Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany

for the degree of
Doctor of Natural Sciences

presented by

Andrea Rocker, Master of Science
born in Emmerich am Rhein, Germany

Oral examination: 22.5.15

Epsilon/Zeta Toxin-Antitoxin Systems in Gram-negative Bacteria

Referees: Prof. Dr. Ilme Schlichting

Dr. Anton Meinhart

Content

Summary

Their genetic repertoire allows bacteria to adapt to constantly changing environments and to survive inhospitable conditions. Such genetic modules are toxin-antitoxin (TA) systems, which increase the survival of a population by inducing dormancy or cell death of a subpopulation of cells. They encode intracellular toxins, which have the potential to kill the cell, and are therefore tightly regulated by cognate antitoxins that inhibit the toxic activity.

Toxicity is exerted by interference with essential cellular processes including replication, translation and homeostasis. Correspondingly, the toxins of the zeta family were found to poison cell wall synthesis by phosphorylation of the peptidoglycan precursor UDP-*N*-acetylglucosamine (UNAG). Zeta toxins are counteracted by small, proteinaceous epsilon antitoxins. These epsilon/zeta TA systems were first discovered as plasmid maintenance modules on the plasmids of Gram-positive streptococci and have long been thought to be confined to Gram-positive bacteria.

This thesis provides the first biochemical and structural characterization of zeta toxin homologs from Gram-negative bacteria - EzeT from *Escherichia coli* and the plasmidic epsilon/zeta systems from *Neisseria gonorrhoeae* - showing that these proteins exhibit zeta-like kinase activity. Moreover, several differences to the hitherto described streptococcal epsilon/zeta systems provide evidence for the diversity of the epsilon/zeta TA family.

The first part of this thesis describes the characterization of the elongated zeta homolog EzeT, which was found to combine toxin and antitoxin functionalities in one polypeptide chain. The protein consists of two domains, one of which was demonstrated to have UNAG kinase activity *in vitro* and *in vivo*. The enzymatic activity is abrogated in the presence of the N-terminal antitoxin domain and mutational analysis indicated that inhibition is performed by an epsilon-like mechanism. Furthermore, EzeT toxicity is temperature dependent, leading to a lytic phenotype and spherical morphology at ambient temperature. The presented novel type of toxin inhibition by a covalently linked antitoxin sets EzeT apart from other TA systems and necessitates the analysis of other elongated or orphan toxin homologs that have been identified in bacterial genomes.

In the second part of this thesis, the X-ray crystal structure of the epsilon/zeta system ε1/ζ1 of *N. gonorrhoeae* is presented, together with biochemical studies that confirmed ζ1 as a UNAG kinase, which is inhibited by binding of ε1. Bioinformatic analysis of ζ1 showed that in the amino acid sequence the catalytically important Walker A motif, commonly found in P-loop kinases, is located closer to the C-terminus than in homologous proteins. Nevertheless, ζ1 adopts a typical zeta-like three-dimensional structure with a conserved architecture of the catalytic center. This is achieved by a rearrangement of the secondary structure elements and a corresponding topological rewiring, forming a fold unusual for P-loop kinases. However, nucleotide binding and hydrolysis

are not impaired and have been observed spectroscopically and by determination of the ADP-bound complex structure.

In addition to the zeta-like domain, ζ1 contains a C-terminal domain that comprises an OB-fold with similarity to DNA-binding modules. This domain is also present in two other gonococcal epsilon/zeta homologs, which are located on the same plasmid. As presented in this work, these systems, called ε2/ζ2 and ε3/ζ3, phosphorylate UNAG, but do not directly interact with ε1/ζ1. Furthermore, the X-ray crystal structure of the heterodimeric ε1/ζ1 complex revealed that inhibition is mediated by ε1 wrapping around the zeta core, indicating a different mode of inhibition compared to the streptococcal antitoxins.

In conclusion, data presented in this thesis show that Gram-negative bacteria encode active UNAG kinases. The characterization of these systems reveals novel inhibition mechanisms by the antitoxins and highlights the structural diversity of zeta toxins.

Zusammenfassung

Die Anpassungsfähigkeit von Bakterien basiert auf ihren im Verlauf der Evolution erworbenen Erbinformationen, die ihnen ermöglichen, sich an wechselnde Umweltbedingungen anzupassen und auch in unwirtlichen Umgebungen zu bestehen. Zu den genetischen Modulen, die das Überleben der Gesamtpopulation sichern, gehören die Toxin-Antitoxin-Systeme (TA-Systeme). Diese Systeme können Zellen in eine Art Ruhezustand versetzen oder bei einem Teil der Bakterien zum Zelltod führen. Um dies zu erreichen kodieren TA-Systeme intrazelluläre Toxine, welche zum bakteriellen Zelltod führen indem sie in lebenswichtige zelluläre Prozesse wie Replikation, Translation oder Homöostase eingreifen. Daher müssen diese Toxine stark reguliert werden, weshalb Bakterien immer auch entsprechende Antitoxine produzieren, welche die Toxine inhibieren.

In der Familie der Epsilon/Zeta Systeme, neutralisieren kleine Proteine, die Epsilon-Antitoxine, die toxische Aktivität der Zeta-Toxine. Zeta-Toxine beeinträchtigen die Zellwandsynthese durch Phosphorylierung von UDP-*N*-acetylglukosamin (UNAG), einem essentiellen Bestandteil des bakteriellen Mureins. Epsilon/Zeta TA-Systeme wurden zuerst als Plasmid-stabilisierende Module in Gram-positiven Streptokokken entdeckt und man ging lange davon aus, dass sie auf Gram-positive Bakterien beschränkt seien.

Die vorliegende Arbeit umfasst die erste biochemische und strukturelle Beschreibung von Zeta-Toxinen Gram-negativer Bakterien (EzeT von *Escherichia coli* und die plasmidischen Epsilon/Zeta-Systeme von *Neisseria gonorrhoeae*). Es konnte gezeigt werden, dass diese Proteine UNAG-Kinase-Aktivität besitzen, aber im Vergleich zu den bisher beschriebenen Epsilon/Zeta-Systemen in Streptokokken eine Reihe von Unterschieden aufweisen. Damit geben sie einen Einblick in die Diversität innerhalb der Epsilon/Zeta-Familie.

Der erste Teil der Arbeit beschreibt das auffällig verlängerte Zeta-Homolog EzeT, welches toxische und antitoxische Funktionen innerhalb einer Polypeptidkette vereint. EzeT besteht aus zwei Proteindomänen, von denen für die C-terminale Domäne UNAG-Kinase-Aktivität *in vitro* und *in vivo* nachgewiesen werden konnte. Die toxische Wirkung des Enzyms wird durch das N-terminale Antitoxin unterbunden. Weitergehende Untersuchungen mittels Mutationsanalyse zeigen, dass dies auf einen Epsilon-ähnlichen Inhibierungsmechanismus zurückzuführen ist. Zudem wurde eine Temperaturabhängigkeit der Toxizität nachgewiesen, die bei niedrigen Temperaturen zur Lyse und zur Ausbildung einer abgerundeten Zellmorphologie führt. Diese neue Form der Inhibierung von Toxinen durch ein kovalent gebundenes Antitoxin unterscheidet EzeT von anderen TA-Systemen. Dies legt die Untersuchung von weiteren verlängerten oder bisher als ´verwaist` geltenden Toxin-Homologen nahe, die bereits in den Genomen von anderen Bakterien identifiziert worden sind.

Im zweiten Teil der Arbeit wird die Röntgenkristallstruktur des Epsilon/Zeta-Systems ε1/ζ1 von *N. gonorrhoeae* vorgestellt, zusammen mit biochemischen Studien, die die UNAG-Kinase-Aktivität von ζ1 und dessen Inhibierung durch ε1 beweisen. Die bioinformatische Untersuchung von ζ1 zeigt, dass in der Aminosäuresequenz das essentielle Walker A Motiv näher beim C-Terminus liegt als in homologen Proteinen. Trotzdem besitzt ζ1 eine typische dreidimensionale Struktur ähnlich den bekannten Zeta-Toxinen und eine konservierte Architektur des katalytischen Zentrums. Dies wird durch eine veränderte Reihung der Sekundärstrukturelemente und einer dementsprechenden topologischen Neuordnung erreicht, was zu einer für P-loop Kinasen unüblichen Faltung führt. Dabei wird der katalytische Mechanismus nicht beeinträchtigt, wie durch spektroskopische Messungen und durch Bestimmung der ADP-gebundenen Komplexstruktur nachgewiesen werden konnte.

Zusätzlich zur Zeta-ähnlichen Domäne enthält ζ1 eine weitere abgegrenzte C-terminale Domäne. Diese beinhaltet eine OB-Faltung, die Ähnlichkeiten mit DNS-bindenden Modulen aufweist. Eng verwandte C-terminale Domänen konnten auch in zwei weiteren Epsilon/Zeta-Systemen von *N. gonorrhoeae* identifiziert werden, die gemeinsam mit ε1/ζ1 auf einem Plasmid vorkommen. Diese Systeme, genannt ε2/ζ2 und ε3/ζ3, phosphorylieren ebenfalls UNAG. Sie treten aber nicht direkt mit ε1/ζ1 in Wechselwirkung. Zudem legt die Röntgenkristallstruktur des ε1/ζ1-Komplexes im Vergleich zu den bereits beschriebenen Epsilon-Antitoxinen einen geänderten Inhibitionsmechanismus nahe, in dem ε1 in gestreckter Konformation um das Zeta-Toxin gewunden ist.

Zusammenfassend beweist diese Arbeit, dass Gram-negative Bakterien aktive Zeta-Toxine kodieren. Die genauere Untersuchung dieser Systeme zeigt nicht nur bisher unbekannte Inhibitionsmechanismen, sondern gibt auch einen Einblick in die strukturelle Vielfalt und enzymatische Variabilität der Zeta-Toxine.

Publications

Results in part of this thesis are also presented in:

Rocker A, Meinhart A; A cis-acting antitoxin inhibits the chromosomal toxin-antitoxin module EzeT of *Escherichia coli*.

Submitted for publication to *Molecular Microbiology*.

The text and the figures in this manuscript have been prepared by Andrea Rocker.

1 Introduction

Bacteria are versatile unicellular organisms that thrive in virtually every habitat on earth including the deep sea [1], hot springs [2] and Antarctica [3]. Under favorable conditions, for example in surface sea water, bacterial densities can reach up to 10^6 cells per milliliter [4]. In these dense communities, bacteria are in constant competition for resources and subject for predation by bacteriophages and larger organisms [5, 6]. In addition, natural habitats are exposed to climatic and meteorological changes and bacteria need to adjust to nutrient fluctuations, oxygen availability and the presence of damaging agents like radiation and oxidizing agents [7]. Moreover, bacteria can live in close association with larger organisms, either as commensals or as pathogens, where they have to cope with attacks by the immune system and antibiotics [8].

In response to these cues, bacteria can regulate gene expression to switch between different metabolic pathways, activate repair systems or synthesize resistance proteins that inactivate stressors such as antibiotics [9]. During evolution, the acquisition of novel genes by horizontal gene transfer can expand the genetic repertoire of bacteria, allowing them to exert new functions or to exploit new ecological niches [10]. To evade stressful conditions, bacteria can form biofilms, matrix-enclosed populations in which the cells are protected against physical or chemical challenges such as desiccation and phagocytosis [11]. When the conditions are too inhospitable to sustain metabolism, bacterial growth ceases and individual cells survive in the form of dormant persister cells [12] or spores [13].

Many of the bacterial strategies to adapt to environmental stresses involve social behaviors that reach beyond the historical perception of bacteria as purely unicellular organisms. Multicellularity has not only been observed in highly specialized species like filamentous cyanobacteria or streptomycetes [13], which differentiate into distinct cell types with specific functions, but in most bacteria that use quorum sensing to coordinate cellular behaviors like predation, sporulation and competence [14]. These strategies generally aim to ensure the survival of a population, instead of individual cells, and bacterial altruistic behavior has been described, for instance in phage abortive infection, where infected cells undergo cell lysis to prevent the spread of phages [15]. Cell death has also been implicated in biofilm formation [16] and sporulation [13] and it was found that these death pathways are regulated by intracellular programs displaying a form of programmed cell death (PCD) in bacteria [17].

1.1 Toxin-antitoxin systems

Systems that can induce bacterial PCD are so-called toxin-antitoxin (TA) systems[18]. TA modules commonly consist of two genes with antagonistic functions [18, 19]. One gene product- the toxin - has the potential to induce cell death or stasis, while the second gene product – the antitoxin - neutralizes this toxicity. Usually both genes are co-regulated, ensuring that sufficient antitoxin levels are produced to inhibit the toxic activity. However, degradation of the antitoxin can free the toxin to fulfil its intracellular function [18]. These functions can range from the stabilization of mobile genetic elements to the induction of dormancy or PCD [20].

1.1.1 Types of TA systems

While all identified toxins are proteins, the nature and the inhibition mechanism by the antitoxins can vary, which is used to classify TA systems into different types [21]. In type I systems the antitoxin is transcribed as a small RNA that binds to the toxin mRNA and thereby interferes with toxin translation [21]. In type II and type III systems the toxin is inhibited by complex formation with a proteinaceous antitoxin [19] or a folded RNA molecule [22], respectively. Recently, three new types of TA systems with only few studied representatives were discovered. In type IV systems the antitoxin is a nuclease that specifically degrades toxin mRNA [23], while in type VI systems the antitoxin targets the toxic protein for degradation [24]. In type V systems no direct interaction between toxin and antitoxin takes place. Instead, the antitoxin binds to the toxin substrate and prevents its degradation [25].

The most extensively studied TA modules are type II systems, which in general consist of two adjacent protein-coding genes [18, 26]. In these bicistronic operons, the antitoxin gene is located upstream of the toxin gene. It encodes a small, proteolytically instable protein that can bind to the toxin and thereby inhibit toxicity (Figure 1.1-1). Degradation of the antitoxin by cellular proteases releases the toxin, which can then interfere with an essential cellular process [21]. Most proteinaceous antitoxins contain a DNA-binding domain and bind to the promoter of the bicistron to autoregulate transcription. The repressor-function is further enhanced in the presence of the toxin, which acts as a co-repressor [18, 26].

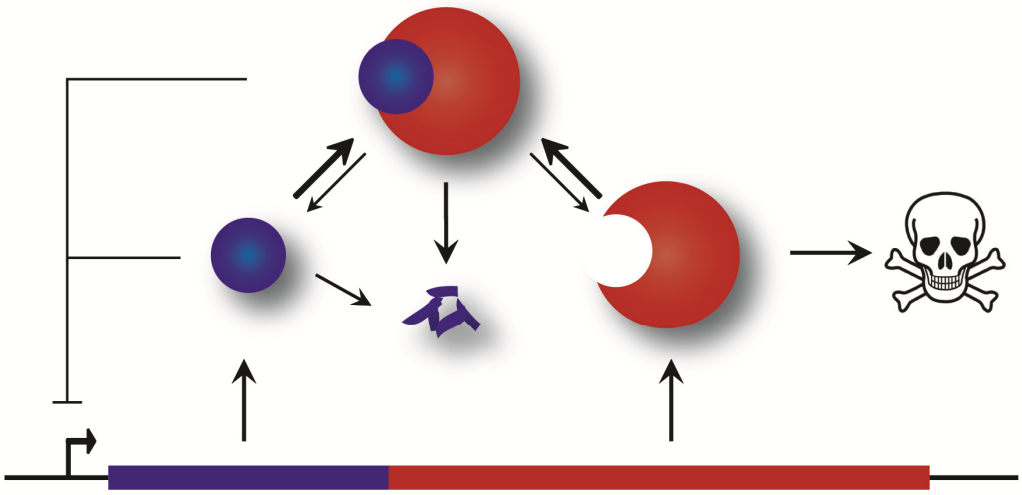


Figure 1.1-1: Schematic of a type II TA system. Toxin (red) and antitoxin (blue) are encoded in a bicistronic operon and form an inactive complex in the cytosol. Expression of both proteins is regulated by transcriptional repression by the antitoxin or the toxin-antitoxin complex. Degradation of the proteolytically unstable antitoxin frees the toxin, which has a bactericidal or bacteriostatic activity.

1.1.2 Functions of TA systems

TA systems were first discovered in 1983 as plasmid stabilization elements on low-copy number plasmids [27] and later shown to act as addiction modules that induce a process called postsegregational killing (PSK) upon plasmid loss [28, 29]. PSK depends on differential proteolytic stabilities of toxin and antitoxin to cause death of plasmid-free offspring. Since the antitoxin is rapidly degraded by cellular proteases it needs to be constantly replenished by protein synthesis. In cells harboring the plasmid, both TA genes are present and an inactive TA complex is formed in the cytosol. However, when the plasmid gets lost during segregation, the daughter cell still inherits the TA protein complex together with other cytosolic compounds. Since the antitoxin cannot be replenished any longer, the toxin is freed and kills the plasmid-free daughter cell [30] (Figure 1.1-2).

With the advances of genomics, a plethora of TA systems and even new toxin and antitoxin families were identified in bioinformatic studies [31-34]. Interestingly, homologs of plasmidic TA systems have been identified on the chromosomes of most eubacteria and archaea [18] where they can be present in large numbers (e.g. 36 reported systems in *Escherichia coli* K-12 [35] and 79 in *Mycobacterium tuberculosis* H37Rv [36]). It has been suggested that these chromosomal systems originated from gene transfer between plasmids and chromosomes [37]. Indeed, chromosomal TA systems are prevalent on mobile genetic elements and some modules were shown to stabilize superintegrons in the genome [38].

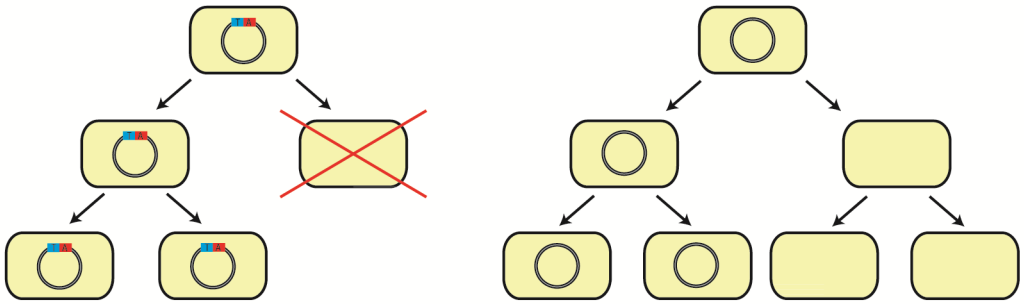


Figure 1.1-2: Stabilization of plasmids by postsegregational killing. Schematic representation of plasmid loss by cells containing low-copy plasmids with (left) or without (right) TA systems. While daughter cells that lose a plasmid encoding a TA system are killed by PSK, plasmid loss in cells without TA systems gives rise to plasmid-free offspring. Figure adapted from Ref. [39].

However, it has been recognized that chromosomal TA systems can exert functions beyond the stabilization of DNA elements [20]. For instance, the simultaneous presence of interacting TA systems on plasmid and chromosome can also enable daughter cells that lose the plasmid to avoid PSK [40]. In this case, the chromosomal antitoxin gene produces sufficient antitoxin to counteract the toxicity of the encoded toxins, even when the antitoxin gene of the plasmid is lost. As plasmids that are actively maintained have a higher chance to get transferred, evolution drives mutational adaptations to diminish this cross-reactivity of plasmidic and chromosomal systems, leaving orphan TA systems on the chromosome that can potentially evolve to fulfill novel functions [40]. One of these novel functions is phage abortive infection, where cells that are infected by phages undergo autolysis before phage particles can assemble and spread [15].

However, toxin activation is not always bactericidal, but can also reduce metabolic activity and evoke dormancy, which can be utilized as a stress response mechanism [19, 41]. For example, the chromosomal hipA toxin of *E. coli* was shown to mediate persister cell formation by phosphorylation of glutamyl-tRNA synthetase [42, 43]. Furthermore, TA systems are implicated in cell differentiation and death during biofilm development [44] and sporulation [45] and as factors that promote the virulence of pathogens [46, 47].

1.1.3 Intracellular poisons

Consistent with their distribution across multiple phyla, toxins target ubiquitous cellular functions [48]. For example, type I pore-forming toxins perforate the plasma membrane and disturb ATP synthesis [49], while the type IV toxin CbtA inhibits MreB and FtsZ polymerization and thereby cell division [50]. The thoroughly studied type II toxins are amongst others DNA gyrase inhibitors that interfere with replication [51], endonucleases that cleave mRNA [18] or kinases that phosphorylate EF-TU and thereby inhibit translation [52]. Furthermore, type II toxins of the zeta family disturb the integrity of the bacterial cell wall and as a result cause autolysis [53].

1.2 Epsilon/zeta systems

1.2.1 Zeta toxins inhibit cell wall synthesis

The type II epsilon/zeta systems contain an epsilon antitoxin that inhibits the toxic action of the zeta toxin by protein complex formation. Zeta toxins are kinases that phosphorylate the peptidoglycan precursor UDP-*N*-acetylglucosamine (UNAG) under consumption of ATP [53] (Figure 1.2-1, top). UNAG is required for the synthesis of cell wall peptidoglycan [54] and glycoconjugates like teichoic acids [55] and lipopolysaccharides [56]. During the cytosolic steps of peptidoglycan synthesis the carboxyvinyltransferase MurA adds an enolpyruvyl to the 3'-hydroxy group of the acetylglucosamine moiety of UNAG [54] (Figure 1.2-1, left). The same hydroxyl group is phosphorylated by zeta toxins yielding UNAG-3P, which competitively inhibits peptidoglycan synthesis and accumulates in the cytosol, because it cannot be incorporated into peptidoglycan [53]. As a consequence the stability of the cell wall is decreased, which ultimately leads to lysis of the cell. UNAG phosphorylation is detrimental to a wide range of cells including eukaryotes, as ζ_{spy} overexpression causes growth inhibition even in the eukaryotic yeast *Saccharomyces cerevisiae* [57].

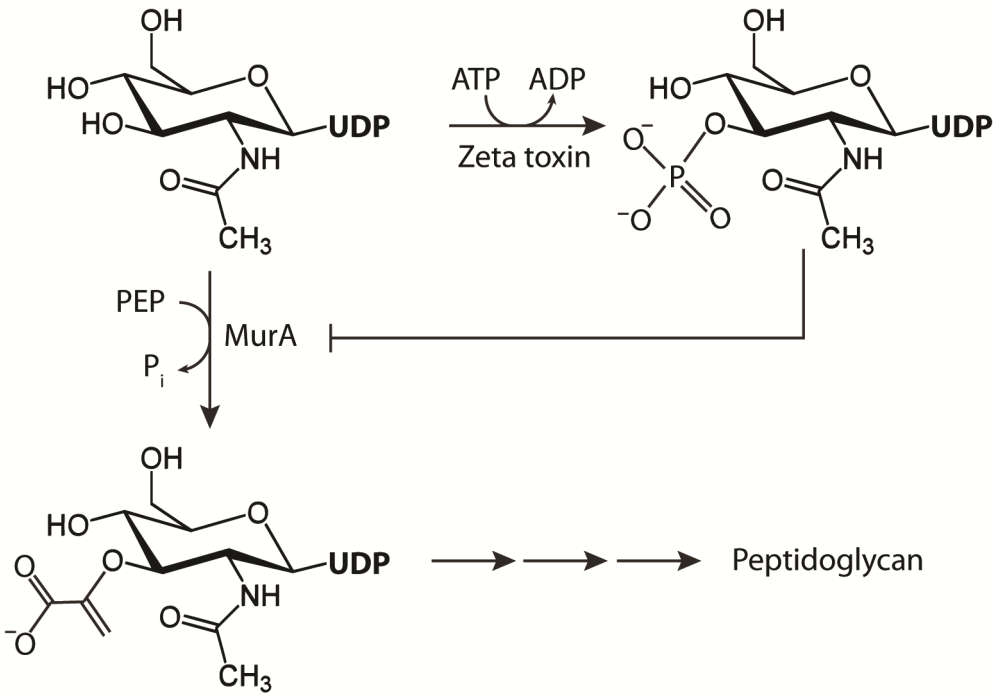


Figure 1.2-1: Reaction catalyzed by zeta toxins. Zeta toxins catalyze the phosphorylation of UNAG under consumption of ATP (top row). The reaction product UNAG-3P inhibits MurA, an enzyme that attaches an enolpyruvyl moiety to the same hydroxyl group (left). Enolpyruvyl-UNAG is an essential component in the synthesis of peptidoglycan.

1.2.2 Streptococcal epsilon/zeta systems

Epsilon/zeta systems were originally discovered as plasmid maintenance modules on low-copy number plasmids and are there denoted with Greek letters: ε/ζ [58]. In addition, homologous systems were found on the chromosomes of various streptococci [59, 60]. One representative of each of these two variants has been extensively studied, the chromosomally encoded pneumococcal epsilon zeta antitoxin toxin (*pezAT*) locus from *Streptococcus pneumonia* and the genes ε and ζ located on plasmid pSM19035 from *Streptococcus pyogenes* (ε/ζ_{spy}) [61]. From the analysis of these streptococcal systems, general assumptions on the functionalities of epsilon/zeta systems have been drawn.

The streptococcal zeta toxins PezT and ζ_{spy} adopt a common fold, which closely resembles that of phosphotransferases, consisting of a central β-sheet surrounded by α-helices (Figure 1.2-2A) [62, 63]. This is also reflected in a similar length of 253 (PezT) or 286 (ζ_{spy}) amino acids. Both proteins share 42 % amino acid sequence identity [62] and contain a Walker A motif characteristic for ATP binding proteins [64, 65]. By homology to other P-loop proteins, the ATP-binding site of zeta toxins could be located in a shallow cleft on one side of the protein. On the opposite side of the β-sheet resides the UNAG-binding site (Figure 1.2-2B), which is characterized by the highly conserved GTXR consensus sequence (where X can be any residue) [66]. Truncation of residues especially at the C-terminus attenuates UNAG kinase activity, suggesting that this globular domain comprises the enzymatically required UNAG kinase core fold [53, 63].

The cognate antitoxins, PezA and ε_{spy}, are less conserved. The carboxy (C)-terminal domain of PezA (residues 70-158) and ε_{spy} share 21 % amino acid sequence identity, while the amino (N)-terminal domain of PezA is predicted to fold into a helix-turn-helix domain typical for transcriptional regulators [62].

In both streptococcal systems, inhibition of UNAG kinase activity is mediated by binding of the cognate epsilon antitoxins. The inhibited complex consists of a dumbbell-shaped heterotetrameric arrangement of two zeta toxins bound to a central epsilon dimer [62, 63] (Figure 1.2-2C). The C-terminal domains of PezA and ε_{spy} both fold into a three-helix bundle. In the inhibited epsilon/zeta complex this bundle binds on top of the ATP binding cleft of zeta and thereby prevents ATP binding and hydrolysis [62, 63]. Several side chains of residues in the amino-terminal helix (Tyr, Phe/Leu, Glu, Glu) protrude into the ATP binding site and sterically block ATP binding (Figure 1.2-2D). Specifically, the large hydrophobic side chains of tyrosine and phenylalanine (ε_{spy}) or leucine (PezA) occupy the binding sites of adenine and the ribose, while the glutamate side chains form hydrogen bonds to residues of the zeta toxin that are required for binding of the phosphoryl-groups of ATP.

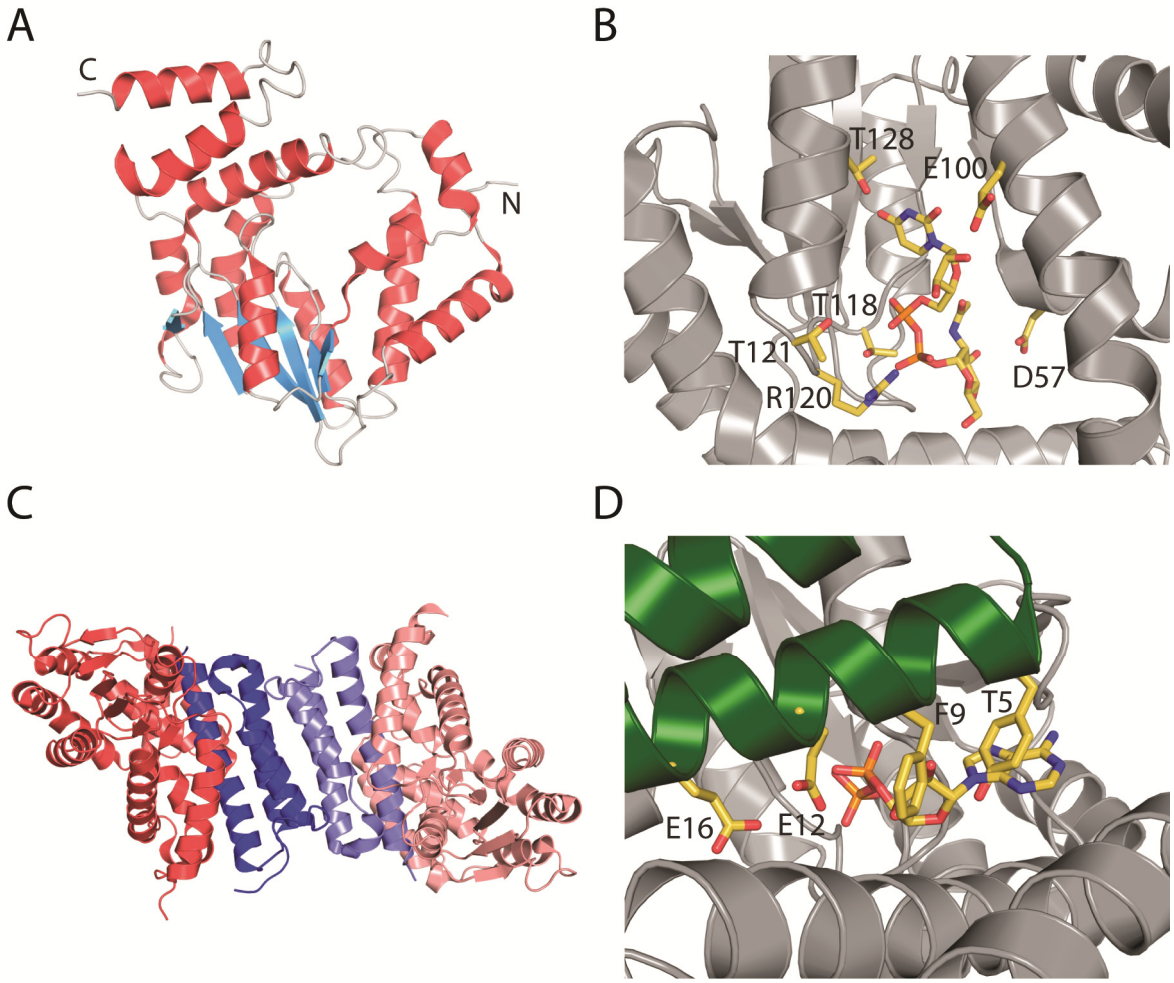


Figure 1.2-2: Structural features of epsilon/zeta systems. (A) Zeta toxin fold of ζ_{spy} (PDB ID: 1GVN). α-helices are shown in red, β-strands in blue, loop regions in gray. **(B)** UNAG binding pocket of ζ_{spy} (gray). Residues previously described to coordinate UNAG and bound UNAG are shown as stick model. PDB ID: 3Q8X **(C)** The heterotetrameric arrangement of the inhibited ε/ζ_{spy} complex. Two ζ_{spy} molecules (red) bind to a central ε_{spy} dimer (blue). The antitoxins adopt a three-helix bundle fold. **(D)** Inhibition by ε_{spy} (green). ADP (stick model) is modelled according to the homologous T4 polynucleotide kinase (PDB ID: 1LTQ) and residues from ε_{spy} that interfere with nucleotide binding are shown as stick models.

Despite their similarities in structure and enzymatic activity the streptococcal epsilon/zeta systems differ in their transcriptional regulation and activation mechanisms and their function. An overview on differences and similarities of these systems can be found in Table 1.1.

The function of the plasmidic ε/ζ_{spy} module has been thoroughly investigated, showing that it acts as a plasmid addiction system that evokes PSK of plasmid-free daughter cells and, together with a plasmid partitioning system, ensures a more than 1000 fold better-than-random segregation efficiency of pSM19035 [58]. In contrast, PezAT is encoded on the hypervariable region of the chromosomal pneumococcal pathogenicity island of some pneumococcal strains [47, 60, 67]. Analysis of pneumococcal strains with different variants of this island with or without active PezT

showed attenuated virulence of strains in the absence of PezT [47, 59]. These results indicate that PezAT is important to establish full virulence of *S. pneumoniae*.

The transcriptional regulation of the chromosomal PezAT system resembles that of other type II TA systems. The N-terminal helix-turn-helix motif of PezA is a DNA-binding domain that enables PezA to bind to the *pezAT* promoter region and to repress transcription of the bicistron (compare Figure 1.1-1). The toxin PezT was shown to act as a co-repressor [62]. In contrast, ε/ζ_{spy} expression is regulated by a third protein, the transcriptional regulator ω [68]. ω is encoded in an operon with ε_{spy} and ζ_{spy} and represses expression from the promoter of the tricistron. In contrast to the helix-turn-helix motif of PezA, the DNA binding domain of ω folds into a ribbon-helix-helix motif [69], indicating a different evolutionary origin of transcriptional regulation.

The binding affinities of ε_{spy} and ζ_{spy} from *S. pyogenes* were reported to be moderate with a dissociation constant of 1.1 µM [70]. This suggests that *in vivo* the heterotetrameric $\varepsilon_2\zeta_{2spy}$ complex is in a dynamic equilibrium with the individual proteins. The protease ClpXP constantly degrades the ε_{spy} antitoxin and leads to activation of the toxin ζ_{spy}, when the antitoxin cannot be replenished [71]. Strikingly, PezAT binding is extremely tight with an apparent binding constant of 65 fM [72]. Furthermore, PezT release was not observed in *E. coli* and PezA was not degraded in the PezAT complex in *Bacillus subtilis* crude extracts [72]. This high stability of the complex suggests that specialized factors are required to activate PezT under certain circumstances. These factors might be host-specific, which is corroborated by BLAST searches that indicate that PezAT systems are restricted to streptococci (unpublished results). A similar host adaptation was found for ε/ζ_{spy} which stabilizes plasmids in firmicutes to varying degrees [73] and is only weakly effective in *E. coli* [74].

Table 1.1: Comparison of the streptococcal epsilon/zeta toxins

	ε/ζ_{spy}	PezAT
Enzymatic activity	UNAG kinase	
Oligomeric State	$\varepsilon_2\zeta_2$	$PezA_2T_2$
Transcriptional regulator	ω	PezA(1-68) (co-repressor PezT)
Function	PSK	virulence?
Binding affinity	1.1 µM	65 fM
Protease/activator	ClpXP	?

1.3 Epsilon/zeta systems and Gram-negative bacteria

The differences in ζ_{spy} activity observed between Gram-negative *E. coli* and Gram-positive *B. subtilis* [74] provoked the question, whether UNAG kinase activity has different effects in Gram-positive and Gram-negative bacteria and might therefore restrict epsilon/zeta systems to Gram-positive strains.

The major difference between Gram-positive and Gram-negative bacteria lies in the composition of their cell wall. While Gram-positive bacteria possess a thick peptidoglycan layer surrounding their cell membrane, the peptidoglycan layer of Gram-negative bacteria is much thinner and surrounded by a second, outer membrane [75]. Nevertheless, both bacterial classes depend on the rigidity of the peptidoglycan layer to preserve the cellular shape and integrity [76]. Furthermore, the basic structure of the glycan strands is conserved [77] and homologs of the enzyme MurA are found in Gram-positive and Gram-negative bacteria [78]. In addition, recombinant overexpression experiments of streptococcal zeta toxins in *E. coli* revealed a growth phenotype and lysis, indicating that UNAG kinase activity is detrimental to *E. coli* [53, 74, 79].

In addition, bioinformatics analysis identified a plethora of zeta toxins in all bacterial phyla [31], which indicates a general functionality of UNAG kinases within the bacterial domain.

1.4 Scope of this thesis

Although bioinformatics studies proof that epsilon/zeta homologs are present in all bacterial phyla [31], profound characterizations were restricted to the two streptococcal systems, ε/ζ_{spy} and PezAT. These experiments revealed that epsilon/zeta systems, although sharing the same mechanism, can show profound differences in binding affinities, regulation and cellular function. Therefore, the **objective** of this thesis was to **investigate novel epsilon/zeta TA systems in Gram-negative bacteria**.

A chromosomal zeta homolog from *E. coli* and the plasmidic epsilon/zeta systems from *N. gonorrhoeae* were chosen for thorough investigation.

1.4.1 *Escherichia coli* EzeT

E. coli is the prokaryotic model organism and has been extensively studied and manipulated. Naturally, this Gram-negative γ-proteobacterium is part of the human gut microbiota, but virulent strains can cause diarrhea, urinary tract infections and meningitis. On the *E. coli* chromosome, a gene with homology to zeta toxins (ECO103_4771 in *E. coli* O103:H2 12009) was identified and termed *ezeT*. Interestingly, no putative antitoxin gene was found in its vicinity. Closer investigation showed that the *ezeT* open reading frame (ORF) is unusually long with 308 amino acids compared to approximately 250 residues that form the zeta toxin core. While the 3'-end of *ezeT* is homologous to other zeta genes, its 5'-end is translated into approximately 80 additional residues. To explain the presence of an active UNAG kinase without an obvious inhibitor the **hypothesis** was proposed that **EzeT is an autoinhibited UNAG kinase in which toxin and antitoxin functionalities are linked.**

To test this hypothesis various microbiological and biochemical experiments were conducted which are described in the first section of the results. The experiments aimed to investigate expression and UNAG kinase activity of EzeT, and to characterize toxicity and inhibition by the two putative domains. Furthermore, experiments were performed to investigate transcriptional regulation, activation and possible functions of this TA system.

1.4.2 *Neisseria gonorrhoeae* epsilon/zeta

N. gonorrhoeae is a human pathogen that causes the sexually transmitted disease gonorrhea. Virulence factors, which contribute to pathogenicity, can be carried and transferred by mobile genetic elements like the gonococcal conjugative plasmid pLE2451 [80]. Sequencing of pLE2451 variants with or without a tetracycline resistance determinant, called *tetM*, identified two adjacent epsilon/zeta homologs on each of these plasmids [81]. In tetracycline-resistance plasmids, these

putative TA systems were denoted ε1/ζ1 and ε2/ζ2, while plasmids without the *tetM* determinant encode ε1/ζ1 and ε3/ζ3, the latter sharing 67 % nucleotide sequence identity with ε2/ζ2 [81].

These plasmids are the first described cases, where two putative epsilon/zeta TA systems reside on one plasmid. To get insights into the functional relevance of these two systems, the **aim** of this thesis was to **characterize the gonococcal epsilon/zeta homologs**, in order to **show that all three modules are functional TA systems comprising active UNAG kinases.**

Therefore, the ability of the recombinantly expressed epsilon/zeta complexes to phosphorylate UNAG was tested, as well as the interactions between the individual proteins. To measure the phosphorylation efficiency and nucleotide binding properties of ζ1, the wild-type protein was separated from the complex and subjected to biochemical and biophysical studies *in vitro*. Furthermore, the structure of the ε1/ζ1 complex was determined and the mechanism of inhibition by ε1 was investigated by measuring enzyme kinetics.

2 Materials and Methods

2.1 Materials

2.1.1 Standard chemicals

Manufacturer	Product
Carl Roth	chloramphenicol, kanamycin
GERBU Biotechnik	dithioerythritol (DTE), ethylenediaminetetraacetic acid (EDTA), glycerol, glycine, isopropyl β-*D*-1-thiogalactopyranoside (IPTG), urea
Merck	$CaCl_2$, HCl, NH_4Cl, $(NH_4)_2SO_4$, KCl, K_2HPO_4, KH_2PO_4, $MgCl_2$, $MgSO_4$, NaCl, Na_2HPO_4, NaH_2PO_4, NaOH, $Na_2S_2O_3$, ammonium acetate, ammonium peroxydisulfate (APS), bromphenol blue, isopropanol, methanol, potassium acetate, tri-sodium citrate, β-mercaptoethanol
Serva	*N*,*N*,*N*′,*N*′-tetramethylethane-1,2-diamine (TEMED), sodium dodecyl sulfate (SDS), Tween 20
Sigma-Aldrich	$MnCl_2$, $RbCl_2$, 2-(*N*-morpholino)ethanesulfonic acid (MES), 2-[4-(2-hydroxyethyl)piperazin-1-yl] ethanesulfonic acid (HEPES), acetic acid, acetone, acetonitrile, L-(+)-arabinose, ampicillin, boric acid, ethanol, imidazole, tris(hydroxymethyl)-aminomethan (Tris), **tri-sodium citrate**[1]
Fluka	**$(NH_4)_2SO_4$**[1]**,** D(+)-Glucose, tetracycline
Riedel-de Haën	formalin

[1]Bold chemicals were used for protein crystallization.

2.1.2 Microorganisms

Strain	Genotype	Manufacturer
DH5α	*E. coli* K-12 F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_K^-, m_K^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1	Life Technologies
BL21-CodonPlus(DE3)-RIL	*E. coli* B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte [*argU ileY leuW* Cam^r]	Stratagene
W	wild-type	DSM1116
OverExpress C43(DE3)	E. coli B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ *gal* λ(DE3)	Lucigen

2.1.3 Plasmids

Plasmid	Resistance	Manufacturer
pMK/RQ	Kanamycin	GeneArt
pET28b	Kanamycin	Novagen
pM11-1a	Kanamycin	Gunter Stier
pBAD/*Myc*-His A	Ampicillin	Invitrogen
pET21d	Ampicillin	Novagen
pETN-Strep	Kanamycin	Gunter Stier
pRedET	Tetracycline	Gene Bridges

2.1.4 Metal affinity resins and chromatography columns

Resin/column	Manufacturer
Ni-NTA agarose	Qiagen
MonoS 5/50 GL	GE Healthcare
MonoQ 5/50 GL	GE Healthcare
MonoQ 10/100 GL	GE Healthcare
Superdex 75 10/300 GL	GE Healthcare
Superdex 200 10/300 GL	GE Healthcare
Partisil-5 SAX RACII	Whatman

2.1.5 Crystallization screens

Screen	Manufacturer
JCSG Core I – IV Suite	Qiagen
AmSO4 Suite	Qiagen

2.1.6 Buffers

Buffer	Composition
TFB-1 buffer	30 mM Potassium actetate-HCl pH 5.8 50 mM $MnCl_2$ 100 mM RbCl, 10 mM $CaCl_2$ 15 % (*v/v*) Glycerol
TFB-2 buffer	10 mM MOPS-NaOH pH 7.0 10 mM RbCl 75 mM $CaCl_2$ 15 % (*v/v*) Glycerol

TBE buffer	89 mM Tris-HCl pH 8.0 89 mM Boric acid 2 mM EDTA
6× DNA loading buffer	30 % (*v/v*) Glycerol 0.25 % (*w/v*) Bromphenol blue 0.25 % (*w/v*) Xylene cyanol FF (Serva)
Separation gel buffer	1.5 M Tris-HCl pH 8.8 0.4 % (*w/v*) SDS
Stacking gel buffer	0.5 M Tris-HCl pH 6.8
SDS running buffer	25 mM Tris-HCl pH 8,0 200 mM Glycine 0.1 % (*w/v*) SDS
5× SDS loading buffer	50 mM Tris-HCl pH 7.0 10 % (*w/v*) SDS 0.1 % (*w/v*) Bromphenol blue 143 mM β-Mercaptoethanol 14 % (*w/v*) DTE 10 % (*v/v*) Glycerol
Fixing solution	50 % (*v/v*) Methanol 12 % (*v/v*) Acetic acid 0.05 % (*v/v*) Formalin
Staining solution	0.2 % (*w/v*) $AgCO_3$ 0.076 % (*v/v*) Formalin
Developing solution	6 % (*w/v*) Na_2CO_3 (Merck) 0.05 % (*v/v*) Formalin 0.0004 % (*w/v*) $Na_2S_2O_3$
Stop solution	50 % (*v/v*) Methanol 12 % (*v/v*) Acetic acid
Friendly transfer buffer	25 mM Tris-HCl pH 8.3 192 mM Glycine 5% (*v/v*) Isopropanol 0.1% (*w/v*) SDS
TBS	10 mM Tris-HCl pH 7.5 150 mM NaCl
TBS-T	25 mM Tris-HCl pH 7.4 137 mM NaCl 2.7 mM KCl 0.1 % (*v/v*) Tween 20
PBS	8 mM Na_2HPO_4 pH 7.4 2 mM KH_2PO_4 137 mM NaCl 2.7 mM KCl
PBS-T	8 mM Na_2HPO_4 pH 7.4 2 mM KH_2PO_4 137 mM NaCl 2.7 mM KCl 0.02 % (*v/v*)Tween 20

Conjugation buffer	20 mM NaH_2PO_4/Na_2HPO_4 pH 7.5 150 mM NaCl
Resuspension buffer	100 mM Tris-HCl pH 8.0 150 mM NaCl 1 mM DTE 2 mM EDTA
Buffer A1	50 mM Tris-HCl pH 8.0 150 mM NaCl 50 mM $(NH_4)_2SO_4$
Buffer A2	50 mM Tris-HCl pH 8.0 10 mM $(NH_4)_2SO_4$ 5 % (*w/v*) Glycerol 250 mM Imidazole 2 mM DTE
Buffer A3	50 mM Tris-HCl pH 8.0 10 mM $(NH_4)_2SO_4$ 5 % (*w/v*) Glycerol 0.5 mM EDTA 2 mM DTE
Buffer A4	50 mM Tris-HCl pH 9.0 2.5 mM $(NH_4)_2SO_4$ 5 % (*w/v*) Glycerol 0.5 mM EDTA 2 mM DTE
Buffer A5	50 mM Tris-HCl pH 9.0 10 mM $(NH_4)_2SO_4$ 5 % (*w/v*) Glycerol 0.5 mM EDTA 2 mM DTE 1 M KCl
Buffer A6	50 mM HEPES-NaOH pH 7.5 250 mM KCl 5 % (*w/v*) Glycerol 2 mM DTE
Buffer B1	50mM Tris pH 7.0 150mM NaCl 50mM $(NH_4)_2SO_4$
Buffer B2	50 mM MES-NaOH pH 6.0 50 mM NaCl 50 mM $(NH_4)_2SO_4$ 500 mM Imidazole 2mM DTE
Buffer B3a	50 mM MES-NaOH pH 6.0 50 mM $(NH_4)_2SO_4$ 0.5 mM EDTA 2 mM DTE
Buffer B3b	50 mM MES-NaOH pH 6.0 150 mM NaCl 50 mM $(NH_4)_2SO_4$ 2 mM DTE

Buffer B4	50 mM MES-NaOH pH 6.0 50 mM $(NH_4)_2SO_4$ 2 mM DTE
Buffer B5	50 mM MES-NaOH pH 6.0 50 mM $(NH_4)_2SO_4$ 2 mM DTE 1 M NaCl
Buffer B6	50 mM MES-NaOH pH 6.0 200 mM NaCl 2 mM DTE
MALS buffer	50 mM MES-NaOH pH 6.0 200 mM NaCl
Phosphorylation buffer	25 mM HEPES-NaOH pH 7.5 100 mM NaCl 5 mM $MgCl_2$
ATPase buffer	50 mM HEPES-NaOH pH 7.5 100 mM NaCl 14 mM KCl 10 mM $MgCl_2$ 1 mM EDTA 1 mM Phosphoenolpyruvate (Roche) 0.35 mM NADH (Roche)
Titration buffer	50 mM MES-NaOH pH 6.0 200 mM NaCl 10 mM $MgCl_2$

2.1.7 Growth media

Bacteria were grown in lysogeny broth (LB) medium [82] unless otherwise indicated. For agar plates, the medium was solidified with 15 g agar-agar (Carl Roth) per liter. Medium and agar plates were produced and autoclaved by the media kitchen of the Max Planck Institute for Medical Research.

