRAILTDSGGFQMVSLLQ-------
Mus                     SLGCRICLGNTYHLGLRPGPELIRKAQGLHGFMNWPHNLLTDSGGFQMVSLFS-------
Homo                    ALGCRICLGNTYHLGLRPGPELIQKANGLHGFMNWPHNLLTDSGGFQMVSLVS-------
                             : : **:*:  :*  . :    . * :      ::*****:* .**
```

```
                                    125    132                    156
Escherichia             -----------IRKITEQGVHFRNPINGDPIFLDPEKSMEIQYDLGSDIVMIFDECTPYP
Vibrio                  -----------IRKITEEGVHFRNPVNGDKIFMDAEKSMEIQKDLGSDIVMIFDECTPYP
Pseudomonas             -----------MRKIKEEGVTFASPVDGAKVFMGPEESMAVQRALGSDIVMIFDECTPYP
Neisseria               -----------MRKLTEEGCTFKSPINGDKLFLSPEISMKIQTVLNSDIAMQLDECTPGE
Agrobacterium           -----------LRKLDEKGVTFKSHVDGSLHHMSPERSIEIQGMLDSDIQMQLDECIALP
Rickettsia              -----------LCKITEEGVSFSSHINGDKYMLTPERSTEIQYLLGSTITMAFDECTPYP
Zymomonas               -----------LTKQSEEGVTFKSHLDGSRHMLSPERSIEIQHLLGSDIVMAFDECTPYP
Helicobacter            -----------NVKLQEDGIVFKSHIDGSKHLFTPAKVLDIQYSLNSDIMMVLDDLVGLP
Gloeobacter             -----------LRTIDDDGVTFREPKSGALVRFTPEHAVAVQNALGADVIMAFDECPPYP
Thermotoga              -----------IR-IDDEGVVFRSPIDGSKVFLNPEISMEVQIALGSDICMVFDHCPVPD
Chlamydia_pneu.         KSCGKKKGMSSLVKITDEGAWFKSYRDGRKLFLSPELSVQAQKDLGADIIIPLDELLPFH
Chlamydia_trac.         KSCGKKKGGNTIIKVNDDGVHFKSYRDGRKLFLSPEISVQAQKDLGADIILPLDELLPFH
Treponema               -----------LRKITEEGVKFQSHIDGSRQFLSPEIAVELQTGFNSDIQMQLDICSSFG
Bacillus                -----------FRNIEEEGVHFRNHLNGDKLFLSPEKAMEIQNALGSDIMMAFDECPPYP
Lactobacillus           -----------NRDITEEGVHFKNHLNGSKMFLSPEKAIQIENDLGPDIMMSLDECPPFF
Clostridium             -----------IRKIQEEGVYFNSHIDGRRIFMGPEESMRIQSNIASTIAMAFDECIPNP
Schizosaccharomyces     -----------LATITEDGVTFLSPRDGTPMLLTPEHSISLQNSIGSDIMMQLDDVVHTL
Caenorhabditis          -----------LMTVDENGVNFESPHTGEMMALPPEKSIEIQQALGADIMMQLDHVIHVL
Drosophila              -----------LAEIDEHGVNFRSPFDNSQCMLTPEHSIEIQNAIGGDIMMQLDDVVKTT
Mus                     -----------LSEVTEEGVHFRSPYDGEETLLSPERSVEIQNALGSDIIMQLDHVVSST
Homo                    -----------LSEVTEEGVRFRSPYDGNETLLSPEKSVQIQNALGSDIIMQLDDVVSST
                                  :.*  * .   .    : .     :  :   : : :*


                                                              203
Escherichia             ADWDYAKRSMEMSLRWAKRSRERFDSLG-------NKNALFGIIQGSVYEDLRDISVKGL
Vibrio                  ATHDEAKKSMEMSLRWAKRSRDHFDKLE-------NPNNLFGIVQGGVYEDLRDVSVKGL
Pseudomonas             ADHDVAKRSMELSLRWAKRSK--IAHGD-------SPSALFGIVQGGMHEDLRLRSLDGL
Neisseria               ATREQARKSLQMSLRWAERSKKAFEDLK-------NPNALFGIVQGAMYEDLREESLKGL
Agrobacterium           AERKEIERAMEMSLRWAERCRVAFGEQ--------PGKAMFGIVQGGDQPDLRIRSAEGL
Rickettsia              ATFEEAKTSMQLTTRWANRSRNAFVKR--------EGYAQFGIIQGSVYEELREQSAKDL
Zymomonas               ATPSRAASSMERSMRWAKRSRDAFDSRKEQA----ENAALFGIQQGSVFENLRQQSADAL
Helicobacter            APLKRLEESIKRSAKWAN-MSLEYHKEKNR-----PSNNLFAIIQG--GTHLKMRSLSVG
Gloeobacter             ADREQVEGAVERTLRWFERCVEAHRRSD---------QALFGIVQGGVWPDLRRRCAEGL
Thermotoga              ADYEEVKEATERTYRWALRSKKAFKTEN---------QALFGIVQGGIYPDLRRESALQL
Chlamydia_pneu.         TDQEYFLTSCSRTYVWEKRSLEYHRKDP-------RHQSMYGVIHGGLDPEQRRIGVRFV
Chlamydia_trac.         ADPTYFHQSSQRTYVWEKRSLDYHLKNP-------GIQSMYGVIHGGTFPDQRKLGCKFV
Treponema               ISKTQTLADLKITMNWLDRAFAAWHNTPHE-----YDGALFPIVQGGFFEDLRLQSLEAI
Bacillus                AEYDYMKRSVERTSRWAERCLNAHNRQ--------DEQGLFGIVQGGEYEDLRTQSAKDL
Lactobacillus           ESYDYVSKSVARTSRWAERGLKVHQHP--------DYQGLFGIVQGAGFKDLREQSAKDL
Clostridium             STREYVENSVARTTRWLERCKKEMDRLNSLPDTINKKQMLFGINQGGTYEDIRKAHAKTI
Schizosaccharomyces     TESKRMEEAMYRSIRWLDRCIQAHKRPE--------TQNLFCIIQGGLDKRLREICCREM
Caenorhabditis          TTGDIVKEAMHRSIRWLDRCKVAHTR-D--------DQAMFPILQGGLNLELRKECAKEM
Drosophila              TTGPRVEEAMERTIRWVDRCIEAHARDD--------DQSLFPIVQGGLDVPLRQRCVSAL
Mus                     VTGPLVEEAMHRSVRWLDRCIAAHKHPD--------KQNLFAIIQGGLNADLRTTCLKEM
Homo                    VTGPRVEEAMYRSIRWLDRCIAAHQRPD--------KQNLFAIIQGGLDADLRATCLEEM
                                   :  *                      : : :*      :


                                230  235                    260
Escherichia             VDIG-FDGYAVGGLAVGEPKADMHRILEHVCPQIPADKPRYLMGVGKPEDLVEGVRRGID
Vibrio                  TEIG-FDGYAVGGLAVGEPKEDMHRVLEHTCPQLPEDKPRYLMGVGKPEDLVEGVRRGID
Pseudomonas             QEIG-FDGLAIGGLSVGEPKEEMIRVLDFLPPQMPADKPRYLMGVGKPEDLVEGVRRGVD
Neisseria               EELD-FPGLAIGGLSVGEPKPEMYRMLRAVGPILPEHKPHYLMGVGTPEDLVYGVAHGVD
Agrobacterium           KELD-LKGYAVGGLAVGEPQDVMLGMLDITLPVLPTEKPRYLMGVGTPDDILKSVARGID
Rickettsia              VELD-FEGYAIGGLAVGEGQELMFKVLDYAPEFLPQNKPRYLMGVGKPVDIIGAVSRGID
Zymomonas               AEIG-FDGYAVGGLAVGEGQDEMFRVLDFSVPMLPDDKPHYLMGVGKPDDIVGAVERGID
Helicobacter            LTHEGFDGYAIGGLAVGESADEMLETIAHTAPLLPKDKPRYLMGVGTPENILDAISLGVD
Gloeobacter             VAAD-LPGYAIGGVSVGEPQTLIERVVRVTAPLLPEHKPRYLMGVGTFREMAQAVAVGVD