LB medium	10 g/l bacto-tryptone (Becton Dickinson) 5 g/l bacto-yeast extract (Becton Dickinson) 10 g/l NaCl Final pH adjusted to 7.0 with NaOH

All media and plates were supplied with appropriate antibiotics to select for the different plasmid constructs (see section 2.1.3) in the following concentrations: 34 µg/ml chloramphenicol, 50 µg/ml kanamycin, 100 µg/ml ampicillin, 3 µg/ml tetracycline. A kanamycin concentration of 15 µg/ml was used for selection for the genomically integrated kanamycin resistance cassette.

To obtain selenomethionine labeled protein cells were grown in minimal medium substituted with SeMet based on Van Duyne *et al*. [83]:

SeMet medium

5.6 g/l K_2HPO_4
1.6 g/l KH_2PO_4
0.8 g/l $(NH_4)_2SO_4$
0.2 g/l tri-sodium citrate · $2H_2O$
0.08 g/l $MgSO_4 \cdot 7H_2O$
5.8 g/l D(+)-Glucose
40 mg/l Alanine
40 mg/l Arginine
40 mg/l Asparagine
40 mg/l Cysteine
40 mg/l Glutamic acid
40 mg/l Glycine
40 mg/l Histidine
40 mg/l Proline
40 mg/l Serine
40 mg/l Tryptophan
40 mg/l Tyrosine
40 mg/l Glutamine
40 mg/l Aspartic acid
0.1 g/l Isoleucine
0.1 g/l Leucine
0.1 g/l Lysine
0.1 g/l Phenylalanine
0.1 g/l Threonine
0.1 g/l Valine
40 mg/l Thiamine
40 mg/l Thymine
50 mg/l Selenomethionine (Acros Organics)
Final pH adjusted to 7.5 with NaOH

All amino acids as well as thiamine and thymine were purchased from Sigma-Aldrich.

Growth experiments with arabinose-induction were performed in glucose-free rich medium (RM) to prevent glucose-mediated repression of the pBAD promoter.

RM medium
- 6 g/l Na_2HPO_4
- 3 g/l KH_2PO_4
- 0.5 g/l NaCl
- 1 g/l NH_4Cl
- 20 g/l casein hydrolysate (Merck)
- 0.2 % (*v/v*) glycerol
- 1 mM $MgCl_2$
- Final pH adjusted to 7.4 with NaOH

2.2 Molecular biology

2.2.1 Transformation of chemically competent *Escherichia coli* cells

A single colony was inoculated into 200 ml LB-medium supplemented with the appropriate antibiotics. The culture was grown at 37 °C overnight and subsequently 1 ml was diluted into 500 ml fresh LB-medium supplemented with the appropriate antibiotics. The cell suspension was incubated at 37 °C until an OD_{600} of 0.5 was reached. 200 ml of culture were harvested by centrifugation (5,500 × g, 10 min, 4 °C) and washed with 50 ml ice cold TFB-1 buffer (the buffer compositions of all buffers can be found in section 2.1.6). After centrifugation the cell pellet was resuspended in 4 ml ice cold TFB-2 buffer and 50 µl aliquots were flash frozen in liquid nitrogen and stored at -80 °C until further use.

For chemical transformation, aliquots of competent cells were thawed on ice before adding 0.5 µl to 4 µl of plasmid DNA. The transformation mix was incubated on ice for 30 min, subjected to a heat shock at 42 °C for 45 s followed by a second incubation on ice for 2 min. 450 µl LB-medium were added to induce cell recovery and the cells were shaken at 750 rpm and 37 °C on a Thermomixer comfort (Eppendorf). After 1 h, cells were plated on pre-warmed LB-agar plates supplemented with the appropriate selection antibiotics and incubated at 37 °C overnight.

2.2.2 Transformation of electrocompetent *Escherichia coli* cells

A single colony was inoculated into 100 ml LB-medium supplemented with the appropriate antibiotics. The culture was grown overnight at 37 °C and subsequently 1 ml was diluted into 100 ml of fresh LB-medium supplemented with the appropriate antibiotics. Cells were grown at 37 °C to an OD_{600} of 0.6 and the cell suspension was harvested by centrifugation (3,320 × g, 4 °C,

10 min). The cells were washed twice in 10 ml ddH_2O and subsequently resuspended in 1 ml 10 % (*v/v*) glycerol. 100 µl aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

For electroporation, one aliquot of competent cells was thawed on ice and mixed with 1 µl to 4 µl plasmid DNA. Cells were directly transferred to a 2 mm electroporation cuvette (biolab products) and subjected to a short electric pulse (~ 5 ms) using an electro cell manipulator 600 (BTX). For general plasmid transformation, a voltage of 1350 V was applied at 50 µF and 128 Ω, transfer of DNA fragments for genomic integration was performed using the same voltage at 25 µF and 480 Ω. Afterwards, the cells were mixed immediately with 1 ml LB-medium, transferred to a microreaction tube and incubated on a Thermomixer comfort (Eppendorf) at 750 rpm and 37 °C. After 1 h, the mixture was plated on pre-warmed LB-agar plates supplemented with the appropriate selection antibiotics and incubated at 37 °C overnight.

2.2.3 Preparation of plasmid DNA

A single colony of *E. coli* DH5α was inoculated into 5 ml LB-medium supplemented with the appropriate antibiotics, grown at 37 °C overnight and harvested by centrifugation. Isolation of plasmid DNA was performed using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions.

The nucleotide concentration was determined spectroscopically by measuring the absorption at 260 nm ($A_{260\,nm}$ =1 ≙ 50 µg/µl) using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Correctness of the nucleotide sequence was confirmed by sequencing at Eurofins Genomics.

2.2.4 Agarose gel electrophoresis

DNA fragments were separated according to electrophoretic mobility in an agarose gel matrix. The agarose gels were prepared by mixing 1 % (*w/v*) universal agarose (Bio&SELL) with TBE buffer. The suspension was heated to dissolve the agarose, cast and 1 µl GelRed Nucleic Acid Gel Stain 10,000× (Biotium) was added per 70 ml gel solution. Prior to loading, DNA loading buffer was added to the DNA samples. In addition, dependent on the size of the desired fragment 100 bp or 1 kb DNA ladders (NEB) were loaded as a standard. 2 µl of a 1:5 dilution were loaded per lane. Electrophoresis was performed in TBE buffer at a voltage of 100 V for 30 min to 90 min.

DNA fragments were visualized by GelRed fluorescence under UV irradiation. If required, DNA fragments of the appropriate size were excised and isolated from the gel slices using the QIAquick Gel Extraction Kit (Qiagen) following the manufacturer's instructions.

2.2.5 Polymerase chain reaction

Amplification of DNA fragments was performed by polymerase chain reaction (PCR) using Vent DNA Polymerase (NEB) in a TPersonal thermocycler (Biometra). Primers were synthesized by Eurofins Genomics and the reactions components were purchased from New England Biolabs (NEB) except for formamide (Sigma-Aldrich).

Primer annealing temperature and elongation time of the three-step protocol (see below) were adjusted to the melting temperature of the primer and the desired product length, respectively. For a typical reaction the following concentrations and cycling protocol were used:

Component	Volume (µl)	Final concentration
10x Thermopol buffer[1]	5	1x
Deoxynucleotide (dNTP) solution mix[2]	1	200 µM
Forward primer (100 µM)	0.5	1 µM
Reverse primer (100 µM)	0.5	1 µM
$MgSO_4$ (100 mM)	1.25	2.5 mM
Formamide	3.5	7% (*v/v*)
Vent DNA polymerase (2000 units/ml)	0.5	0.02 units/µl
DNA template	x	1-4 ng/µl
Add H_2O to a final volume of 50 µl		

[1] 200 mM Tris-HCl pH 8.8, 100 mM $(NH_4)_2SO_4$, 100 mM KCl, 20 mM $MgSO_4$, 1% (*v/v*) Triton® X-100
[2] dATP, dCTP, dGTP, dTTP (10 mM each)

Step	Temperature (°C)	Time (s)	Cycles
Initial denaturation	95	300	1
Denaturation	95	30	30
Primer annealing	45-65	30	
Elongation	72	60 s per 1000 bp	
Final elongation	72	600	1
Hold	4	-	

PCR products were either purified using the QIAQuick PCR Purification Kit (Qiagen) or separated on a preparative 1 % (*w/v*) agarose gel (see section 2.2.4) and extracted from excised gel slices.

2.2.6 Restriction digestion

Restriction digestions of PCR products or DNA plasmids were performed with restriction enzymes purchased from New England Biolabs (NEB) using the buffer provided. The following reaction mixture was incubated at 37 ºC over night:

Component	Volume (µl)	Final concentration
10x CutSmart buffer[1]	5	1x
Restriction enzyme 1 (5-20 units/µl)	1	0.4-0.1 units/µl
Restriction enzyme 2 (5-20 units/µl)	1	0.4-0.1 units/µl
DNA	x	0.02 µg/µl
Add H_2O to a final volume of 50 µl		

[1] 500 mM potassium acetate pH 7.9, 200 mM Tris-acetate, 100 mM magnesium acetate, 1 mg/ml bovine serum albumin

Samples of preparative restriction digestions were either purified using the QIAQuick PCR Purification Kit (Qiagen) or separated on a 1 % (*w/v*) agarose gel (see section 2.2.4) and extracted from excised gel slices.

2.2.7 DNA ligation

Digested DNA inserts were ligated into a vector with compatible cohesive or blunt ends using the Rapid DNA Ligation Kit (Fermentas). The vector and insert concentrations were adjusted to a final vector amount of 50 ng and a molar vector:insert ratio of 1:3. The reaction mixture was incubated at room temperature for 15 min.

Component	Volume (µl)	Final concentration
5x Rapid ligation buffer	2	1x
Rapid T4 DNA Ligase (5 units/µl)	1	0.1 units/ µl
Vector DNA	x	1 ng/µl
Insert DNA	x	1-4 ng/µl
Add H_2O to a final volume of 10 µl		

4 µl of the ligation mixture were transformed into *E. coli* DH5α cells and single colonies were inoculated into 5 ml LB medium for the preparation of plasmid DNA (see section 2.2.3). To verify successful ligation of insert DNA into plasmid DNA, a restriction digest (see section 2.2.6) was performed in a volume of 10 µl and incubated at 37 ºC for 2 h. Digestion mixtures were loaded on an analytical 1 % (*w/v*) agarose gel (see section 2.2.4) for separation and fragments were visualized by GelRed fluorescence.

2.2.8 Site directed mutagenesis

Nucleotide alterations, insertions or deletions were introduced by site-directed mutagenesis according to the QuikChange protocol (Agilent Technologies). Primers were synthesized by

Eurofins Genomics and the reaction components including Herculase II Fusion DNA polymerase were purchased from Agilent Technologies and deoxynucleotides were purchased from NEB.

For a typical reaction the following protocol was used:

Component	Volume (µl)	Final concentration
5x Herculase II reaction buffer[1]	10	1x
Deoxynucleotide (dNTP) solution mix (10 mM each)[2]	1	200 µM
Forward primer (10 µM)	1.25	250 nM
Reverse primer (10 µM)	1.25	250 nM
DMSO	2	4 % (*v/v*)
Herculase II Fusion DNA polymerase	1	N.D.[1]
DNA template	x	0.1-1 ng/µl
Add H_2O to a final volume of 50 µl		

[1] Buffer composition and Polymerase concentration not stated by manufacturer. N.D. = not determined
[2] dATP, dCTP, dGTP, dTTP

Step	Temperature (°C)	Time (s)	Cycles
Initial denaturation	95	120	1
Denaturation	95	20	30
Primer annealing	55	20	
Elongation	72	300	
Final elongation	72	600	1
Hold	4	-	

Subsequently, methylated template DNA was removed by digestion with *Dpn*I (20 units, New England Biolabs) at 37 °C for 2 h and 4 µl of the digestion mixture were used for transformation into DH5α cells. Plasmid DNA was isolated from cultures grown from a single colony (see section 2.2.3) and successful mutation was verified by sequencing.

2.2.9 Cloning of *ezeT*

The nucleotide sequence of the genomic region of open reading frame (ORF) ECO103_4771 including 250 base pairs upstream of the start codon was synthetically manufactured (GeneArt, Life Technologies). To facilitate subcloning and protein purification the locus was designed to include DNA sequences encoding for a N-terminal hexahistidine (His_6)-tag, a tobacco etch virus (TEV) protease recognition site and an *Nco*I restriction site overlapping with the *ezeT* start codon. The synthetic construct was ligated into pMK/RQ by the manufacturer (pMK/RQ(His-EzeT)).

Cloning of full-length ezeT

For recombinant overexpression, the *ezeT* ORF was subcloned into an IPTG-inducible expression vector. Therefore, the ORF was PCR amplified from pMK/RQ(His-EzeT) with the primers T7_for[1] (priming within the promoter region) and EzeT_3'NotI_His_r, which primes upstream of the stop codon and introduces a *Not*I recognition site at the 3'-end of the ORF. After restriction digestion with *Nco*I and *Not*I the DNA fragment was ligated into pET28b (pET28b(EzeT)), thereby exchanging the coding sequence of an N- for a C-terminal His_6-tag. Note, that expression tests with full-length EzeT showed that the solubility of N-terminally tagged protein is significantly decreased when compared to C-terminally tagged protein.

In addition, EzeT was expressed from an arabinose-inducible vector (pBAD(EzeT)). Therefore, the *ezeT* ORF was PCR-amplified from pET28b(EzeT) using pET_bone_for (priming upstream of the promoter region) and EzeT_3'HindIII_r (introducing a *Hind*III recognition site downstream of the stop codon), digested with *Nco*I and *Hind*III and ligated into pBAD/Myc-His A.

Cloning of truncated ezeT

The DNA fragment containing the ORF encoding for the N-terminal domain of EzeT (residues 1 to 82) was PCR-amplified from pET28b(EzeT) with T7_for (priming within the promoter region) and EzeT_83stop_HindIII_r (which exchanges the codon for residue 83 to a stop codon and introduces a *Hind*III restriction site). After restriction digestion with *Nco*I and *Hind*III this fragment was ligated into pM11-1a (pET(EzeT(1-82))).

The N-terminal truncation construct EzeTΔN83 and the two longer constructs EzeTΔN62 and EzeTΔN36 were created by PCR-amplification of the corresponding fragment from pET28b(EzeT) with EzeT_5'NcoI_start82_f, EzeT_5'NcoI_start62_f or EzeT_5'NcoI_start36_f, respectively, and EzeT_3'SalI_r (which introduces a *Sal*I site downstream of the stop codon), digestion with *Nco*I and *Sal*I and ligation into pET28b. EzeT_5'NcoI_start82_f introduces an *Nco*I site overlapping with a start codon at position Thr82 and mutates the His83 codon to an Asp codon (pET(EzeTΔN83)). EzeT_5'NcoI_start62_f mutates the codon for Gln62 to an *Nco*I site overlapping with a start codon (pET(EzeTΔN62)) and EzeT_5'NcoI_start36_f mutates the codon for Ser36 to an *Nco*I site overlapping with a start codon (pET(EzeTΔN36)).

To be able to induce expression of the N- and C-terminal domains with two different inducers, the ORF encoding for EzeTΔN83 was excised from pET28b(EzeTΔN83) with *Nco*I and *Hind*III and cloned into pBAD/Myc-His A (pBAD(EzeTΔN83)).

[1] All primer sequences can be found in the Appendix (section 7.1).

Cloning of catalytically inactive variants

The Walker A lysine (K121) was exchanged for an alanine by site directed mutagenesis (see 2.2.8) in the vector constructs pET28b(EzeT), pBAD(EzeT) and pET28b(EzeTΔN83) using EzeT_K121A_f and EzeT_K121A_r (pET28b(EzeT_K121A), pBAD(EzeT_K121A) and pET28b(EzeTΔN83_K121A)). Similarly, codons in the N-terminal domain were mutated to alanines by two consecutive site-directed mutagenesis steps on pET28b(EzeT) using the primer pairs EzeT_M5A_D8A_f / EzeT_M5A_D8A_r and EzeT_N12A_Q16A_f / EzeT_N12A_Q16A_r (pET28b(EzeT_M5A_D8A_N12A_Q16A)).

Cloning of Strep-ezeT-His

A plasmid encoding N-terminally Strep-tagged and C-terminally His-tagged EzeT (pET28b(Strep-EzeT-His)) was generated by restriction digestion of pET28b(EzeT) with *Nco*I and *Blp*I to obtain a DNA fragment encoding for C-terminally His_6-tagged EzeT and ligated into similarly digested pETN-Strep (constructed and provided by Gunter Stier, University of Heidelberg, Germany), providing the coding sequence for an N-terminal Strep-tag.

2.2.10 Cloning of *ε1/ζ1*

The plasmid DNA of pEP5289 from *Neisseria gonorrhoeae* encoding for ε1/ζ1 and ε2/ζ2 was kindly provided by Chris van der Does from the Max Planck Institute for Terrestrial Microbiology in Marburg, Germany. The bicistronic operon was PCR-amplified from this plasmid and cloned into inducible expression vectors for recombinant protein expression in *E. coli*.

Cloning of the ε1/ζ1_K115A *bicistron*

To facilitate cloning of a potentially toxic protein, a two-step protocol was applied to introduce a Walker A lysine (K115) to alanine mutation during subcloning of the bicistronic *ε1/ζ1* operon. In the first step, the operon was amplified from pEP5289 with E1_5'NdeI_f and Z1_3'HindIII_r to introduce an *Nde*I restriction site on the 5'-end of the ε1 ORF. The PCR product was digested with *Nde*I and *Eco*RI (present in the ζ1 ORF close to the Walker A lysine) to obtain the 5'-fragment of the operon and ligated into pET28b (pET28b(His-TEV-epsilon)). Thereby, the coding sequence for a TEV-cleavable His_6-tag was introduced at the 5'-end of *ε1*. In a second step, the 3'-half of the operon was amplified from pEP5289 with Z1_EcoRI_K115A_f and Z1_3'HindIII_r. The former primer contained an *Eco*RI recognition sequence and altered the Walker A lysine codon into an alanine codon. The latter primer provided a *Hind*III recognition site to allow for cloning into pET28b(His-TEV-epsilon) (obtained in the first cloning step) with *Eco*RI and *Hind*III (pET28b(ε1/ζ1_K115A)). Cloning of this construct was performed by Marion Gradl.

Cloning of the individual genes

The two ORFs of the bicistronic *ε1/ζ1_K115A* operon were cloned individually into pET28b by Roman Sakson. To obtain pET(ε1), the *ε1* ORF was amplified from pET28b(ε1/ζ1_K115A) by PCR using pETbone_for and E1_3'HindIII_r (which introduces a *Hind*III restriction site downstream of the stop codon of *ε1*), digested with *Nco*I and *Hind*III and ligated into pET28b. For pET28b(ζ1_K115A-His) the *ζ1_K115A* ORF was amplified from pET28b(ε1/ζ1_K115A) by PCR using Z1_5'NcoI_f (introducing an *Nco*I restriction site upstream of the start codon of *ζ1*) and Z1_3'XhoI_His_r. The latter primer omits the stop codon of the ORF and allows the fusion of the coding sequence for a His_6-tag to the gene product upon ligation into pET28b.

For pull-down experiments, a DNA fragment encoding ζ1_K115A without an affinity tag was ligated into pET28b after PCR amplification from pET28b(ε1/ζ1_K115A) using Z1_5'NcoI_f and T7_rev and restriction digestion with *Nco*I and *Xho*I (pET28b(ζ1_K115A)). In addition, the DNA fragment encoding His_6-tagged ε1 was cloned into pET21d (pET21d(ε1)) and an arabinose-inducible pBAD vector (pBAD(ε1)). The former construct was prepared by restriction digestion of pET21d and pET28b(ε1) with *Nco*I and *Not*I and subsequent ligation of the two fragments. Similarly, the latter construct was prepared by restriction digestion of pBAD/*Myc*-His A and pET28b(ε1) with *Nco*I and *Hind*III and subsequent ligation.

Cloning of wild-type ε1/ζ1

To provide sufficient amounts of ε1 in the competent cells used for cloning of wild-type *ε1/ζ1*, ε1 was provided from an arabinose-inducible vector. Therefore, DH5α cells were transformed with pBAD(ε1) (see former paragraph), a single colony was picked and the cells were grown at 37 ºC in LB medium supplemented with ampicillin until an OD_{600} of 0.2 was reached. Then, protein expression was induced by addition of L-(+)-arabinose to a final concentration of 0.15 % (*w/v*). Cells were further grown at 37 ºC until an OD_{600} of 0.6 was reached and electrocompetent cells were prepared as described in section 2.2.2.

These cells were transformed with 2 µl of a site directed mutagenesis reaction on pET28b(ε1/ζ1_K115A) using Z1_A115K_f and Z1_A115K_r to yield pET28b(ε1/ζ1). The transformation mixture was then plated on LB-agar plates containing ampicillin, kanamycin and 0.02 % (*w/v*) L-(+)-arabinose. Plasmid DNA isolated from single colonies contained a mixture of pET28b(ε1/ζ1) and pBAD(ε1) introduced during preparation of competent cells. To obtain pure pET28b(ε1/ζ1), this plasmid DNA mixture was retransformed into plasmid-free DH5α cells, which were then plated on LB-agar plates containing only kanamycin. After incubation, two morphologically different colony types were visible: colonies with normal size and unusually small colonies, probably due to low levels of ζ1 kinase activity. These smaller colonies were not able to grow on ampicillin and plasmid DNA preparations from these colonies contained exclusively pET28b(ε1/ζ1).

2.2.11 Cloning of *ε2/ζ2*

Similar to the cloning of *ε1/ζ1* the bicistron of *ε2/ζ2* was amplified by PCR from pEP5289 and cloned into inducible expression vectors for recombinant expression in *E. coli*.

Cloning of ε2/ζ2_K46S

Analogous to the preparation of *ε1/ζ1_K115A* (compare 2.2.10, *Cloning of the inactive complex*) a two-step protocol was used to introduce a Walker A lysine (K46) to serine mutation during cloning of *ε2/ζ2* (pET28b(ε2/ζ2_K46S)).

Exchange of the lysine for a serine codon was coupled to the introduction of a *Bam*HI restriction site. The mutation was introduced by PCR-amplification from pEP5289 with the primer pairs E2_5'NdeI_f (introduces an *Nde*I restriction site at the start codon) / Z2_K46S_r (introduces the mutation) and Z2_K46S_f (introduces the mutation) / Z2_3'XhoI_r (introduces an *Xho*I restriction site downstream of the stop codon) to obtain fragments for the 5'- and 3'-fragement of the operon, respectively. A second PCR was performed on these two DNA fragments using E2_5'NdeI_f and Z2_3'XhoI_r. The resulting fragment was digested with *Nde*I and *Xho*I and ligated into pET28b. Thereby the sequence coding for a thrombin-cleavable His_6-tag is fused 5' to the *ε2* ORF.

Cloning of the individual genes

A DNA fragment encoding N-terminally His_6-tagged ε2 was cloned into pET28b by PCR-amplification of the ORF from pEP5289 using E2_5'NdeI_f and E2_3'HindIII_r (introduces a *Hind*III restriction site downstream of the ε2 stop codon), restriction digestion with *Nde*I and *Hind*III and ligation into pET28b (pET28b(ε2)). For pull-down experiments, the *ε2-His* ORF was subsequently cloned into pET21d by digestion of pET28b(ε2) with *Nco*I and *Not*I and ligation to yield pET21d(ε2).

In addition, a DNA fragment encoding ζ2_K46A without an affinity tag was cloned into pET28b by Florence Jungblut. The ORF was amplified by PCR from (pET28b(ε2/ζ2_K46S)) using Z2_5'NcoI_f (introduces an *Nco*I restriction site at the start codon of ζ2 and changes the codon for isoleucine 2 to a valine) and Z2_3'XhoI_r, digested with *Nco*I and *Xho*I and ligated into pET28b (pET28b(ζ2_K46A)).

Cloning of wild-type ε2/ζ2

To provide sufficient amounts of ε2 in the competent cells used for cloning of wild-type *ε2/ζ2*, a cloning strategy similar to cloning of *ε1/ζ1* was applied (see former paragraph), but ε2 was provided from an IPTG-inducible vector (pET21d(ε2)) that confers kanamycin resistance. OverExpress C43(DE3) cells were transformed with pET21d(ε2) (see former paragraph), a single colony was picked and the cells were grown at 37 ºC in LB medium supplemented with ampicillin

until an OD_{600} of 0.6 was reached. Then protein expression was induced by addition of 0.5 mM IPTG and cells were further grown at 37 ºC. As IPTG induction led to an expression phenotype the culture was grown over night. The cultures were diluted to OD_{600}=0.1 in fresh LB medium and grown at 37 ºC until an OD_{600} of 0.6 was reached. Then, chemically competent cells were prepared from these cultures as described in 2.2.1.

These competent cells were transformed with 5 µl of a site directed mutagenesis reaction on pET28b(ε2/ζ2_K46A) using Z2_S46K_f and Z2_S46K_r to yield pET28b(ε2/ζ2) and plated on LB-agar plates containing kanamycin. However, separation of pET28b(ε2/ζ2) from pET21d(ε2) as reported for pET28b(ε1/ζ1) (section 2.2.10, *Cloning of wild-type complex*) was not successful.

2.2.12 Cloning of ε3/ζ3

The nucleotide sequence of the *ε3/ζ3* bicistron from the *N. gonorrhoeae* plasmid pEP5233 was ordered from GeneArt (Life Technologies). To allow subcloning, an *Nco*I restriction site was added to the 5'-end of *ε3* and a *Not*I restriction site to the 3'-end of *ζ3* and the stop codon of *ζ3* was omitted. The synthetic construct was ligated into pMK/RQ by the manufacturer (pMK/RQ(ε3/ζ3)). All cloning steps regarding *ε3/ζ3* were performed by Florence Jungblut.

Cloning of the ε3/ζ3 *bicistron*

In contrast to cloning of *ε1/ζ1* and *ε2/ζ2*, cloning of wild-type *ε3/ζ3* did not require additional expression of ε3 in the competent cells. The bicistron was excised from pMK/RQ(ε3/ζ3) by restriction digestion with *Nco*I and *Not*I and cloned into pET28b (pET28b(ε3/ζ3)) whereby the coding sequence for a non-cleavable His_6-tag was added at the 3'-end of the *ζ3* ORF.

A plasmid encoding the inactive variant (pET28b(ε3/ζ3_K46A)) was generated by side directed mutagenesis using the primer pair Z3_K46A_f / Z3_K46A_r.

Cloning of the individual genes

A DNA fragment encoding N-terminally His_6-tagged ε3 was cloned into pET21d by PCR-amplification of the ORF from pET28b(ε3/ζ3) using E3_5'NdeI_f (introduces an *Nde*I restriction site at the ε3 start codon) and E3_3'HindIII_r (introduces a *Hind*III restriction site downstream of the ε3 stop codon), restriction digestion with *Nde*I and *Hind*III and ligation into pET21d (pET21d(ε3)).

The ORF encoding ζ3_K46A was cloned into pET28b by PCR-amplification from pET28b(ε3/ζ3_K46A) using Z3_5'NcoI_f (introduces an *Nco*I restriction site at the ζ3 start codon and changes S2 to a glycine) and Z3_404stop_HindIII_r (re-introduces the original ζ3 stop codon followed by a *Hind*III restriction site), restriction digestion with *Nco*I and *Hind*III and ligation into pET28b (pET28b(ζ3_K46A)).

2.2.13 Cloning of GFP reporter plasmids and promoter mutants

For reporter assays, the T7 promoter region was removed from pMK/RQ(EzeT). Therefore, plasmid DNA was digested with the *Fsp*I and *Asc*I restriction enzymes, single stranded overhangs were filled in with the Klenow fragment and the blunted ends were religated. Subsequently, an *Nde*I restriction site was introduced at the start codon upstream to the His_6-tag coding sequence by site directed mutagenesis (see 2.2.8) using EzeT_5'NdeI_His_f and EzeT_5'NdeI_His_r. The coding region for superfolder GFP (sfGFP) [84] was amplified from the pET expression vector of sfGFP using T7_for (priming in the promoter region) and GFP_3'EcoRI_r (priming downstream of the T7 terminator region of pET28b). Both, the PCR product for sfGFP and the modified pMK/RQ(EzeT) vector DNA, were digested with the *Nde*I and *Eco*RI restriction enzymes and ligated (pREP(sfGFP)).

All alterations of the *ezeT* promoter region were introduced by site directed mutagenesis (see 2.2.8) on pREP(sfGFP). The Shine-Dalgarno sequence in *ezeT* (pREP(delSD) was mutated using GFP_delSD_f and GFP_delSD_r. The region downstream of the Pribnow-Schaller box (pREP(delPSdown) was mutated using GFP_delPSdown_f and GFP_delPSdown_f. The Pribnow-Schaller box in *ezeT* (pREP(delPS) was mutated using GFP_delPS_f and GFP_delPS_r. The putative -35 region (pREP(s70)) was altered to the -35 consensus sequence of the σ^{70} factor using GFP_s70_f and GFP_s70_r or a random sequence (pREP(del35)) using GFP_del35_f and GFP_del35_r. The region upstream of the putative -35 region (pREP(delUPE)) was deleted after introduction of *Hind*III restriction sites in the vector backbone (GFP_delUPE_1_f / GFP_delUPE_1_r) and immediately upstream to the putative -35 region (GFP_delUPE_2_f / GFP_delUPE_2_r) into pREP(sfGFP) by site directed mutagenesis, subsequent excision by restriction digestion and re-ligation of the residual vector. Similarly, the region upstream of the Pribnow-Schaller box (pREP(delprom)) was deleted by introduction of a second *Hind*III restriction site immediately upstream to the putative -10 region (GFP_delprom_f / GFP_delprom_r), restriction digestion and re-ligation.

2.2.14 Genomic integration

For the isolation of endogenous EzeT, an epitope-tagged version was integrated into the *E. coli* W genome into the *ezeT* locus to give *E. coli* W(EzeT-HA). To this end, the coding sequence of the His_6-tag of pET(EzeT) was exchanged for one of an influenza hemagglutinin (HA)-tag in two consecutive site-directed mutagenesis reactions (pET28b(EzeT-HA)) using the primer pairs HA-tag_1_f / HA-tag_1_r and HA-tag_2_f / HA-tag_2_r. This ORF was PCR-amplified using EzeT_5'NdeI_f (introduces an *Nde*I restriction site at the start codon) and HA-tag_3'BamHI_r (introduces a *Bam*HI restriction site downstream of the stop codon), digested with *Nde*I and *Bam*HI and cloned into pREP(sfGFP) from which the *gfp* ORF was excised by *Nde*I and

*Bam*HI restriction digestion (pREP(EzeT-HA)). A kanamycin/neomycin resistance cassette was PCR-amplified from a FRT-PGK-gb2-neo-FRT fragment (*E. coli* gene deletion kit, Gene Bridges) using Neo_5'MfeI_f and Neo_3'Mfe_r. This PCR product was digested using *Mfe*I and inserted between EzeT-HA and the residual EzeT coding fragment of pREP(EzeT-HA). For that purpose, the plasmid was digested with *Eco*RI to create compatible single stranded overhangs for ligation yielding pINT(*ezeT::ezeT-HA,neo*).

Similarly, a second integration construct was prepared by insertion of the kanamycin/neomycin resistance cassette into pREP(sfGFP) yielding pINT(*ezeT::GFP, neo*). Integration of this DNA fragment into *E. coli* W creates an *ezeT* knock-out strain (*E. coli* W *ezeT*$^-$) in which the *ezeT* ORF is replaced by *gfp*. Both constructs contain the *ezeT* promoter region and the C-terminal part of EzeT that served as homology arms for homologous recombination into *E. coli* W.

Genomic integration was performed according to the Red/ET recombination protocol from Gene Bridges. Recombination competent *E. coli* W were prepared by electroporation with the thermosensitive plasmid pRedET (Gene Bridges). This plasmid encodes for the λ phage recombination proteins Redα, Redβ and Redγ and is maintained at 30 ºC but cured at 37 ºC to allow temporal control of recombination. Therefore, plasmid transformation and incubation of the plates were performed at 30 ºC. Single colonies were used to inoculate LB-medium containing tetracycline and cultures were grown at 30 ºC. The expression of the recombination proteins was induced at OD_{600}=0.3 with 0.3 % (*w/v*) L-(+)-arabinose and after 1 h cells were harvested and made electrocompetent.

The integration constructs were PCR-amplified from pINT(*ezeT::ezeT-HA,neo*) using Int_f and M13_rev and purified following the QIAquick PCR purification kit (Qiagen) protocol. Cells were transformed with 800 ng purified, linear PCR product, incubated at 37 ºC for 3 h to allow recombination to occur and streaked on LB-agar plates containing 15 µg/ml kanamycin. Successful integration was confirmed by sequencing of the PCR-amplified genomic DNA region using the Int_control_f (priming upstream of the *ezeT* locus in *lysC*) and Int_control_f (priming downstream of *ezeT* in the hypothetical gene *ECW_m4384*).

2.3 Microbiology

2.3.1 Growth kinetics

To investigate the toxicity *in vivo*, *E. coli* BL21-CodonPlus(DE3)-RIL were transformed with plasmids expressing full-length, truncated or mutated EzeT. Single colonies were used to inoculate 100 ml LB medium supplemented with kanamycin and chloramphenicol, and were grown over night in 250 ml baffled flasks at 37 °C. The pre-cultures were diluted into 100 ml fresh

medium to an OD_{600} of 0.05 and cultures were grown to OD_{600}=0.1 in 250 ml baffled flasks at various temperatures before protein expression was induced with 1 mM IPTG. Cell proliferation was monitored regularly by measuring the OD_{600} and by determination of colony forming units (CFUs).

For CFU determination, ten-fold serial dilutions in LB medium were prepared and 10 µl of the appropriate dilutions were plated on pre-warmed LB-agar plates supplemented with kanamycin. The plates were incubated at the respective incubation temperature of the expression culture until visible colonies formed. All experiments were performed in triplicates.

Experiments studying the inhibition by the N-terminal domain *in trans* were performed in glucose-free RM medium to prevent catabolite repression of the arabinose-inducible promoter. *E. coli* BL21-CodonPlus(DE3)-RIL cells were co-transformed with pET28b(EzeT(1-82)) and pBAD(EzeTΔN83) and the expression culture was grown at 37 °C to OD_{600}=0.1-0.2, cooled to 16 °C and expression of the C-terminal domain from pBAD(EzeTΔN83) was induced with 0.2 % (*w/v*) L-(+)-arabinose after 1 h. Expression of the N-terminal domain from pET28b(EzeT(1-82)) was induced by addition of 1 mM IPTG at various time points.

2.3.2 Propidium iodide influx

The break-down of the osmotic barrier was followed using the membrane-impermeable DNA-intercalator propidium iodide. As described in section 2.3.1 pre-cultures of *E. coli* BL21-CodonPlus(DE3)-RIL cells carrying plasmids encoding full-length, truncated or mutated EzeT were prepared. These were diluted into 50 ml fresh medium to an OD_{600} of 0.025 and grown at 16 °C. When an OD_{600} ~ 0.1 was reached, these cultures were mixed with an equal volume of LB medium supplemented with 0.1 mg/ml propidium iodide (Sigma-Aldrich) and 100 µg/ml ampicillin or 2 mM IPTG. 200 µl of these mixtures were transferred into a 96-well plate (Corning 3651, Corning) and OD_{600} and propidium iodide fluorescence (excitation: 520 nm, emission: 620 nm) were monitored over 5 h using a Varioscan Flash Multimode Reader (Thermo Scientific).

2.3.3 GFP reporter assay

E. coli DH5α or *E. coli* W cells were transformed with plasmids expressing sfGFP (pREP(sfGFP)) under control of various versions of the endogenous promoter region of the *ezeT* operon from *E. coli* W (see section 2.2.13). A single colony was used to inoculate 5 ml LB medium and grown at 37 °C overnight. Subsequently, cell cultures were diluted to an OD_{600} of 0.1 and 200 µl aliquots of each culture were further incubated in 96-well plates (Corning 3651, Corning) under constant shaking at 37 °C until stationary phase was reached. GFP fluorescence (excitation:

485 nm, emission: 510 nm) was monitored using a Varioscan Flash Multimode Reader (Thermo Scientific). Untransformed *E. coli* DH5α were used for background correction.

To test for autorepression by EzeT, *E. coli* DH5α or *E. coli* W cells were co-transformed with an additional pBAD(EzeT) plasmid and grown in RM medium. Expression was induced by addition of 0.2 % (*w/v*) L-(+)-arabinose. As a control, DH5α cells only transformed with pREP(sfGFP) were similarly exposed to arabinose.

To measure the influence of different stress conditions, *E. coli* DH5α, *E. coli* BL21-CodonPlus(DE3)-RIL, *E. coli* W or *E. coli* W *ezeT*$^-$ cells were transformed with pREP(sfGFP). A single colony was used to inoculate 5 ml LB medium and grown at 37 °C overnight. Subsequently, cell cultures were diluted to an OD_{600} of 0.1 and 200 µl aliquots of each culture were further incubated in 96-well plates (Corning 3651, Corning) under constant shaking at 37 °C. GFP fluorescence (excitation: 485 nm, emission: 510 nm) was monitored using a Varioscan Flash Multimode Reader (Thermo Scientific).

The cultures in the individual wells were treated with different chemicals ($Fe(III)Cl_3$, DTE, sodium cacodylate, sodium citrate, hydrogen peroxide, chloramphenicol, bipyridyl), high temperature (95 ºC for 1 min, or 45 ºC for 60 min) or UV irradiation, or grown in medium without nutrients or with glycerol as a carbon source.

2.3.4 Microscopy

Phase contrast images were taken from *E. coli* BL21-CodonPlus(DE3)-RIL cells transformed with plasmids expressing EzeTΔN83 and EzeTΔN83. Expression cultures were grown as described in section 2.3.1 at 20 ºC. Phase contrast and fluorescence microscopy was performed on an Axiovert 135 microscope (Zeiss) using a 100 x oil immersion objective lens (Plan-Neofluar, N.A. 1.3, Zeiss). Images were captured with a ProgRes C3 CCD camera (Jenoptic). Life/dead staining was performed using the LIVE/DEAD BacLight bacterial viability kit (Molecular Probes). 20 µl cell suspension was mixed with 0.4 µl of a solution containing 1.67 mM SYTO 9 and 9.99 mM propidium iodide. SYTO 9 (live) fluorescence was visualized with a 450 nm to 490 nm excitation filter and a 520 nm long pass emission filter, while propidium iodide (dead) fluorescence was visualized using 545 nm band pass excitation filter and a 590 nm long pass emission filter.

Electron microscopic images were taken from *E. coli* BL21-CodonPlus(DE3)-RIL cells transformed with plasmids expressing full-length or mutated EzeT (EzeT_M5A_D8A_N12A_Q16A). Expression cultures were grown as described in section 2.3.1 and protein expression was induced at 16 ºC for 16 h. Samples were harvested by mild centrifugation (300 × g, 4 ºC, 10 min) and resuspended in 100 µl LB medium. Cells were further

concentrate by centrifugation (300 × g, 4 ºC, 10 min) and the pellet was resuspended in 50 µl LB medium containing 40 % (*w/v*) bovine serum albumin (Sigma-Aldrich).

Cryo-electron microscopy and high-pressure freezing was performed by Ulrike Mersdorf. Cells were cryopreserved using a Leica EM HPM100 high pressure freezer and transferred under liquid nitrogen into a Leica AFS system for freeze-substitution. Samples were incubated in acetone containing 0.5 % (*w/v*) osmium tetroxide (Serva), 0.5 % (*w/v*) tannic acid (Sigma-Aldrich), 0.5 % (*v/v*) glutaraldehyde (Sigma-Aldrich) and 0.2 % (*w/v*) uranyl acetate (Ted Pella) for 50 h using a temperature gradient from -90 ºC to -30 ºC. Subsequently, the samples were washed three times for 1 h with 100 % acetone, equilibrated to room temperature and gradually infiltrated with Epon embedding medium according to Luft (glycidyl ether 100, dodecenyl succinic anhydride, methyl nadic anhydride, dimethylbenzylamine; all Serva) [85]. Samples were polymerized for 2 days at 60 ºC and 50 nm thin sections were obtained using a diamond knife (DiATOME) and an Ultracut E ultramicrotome (Reichert-Jung). The sections were stained for 15 min in 7 % (*w/v*) aqueous uranyl acetate (Ted Pella) and 5 min in Reynold's lead citrate (26.6 g/l lead nitrate (Sigma-Aldrich), 35.2 g/l sodium citrate (Merck), 160 mM sodium hydroxide (Merck)). Image data were recorded on a Zeiss EM 912 transmission electron microscope at 120 kV. Adobe Photoshop CS5 was used to compile the transmission electron microscopy images.

2.3.5 Plasmid stabilization assay

To test whether EzeT is able to stabilize plasmids, the *ezeT* ORF was cloned into a derivative of the enteric low-copy number plasmid R1, pOU82 [86]. pOU82 contains an ampicillin resistance gene, the *lacZYA* operon and the origin of replication of R1. Replication of pOU82 is controlled in two ways. On the one hand, the replication control genes are under control of the λPr promoter that is regulated by the temperature-sensitive lambda repressor λcI 857. This results in a low copy number at a growth temperature of 30 ºC and run-away replication at 42 ºC. On the other hand, the replication control region does not encode for CopA. CopA is an antisense RNA that negatively regulates RepA and thereby downregulates replication. Overexpression of CopA from an IPTG-inducible promoter on a second plasmid, pKG339 [86], prevents plasmid replication. pKG339 confers tetracycline resistance. pOU82 and pKG339 were generously provided by Kenn Gerdes, University of Copenhagen, Denmark.

ezeT was cloned into pOU82 under control of its endogenous promoter or the constitutive *lacI* promoter. *E. coli* W *ezeT*$^-$ were co-transformed with one of these two vectors or the empty vector together with pKG339 and plated on LB-agar plates containing ampicillin and tetracycline. Cells were grown in 5 ml RM medium supplemented with ampicillin and tetracycline at 30 ºC over night. This pre-culture was diluted into fresh medium and grown at 30 ºC. At OD_{600}=0.5 the cultures were diluted again to OD_{600}=0.01 in RM medium containing tetracycline and 2 mM IPTG

to induce pOU82 replication arrest. The OD_{600} values and CFUs were measured every hour over the course of 6 h. CFUs were plated on LB-agar plates with and without ampicillin to distinguish between plasmid-containing and plasmid-free cells. In addition, blue-white screening allows to determine the percentage of colonies that retained pOU82.

2.3.6 Plaque assay

To determine whether *ezeT* contributes to phage resistance, *E. coli* cells transformed with pBAD, pBAD(EzeT) or pBAD(EzeT_K121A) were grown in the presence of T4 phages. T4 phage solutions were kindly provided by Elisabeth Hartmann and used in hundred-fold dilution series. For the assay, *E. coli* pre-cultures were grown at 37 ºC in LB medium supplemented with ampicillin and diluted to reach OD_{600}=0.2. After 30 min, 0.2 % (*w/v*) L-(+)-arabinose was added. 30 min after induction (OD_{600}≈0.6) cells were diluted to OD_{600}=0.4 in LB medium and 100 µl of these cultures were mixed with 100 µl phage solution and 3 ml prewarmed (~ 45 ºC) LB medium containing 0.75 % (*w/v*) agar-agar (Carl Roth). These reaction mixtures were poured immediately onto LB-agar plates and incubated at 37 ºC over night. Bacteria-devoid regions (plaques) were counted on the next day.

2.3.7 Persister assay

The ability to form persister cells was tested by antibiotic treatment with ciprofloxacin, gentamicin and tetracycline. *E. coli* W, *E. coli* W *ezeT*$^-$ and *E. coli* W *ezeT*$^-$ harboring pBAD(EzeT) were grown in LB medium over night. This pre-culture was diluted to OD_{600}=0.05 and incubated at 37 ºC. 0.1 % (*w/v*) L-(+)-arabinose was added after 3 h and cells were further incubated. 1 h later, 2 ml of culture were harvested by centrifugation (6,000 × g, 4 ºC, 5 min), washed with 0.85 % (*w/v*) NaCl and resuspended in 0.85 % (*w/v*) NaCl to an OD_{600} of 1.5. Ciprofloxacin (Sigma-Aldrich), gentamicin (Life Technologies) or tetracycline (Sigma-Aldrich) were added to final concentrations of 0.1 µg/ml, 4 µg/ml or 10 µg/ml, respectively. The cells were incubated at 37 ºC and serial dilutions were prepared after 4 h and 24 h and plated on LB-agar plates for CFU determination.