Thermotoga              TSIG-FDGYAIGGLSIGEERSLTLEMTEVTVEFLPEDKPRYFMGGGSPELILELVDRGVD
Chlamydia_pneu.         EDEP-FDGSAIGGS-LGRNLQEMSEVVKITTSFLSKERPVHLLGIGDLPSIYAMVGFGID
Chlamydia_trac.         EDLP-FDGSAIGGS-LGKNLQDIVEVVGVTAANLSAERPRHLLGIGDLPSIWATVGFGID
Treponema               LKHE-PRGIAIGGLSIGEPKDLYQEYLSFTAKHIPKNKPLYVMGIGTPDYILEAVKNGVD
Bacillus                ISLD-FPGYAIGGLSVGEPKDVMNRVLEFTTPLLPKDKPRYLMGVGSPDALIDGAIRGVD
Lactobacillus           VSLD-FPGYSIGGLSVGESKAEMNHVLDFTTPLLPENKPRYLMGVGSADALIDGAIRGVD
Clostridium             VDMD-LDGYAIGGLAVGETHEEMYRVIDEVAPIFPDNKPLYLMGVGLPSNILEAVDRGVD
Schizosaccharomyces     VKRN-TPGIAVGGLSGGEEKHAFCETVYTCTSILPDNKPRYLMGVGYAEDLVVCVALGMD
Caenorhabditis          AKRA-KVGIAIGGLSGGEEKDHFWRVVAACCAALPPHLPRYVMGVGFPVDLVICSFLGAD
Drosophila              MERQ-VRGFAVGGLSGGESKHDFWRMVDVCTGYLPKDKPRYLMGVGFAADLVVCVALGID
Mus                     TKRD-VPGFAIGGLSGGESKAQFWKMVALSTSMLPKDKPRYLMGVGYATDLVVCVALGCD
Homo                    TKRD-VPGFAIGGLSGGESKSQFWRMVALSTSRLPKDKPRYLMGVGYATDLVVCVALGCD
                               * ::**   *.              :. . * :.:* *    :      * *
```

```
                          280                                     318  323
Escherichia               MFDCVMPTRNARNGHLFVTDG----VVKIRNAKYKSDTGPLDPECDCYTCRNYSR-----
Vibrio                    MFDCVMPTRNARNGHLFVTGG----VIKIRNAAHKTDTTPLDPHCDCYTCKNYSK-----
Pseudomonas               MFDCVMPTRNARNGHLFVDSG----VIKIRNSVHKHDDSTLDPTCDCYTCKHFSR-----
Neisseria                 MFDCVMPTRNARNGWLFTRFG----DLKIKNAKHKLDKRPIDESCTCYACQNFSR-----
Agrobacterium             MFDCVMPTRSGRHGLAFTRRG----RVNIRNARHAEDMRPLDEQSNCPASRDYSR-----
Rickettsia                MFDCVIPTRSGRNGQAFTKYG----TVNIRNSKYADDNKPLEHDCLCPACRNYSK-----
Zymomonas                 MFDCVLPTRSGRNGQAFTWDG----PINIRNARFSEDLKPLDSECHCAVCQKWSR-----
Helicobacter              MFDCVMPTRNARNATLFTHSG----KISIKNAPYKLDNTPIEENCACYACKRYSK-----
Gloeobacter               LFDCVMPTRVARHGSALLLGTGGDRRINLKNAQFRRDYEPLDCVCPCYTCRHFSR-----
Thermotoga                MFDSVFPTRIARHGTALTWNG----KLNLKASYNKRSLEPVDERCGCYTCKNFTR-----
Chlamydia_pneu.           SFDSSYPTKAARHGLILSKAG----PIKIGQQKYSQDSSTIDPSCSCLTCLSGISR----
Chlamydia_trac.           SFDSSYPTKAARHGMILTSQG----PLKINNQRYSSDLNPIEPGCSCLACSQGITR----
Treponema                 IFDCVLPSRNARNGNLFTHEG----AISIKRKEYEFDFNPIDSQCKCKVCRQYTR-----
Bacillus                  MFDCVLPTRIARNGTVFTAEG----RLNMKNAKFERDFRPIDEECDCYTCKNYTR-----
Lactobacillus             MFDCVLPTRIARNGTCMTSHG----RLVVKNAAYAHDFTPLDDNCDCYTCRNFTR-----
Clostridium               FFDCVLPARNGRHGHVFTKYG----KINLMNAKFELDGNPIDEGCECPACKHYSR-----
Schizosaccharomyces       MFDCVYPTRTARFGNALTRKG----VINLRNQKFRNDIGPLEEGCSCPCCKTELEGGWGI
Caenorhabditis            MFDCVYPTRTARFGTAMVRRGG---LMQLNQKRYKEDFLPIDKKCECNTCKNYTR-----
Drosophila                MFDCVFPTRTARFGCALVDSG----QLNLKQPKYKLDMEPIDKDCDCSTCRRYTR-----
Mus                       MFDCVYPTRTARFGSALVPTG----NLQLKKKQYAKDFSPINPECPCPTCQTHSR-----
Homo                      MFDCVFPTRTARFGSALVPTG----NLQLRKKVFEKDFGPIDPECTCPTCQKHSR-----

                            330334   339   345 349
Escherichia               --AYLHHLDRCNEILGARLNTIHNLRYYQRLMAGLRKAIEEGKLESFVTDFYQRQGREVPP
Vibrio                    --SYLHHLDRCNEILGARLNTIHNLRYYQRLMESIRKAIDEDRFDQFVAEFYARRNREVPP
Pseudomonas               --AYLHHLDKCGEMLGSMLNTIHNLRHYQRVMAGLREAIQQGTLAAFVDAFYAKRGLPTPP
Neisseria                 --AYLHHLHRAGEILGAQLNTIHNLHFYQVIMAEMREAVEQGKFADWQAQFHENRARGTD-
Agrobacterium             --AYLHHLTRSNEALGGMLLSWHNLAYYQELMQGIRTSIEEGRFADFYAETIEMWARGDID
Rickettsia                --AYLHHLVRIGEILGSMLMTWHNLTYFQNLMSRIRAYIKLGKDFDFDS------------
Zymomonas                 --AYIHHLIRAGEILGAMLMTEHNIAFYQQLMQKIRDSISEGRFSQFAQDFRARYFARNS-
Helicobacter              --AYLHHLFRAKELTYARLASLHNLHFYLELVKNARNAILEKRFLSFKKEFLEKYNSRSH-
Gloeobacter               --AYLAHLVRSEEILAMTLLSIHNVATLTRFAALLRCAIATGSFAQEFAHYLQSGPEPVLS
Thermotoga                --SYIHHLFDRGEVLGQILLTIHNINFMISLMKEVRRSIESGTFKELKSKVVEVYSSGGVN
Chlamydia_pneu.           --AYLRHLFKVREPNAAIWASIHNLHHMQQVMKEIREAILKDEI----------------
Chlamydia_trac.           --AYLRHLFKVHEPNAGIWASIHNMHHMQKVMREIREGILNDRI----------------
Treponema                 --AYLRHLFRTKEILYSMLATYHNLAFLYSMVQDIREAIQNDSFNDYYKNFLKKYENRLD-
Bacillus                  --AYIRHLIRCNETFGLRLTTYHNLHFLLHLMEQVRQAIREDRLGDFREEFFERYGYNKPN
Lactobacillus             --AYIRHLIKADETFGLRLTSYHNLYFLLHLMKQVRQAIMDDNLLEFRQNFFEMYGFNDKN
Clostridium               --AYIRHLFKAKEMLAMRLCVLHNLYFYNKLMEDIRKAIEGDYFKEFKEEKLHNWSGKA--
Schizosaccharomyces       TRAYF-NSLVSKETVGANLMTIHNVHFQLQLMRDMRESIIKDEFPSFVKNFFHEWNHGDKS
Caenorhabditis            --AYI-HSIVGKETVGCHLVSVHNIKHQLDLMRDVRQAIQSNSVEQFLKQFLYDYYGPIQS
Drosophila                --SYLHHIATN-ESVSSSLLSIHNVAYQLRLMRSMREAIQRDEFPQFVA-DFMARHF-KAE
Mus                       --AFLHALLHSDNTTALHHLTVHNIAYQLQLLSAVRSSILEQRFPDFVR-NFMRTMYGDHS
Homo                      --AFLHALLHSDNTAALHHLTVHNIAYQLQLMSAVRTSIVEKRFPDFVR-DFMGAMYGDPT

Escherichia               LNVD----------------------------------------------------------