2.3.8 Biofilm assay

The ability to form biofilms was assessed by crystal violet staining. *E. coli* W, *E. coli* W *ezeT*$^-$ and *E. coli* W *ezeT*$^-$ harboring pBAD(EzeT), pBAD(EzeTΔN83), pBAD(EzeTΔN83_K121A) or an empty pBAD vector were grown in LB medium over night. This pre-culture was diluted to OD_{600}=0.05 in LB medium containing 0.1 % (*w/v*) L-(+)-arabinose. Per culture, 200 µl each were aliquoted into four wells of a sterile, transparent 96-well flat-bottom

polystyrene plate (Costar tissue-culture treated cell culture plates, Corning, or Nunc MaxiSorp, affymetrix). These plates were incubated without shaking at 30 ºC for 48 h. To reduce evaporation, surrounding wells were filled with medium and the plate was closed with a lid.

Subsequently, the OD_{600} was measured and the biofilm formation was quantified. For staining of organic material on the walls of the wells, the remaining LB medium and non-attached cells were removed. The wells were washed three times with 0.85 % (*w/v*) NaCl and left open to dry. 125 µl of 0.1 % (*w/v*) crystal violet (Sigma-Aldrich) were applied and the wells were washed again three times with 0.85 % (*w/v*) NaCl and dried. The bound stain was dissolved in 200 µl ethanol/acetone (80:20, *v/v*) and this solution was transferred to a fresh 96-well plate for absorption measurements at 590 nm using a Varioscan Flash Multimode Reader (Thermo Scientific).

2.4 Biochemistry

2.4.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated according to their electrophoretic mobility by discontinuous SDS-PAGE as initially described by Laemmli [87]. Gels were cast and run using the Mini-PROTEAN system (Bio-Rad). The acrylamide concentration in the separation gel was adjusted to the molecular weight of the protein of interest for sufficient separation. 15 %, 17 % and 20 % (*w/v*) acrylamide solutions for separation gels (see below) were prepared and polymerization of the acrylamide was initiated by adding APS and TEMED to a final concentration of 0.1 % (*w/v*) and 1 % (*v/v*), respectively. After polymerization, a 5 % (*w/v*) stacking gel was cast on top in a similar fashion.

Gel type	H_2O	Buffer	Acrylamide solution[1]
15 % separation gel	7.3 ml	7.5 ml separation gel buffer	15 ml
17 % separation gel	5.3 ml	7.5 ml separation gel buffer	17 ml
20 % separation gel	2.3 ml	7.5 ml separation gel buffer	20 ml
5 % stacking gel	8.9 ml	3.8 ml stacking gel buffer	2.3 ml

[1] Rotiphorese®Gel 30 (37.5:1): 30 % Acrylamide/Bisacrylamide in the ratio 37.5:1 (Carl Roth)

Prior to loading, SDS loading buffer (the buffer compositions of all buffers can be found in section 2.1.6) was added to the samples and the mixture was heated at 95 °C for 5 min for protein denaturation. Dependent on the protein concentrations, 5 µl to 15 µl of the samples were loaded per lane. In addition, 5 µl of the Low Range SDS-PAGE standard (Bio-Rad) diluted 1:20 in SDS loading buffer were loaded for size estimation. Electrophoresis was performed in SDS running buffer at 35 mA for about 45 min. Proteins were visualized by Coomassie staining with Instant*Blue* (Expedeon).

2.4.2 Silver staining

Silver staining [88] was used to detect low amounts of proteins in polyacrylamide gels. Proteins species were separated by SDS-PAGE and fixed in the gel with fixing solution at RT for 2 h. After washing three times for 20 min with 35 % (*v/v*) ethanol, a solution containing 0.02 % (*w/v*) of the sensitizing agent $Na_2S_2O_3$ was applied for 2 min followed by washing with H_2O (3 x 5 min). Silver ions were allowed to bind to the proteins by incubating in staining solution for 20 min. After three 1 min washes in H_2O, developing solution was added to induce silver reduction. Development was stopped by adding stop solution as bands became visible. The gel was stored at 4 °C in 1 % (*v/v*) acetic acid.

To identify proteins by mass spectrometry after silver staining, the reduced silver ions were removed from the sample [89]. Gel slices containing the protein of interest were cut from the gel and soaked for 5 min in 0.75 ml 15 mM $K_3[Fe(CN)_6]$ (Sigma-Aldrich), 50 mM $Na_2S_2O_3$. After washing three times with 1 ml H_2O, the gel pieces were further analyzed by mass spectrometry (see section 2.4.5).

2.4.3 Western Blot

Western Blotting was used to detect proteins labeled with His_6-, HA- and Strep-epitope tags. Protein species were separated by SDS-PAGE and then transferred onto a PVDF membrane (Porablot PVDF, Macherey-Nagel). Prior to blotting the membrane was activated with methanol and then transferred into Friendly transfer buffer together with the polyacrylamide gel and two appropriately sized sheets of gel blotting paper (GB005, Whatman). Between the electrodes of a Semidry blotter SD1 (cti) a stack of (anode side) gel blotting paper – PVDF membrane – polyacrylamide gel – gel blotting paper (cathode side) was assembled. The blotter was connected to a PowerPack Basic (Bio-Rad) and transfer was performed at 60 mA for 1 h.

To prevent unspecific antibody binding, the hydrophobic membrane was incubated in 5 % (*w/v*) nonfat dried milk in TBS-T at 4 ºC over night. Subsequently, the membrane was washed twice in TBS-T for 10 min and incubated with the primary antibody at room temperature for 2 h. The membrane was washed again in TBS-T before adding the secondary antibody at room temperature for 1 h. Next, the membrane was washed three times in TBS-T and three times in TBS for 10 min each. Horse-radish peroxidase coupled antibodies were used for protein detection by chemiluminescence. To start the reaction the membrane was incubated for 2 min in freshly prepared detection reagent (Amersham ECL Prime Western Blotting Detection Reagent, GE Healthcare). The solution was removed and the membrane was either exposed to a sheet of BioMax light film (Kodak) for 30 - 300 s which was developed using an Optimax TR X-ray film processor (Protec) or chemiluminescence was directly detected with a ChemiDoc MP Imaging System (Bio-Rad).

Both primary and secondary antibodies were diluted in 5 % (*w/v*) nonfat dried milk in TBS-T and stored at -20 ºC between repeated uses. For the detection of the HA epitope the rabbit anti-HA antibody (abcam ab9110, dilution 1: 4000) was used as primary antibody. The horseradish peroxidase-conjugated anti-rabbit IgG antibody (Delta Biolabs, dilution 1: 5000) was used as secondary antibody. For detection of His_6-tagged proteins the Penta·His HRP conjugate (Qiagen, mouse monoclonal IgG1 horse-radish peroxidase conjugate, final dilution 1:2000) was used. This primary antibody is already coupled to HRP and therefore the secondary antibody step was omitted and antibody incubation and washing was directly followed by chemiluminescence detection.

For detection of proteins with a Strep-tag the protocol had to be modified as the streptactin used for detection also binds biotinylated proteins. Therefore, after protein transfer the membrane was incubated in PBS containing 3 % (*w/v*) bovine serum albumin (Sigma-Aldrich) and 0.5 % (*v/v*) Tween 20 at 4 ºC over night. The next day the membrane was washed twice in PBS-T for 10 min and then biotinylated proteins were blocked by avidin by washing with 0.5 ml chicken egg white in PBS-T at room temperature for 10 min. After washing in PBS-T the membrane was incubated with Strep-Tactin-HRP (iba, dilution 1:100000 in PBS-T) at room temperature for 1 h. Next, the membrane was washed three times in PBS-T and three times in PBS for 10 min each and chemiluminescence detection was performed as described above.

2.4.4 Determination of protein concentration

Protein concentration was determined by measuring the absorbance of the sample at 280 nm using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). At this wavelength the absorption of a protein is dominated by the amino acids tryptophan, tyrosine and cysteine [90]. Based on the amino acid sequence the molar extinction coefficient of a linear polypeptide chain at 280 nm ($\varepsilon_{P,\,280\,nm}$) can be calculated [91]. Using the online tool ProtParam [92] (http://web.expasy.org/protparam/) the following extinction coefficients were calculated for each construct:

Table 2.1: Molar extinction coefficient and molecular weight of purified proteins

Protein	Extinction coefficient at 280 nm ($M^{-1}cm^{-1}$)	Molecular weight (Da)
EzeT-CHis	25,900	35,806
ε1/ζ1	27,390	51,902
ε1-His	8,480	9,583
ζ1	20,400	44,618
ε2/ζ2	31,860	54,468
ε2/ζ2-His	31,860	56,351
ε3/ζ3	28,880	54,454

Protein concentrations (c_p) were calculated according to the law of Lambert-Beer:

$$c_P[g/l] = \frac{A_{P,280\,nm} \cdot MW}{\varepsilon_{P,280\,nm} \cdot d} \qquad \textbf{Equation 1}$$

Where $A_{P,\,280\,nm}$ is the absorbance of the protein at 280 nm at path length d and $\varepsilon_{P,\,280\,nm}$ and MW are the molar extinction coefficient and molecular weight of the protein as given in Table 2.1.

2.4.5 Mass spectrometry

Peptide mass fingerprinting [93] for the identification of proteins by mass spectrometry was performed by Melanie Müller and Marion Gradl.

Proteins were separated on an SDS-PAGE gel and protein containing gel slices were subjected to in-gel digestion by trypsin (proteomics grade, Sigma-Aldrich). The mass of the eluted peptides was determined by MALDI-TOF using a Shimadzu Axima Performance mass spectrometer (Shimadzu/Kratos Analytical) in reflectron mode. The combination of characteristic peptide masses was used for protein identification using Mascot PMF (Matrix Science) by comparison against a library of *E. coli* and recombinant proteins.

2.4.6 Immunoprecipitation

Prior to immunoprecipitation, 5 µg antibody (rat anti-HA high affinity antibody 3F10, Roche) was chemically cross-linked to 50 µl protein G-Dynabeads (Life Technologies) using 5 mM bis(sulfosuccinimidyl)suberate (BS3, Thermo Scientific) in conjugation buffer at ambient temperature. The reaction was quenched after 1 h by adding Tris-HCl pH 7.5 to a final concentration of 50 mM. The beads were thoroughly washed with PBS-T to remove non-cross-linked antibodies.

E. coli W (control) and *E. coli* W(EzeT-HA) were grown at 37 °C to OD_{600}=1, 1 l of culture was harvested by centrifugation and incubated in 10 ml resuspension buffer in the presence of 10 mg/ml lysozyme at 4 °C for 30 min. After addition of 0.2 % (*v/v*) Tween 20, the cells were lysed in an M-110S microfluidizer (Microfluidics) and cell debris was cleared by centrifugation. The cross-linked antibody was added to the supernatant and incubated under constant rotation at 4 °C for 2 h. Beads were retrieved using a magnetic field and washed several times with PBS-T. Bound proteins were eluted by denaturation with 8 M urea.

For co-precipitation of bound proteins, the antibody was not cross-linked to the beads. Instead, 10 µg antibody (rat anti-HA high affinity antibody (3F10, Roche)) was added to the cleared lysate and allowed to bind under constant rotation at 4 °C for 1 h before 200 µl protein G-Dynabeads (Life Technologies) were added and incubated for additionally 1 h. After washing, the proteins were eluted in 40 µl 5x SDS loading buffer.

The elutions were separated by SDS PAGE (see 2.4.1) and protein expression was analyzed by Western blotting using an anti-HA antibody (abcam ab9110, dilution 1: 4000) as primary antibody and a horseradish peroxidase-conjugated anti-rabbit IgG antibody (Delta Biolabs, dilution 1: 5000) as secondary antibody. In addition, gels from co-immunoprecipitation were stained by silver staining and potential binding partners were identified by mass spectrometry.

2.4.7 Recombinant expression of native proteins

Proteins were recombinantly overexpressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells, which carry a λ DE3 construct encoding IPTG-inducible T7 polymerase, were transformed with the expression constructs. Cells were grown in LB medium supplemented with the appropriate selection antibiotic and chloramphenicol to stabilize the pRIL plasmid. Freshly transformed cells were used to inoculate a 100 ml pre-culture which was grown at 37 °C for 5 h to 6 h. 10 ml of this pre-culture were used for inoculation per liter of medium and incubated under constant shaking in baffled flasks at 37 °C until an OD_{600} of 1 (EzeT) or 0.6 (gonococcal proteins) was reached. At this time point, the temperature was reduced to 16 ºC (EzeT) or 18 ºC (gonococcal proteins) and expression was induced by addition of 1 mM (EzeT) or 0.5 mM (gonococcal proteins) IPTG. The next day, cells were harvested by centrifugation at 5,000 × g and 4 ºC for 15 min and the cell pellet was resuspended in lysis buffer (for EzeT buffer A1, for the gonococcal proteins buffer B1; the buffer compositions of all buffers can be found in section 2.1.6). The suspension was either immediately subjected to protein purification or flash frozen in liquid nitrogen and stored at -80 °C.

2.4.8 Recombinant expression of selenomethionine labeled protein

Expression of selenomethionine labeled ε1/ζ1 was performed according to the protocol from Van Duyne *et al*. [83]. The protocol for overexpression and harvest of the cells was identical to the native protein expression (see former section), except that LB medium was replaced with SeMet medium (see section 2.1.7) containing selenomethionine. As *E. coli* BL21-CodonPlus(DE3)-RIL are prototrophic for all amino acids including methionine, the medium contains an excess of amino acids which suppress the bacterial methionine biosynthesis by negative feedback inhibition.

2.4.9 Cell disruption and lysate clarification

Cell lysis and all following steps of protein purification and handling were performed at 4 °C and with pre-chilled buffers. Prior to cell lysis, β-mercaptoethanol was added to the resuspended cells to a final concentration of 5 mM. Cells walls were disrupted by 20 min sonication using a W-450D Sonifier (Branson) at 50 % amplitude and cycles of 0.5 s sonication followed by 0.5 s recovery. Subsequently, bacterial lysates were clarified by centrifugation at 48,000 × g and 4 °C for

1 h. To analyze protein solubility, samples of pellets and supernatant were taken. The pellets were resuspended in an equal volume of 8 M urea for solubilization of insoluble proteins.

2.4.10 Purification of recombinant EzeT

The clarified lysate from 2 l expression culture resuspended in 40 ml buffer A1 was loaded on 2 ml Ni-NTA agarose (Qiagen) equilibrated with buffer A1. After washing with 20 ml buffer A1, the bound proteins were eluted with 20 ml buffer A2 into a reservoir containing 25 ml buffer A3 and dialyzed over night against 1 l buffer A4 using a 14,000 Da molecular weight cut-off (MWCO) dialysis tube (Visking, Medicell). The next day, the dialyzed protein was loaded onto a MonoQ 10/100 GL column (GE Healthcare) equilibrated with buffer A4 using an ÄKTApurifier system (GE Healthcare). Proteins were eluted with a linear gradient of 7.5 column volumes (CV) to buffer A5. Eluting proteins were collected in 0.8 ml fractions in tubes prefilled with 0.2 ml buffer A5. Fractions were analyzed on a 15 % (*w/v*) SDS-PAGE gel and EzeT containing fractions were pooled and concentrated using an Amicon-Ultra centrifugal filter unit (10,000 MWCO, Millipore). The concentrated protein was applied to a Superdex75 10/300 GL column (GE Healthcare) equilibrated with buffer A6. Protein purity was verified by SDS-PAGE and EzeT containing fractions were pooled and concentrated using an Amicon-Ultra centrifugal filter unit (10,000 MWCO, Millipore). The purified protein proved to be instable at higher concentrations (> ~1 mg/ml) and could not be stored for extended periods of time.

The doubly-tagged Strep-EzeT-His was purified according to the same protocol, except that the final size exclusion step was omitted and that 0.5 ml Ni-NTA agarose were used for 6 l of expression culture to account for the decreased solubility of this protein variant.

2.4.11 Purification of recombinant epsilon/zeta complexes from *N. gonorrhoeae*

For ε1/ζ1, ε2/ζ2 and ε3/ζ3 complex purification, the clarified lysate from 4 l expression culture resuspended in 80 ml buffer B1 was loaded on 1 mL Ni-NTA agarose (Qiagen) equilibrated with buffer B1. After washing with 10 ml buffer B1, the bound proteins were eluted with 10 ml buffer B2. The eluate was collected in 10 fractions which were loaded on a 20 % (*w/v*) SDS-PAGE gel. Protein containing fractions were pooled and dialyzed against 1 l buffer B3a over night. During dialysis, the His_6-tag of epsilon was cleaved off by protease digest depending on the construct. Undigested protein and the protease were removed by re-chromatography on Ni-NTA agarose (Qiagen). Hereby, the His_6-tag of ε1 was cleaved off by TEV digest, while ε2 contained a thrombin recognition site. Cleavage by thrombin required the addition of 2 mM $CaCl_2$ during dialysis. ζ3

contained an un-cleavable C-terminal His_6-tag and therefore no protease digestion and re-chromatography was performed.

The conductivity of the sample was adjusted to less than 10 millisiemens prior to loading onto a MonoS 10/100 GL column (GE Healthcare) equilibrated with buffer B4 using an ÄKTApurifier system (GE Healthcare). Proteins were eluted with a linear gradient of 30 CV to buffer B5. 1 ml fractions were collected and analyzed on a 20 % (*w/v*) SDS-PAGE gel and epsilon/zeta complex containing fractions were pooled and concentrated using a Vivaspin centrifugal filter unit (10,000 MWCO, Sartorius stedim biotech). The concentrated proteins were applied to a Superdex75 10/300 GL column (GE Healthcare) equilibrated with buffer B6. Protein purity and complex composition was verified by SDS-PAGE. Epsilon/zeta complex containing fractions were pooled and concentrated using a Vivaspin centrifugal filter unit (10,000 MWCO, Sartorius stedim biotech) and aliquots were flash frozen in liquid nitrogen and stored at -80 ºC.

Wtε/1ζ1, ε1/ζ1_K115A, wtε2/ζ2 and wtε3/ζ3 were purified according to the same protocol. SeMet-labeled ε1/ζ1_K115A was purified similar to the native proteins, except that in each buffer the concentration of the reducing agent was increased to 5 mM DTE.

For initial crystallization trials of the ε1/ζ1 complex, the His_6-tag was not proteolytically removed from ε1. Therefore, TEV protease was omitted from the dialysis and the dialyzed proteins were loaded onto the MonoS column immediately.

2.4.12 Purification of recombinant ε1

Purification of His_6-tagged ε1 was performed similar to purification of the complexes (see section 2.4.11). Due to decreased protein stability at low ionic strength, dialysis was performed against 1 l buffer B3b. Dialyzed protein was concentrated using a Vivaspin centrifugal filter unit (3,000 MWCO, Sartorius stedim biotech) and loaded onto a MonoS 10/100 GL column (GE Healthcare) without prior dilution. After elution, protein containing fractions were pooled and concentrated before loading on a Superdex75 10/300 GL column (GE Healthcare) equilibrated with buffer B6. Protein purity was verified using 20 % SDS-PAGE gels and ε1 containing fractions were pooled and concentrated to 4.7 mg/ml (558 µM) using a Vivaspin centrifugal filter unit (3,000 MWCO, Sartorius stedim biotech).

2.4.13 Purification of recombinant wild-type ζ1

Wild-type ζ1 was isolated from the ε1/ζ1 complex by complex separation and subsequent purification using a protocol modified from section 2.4.11. The complex was purified using metal affinity chromatography and cation exchange chromatography as described before. After elution from the MonoS column, the pooled protein fractions were re-loaded onto 1 ml Ni-NTA agarose

(Qiagen) equilibrated with buffer B1. The flow-through from loading as well as from the following wash with 10 ml buffer B5 was pooled and concentrated. The protein solution was applied to a Superdex75 10/300 GL column (GE Healthcare) equilibrated with buffer B6. However, free ζ1 tended to stick to the column material and had to be removed by washing the column with free ε1.

Protein purity was verified by SDS-PAGE and Western Blotting to detect remnants of His_6-tagged ε1. ζ1 containing fractions were pooled and concentrated to 25.4 mg/ml (570 µM) using a Vivaspin centrifugal filter unit (10,000 MWCO, Sartorius stedim biotech) and aliquots were flash frozen in liquid nitrogen and stored at -80 ºC.

2.4.14 Interaction of gonococcal epsilon and zeta proteins

E. coli BL21-CodonPlus(DE3)-RIL were co-transformed with two expression plasmids, a pET21d plasmid carrying a His_6-tagged epsilon (pET21d(ε1), pET21d(ε2) or pET21d(ε3)) and a pET28b plasmid carrying a untagged zeta (pET28b(ζ1), pET28b(ζ2) or pET28b(ζ3)) and plated on LB-agar plates containing both ampicillin and kanamycin. 100 ml pre-cultures were grown at 37 ºC in LB medium supplemented with ampicillin and kanamycin over night. The next morning, the cultures were diluted 1:100 into 2 l fresh medium and cells were grown at 37 ºC before protein expression was induced with 0.5 mM IPTG at OD_{600}=0.6 and the cultures were further incubated at 20 ºC.

The next morning, cultures were harvested by centrifugation (5,000 × g, 4 ºC, 15 min), resuspended in 30 ml buffer B1 and lysed as described in section 2.4.9. The cleared lysate was applied to 0.2 ml Ni-NTA agarose (Qiagen) equilibrated with buffer B1. After washing with 10 ml buffer B1, bound proteins were eluted with 2 ml buffer B2. Samples of the individual purification steps were analyzed using 20 % (*w/v*) SDS-PAGE gels.

Interaction studies were performed by Florence Jungblut.

2.4.15 Low-molecular-weight-metabolite extraction

Low-molecular-weight-metabolite extraction and analysis was performed similarly as described by Mutschler *et al.* [53]. *E. coli* BL21-CodonPlus(DE3)-RIL cells were transformed with plasmids expressing either full-length, truncated or inactive EzeT. Single colonies were used to inoculate 100 ml LB medium and grown over night at 37 ºC. 20 ml of this pre-culture were used to inoculate 2 l of LB-medium supplemented with kanamycin and chloramphenicol and the cultures were grown in unbaffled flasks at 37 °C. At OD_{600}=0.2, cell cultures were cooled to 16 °C and protein expression was induced by addition of 1 mM IPTG at OD_{600}=0.6. After 2 h of incubation, cells were harvested by centrifugation at 5,000 × g and 4 ºC for 15 min. The cell pellets were resuspended in 20 ml ice-cold 80 % (*v/v*) aqueous acetonitrile and incubated for 30 min on ice with

regular agitation. Subsequently, the insoluble material was removed by centrifugation at 20,000 × g and 4 ºC for 30 min and the supernatant was flash frozen in liquid nitrogen and stored at -80 °C.

For high-performance liquid chromatography (HPLC) analysis the solvent was evaporated by vacuum concentration using a Speed Vac Concentrator (Bachofer) and the samples resuspended in 1.5 ml deionized H_2O. Insoluble compounds were removed by high-speed centrifugation at 20,000 × g and 4 ºC for 30 min and filtration with Ultrafree-MC centrifugal filter units (0.22 µm cut-off, Millipore). The absorption at 260 nm was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies) and adjusted to a value of 10 AU. 30 µL of sample was applied to a Partisil-5 SAX RACII column (Whatman) equilibrated with 5 mM KH_2PO_4. Bound metabolites were eluted with a linear gradient of 36 CV to 500 mM KH_2PO_4 at a flow rate of 1 mL/min using a Prominence liquid chromatography system (Shimadzu). The absorbance at 260 nm was monitored. Under these conditions, UNAG elutes at 12 ml, UNAG-3P at 22.5 ml and ATP at 48 ml, respectively.

2.4.16 EzeT stability in *E. coli* crude extract

Crude extracts were obtained from 150 ml stationary phase *E. coli* W culture. The culture was grown in LB medium at 37 ºC over night and harvested by centrifugation (5,000 × g, 4 ºC, 15 min). The cell pellet was resuspended in 20 ml buffer A1 and lysed by passing it three times through an M-110S microfluidizer (Microfluidics). 450 µl of the resulting crude extract was supplemented with an ATP regeneration system [1 mM ATP (Jena Bioscience), 2 µl pyruvate kinase / lactate dehydrogenase (Pyruvate Kinase/Lactic Dehydrogenase enzymes from rabbit muscle (682 U/ml / 990 U/ml), Sigma-Aldrich), 0.8 mM phosphoenolpyruvate (Roche)] and Strep-EzeT-His to a final concentration of 0.05 mg/ml. The mixture was incubated at 30 ºC and 20 µl samples were taken in regular intervals. Samples were separated on a 17 % (*w/v*) SDS-PAGE gel and protein amount and electrophoretic mobility were monitored by Western Blot against the N-terminal Strep-tag and the C-terminal His_6-tag.

2.5 Biophysics

2.5.1 CD spectroscopy

CD spectroscopy measurements were performed with a J-810 CD spectropolarimeter (Jasco) in quartz glass cuvettes (Hellma) with a path length of 1 mm. Protein melting curves were measured by recording the temperature-dependent change in ellipticity θ of the sample. Unfolding of 10 µM EzeT in 200 µl buffer A6 was recorded by thermal denaturation from 10 ºC to 70 ºC at a wavelength of 222 nm and a heat rate of 1 ºC/min.

The melting temperature was determined in Prism (GraphPad) by assuming a two state unfolding process [94]. Data were fitted to the following equation:

$$\text{signal} = \frac{(aN + bN \times T) + (aD + bD \times T) \times \exp\left[\frac{\Delta H}{R} \times \left(\frac{1}{T_m} - \frac{1}{T}\right)\right]}{1 + \exp\left[\frac{\Delta H}{R} \times \left(\frac{1}{T_m} - \frac{1}{T}\right)\right]} \qquad \textbf{Equation 2}$$

Where aN and bN are the baseline intercept and slope for the native state of the protein and aD and bD are those for the denatured state. R is the gas constant (8.31 J K^{-1} mol^{-1}) and ΔH is the van't Hoff enthalpy as a function of the temperature T. T_m is the melting temperature which is defined as the temperature at the transition midpoint where 50 % of the protein is unfolded.

2.5.2 Dynamic light scattering

Dynamic light scattering was used to determine the hydrodynamic radius of epsilon-zeta complexes in order to estimate their oligomeric state. Samples with a protein concentration of approximately 5 mg/ml in buffer B6 were measured in a quartz cuvette with a path length of 1.5 cm using a Viscotec 802DLS instrument (Malvern Instruments). Measurements were performed at a temperature of 20 ºC. The scattering signal was collected 20 times over a duration of 5 s each and averaged. The resulting autocorrelation curve was analyzed using the supplied OmniSIZE 3.0 software (Malvern Instruments).

2.5.3 Static light scattering

By measuring the scattering intensity at different angles during static light scattering the molar mass of the epsilon-zeta complexes was determined. Therefore, the samples were separated on a Superdex 200 10/300 GL column connected to a Waters HPLC system (Milford). After size exclusion chromatography, the protein species were analyzed with a refractive index detector 2414, a photodiode array detector 2996 and a multiangle light scattering detector (Dawn Heleos, Wyatt).

40 µl of sample with a protein concentration of approximately 5 mg/ml were injected at a flow rate of 0.5 ml/min onto the column equilibrated with MALS buffer. A gel-filtration standard (Bio-Rad, Order no. 151-1901) containing 5 mg/ml protein was included for calibration. Multi-angle light scattering data were analyzed using the ASTRA software provided by the manufacturer (Wyatt).

2.5.4 Chromatographic UNAG phosphorylation assay

UNAG phosphorylation by UNAG kinases was investigated *in vitro* following a protocol modified from Mutschler *et al.* [53]. 3 µM protein in phosphorylation buffer was incubated with 1 mM ATP (Jena Bioscience) and 0.25 mM UNAG (Sigma-Aldrich) at 25 °C. After 3 h and 24 h of incubation, 200 µl aliquots were removed and frozen. For analysis, the samples were thawed, diluted 1:2 with deionized H_2O and centrifuged at 21,000 × g and 4 ºC for 5 min. The mixture was applied to a MonoQ 5/50 GL (GE Healthcare) equilibrated with deionized H_2O and bound species were eluted in a linear gradient of 32.5 CV to 1 M ammonium acetate pH 8.0.

2.5.5 Steady-state UNAG phosphorylation assay

The ATPase activity of free wtζ1 was measured by Annika Grimmer using a coupled colorimetric assay [95]. In this assay the ATP hydrolysis, that is required for UNAG phosphorylation, is coupled to the oxidation of NADH, which can be followed spectroscopically as a decrease in absorption at 340 nm. By measuring this decrease in absorption over time, the rate of NADH production and accordingly the rate of ATP consumption by the UNAG kinase can be calculated:

$$v = \frac{\Delta Abs_{340nm}/\Delta t}{\varepsilon_{340nm} \cdot c_p \cdot d} \qquad \textbf{\textit{Equation 3}}$$

The apparent turnover rate v at a given substrate concentration is dependent on the wtζ1 concentration c_p (0.2 µM), the optical path length d of the cuvette (1 cm) and ε_{340nm} the extinction coefficient of NADH at 340 nm (6220 M^{-1} cm^{-1}). The change in absorption at 340 nm ΔAbs_{340nm} over the time interval Δt is given by the slope of a linear fit to the absorption graph.

Absorption at 340 nm was measured in quartz cuvettes (Hellma) at 25 °C using a V-650 UV/VIS spectrophotometer (JASCO). Data were collected and analyzed with the Spectra-Manager software provided by the manufacturer (JASCO).

Reactions were prepared with 100 µl of a reaction mix containing ATPase buffer supplemented with 7 U pyruvate kinase / 10 U lactate dehydrogenase (Pyruvate Kinase/Lactic Dehydrogenase enzymes from rabbit muscle, Sigma-Aldrich) and varying concentrations of ATP and UNAG and incubated until a stable baseline was reached. The reaction was started by addition of wtζ1 to a final concentration of 0.2 µM.

To determine the Michaelis-Menten parameters, turnover rates were determined at a constant ATP concentration of 10 mM while varying the UNAG concentration or at a constant UNAG concentration of 10 mM while varying the ATP concentration. The rates v were plotted against the

substrate concentration S and fitted with the GraFit data analysis software (Erithacus) using the Michaelis-Menten equation:

$$v = \frac{v_{max} \cdot S}{S + K_m} \qquad \textbf{Equation 4}$$

Thereby, the apparent Michaelis-Menten constant K_m and the maximal apparent reaction velocity v_{max} were determined. Inhibition of ζ1 by ε1 was determined by measuring the Michaelis-Menten parameters for ATP at different ε1 concentrations (0 µM, 0.1 µM, 0.2 µM).

2.5.6 Fluorescence equilibrium titration

Fluorescence measurements were performed at 25 ºC in 600 µl titration buffer in quartz cuvettes (Hellma) using a FP-8500 fluorescence spectrometer (JASCO). Fluorescence data were collected at a photomultiplier voltage of 780 V.

To account for the dilution effects during titration the observed fluorescence intensity F_{obs} was corrected for the increase in volume:

$$F_{corr} = F_{obs} \frac{v_{sample}}{v_{start}} \qquad \textbf{Equation 5}$$

Where v_{start} gives the starting volume (600 µl) and v_{sample} the volume at the time of measurement. The corrected fluorescence intensity F_{corr} was plotted against the nucleotide concentration and the titration curves were fitted in Prism (GraphPad) with a quadratic binding equation [96]:

$$F = F_0 + \Delta F_{max} \cdot \frac{(X + B_0 + K_d)/2 - \sqrt{[(X + B_0 + K_d)/2]^2 - X \cdot B_0}}{B_0} \qquad \textbf{Equation 6}$$

Where F is the measured spectroscopic signal, F_0 the offset and ΔF_{max} the total signal change. A_0 and B_0 are the total concentrations of protein and nucleotide, respectively, and K_d is the dissociation constant obtained for protein and nucleotide.

Tryptophan fluorescence

Tryptophan fluorescence was measured at an excitation wavelength of 300 nm and an emission wavelength of 340 nm with a band width of 5 nm. 3 µM to 5 µM protein were titrated with nucleotides, doubling the concentration in each step from 1 µM to 8192 µM. ATP titrations

were performed in the presence of an ATP regeneration system (8 µg pyruvate kinase (Roche), 200 µM phosphoenolpyruvate (Roche)) and 15 mM KCl.

Mant fluorescence

2'-/3'-O-(N'-Methylanthraniloyl) (mant) fluorescence was measured at an excitation wavelength of 355 nm and an emission wavelength of 448 nm with a band width of 5 nm. 2 µM mant-labeled nucleotides were titrated with increasing protein concentrations between 1 µM and 15 µM. Mant-ADP and mant-ATP were purchased from Biolog, mant-APPNHP was purchased from Jena Biosciences.

Fluorescence displacement titrations of mant-ADP from the mant-ADP-ε1/ζ1 complex using ADP were performed to determine the binding affinity for the unlabeled nucleotide. After correction for dilution the data were fitted in the GraFit data analysis software (Erithacus) with a cubic binding equation as described by Thrall *et al*. [97].

2.5.7 Nucleotide binding kinetics

The kinetics of mant-ADP binding to ε1/ζ1 were measured at 25 ºC using a KinetAsyst stopped-flow instrument (HI-TECH). The mant-chromophore was excited at a wavelength of 355 nm and fluorescence was detected using a high-pass filter at 420 nm and a photomultiplier voltage of 780 V.

60 µl of each 2 µM mant-ADP (Biolog) and 2 µM, 4 µM or 8 µM ε1/ζ1 in titration buffer were mixed and mant-fluorescence was monitored over 200 s. For each protein concentration three mixing reactions were performed and averaged. The resulting traces were fitted to single exponential functions using the Kinetic Studio software (HI-TECH):

$$f(t) = -A \cdot e^{-k \cdot t} + C \qquad \textbf{Equation 7}$$

Where A and C are amplitude and offset of the exponential function that describes the change in fluorescence intensity over time t. The observed rate constants k_{obs} were plotted against the respective concentrations of ε1/ζ1 after mixing ([x]). The data points could be fitted with by linear regression giving the association (k_{on}) and dissociation (k_{off}) rates of the binding:

$$k_{obs} = k_{off} + k_{on} \cdot [x] \qquad \textbf{Equation 8}$$

From which the dissociation constant K_d could be calculated by:

$$K_d = k_{off}/k_{on} \qquad \textbf{Equation 9}$$

2.5.8 Fluorescence anisotropy

Fluorescence anisotropy of chromophore-labeled nucleotides was measured in the presence and absence of protein using a Fluorolog-3 (HORIBA Jobin Yvon). A T-format configuration was used for the simultaneous detection of horizontally and vertically polarized light at a photomultiplier voltage of 950 V.

Mant-fluorescence was excited at 355 nm (bandwidth 2 nm) and anisotropy was measured at 448 nm (bandwidth 5 nm). 1,N^6-etheno-adenosine diphosphate (etheno-ADP) was provided by Jochen Reinstein. The chromophore was excited at 310 nm (bandwidth 5 nm) and anisotropy was measured at 410 nm (bandwidth 5 nm). 2 µM labeled nucleotides in titration buffer were incubated at 25 ºC before 4 µM ε1/ζ1 or ζ1 was added and the anisotropy signal was analyzed with the FluorEssence software (HORIBA Jobin Yvon).

2.6 X-ray crystallography

2.6.1 Crystallization and cryoprotection

Possible crystallization conditions were screened using a sitting drop vapor diffusion setup. 70 µl of commercially available sparse matrix screening solutions (see section 2.1.5) were used as reservoir solutions in 96-well CrystalQuick low-profile plates (Greiner bio-one). Screens were set up using a Mosquito nanoliter pipetter (TTP LabTech Ltd) by mixing 100 nl each of protein and reservoir solution. Plates were stored at 20 ºC and crystal growth was monitored over time. Images were taken using a Rock Imager 1000 (Formulatrix) and examined using the provided Rock Maker software.

To produce larger individuals, crystals were reproduced using the hanging drop vapor diffusion method. The drop volume was scaled up to 2 µl in 24-well Linbro plates (Crystalgen Inc) with a corresponding reservoir volume of 700 µl. Plates were again stored at 20 ºC and crystal growth was monitored manually over time.

Crystallization and cryoprotection of ε1/ζ1 complex

Initial crystallization and optimization trials were performed by Roman Sakson. In the sparse matrix screens (JCSG Core I – IV Suite, Qiagen) crystalline material was observed for several ammonium sulfate conditions. Therefore, different ammonium sulfate containing conditions were tested using the AmSO4 Suite screen (Qiagen) and reservoir solutions with varying pH and ammonium sulfate concentrations. Crystals grew as stacked plates, but these plates were larger and more separated for protein batches where the His_6-tag was proteolytically removed. To obtain larger crystals the hanging drop vapor diffusion method was used.

Crystals used for data collection were grown in 24-well Linbro plates (Crystalgen Inc) using protein concentrated to 21.5 mg/ml and the following condition:

1.8 M $(NH_4)_2SO_4$
100 mM sodium citrate pH 4.5

ε1/ζ1 crystals appeared as clusters of stacked plates within two to three days and were harvested after approximately two weeks. For cryoprotection, crystals were soaked in 2 M $(NH_4)_2SO_4$, 100 mM sodium citrate pH 4.5 and 20 % (*v*/*v*) propylene glycol (SACF) and subsequently flash frozen in liquid nitrogen.

To obtain nucleotide-bound structures these conditions and the buffers used for cryoprotection were supplemented with 15 mM $MgCl_2$ and 10 mM of the respective nucleotide.

SeMet-labeled ε1/ζ1_K115A was crystallized at a concentration of 21.7 mg/ml using the same condition and harvested according to the same protocol.

2.6.2 X-ray data collection and processing

Prior to data collection at the X10SA beam line of the Swiss Light Source (Paul-Scherer Institute, Villigen, Switzerland), single diffraction images of fished crystals were measured at the institute to assess the diffraction quality. X-rays were generated with a MicroMax 007 HF (Rigaku) microfocus rotating copper anode (40 kV/30 mA), equipped with VariMax HF focusing mirrors (Confocal Max-FluxR CMF optics, Rigaku) and detected with a Mar345 image plate detector (MarResearch). Diffraction data of all crystals were collected at -173.15 °C.

At the Swiss Light Source diffraction data were collected with a Pilatus 6M detector (Dectris) in a continuous shutter-less φ-scan with oscillation ranges of 0.2 ° and exposure times of 0.2 s. The measurement parameters for the different data sets can be found in Table 2.2.

All data sets were processed with the XDS software package [98]. Reflections were identified, indexed and integrated with XDS and scaled and merged with XSCALE. XDSCONV was used for file conversion into formats suitable for successive programs.

Table 2.2: Data collection parameters

Parameters	**ε1/ζ1_K115A**		**wtε1/ζ1**	
	Native	**SeMet**	***apo***	**ADP**
X-ray wavelength (Å)	0.9763	0.9763	0.9780/1.000	0.9780
Detector distance (mm)	500	500	500/555	500
Oscillation range (°)	0.2	0.2	0.2	0.2
Number of images	1800	7200	1800/900	1800
Resolution range (Å)	50-2.6	50-2.8	50-2.9/50-2.9	50-2.9

2.6.3 Structure determination

Single wavelength anomalous diffraction (SAD) phasing

To obtain initial phases for structure determination of ε1/ζ1, SAD diffraction data [99] were recorded for crystals of SeMet-labeled ε1/ζ1_K115A. A highly redundant data set was collected at the K absorption edge of selenium (λ = 0.9763 nm). SAD phasing was performed using the SHELX-suite [100] in HKL2MAP [101]. SHELXC was used to extract and evaluate the anomalous signal. Estimates for the anomalous structure factor amplitudes were generated to a resolution of 3.5 Å, at which the anomalous correlation coefficient was 33 %. SHELXD was used to locate the heavy atom sites using the dual space recycling method with 20 cycles. 32 selenium sites were searched in 100 trials to a resolution of 2.4 Å. The heavy atom substructure obtained by SHELXD was used for phasing of the entire structure in SHELXE. In addition, SHELXE performed subsequent density modification assuming a solvent content of 60 % but this did not result in an interpretable electron density. Pseudo-free correlation coefficients (64 % / 60 %) and map correlation coefficients (0.776 / 0.750) were very similar for both substructure enantiomers.

Instead, the anomalous substructure was used for SAD phasing in Phaser [102] to a resolution of 3.0 Å. Phaser applies likelihood-based methods for phase determination by refining the phase probability (Hendrikson-Lattman coefficients) and log-likelihood gradient maps for substructure improvement. In the same step density modification with Parrot [103] was performed, which was improved by defining non-crystallographic symmetry restraints. The obtained phases were extended to a resolution of 2.4 Å using DM [104]. All three programs are implemented in the CCP4 software suite [105].

For the isomorphous crystals of wtε1/ζ1 the set of test reflections excluded in model building was inherited and initial phases were obtained using a rigid body refinement protocol [106].

Phasing by molecular replacement

For crystals of native ε1/ζ1_K115A phase extension was not possible due to the non-isomorphic unit cells. Therefore, the native dataset was phased by molecular replacement in Phaser [102] using one heterodimer (chains G and H, without water molecules or ions) of the SeMet-labeled ε1/ζ1_K115A model as a search structure.

2.6.4 Structure refinement

For the ε1/ζ1_K115A structure an initial model was manually built into the electron density obtained from SAD phasing with the modified structure factor amplitudes (FDM) and phases (PHIDM) obtained from DM. Models were built in *Coot* [107] and initially refined to a resolution of 2.8 Å by a simulated annealing protocol performed with CNS [108, 109] and taking into account the experimental phase information. 100 cycles of energy minimization were followed by melting

to a starting temperature of 4726.85 °C (5000 K). Subsequently, the model was cooled to 26.85 °C (300 K) in steps of 25 °C. All subsequent refinement and improvement of the models obtained by molecular replacement or rigid body refinement was performed by iterative cycles of manual building and refinement using phenix.refine [110] with non-crystallographic symmetry torsion-angle restraints [111]. In addition, the model of the SeMet-labeled protein was refined taking into account the anomalous scattering contributions.

As an indicator of the model quality, an R-factor was calculated, taking into account the differences between the observed $|F_0|$ and the calculated $|F_c|$ structure factor amplitudes (R_{work}):

$$R_{work} = \frac{\sum ||F_0| - |F_c||}{\sum |F_0|} \qquad \textbf{Equation 10}$$

To obtain a measure of the model bias an R-factor for the omitted reflection (R_{free}) [112] was calculated similarly, taking in account only a randomly generated set of 5% of the measured reflections. These were excluded from the refinement and kept constant for all related datasets.

The stereochemistry of the model for intermediate and final structures was evaluated using the subroutines implemented in *Coot* [107] and final structure validation was performed using MolProbity [113].

2.7 Bioinformatics

2.7.1 Homology search and sequence extraction

The NCBI non-redundant nucleotide or protein collections [114] were searched for homologous sequences using BLAST [115, 116]. The Bio::Seq module of BioPerl [117] was used to automatically extract nucleotide or protein sequences that correspond to identified NCBI Genbank IDs as well as adjacent sequences or ORFs. The Perl script was written by Robert Lindner.

2.7.2 Sequence alignments and secondary structure prediction

Multiple sequence alignments and sequence identity calculations were generated with ClustalOmega [118, 119]. Sequence conservation values were calculated with the AMAS server [120] and protein secondary structure elements were predicted using PSIPRED v3.3 [121]. Alignments that compile this information were prepared with ALSCRIPT [122].

Sequence logos to highlight the conservation of specific regions were generated with WebLogo [123].

2.7.3 Protein structure analysis

Structurally similar proteins were identified using DaliLite v.3 [124] searching the RCSB Protein Data Bank (PDB) [125] for matches to a query structure. Superposition of protein chains was performed with LSQKAB [126] implemented in the CCP4 software suite [105] or by Secondary-Structure-Matching (SSM) [127] in *Coot* [107].

Contact areas of molecular interaction surfaces were calculated using the PISA server [128] and electrostatic surface potentials were calculated using the APBS plugin [129] in PYMOL (DeLano Scientific LLC).

All protein structure figures were prepared using PYMOL (DeLano Scientific LLC).

3 Results

3.1 *Escherichia coli* EzeT

This section describes the identification and characterization of the type II-like TA system EzeT, a zeta homolog in *E. coli*. Bioinformatic studies showed the prevalence of EzeT in Gram-negative bacteria and helped to analyze the *ezeT* promoter. Microbiological and biochemical studies on the expression and enzymatic activity *in vivo* and *in vitro* confirmed the presence of an active UNAG kinase in Gram-negative bacteria. Finally, growth analyses allowed to deduce a new type of toxin inhibition that might be extended to other TA systems.

3.1.1 N-terminally elongated Zeta toxins in Gram-negative bacteria

So far, only epsilon/zeta systems from Gram-positive bacteria have been characterized. To identify zeta toxin homologs in the genomes of Gram-negative bacteria, the nucleotide sequences of ζ_{spy} and PezT were used as query sequences in a BLAST search [115, 116]. When probing all available sequenced *E. coli* genomes in the NCBI Nucleotide collection [114], highly conserved, uncharacterized open reading frames (ORFs) were found in 364 strains. Interestingly, in spite of their homology to zeta toxins, no epsilon-like antitoxin gene could be identified in their vicinity.