Vibrio                    LQKDKA--------------------------------------------------------
Pseudomonas               LDA-----------------------------------------------------------
Neisseria                 ----------------------------------------------------------
Agrobacterium             PV----------------------------------------------------------
Rickettsia                ----------------------------------------------------------
Zymomonas                 ----------------------------------------------------------
Helicobacter              ----------------------------------------------------------
Gloeobacter               N----------------------------------------------------------
Thermotoga                V----------------------------------------------------------
Chlamydia_pneu.           ----------------------------------------------------------
Chlamydia_trac.           ----------------------------------------------------------
Treponema                 ----------------------------------------------------------
Bacillus                  AKSF---------------------------------------------------------
Lactobacillus             PKNF---------------------------------------------------------
Clostridium               ----------------------------------------------------------
Schizosaccharomyces       N------------YPSWAVDALRMVNIDLLA------------------------------
Caenorhabditis            ENPSKQDSEKMREVPQWVRDAVDHMGYKLDF------------------------------
Drosophila                P------------VPAWIREALSAVNIQLPADPERIDEQDQKPKTEKRRETEDVAEEQVA
Mus                       L------------CPAWAVEALASVGIMLT------------------------------
Homo                      L------------CPTWATDALASVGITLG------------------------------
```

# 6.2 Crystal data

## 6.2.1 Crystal data for TGT(E235Q)

| **Crystal data** | **E235Q** pH 6.0 | **E235Q** pH 8.5 | **E235Q**· **preQ$_0$** (pH 6) |
|---|---|---|---|
| A. *Data collection and processing* | | | |
| No. crystals used | 1 | 1 | 1 |
| Wavelength (Å) | 1.5418 | 1.5418 | 1.5418 |
| Space group | *C*2 | *C*2 | *C*2 |
| Unit cell parameters | | | |
| *a* (Å) | 90.66 | 90.76 | 90.45 |
| *b* (Å) | 65.22 | 65.18 | 65.22 |
| *c* (Å) | 70.38 | 70.46 | 70.44 |
| β (deg.) | 96.28 | 96.12 | 96.33 |
| B. *Diffraction data* | | | |
| Resolution range (Å) | 20–1.57 (1.60–1.57) | 20–1.55 (1.58–1.55) | 20–1.7 (1.73–1.70) |
| Unique reflections | 51,910 | 56,732 | 44,097 |
| $R(I)_{sym}$ (%) | 3.9 (16.3) | 5.8 (19.1) | 7.5 (42.8) |
| Completeness (%) | 91.1 (71.0) | 95.6 (78.5) | 98.3 (96.7) |
| Redundancy | 2.5 (2.1) | 2.3 (1.6) | 2.8 (2.6) |
| I/σ(I) | 22.4 (3.6) | 15.6 (3.8) | 12.3 (2.3) |
| C. *Refinement* | | | |
| Program used for refinement | SHELXL | SHELXL | SHELXL |
| Resolution range (Å) | 20–1.57 | 20–1.55 | 10–1.7 |
| Reflections used in refinement | 51,482/2,587 | 56,092/2,843 | 42,493/2,125 |
| Final *R* values | | | |
| $R_{free}$ (%) | 20.2 | 21.4 | 24.4 |
| $R_{work}$ (%) | 16.5 | 17.4 | 19.6 |
| No. of atoms (non-hydrogen) | | | |
| Protein atoms | 2,871 | 2,883 | 2,883 |
| Water molecules | 292 | 304 | 246 |
| Ligand atoms | --- | --- | 13 |
| RMSD, angle (deg.) | 2.4 | 2.4 | 2.3 |
| RMSD, bond (Å) | 0.011 | 0.011 | 0.008 |
| Ramachandran plot | | | |
| most favoured regions (%) | 94.9 | 95.3 | 92.5 |
| additionally allowed regions (%) | 4.8 | 4.0 | 6.5 |
| generously allowed regions (%) | 0.3 | 0.6 | 0.9 |
| Mean *B*-factors (Å$^2$) | | | |
| Protein atoms | 23.1 | 21.6 | 23.4 |
| Water molecules | 30.2 | 29.9 | 28.9 |
| Ligand atoms | --- | --- | 27.6 |

## 6.2.2 Crystal data for 6-amino-quinazolinones

| **Crystal data** | **TGT·Q19** (pH 5.5) | **TGT·Q21** (pH 5.5) |
|---|---|---|
| A. *Data collection and processing* | | |
| No. crystals used | 1 | 1 |
| Wavelength (Å) | 1.5418 | 1.5418 |
| Space group | *C*2 | *C*2 |
| Unit cell parameters | | |
| *a* (Å) | 88.09 | 88.64 |
| *b* (Å) | 63.44 | 64.03 |
| *c* (Å) | 71.12 | 70.70 |
| β (deg.) | 93,297 | 93.00 |
| B. *Diffraction data* | | |
| Resolution range (Å) | 20–2.15 (2.20–2.15) | 20–2.0 (2.05–2.00) |
| Unique reflections | 20,948 | 25,282 |
| $R(I)_{sym}$ (%) | 3.9 (47.6) | 5.5 (44.8) |
| Completeness (%) | 97.0 (99.5) | 94.2 (97.5) |
| Redundancy | 2.5 (2.4) | 2.5 (2.3) |
| I/σ(I) | 19.2 (2.3) | 15.0 (1.8) |
| C. *Refinement* | | |
| Program used for refinement | SHELXL | SHELXL |
| Resolution range (Å) | 10–2.15 | 10–2.0 |
| Reflections used in refinement | 18,987/918 | 23,226/1,123 |
| Final *R* values | | |
| $R_{free}$ (%) | 35.0 | 30.3 |
| $R_{work}$ (%) | 23.8 | 21.5 |
| No. of atoms (non-hydrogen) | | |
| Protein atoms | 2,804 | 2,773 |
| Water molecules | 47 | 73 |
| Ligand atoms | 15 | 21 |
| RMSD, angle (deg.) | 1.6 | 1.9 |
| RMSD, bond (Å) | 0.005 | 0.016 |
| Ramachandran plot | | |
| most favoured regions (%) | 90.2 | 92.7 |
| additionally allowed regions (%) | 8.9 | 7.3 |
| generously allowed regions (%) | 1.0 | 0.0 |
| Mean *B*-factors ($Å^2$) | | |
| Protein atoms | 58.4 | 43.5 |
| Water molecules | 51.7 | 39.8 |