The encoded protein sequences could be subdivided into 37 groups of identical sequences with an overall sequence identity of 88 % and pairwise sequence identities above 96 % (see Appendix, section 7.2). Due to the computational assignment of different start codons the annotated protein sequences varied slightly in length (between 304 and 310 amino acids (aa)). However, the region at the 5'-end of all ORFs, including the putative promoter sequence, is highly conserved and different translational start points are therefore unlikely. Within this region, a Shine-Dalgarno consensus sequence was identified. This ribosome binding site is usually located 4 to 18 nucleotides upstream of the start codon with an optimal spacing of 8 to 10 nucleotides [130]. The GTG start codon of the putative 310 aa polypeptide lies directly next the Shine-Dalgarno sequence, which is most likely too short for translation initiation, while the proposed start codon of the 304 aa gene product is 20 nucleotides downstream of the Shine-Dalgarno sequence, which would result in inefficient translation initiation. In contrast, the spacing is optimal for a 308 aa long polypeptide starting 8 nucleotides downstream of the Shine-Dalgarno sequence (compare Figure 3.1-10B). Therefore, the annotated 308 aa protein from *E. coli* O103:H2 12009 (locus ECO103_4771) was used in subsequent studies.

When the sequence of this zeta homolog was compared to the streptococcal zeta toxins, a region with high homology to ζ_{spy} and PezT having 38 % and 47 % aa similarity, respectively (Figure 3.1-1A), was identified in the C-terminal part of the assigned protein sequence (aa 80 to 308). This zeta-like region comprises the catalytically important Walker A motif, the magnesium coordinating aspartate of a degenerate DxD motif and the UNAG binding loop characterized by the GTXR consensus sequence [61, 66], indicating that the protein has UNAG kinase activity.

In addition, the protein contains 80 aa at the N-terminus, which are not found in the shorter streptococcal homologs (ζ_{spy} consists of 287 aa and PezT of 253 aa). Interestingly, these residues are predicted to form three α-helices, which are similar to the secondary structure elements that form the three-helix bundle of epsilon antitoxins (Figure 3.1-1A). Together, these findings led to the hypothesis that the identified zeta homolog is an autoinhibited UNAG kinase in which toxin and antitoxin functionalities are encoded in one polypeptide chain. Therefore the gene product was called *E. coli* epsilon/zeta, abbreviated as EzeT.

Currently, genomic sequences of 2914 completely or partially sequenced *E. coli* isolates are available in the NCBI Nucleotide collection and the *ezeT* ORF was found in 364 of these genomes. Sorting of the sequenced strains by phylogroup revealed a clear prevalence of *ezeT* in phylogroup B1. Except for *E. coli* HS and *E. coli* N1 (phylogroup A), all of the 138 strains that could be assigned to a phylogroup were B1 strains. These include the laboratory strain *E. coli* W as well as the pathogenic strains *E. coli* O103:H2 and O104:H4, which can cause lethal hemolytic-uremic syndrome as demonstrated in the German 2011 outbreak. In all cases, *ezeT* was found in a conserved 3.4 kb islet, located between *lysC*, encoding for lysine sensitive aspartokinase III, and *yjbD*, a gene of unknown function (Figure 3.1-1B). This islet contains two additional hypothetical ORFs further downstream on the complementary strand; a putative nucleotidyl transferase and a homolog of SPFH domain/band 7 proteins, which are integral membrane proteins [131].

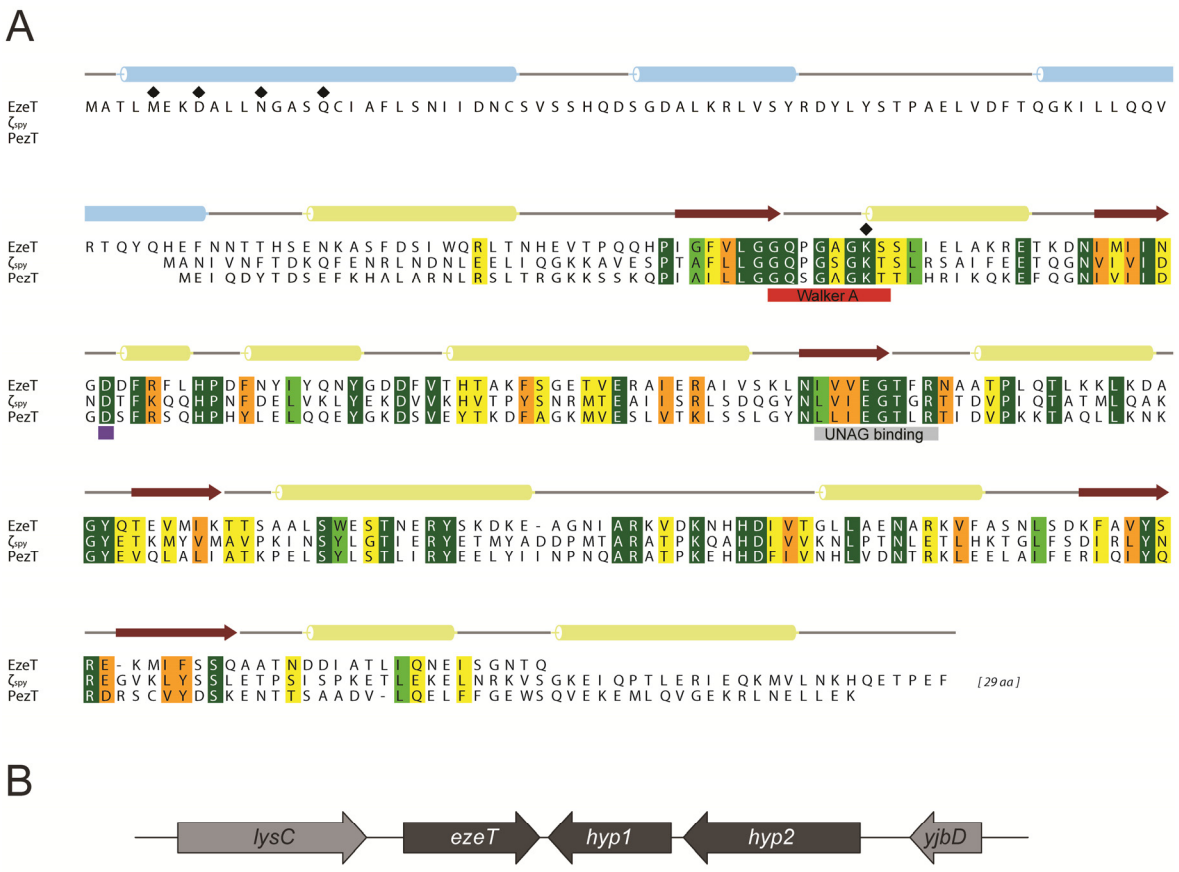


Figure 3.1-1: A zeta homolog in *E. coli*. **(A)** Amino acid sequence alignment of EzeT (ECO103_4771) and the two streptococcal zeta toxins ζ_{spy} and PezT shows the homology of the C-terminal domain of EzeT to this toxin family. Residues are colored according to conservation (dark green: identity to yellow: low homology). Secondary structure elements of ζ_{spy} are indicated above the alignment as olive bars (α-helices) and red arrows (β-strands). The catalytically important Walker A motifs (red), the magnesium-coordinating aspartate of the DxD motif (purple) and the UNAG binding region (gray) are conserved. The putative N-terminal domain of EzeT comprises about 80 amino acids and is predicted to fold into three α-helices (blue bars). Residues that were mutated in this study are indicated with diamonds (♦). **(B)** Genetic organization of the *ezeT* containing 3.4 kb islet found in 364 *E. coli* strains. The islet is inserted between the genes encoding lysine sensitive aspartokinase III (*lysC*) and *yjbD*, a gene of unknown function (light gray) and contains three ORFs (dark gray), *ezeT* and two hypothetical genes encoded on the complementary strand. *Hyp1* denotes a putative nucleotidyl transferase and *hyp2* a homolog of membrane proteins belonging to the SPFH domain/band 7 family.

3.1.2 EzeT purification and stability

The *ezeT* ORF was cloned into an IPTG-inducible expression vector (pET28b(EzeT)) thereby adding the coding sequence for a His_6-tag to the 3'-end. *E. coli* BL21 CodonPlus(DE3)-RIL cells, which do not harbor an endogenous *ezeT* ORF, were transformed with this vector and the protein was recombinantly expressed at a growth temperature of 16 ºC. Recombinant EzeT was purified by metal affinity chromatography, anion exchange chromatography and size exclusion chromatography to yield a homogenous sample. Protein purity was monitored by SDS-PAGE, which revealed that EzeT migrated with a velocity corresponding to an apparent electrophoretic

mobility of 35 kDa (Figure 3.1-2A). Moreover, the elution profile from the Superdex 75 10/300 GL used for size exclusion chromatography showed a single peak eluting at a retention volume of 11.5 ml (Figure 3.1-2B), which corresponds to a molecular mass of approximately 35 kDa. Since full-length His_6-tagged EzeT has a theoretical molecular weight of 35.8 kDa, these results indicate that the entire ORF gets translated as a single polypeptide chain when expressed recombinantly and that the protein forms a monomer in solution.

However, under the conditions used, the protein solubility was impaired and the purified protein tended to precipitate over time, especially at high protein concentrations. This instability became already evident upon protein overexpression, during which EzeT accumulated in inclusion bodies. The ratio between soluble protein and protein in inclusion bodies depended on the expression temperature; at 16 ºC more than half of the protein was found in the soluble fraction after cell lysis, while upon expression at 25 ºC 80 % of the protein proved to be insoluble (Figure 3.1-2C). Furthermore, additional experiments showed that the presence of an N-terminal His_6-tag is detrimental to protein solubility even at 16 ºC.

To closer investigate the stability of EzeT, the melting temperature of purified full-length EzeT was determined by CD spectroscopy. Thermal denaturation was followed by detection of light polarization at 222 nm and the data points were fitted in Prism assuming a two-state unfolding process (Figure 3.1-2D). The unfolding was irreversible, giving an apparent melting temperature of 38 ºC, which is close to the standard growth temperature of *E. coli* of 37 ºC. In conclusion, the results indicate a low thermal stability of EzeT, which might have implications for EzeT functionality.

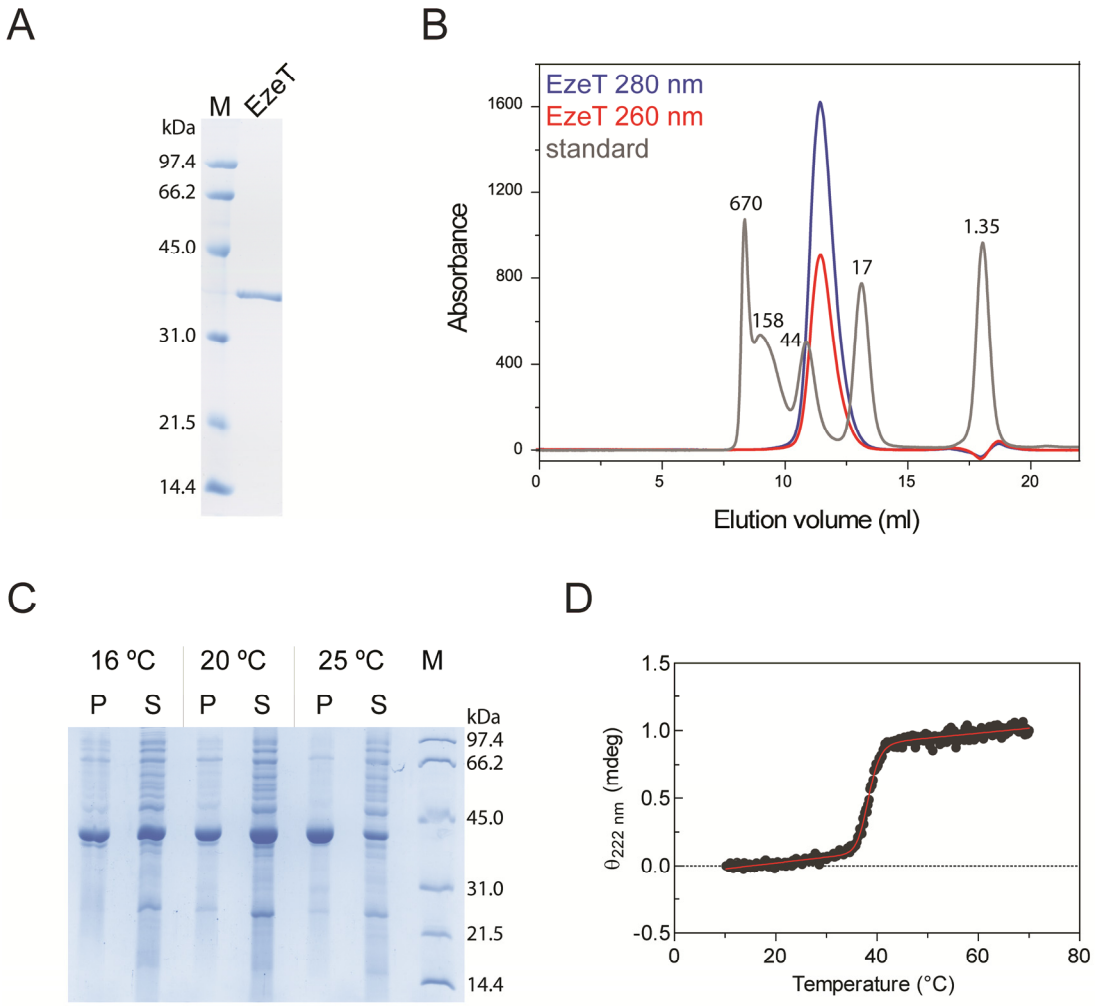


Figure 3.1-2: EzeT purification and stability. SDS-PAGE analysis **(A)** and gel filtration chromatography **(B)** of EzeT indicate that the protein is expressed as a single 35.8 kDa polypeptide chain and is a monomer in solution. Molecular weight estimates (in kDa) of the components of the Low Range SDS-PAGE standard (M) and the Gel Filtration standard (gray) are given for comparison. **(C)** EzeT is better soluble when overexpressed at lower temperatures as shown by a comparison of soluble (S) and insoluble (P) lysate fractions after expression at 16 °C, 20 °C and 25 °C. While 80 % of the protein is in the insoluble fraction at 25 °C, more than half of the protein is soluble upon overexpression at 16 °C. **(D)** Melting curve of EzeT measured by circular dichroism spectroscopy. The data were fitted assuming a two-state unfolding process (red line), giving an apparent melting temperature of 38 °C.

3.1.3 EzeT phosphorylates UNAG *in vitro*

The conservation of the catalytically important residues in EzeT suggested that the protein has UNAG kinase activity, which was tested in an *in vitro* phosphorylation assay using purified full-length protein.

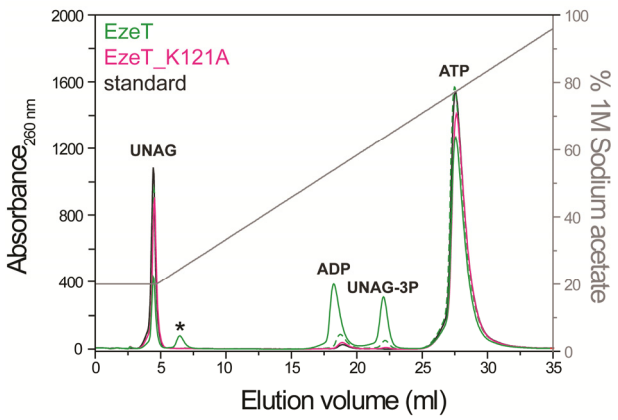


Figure 3.1-3: UNAG phosphorylation by EzeT *in vitro*. Nucleotides/ nucleotide sugars were separated by anion exchange chromatography as shown for a standard containing UNAG and ATP (black line). Incubation of EzeT with UNAG and ATP leads to the formation of two products, ADP and UNAG-3P (green lines), which could not be detected for the Walker A lysine mutant EzeT_K121A (pink line). After 3 hours of incubation, EzeT had phosphorylated 9 % of the substrate UNAG (dashed line) and after 24 hours 60 %, indicating a low turn-over rate. The asterisk (*) denotes the retention time of AMP which is probably the result of a minor contamination.

Freshly purified EzeT was incubated together with $MgCl_2$, ATP and UNAG at 25 ºC. The resulting nucleotide/nucleotide sugar species were separated by anion exchange chromatography and monitored spectroscopically (Figure 3.1-3). Enzymatic activity of EzeT was observable as a decrease in UNAG levels and formation of an equimolar amount of ADP. Simultaneously, an additional species appeared with a retention time corresponding to UNAG-3P, showing that EzeT phosphorylates UNAG and encodes an active UNAG kinase in Gram-negative bacteria. The reaction was specific to EzeT, as no UNAG-3P production was observed for an EzeT variant in which the essential Walker A lysine (K121) was mutated to an alanine (EzeT_K121A). However, turn-over by the full-length protein was slow, as after 3 h only 9 % and after 24 h only 60 % of the substrate was phosphorylated.

This enzymatic activity is much slower than for free PezTΔC242 (nearly complete turn-over after 1 h), but comparable to the activity observed for the ε/ζ_{spy} complex, in which UNAG kinase activity is inhibited by equimolar amounts of the cognate antitoxin ε_{spy} [53]. Similarly, the phosphorylation rate observed for EzeT is in agreement with an inhibitory function of the N-terminal region, although a low intrinsic phosphotransferase activity of EzeT cannot be excluded.

3.1.4 The C-terminal domain of EzeT is toxic *in vivo*

TA systems usually consist of two parts, a toxin which is toxic upon overexpression and an antitoxin that inhibits toxicity. In EzeT, these two functionalities seem to be located in two domains connected by a linker. Based on homology to the characterized epsilon/zeta systems and analysis of the predicted secondary structure, the domain boundary was assigned to the region around residue 82 (NNTTHS, see Figure 3.1-1A) and the corresponding domains were separately cloned into different vectors (pbad(EzeT(1-82)) and pET28b(EzeTΔN83)). However, domain separation

reduced the solubility of the protein even further and overexpression and purification attempts of the C-terminal domain alone did not yield any protein suitable for *in vitro* analysis.

Therefore, toxicity of EzeT was investigated *in vivo* by comparing the growth phenotype of cultures overexpressing full-length EzeT or the truncated C-terminal zeta-like domain (EzeTΔN83). Protein expression was induced with IPTG at 20 ºC and cell propagation was monitored by measuring the optical density at 600 nm (OD_{600}) and the number of colony forming units per ml (CFUs) (Figure 3.1-4).

As expected from the low apparent reaction rate *in vitro*, expression of full-length EzeT showed no apparent phenotype *in vivo*, similar to the inactive Walker A lysine mutant (EzeTΔN83_K121A) and the empty pET28b vector. In contrast, expression of EzeTΔN83 caused a rapid drop in CFUs. Within 180 min the number of CFUs decreased by 4 orders of magnitude, whereas a decrease in OD_{600} by approximately 50 % got visible after a lag phase of 90 min. This difference between both methods is probably due to scattering from dead cells and cell debris that increases the OD_{600}, while, on the other hand, continuing intoxication after plating might decrease the ability of individual cells to grow into visible colonies. Furthermore, the transfer to solid media might impose additional stresses like the mechanical forces during plating or the lower humidity. Nevertheless, these results show that, like in other TA systems, overexpression of the toxin leads to a growth phenotype.

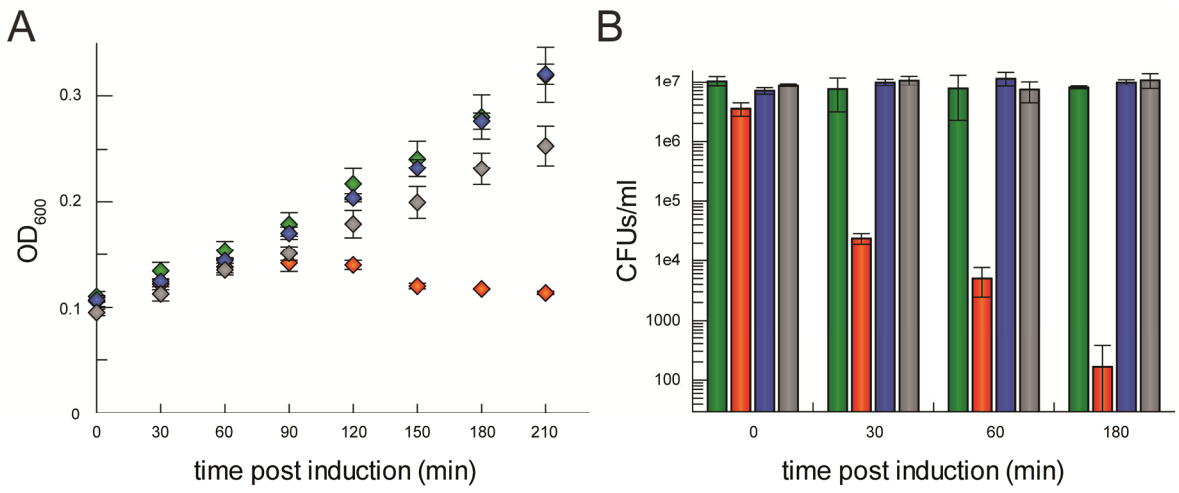


Figure 3.1-4: Toxicity of EzeTΔN83 overexpression *in vivo*. Growth curves **(A)** and CFUs **(B)** of cultures expressing EzeT (green), EzeTΔN83 (red) or EzeTΔN83_K121A (blue) at 20 ºC. Expression from an empty pET28b (gray) was used as control. While overexpression of full-length protein or the Walker A mutant did not affect the growth rate compared to the control, overexpression of EzeTΔN83 caused a decrease in OD_{600} after 90 min and a rapid drop in CFUs. Measurements were performed in triplicates.

3.1.5 Toxicity is temperature dependent

Cloning attempts for other epsilon/zeta systems failed to separate the toxin gene from the epsilon antitoxin gene without introducing mutations that impair enzymatic activity [53]. Most likely, already minor amounts of free toxin are detrimental to cell growth. Surprisingly, the enzymatically active C-terminal domain of EzeT could be cloned without its inhibiting N-terminal domain using standard cloning methods, although toxicity was detectable at 20 ºC. Together with the low thermal stability of full-length EzeT, this indicated that the toxicity of the C-terminal domain might be dependent on the growth temperature. Therefore, the effect of overexpression of the C-terminal domain was compared at different temperatures in *E. coli* BL21 CodonPlus(DE3)-RIL (Figure 3.1-5).

At 37 ºC, the temperature used for culture growth during cloning, overexpression of EzeTΔN83 did not affect cell viability, as OD_{600} and CFUs were identical to those for the inactive Walker A lysine mutant EzeTΔN83_K121A. Notably, after 180 min the number of CFUs decreases slightly upon expression of both protein variants, which is indicative for an overexpression artifact. Presumably, this acquired sensitivity to plating is caused by an accumulation of insoluble protein and inclusion bodies in the cytosol due to the large amount of protein synthesized upon IPTG-induction.

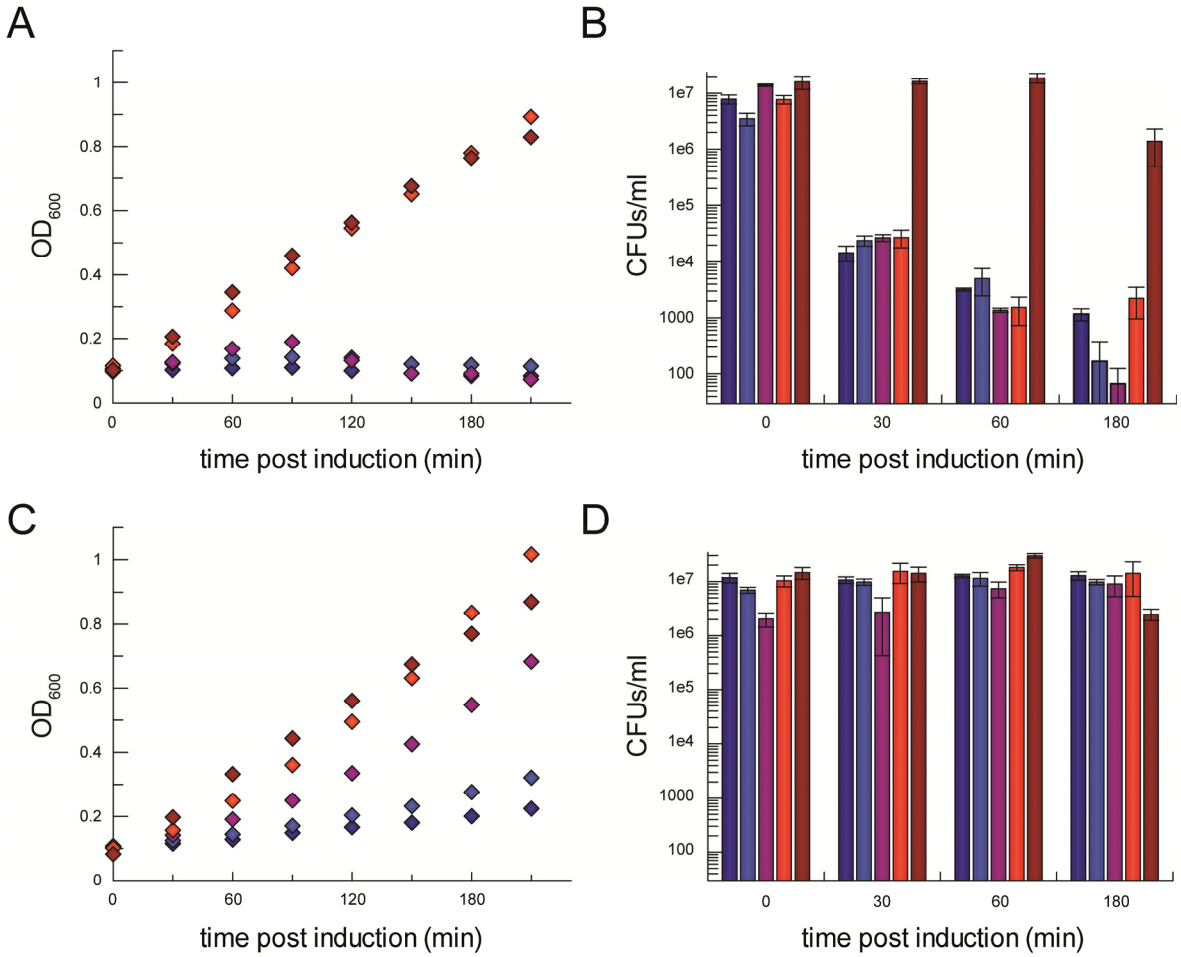


Figure 3.1-5: Temperature dependence of EzeTΔN83 toxicity. Growth curves **(A, C)** and CFUs **(B, D)** of cultures expressing EzeTΔN83 (A, B) or enzymatically inactive EzeTΔN83_K121A (C, D) at different temperatures (16 ºC: dark blue; 20 ºC: blue; 25 ºC: purple; 30 ºC: red; 37 ºC: dark red). Cultures expressing EzeTΔN83_K121A grew steadily, albeit with temperature dependent doubling times. Cultures expressing EzeTΔN83 at 16 ºC, 20 ºC and 25 ºC showed a decrease in OD_{600} 90 min after induction and a rapid drop in CFUs. After 180 min the number of CFUs was decreased by approximately 4 orders of magnitude. Expression at 30 ºC led to a decrease in CFUs, but not in OD_{600} and at 37 ºC EzeTΔN83-expressing cultures grew similar to the EzeTΔN83_K121A control. Note the slightly reduced number of CFUs for both constructs at 180 min at 37 ºC, which is probably caused by an overexpression artifact. Measurements were performed in triplicates.

In contrast, cell death was observed when overexpression of EzeTΔN83 was induced at 16 ºC or 20 ºC, as described before. A decrease in OD_{600} and CFUs was also detected upon expression at 25 ºC, whereas at 30 ºC the OD_{600} constantly increased. However, although no sign of toxicity was evident by monitoring the turbidity of the culture, CFUs immediately dropped to vanishingly low levels after induction at 30 ºC. This discrepancy between the two methods used to study viability was already seen as a difference in lag phase at 20 ºC, but is probably even more pronounced at 30 ºC due to the low protein stability at this temperature. At all of these temperatures, expression of EzeTΔN83_K121A did not cause the loss of cell viability showing that the enzymatic activity is responsible for the growth phenotype. The temperature dependence and the low *in vitro* melting point suggest a temperature-sensing function of EzeT.

3.1.6 The C-terminal domain of EzeT causes a lytic phenotype

UNAG-3P cannot be converted by enzymes of the peptidoglycan synthesis pathway and was shown to accumulate in the cytosol of PezTΔC242-expressing cells [53]. To prove that EzeT toxicity is similarly caused by the phosphorylation of UNAG, the accumulation of UNAG-3P was measured *in vivo* by analyzing low-molecular-weight-metabolite extracts of cells expressing EzeTΔN83. Therefore, protein expression was induced by IPTG at 16 ºC and cultures were harvested after 2 hours. Low-molecular-weight compounds were solubilized using 80 % aqueous acetonitrile and separated by anion exchange chromatography. Eluents were monitored spectroscopically and compared to authentic standards (Figure 3.1-6A).

An accumulation of a compound with the same retention time as UNAG-3P was detected in extracts from cells expressing EzeTΔN83, but not from cells expressing the inactive Walker A lysine mutant EzeTΔN83_K121A. This shows that the UNAG kinase domain of EzeT is enzymatically active under physiological conditions. In agreement with the low *in vitro* activity and the inability in evoking a growth phenotype, no UNAG-3P formation was detected for the full-length protein *in vivo* showing that the N-terminal domain influences the phosphorylation efficiency.

Phosphorylation of UNAG was shown previously to decrease the stability of the bacterial cell wall and to lead to cell lysis [53]. Therefore, the break-down of the osmotic barrier was measured in *E. coli* BL21 CodonPlus(DE3)-RIL cells expressing truncated EzeT by following the influx of propidium iodide into the cell (Figure 3.1-6B). Intercalation of propidium iodide into DNA was measured by detecting the fluorescence signal at 620 nm upon excitation at 520 nm. Indeed, cultures expressing EzeTΔN83 showed a strong increase in propidium iodide fluorescence simultaneously with a decrease in OD_{600}, comparable to the effect of administration of 50 µg/ml of the cell wall perturbing antibiotic ampicillin. In contrast, propidium iodide fluorescence of cultures expressing the Walker A lysine mutant was comparable to fluorescence of the non-induced control culture, indicating that the bacterial cell wall remains intact. These results suggest that the C-terminal domain of EzeT is a UNAG kinase that causes a lytic phenotype due to the inhibition of peptidoglycan synthesis.

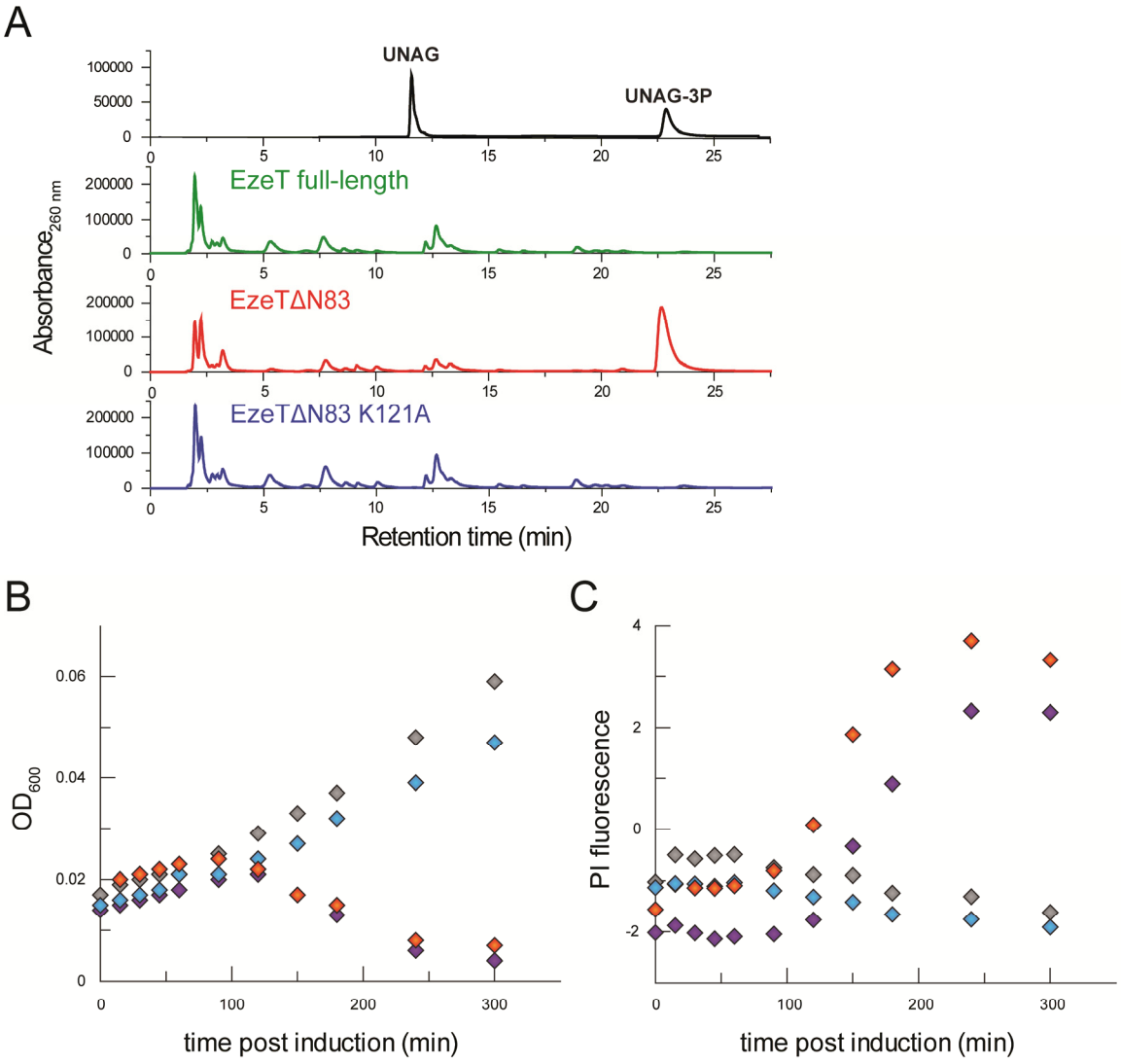


Figure 3.1-6: EzeTΔN83 causes UNAG-3P accumulation and break-down of the osmotic barrier. (A) HPLC analysis of small-metabolite-extracts from cultures expressing EzeT (green), EzeTΔN83 (red) or EzeTΔN83_K121A (blue). Only expression of EzeTΔN83 leads to the accumulation of a compound with the same retention time as UNAG-3P, which is included in the standard (black). For the full-length protein no UNAG-3P formation is observed, indicating that the N-terminal domain decreases the enzymatic activity. Growth curves **(B)** and propidium iodide (PI) fluorescence measurements **(C)** of cultures expressing EzeTΔN83 (red) or EzeTΔN83_K121A (blue) or non-induced cells (gray) grown in a medium containing 0.05 mg/ml propidium iodide. Cells treated with ampicillin (purple) show the typical increase in propidium iodide fluorescence caused by the break-down of the osmotic barrier. A similar increase in propidium iodide fluorescence is also observed for the cultures expressing EzeTΔN83, indicating that the protein also disturbs the bacterial cell wall.

3.1.7 Phenotype of EzeT expression

The destabilization of the bacterial cell wall is not only evident as a loss of viability and a break-down of the osmotic barrier of an entire culture, but can be also followed as changes in the morphology of individual cells. In *E. coli* cells overexpressing PezTΔC242 membrane bulges were

formed at the septum within 30 min and rupture of those bulges led to cell lysis and the formation of empty ghost cells [53].

The time course of morphological changes upon overexpression of EzeTΔN83 at 20 ºC was followed by phase contrast microscopy. The membrane integrity was assessed by fluorescent live/dead staining using the cell-permeant SYTO 9 and the cell-impermeant propidium iodide nucleic acid stains (Figure 3.1-7A).

Microscopic images taken 30 min after induction revealed that the majority of cells had a rod-shaped morphology and an intact, propidium-iodide impermeable membrane. Later on, surviving cells adopted a spherical shape, while a significant number of propidium iodide-permeable 'ghosts' and lysed cells became visible. During the course of the experiment, cells from the control culture expressing EzeTΔN83_K121A maintained their rod-like morphology and membrane integrity and dividing cells were commonly observed. It is therefore likely that the viable, spherical cells in cultures expressing EzeTΔN83 are a result of the UNAG kinase activity. The accumulation of UNAG-3P probably decreases the stability of the peptidoglycan layer, which is required to maintain the typical rod-like shape of *E. coli* against the internal turgor pressure. Interestingly, the observed morphology is reminiscent of spheroplast formation after treatment of cells with cell wall perturbing agents [132, 133]. These agents perturb the rigid peptidoglycan scaffold and thereby cause the cells to adopt a spherical morphology.

Notably, the morphological studies indicate, that the onset of cell death is after approximately 90 min, while cells are mostly intact after 60 min. These results are in good agreement with the bulk measurements of propidium iodide influx (former section), which indicated an onset of membrane permeabilization after approximately 100 min. Together, the data show that OD_{600} measurements give a better indication of the time point of killing by the toxin than determination of CFUs. It is, however, conceivable that the stability of the peptidoglycan layer is already impaired at earlier time points.

Despite their unusual shape, spherical survivors of the toxic activity seem to be viable for prolonged times. In cryo-electron microscopic images of cultures expressing EzeTΔN83 at 16 ºC for 16 h, bloated cells with irregular shapes are visible, which seem to be surrounded by an intact membrane (Figure 3.1-7B). Most of these cells contain one or more large, electron-lucent structures, which most likely are inclusion bodies. After incubation for more than 24 hours, some of the cells seem to be able to overcome the toxic effect of protein expression and culture growth resumes (data not shown).

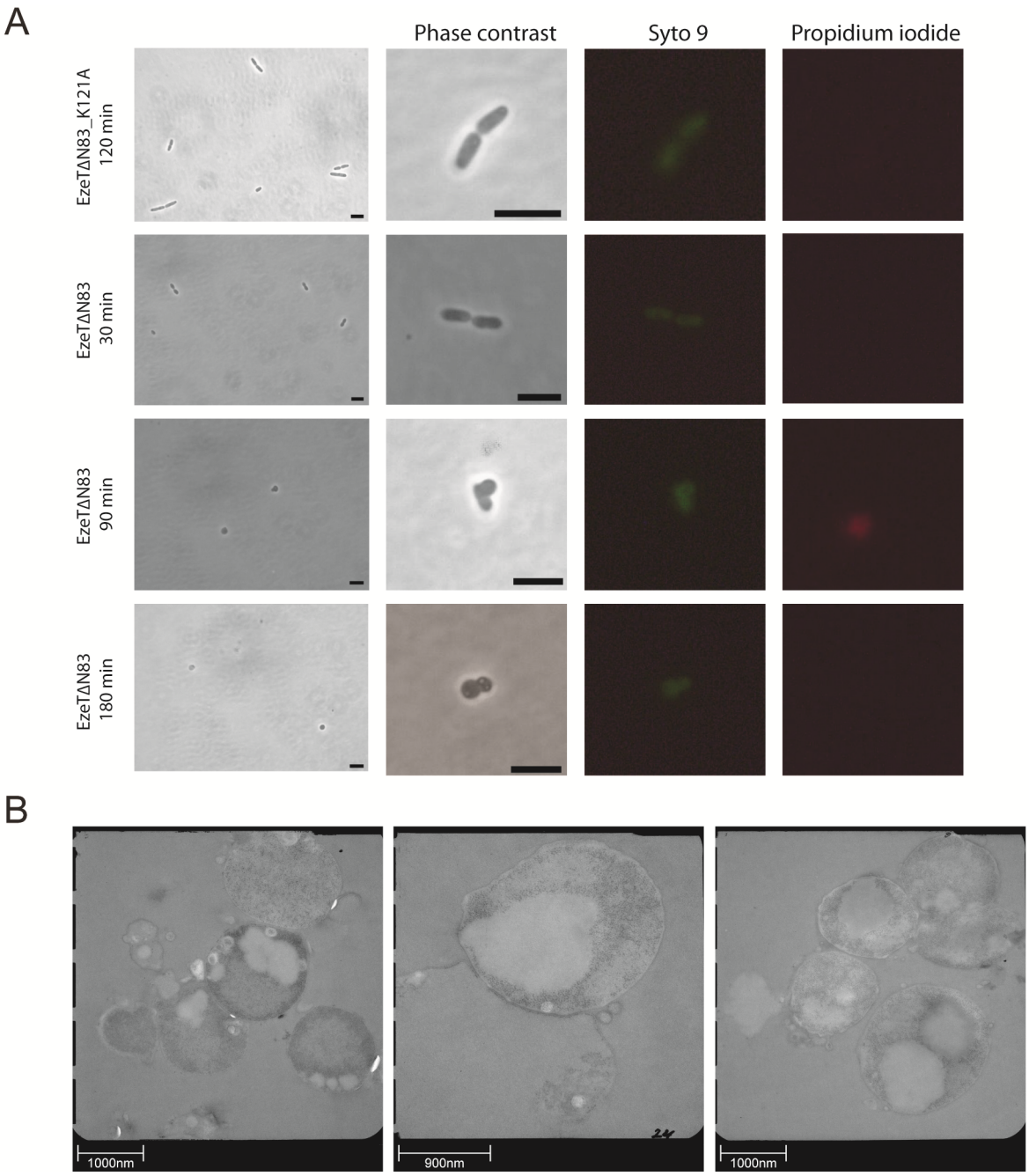


Figure 3.1-7: Phenotype of EzeTΔN83 expression. (A) Representative phase contrast and fluorescence microscopy images of cultures expressing EzeTΔN83 at 20 ºC. Live cells are visualized by green fluorescence using membrane-permeable SYTO 9, dead cells are visualized be red fluorescence using membrane-impermeable propidium iodide. Images were taken at indicated time points. A reduction in the total cell number and rounding of cells is visible after 90 min. In contrast, cells expressing EzeTΔN83_K121A (top) are viable, rod-shaped and divide during the course of the experiment. The black bar has a length of 5 µm. **(B)** Cryo-electron microscopy images of cultures expressing EzeTΔN83 for 16 h at 16 ºC. Cells adopt irregular shapes, but seem to form intact membrane-enclosed compartments. The large electron lucent structures most likely are inclusion bodies.

3.1.8 The N-terminal domain of EzeT inhibits toxicity

In conventional TA systems co-expression of the antitoxin in *trans* inhibits toxicity of the toxin. In this study, the toxicity of EzeT was shown to be located in the C-terminal domain, while the presence of the N-terminal domain in the full-length protein seemed to inhibit both UNAG kinase activity and the lytic phenotype. To test whether the separated domains of EzeT indeed can function like a conventional type II TA system, *E. coli* BL21 CodonPlus(DE3)-RIL were co-transformed with plasmids encoding for EzeT(1-82) and EzeTΔN83 under control of two separately-inducible promoters (Figure 3.1-8).

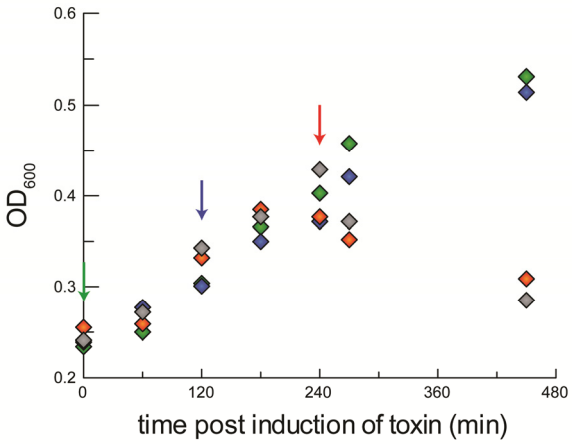


Figure 3.1-8: Co-expression of the N-terminal domain inhibits EzeTΔN83 toxicity. Growth curves of cultures expressing EzeTΔN83 from an arabinose-inducible pBAD-vector at 16 ºC. Expression of EzeT(1-82) from a co-transformed IPTG-inducible pET28b-vector was induced separately at different time points (arrows; 0 min (green), 120 min (blue) or 240 min (red)). EzeTΔN83 expression only (gray) leads to a decrease in OD_{600} after 240 min. Co-expression of the N-terminal domain can rescue this phenotype when induced during the lag phase. Induction of EzeT(1-82) after 240 min, when OD_{600} started to decrease was not effective anymore, indicating a 'point-of-no-return' [134].

As expected, no lytic phenotype was detectable upon co-expression of both domains from pET28b(EzeT(1-82)) and pBAD(EzeTΔN83). In contrast, upon induction of EzeTΔN83 only with 0.2 % L-(+)-arabinose, a lytic phenotype was observed, similar as upon overexpression from the IPTG-inducible T7-promoter. However, the lag time until a decline in OD_{600} was detectable, increased from 90 min to 240 min. This is probably due to a lower level of protein expression from the arabinose-inducible promoter, as well as low-level expression of EzeT(1-82) due to leakiness of the T7 promoter. Induction of EzeT(1-82) within this lag time was sufficient to restore normal growth. In contrast, induction of EzeT(1-82) after 240 min, when the OD_{600} already decreased, could not rescue cell viability anymore. This phenomenon, known as the 'point-of-no-return', has been described earlier for other TA systems including the homologous ε/ζ_{spy} system [79, 135] and is thought to occur when the damage caused by the toxin reaches irreversible levels [134].

In conclusion, the two domains of EzeT show toxin and antitoxin properties similar to the two individual proteins of type II systems, but are covalently linked into one polypeptide chain. This type of TA system inhibition has not been described earlier.

3.1.9 Residues at the N-terminus are required for inhibition

In the structurally characterized epsilon/zeta systems, the three-helix bundle of epsilon antitoxins binds to zeta and inhibits enzymatic activity by steric hindrance of ATP binding [62, 63]. This is accomplished by side chains of the N-terminal helix of epsilon that protrude into the active site of zeta. The conservation of predicted secondary structure elements (see Figure 3.1-1) raises the possibility that this mechanism of inhibition is conserved in EzeT.