| Ligand atoms | 60.6 | 67.9(55.3 ;93.1) |

## 6.2.3 Crystal data for *lin*-benzoguanines

| **Crystal data** | **TGT·L1**(pH 5.5) | **TGT·L2**(pH 5.5) | **TGT·L3**(pH 5.5) | **TGT·L4**(pH 5.5) |
|---|---|---|---|---|
| A. *Data collection and processing* | | | | |
| No. crystals used | 1 | 1 | 1 | 1 |
| Wavelength (Å) | 1.5418 | 1.5418 | 1.5418 | 1.5418 |
| Space group | *C*2 | *C*2 | *C*2 | *P*2 |
| Unit cell parameters | | | | |
| *a* (Å) | 91.20 | 90.59 | 90.92 | 71.11 |
| *b* (Å) | 65.59 | 64.93 | 65.25 | 64.21 |
| *c* (Å) | 69.81 | 70.12 | 69.97 | 88.08 |
| β (deg.) | 96.07 | 95.83 | 95.98 | 95.11 |
| B. *Diffraction data* | | | | |
| Resolution range (Å) | 20–1.7 (1.76–1.70) | 20–1.58 (1.61–1.58) | 20–1.58 (1.61–1.58) | 20–2.1 (2.14–2.10) |
| Unique reflections | 44,525 | 53,824 | 53,930 | 44,873 |
| $R(I)_{sym}$ (%) | 4.3 (16.7) | 6.9 (39.4) | 4.2 (22.1) | 5.9 (35.7) |
| Completeness (%) | 98.7 (99.3) | 97.1 (78.4) | 96.4 (69.3) | 96.4(98.5 ) |
| Redundancy | 2.3 (1.7) | 2.2 (1.6) | 2.4 (1.8) | 2.4 (2.1) |
| I/σ(I) | 20.6 (4.2) | 12.4 (1.7) | 21.4 (3.6) | 13.4 (2.4) |
| C. *Refinement* | | | | |
| Program used for refinement | CNS | SHELXL | SHELXL | CNS |
| Resolution range (Å) | 20–1.7 | 10–1.58 | 20–1.58 | 20–2.1 |
| Reflections used in refinement | 44,059/4450 | 51,398/2624 | 52,915/2681 | 42,047/2118 |
| Final *R* values | | | | |
| $R_{free}$ (%) | 22.0 | 21.2 | 19.6 | 25.5 |
| $R_{work}$ (%) | 19.6 | 17.0 | 15.9 | 22.1 |
| No. of atoms (non-hydrogen) | | | | |
| Protein atoms | 2870 | 2880 | 2823 | 5568 |
| Water molecules | 368 | 275 | 333 | 243 |
| Ligand atoms | 15 | 23 | 24 | 25 |
| RMSD, angle (deg.) | 1.2 | 2.3 | 2.4 | 1.2 |
| RMSD, bond (Å) | 0.005 | 0.009 | 0.011 | 0.006 |
| Ramachandran plot | | | | |
| most favoured regions (%) | 94.1 | 95.8 | 95.7 | 92.8 |
| additionally allowed regions (%) | 5.6 | 3.9 | 4.0 | 6.9 |
| generously allowed regions (%) | 0.3 | 0.3 | 0.3 | 0.3 |
| Mean *B*-factors ($Å^2$) | | | | |
| Protein atoms | 20.6 | 20.7 | 21.0 | 32.6 |
| Water molecules | 34.1 | 31.0 | 31.4 | 34.6 |
| Ligand atoms | 20.2 | 18.7 | 29.4 | 37.9 |
| PDB-code | 2BBF | 1Y5V | 1Y5W | 1Y5X |

## 6.2.4 Crystal data for 2-amino-*lin*-benzoguanines

| **Crystal data** | **TGT·L12** (pH 5.5) |
|---|---|
| A. *Data collection and processing* | |
| No. crystals used | 1 |
| Wavelength (Å) | 1.5418 |
| Space group | *C*2 |
| Unit cell parameters | |
| *a* (Å) | 89.87 |
| *b* (Å) | 64.97 |
| *c* (Å) | 70.94 |
| β (deg.) | 93.20 |
| B. *Diffraction data* | |
| Resolution range (Å) | 20–1.95 (2.00–1.95) |
| Unique reflections | 29,427 |
| $R(I)_{sym}$ (%) | 8.2 (43.5) |
| Completeness (%) | 98.8 (97.8) |
| Redundancy | 2.4 (2.4) |
| I/σ(I) | 10.9 (2.0) |
| C. *Refinement* | |
| Program used for refinement | SHELXL |
| Resolution range (Å) | 10–1.95 |
| Reflections used in refinement | 27,447/1,336 |
| Final *R* values | |
| $R_{free}$ (%) | 25.9 |
| $R_{work}$ (%) | 19.8 |
| No. of atoms (non-hydrogen) | |
| Protein atoms | 2,796 |
| Water molecules | 187 |
| Ligand atoms | 27 |
| RMSD, angle (deg.) | 2.0 |
| RMSD, bond (Å) | 0.006 |
| Ramachandran plot | |
| most favoured regions (%) | 94.4 |
| additionally allowed regions (%) | 5.3 |
| generously allowed regions (%) | 0.3 |
| Mean *B*-factors ($Å^2$) | |
| Protein atoms | 28.2 |
| Water molecules | 32.4 |
| Ligand atoms | 59.5(40.6 ;86.9) |

# 6.3 Nomenclature for amino acids

| amino acid | three letter code | one letter code |
|---|---|---|
| alanine | Ala | A |
| arginine | Arg | R |
| asparagine | Asn | N |
| aspartic acid (aspartate) | Asp | D |
| cysteine | Cys | C |
| glutamine | Gln | Q |
| glutamic acid (glutamate) | Glu | E |
| glycine | Gly | G |
| histidine | His | H |
| isoleucine | Ile | I |
| leucine | Leu | L |
| lysine | Lys | K |
| methionine | Met | M |
| phenylalanine | Phe | F |
| proline | Pro | P |
| serine | Ser | S |
| threonine | Thr | T |
| tryptophan | Trp | W |
| tyrosine | Tyr | Y |
| valine | Val | V |

# Literature


1. **Brenk R, Meyer EA, Reuter K, Stubbs MT, Garcia GA, Diederich F and 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.
2. **Brenk R, Naerum L, Gradler U, Gerber HD, Garcia GA, Reuter K, Stubbs MT and Klebe G.** (2003). Virtual screening for submicromolar leads of tRNA-guanine transglycosylase based on a new unexpected binding mode detected by crystal structure analysis. *J Med Chem*. 46, 1133-43.
3. **Grädler U, Gerber HD, Goodenough-Lashua DM, Garcia GA, Ficner R, Reuter K, Stubbs MT and Klebe G.** (2001). A New Target for Shigellosis: Rational Design and Crystallographic Studies of Inhibitors of tRNA-guanine Transglycosylase. *J Mol Biol*. 306, 455-467.
4. **Xie W, Liu X and Huang RH.** (2003). Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate. *Nat Struct Biol*. 10, 781-788.
5. **Sansonetti PJ.** (2001). Rupture, invasion and inflammatory destruction of the intestinal barrier by Shigella, making sense of prokaryote-eukaryote cross-talks. *FEMS Microbiology Reviews*. 25, 3-14.