To narrow down the region within the N-terminal domain of EzeT required for inhibition, constructs with truncations after the first (EzeTΔ36) or second (EzeTΔ62) predicted α-helix were expressed in *E. coli* BL21 CodonPlus(DE3)-RIL and OD_{600} and CFUs were compared to cultures expressing the full-length protein or EzeTΔN83 (Figure 3.1-9). For both truncation constructs, a decrease in OD_{600} was detected, although with a longer lag phase of 360 min or 210 min, respectively, when compared to cultures expressing the C-terminal domain alone (90 min). Determination of CFUs showed a strong drop in viability after 180 min (2500 fold) for EzeTΔN62 and a moderate 30 fold decrease for EzeTΔN36, which was more pronounced after 300 min (300 fold). The weaker toxicity could be caused by differences in protein stability of the truncated constructs, leading to lower levels of active enzyme in the cell for EzeTΔ36 and EzeTΔ62. However, the loss of inhibitory function indicates that the N-terminal helix is required for inhibition, either because residues in this helix are directly involved in blocking ATP binding or because the helix is required for domain stability or binding to the toxin.

In order to discern the effects on tertiary structure destabilization from the requirement of individual side chains, residues that presumably interfere with ATP binding were mutated to alanine residues. In ε_{spy} and PezA these residues (tyrosine, phenylalanine (ε_{spy}) / leucine (PezA) and two glutamates [62, 63]) are located on one side of the N-terminal α-helix in the three-helix bundle. Correspondingly, methionine 5, aspartate 8 and the carbamoyl-containing amino acids asparagine 12 and glutamine 16 of EzeT are predicted to be located in successive turns of the first α-helix and might therefore protrude into the ATP binding site. Therefore, these residues were mutated to alanine and the corresponding construct (EzeT_M5A_D8A_N12A_Q16A) was expressed in *E. coli* BL21 CodonPlus(DE3)-RIL. Measurement of OD_{600} and CFUs showed a steady increase in OD_{600}, but similarly to EzeTΔN36 CFUs dropped 30 fold after 180 min and 200 fold after 300 min (Figure 3.1-9). This shows that these residues in the N-terminal domain are indeed required to block substrate binding to the C-terminal domain. It is likely that the N-terminal domain inhibits UNAG kinase activity by preventing ATP binding similar as described for PezA and ε_{spy} [62, 63].

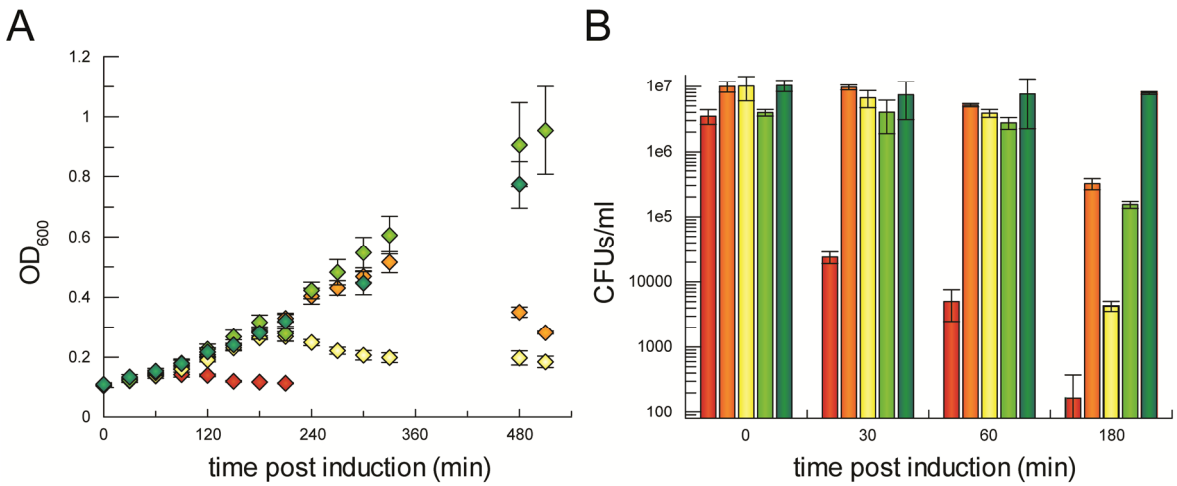


Figure 3.1-9: Truncation or mutation of the ultimate N-terminal residues relieves inhibition. Growth curves **(A)** and CFUs **(B)** of cultures expressing EzeTΔN83 (red), EzeTΔN62 (orange), EzeTΔN36 (yellow), EzeT(M5A_D8A_N12A_Q16A) (light green) or full-length EzeT (dark green) at 20 ºC. Similar to expression of the C-terminal domain (EzeTΔN83), expression of EzeTΔN36 and EzeTΔN62 leads to a decrease in OD_{600}, albeit with a longer lag phase of 360 min and 210 min, respectively. A reduction in CFUs is visible for both cultures after 180 min. In addition, a decrease in CFUs is visible for EzeT(M5A_D8A_N12A_Q16A), while the OD_{600} increases similarly to non-toxic full-length EzeT.

3.1.10 EzeT is expressed from its endogenous promoter

To monitor expression from the endogenous *ezeT* promoter an HA-tagged version of EzeT was integrated into the genome of *E. coli* W in place of the wild-type *ezeT* locus yielding *E. coli* W(EzeT-HA). This strain and wild-type *E. coli* W were grown at 37 ºC to an OD_{600} of 1, harvested and lysed. The cleared lysate was then subjected to immunoprecipitation with antibodies against the HA-epitope coupled to magnetic beads. Bound proteins were eluted by denaturation with 8 M urea and analyzed by Western blotting. A species with the same electrophoretic mobility as recombinant HA-tagged EzeT was detected in the lysate of *E. coli* W(EzeT-HA), but was not present in the lysate of *E. coli* W (Figure 3.1-10A). This proved the expression of full-length EzeT *in vivo* and showed that the protein is indeed synthesized as an autoinhibited UNAG kinase. However, under the applied growth conditions, the expression of EzeT was low as the protein could hardly be detected without enrichment.

3.1.11 Regulation of *ezeT* transcription

To identify factors that influence the expression from the *ezeT* promoter, the genomic region comprising 250 nucleotides upstream of the EzeT start codon was cloned into a reporter plasmid (pREP(sfGFP). In this plasmid, *ezeT* promoter activity can be probed by spectroscopically following the synthesis of superfolder GFP (sfGFP) [84]. Indeed, GFP fluorescence was detected in *E. coli* DH5α cells transformed with this reporter plasmid. The fluorescent signal was followed from initial inoculation at an OD_{600} = 0.1 till stationary phase, but was only discernible after more

than 5 h, when a high cell density was reached and growth had ceased. This probably depends on a number of different factors that influence the amount of mature GFP: the increasing number of cells, the time needed for GFP maturation and, potentially, a dependence of GFP expression on growth phase.

Bioinformatic analysis of the promoter region (Figure 3.1-10B) indicated the presence of a conventional Pribnow-Schaller box (TATAAT) [136, 137] 31 nucleotides upstream of the start codon of EzeT. This Pribnow-Schaller box is usually bound by the housekeeping sigma factor σ^{70} [138]. In addition, this sigma factor recognizes a region approximately 25 nucleotides further upstream (-35 region). However, the consensus sequence TTGACA is not present at this position in the *ezeT* promoter. To verify the regulatory function of the Pribnow-Schaller box, this sequence was mutated to random nucleotides (pREP(delPS)) and GFP fluorescence was measured (Figure 3.1-10C). No fluorescence could be detected, showing that this region is essential for transcription factor binding. Similarly, randomization of the Shine-Dalgarno sequence (AGGAG, pREP(delSD)) required for ribosome binding to the mRNA completely abolished GFP fluorescence. In contrast, mutation of the -35 region to a conventional binding site for σ^{70} (pREP(s70)) strongly increased fluorescence levels. These results show the functionality of the Pribnow-Schaller box and the Shine-Dalgarno sequence and suggest that the promoter is weak under standard growth conditions, but might be activated in specific situations.

To further investigate transcriptional regulation, individual parts of the sequence were mutated subsequently. Most alterations in the promoter region did not significantly change the GFP fluorescence levels, excluding that the mutated nucleotides are important for transcriptional regulation. Nevertheless the promoter sequence upstream of the Pribnow-Schaller box seems to contain elements required for transcription, as deletion of this entire region (pREP(delprom)) reduced fluorescence. Notably, GFP fluorescence was increased upon mutation directly downstream of the Pribnow-Schaller box (pREP(delPSdown)) and for random mutations in the -35 region (pREP(del35)).

It is conceivable, that the *ezeT* promoter evolved to restrict protein expression to a low level under normal growth conditions and that transcription is only activated under certain conditions, for instance by binding of the alternative sigma factor σ^{38} at high cell densities or under stress. σ^{38} recognizes promoters with similar consensus sequences as σ^{70} but has a higher tolerance to deviations from the ideal sequence and spacing of the promoter elements [139]. However, when *E. coli* cells carrying the reporter plasmid were subjected to different stress conditions no significant alteration in GFP fluorescence levels was detected. The stress conditions tested included different osmolarity, pH or nutrient availability and the exposure to reducing/oxidizing agents, UV irradiation or antibiotics.

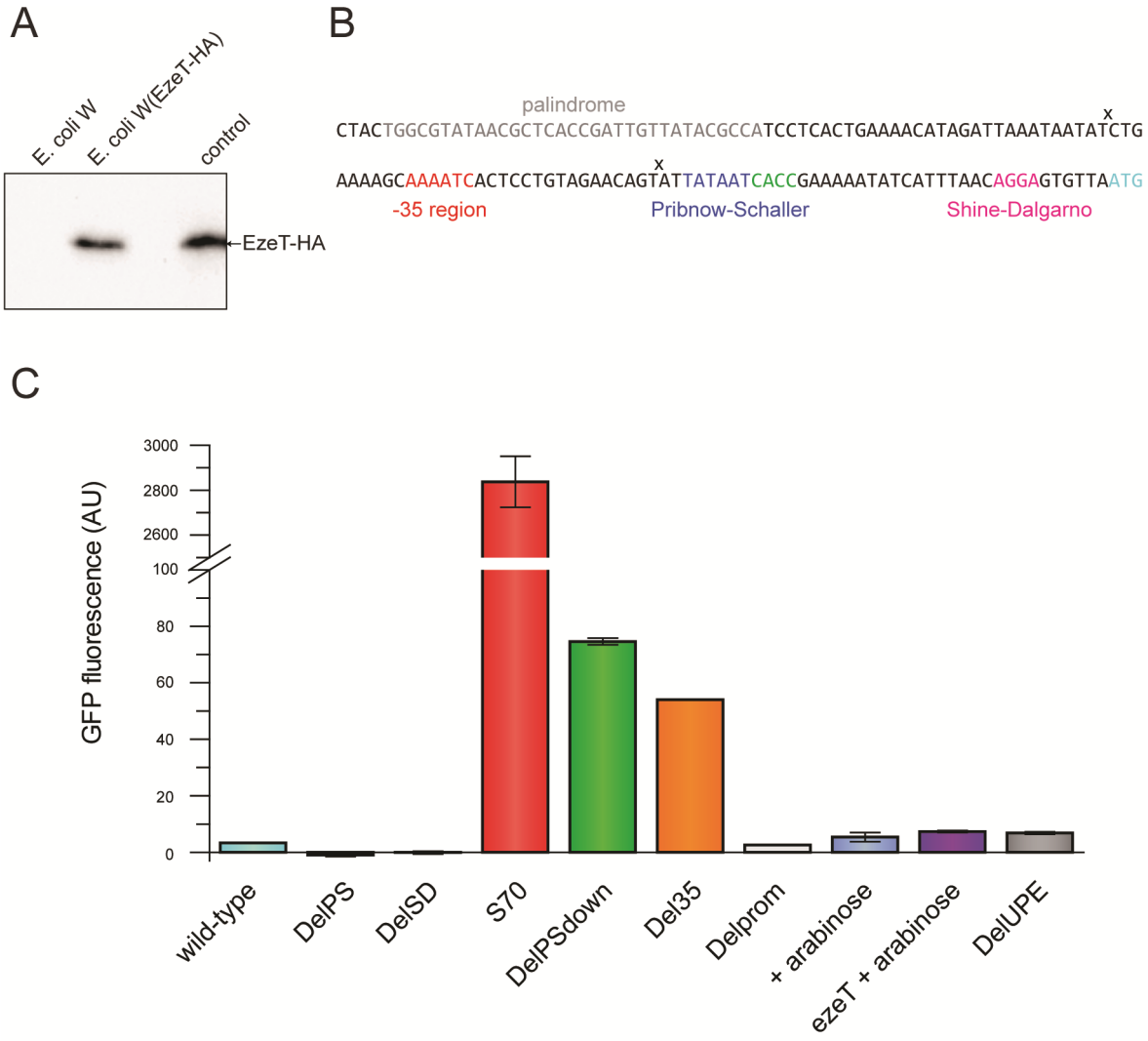


Figure 3.1-10: EzeT expression from its endogenous promoter. (A) Western Blot detection of hemagglutinin-epitopes after immunoprecipitation of tagged proteins from *E. coli* W and *E. coli* W(EzeT-HA). A protein with the same electrophoretic mobility as recombinant EzeT-HA (control) is only present using *E. coli* W(EzeT-HA). **(B)** *ezeT* promoter sequence upstream of the start codon (light blue). The putative Shine-Dalgarno sequence (pink), the Pribnow-Schaller box (dark blue) as well as the corresponding location of the -35 region (red) and the palindrome (gray) are highlighted. The 'x' denotes deletion sites, indicating the 3'-end of the putative promoter region that has been deleted. **(C)** Analysis of ezeT promoter activity measured as GFP fluorescence intensity of dense cultures. The colors indicate the region of the promoter which was mutated according to (B) (Pribnow-Schaller box, DelPS, blue; Shine-Dalgarno, DelSD, pink; -35 region, S70 (to ideal σ70 sequence), red, and Del35 (to random sequence), orange; region downstream of the Pribnow-Schaller box, DelPSdown, green; the wild-type *ezeT* promoter is colored light blue. Delprom denotes a deletion upstream of X, DelUPE a deletion upstream of x. No difference is seen in GFP fluorescence after addition of arabinose in the presence (purple) or absence (violet) of EzeT under control of an arabinose-inducible promoter excluding autorepression of the promoter.

In addition to sigma factors and the RNA polymerase, DNA-binding proteins that act as activators or repressors can alter transcriptional rates. In many type II TA systems protein synthesis is regulated by autorepression by the antitoxin or the toxin-antitoxin complex [18]. A DNA-recognition motif of the antitoxin binds to palindromic sequences in the promoter region, e.g. the *E. coli* TA systems MazEF (overlapping with the core promoter, nucleotides -34 to +6 [140]) and RelBE (overlapping with the transcription start site, nucleotides -5 to +19 [141]). A palindromic sequence was found upstream of the *ezeT* start codon. However, this sequence could not be attributed to binding of a known bacterial transcription factor and no homologous palindromes could be identified in the *E. coli* genome. This could in principle imply gene-specific regulation by autorepression of EzeT, but the palindromic sequence is located approximately 120 nucleotides upstream of the start codon, which is presumably to far for the bound protein to interact with the RNA polymerase complex.

To investigate a possible repressor function of EzeT, cells were co-transformed with the GFP-reporter plasmid and a plasmid encoding for arabinose-inducible EzeT and the fluorescence intensity was compared to the fluorescence of cells carrying only the reporter plasmid (Figure 3.1-10C). No difference in intensity was measurable between both cultures upon addition of arabinose, thereby excluding autorepression of the *ezeT* promoter. In addition, GFP fluorescence of cells carrying a reporter plasmid in which the entire region more than 70 nucleotides upstream of the start codon was removed (pREP(delUPE)) did not change when compared to the GFP fluorescence observed for the original plasmid. The experiments were performed with the *E. coli* DH5α laboratory strain, which does not have an endogenous *ezeT* locus, and with *E. coli* W, which harbors the *ezeT* locus, thereby excluding strain specific regulation.

3.1.12 EzeT activation

Expression from the endogenous *ezeT* promoter yields full-length protein, in which the N- and C-terminal domain are covalently bound. This protein was demonstrated to be non-toxic, while the separated C-terminal domain alone showed UNAG kinase activity *in vivo*, suggesting that the inhibition by the N-terminal domain has to be relieved to activate the toxic function.

In type II TA systems, the inactive protein-protein complex dissociates and the antitoxin is degraded to activate the toxin [18]. In EzeT, complete dissociation of the two domains is prevented by the connecting linker. Instead, toxin activation could involve binding of a factor that favors an ‘open’ complex in which the active site is not blocked or chaperones that induce unfolding of the N-terminal domain. Furthermore, proteases could degrade the N-terminal part of EzeT or cleave the interdomain linker followed by domain separation.

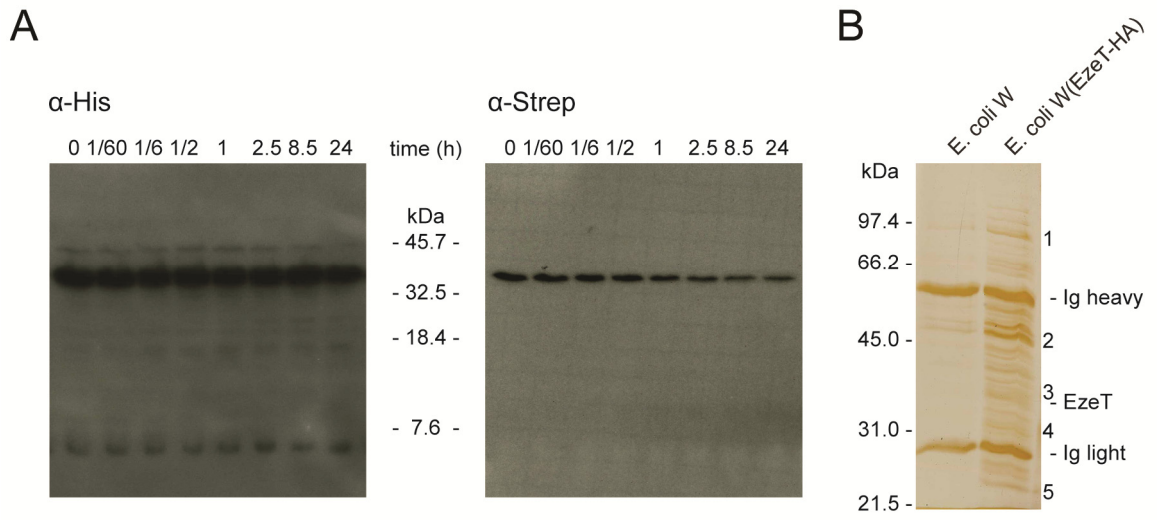


Figure 3.1-11: EzeT activation. (A) Western Blot detection of the His_6- (left) and Strep-epitopes (right) of N-terminally Strep- and C-terminally His_6-tagged EzeT after incubation in *E. coli* W lysate for the indicated time periods. **(B)** Silver stained SDS-PAGE gel of a co-immunoprecipitation using HA-tagged EzeT from *E. coli* W(EzeT-HA). The amount of unspecific binding was monitored following the same protocol using *E. coli* W lysate that does not contain tagged EzeT. Additional and more intense bands are observed when HA-tagged EzeT is present. Five of the co-precipitated proteins could be identified by mass spectrometry. 1: Formate C-acetyltransferase 1, 2: EF-TU, 3: Threonine dehydratase, 4: MinD, 5: 50S ribosomal protein L6.

Proteolytic degradation of EzeT in *E. coli* lysate was monitored by mixing purified protein with cell lysate. The protein contained an N-terminal Strep-tag and a C-terminal His_6-tag to be able to independently detect N- and C-terminally truncated fragments in a Western Blot (Figure 3.1-11A). However, EzeT proved to be very stable over 24 h and no faster migrating, supposedly cleaved fragment was observed.

In addition, co-immunoprecipitation of proteins interacting with endogenous EzeT-HA was performed. In a silver stained gel (Figure 3.1-11B) additional protein species were observed that were more prominent than in the control in which the HA-tag was not present. Five co-eluting proteins could be identified by mass spectrometry as formate C-acetyltransferase 1 (*pflB*), EF-Tu, threonine dehydratase (*tdcB*), MinD and the 50S ribosomal protein L6 (*rplF*). Of these, formate C-acetyltransferase 1 and the proteins of the translation machinery EF-Tu and L6 are common cytosolic proteins [142] that might have bound unspecifically. Nevertheless, it should be mentioned that an interaction of the homolog PezT with ribosomes was observed previously [143], although the functional implications of this binding are unclear. MinD is required for correct positioning of the septum at midcell [144] and threonine dehydratase is involved in the synthesis of branched-chain amino acids [145]. None of the identified proteins is a protease or chaperone that could be involved in EzeT activation and the *in vivo* functionality of an interaction with those co-purified proteins is enigmatic.

In conclusion, no activation of EzeT was observed under the tested conditions. As only standard growth conditions were applied, it is still conceivable that EzeT activity is tightly regulated and only evoked under certain circumstances. However, how and when this occurs is unknown both for its transcriptional as well as enzymatic regulation.

3.1.13 EzeT function

Originally, TA systems were discovered as plasmid addiction systems that ensure the persistence of a plasmid in a bacterial population [27] and even some chromosomally encoded TA systems were found to stabilize genetic loci [38].

To test whether EzeT is able to stabilize plasmids, the *ezeT* ORF was cloned into a derivative of the enteric low-copy number plasmid R1, pOU82 [86]. pOU82 contains an ampicillin resistance gene, the *lacZYA* operon and the origin of replication of R1, but does not encode CopA. CopA is an antisense RNA that negatively regulates RepA and thereby downregulates replication. To study stabilization efficiencies, CopA is provided from an IPTG-inducible promoter on a second plasmid, thereby arresting pOU82 replication. After the addition of IPTG, no further copies of pOU82 are synthesized and plasmid numbers stay constant. After a few rounds of cell divisions each mother cell will only contain a single plasmid, such that only one daughter cell can inherit the encoded genetic information. If the plasmid contains a module that induces PCD upon loss of the plasmid, all plasmid-free cells will be killed, which limits the number of cells to the number of plasmids present. In this case, the number of CFUs should remain constant after addition of IPTG, while they will steadily increase, when PCD is not induced and the plasmid-free daughter cells survive.

The loss rates of plasmids with and without *ezeT* were determined, but no difference in plasmid stabilization efficiency was observed (Figure 3.1-12). Instead, CFUs steadily increased, while the number of colonies containing pOU82 remained constant. Furthermore, there was no significant difference detectable, whether EzeT was under control of its own *ezeT* promoter or a constitutive *lacI* promoter. Apparently, EzeT belongs to the group of chromosomal TA systems that do not stabilize their genetic loci, like the ccdAB modules of *Erwinia chrysanthemi* and *E. coli* [40, 146].

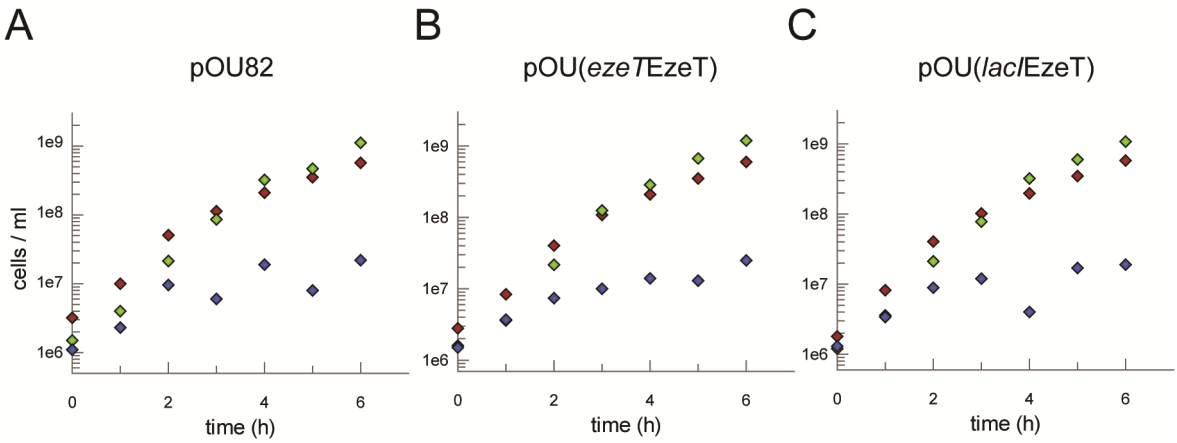


Figure 3.1-12: Plasmid maintenance assay. pOU82 **(A)** is a conditionally unstable plasmid and the stabilizing effect of EzeT under control of the *ezeT* promoter **(B)** or the constitutive *lacI* promoter **(C)** was tested. Cell numbers per ml were determined by CFUs (green) and calculated from OD_{600} (red) and compared to the number of bacteria carrying pOU82 (blue). Due to the replication arrest of the pOU82 derivatives the number of plasmids remains constant, which is equivalent to a constant value for the plasmid containing cells (blue). If a plasmid stabilizing TA system is present, the plasmid-free cells are killed and culture growth ceases. The total number of viable cells (green) will then equal the number of plasmid containing cells (blue). Without a plasmid stabilization system, total cell numbers increase, while the number of plasmid containing cells remains constant, revealing the emergence of viable plasmid-free cells. This scenario is observed for all three plasmids, with or without EzeT, indicating that the protein does not induce post-segregational killing.

Chromosomal TA systems are often regulated in response to external stimuli and induce phage abortive infection, biofilm and persister cell formation. To test whether EzeT is involved in any of these processes, assays were performed with *E. coli* W and an *ezeT* knock-out strain (*E. coli* W *ezeT*$^-$).

The ability of EzeT to cause abortive infection upon exposure to T4 phage was tested in a plaque assay. However, *E. coli* W was not sensitive to T4 phage infection. Therefore, *E. coli* DH5α cells were transformed with pBAD vectors containing EzeT or EzeT_K121A or an empty control vector and exposed to the phages. Although a considerable number of plaques was observed, no difference was seen between the different constructs. In conclusion, EzeT does not induce abortive infection in response to T4 infection.

Similarly, the propensity to form biofilms can depend on the expression of certain TA systems [147], which was compared for *E. coli* W, *E. coli* W *ezeT*$^-$ and *E. coli* W *ezeT*$^-$ harboring pBAD(EzeT). Cultures were grown in 96-well plates at 30 ºC for 48 h without agitation to allow attachment. Biofilm formation was quantified by crystal violet staining of adherent cells after rigorously washing the wells. No difference was seen between the amounts of biofilm formed for both strains, indicating that EzeT does not influence biofilm formation.

Persisters are dormant cells with a reduced metabolism, which makes them tolerant to antibiotics [12]. Their frequency is thought to be regulated by various factors including

TA systems [148] and can be assessed by determination of the number of CFUs after exposure to different antibiotics. To test whether EzeT can influence the number of persisters in a population, stationary phase cultures of *E. coli* W, *E. coli* W *ezeT*⁻ and *E. coli* W *ezeT*⁻ harboring pBAD(EzeT) were treated with gentamicin, tetracycline and ciprofloxacin and CFUs were determined. No difference was seen between all three strains, indicating that EzeT does not change the relative amount of persister cells in response to these antibiotics.

3.1.14 Summary

The results described in this section provide the characterization of EzeT, a type II-like TA system with a novel mechanism of inhibition. The protein is the first characterized UNAG kinase in Gram-negative bacteria and expression of the C-terminal domain leads to a lytic phenotype caused by the accumulation of UNAG-3P. EzeT toxicity is temperature dependent, which is in accordance with low protein stability and a thermal melting point of 38 ºC *in vitro*. In the presence of the N-terminal domain the UNAG kinase activity is abrogated, presumably by in an epsilon-like mechanism of inhibition. EzeT is the first described TA system, in which the toxin and antitoxin functionalities are linked in one polypeptide chain, thereby constituting a new type of proteinaceous TA system.

3.2 *Neisseria gonorrhoeae* epsilon/zeta systems

This section describes the characterization of the three plasmidic epsilon/zeta systems of *N. gonorrhoeae*, ε1/ζ1, ε2/ζ2 and ε3/ζ3, all of which were found to form protein-protein complexes and to be able to phosphorylate UNAG *in vitro*. Beyond these characteristics of epsilon/zeta TA systems, several variations to the streptococcal systems have been observed. In particular, structural and biophysical studies on ε1/ζ1 showed a novel zeta toxin topology and a different mode of complex formation and inhibition.

3.2.1 Multiple epsilon-zeta systems on *N. gonorrhoeae* plasmids

Antibiotic resistance genes in the Gram-negative bacterium *N. gonorrhoeae* are often carried and transferred by plasmids and spread by horizontal gene transfer [149]. Conjugative plasmids like those of the pLE2451-group contribute on the one hand to the mobilization of genetic material [150]; on the other hand, they themselves contain genetic load regions that encode genes, which increase pathogen survival or virulence, like the tetracycline-resistance gene *tetM* [81]. Sequencing of pLE2451-like plasmids with and without the *tetM* determinant revealed the presence of two putative epsilon/zeta systems in their genetic load regions [81].

One of these systems, termed ε1/ζ1, was found on all pLE2451 variants. The closest homolog of ζ1, as identified with 50 % aa sequence identity, is CAMGR0001_1552 of *Campylobacter gracilis*, a protein that has been annotated as putative zeta toxin. Therefore, the ε1/ζ1 locus has been proposed to encode a hitherto uncharacterized epsilon/zeta system [81].

The epsilon/zeta homolog encoded adjacent to ε1/ζ1 differed between plasmids with or without the *tetM* determinant. In plasmids without antibiotic resistance gene it was designated ε2/ζ2 and in those that carry *tetM* ε3/ζ3, respectively (Figure 3.2-1). Nevertheless, both systems share 67 % nucleotide sequence identity and were proposed to be closely related [81].

ζ2 and ζ3 contain a region of 260 aa with 38 % sequence similarity to PezT and ζ_{spy} (Figure 3.2-2D). In particular, the Walker A motif, the UNAG binding motif and the magnesium coordinating DxD motif are conserved [66], indicating that both proteins are functional UNAG kinases. Furthermore, the predicted secondary structure of ζ2 and ζ3 corresponds well to the structure of the streptococcal zeta toxins (Figure 3.2-2D).

In contrast, the secondary structure prediction of ζ1 differs from the structure of these toxins, especially in the N-terminal part of the polypeptide chain (Figure 3.2-2B). Structural differences are also evident in a shifted position of the Walker A motif in the aa sequence, which was identified as residue 115 in ζ1, whereas it can be found at position 64 in ζ2 and ζ3. Regions with homology to the DxD motif or the UNAG binding motif could not be identified by sequence alignment.

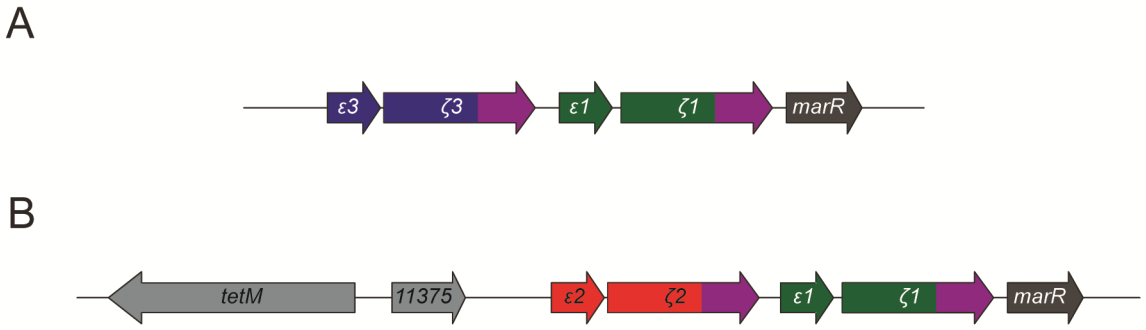


Figure 3.2-1: Comparison of the genetic organization of the gonococcal epsilon/zeta systems. Schematic representation of the epsilon/zeta systems of pLE2451 **(A)** and a variant carrying the *tetM* determinant **(B)**. ε1/ζ1 (green) and the transcriptional repressor MarR (dark gray) are encoded on both plasmids, while ε3/ζ3 (blue) is found only in plasmids, which do not carry an antibiotic resistance gene. The tetracycline-resistance determinant *tetM* and the hypothetical ORF 11375 (light gray) were integrated into the genetic load region adjacent to ε2/ζ2 (red). The conserved C-terminal domain present in all gonococcal zeta proteins is marked in purple.

All three proposed zeta toxins are located downstream of ORFs encoding small, 61 to 84 aa proteins, which were called ε1, ε2 and ε3, respectively. The close proximity of the corresponding epsilon and zeta genes suggests that they are transcribed in one mRNA.

Similar to the toxin homologs ζ2 and ζ3, ε2 and ε3 share 56 % aa sequence identity (Figure 3.2-2C), while ε1 is different both in aa sequence and in length (Figure 3.2-2A). Although the aa sequence similarity of all three proteins to ε_{spy} and PezA is low (approximately 20 %), secondary structure prediction indicates the presence of three α-helices in all gonococcal epsilon homologs (Figure 3.2-2A/C), that might adopt the three-helix bundle fold described for the streptococcal antitoxins [62, 63].

Figure 3.2-2: Amino acid sequences of gonococcal epsilon/zeta systems. (A) Amino acid sequence of ε1. The protein was predicted to fold into three α-helices (blue bars). **(B)** Amino acid sequence of ζ1 excluding the 130 residues at the C-terminus. The identified Walker A motif is highlighted in red, the Walker A lysine (K115), which was mutated in this study, is marked with a diamond (♦). Predicted secondary structure elements are shown as green bars (α-helices) and pink arrows (β-strands). **(C)** Amino acid sequence alignment of ε2 and ε3. Residues are colored according to conservation (dark green: identity to yellow: low homology). The proteins were predicted to fold into three α-helices (blue bars). **(D)** Amino acid sequence alignment of ζ2, ζ3, ζ_{spy} and PezT. Secondary structure elements of ζ_{spy} are indicated above the alignment as olive bars (α-helices) and red arrows (β-strands). The catalytically important Walker A (red) and DxD (purple) motifs and the UNAG binding region (gray) are conserved and the Walker A lysines (K46 in ζ2 and ζ3), which were mutated in this study, are marked with a diamond (♦). Note that the additional C-terminal domains of ζ2 (residues 266 to 402) and ζ3 (residues 266 to 403) are excluded from this alignment and are shown in **(E)** together with the C-terminal domain of ζ1 (residues 271 to 401) to show the high degree of conservation in this region.


A
ε1 MNKVEPQESNAIRMIKEACEKNRRMMTDEAFRKEVEKRLYAGPSPELLAKLRVLWAANKEQ
B
ζ1 MVKLSSDINLRDFGNNEYLSSVQDEAIRFATEQTDEILSLYSQHADTEGGRYVCADTFKELFPAFENKED
ζ1 RATVNNAIHNSAAVLSSTQFDEVLKRDEPQKKEVIFVTGIPGSGKTSTVKNMMMQDTTKLLFEGQLARPQ
Walker A
ζ1 SAFRKIEQCLERNLEVTIVAVSMRAERASDNTYKRFNEYGRGASIGIMADIQANLPDGLKQIRDKFGDAV
ζ1 KIVGINQDRNSEFIDKFDDVIKMLSLGSQEQILGRLAEKIQSDFDSGKISRECFNQAKGSM [130 aa]
C
ε2 MQMQETIPGINYEVALELLGDSMRPLIQARESEFVKDKPNQAFIEYCTARLEAIDMLRDSLSPEDTDIIA
ε3 MKEVIPGINYEVALEVLGNAAQPFIQARETERKKSHPNEAFIQYCSDSIKAITVFRHNLRATDEDVIA
ε2 RIFNPDDQTVGRVK
ε3 RILDPNDTIVGGSV
D
ζ2 MIEYTPLSAEGKARILNGFMKPRLSRTQSVEQPKIVLVGAQPGAGKSKAASLAKSELRQEGGYIHVDADI
ζ3 MSDYQKLSDAGRAEIVAEYMSALLEITQAVDVPQIALVAAQPGAGKSKTADIVKEEFASKGGHIHVDADI
PezT MEIQDYTDSEFKHALARNLRSLTRGKKSSKQPIAILLGGQSGAGKTTIHRIKQKEF--QGNIVIIDGDS
ζspy MANIVNFTDKQFENRLNDNLEELIQGKKAVESPTAFLLGGQPGSGKTSLRSAIFEET--QGNVIVIDNDT
Walker A
ζ2 MRALIPAPEGVVYSS-----EQTQKDAGALAISVRNSAKENRRNIVEEGTFRNAASISQFIRDRKSEGYG
ζ3 MRQKIPVPPGVVYSS-----QQTQEDAGKLAVGVRKSALENSRNVLEEGTFRNAEAVGMSIKAARERAGLK
PezT FRSQHPHYLELQQEYGKDSVEYTKDFAGKMVESLVTKLSSLGYNLLIEGTLRTIDVPKKTAQLLKNKGYE
ζspy FKQQHPNFDELVKLYEKDVVKHVTPYSNRMTEAIISRLSDQGYNLVIEGTGRTTDVPIQTATMLQAKGYE
UNAG binding
ζ2 VEMLAVATASEESVAGIFKRYEEQHAKGVSQPRFVEESYHNEAMAGFKDTLSQ--CESSFDRVRVTNRAG
ζ3 IEMLAVATAPEESLAGIFKRYEDQYLTKNIQPRFVDEDFHNKAFEGFKNTVAT--HEAEFDRIRVTNRPG
PezT VQLALIATKPELSYLSTLIRYEELYIINPNQARATPKEHHDFIVNHLVDNTRKLEELAIFERIQIYQRDR
ζspy TKMYVMAVPKINSYLGTIERYETMYADDPMTARATPKQAHDIVVKNLPTNLETLHKTGLFSDIRLYNREG
ζ2 DILYDSLNRRQNQHETAKDALSAYQEITPKRLKQVVKAWDEIQLQAESRSIDPIPNYLGMVKQHSEA [132 aa]
ζ3 EILYDSLNKQQNKQASAKDAMEFYQQITPERLKQVAQVWDVIQLQADRRSQDPVPNYFDKVKQHREE [134 aa]
PezT SCVYDSKENTTSAADV-LQE-LFF----GE---------WSQVEKEMLQVGEKRL--------NELLEK
ζspy VKLYSSLETPSISPKETLEK-ELNRKVSGK---EIQPTLERIEQKMVLNKHQETPEFKA-IQQKLES [17 aa]
E
ζ1Δ270 MDLESVF--AKKEYSQQRVVTNSKGVTLETKSANELWSKVEQIPVTGMKAGIYLLGQAKKAE-TGQTYSG
ζ2Δ265 QHSEAIYQRVEEIYRQERVVANSEGATLQRKS-GDTWQDIEKAEAKGMKAGIHMLGTAKPAKS-GREYSG
ζ3Δ265 QHREEIYQRVEEIYRQERVVANSEGATLQRKS-GDTWQDIEKAEAKGMKAGIHMLGTGETGRIPAKSTVE
ζ1Δ270 EIIYKDAAAVFQKTKNGLVRHNATHNE----ERLAKLVEIGQNVSIGSNKGKLIVKSLEYSAKKSISR
ζ2Δ265 EIVHKDEASVFQKTDQGLIRHKAVQGMAEGKFSSLSEQVEIGQKVSIKREGNELSVKPADASLKKTMKR
ζ3Δ265 EIVHKDEASVFQKTDQGLIRHKAVQGMSEGKFSSLSEQVEIGQKVSIKREGNGLSVKASDASVKKTMKR

Further investigation of the sequences of the gonococcal zeta proteins indicated that - similar to EzeT – all three polypeptide chains are longer when compared to their streptococcal homologs, with 401 aa, 402 aa and 403 aa for ζ1, ζ2 and ζ3, respectively. In contrast, PezT consists of 253 aa and ζ_{spy} 287 aa. The approximately 140 additional residues are located at the C-terminal end of the three proteins and form a conserved putative domain with high mutual sequence identity (43 % to 86 %, Figure 3.2-2E). The function of these C-terminal domains has not been investigated yet, but they show homology to the plasmid stable inheritance protein KfrB and the conjugal transfer protein TraB.

Except for the additional C-terminal domain, ε2/ζ2 and ε3/ζ3 are highly similar to the streptococcal epsilon/zeta systems; in contrast, ε1/ζ1 differs much more in its primary and secondary structure. Research on the gonococcal epsilon/zeta homologs was, therefore, primarily focused on this module, aiming to confirm ε1/ζ1 as a *bona fide* epsilon/zeta system consisting of a UNAG kinase and an inhibiting antitoxin. This might help to answer the question why two epsilon/zeta homologs coexist on one plasmid.

3.2.2 ε1/ζ1 complex formation

In type II TA systems, inhibition is performed by protein-protein complex formation [18]. For the streptococcal epsilon/zeta systems, these complexes were shown to adopt a heterotetrameric $epsilon_2zeta_2$ arrangement [62, 70]. To analyze the oligomeric state of ε1/ζ1, the complex was purified by metal affinity chromatography, cation exchange chromatography and size exclusion chromatography, which yielded pure recombinant complex used in light scattering experiments.

Dynamic light scattering experiments yielded an apparent molecular mass of 56.0 kDa for the complex (Figure 3.2-3A/B), while static light scattering indicated a mass of 43.4 kDa (Figure 3.2-3C). Both experiments indicate that ε1/ζ1 forms a heterodimer in solution given a theoretical molecular mass of 51.9 kDa.

3.2.3 UNAG kinase activity of ε1/ζ1

The purified complex containing wild-type ε1/ζ1 was subjected to an *in vitro* phosphorylation assay as described for EzeT (see section 3.1-3). The reaction mixture was separated by anion exchange chromatography after 3 h of incubation and monitored spectroscopically (Figure 3.2-3D). A species with a retention time corresponding to UNAG-3P and the typical uridine absorption ratio of 2.8:1 (absorption at 260 nm : absorption at 280 nm) was formed and 50 % of the substrate was turned-over after 3 h. Surprisingly, a significant amount of

ADP was produced. When integrating the area under the curves of the absorption at 260 nm for ADP and UNAG-3P and scaling those values by the extinction coefficients of ADP (15,400 M^{-1} cm^{-1}) and UNAG-3P (10,100 M^{-1} cm^{-1}), respectively, a 1.5 fold higher production of ADP when compared to UNAG-3P was observed. In contrast, UNAG-3P formation by PezT and EzeT involved equimolar amounts of UNAG and ATP.

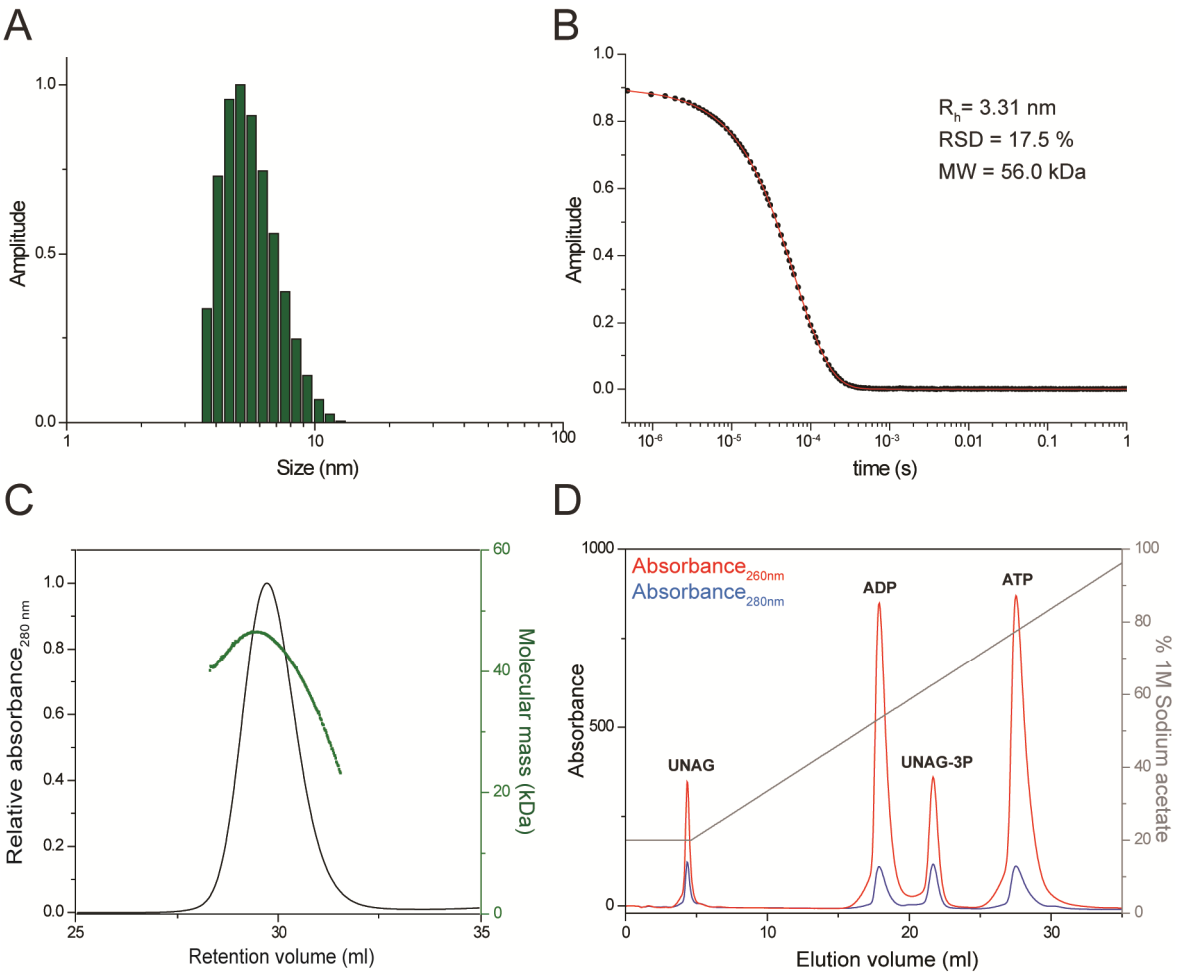


Figure 3.2-3: ε1/ζ1 forms a heterodimer in solution and phosphorylates UNAG. (A) Relative distribution of the hydrodynamic radius (R_h) of ε1/ζ1 as determined by dynamic light scattering. **(B)** shows the experimentally determined autocorrelation function (black) and the corresponding fit (red), which allowed to determine the molecular mass (MW) as 56.0 kDa with a relative standard deviation (RSD) of 17.5 % **(C)** Elution profile of purified ε1/ζ1 (black line) from a Superdex 200 size exclusion column. The molecular masses of the species in the peak have been determined by static light scattering (green dots). A slight tailing of the eluting peak suggests that the complex dissociates during size exclusion chromatography at the concentration used (final concentration in peak approximately 2 µM), pointing at a K_d in the micromolar range. **(D)** Separation of nucleotides/ nucleotide sugars by anion exchange chromatography after incubation of ε1/ζ1 with UNAG and ATP for 3 h. The ratio of the absorption at 260 nm and 280 nm helps to distinguish compounds containing adenine (ratio ~7:1) from compounds containing uridine (ratio ~2.8:1).