6. **Mathan MM and Mathan VI.** (1991). Morphology of rectal mucosa of patients with shigellosis. *Rev. Infect. Dis.* 13, S314-S318.
7. **Bennish ML.** (1991). Potentially lethal complications of shigellosis. *Rev. Infect. Dis.* 13, S319-S324.
8. **Kotloff K, Winickoff J, Ivanoff B, Clemens J, Swerdlow D, Sansonetti P, Adak G and Levine M.** (1999). Global burden of Shigella infections: implications for vaccine development and implementation of control strategies. *Bulletin of the World Health Organization (WHO Bull.)*. 77, 651-666.
9. **Cohen D, Green M, Block C, Slepon R, Ambar R, Wassermann SS and Levine MM.** (1991). Reduction of transmission of shigellosis by control of houseflies (Musca domestica). *Lancet*. 337, 993-997.
10. **Jennison AV and Verma NK.** (2004). *Shigella flexneri* infection: pathogenesis and vaccine development. *FEMS Microbiology Reviews*. 28, 43–58.
11. **Heikkila E.** (1990). Increase of trimethoprim resistance among *Shigella* species, 1975-1988: analysis of resistance mechanism. *J. of Infectious Diseases*. 161, 1242-1248.
12. **Ashkenazi S, Levy I, Kazaronovski V and Samra Z.** (2003). Growing antimicrobial resistance of *Shigella* isolates. *J Antimicrob. Chemother.* 51, 427–429.
13. **Goma Epidemiology Group.** (1995). Public health impact of Rwandan refugee crisis: what happened in Goma, Zaire, in July, 1994? *Lancet*. 345, 339-344.
14. **Sansonetti PJ.** (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-G323.
15. **O'Loughlin EV and Robins-Browne RM.** (2001). Effect of Shiga toxin and Shiga-like toxins on eukaryotic cells. *Microbes and Infection*. 3, 493-507.
16. **Escobar-Páramo P, Clermont O, Blanc-Potard AB, Bui H, Le Bouguénec C and Denamur E.** (2004). A Specific Genetic Background Is Required for Acquisition and Expression of Virulence Factors in *Escherichia coli*. *Molecular Biology and Evolution*. 21, 1085-1094.

17. **Fernandez MI and Sansonetti PJ.** (2003). *Shigella* interaction with intestinal epithelial cells determines the innate immune response in shigellosis. *Int. J. Med. Microbiol.* 293, 55-67.
18. **Van Nhieu GT, Bourdet-Sicard R, Duménil G, Blocker A and Sansonetti PJ.** (2000). Bacterial signals and cell responses during *Shigella* entry into epithelial cells. *Cellular Microbiology*. 2, 187-193.
19. **Dorman CJ and Porter ME.** (1998). The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. *Molecular Microbiology*. 29, 677-684.
20. **Durand JM and Bjork GR.** (2003). Putrescine or a combination of methionine and arginine restores virulence gene expression in a tRNA modification-deficient mutant of *Shigella flexneri*: a possible role in adaptation of virulence. *Mol Microbiol*. 47, 519-527.
21. **Durand JM, Dagberg B, Uhlin BE and Bjork GR.** (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.
22. **Björk G.** (1996). Stable RNA modification. In: Neidhardt FC, Curtiss I, R., Ingraham JL, Lin CCC, Low J, K.B., Magasanik Bea, eds. *Escherichia coli and Salmonella: Cellular and Molecular Biology*. 2 ed. Washington, DC: American Society for Microbiology Press, 861-886.
23. **Durand JM, Bjork GR, Kuwae A, Yoshikawa M and Sasakawa C.** (1997). The modified nucleoside 2-methylthio-N6-isopentenyladenosine in tRNA of *Shigella flexneri* is required for expression of virulence genes. *J Bacteriol*. 179, 5777-5782.
24. **Durand JM, Okada N, Tobe T, Watarai M, Fukuda I, Suzuki T, Nakata N, Komatsu K, Yoshikawa M and Sasakawa C.** (1994). vacC, a virulence-associated chromosomal locus of *Shigella flexneri*, is homologous to tgt, a gene encoding tRNA-guanine transglycosylase (Tgt) of *Escherichia coli* K-12. *J Bacteriol*. 176, 4627-4634.
25. **Romier C, Reuter K, Suck D and Ficner R.** (1996). Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange. *Embo J*. 15, 2850-2857.
26. **Björk GR.** (1995). Biosynthesis and function of modified nucleosides in tRNA. In: Söll D, Rajbhandary UL, eds. *tRNA: Structure, Biosynthesis, and Function*. Washington, DC: American Society for Microbiology Press, 165-205.
27. **Björk GR, Durand JM, Hagervall TG, Leipuviene R, Lundgren HK, Nilsson K, Chen P, Qian Q and Urbonavicius J.** (1999). Transfer RNA modification: influence on translational frameshifting and metabolism. *FEBS Lett*. 452, 47-51.
28. **Iwata-Reuyl D.** (2003). Biosynthesis of the 7-deazaguanosine hypermodified nucleosides of transfer RNA. *Bioorg Chem*. 31, 24-43.
29. **Reader JS, Metzgar D, Schimmel P and de Crécy-Lagard V.** (2004). Identification of Four Genes Necessary for Biosynthesis of the Modified Nucleoside Queuosine. *J Biol Chem*. 279, 6280-6285.
30. **Okada N and Nishimura S.** (1979). Isolation and characterization of a guanine insertion enzyme, a specific tRNA transglycosylase, from *Escherichia coli*. *J Biol Chem*. 254, 3061-3066.
31. **Okada N, Noguchi S, Kasai H, Shindo-Okada N, Ohgi T, Goto T and Nishimura S.** (1979). Novel Mechanism of Post-transcriptional Modification of tRNA. *J Biol Chem*. 254, 3067-3073

32. **Noguchi S, Nishimura Y, Hiroto Y and Nishimura S.** (1982). *J. Biol. Chem.* 257, 6544-6550.
33. **Marks T and Farkas WR.** (1997). *Biochem. Biophys. Res. Commun.* 2, 233-237.
34. **Nakanishi S, Ueda T, Hori H, Yamazaki N, Okada N and Watanabe K.** (1994). A UGU sequence in the anticodon loop is a minimum requirement for recognition by *Escherichia coli* tRNA-guanine transglycosylase. *J Biol Chem*. 269, 32221-32225.
35. **Curnow AW and Garcia GA.** (1995). tRNA-guanine transglycosylase from *Escherichia coli.* Minimal tRNA structure and sequence requirements for recognition. *J Biol Chem*. 270, 17264-17267.
36. **Kuchino Y, Kasai H, Nihel S and Nishimura S.** (1976). Biosynthesis of the modified nucleoside Q in transfer RNA. *Nucleic Acids Res.* 3, 393-398.
37. **Gaur R and Varshney U.** (2005). Genetic Analysis Identifies a Function for the *queC* (*ybaX*) Gene Product at an Initial Step in the Queuosine Biosynthetic Pathway in *Escherichia coli*. *J Bacteriol*. 187, 6893-6901.
38. **Van Lanen SG, Reader JS, Swairjo MA, de Crécy-Lagard V, Lee B and Iwata-Reuyl D.** (2005). From cyclohydrolase to oxidoreductase: Discovery of nitrile reductase activity in a common fold. *Proc Natl Acad Sci U S A*. 102, 4264–4269.
39. **Hoops GC, Townsend LB and Garcia GA.** (1995). tRNA-guanine transglycosylase from Escherichia coli: structure-activity studies investigating the role of the aminomethyl substituent of the heterocyclic substrate PreQ1. *Biochemistry*. 34, 15381-7.