3.2.4 Structure determination of the ε1/ζ1 complex

Despite the differences in amino acid sequence and predicted secondary structure of ζ1 when compared to the streptococcal zeta toxins, it could be shown that ε1/ζ1 is an active UNAG kinase encoded on a plasmid of the Gram-negative bacterium *N. gonorrhoeae*,. To gain insights into the structural characteristics that allow UNAG binding and the phosphoryl transfer reaction, the structure of the ε1/ζ1 complex was determined by X-ray crystallography.

Crystallization of ε1/ζ1_K115A

Purified ε1/ζ1_K115A complex was subjected to initial crystallization trials by Roman Sakson, revealing $(NH_4)_2SO_4$ as a suitable precipitant for crystal formation. Initially, crystals grew as clusters of thin needles, but stacks of thin plates were obtained after proteolytic removal of the His_6-tag of ε1. Further optimization of the crystallization conditions in order to slow down crystal growth yielded larger and more separated individuals (Figure 3.2-4). For data collection, crystals were grown in 1.8 M $(NH_4)_2SO_4$ and 100 mM sodium citrate pH 4.5 using a protein concentration of 21.5 mg/ml. Individual plates were broken from the stacks and transferred to solutions containing cryoprotectants before flash-cooling in liquid nitrogen. Different substances with the ability to prevent ice crystal formation during flash-cooling were investigated for their influence on diffraction quality, but yielded comparable results. In several cases, multiple diffraction lattices were observed, most likely due to slightly different orientations of the crystal layers within one morphological unit. Moreover, crystals diffracted anisotropically and strong variations in diffraction quality were observed, which required harvesting and testing of multiple crystals for optimization.

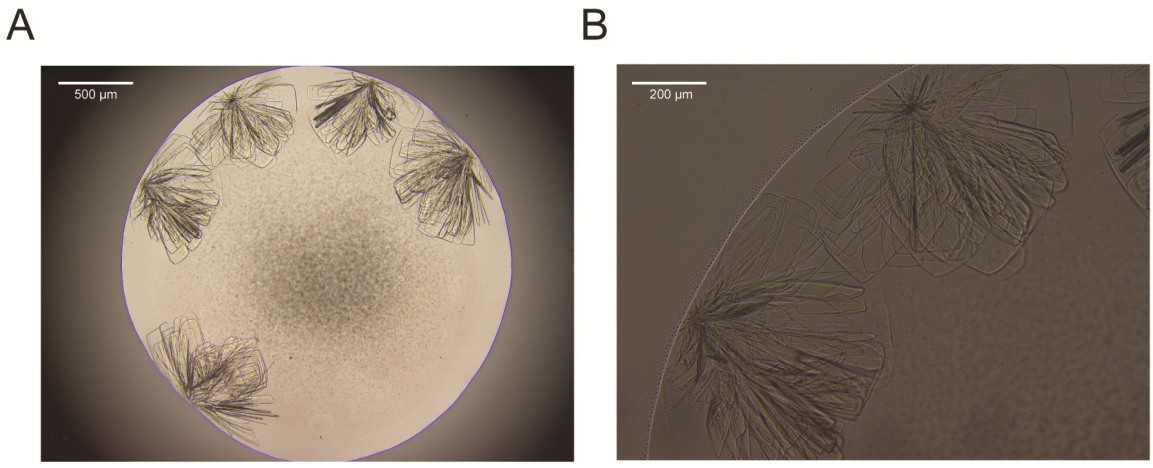


Figure 3.2-4: Crystals obtained for ε1/ζ1_K115A. (A) Crystals grew as clusters of large, thin plates in 1.8 M $(NH_4)_2SO_4$ and 100 mM sodium citrate pH 4.5 after proteolytic removal of the His_6-tag of ε1. **(B)** shows a 2.5 fold magnification from (A).

Eventually, diffraction data for crystals of native ε1/ζ1_K115A were collected to a resolution of 2.6 Å at the X10SA beam line of the Swiss Light Source. The crystal symmetry belonged to space group $P2_1$ with unit cell dimensions $a = 80.99$ Å, $b = 148.89$ Å, $c = 95.06$ Å, $\alpha = 90$ °, $\beta = 98.32$ °, $\gamma = 90$ °. The Matthews coefficient of 2.73 Å^3/Da indicated 4 dimers per asymmetric unit with a solvent content of 54.9 % and molecular replacement with search models derived from PezAT or ε/ζ_{spy} was attempted. However, the obtained phase information was not sufficient to build an initial model and, instead, phases were determined experimentally using SAD of SeMet-labeled protein crystals.

SeMet-labeled ε1/ζ1_K115A crystallized from the same conditions and in a similar plate-like morphology as native ε1/ζ1_K115A. Highly-redundant SAD diffraction data were collected to a resolution of 2.4 Å. To reduce the effects of radiation damage during the long exposure, the crystal was translated during data collection and diffraction images were collected in 4 wedges of 1800 images each. Although the crystal symmetry belonged to the same space group, $P2_1$, the crystals of the SeMet-labeled protein were non-isomorphous to the native protein crystals with unit cell parameters $a = 79.76$ Å, $b = 149.58$ Å, $c = 125.09$ Å, $\alpha = 90$ °, $\beta = 94.82$ °, $\gamma = 90$ °. The data collection and refinement statistics of ε1/ζ1_K115A are summarized in Table 3.2-1.

Crystallization of wild-type ε1/ζ1

In order to closer investigate the active site region of ε1/ζ1, wild-type protein was purified and crystallized in the *apo*-form and in the presence of nucleotides.

Apo-wtε1/ζ1 crystallized from the same conditions as ε1/ζ1_K121A in a similar plate-like morphology and diffraction data were collected to a resolution of 2.9 Å. The crystal symmetry belonged to space group $P2_1$ with unit cell parameters similar to the crystals of SeMet-labeled ε1/ζ1_K115A with $a = 79.72$ Å, $b = 150.72$ Å, $c = 124.59$ Å, $\alpha = 90$ °, $\beta = 94.95$ °, $\gamma = 90$ °. Nevertheless, the unit cell parameters, especially the monoclinic angle β, varied slightly between crystals even within the same crystal shape indicating some degree of flexibility in the protein or the molecular contacts formed within the crystal.

Table 3.2-1: X-ray data collection and refinement statistics of ε1/ζ1_K115A

Crystal	**Native**	**SeMet-labeled**
Space group	$P2_1$	$P2_1$
Cell parameters		
a, *b*, *c* [Å]	80.99, 148.89, 95.06	79.76, 149.58, 125.09
α, *β*, *γ* [°]	90, 98.32, 90	90, 94.82, 90
Resolution [Å][1]	50.0-2.6 (2.7-2.6)	50.0-2.4 (2.8-2.4)
Number of reflections, all[1]	469,491 (50,862)	3,128,956 (1,146,656)
Number of reflections, unique[1]	133,525 (14,322)	224,611 (83,134)
Completeness [%][1]	98.8 (98.5)	99.7 (99.6)
Redundancy[1]	3.5 (3.6)	13.9 (13.8)
$\langle I/\sigma(I)\rangle$[1]	13.27 (2.9)	11.8 (2.5)
R_{meas}[1,2]	8.3 (64.5)	21.5 (152.2)
$CC_{1/2}$[1,3]	99.8 (82.9)	99.8 (92.6)
Wilson B-factor [Å^2]	51.26	51.77
Phasing		
Selenium sites expected / observed	-	32 / 45
Correlation coefficient all / weak [%]	-	44.85 / 27.20
Rotational / Translational Z-score	9.75 / 22.87	-
Log likelihood gain after final refinement	3169	-
Refinement		
Resolution range in refinement [Å]	49.3 – 2.6	48.9 – 2.4
R_{work}/R_{free}[4]	0.211 / 0.262	0.218 / 0.261
Number of:		
Protein atoms	13986 (4 dimers)	13,973 (4 dimers)
Water molecules	33	438
Ions	20 SO_4^{2-}	51 SO_4^{2-}
Other solvent molecules	3 glycerol	-
Average B-factor [Å^2]:		
Protein atoms	53.6	51.7
Water molecules	45.0	49.4
Ions	65.5	74.8
Other solvent molecules	48.7	-
RMSD from target values:		
Bond lengths [Å]	0.010	0.011
Bond angles [°]	1.352	1.249
Ramachandran statistics:		
Preferred [%]	98.1	98.7
Allowed [%]	1.9	1.3
Outliers [%]	0.0	0.0

Values in parentheses are for the highest resolution shell

[2] $R_{meas} = \sum_{hkl} \sqrt{n/n-1} \sum_{j=1}^{n} |I_{hkl,j} - \langle I_{hkl}\rangle| / \sum_{hkl} \sum_{j} I_{hkl,j}$ where $\langle I_{hkl}\rangle$ is the average of symmetry related observations of a unique reflection.

[3] $CC_{1/2}$ gives the Pearson correlation coefficient of two 'half' data sets derived by averaging half of the observations for a given reflection.

[4] $R = \sum ||F_o| - |F_c|| / \sum |F_o|$ where R_{work} is calculated for the working set, while R_{free} is calculated for the test set excluded from refinement.

To get closer insights into the nucleotide binding properties of ε1/ζ1 the crystallization solution was supplemented with 15 mM $MgCl_2$ and different nucleotides to final concentrations of 10 mM. UNAG, UNAG-3P and different adenine nucleotides, i.e. ADP, ATP, ATPγS and APPNHP, were used as well as the transition state analog aluminium fluoride together with ADP and UNAG. To prevent ligand dissociation during harvest, the cryoprotectant solution was similarly supplemented with the respective nucleotides and $MgCl_2$. Crystals formed in all conditions, but data analysis revealed that either no nucleotide was present (UNAG, UNAG-3P, APPNHP) or the bound nucleotide was identified as ADP in cases where ADP was applied to the crystallization conditions or likely arose form hydrolysis of ATP or ATPγS. Similarly, *apo*-crystals that were incubated in nucleotide containing solution for five to 15 minutes did not take up the nucleotide or started to dissolve.

Diffraction data of a crystal of wtε1/ζ1 bound to ADP were collected to a resolution of 2.9 Å. The crystal symmetry belonged to space group $P2_1$ with unit cell parameters of $a = 79.94$ Å, $b = 149.82$ Å, $c = 124.21$ Å, $\alpha = 90$ °, $\beta = 95.04$ °, $\gamma = 90$ °. The data collection and refinement statistics of wtε1/ζ1 in the *apo*- and ADP-bound form are summarized in Table 3.2-2.

De novo phasing of ε1/ζ1_K115A by single wavelength anomalous diffraction

The anomalous signal measured from the SeMet-labeled ε1/ζ1_K115A allowed to determine initial phases by SAD. 45 selenium atoms were placed in the heavy atom substructure by SHELXD [100] with correlation coefficients of 44.85 % (all) and 27.20 % (weak), but the electro density obtained from SHELXE [100] did not allow model building. Instead, the heavy atom substructure was used for SAD phasing in Phaser [102] followed by density modification in Parrot [103] and DM [104] to obtain an interpretable electron density map.

Four ε1/ζ1 dimers were identified in the asymmetric unit with a Matthews coefficient of 3.58 $Å^3$/Da and an estimated solvent content of 65.6 %. The model was built manually and improved by refinement with torsion-angle non-crystallographic symmetry restraints [111] in phenix.refine [110].

A continuous polypeptide chain spanning almost the entire molecule could be modeled for all eight proteins, comprising residues 5 to 61 (of 61) for the longest ε1 chain and residues 1 to 396 (of 401) for the longest ζ1 chain. In addition, water molecules and sulfate ions were placed into the electron density, considering the signal intensity, coordination distances and charge interactions. In particular, well-defined electron density suggested the presence of a sulfate atom bound to the backbone amides of residues in the P-loop. The refinement converged at a free R-factor of 26.1 % and a R_{work} of 21.8 %.

Table 3.2-2: X-ray data collection and refinement statistics for wtε1/ζ1, wtε1/ζ1+ADP

Crystal	*Apo*-wtε1/ζ1	wtε1/ζ1+ADP
Space group	$P2_1$	$P2_1$
Cell parameters		
a, *b*, *c* [Å]	79.72, 150.72, 124.59	79.94, 149.82, 124.21
α, *β*, *γ* [°]	90, 94.95, 90	90, 95.04, 90
Resolution [Å][1]	50.0-2.9 (3.0-2.9)	50.0-2.9 (3.0-2.9)
Number of reflections, all[1]	445,976 (44,363)	444,481 (44,059)
Number of reflections, unique[1]	64,694 (6,207)	64,331 (6,202)
Completeness [%][1]	99.6 (99.5)	99.6 (99.6)
Redundancy[1]	6.9 (7.1)	6.9 (7.1)
$\langle I/\sigma(I)\rangle$[1]	14.36 (2.71)	12.74 (2.70)
R_{meas}[1,2]	11.4 (93.3)	15.0 (91.1)
$CC_{1/2}$[1,3]	99.8 (84.1)	99.7 (84.6)
Wilson B-factor [Å^2]	53.98	45.41
Refinement		
Resolution range in refinement [Å]	49.0 – 2.9	48.3 – 2.9
R_{work}/R_{free}[4]	0.201 / 0.246	0.217 / 0.260
Number of:		
Protein atoms	14,087 (4 dimers)	13,958 (4 dimers)
Water molecules	53	51
Ions	54 SO_4^{2-}	39 SO_4^{2-} / 4 Mg^{2+}
Ligands		4 ADP
Average B-factor [Å^2]:		
Protein atoms	54.2	49.9
Water molecules	45.1	40.9
Ions	78.6	78.2 / 36.0
Ligands		50.62
RMSD from target values:		
Bond lengths [Å]	0.011	0.011
Bond angles [°]	1.347	1.360
Ramachandran statistics:		
Preferred [%]	98.8	99.0
Allowed [%]	1.2	1.0
Outliers [%]	0.0	0.0

[1]Values in parentheses are for the highest resolution shell

[2]$R_{meas} = \sum_{hkl}\sqrt{n/n-1}\sum_{j=1}^{n}\left|I_{hkl,j} - \langle I_{hkl}\rangle\right| / \sum_{hkl}\sum_{j} I_{hkl,j}$ where $\langle I_{hkl}\rangle$ is the average of symmetry related observations of a unique reflection.

[3]$CC_{1/2}$ gives the Pearson correlation coefficient of two 'half' data sets derived by averaging half of the observations for a given reflection.

[4]$R = \sum\left||F_o| - |F_c|\right| / \sum|F_o|$ where R_{work} is calculated for the working set, while R_{free} is calculated for the test set excluded from refinement.

Phasing of apo- or nucleotide-bound wtε1/ζ1

The model obtained for SeMet-labeled ε1/ζ1_K115A, excluding water molecules and ions, was used for phasing of the diffraction data of the isomorphous wtε1/ζ1 crystals. Initial phases were obtained using a rigid body refinement protocol [106] and the models were subsequently improved by iterative cycles of manual model building and refinement till convergence was reached (see refinement statistics in Table 3.2-1).

The final models were very similar to the structure obtained for SeMet-labeled ε1/ζ1_K115A, but as expected from the mutation at residue 115, which restored the Walker A lysine of the wild-type sequence, additional electron density was observed at this position that allowed to unambiguously model the lysine side chain.

In the electron density of the nucleotide-bound structure, additional density in the vicinity of the P-loop of this structure allowed to unambiguously model a magnesium ion and ADP into each heterodimer. The B-factor for ADP was refined to 51 $Å^2$, which is comparable to the overall B-factor for the protein atoms (50 $Å^2$), indicating full occupancy of the nucleotide. The water molecules coordinating the magnesium ion were not sufficiently defined in the electron density to be included in the model, which resulted in a relatively low B-factor of the magnesium ion (36 $Å^2$). In the UNAG binding pocket, residual density was observed, but could not be interpreted unambiguously.

In contrast to the other models, a loop region in the C-terminal region of ζ1 (residues 381 and 386) could not be build, due to a lack of well-defined electron density, indicating a high degree of flexibility of this loop.

Phasing of native ε1/ζ1_K115A by molecular replacement

As described above, native ε1/ζ1_K115A crystallized with unit cell parameters that differed from those of the other collected data sets. Therefore, phase extension was not possible. Instead, initial phases were obtained by molecular replacement using one heterodimer (chains G and H, without water molecules or ions) of the SeMet-labeled ε1/ζ1_K115A model as a search structure. As indicated by a Matthews coefficient of 2.73 $Å^3$/Da and the estimated solvent content of 54.9 %, Phaser_for_MR [102] placed four heterodimers into the asymmetric unit (see Table 3.2-1 for phasing and refinement parameters).

Inspection of the electron density revealed a different packing of those four dimers when compared to the non-isomorphous structures. In addition, a different positioning of the C-terminal region of ζ1 in two of the four dimers required manual rigid body placement of the altered region into the observed electron density. The initial model was then further optimized and refined till convergence was reached at a final R_{free} of 26.2 % and a R_{work} of 21.1 %. All models showed excellent stereochemistry as indicated by the Ramachandran statistics and by analysis with the MolProbity server [113].

3.2.5 Structure of the ε1/ζ1 complex

An analysis of the molecular contacts formed between the ε1/ζ1 heterodimers in the non-isomorphous unit cells revealed that these contacts were relatively weak. They comprised a maximum of 2.3 % of the surface area of the heterodimer, corresponding to a solvation free energy gain of 5 kcal/mol, and involved a limited number of hydrogen bonding interactions or salt bridges as determined using the PISA server [128]. In addition, in each of the asymmetric units of the two non-isomorphous crystals slightly different relative dimer orientations were observed. This indicated that no higher-order oligomers were formed within the crystals, which was already suggested by light scattering experiments in solution.

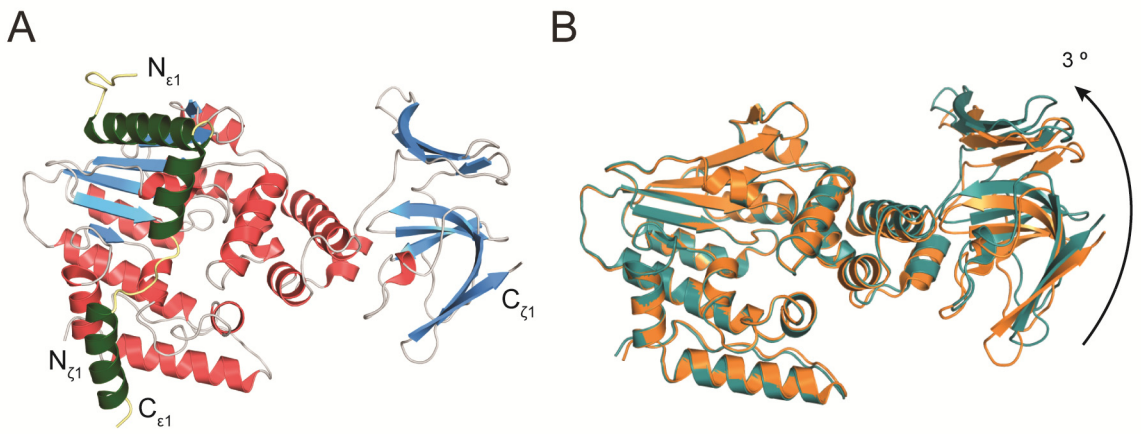


Figure 3.2-5: Crystal structure of the ε1/ζ1 heterodimer. (A) ε1/ζ1 heterodimer in ribbon representation. ε1 is colored green and yellow, ζ1 red, blue and gray. The N-terminal UNAG kinase domain is shown on the left, the C-terminal domain on the right. **(B)** Superposition of two ζ1 conformations (chain F (teal) and chain H (orange)) observed in one unit cell of native ε1/ζ1_K115A crystals shows the movement of the C-terminal domain of ζ1 relative to the UNAG kinase domain.

Structural flexibility of ε1/ζ1

Each heterodimer consisted of one epsilon antitoxin bound to one zeta toxin (Figure 3.2-5A). All four heterodimers in the structure of SeMet-labeled ε1/ζ1_K115A were very similar with RMSD values around 0.2 Å for the C_α-atoms of residues 2 to 395 of ζ1_K115A and between 0.3 Å and 0.5 Å for the C_α-atoms of residues 7 to 59 of ε1. In contrast, in the more tightly packed crystals of native ε1/ζ1_K115A differences between the individual ζ1_K115A polypeptide chains were observed. While the zeta toxins in chain B and H and chain D and F showed RMSD values for the C_α-atoms of 0.44 Å and 0.48 Å, respectively, superposition of chains between these two groups yielded RMSD values for the C_α-atoms around 1 Å. These differences were largely due to a movement of the C-terminal part of the protein (residues 272 to 401) relative to the remaining part of the dimer. From superposition of these chains it became evident that a hinge located at methionine 271 allowed for different relative orientations of the N- and C-terminal regions by an angle of at least 3 ° (Figure 3.2-5B).

The hinge separates the two regions of ζ1 that have been suggested by sequence analysis to form an N-terminal UNAG kinase and a conserved C-terminal region that is present in all gonococcal zeta homologs.

The additional C-terminal domain of ζ1 contains an OB-fold

The C-terminal part of the ζ1 polypeptide chain folds into an independent, compact domain. It consists of a five-stranded Greek-key β-barrel capped by a short α-helix, resembling the so called oligonucleotide/oligosaccharide binding fold (OB-fold) [151], and contains an additional small, four-stranded, antiparallel β-sheet and an α-helix (Figure 3.2-6A). The β-barrel has an open, distorted conformation, similar to the one found in the OB-fold of the RNA-binding S1 domain of *E. coli* polynucleotide phosphorylase [152].

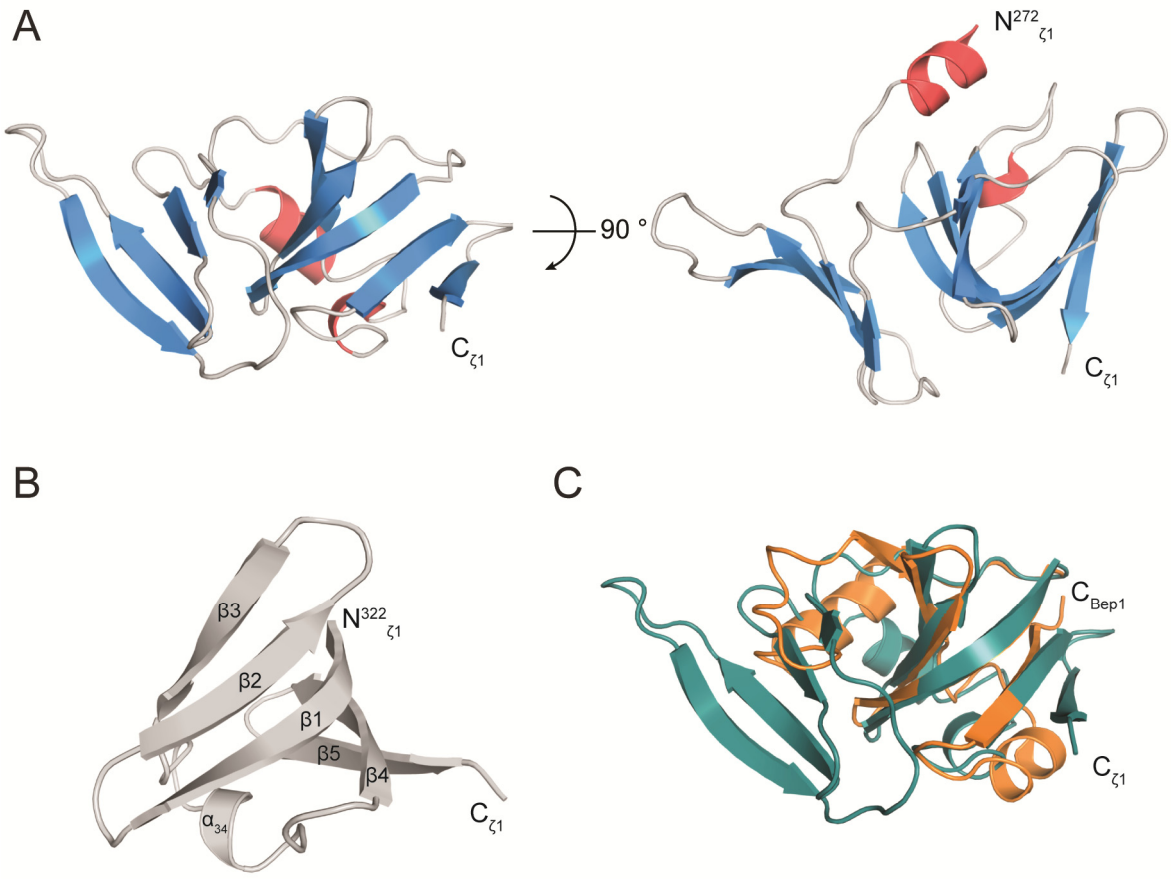


Figure 3.2-6: The C-terminal domain of ζ1 contains an OB-fold. (A) Ribbon representation of the C-terminal domain of ζ1 (residues 272-401). The domain consists of a small, four-stranded, antiparallel β-sheet (right), a helix (middle back/top) and a so called OB-fold (right). **(B)** The OB-fold of ζ1 (residues 322-401) consists of a five-stranded, antiparallel β-barrel with a short α-helical turn (α_{34}) inserted in the loop between β-strand 3 and β-strand 4. **(C)** Superposition of the C-terminal domain of ζ1 (teal, depicted as in (A)) with the C-terminal domain of Bep1 (orange, residues 230-312, PDB ID 4NPS).

In a DALI search [124] for structurally similar domains in the PDB database [125], several DNA binding proteins were identified. Highest homology was found to the human γ-interferon-inducible protein 16 (PDB ID 3RNU) and the related inflammasome inducer AIM2 (PDB ID 3RN5), the homologous recombination proteins RecO (PBD ID 1U5K) and RecR (PDB ID 4JCV) from *Deinococcus radiodurans* and a number of other DNA and RNA binding proteins. It seems thus likely, that the C-terminal domain of the gonococcal zeta homologs is involved in the binding of polynucleotides.
Notably, Bep1, a VirB-translocated effector protein from *Bartonella clarridgeiae* (PDB ID 4NPS) contains a similar OB-fold, of which 71 amino acids were superimposed on the OB-fold of ζ1 with a RMSD value for the C_α-atoms of 2.9 Å (Figure 3.2-6B).

The UNAG kinase domain has a zeta-like fold but a different topology
The N-terminal part of ζ1 folds into a separate domain that is structurally similar to the streptococcal zeta toxins (Figure 3.2-7A), which is reflected in a RMSD value for the C_α-atoms of 2.6 Å (PezT) and 3.7 Å (ζ_{spy}) obtained by SSM superposition [127] in *Coot* [107]. Like these proteins, ζ1 consists of a central β-sheet surrounded by α-helices [62, 63]. The β-sheet is composed of five, parallel β-strands complemented by a short, antiparallel β-strand. The α-helices are arranged on both sides of the β-sheet to form a shallow ATP-binding cleft on one side and a cleft for binding a second substrate, most likely UNAG, on the opposite side of the sheet.

Interestingly, aa sequence analysis of ζ1 (see section 3.2.1) revealed, that the Walker A motif is located further towards the C-terminus than in the streptococcal zeta toxins and the DxD and UNAG binding motives could not be predicted from aa sequence alignment. Nevertheless, the architecture of the catalytic center is conserved between ζ_{spy} and ζ1 (Figure 3.2-7B).
Superposition of the structures of ζ_{spy} and ζ1 showed that this is due to a different order of secondary structure elements in the polypeptide chain; β-strand 2 of ζ_{spy} is structurally equivalent to β-strand 1 of ζ1 and β-strand 1 of ζ_{spy} is homologous to β-strand 2 of ζ1. To adopt the same overall fold, the loops connecting the β-strands in the N-terminal part of the protein are differently arranged, changing a 2-3-1-4-5-6 topology (ζ_{spy}) in the β-sheet to a 1-3-2-4-5-6 arrangement (ζ1). The corresponding change in the aa sequence results in both cases in a location of the Walker A motif after the third β-strand of the β-sheet (strand 1 or 2, respectively) and the magnesium coordinating aspartate after the first β-strand of the β-sheet (strand 2 or 1). Such a reversed order of catalytic residues, placing the DxD motif N-terminal to the Walker A motif, is unusual for P-loop kinases [153].

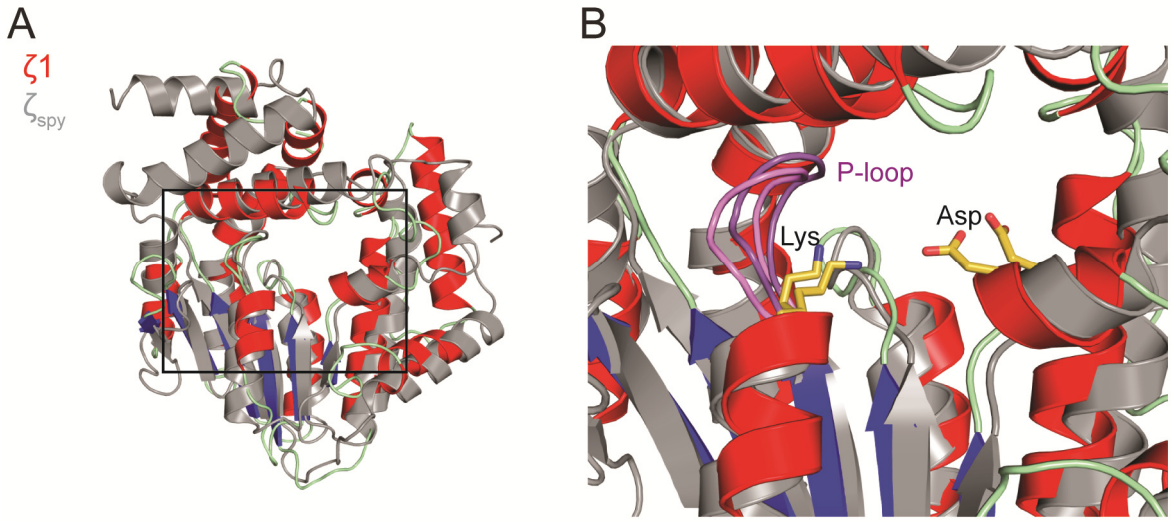


Figure 3.2-7: The UNAG kinase domain of ζ1. (A) Superposition of the UNAG kinase domain of ζ1 (blue, red, light green) with ζ_{spy} (gray). All secondary structure elements surrounding the active site (center) possess close structural equivalents. **(B)** Close-up of the active site. The P-loop is highlighted in violet (ζ1) or purple (ζ_{spy}) and the Walker A lysine and the second aspartate of the DxD motif are shown as stick models. Despite the different location in the polypeptide chain, all active site components are located at similar positions in the three-dimensional structure.

Substrate binding by ζ1

Crystallization attempts with the ε1/ζ1 complex and different nucleotides were performed to gain insights into the influence of the structural rearrangements on substrate binding and the phosphoryl transfer reaction. Of these trials, co-crystallization with ADP and $MgCl_2$ was successful and a well-defined electron density in the putative ATP binding pocket of ζ1 allowed to model an ADP nucleotide and a magnesium ion into the structure of wtε1/ζ1.

In this ADP-bound structure, the β-phosphate of ADP replaced the sulfate ion bound to the P-loop in the *apo*-structures, as observed for other kinases [154]. This, together with the relatively low B-factor and the well-defined electron density, indicates correct placement of the nucleotide into the ATP-binding pocket.

This nucleotide bound structure allowed to assign the ATP-binding residues in ζ1, which are reminiscent to the residues in other P-loop kinases [64, 153]. Except for the Walker A motif or P-loop, which has been identified in ζ1 by sequence comparison, these motifs include a degenerate Walker B motif, with glutamate 133 coordinating the magnesium ion likely via a water molecule that could not be resolved in the electron density, and a degenerate DxD motif, with aspartate 56 located after the first strand in the β-sheet. To initiate the phosphoryl transfer reaction, this aspartate probably deprotonates a hydroxyl group of the phosphate acceptor to facilitate a nucleophilic attack on the γ-phosphate of ATP [155].

Similar to other P-loop kinases [64, 153], interactions of the phosphate groups of ADP are mainly established by residues in the P-loop, including threonine 116, and serine 117, arginine 175 of a putative lid module (see below) and the magnesium ion (Figure 3.2-8A). The ε-amino group of

the Walker A lysine is positioned approximately 4.3 Å from the β-phosphate and might thus interact with the transferred γ-phosphate group of ATP [64].

Arginine 175 protrudes into the catalytic center from an α-helix inserted between β-strand 4 and β-strand 5 (Figure 3.2-8A). This residue and the entire α-helix are reminiscent of the so called lid module [153]. This lid helix is proposed to move during the phosphoryl transfer reaction to allow the arginine residue to stabilize the transition state [153, 156]. The 2'-hydroxy group of the ribose is in hydrogen bonding distance to lysine 174, which is also part of the lid helix, but further contacts of the ribose or the adenine base to ζ1 were not observed.

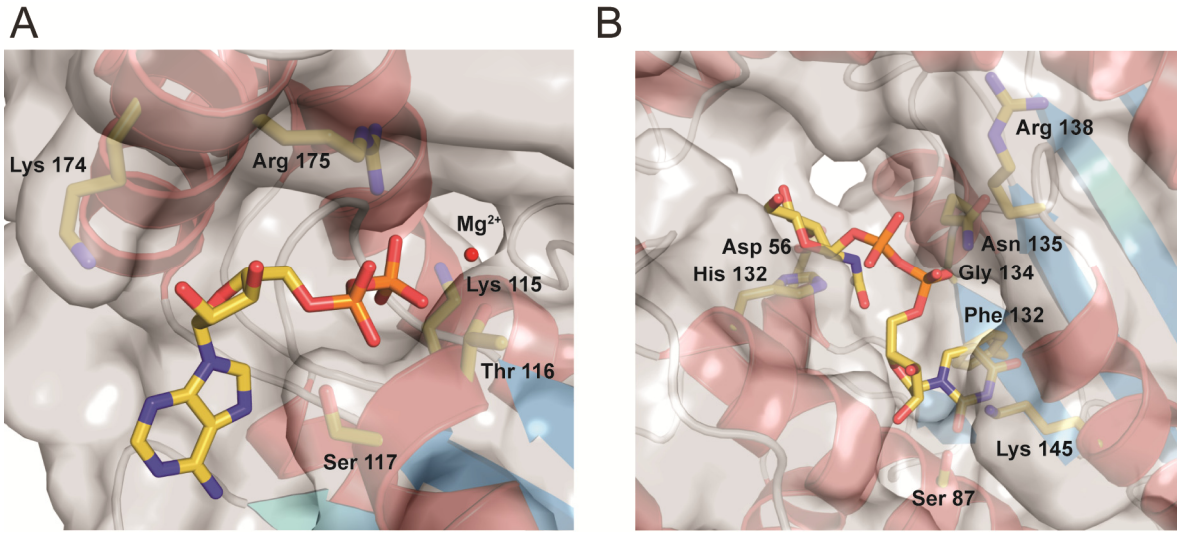


Figure 3.2-8: Substrate binding by ζ1. (A) ATP-binding pocket of ζ1 (transparent surface representation with ribbon representation underneath) with bound ADP (stick model) and a magnesium ion (sphere). Residues of ζ1 that are in hydrogen bonding distance to ADP are shown as stick model. **(B)** UNAG (sticks) has been modeled into the UNAG binding pocket of ζ1 by superposition with the ε/ζ_{spy} structure in complex with UNAG (PDB ID 3Q8X). ATP enters the active site through the tunnel in the top center. Residues that might form interactions with the substrate are shown as sticks.

To gain insights into how UNAG-binding is performed by ζ1, the ε1/ζ1 model was superimposed to the ε/ζ_{spy} structure in complex with UNAG (Figure 3.2-8B, PDB ID 3Q8X). The most pronounced difference in the two structures is the width of the UNAG binding cleft, which is significantly narrower in ζ1 when compared to ζ_{spy}, introducing a number of clashes between UNAG and ζ1 in this superposition. In addition, the distance between the β-phosphate of ADP and the 3'-hydroxyl group of the modeled UNAG is approximately 8 Å. Even when the two phosphate-oxygen bond to the γ-phosphate of approximately 1.6 Å [157] are fully extended, this leaves a distance of 5 Å, which is too far for a nucleophilic attack to take place. Both observations might be due to a more open conformation of ζ_{spy}, which is probably induced by binding to the antitoxin ε_{spy}. It can be speculated that in the absence of the antitoxin and upon binding of ATP a conformational change takes places in ζ_{spy}, such that the two substrates approach each other at the catalytic center. This 'closed' structure might resemble the tighter active site observed in ζ1.

Nevertheless, residues that line the UNAG binding pocket and are likely to interact with the substrate can be identified. These residues include aspartate 56 and histidine 79, which are in hydrogen bonding distance to the 3'- and 5'-hydroxy groups, respectively.

At the position of the glycine (Gly 118 in ζ_{spy}) of the previously identified GTXR UNAG binding motif [66], a conserved glycine (Gly 134) can be found. However, the consecutive threonine involved in phosphate binding is exchanged for a glutamine (Gln 135). The arginine within the same motif seems to be conserved in ζ1, although the side chain (Arg 138) adopts a different conformation. However, it might re-orientate towards the phosphate groups once UNAG is bound (Figure 3.2-8B). Interestingly, insertion of an additional amino acid between the glutamine and the arginine changes the GTXR motif in conventional zeta toxins to a GQXXR motif in ζ1.

The most noticeable residue that might interact with the uracil base is phenylalanine 132, which might form a pi-stacking interaction. In addition, serine 87 and lysine 145 might form hydrogen bonding interactions with the uridine.

In conclusion, although many of the residues in the UNAG binding pocket are not identical, homologous residues are exposed to form interactions with UNAG and position the nucleotide for phosphoryl transfer.

ε1 binds in an elongated conformation

ε1 was predicted to fold into three α-helices and indeed an α-helical structure is observed for ε1 in the crystal structure (Figure 3.2-9A). However, instead of forming a three-helix bundle like the streptococcal epsilon antitoxins [62, 63], ε1 consists of three α-helices that lie extended across the surface of ζ1. Due to this wrapping a relatively large contact area of 1730 $Å^2$ is established as determined by the PISA server [128], indicating stable heterodimer formation. The binding interface is mainly established by hydrogen bonds and salt bridges, which is reflected by the complementary electrostatic surface potentials (Figure 3.2-9B), similar as it has been described for the PezAT complex [72].

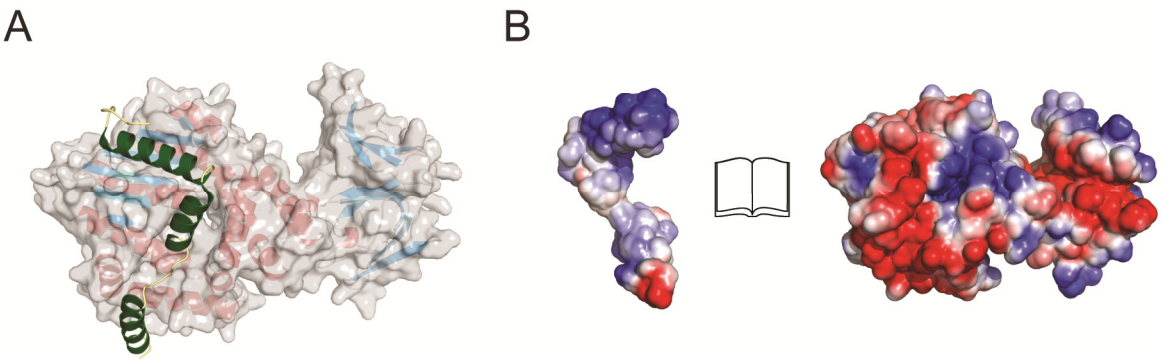


Figure 3.2-9: Binding of ε1 to ζ1. (A) ε1 forms three α-helices (dark green), which are stretched out over the surface of ζ1 (opaque, revealing the underlying secondary structure elements). The middle α-helix of ε1 lies on top of the active site (center). **(B)** Open book representation of the heterodimer interface of ε1/ζ1. The electrostatic surface potential is mapped on the surface of ε1 (left) and ζ1 (right) and is contoured from + 2 kT/e (blue) to – 2 kT/e (red).

The ε1/ζ1 complex binds ADP

In the streptococcal epsilon/zeta complexes ε/ζ_{spy} and PezAT, binding of ATP is sterically blocked by residues of the antitoxin [62, 63]. In contrast, binding of ADP to the ε1/ζ1 complex was observed by X-ray crystallography.

In the elongated conformation, the second helix of ε1 resides on top of the ATP-binding site of the kinase, similar to the position of the N-terminal helix of PezA/ε_{spy} in the complex with their cognate zeta toxins [62, 63]. However, the angle of this helix with respect to the zeta core is different and as a result the side chains of residues within this helix do not penetrate deeply into the ATP binding site. Instead, glutamate 36 and tyrosine 40 of ε1 are in hydrogen bonding distance to the 2'- and 3'-hydroxy groups of the ribose, respectively, and the methionine is in close proximity of the adenine base (Figure 3.2-10A).

When comparing the *apo*-ε1/ζ1 structure with the ADP-bound structure, most differences that are observed are changes in side chain conformations of residues that interact with the nucleotide (Figure 3.2-10B). Movements of the main chain atoms are only minor, indicating that binding of ε1 to ζ1 does not interfere with binding of the adenine nucleotide. Nevertheless, no crystals of ε1/ζ1 bound to ATP or a non-hydrolyzable ATP analog were obtained, suggesting that some differences between ADP and ATP binding might exist.

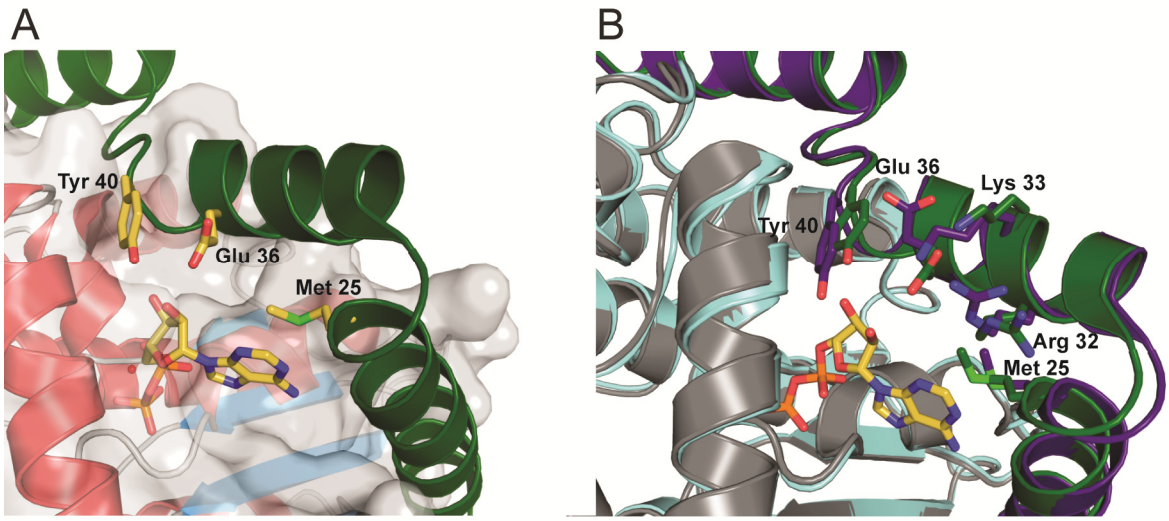


Figure 3.2-10: Interactions of ε1 with ADP. (A) ADP bound to ε1/ζ1 is shown as stick model in the ATP binding cleft of ζ1 (transparent surface representation with ribbon representation underneath), which is covered by the middle α-helix of ε1 (green). Glutamate 36 and tyrosine 40 are in hydrogen bonding distance to the hydroxy groups of the ribose, while methionine 25 is in close proximity of the adenine base. **(B)** Superposition of the ATP binding cleft of models obtained with (ε1: green, ζ1: light blue) and without (ε1: purple, ζ1: gray) bound ADP. ADP and side chains of ε1 that are altered upon nucleotide binding are shown as stick model. While conformational changes in the side chains are observed, no larger movements of the main chain are visible.

3.2.6 Wild-type ζ1 can be separated from its antitoxin

The observed difference in inhibition by ε1 compared to the streptococcal epsilon antitoxins raised further questions regarding the inhibition of ζ1 by ε1. In particular, inhibition by ε1 had to be proven experimentally and the mechanism of inhibition had to be clarified.

Therefore, experiments were performed to determine the inhibition by ε1 by the practical student Annika Grimmer during an internship. To obtain quantitative data, Michaelis-Menten parameters were determined for ζ1 alone and in the presence of different concentrations of ε1. However, cloning of wild-type ζ1 (wtζ1) alone was not successful, most likely because of leakiness from the promoter that led to the formation of minor amounts of ζ1 and the formation of a toxic enzymatic reaction product. However, cloning, overexpression and purification of the ε1/ζ1 complex was possible and it was recognized that the heterodimer can be disrupted by varying the sodium chloride concentration. This was used to separate wtζ1 from ε1 in order to perform *in vitro* studies.

ε1/ζ1 complex obtained after nickel affinity chromatography was loaded on a cation exchange column and eluted by linearly increasing the salt concentration in the buffer. Dependent on the buffer conductivity the stoichiometry of eluted ε1 and ζ1 changed from a majority of ζ1 which eluted at lowest salt concentration, to equimolar and free ε1 at higher conductivity (Figure 3.2-11A). Fractions containing mainly free wtζ1 were loaded again on a nickel affinity

column to remove remaining His_6-tagged ε1 and the ε1/ζ1 complex. Free wtζ1 was obtained in the flow-through and after washing with a buffer containing 1 M NaCl. Protein purity was verified by SDS-PAGE and Western Blotting to exclude impurities due to residual His_6-tagged ε1 (Figure 3.2-11B/C).

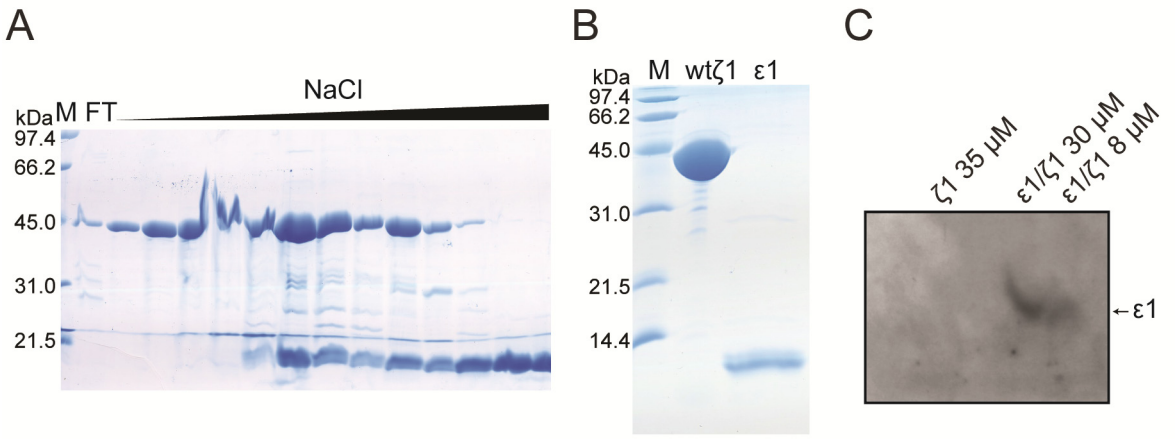


Figure 3.2-11: Isolation of free wild-type ζ1. (A) 17 % (*w/v*) SDS-PAGE gel of the separation of ε1 and ζ1 on a cation exchange column. When increasing the salt concentration, ζ1 elutes first, before the ε1/ζ1 complex and free ε1. Fractions containing mainly ζ1 were further purified yielding free wild-type ζ1 as determined by Coomassie staining **(B)** and a Western Blot against the His_6-tag of ε1 **(C)**. Data in this Figure have been contributed by Annika Grimmer.

3.2.7 Enzymatic activity of ζ1

The enzymatic activity of free wtζ1 was measured using the *in vitro* phosphorylation assay described for EzeT and a coupled ATPase assay to determine the enzymatic activity.

For the *in vitro* phosphorylation assay 1 µM wtζ1 was incubated together with ATP and UNAG for 30 min (Figure 3.2-12A). Analysis of this reaction revealed 76 % turn-over of UNAG, compared to a 50 % turn-over at a three times higher concentration of wtε1/ζ1 after 3 h, hinting at an inhibition of the enzymatic activity by ε1.

As observed for the ε1/ζ1 complex, a 1.8 fold excess of ADP over UNAG-3P was produced. What caused this non-stoichiometric relationship is unclear; especially as in the coupled ATPase assay no significant basal ATPase activity was detected. Therefore, ATP hydrolysis by ζ1 seems to be stimulated by UNAG or modified UNAG.

In the coupled ATPase assay used [95], the ATP hydrolysis of an ATPase of interest is coupled to the oxidation of NADH, which can be followed spectroscopically. Steady state reaction velocities were determined at different concentrations of UNAG and ATP (Figure 3.2-12B/C). 0.2 µM wtζ1 were titrated with UNAG (0.05 mM to 10 mM) or ATP (0 mM to 10 mM) while keeping the other substrate constant (10 mM ATP or 10 mM UNAG). These data were plotted according to Michaelis and Menten [158] to calculate the apparent maximal velocity (v_{max}) and the apparent Michaelis-Menten constant (K_m) of wtζ1.

By varying the ATP concentration, a v_{max} of 0.2±0.005 mM/min and a K_m(ATP) of 0.2±0.03 mM was obtained at an enzyme concentration normalized to 1 µM (Figure 3.2-12D). The corresponding turn-over number at maximal reaction velocity is k_{cat}=3.3 s^{-1}.

Varying the UNAG concentrations yielded a v_{max} of 0.2±0.01 mM/min and a K_m(UNAG) of 1.7±0.2 mM. In the Michaelis-Menten plot, v_{max} could not be reached, as UNAG concentrations above 10 mM caused a decrease in reaction rates. This is probably caused either by an inhibition of the enzymatic activity of ζ1 or the coupling enzymes by UNAG or other contaminants in the UNAG stock solution, or by chelation of metal ions by UNAG. Nevertheless, an excellent agreement of v_{max} determined by both titration experiments indicated that the Michaelis-Menten constants can be determined using UNAG concentrations up to 10 mM.

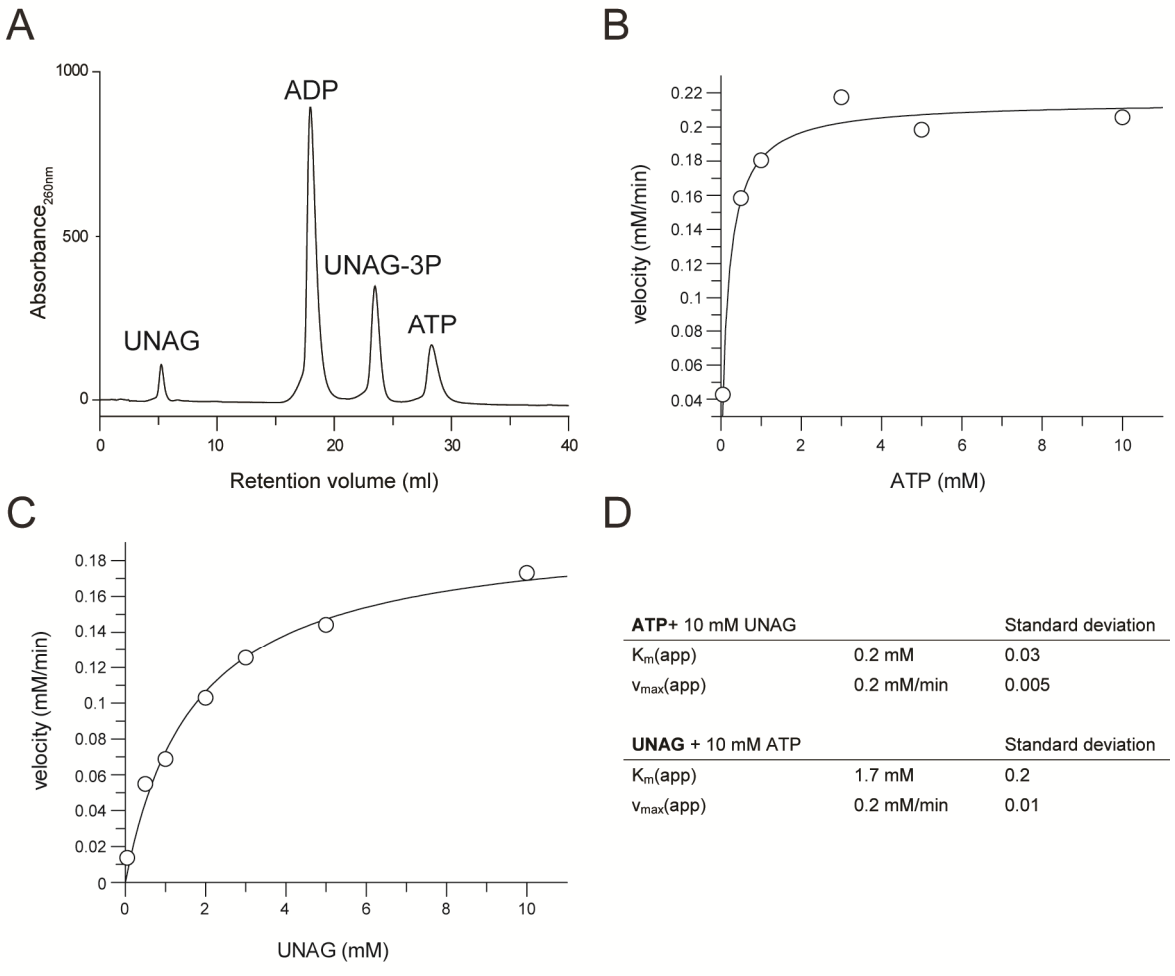
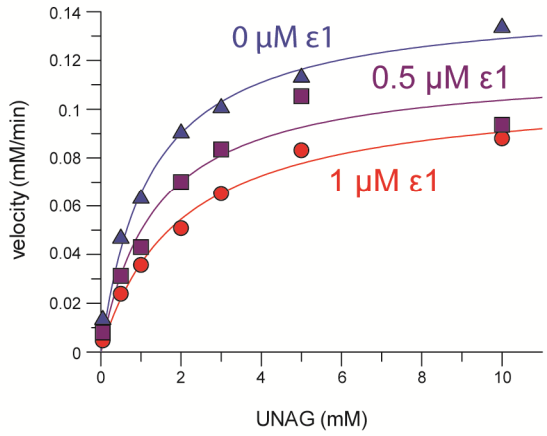


Figure 3.2-12: Enzymatic activity of wtζ1. (A) *In vitro* phosphorylation assay of wtζ1 showing the synthesis of non-stoichiometric amounts of ADP and UNAG-3P. 76 % of UNAG were modified after 30 min incubation. **(B, C)** Michaelis-Menten plots showing the initial reaction velocities determined at different ATP (B) or UNAG (C) concentrations. The experimental reaction rates were normalized to an enzyme concentration of 1 µM prior to plotting. The corresponding fits of the Michaelis-Menten equation (black line) were used to calculate the kinetic parameters given in **(D)**. Data in this Figure have been contributed by Annika Grimmer.

3.2.8 Inhibition by ε1

Using the coupled ATPase assay, the inhibition of ζ1 by the putative antitoxin ε1 was determined. Therefore, wtζ1 was pre-incubated with different amounts of ε1 and the reaction velocities at different UNAG concentrations were measured to calculate apparent v_{max} and K_m values for each concentration of ε1 (Figure 3.2-13A). Measurements were performed in the presence of an equimolar amount of ε1 (0.2 µM) or with 0.1 µM ε1 (half of the ζ1 concentration). In these experiments the apparent K_m for UNAG nearly doubled and the v_{max} decreased by approximately one third (Figure 3.2-13B) showing that ε1 acts as an inhibitor of the UNAG kinase activity. Nevertheless, UNAG phosphorylation was clearly detectable with stoichiometric amounts of ε1 and ζ1, indicating that either ζ1 retains some activity in complex with ε1, or that the affinities of the two proteins are low such that at the concentrations used in the assay the ε1/ζ1 complex partially dissociates.

A



B

0 µM ε1 + 1 µM ζ1		Standard deviation
K_m(app)	1.2 mM	0.2
v_{max}(app)	0.14 mM/min	0.006
0.5 µM ε1 + 1 µM ζ1		Standard deviation
K_m(app)	1.4 mM	0.4
v_{max}(app)	0.12 mM/min	0.01
1 µM ε1 + 1 µM ζ1		Standard deviation
K_m(app)	2 mM	0.3
v_{max}(app)	0.1 mM/min	0.005

Figure 3.2-13: Inhibition of ζ1 by ε1. (A) Michaelis-Menten plots showing the initial reaction velocities at different UNAG concentrations in the presence of 0 µM (blue), 0.5 µM (purple) and 1 µM (red) ε1. The experimental reaction rates were normalized to an enzyme concentration of 1 µM prior to plotting. The corresponding fits (colored lines) were used to calculate the kinetic parameters given in **(B)**.

3.2.9 Nucleotide binding to ε1/ζ1

As inhibition by ε1 could be shown, further experiments were performed to gain insights into how this inhibition is achieved. Therefore, binding constants for ADP and ATP to ε1/ζ1 were determined spectroscopically.

Equilibrium binding measurements were performed by measuring the intrinsic tryptophan fluorescence of a constant amount of ε1/ζ1 or free ζ1 complex, while titrating increasing amounts of nucleotide. However, no signal change was observed when using ζ1, indicating that the local

environment of the single tryptophan of ζ1, Trp305 in the C-terminal domain, is not affected by nucleotide binding.

In the ε1/ζ1 complex, ε1 provides a second tryptophan, Trp54 located in the third helix, and the fluorescence change upon nucleotide titration to ε1/ζ1 allowed to determine binding constants of K_d(ADP)=3.5±0.8 μM and K_d(ATP)=7.2±1.9 μM (Figure 3.2-14A/B).

In a second approach, the fluorescence of a constant amount of mant-labeled nucleotide was followed upon protein titration (Figure 3.2-14C). For ε1/ζ1, a K_d(mant-ADP)=6.9±0.6 μM could be determined, which corresponds well to a K_d(mant-ADP)=5.4±0.3 μM obtained by measuring fast kinetics with the stopped-flow technique (Figure 3.2-14D/E). Correspondingly, the association (k_{on}) and dissociation (k_{off}) rate constants of mant-ADP were determined as k_{on}=0.00292±0.00001 M^{-1} s^{-1} and k_{off}=0.01572±.0.00003 s^{-1}.

The ε1/ζ1-mant-ADP complex was titrated with an excess of ADP to determine the binding constant of ADP, giving a K_d(ADP)=0.15±0.03 μM (Figure 3.2-14F). This value clearly deviates from the K_d determined by intrinsic tryptophan fluorescence, but without further information it cannot be decided, which binding constant is more appropriate.

No suitable signal change was observed using a non-hydrolyzable ATP-analog (mant-APPNHP) in titrations of the ε1/ζ1 complex. In addition, titration of free ζ1 into mant-ADP yielded only very weak signals that did not allow the reliable determination of a binding constant. It can be concluded that ε1 seems to be required to cause a sufficiently strong shift in the mant-environment upon nucleotide binding. This is probably caused by the hydrogen bonding interactions of glutamate 36 and tyrosine 40 of ε1 to the 2'- and 3'-hydroxy groups of the ribose, to which the mant-group is attached.

Using fluorescence anisotropy as an alternative technique to observe protein-nucleotide complex formation with mant-ADP and mant-APPNHP similar results as for the fluorescence intensity measurements were obtained. No signal was observed for ζ1, while the increase in mant-fluorescence for ε1/ζ1-mant-ADP complex formation was associated with a change in anisotropy. Moreover, fluorescence anisotropy measurements with the more rigid etheno-ADP yielded low, not-interpretable signal amplitudes.

In conclusion, it was not possible to compare the nucleotide binding affinities of ε1/ζ1 to affinities of free ζ1 as none of the performed experiments allowed for the determination of binding rates or constants of ζ1 on its own. Binding constants of ATP and ADP to ε1/ζ1 as determined by tryptophan fluorescence titration were comparable, indicating that both nucleotides bind equally well to the ε1/ζ1 complex.

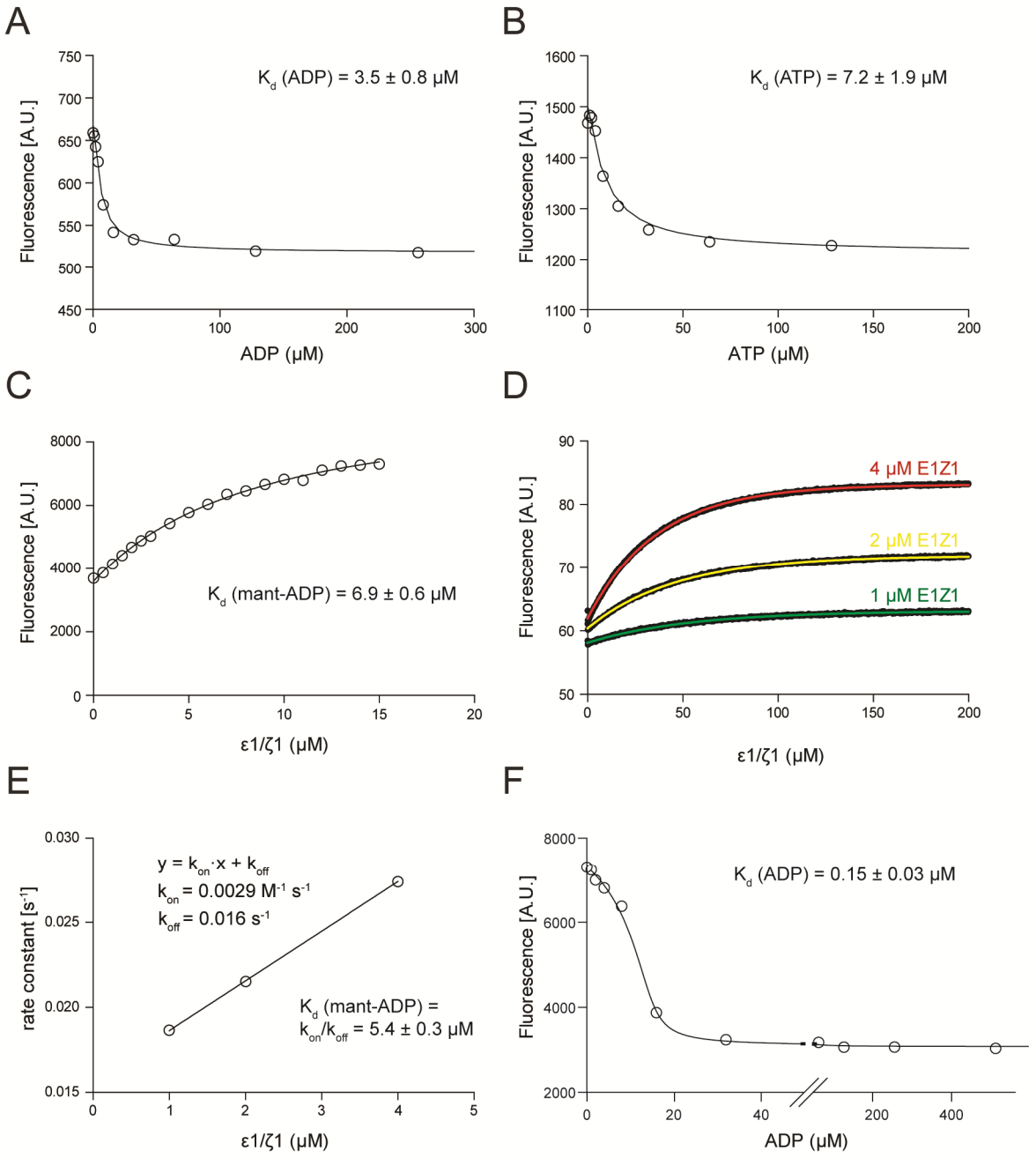


Figure 3.2-14: Nucleotide binding kinetics to ε1/ζ1. (A/B) Intrinsic tryptophan fluorescence of ε1/ζ1 was measured upon titration of ADP (A) or ATP (B). The data points were fitted with a quadratic equation (black line) to obtain the stated binding constants. **(C)** Fluorescence measurements of the mant-group attached to ADP upon titration of ε1/ζ1. The data points were fitted with a quadratic equation (black line) to obtain the stated binding constant. **(D)** Kinetic traces of direct mixing experiments using mant-ADP and different concentrations of ε1/ζ1. Fits to the exponential function are shown as colored lines, the resulting rate constants have been plotted against the protein concentration in **(E)**. Linear regression analysis gave the stated binding constants. **(F)** Chase experiment of the mant-ADP-ε1/ζ1 complex from (C) with increasing ADP concentrations. The data points were fitted with a cubic equation (black line) to obtain the stated binding constant.

3.2.10 ε2/ζ2 and ε3/ζ3 phosphorylate UNAG *in vitro*

After characterizing ε1/ζ1 and finding a functional zeta toxin and an inhibiting epsilon antitoxin, the other two putative epsilon/zeta systems were also characterized to determine their UNAG kinase activity and possible interactions with ε1/ζ1.

To test whether all three putative zeta toxins have indeed UNAG kinase activity, *in vitro* phosphorylation assays were performed. Therefore, active zeta variants had to be expressed and purified. Unlike for ζ1, complex separation was not feasible, instead, the zeta homologs were expressed and purified together with their cognate epsilon proteins.

The purified ε1/ζ1, ε2/ζ2 and ε3/ζ3 complexes were subjected to an *in vitro* chromatographic UNAG kinase assay as described for EzeT (see section 3.1.3). For all three epsilon/zeta systems the formation of a novel species from UNAG was observed (Figure 3.2-15). These species eluted at higher conductivity than UNAG, indicating a higher surface charge of the molecule, and showed a typical uridine absorption ratio of the values at 260 nm and 280 nm of 2.8:1. The modification of UNAG introduced by ζ2 and ζ3 is most likely the attachment of a negatively charged phosphoryl group, as is indicated by the hydrolysis of ATP and the conservation of aa sequence motifs typical for P-loop kinases.

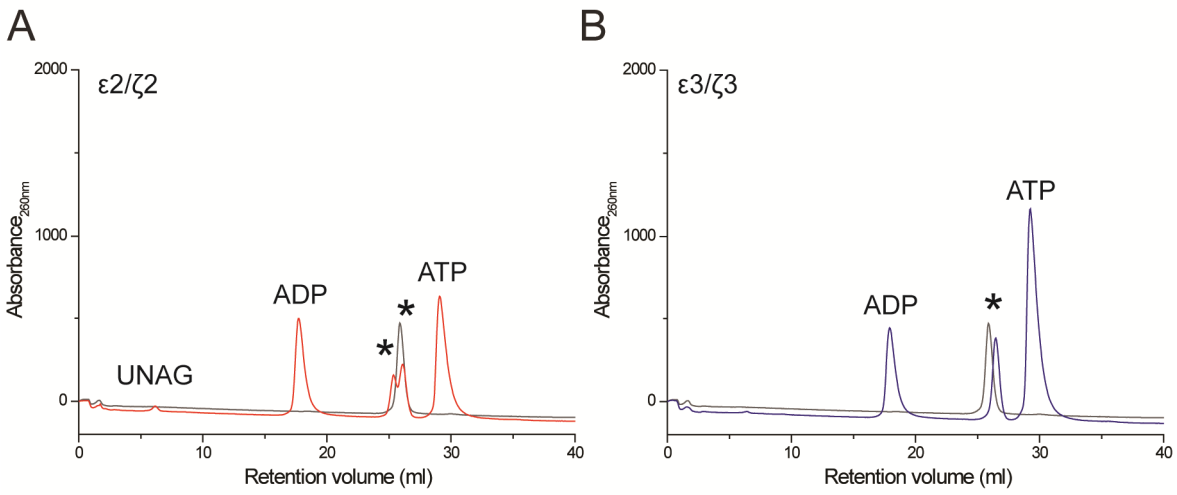


Figure 3.2-15: In vitro UNAG kinase assay. Separation of the phosphorylation assays of ε2/ζ2 (**(A)**, red) and ε3/ζ3 (**(B)**, blue) after 3 h of incubation with ATP and UNAG. The separation of purified UNAG-3P (gray) is shown for comparison. Formation of ADP and one or two additional species (*) with a retention volume similar to that of UNAG-3P is observed.

UNAG-3P, the product of UNAG phosphorylation by PezT, was used as a standard and elutes at an ionic strength of 49.7 mS. While the reaction product of ε1/ζ1 elutes at a ionic strength that corresponds well to UNAG-3P, slight differences in the retention volume and the corresponding conductivity were observed for the reaction products of ε2/ζ2 and ε3/ζ3 (Table 3.2-3), pointing at potential differences in the surface charges of the reaction products. Most surprisingly, ε2/ζ2 synthesized two reaction products with nearly identical retention volumes. Using mass spectrometry the identity of the individual compounds could not be determined unambiguously. Whether the differences indicate a different site of phosphorylation or some additional modification and their functional implications are not known.

Table 3.2-3: Products of the gonococcal zeta toxins

	ε1/ζ1	ε2/ζ2	ε3/ζ3	UNAG-3P
Peak position (mS)	50.0	48.9 / 50.0	50.7	49.7
Turn-over (%)	7.8	95.9	> 99	-

Even in complex with their cognate epsilon antitoxins, ζ2 and ζ3 achieved nearly complete turn-over of UNAG within 3 h (Figure 3.2-15 and Table 3.2-3), showing that the enzymes are either more efficient UNAG kinases than ζ1, that the epsilon/zeta complexes are less stable or that inhibition is less pronounced. Nevertheless, difficulties in cloning free toxin and the apparently positive effect of epsilon co-expression indicate inhibition by epsilon, such that all three systems can be considered type II epsilon/zeta TA systems.

3.2.11 Interactions of the gonococcal epsilon and zeta proteins

In a first approach to deduce the functional relevance of multiple epsilon/zeta systems on one plasmid, the physical interactions between the individual proteins were tested by pull-down experiments with the help of Florence Jungblut. Therefore, *E. coli* BL21(DE3)-CodonPlusRIL cells were co-transformed with two plasmids carrying a His_6-tagged epsilon and an untagged zeta, respectively. As cloning of solitary wild-type toxins was not possible, the Walker A lysines of all three putative toxins were mutated to alanines, yielding ζ1_K115A, ζ2_K46A and ζ3_K46A, which were used in this assay. After protein expression was induced, cells were harvested, lysed and the cleared lysate was applied to nickel affinity resin to bind the His_6-tagged antitoxin homologs and their potential binding partners.

Recombinant expression of all proteins was successful, as can be seen in the cleared lysate fractions (Figure 3.2-16) and all His_6-tagged epsilon homologs could be bound to nickel affinity resin. When ε1, ε2 and ε3, were co-expressed with their cognate zeta mutants, co-elution of both

proteins was observed, indicating stable complex formation, which was already observed during protein purification. However, binding to the non-cognate zeta homologs (e.g. ε1 to ζ2 and ε2 to ζ1 etc.) was not detected (Figure 3.2-16). Furthermore, expression of ε1 in combination with ζ2 or ζ3 resulted in very low protein levels and correspondingly in unspecific binding to the nickel affinity resin. Most likely this is caused by proteolytic degradation of free ε1, which gets stabilized in complex with ζ1, which was described for several other TA systems including the homologous ε/ζ_{spy} [71]. Interestingly, ε2 and ε3 were highly expressed even in combination with non-cognate zeta toxins, indicating higher proteolytic stability.

In conclusion, this shows that the three putative epsilon/zeta systems form protein-protein complexes like typical type II systems, where complex formation most likely leads to inhibition of toxin activity. Furthermore, no interaction between non-cognate epsilon/zeta pairs could be detected, suggesting that the protein-protein interfaces are sufficiently different to prevent mutual complex formation. Therefore, the two systems on each plasmid have the potential to independently stabilize DNA and lead to post-segregational killing. Alternatively, one or both of the systems might have evolved to fulfil different tasks within the cell.

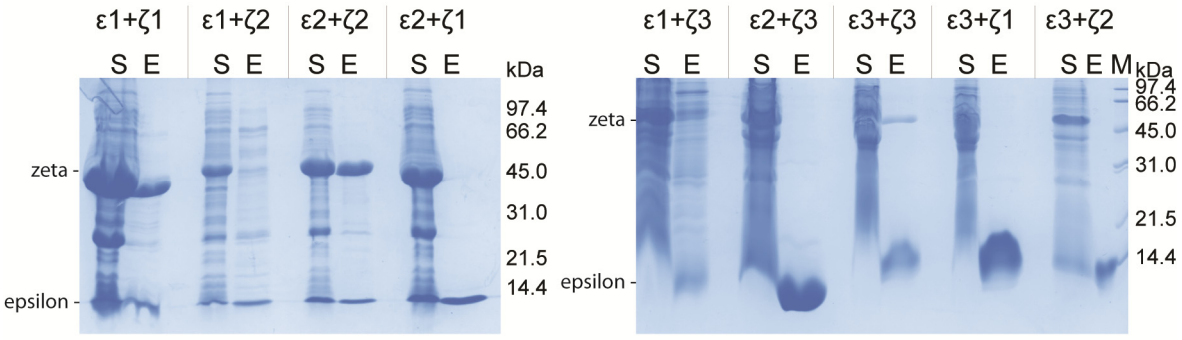


Figure 3.2-16: Interactions of the gonococcal epsilon and zeta proteins. The gonococcal zeta homologs ζ1_K115A, ζ2_K46A and ζ3_K46A were co-expressed with their putative cognate and non-cognate epsilon antitoxins ε1, ε2 and ε3 in the indicated combinations. The cleared lysate (S) was applied to nickel affinity resin and bound proteins were eluted with imidazole (E). While the His_6-tagged epsilon homologs were always bound to the resin, co-elution of each of the zeta mutants was only observed in combination with its cognate epsilon partner. Molecular weight estimates of the components of the Low Range SDS-PAGE standard (M) are given for comparison.

3.2.12 Summary

The results described in this section provide a characterization of three epsilon/zeta TA systems in the Gram-negative bacterium *N. gonorrhoeae*. For ε2/ζ2 and ε3/ζ3 UNAG kinase activity could be confirmed and ζ2 was shown to synthesize two different reaction products, most likely by phosphorylation of two different hydroxyl groups in UNAG.

The separation of the ε1/ζ1 dimer allowed to measure the enzymatic activity of free ζ1 toxin and to determine inhibition by ε1. ε1/ζ1 forms a heterodimeric protein-protein complex, the structure of which could be determined using X-ray crystallography. The ε1 antitoxin binds in an extended conformation and stretches out on the surface of ζ1, this suggests a different mode of inhibition compared to the streptococcal epsilon antitoxins. In addition, topological rewiring in the central β-sheet of ζ1 gave rise to a novel UNAG kinase fold in which the spatial organization is conserved despite re-arrangements in the primary sequence.

4 Discussion

The characterization of epsilon/zeta homologs from Gram-negative bacteria - EzeT from *E. coli* and the epsilon/zeta systems from *N. gonorrhoeae* - described in this thesis, shows that these species encode for active UNAG kinases, which are inhibited by proteinaceous antitoxins. These results unambiguously demonstrate that UNAG kinases are found in both Gram-positive and Gram-negative bacteria, in contradiction of the original notion that UNAG kinases might be only active in Gram-positive bacteria. Specific features of chromosomal EzeT and the gonococcal plasmid-encoded systems will be discussed in the following two sections.

4.1 *Escherichia coli* EzeT

As described in chapter 3.1, a zeta homolog has been identified on the chromosome of numerous *E. coli* strains. This protein, called EzeT, phosphorylates UNAG and is expressed from its endogenous promoter. It is the first thoroughly characterized UNAG kinase encoded in Gram-negative bacteria. Most remarkable is the mechanism of toxin inhibition, which is mediated by the N-terminal domain of EzeT. Such an autoinhibition by a covalently linked antitoxin has not been described for any TA system, so far.

4.1.1 A UNAG kinase regulated by a novel type of inhibition

The C-terminal domain of EzeT is homologous to the streptococcal zeta toxins, both in primary as well as in secondary and thus most likely also in tertiary structure. Moreover, all residues previously identified to be essential for the catalytic activity [61] are conserved and in the course of this study the protein was shown to phosphorylate UNAG *in vitro*. *In vivo*, overexpression of the C-terminal domain leads to the accumulation of UNAG-3P and evokes cell lysis, similar as it was shown for the attenuated streptococcal toxin PezTΔC242 [53]. EzeT is thus an active, zeta-like UNAG kinase.

In the presence of the N-terminal domain of EzeT, this kinase activity is inhibited, most likely by sterically blocking ATP binding, as it was described for the streptococcal epsilon antitoxins [62, 63]. In conclusion, the here described results show that EzeT is expressed as a single polypeptide chain comprising two domains with UNAG kinase and inhibitory functions, respectively. While different types of inhibition by the antitoxin have been reported [21, 24], an antitoxin provided *in cis* has not been described before. Therefore, EzeT is the first member of a novel type of TA system.

4.1.2 EzeT homologs in Gram-negative bacteria

When searching for homologs of EzeT in other species, related proteins were found in *Escherichia fergusonii* (93 % aa similarity) and the γ-proteobacteria *Actinobacillus minor* and *Succinatimonas* sp. (50 % aa similarity) (Figure 4.1-1). From sequence similarity it can be assumed that the *E. fergusonii* and *A. minor* proteins can phosphorylate UNAG and are inhibited by their N-terminal domain, thus sharing the novel mechanism of inhibition by an antitoxin acting *in cis*. Notably, all of the EzeT homologs contain a degenerate DxD motif, in which only the second, magnesium-coordinating aspartate is conserved, which was also observed for ζ1 (see section 3.2.1).

In contrast, the *Succinatimonas* homolog contains a degenerate Walker A motif and is thus most likely enzymatically inactive. Interestingly, the biggest differences between those proteins are found in the region identified as potential domain linker in EzeT, which shows a different arrangement in *A. minor* and *Succinatimonas* sp..

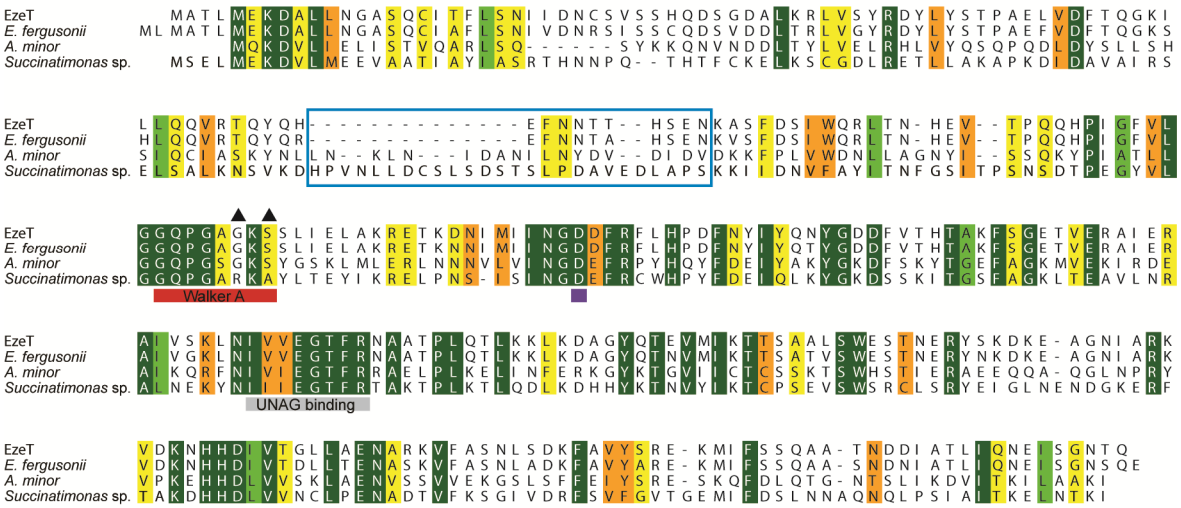


Figure 4.1-1: Full-length EzeT homologs. Amino acid sequence alignment of EzeT and full-length homologs of *Escherichia fergusonii* ATCC 35469 (YP_002385128), *Actinobacillus minor* 202 (EEV25437) and *Succinatimonas* sp. CAG:777 (CCX90861). Residues are colored according to conservation (dark green: identity to yellow: low homology). The catalytically important Walker A motif (red), the aspartate of the degenerate DxD (purple) motif and the UNAG binding region (gray) are conserved, except for two mutations in the Walker A motif of *Succinatimonas* (▲) that assumedly prevent ATP binding and hydrolysis in this protein. The blue box highlights the location of the interdomain linker.

Additionally, homologs of EzeT lacking the N-terminal polypeptide stretch were identified in other host-associated Gram-negative genera (*Neisseria*, *Haemophilus*, *Veillonella,* Figure 4.1-2A). Interestingly, many of these shorter homologs, which share between 39 % to 43 % aa sequence identity with the C-terminal domain of EzeT, but only 27 % to 34 % identity with ζ_{spy} and PezT, are encoded together with a separate ORF of a putative antitoxin in a conventional bicistronic TA operon. These antitoxins are small, about 80 aa long, and predicted to fold into three helices. Moreover, all of them, including the N-terminal domain of EzeT, are conserved at their

N-terminus (Figure 4.1-2B). This suggested that, similar to inhibition by epsilon antitoxins in the streptococcal systems, these residues of the N-terminal helix of the three-helix bundle are important for toxin inhibition.

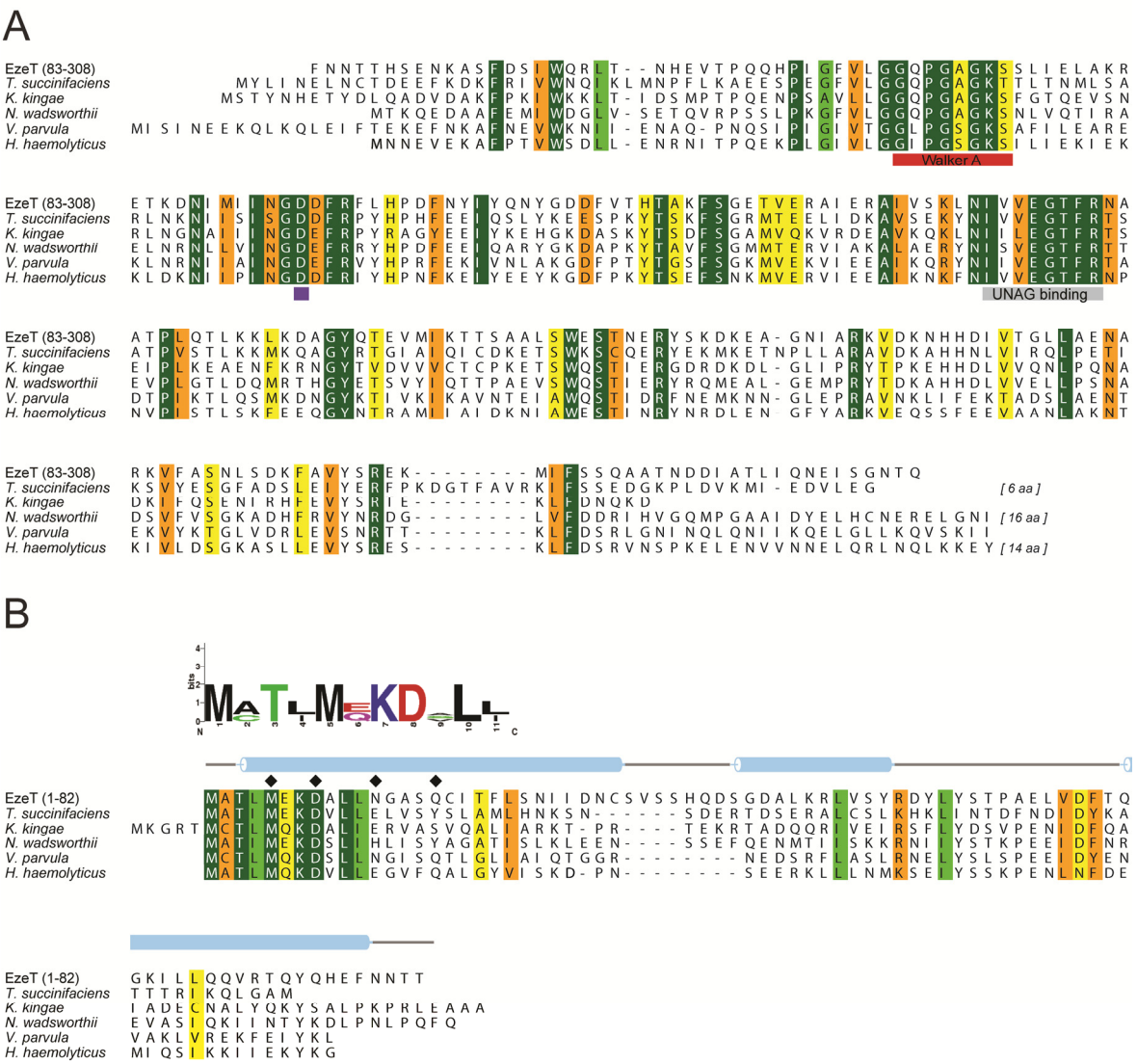


Figure 4.1-2: EzeT homologs organized in bicistronic operons. Amino acid sequence alignment of EzeT (ECO103_4771) and bicistronic EzeT homologs, color coding as in Figure 4.1-1. **(A)** Alignment of residues 78 to 308 of EzeT with putative UNAG kinase proteins of *Treponema succinifaciens* DSM 2489 (AEB13801), *Kingella kingae* KKC2005004457 (EQB59621), *Neisseria wadsworthii* (WP_009115638), *Veillonella parvula* HSIVP1 (EQC63500) and *Haemophilus haemolyticus* (EGT79785). The catalytically important sites are conserved. **(B)** Alignment of residues 1 to 82 of EzeT with the small proteins upstream of the putative UNAG kinases shown in (A). The high degree of conservation at the N-terminus of these proteins is also reflected by the consensus logo for the first 11 amino acids created with WebLogo [123] and shown above the alignment. Regions predicted to be α-helical in EzeT are indicated by blue bars. Residues marked with diamonds (♦) are suspected to block ATP binding similar to the streptococcal epsilon antitoxins and were analyzed by mutagenesis.

In analogy to the streptococcal epsilon antitoxins, four residues potentially located on one side of the N-terminal α-helix were expected to be required for the sterical blocking of ATP binding[62, 63]. Residues that may fulfil this function were the highly conserved methionine 5 and aspartate 8 and the carbamoyl-containing amino acids asparagine 12 and glutamine 16 (Figure 4.1-2B). Truncation or mutation of these amino acids in EzeT indeed diminished inhibition.

The conservation of the residues important for catalysis and inhibition in these bicistronic systems suggests that the individual proteins encode for a UNAG kinase and a separate antitoxin like conventional type II TA systems.

It can be speculated that during the evolution of EzeT the adjacent genes in one of these bicistronic operons fused and gave rise to a two-domain protein. This fusion is likely the result of a mutation, which caused the disruption of the stop codon of the upstream gene, while placing the second ORF in frame with the first. Thereby, most likely a short linker was established between two independently folded domains. Determination of the precise fusion site in EzeT is hampered by the absence of a methionine in the linker region, which could otherwise resemble the start codon of the original toxin gene. This codon might either have been deleted during fusion or mutated subsequently to optimize the linker properties.

4.1.3 A covalent linker between toxin and antitoxin

In EzeT, the toxin and the antitoxin domain are translated as a single polypeptide chain, but can interact with each other when expressed separately, as demonstrated in the rescue experiments (see section 3.1.8). This indicates the presence of a specific binding interface between the two domains, which can be established in presence or absence of a covalent connection. However, to allow for the proposed epsilon-like inhibition, the domain linker between the N- and C-terminal domain has to be sufficiently long to enable a positioning of the N-terminal domain in the vicinity of the ATP-binding site.

To estimate the required linker length the structures of PezAT and ε/ζ_{spy} were analyzed. In both complexes, the C-terminus of epsilon and the N-terminus of zeta are located on the same side of the dimer. The C-terminal α-helix of PezA and the N-terminal α-helix of PezT are 24 Å apart (Figure 4.1-3). This distance is slightly shorter (22 Å) in ε/ζ_{spy} as a result of a minor difference in the orientation of the three-helix bundle of ε_{spy} to the toxin. The minimum number of amino acids necessary to span this distance was calculated by assuming a distance of 3.4 Å spanned between the C_α-atoms of neighboring residues in an elongated β-strand conformation. In this case, at least 6 amino acids are required to connect the two proteins.

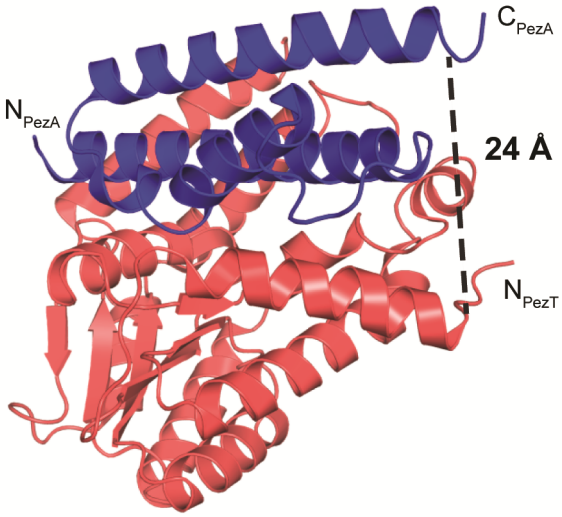


Figure 4.1-3: Distance between the C-terminus of PezA and the N-terminus of PezT. Model of an epsilon/zeta dimer (PDB ID 2P5T) showing the distance between α-helices in the C-terminus of PezA (blue) and the N-terminus of PezT (red). Both termini are located at one side of the complex with a spacing of 24 Å.

According to secondary structure predictions of EzeT, the third predicted α-helix of the N-terminal domain and the first α-helix of the zeta toxin core are linked by a stretch of 6 to 8 amino acids. Thus, it is likely that this linker length is sufficient to allow EzeT to adopt an epsilon/zeta-like conformation, but that the linker is rather a short tether than a long flexible loop.

This covalent linkage has important consequences for the activation of these TA systems, when compared to separately expressed type II systems.

4.1.4 Activation of covalently linked TA systems

For bacteria carrying TA systems, the tight regulation of toxin activity is crucial for survival. Activation has to be restricted to situations when toxin induced cell death or stasis is evolutionary beneficial. In type II systems, toxin activation is achieved by complex dissociation and degradation of the antitoxin by cellular proteases [18].

In TA modules that are involved in plasmid maintenance proteolytic degradation is a continuous process and needs to be counteracted by constant antitoxin synthesis. Activation occurs when the rate of antitoxin synthesis is smaller than the degradation rate, e.g. in the case of gene loss [159]. In contrast, chromosomal TA systems might be induced upon exposure to stress [160-162] or activated by small molecules [163].

In EzeT, the covalent tethering of the inhibitory domain to the UNAG kinase domain increases the effective local concentration of the inhibitor and, moreover, prohibits complete dissociation of the two domains. Activation of the UNAG kinase activity is most likely accomplished by cleavage of the linker, proteolytic degradation of parts of the N-terminal domain or the binding of factors that induce structural rearrangements or partial unfolding of the N-terminal domain.