40. **Van Lanen SG, D. KS, Matthieu S, Link T, Culp J and Iwata-Reuyl D.** (2003). tRNA Modification by S-Adenosylmethionine:tRNA Ribosyltransferase-Isomerase. *J Biol Chem*. 278, 10491–10499.
41. **Campanacci V, Dubois DY, Becker HD, Kern D, Spinelli S, Valencia C, Pagot F, Salomoni A, Grisel S, Vincentelli R, Bignon C, Lapointe J, Giegé R and Cambillau C.** (2004). The *Escherichia coli* YadB Gene Product Reveals a Novel Aminoacyl-tRNA Synthetase Like Activity. *J Mol Biol*. 337, 273–283.
42. **Blaise M, Becker HD, Keith G, Cambillau C, Lapointe J, Giegé R and Kern D.** (2004). A minimalist glutamyl-tRNA synthetase dedicated to aminoacylation of the $tRNA^{Asp}$ QUC anticodon. *Nucleic Acids Res*. 32, 2768-2775.
43. **Dubois DY, Blaise M, Becker HD, Campanacci V, Keith G, Giegé R, Cambillau C, Lapointe J and Kern D.** (2004). An aminoacyl-tRNA synthetase-like protein encoded by the *Escherichia coli yadB* gene glutamylates specifically $tRNA^{Asp}$. *Proc Natl Acad Sci U S A*. 101, 7530–7535.
44. **Salazar JC, Ambrogelly A, Crain PF, McCloskey JA and Söll D.** (2004). A truncated aminoacyl–tRNA synthetase modifies RNA. *Proc Natl Acad Sci U S A*. 101, 7536–7541.
45. **Lin YL, Elias Y and Huang RH.** (2005). Structural and Mutational Studies of the Catalytic Domain of colicin E5: A tRNA-Specific Ribonuclease. *Biochemistry*. 44, 10494-10500.
46. **Swairjo MA, Reddy RR, Lee B, Van Lanen SG, Brown S, De Crécy-Lagard V, Iwata-Reuyl D and Schimmel P.** (2005). Crystallization and preliminary X-ray characterization of the nitrile reductase QeuF: a queuosine-biosynthesis enzyme. *Acta Crystallog. sect. F*. F61, 945-948.

47. **Ishitani R, Nureki O, Fukai S, Kijimoto T, Nameki N, Watanabe M, Kondo H, Sekine M, Okada N, Nishimura S and Yokoyama S.** (2002). Crystal structure of archaeosine tRNA-guanine transglycosylase. *J Mol Biol*. 318, 665-677.
48. **Ishitani R, Nureki O, Nameki N, Okada N, Nishimura S and Yokoyama S.** (2003). Alternative tertiary structure of tRNA for recognition by a posttranscriptional modification enzyme. *Cell*. 113, 383-394.
49. **Sprinzl M, Horn C, Brown MJ, Ioudovitch A and Steinberg S.** (1998). Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*. 26, 148-153.
50. **Gregson JM, Crain PF, Edmonds CG, Gupta R, Hashizume T, Phillipson DW and McCloskey JA.** (1993). Structure of the archaeal transfer RNA nucleoside G*-15 (2-amino-4,7-dihydro- 4-oxo-7-beta-D-ribofuranosyl-1H-pyrrolo[2,3-d]pyrimidine-5-carboximi dam ide (archaeosine)). *J Biol Chem*. 268, 10076-10086.
51. **Grädler U.** *De Novo*-Design und Strukturbestimmung von Inhibitoren der tRAN-Guanin Transglykosylase aus *Zymomonas mobilis* als neues Target der Bakterienruhr [Doktorarbeit]. Philipps-University of Marburg; 2000.
52. **Brenk R.** Virtuelles Screening, strukturbasiertes Design und Kristallstrukturanalyse von Inhibitoren der tRNA-Guanin Transglykosylase, ein Target der Bakterienruhr [Doktorarbeit]. Philipps-University of Marburg; 2003.
53. **Goodenough-Lashua DM and Garcia GA.** (2003). tRNA-guanine transglycosylase from E. coli: a ping-pong kinetic mechanism is consistent with nucleophilic catalysis. *Bioorg Chem*. 31, 331-44.
54. **Stengl B, Reuter K and Klebe G.** (2005). Mechanism and Substrate Specificity of tRNA–Guanine Transglycosylases (TGTs): tRNA Modifying Enzymes from the Three Different Kingdoms of Life Share a Common Catalytic Mechanism. *Chembiochem*. 6, 1-15.
55. **Romier C, Meyer JE and Suck D.** (1997). Slight sequence variations of a common fold explain the substrate specificities of tRNA-guanine transglycosylases from the three kingdoms. *FEBS Lett*. 416, 93-8.
56. **Andreeva A, Howorth D, Brenner SE, Hubbard TJP, Chothia C and Murzin AG.** (2004). SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res*. Vol. 32, D226-D229.
57. **Pearl FMG, Lee D, Bray JE, Sillitoe I, Todd AE, Harrison AP, Thornton JM and Orengo CA.** (2000). Assigning genomic sequences to CATH. *Nucleic Acids Res*. 28, 277-282.
58. **Shindo-Okada N, Okada N, Ohgi T, Goto T and Nishimura S.** (1980). *Biochemistry*. 19, 395-400.
59. **Bai Y, Fox DT, Lacy JA, Van Lanen SG and Iwata-Reuyl D.** (2000). Hypermodification of tRNA in Thermophilic archaea. Cloning, overexpression, and characterization of tRNA-guanine transglycosylase from *Methanococcus jannaschii*. *J Biol Chem*. 275, 28731-28738.
60. **Watanabe M, Matsuo M, Tanaka S, Akimoto H, Asahi S, Nishimura S, Katze JR, Hashizume T, Crain PF, McCloskey JA and Okada N.** (1997). Biosynthesis of archaeosine, a novel derivative of 7-deazaguanosine specific to archaeal tRNA, proceeds via a pathway involving base replacement on the tRNA polynucleotide chain. *J Biol Chem*. 272, 20146-20151.
61. **Deshpande KL and Katze JR.** (2001). Characterization of cDNA encoding the human tRNA-guanine transglycosylase (TGT) catalytic subunit. *Gene*. 265, 205-12.

62. **Ferré-D'Amaré AR.** (2003). RNA-modifying enzymes. *cur Opin Struct Biol*. 13, 49-55.
63. **J. S and Söll D.** (2006). The RNA-binding PUA domain of archaeal tRNA-guanin transglycosylase is not required for archaeosine formation. *J Biol Chem*. published online, http://www.jbc.org/cgi/doi/10.1074/jbc.M512841200.
64. **Brenk R, Stubbs MT, Heine A, Reuter K and Klebe G.** (2003). Flexible adaptations in the structure of the tRNA-modifying enzyme tRNA-guanine transglycosylase and their implications for substrate selectivity, reaction mechanism and structure-based drug design. *Chembiochem*. 4, 1066-77.
65. **Meyer EA, Donati N, Guillot M, Schweizer WB, Diederich F, Stengl B, Brenk R, Reuter K and Klebe G.** (2006). Synthesis, Biological Evaluation, and Crystallographic Studies of Extended Guanine based (*lin*-Benzoguanine) Inhibitors for tRNAGuanine Transglycosylase (TGT). *Helv. Chim. Acta*. 89, 573-597.
66. **Meyer EA, Furler M, Diederich F, Brenk R and Klebe G.** (2004). Synthesis and *In Vitro* Evaluation of 2-Aminoquinazolin-4(3*H*)-one-Based Inhibitors for tRNA-Guanine Transglycosylase (TGT). *Helv. Chim. Acta*. 87, 1333-1356.
67. **Todorov KA and Garcia GA.** (2006). Role of Aspartate 143 in *Escherichia coli* tRNA-Guanine Transglycosylase: Alteration of Heterocyclic Substrate Specificity. *Biochemistry*. 45, 617-625.
68. **Reuter K and Ficner R.** (1995). Sequence analysis and overexpression of the Zymomonas mobilis tgt gene encoding tRNA-guanine transglycosylase: purification and biochemical characterization of the enzyme. *J Bacteriol*. 177, 5284-8.
69. **Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilbout S and Schneider M.** (2003). the SWISS-PROT protein knowlegebase and its supplement TrEMBL in 2003. *Nucleic Acids Res*. 31, 365-370.
70. **Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B and Bork P.** (2006). Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science*. 311, 1283-1287.
71. **Gerlt JA.** (2000). New wine from old barrels. *Nature*. 7, 171-173.
72. **Nagano N, Orengo CA and Thornton JM.** (2002). One Fold with Many Functions: The Evolutionary Relationships between TIM Barrel Families Based on their Sequences, Structures and Functions. *J Mol Biol*. 321, 741-765.
73. **Farber GK and Petsko GA.** (1990). The evolution of $\alpha$/ß barrel enzymes. *Trends Biochem. Sci.* 15, 228-234.
74. **Lesk AM, Branden CI and Chothia C.** (1989). Structural principles of $\alpha$/ß barrel proteins: The packing of the interior of the sheet. *Protein Struct. Funct. Genet.* 5, 139-148.
75. **Copley RR and Bork P.** (2000). Homology among $(ß\alpha)_8$ Barrels: Implications for the Evolution of Metabolic Pathways. *J. Mol. Biol.* 303, 627-640.
76. **Martin W and Koonin EV.** (2006). Introns and the origin of nucleus-cytosol compartmentalization. *Nature*. 440, 41-45.
77. **McGovern SL, Caselli E, Grigorieff N and Shoichet BK.** (2002). A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening. *J. Med. Chem*. 45, 1712-1722.
78. **Ryan AJ, Gray NM, Lowe PN and Chung C.** (2003). Effect of Detergent on "Promiscuous" Inhibitors. *J. Med. Chem.* 46, 3448-3451.

79. **Nonekowski ST and Garcia GA.** (2001). tRNA recognition by tRNA-guanine transglycosylase from *Escherichia coli*: the role of U33 in U-G-U sequence recognition. *Rna*. 7, 1432-1441.
80. **Kung FL and Garcia GA.** (1998). tRNA-guanine transglycosylase from Escherichia coli: recognition of full-length 'queuine-cognate' tRNAs. *FEBS Lett*. 431, 427-32.
81. **Bisswanger H.** (2000). *Enzymkinetik Theorie und Methoden*. 3. ed Weinheim, New York, Chichester, Brisbane, Singapore, Toronto: WILEY-VCH.
82. **Segel IH.** (1993). *Enzyme Kinetics Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* New York: WILEY.
83. **Rich RL and Myszka DG.** (2000). Advances in surface plasmon resonance biosensor analysis. *Curr Opin Biotechnol*. 11, 54-61.
84. **Brenk R, H.-D. G, Kittendorf JD, Garcia GA, Reuter K and Klebe G.** (2003). From Hit to Lead: *De Novo* Design Based on Virtual Screening Hits of Inhibitors of tRNA-Guanine Transglycosylase, a Putative Target of Shigellosis Therapy. *Helv. Chim. Acta*. 86, 1435-1452.
85. **Meyer EA, Brenk R, Castellano RK, Furler M, Klebe G and Diederich F.** (2002). De Novo Design, Synthesis, and In Vitro Evaluation of Inhibitors for Prokaryotic tRNA-Guanine Transglycosylase: A Dramatic Sulfur Effect on Binding Affinity. *Chembiochem*. 3, 250-253.
86. **Howard EI, Sanishvili R, Cachau RE, Mitschler A, Chevrier B, Barth P, Lamour V, Van Zandt M, Moras D, Schneider TR, Joachimiak A and Podjarny A.** (2004). Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase–Inhibitor Complex at 0.66 Å. *Proteins: Structure, Function and Bioinformatics*. 55, 792-804.
87. **Sotriffer C, Krämer O and Klebe G.** (2004). Probing flexibility and 'induced-fit' phenomena in aldose reductase by comparative crystal structure analysis and molecular dynamics simulations *Proteins: Structure, Function and Genetics*. 56, 52-66.
88. **Urzhumtsev A, Tete-Favier F, Mitschler A, Barbanton J, Barth P, Urzhumtseva L, Biellmann JF, Podjarny A and Moras D.** (1997). A 'specificity' pocket inferred from the crystal structures of the complexes of aldose reductase with the pharmaceutically important inhibitors tolrestat and sorbinil. *Structure*. 5, 601-612.
89. **Rauh D, Klebe G, Sturzebecher J and Stubbs MT.** (2003). ZZ made EZ: influence of inhibitor configuration on enzyme selectivity. *J. Mol. Biol.* 330, 761-770.
90. **Rauh D, Reyda S, Klebe G and Stubbs MT.** (2002). Trypsin mutants for structure-based drug design: expression, refolding and crystallisation. *J. Biol. Chem.* 383, 1309-1314.
91. **Bohm H and Klebe G.** (1996). Was läßt sich aus der molekularen Erkennung in Protein-Ligand-Komplexen für das Design neuer Wirkstoffe lernen? *Angew Chem*. 108, 2750-78.
92. **Albert A, Goldacre R and Phillips J.** (1948). *J. Chem. Soc*, 2240.
93. **Brown HC.** (1955). In: Braude EA, Nachod FC, eds. *Determination of Organic Structures by Physical Methods*. New York: Academic Press.
94. **Bruice TC and Schmir GL.** (1958). *J. Am. Chem. Soc.* 80, 148.
95. **Lenz J.** "Fischen" von Liganden - Ein supramolekularer Ansatz zum Auffinden biologisch aktiver Verbindungen aus komplexen Gemischen [Doktorarbeit]. Philipps-University of Marburg; 1999.

96. **Heller D.** Auffinden von Inhibitoren der tRNA-Guanin Transglykosylase (TGT) und der Insulin-like Growth Factor 1 Receptor Tyrosine Kinase (IGF-1-RTK) aus Pflanzenextrakten durch Ligandenfischen sowie deren Isolierung und Identifizierung [Doktorarbeit]. Philipps University of Marburg; 2005.
97. **Sambrook J, Fritsch EF and Maniatis T.** (1989). *Molecular Cloning - a laboratory manual*. 2nd ed New York: Cold Springer Harbor.
98. **Curnow AW, Kung FL, Koch KA and Garcia GA.** (1993). tRNA-guanine transglycosylase from *Escherichia coli*: gross tRNA structural requirements for recognition. *Biochemistry*. 32, 5239-46.
99. **Cantor CR and Schimmel PR.** (1980). Part 11: Techniques for the Study of Biogical Structure and Function. In: Chemistry B, ed. *Freeman, W. H.* San Francisco.
100. **Romier C, Ficner R, Reuter K and Suck D.** (1996). Purification, crystallization, and preliminary x-ray diffraction studies of tRNA-guanine transglycosylase from Zymomonas mobilis. *Proteins*. 24, 516-9.
101. **Reuter K, Chong S, Ullrich F, Kersten H and Garcia GA.** (1994). Serine 90 is required for enzymic activity by tRNA-guanine transglycosylase from Escherichia coli. *Biochemistry*. 33, 7041-6.
102. GraFit program]. Version 4.09. Erithacus Software Limited, USA; 1999.
103. **Fischer W and Hannappel E.** (1996). *Praktikum für Biochemie, 11. neubearbeitete und erweiterte Auflage* Erlangen: Verlag Ludwig Müller.
104. **Otwinowski Z and Minor W.** (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307-326.
105. **Navaza J.** (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr. sect. A*. 50, 157-163.
106. **Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T and Warren GL.** (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*. 54, 905-921.
107. **Sheldrick GM and Schneider TR.** (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* 277b, 319-343.
108. **Jones TA, Zou JY, Cowan SW and Kjeldgaard.** (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*. 47, 110-119.
109. **Laskowski RA, MacArthur MW, Moss DS and Thornton JM.** (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 26, 283-291.
110. **Gerber PR and Muller K.** (1995). MAB, a generally applicable molecular force field for structure modelling in medicinal chemistry. *J Comput Aided Mol Des*. 9, 251-68.

## Danksagungen

Herrn Prof. Dr. *Gerhard Klebe* danke ich sehr herzlich für das spannende Thema, wertvolle Anregungen in Diskussionen und insbesondere für die gewährte Freiheit bei der Bearbeitung des Themas.

Bei PD Dr. *Klaus Reuter* möchte ich mich für hilfreiche Diskussionen bedanken, die mit dazu beitrugen, die Bestimmung der Inhibitionskonstanten in Einklang mit dem Reaktionsmechanismus der TGT zu bringen.

*Ruth Brenk* und *Tanja Sgraja* danke ich für die Einarbeitung in die praktischen Methoden dieser Arbeit.

*Naomi Tidten* danke ich für ihre Hilfestellung bei der Herstellung der TGT(E235Q)-Mutante. Ihre Laborerfahrung hat es möglich gemacht, tRNA wesentlich effizienter herzustellen und ein Verfahren zu etablieren, tRNA radioaktiv zu markieren.

Dr. *Andreas Heine* möchte ich für seine enorme Hilfe und unendliche Geduld danken bei allem, was mit Röntgenkristallographie zu tun hat.

*Emmanuel Meyer*, *Simone Hörtner* und Prof. *François Diederich* von der ETH Zürich danke ich für die gute Zusammenarbeit. Besonders die beiden Besuche in Zürich, die nicht nur das Projekt sehr bereicherten, werden mir noch lange in Erinnerung bleiben.

*Tim Larsen* und Prof. *Andreas Link* möchte ich für die gute Kooperation danken, die auch nach ihrem Umzug von Marburg nach Greifswald reibungslos funktionierte.

Auch bei *Daniela Heller* und Prof. *Rudolf Matusch* möchte ich mich für die gute Zusammenarbeit bedanken. Obwohl wir noch keine Kristallstrukturen der bisher „gefischten“ Verbindungen erhalten konnten, ist der Ansatz doch sehr innovativ.

Bei *Wolfgang Hartmann* und Prof. Dr. *Michael Keusgen* möchte ich mich dafür bedanken, an ihrem SPR-Gerät Messungen durchführen zu können.

Mein Dank gilt auch *Hans-Dieter Gerber* für seine unermüdliche Mühe, die Synthese von $preQ_1$ wesentlich zu optimieren.

*Tina Ritschel* danke ich für die spannenden Diskussionen bei der Übergabe des Projektes. Ich wünsch Ihr sehr viel Erfolg und Spaß mit ihren Vorhaben.

Ich danke meinen Literaturarbeitern *Tilo Jansen* und *Philipp Hoffmann* und meiner Vertieferin *Valerie Hondorf* für ihre geleistete Arbeit.

Allen Mitgliedern und Ehemaligen der Arbeitsgruppe Klebe danke ich für die Hilfsbereitschaft, viele interessante Diskussionen und die angenehme Arbeitsatmosphäre.

Bei *Angela Scholz* möchte ich mich bedanken für die Hilfe bei allen bürokratischen Angelegenheiten. Ihr offenes Ohr für Probleme und ihre sympathische Art trugen wesentlich zum angenehmen Klima in der Gruppe bei.

*Christian Sohn* danke ich für die Unterstützung bei Laborarbeiten, sowie für die Betreuung des Röntgengeräts.

Ein besonderer Dank geht an die Administratoren für stets einsatzbereite Rechner und die Bereitschaft bei Problemen kurzfristig die eigene Arbeit zu unterbrechen.

Für die sorgfältige Durchsicht des Manuskripts danke ich PD Dr. *Klaus Reuter*, Dr. *Andreas Heine*, *Naomi Tidten, Tina Ritschel* und *Alexander Hillebrecht*.

Bedanken möchte ich mich bei allen Freunden und Gruppenmitgliedern aus Marburg, insbesondere bei *Tobias Samusch, Bettina Beckmann, David Hartmann* und den anderen lieben Leuten aus der *KHG Marburg*, sowie *Jasmine Fokkens* und *Christoph Sotriffer*. Viele Erlebnisse mit ihnen werden schöne Erinnerungen an Marburg bleiben.

Meinen Eltern danke ich für ihre Geduld, mich die langen Jahre des Studiums und der Promotion immer interessiert zu unterstützen.

Ganz herzlich danken will ich *Hanna Blumenschein*, die ich in Marburg kennen und lieben lernen durfte. Sie hat mir immer wieder Mut gemacht weiterzumachen, wenn ich vor mehr Fragen als Antworten stand.

Der Deutschen Forschungsgemeinschaft und dem Graduiertenkolleg *Proteinfunktion auf atomarer Ebene* danke ich für die finanzielle Unterstützung.

## Erklärung

Ich versichere, dass ich meine Dissertation

**Structural and Functional Studies of**

**tRNA-Guanine Transglycosylase:**

**A putative Drug Target for Shigellosis Therapy**

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe.

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

Marburg, den 31. Mai 2006 _____

( Bernhard Stengl )

## Lebenslauf

Bernhard Josef Stengl

| | |
|---|---|
| 04.Oktober 1973 | geboren in Roth bei Nürnberg |
| Sept.1980 – Juli 1984 | Grundschule Allersberg |
| Sept. 1984 – Juli 1993 | Gymnasium Hilpoltstein |
| Juli 1993 | Abitur |
| Juli 1993 – Juni 1994 | Grundwehrdienst in Roth bei Nürnberg |
| Okt. 1994 – Sept. 1995 | Studium der Geodäsie an der TU Karlsruhe |
| Okt. 1995 – Juli 2002 | Studium von Biologie und Chemie für das Lehramt an Gymnasien an der FAU Erlangen-Nürnberg |
| Mai 1997 | Aufnahme in das Cusanuswerk (Bischöfliche Studienförderung) |
| Aug. 1997 | Zwischenprüfung Chemie |
| Feb. 1998 | Zwischenprüfung Biologie |
| Okt. 1999 – April 2000 | Zulassungsarbeit im Fach Biologie im Arbeitskreis von Prof. Dr. Jürgen Heinze: *Kutikulare Kohlenwasserstoffe und Sozialstruktur bei der Ameisenart Pachycondyla inversa* |
| Dez. 1999 – Feb. 2000 | Forschungsaufenthalt an der Universität Florenz im Rahmen der Zulassungsarbeit |
| Juli 2002 | Erste Staatsprüfung für das Lehramt an Gymnasien |
| Okt. 2002 – Juli 2006 | Promotion an der Philipps-Universität Marburg in der Arbeitsgruppe von Prof. Dr. Gerhard Klebe mit dem Titel: *Structural and Functional Studies of tRNA-Guanine Transglycosylase: A putative Drug Target for Shigellosis Therapy* |