Proteolytic cleavage to activate precursor enzymes has been described for so called zymogens [164]. These proteins are generally toxic when occurring in the cytosol as an enzymatically active form, including the eukaryotic digestive and coagulation enzymes and secreted bacterial proteases [165]. Therefore, these proteins are translated together with an inactivating pro-domain, which can be also involved in correct protein folding [165]. Removal of the pro-domain, mostly after secretion, leads to activation of the enzyme [164]. Similarly, such an activation of EzeT could occur in the cytosol, once it is beneficial for the bacterium. Noteworthy, a similar activation strategy for artificially fused TA systems has been described recently [166, 167]. Here, fusion proteins of the TA system MazEF have been designed by connecting the toxin MazF to the antitoxin MazE using a liker that contains a protease recognition site. Activation of these TA fusions by protease expression was shown to be effective *in vitro* and led to cell killing *in vivo* [166, 167].

Although proteolytic cleavage is a convincing strategy for EzeT activation neither cleavage nor binding of factors, which are likely to activate EzeT, has been observed in the experiments described in this work.

However, these experiments have been performed with cultures grown in rich medium at 37 ºC, providing ideal growth conditions for *E. coli*. In these situations, activation of a TA system is most likely not beneficial to the cell. Therefore, it can be suggested that activation of EzeT is performed by specific factors that are only effective under certain conditions.

Further evidence for a specific regulation is provided by the apparent inability of EzeT to stabilize its genetic locus, a finding that has also been reported for several other chromosomal TA systems [40, 146]. In contrast, in TA systems involved in plasmid maintenance the antitoxin is constantly degraded, which requires constant gene synthesis to keep the toxin in an inactive state. In these systems, activation is only influenced by the presence or absence of the genetic information, which is the basis for plasmid stabilization [30].

A potential draw-back of EzeT activation by proteolytic cleavage is the irreversibility of this process. Although an inhibition of the C-terminal domain by the N-terminal domain provided *in trans* was possible, in these experiments a significantly higher expression level of the N-terminal compared to the C-terminal domain was reached. This would require separate high-level expression of the N-terminal domain, which is not feasible when both domains are encoded as a single polypeptide chain. Therefore, rescue after toxin activation would only be possible by toxin degradation.

4.1.5 Possible functions of EzeT

The functions of chromosomal TA systems can be very diverse and include many stress adaptation strategies. To gain insights into the role of EzeT and the conditions under which it is activated, the knock-out strain E. coli W *ezeT* $^{-}$ was tested for its ability to induce persister cell or biofilm formation or phage abortive infection. No difference to the wild-type strain could be detected. This does not exclude a role of EzeT in these effects, as it is known from previous studies with TA system knock-out strains that TA systems can be highly redundant. In *E. coli* K-12 MG1655 the knock-out of 10 TA modules was required to see an effect on persister cell formation compared to the wild-type strain [168].

Analysis of the *ezeT* promoter and expression levels of EzeT in *E. coli* W showed low-level expression upon incubation at 37 ºC. Furthermore, no autorepression by EzeT could be observed, a mechanism that is utilized by most type II TA systems to ensure constant, low-level TA complex synthesis [169, 170]. Instead, transcription from *ezeT* seems to be regulated by the cellular transcription machinery and might be increased under certain circumstances. Such regulation has been described for the MazEF TA system of *Staphylococcus aureus* which is transcribed by the alternative sigma factor σ^B and regulated in response to environmental and antibiotic stresses by additional DNA binding proteins [171]. Moreover, some type I TA systems and the *E. coli* type II system yafNO are induced by the SOS-response [41, 172, 173].

As no specific regulatory elements were found in the promoter region, it is likely that promoter recognition is not performed by the housekeeping sigma factor σ^{70} but the alternative sigma factor σ^{38} that is activated under stress and at high cell densities [139]. However, which kind of stress might induce transcription remains unknown, as no significant promoter activation was detected under a variety of stress conditions tested.

The enzymatic activity of the C-terminal domain leads to the accumulation of UNAG-3P, membrane permeabilization and cell lysis when it is overexpressed alone. In addition, the formation of long-lived spherical cells was observed after 90 min of induction. In adherently growing *E. coli* cells, recombinantly overexpressing the attenuated streptococcal zeta toxin variant, PezTΔC242, similar bloated morphologies were described [53]. However, lysis upon overexpression of PezTΔC242 in liquid culture was observed earlier after induction [53], which might be a consequence of different rates of UNAG-3P formation due to different levels of active, folded enzyme, different kinase activity or differences in bacterial growth rate due to the different incubation temperatures.

The spherical cells observed after EzeTΔN83 expression (see section 3.1.7) resemble spheroplasts, which form after administration of cell-wall damaging antibiotics, for example ampicillin [133]. Cell wall-deficient bacteria or L-forms have been studied in the laboratory, but

are also considered an important cause of bacterial persistence in chronic infections [132]. EzeT might thus induce the L-form state by down-regulation of the peptidoglycan layer as well as potentially lipopolysaccharides constituting the outer capsule. This down-regulation was proposed to reduce the amount of antigens presented on the cell surface and thereby help bacteria to evade the immune system [174], which would make EzeT a virulence factor that might contribute to the infections caused by pathogenic strains of *E. coli*.

An interesting finding regarding the function of EzeT is the clear phylogenetic restriction to phylogroup B1. On the one hand, this indicates that the occurrence can be dated back to the time of separation of this phylogroup from the sister phylogroup A, which is estimated to have happened roughly 0.5 to 1.5 million years ago [175]. On the other hand, it might hint at the function of EzeT. With respect to the low thermal stability of EzeT, it is remarkable that B1 isolates are predominantly found in locations with ambient temperature e.g. soil [176], on plants [177] and in the gastrointestinal tract of ectotherms [178]. In addition, B1 strains show prolonged survival outside the host [179] and are more prone to shedding from the gastrointestinal tract than strains of phylogroup A [180]. It is thus possible, that EzeT has a thermosensor function, which regulates bacterial survival upon excretion from the host.

The importance of such genes, which ensure survival of enteric bacteria in the environment, has been already proposed 30 years ago [181]. In this report the author suggested that mutations that increase the survival time of *E. coli* outside of the host will spread, because they increase the probability of ingestion and thus distribution to novel hosts. In addition, the author calculated that 50 % of all *E. coli* cells are found in the environment, marking the importance of this habitat. Interestingly, *E. coli* W, the only laboratory strain that encodes EzeT, has been isolated from the soil of a cemetery [182].

Furthermore, another toxin with increased activity at ambient temperature, GraT of the GraTA TA system, was identified in the soil bacterium *Pseudomonas putida*[183], indicating that temperature-regulated TA systems might be beneficial for environmental bacteria.

4.1.6 A novel type of TA system?

Although close EzeT homologs have been identified, which are likely to be regulated in a similar fashion, it can only be speculated about the prevalence of autoinhibited TA systems in other families. Most of the published bioinformatic studies to identify proteinaceous TA systems have been restricted on TA systems consisting of two small neighboring genes [31, 33, 184], therefore, elongated homologs without an additional ORF for the antitoxin might have been overlooked. Indeed, homology searches that use more flexible scoring criteria regarding size, genomic context and domain conservation identified a plethora of putative single or unusually long toxins [32, 34].

The identification of covalently linked TA systems prompt further systematic bioinformatics studies to identify putative autoinhibited TA modules and the characterization of other solitary, elongated toxin homologs. These might help to identify highly-specialized and highly-regulated TA systems, which have been overlooked until now.

4.2 *Neisseria gonorrhoeae* epsilon/zeta systems

Three putative epsilon/zeta systems have been identified on conjugative plasmids from *N. gonorrhoeae*, denoted ε1/ζ1 and ε3/ζ3 in a variant without the *tetM* determinant and ε1/ζ1 and ε2/ζ2 in the presence of *tetM*. The characterization of these gonococcal epsilon/zeta systems revealed, that all three zeta homologs can phosphorylate UNAG, but form different reaction products. In addition, the three-dimensional structure of the module with the lowest sequence homology to zeta toxins, ε1/ζ1, was determined by X-ray crystallography. It was shown that the observed shift in the location of the Walker A motif by 51 amino acids is compensated by a different topology, such that ζ1 adopts a zeta-like three-dimensional architecture of the active site. Furthermore, residues important for substrate binding were identified and the structure of the C-terminal domain, common to all gonococcal zeta toxins, was elucidated. Finally, an unusual mode of inhibition by ε1 was discovered.

4.2.1 A UNAG kinase in Gram-negative *N. gonorrhoeae*

As described in chapter 3.2, ζ1 has UNAG kinase activity and is therefore, together with EzeT, the second characterized UNAG kinase in Gram-negative bacteria. Surprisingly, a 1.8 fold higher amount of ADP than UNAG-3P was produced during the phosphorylation of UNAG by wtζ1 and the independently purified ε1/ζ1 complex. However, no ATPase activity was detected in the coupled ATPase assay when UNAG was omitted, indicating that ATP hydrolysis is stimulated by UNAG. In contrast, reactions performed with other zeta toxins (PezTΔC242 (not shown), EzeT, ε2/ζ2, ε3/ζ3) showed equimolar conversion of ATP and UNAG.

Interestingly, a similar stimulation has also been reported for purified wild-type ζ_{spy} [185], which likewise does not hydrolyze ATP in the absence of UNAG, but consumes twice as much ATP as UNAG in a phosphorylation reaction. How and when the uncoupling of ATP hydrolysis and UNAG phosphorylation takes place has to be investigated further.

The separation of ζ1 from the ε1/ζ1 complex allowed to measure enzyme kinetics and the *in vitro* activity of a wild-type zeta toxin. Hitherto, Michaelis-Menten constants could be only determined for toxin mutants due to problems in cloning and expression of solitary wild-type toxins. Therefore, the obtained constants can only be compared to values determined for an attenuated version of PezT, PezTΔC242 [143].

The maximal reaction velocity v_{max} of ζ1 was found to be 0.2±0.01 mM/min at an enzyme concentration of 1 µM (k_{cat}=3.3 s^{-1}), which is 8 fold slower than the maximum rate of PezTΔC242. However, this maximal velocity is achieved at lower substrate concentrations as indicated by lower K_m values when compared to the K_m of PezTΔC242 for UNAG of 8 mM and for ATP of 12 mM [143].

Eventually, the *in vivo* reaction rate depends on the intracellular substrate concentrations, which are 100 µM to 400 µM UNAG [186-188] and 3 mM ATP [189]. Assuming an intracellular concentration of 250 µM UNAG the reaction velocity of PezTΔC242 (0.04 mM/min) is only twice as high as the reaction velocity of wtζ1 (0.02 mM/min). It can, however, be assumed that the non-truncated PezT has a higher enzymatic activity which makes it more toxic *in vivo*. The differences in reaction velocities might indicate functional adaptations in different species to either cause cell lysis or dormancy and a reduction in the UNAG-containing cell wall components like peptidoglycan, lipopolysaccharides or teichoic acids.

4.2.2 ζ1 differs from conventional zeta toxins

The N-terminal UNAG kinase domain of ζ1 is structurally similar to the streptococcal zeta toxins. However, in the primary sequence the Walker A motif of ζ1 is located further away from the N-terminus than in ζ2, ζ3, PezT or ζ_{spy}. This is reflected in a change in the topology of ζ1 compared to the streptococcal proteins.

To conserve the spatial organization of the active site and to accommodate this rearrangement in the overall fold, two β-strands in the central β-sheet are exchanged (Figure 4.2-1). This change in topology results in a different linkage of the connecting loops, while the positioning of the secondary structure elements is conserved. Despite the structural rearrangement, ζ1 is competent for ATP binding and phosphorylation, as has been shown in enzymatic activity assays.

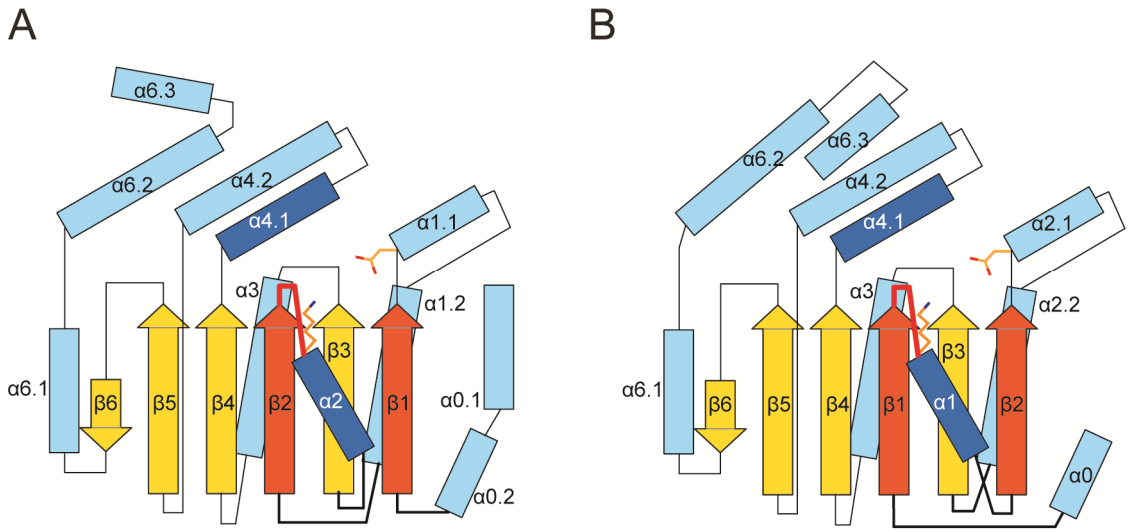


Figure 4.2-1: Topology diagrams of UNAG kinases. Topology diagrams of ζ1 **(A)** and ζ_{spy} **(B)** based on the crystal structures (ζ_{spy}: PDB ID 1GVN; ζ1: this study). Both structures contain a central six-stranded β-sheet and similarly positioned α-helices. β-strand 1 and β-strand 2 (orange) and the successive α-helices (α1/α2, α2.1/α1.1 and α2.2/α1.2) are encoded in opposite order in the polypeptide chain, changing a 2-3-1-4-5-6 topology in ζ_{spy} to a 1-3-2-4-5-6 topology in ζ1. The required topology rewiring is shown in bold, the P-loop as red line, Walker A lysine and the magnesium-coordinating aspartate of the DxD motif in stick representation. α-helices lying on top of the β-sheet are shown in dark blue, α-helices lying behind the β-sheet in light blue, identical β-strands in yellow.

The 2-3-1-4-5-6 topology observed in conventional zeta toxins corresponds to the common five-stranded, parallel β-sheet found in P-loop kinases [153, 190], with an additional sixth antiparallel β-strand typical for zeta toxins [62, 63]. In contrast, the 1-3-2-4-5-6 topology of ζ1 is rarely found in kinases [153, 190]. The conservation of the sixth β-strand suggests that the novel UNAG kinase fold evolved from a conventional zeta toxin by divergent evolution, although it cannot be fully excluded that a P-loop kinase with an uncommon topology acquired a similar arrangement of secondary structure elements and the ability to bind UNAG.

The exchange of two β-strands in a β-sheet is a complex evolutionary process, especially when it involves catalytically important residues [191]. Direct in-frame excision and insertion of the nucleotides encoding for one strand and several helices within one gene is unlikely. Instead, topological rewiring is usually assumed to be a consequence of gene duplications and subsequent reductions [192]. However, such evolutionary processes have been only described in detail for domains consisting of a limited number of secondary structure elements [193]. In homology to these systems, a partial or complete duplication of ζ1 or its N-terminal domain would be required, that results in a functional fusion protein of two toxins. In this fusion the β-strand of one protein half then takes over the function of the original strand of the other half which allows deletion of excessive parts. How the process actually took place is not certain, but the genetic load region of plasmids are known for their high plasticity and frequent genomic rearrangements.

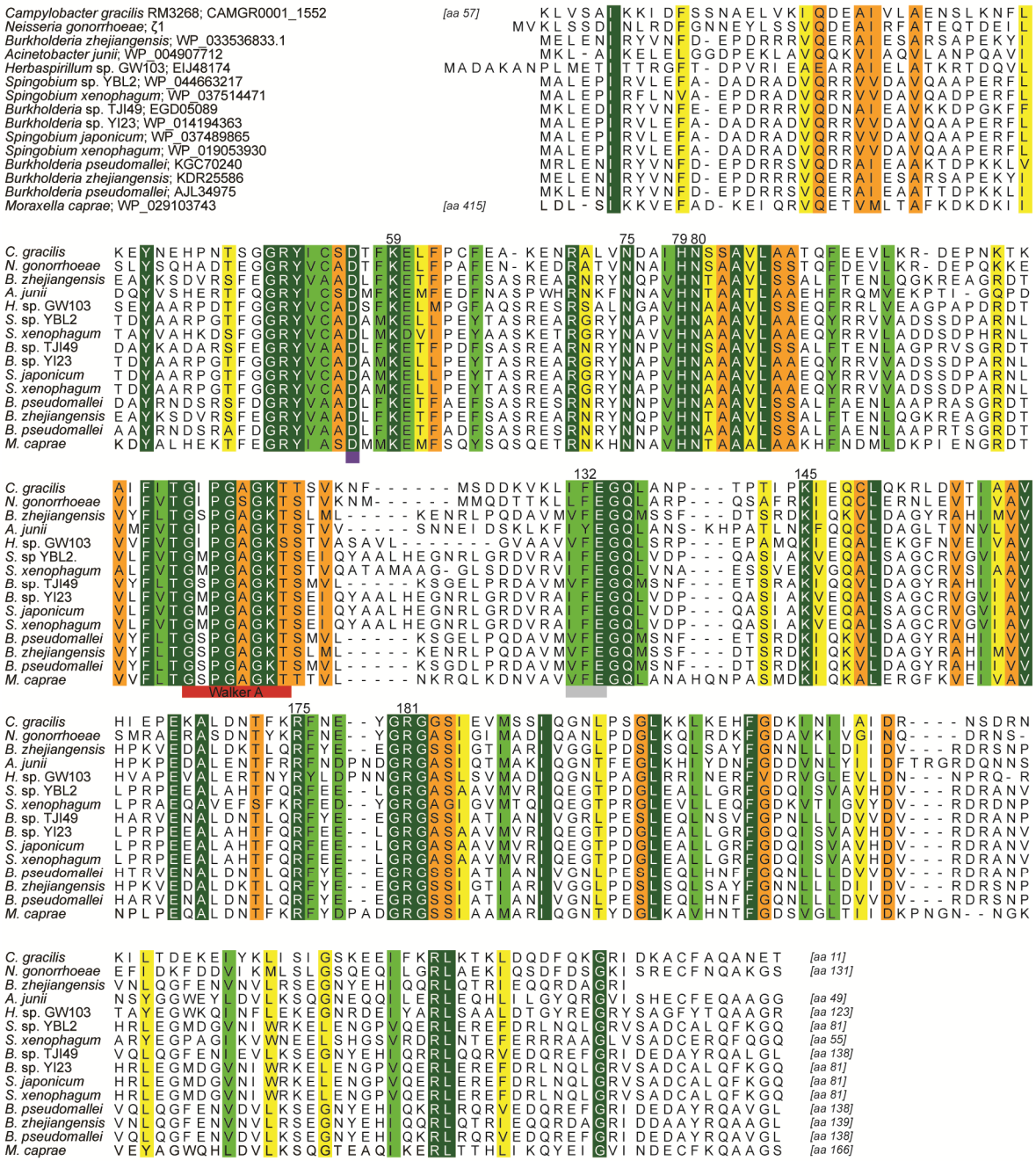


Figure 4.2-2: Alignment of proteins homologous to ζ1. Amino acid sequence alignment of ζ1 (*N. gonorrhoeae*) and homologous proteins. Protein identifier are given in the top right, the numbers of residues at the N- and C-termini that have been omitted from the alignment are indicated. Residues are colored according to conservation (dark green: identity to yellow: low homology). The catalytically important Walker A motif (red) and the degenerate DxD (purple) and Walker B (gray) motifs are conserved. Residues that were suggested to bind the substrates are numbered accordingly above the alignment.

4.2.1 Residues involved in substrate binding

Analysis of the X-ray crystal structure of ADP-bound ε1/ζ1 revealed several residues that interact with the adenine nucleotide or the magnesium ion, including the Walker A motif, degenerate Walker B and DxD (mutated to CAD) motifs and a putative lid helix. In addition, superposition with the structure of ε/ζ_{spy} in complex with UNAG [53] allowed to identify residues that might be important for binding of the nucleotide sugar (see section 3.2-5). However, it has to be kept in mind that the ε/ζ_{spy} complex seemed to be in a relatively open conformation compared to ζ1, such that UNAG was probably not ideally positioned.

When comparing the amino acid sequence of ζ1 to the NCBI protein collection [114] using BLAST [115, 116], several homologous proteins were found in proteobacterial genera such as *Burkholderia*, *Spingobium*, *Acinetobacter* and *Campilobacter* (Figure 4.2-2).

A depiction of the amino acid conservation of these proteins plotted onto the surface representation of ζ1, demonstrates a high level of conservation in the substrate binding pockets, but not on the remaining solvent-exposed surface (Figure 4.2-3). The conserved residues include the degenerate DxD motif and the adjacent β-strand 2, the Walker A motif and glutamate 133 of the degenerate Walker B motif. Additionally conserved residues that have been suggested for substrate binding are arginine 175 and arginine 181 of the lid helix, histidine 79 and lysine 145. Phenylalanine 132 is in several cases replaced by a tyrosine, which might still stack to the uracil base.

Additionally, the side chains of the conserved residues lysine 59, asparagine 75 and asparagine 80 are located deep in the UNAG binding pocket and might get in hydrogen bonding distance to the glucosamine, once it is correctly positioned for the phosphoryl transfer reaction. However, the arginine of the proposed GQXXR motif is not conserved, suggesting that this interaction is not required for UNAG binding.

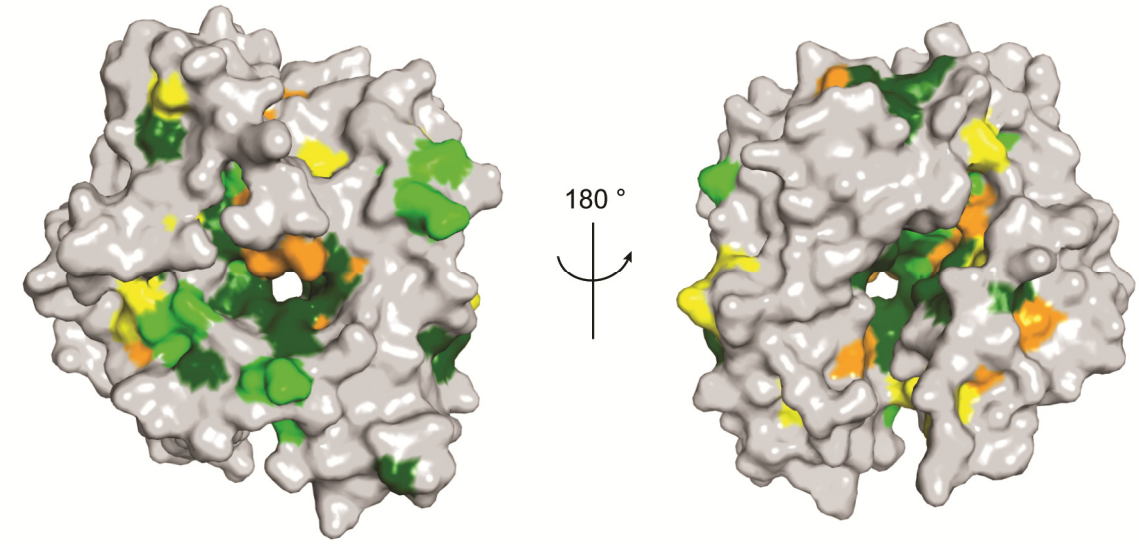


Figure 4.2-3: Catalytically important residues are conserved in ζ1 homologs. Amino acid conservation of proteins homologous to ζ1 mapped on the surface representation of ζ1 colored as in Figure 2. The molecule is rotated by 180 ° to show the ATP binding site (left) and the UNAG binding pocket (right). Highest homology is found in the substrate binding pockets as depicted by the degree of conservation.

4.2.2 Inhibition by ε1

Inhibition of ζ1 by ε1 was proven by monitoring the decrease in enzymatic activity upon addition of different concentrations of ε1. Like other type II TA systems, ε1 and ζ1 form a protein-protein complex, which is most likely required for inhibition of the enzymatic activity. This complex exists as a heterodimer in solution, as was determined by static and dynamic light scattering experiments. The apparent absence of oligomer interfaces between heterodimers was also confirmed by analysis of the molecular contacts in the two different crystal forms. In contrast, the streptococcal epsilon/zeta systems form heterotetramers, in which a central epsilon dimer is bound by two zeta toxins [61].

Furthermore, ε1 does not fold into a compact structure, but instead binds extended across the surface of ζ1. This arrangement of three α-helices looks rather similar to the RelB antitoxin of *Pyrococcus horikoshii* [194] (Figure 4.2-4A), which wraps around its cognate RelE toxin. In contrast, the streptococcal epsilon antitoxins form three-helix bundles with a defined hydrophobic core (Figure 4.2-4B). Like the RelB antitoxins and the related MazE, HigA and ParD antitoxins, ε1 might exist unstructured in solution and adopt an ordered structure upon binding to ζ1 as described by the order-disorder binding model[18, 159]. According to this model, the disordered state of the free antitoxin is thought to present regions for the recognition by proteases, which are shielded in the toxin-bound form. This is corroborated by the low proteolytic stability observed during protein expression with non-cognate zeta toxins (see section 3.2.11).

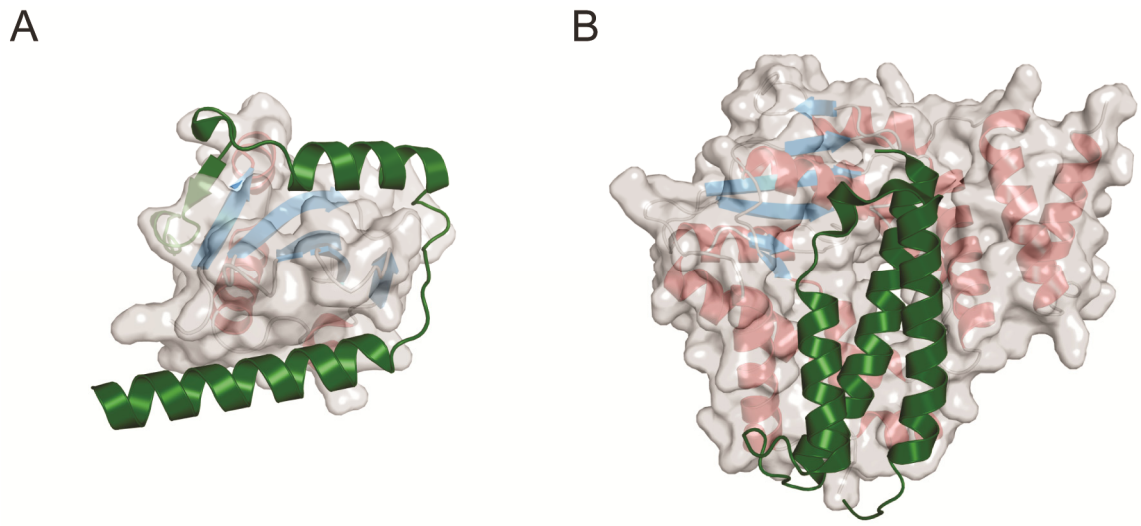


Figure 4.2-4: Antitoxin folds. (A) Elongated conformation of the RelB antitoxin (green) from *Pyrococcus horikoshii* stretched out on its cognate RelE toxin (surface representation) (PDB ID 1WMI). **(B)** Compactly folded ε_{spy} antitoxin (green) on top of the ATP binding site of ζ_{spy} (surface representation) (PDB ID 1GVN).

The origin of the elongated ε1 fold is unknown, but it is conceivable, that either the original three-helix bundle antitoxin evolved to adopt a different binding mode, or that *N. gonorrhoeae* acquired a protein unrelated to the streptococcal ε1 antitoxins that evolved to inhibit ζ1. Such an interacting protein might be a member of another antitoxin family, like the RelE or ParD antitoxins, although the sequence conservation of ε1 to these proteins is low. Nevertheless, recent reports have shown that toxins and antitoxins of different classes can interact to form functional TA pairs [31, 195-197].

Only minor structural rearrangements of the ε1/ζ1 complex and in particular of residues in ε1 were observed upon ADP binding, suggesting that binding of ε1 to ζ1 does not sterically block ATP binding, which is in strong contrast to the mode of inhibition of the streptococcal epsilon/zeta systems [61]. In addition, nucleotide binding affinities for ADP and ATP were determined for the ε1/ζ1 complex, showing K_d values in the low micromolar range. These values are well below the intracellular concentrations of ATP (3 mM) and ADP (250 µM) [189], indicating that the complex should be able to bind these nucleotides within the cell. This suggests that the inhibition performed by ε1 is different from a competitive binding to the ATP site, which was observed for the streptococcal systems [61].

Instead, binding of ε1 might inhibit progression through the transition state, for example by inhibiting movements of the kinase domain that are required for the phosphoryl-transfer. Such extensive movements have been described for NMP kinases, where the lid domain approaches the active site during catalysis [198]. This lid is located in the loop between β-strand 4 and β-strand 5 and contains a conserved arginine (residue 175 in ζ1) that is likely required to stabilize the

transition state [153]. In the ε1/ζ1 complex this part of ζ1 is located in the vicinity of the ε1 middle helix, which might prevent the contraction of the active site to bring arginine 175 and the neighboring arginine 181 closer to the transferred phosphoryl group and thereby prohibit charge stabilization of the transition state (Figure 4.2-5).

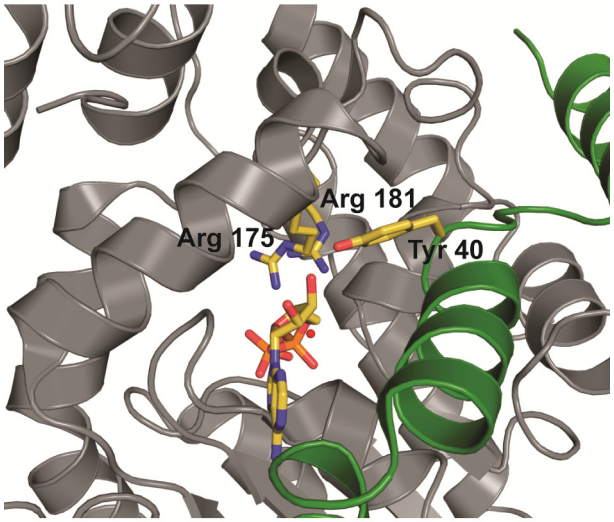


Figure 4.2-5: Putative lid domain of ζ1. Arginine 175 and arginine 181 of α-helix 4.1 of ζ1 (gray) are in close proximity to the phosphoryl groups of ADP. Tyrosine 40 of ε1 (green) forms a hydrogen bond to the 3'-hydroxy group of the ribose and sterically blocks movement of the lid helix.

Alternatively, binding of ε1 might alter the kinetics of ATP and ADP binding, which could be fix ζ1 in a product-bound state. Unfortunately, the nucleotide binding properties of free ζ1 could not be characterized with the fluorescence techniques used; therefore it is impossible to directly compare the affinities of free and complexed ζ1 in order to determine the influence of ε1. Interestingly, no crystal structures of ε1/ζ1 bound to ATP, an ATP analog or UNAG could be obtained.

From the previous considerations, it seems likely that ε1 does not compete with ATP for the same binding site, but regulates toxin activity by allosteric binding. Therefore, mixed inhibition was assumed, in which the inhibitor does not interfere with substrate binding but with the enzymatic conversion. In this case the type of inhibition and the underlying inhibition constants can be determined from the measured Michaelis-Menten kinetics for UNAG at different inhibitor concentrations. Nevertheless, direct competition experiments with varying ATP and ε1 levels would help to exclude competitive inhibition of ATP binding by ε1.

The determined Michaelis-Menten kinetics showed a decrease in the apparent v_{max} and an increase in the apparent K_m values for UNAG, indicating mixed inhibition, where the inhibitor binds to both the enzyme as well as the enzyme-substrate complex with inhibition constants K_i and

K_i', respectively [199]. To confirm this assumption, a Dixon plot and a Cornish-Bowden plot were drawn to determine the type of inhibitor and the corresponding inhibition constants.

According to the Dixon method the inverse of the reaction velocities at different substrate concentrations is plotted against the inhibitor concentration [200]. -K_i is given by the inhibitor concentration at the intersection of the linear regression lines through these data points. The Cornish-Bowden plot is drawn similarly, but instead of using the inverse of the reaction velocities directly, these are multiplied by the respective substrate concentrations [199]. In this plot, competitive inhibitors are described by a set of parallel lines, while for all other inhibitor types -K_i' can be determined from the inhibitor concentration at the intersection of the linear regression lines through the data points.

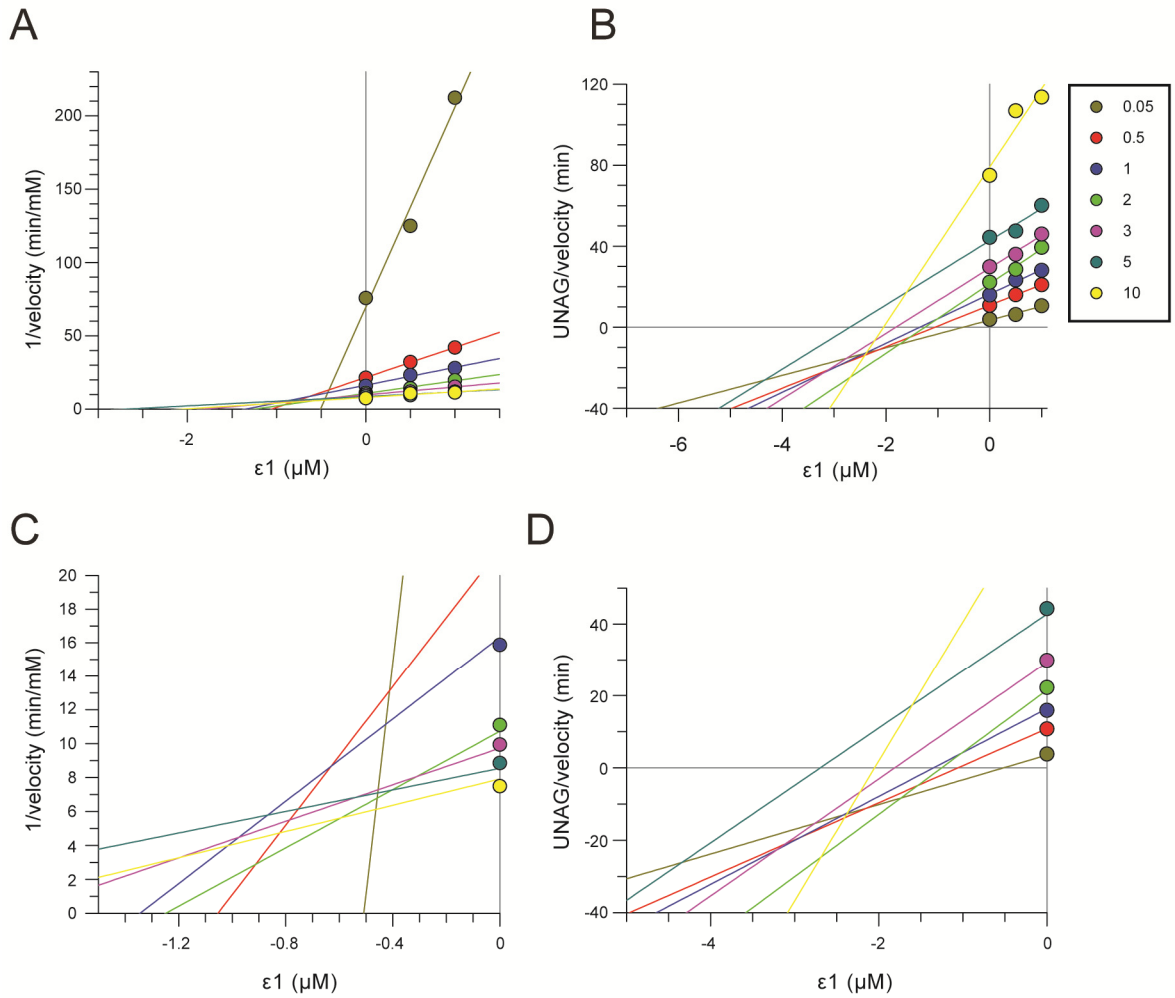


Figure 4.2-6: Inhibition of ζ1 by ε1. (A) The inverse of the reaction velocities for each set of UNAG concentrations was plotted against the ε1 concentration according to Dixon. The intersection region of the regression lines is magnified in **(C)** and allows to approximate the inhibition constant K_i. **(B)** The UNAG concentration divided by the reaction velocity is plotted against the ε1 concentration according to Cornish-Bowden. The intersection region of the regression lines is magnified in **(D)** and allows to approximate the inhibition constant K_i'. For both plots the experimental reaction rates were normalized to an enzyme concentration of 1 μM prior to plotting. The legend shown on the top right gives the different UNAG concentrations in mM.

Using these graphical methods, a mixed inhibition by ε1 could be confirmed, as in both plots the regression lines form an intersection (Figure 4.2-6). The intersections are not well defined, due to the limited number of inhibitor and substrate concentrations used, but nevertheless a rough estimate of the inhibition constants can be made. K_i, describing the inhibitor binding to the enzyme, i.e. the K_d, is given by the inhibitor concentration at the intersection in the Dixon plot (Figure 4.2-6A/C), which is approximately 0.7 µM. K_i', describing the inhibitor binding to the enzyme-substrate complex, is given in a similar fashion by the inhibitor concentration at the intersection in the Cornish-Bowden plot (Figure 4.2-6B/D). This value is slightly higher, between 2 µM to 3 µM, indicating that ε1 binding to the ζ1-substrate complex is only mildly impaired by UNAG binding.

The obtained K_d for binding of ε1 to ζ1 is in good agreement with the K_d of the plasmidic ε/ζ_{spy} system of 1 µM [70]. It can be assumed that a K_d around 1 µM is a an optimal affinity ensuring reliable toxin inhibition when toxin and antitoxin genes are present and transcribed, while allowing activation by epsilon release and degradation in case of plasmid loss. In order to determine the K_d for ε1/ζ1 more precisely additional inhibitor concentrations have to be used to calculate the K_i. A direct determination of binding constants was attempted using isothermal titration calorimetry, but was not possible due to the apparent interference of other effects caused by dilution of ε1 (data not shown).

4.2.3 The C-terminal domain might bind to DNA

All three gonococcal zeta toxins contain an additional C-terminal domain with high sequence similarity in all three proteins. This domain has been structurally characterized in ζ1 and was found to fold into an OB-fold, a small, additional β-sheet and an α-helix. OB-folds are known for interactions with other molecules, in particular oligonucleotides and oligosaccharides, but also proteins and small metabolites [201]. In addition, OB-folds can establish catalytic sites as is demonstrated in inorganic pyrophosphatases [202].
The OB-fold of ζ1 is structurally related to the family of nucleotide-binding OB-folds, like those found in transcriptional regulators, DNA-repair proteins or tRNA synthetases. The primary sequence shares highest homology with the conjugal transfer proteins KfrB (44.2 % amino acid sequence similarity), which is located on the same plasmid, and TraB from *Citrobacter freundii* (43.6 % similarity). Structures of these proteins have not been determined, but it seems conceivable that they also adopt an OB-fold that might be involved in DNA binding.

The DNA-binding activity of the OB-fold could theoretically be involved in transcriptional regulation of ε1/ζ1 expression, especially as no ω-like repressor [68] could be identified on the

plasmid and ε1 does not contain a separate DNA-binding domain, which is responsible for transcriptional regulation in PezAT [62]. However, these assumptions and the nucleotide binding region have to be tested experimentally. Theoretically, the promoter could be either located directly upstream of the *ε1/ζ1* bicistron, or - as has been proposed by Pachulec *et al*. [81] - ε2/ζ2 or ε3/ζ3, ε1/ζ1 and a transcriptional regulator *marR* could constitute a single operon. The promoter of the operon might than be either regulated by MarR or by one or both of the OB-fold containing zeta toxins.

4.2.4 ε2/ζ2 and ε3/ζ3 are functional epsilon/zeta systems

The gonococcal conjugative plasmids carry two adjacent epsilon/zeta systems, ε1/ζ1 and either ε2/ζ2 or ε3/ζ3 [81], which were shown to not physically interact with each other (see section 3.2.11). This suggests that these epsilon/zeta systems function independently and mutation or removal of one of these systems would not disturb the functioning of the other. It is thus possible that one or both of the systems have evolved to fulfil different tasks within the cell.

Bioinformatic analysis indicated that ε2/ζ2 or ε3/ζ3 contain an N-terminal domain high sequence similarity to PezT and ζ_{spy}. In addition, conservation of the predicted secondary structure elements suggests a typical zeta-like fold. Furthermore, the epsilon antitoxins likely have a higher proteolytic stability than ε1 (see section 3.2.11) and are slightly longer, suggesting that ε2 and ε3 might fold into a three-helix bundle like the streptococcal zeta toxins [62, 63]. Structural information for these systems is not available, as crystallization trials with ε2/ζ2 and ε3/ζ3 did not yield conditions that were deemed suitable for optimization.

As inferred from the homology to ζ_{spy} and PezT, ζ2 and ζ3 phosphorylate UNAG. The reaction product of the enzymatic reaction with ζ3 seems to resemble UNAG-3P, whereas two different phosphorylated UNAG species are formed by ζ2. One of these reaction products is most likely UNAG-3P, while the second product possesses minor chemical differences, which leads to weaker interaction with the resin of the anion exchange column used for separation. This difference might be caused by additional modifications or an alternative phosphorylation site, e.g. the 4'-hydroxy group of the *N*-acetyl glucosamine moiety. Further experiments using, for example, nuclear magnetic resonance spectroscopy could help to deduce the structure of the products.

In kinases, the site of phosphorylation is determined by the orientation of the substrate in the binding pocket, which positions the phosphoryl acceptor group in proximity of the γ-phosphate of ATP. Thus, synthesis of two different reaction products indicates some degree of flexibility in substrate recognition, which might be an evolutionary pathway towards a change in substrate specificity and could ultimately lead to emergence of a new enzyme class.

4.2.5 Why do plasmids contain two epsilon/zeta systems?

The gonococcal epsilon/zeta systems are located in the genetic load region of a conjugative plasmid, pLE2451. These load regions are known to acquire novel genes and rapidly exchange genetic material, leading to a flexible genetic content [203]. The associated higher mutation rates are thought to be important factors in the evolution of novel protein families [204].

It is possible that such evolutionary processes are observable in pLE2451. On the one hand, gene duplications and deletions are a likely cause for the topological rewiring observed in ζ1, on the other hand, ζ2 and ζ3 seem to have diverged from a common ancestor that was present at the time of *tetM* integration [81]. The observed flexibility in the enzymatic reaction performed by ζ2 indicates a lack of evolutionary pressure to preserve the UNAG kinase activity, which might be due to ε1/ζ1, located on the same plasmid, performing an identical enzymatic reaction.

In conclusion, the presence of two epsilon/zeta systems on the conjugative plasmids might be rather a sign of genetic load regions being evolutionary playgrounds than a functional requirement.

4.3 Conclusions and Outlook

The data presented in this thesis provide the first characterization of active UNAG kinases encoded in the genomes of Gram-negative bacteria, demonstrating the extensive spread of epsilon/zeta systems in bacteria.

Overexpression of a toxic form of EzeT led to accumulation of intracellular UNAG-3P and cell lysis, similar to overexpression of PezTΔC242 [53], indicating that under these conditions the peptidoglycan layer is severely disturbed. However, under physiological concentrations, toxin activation might rather lead to dormancy, as has been proposed for ζ_{spy} [135]. The result of toxin activation as either dormancy or cell death is an ongoing debate in the toxin-antitoxin field [19, 205, 206].

When inducing dormancy, zeta expression in Gram-negative bacteria might not only decrease the level of peptidoglycan, but could also disturb the synthesis of other carbohydrates dependent on UNAG. Such structures include lipopolysaccharides [56] and several capsular polysaccharides [207], which are exposed on the outer surface of Gram-negative bacteria. The formation of lipopolysaccharides depends on the addition of fatty acid chains to the 3'-hydroxy group of UNAG and is, therefore, likely disturbed by UNAG-3P synthesis. A reduction in the amount of lipopolysaccharides could enhance bacterial interactions, like described for hyaluronic acid down-regulation [208, 209], and therefore contribute to cellular clumping and biofilm formation.

The implication different cellular targets of UNAG-3P formation does encourage further studies on the functionality of zeta homologs in other species, including archaea, which encode zeta homologs[31]. Archaea do not form a peptidoglycan layer in their cell wall [210], but utilize UNAG for example in the formation of sialic acid [211], which can serve as nutrient, or as surface coating [212].

In conclusion, the results of this thesis demonstrate that epsilon/zeta modules are a versatile, highly abundant TA system family, which can be utilized by bacteria in multiple ways.

5 Acknowledgments

This work would not have been possible without the support of many people inside and outside the lab and I am thankful to everybody who contributed to the completion of this thesis.

I would like to thank…

… my supervisor Anton Meinhart for giving me the opportunity to work on this challenging and exciting project, for sharing his passion for toxin-antitoxin systems, for his guidance and the numerous fruitful discussions.

… Ilme Schlichting for her willingness of being my first appraiser, for her clear-cut thoughts and the continuous support.

… Jochen Reinstein for guiding me through the field of enzyme kinetics and elaborate data analysis, for all his support and for proof-reading my manuscripts.

… Robert Shoeman, Max Cryle, Thomas Barends and Matthias Fischer for their excellent advice regarding a variety of different subjects.

… Hannes Mutschler for sharing his knowledge on epsilon/zeta systems and introducing me to the topic and a variety of methods.

… Chris Roome for solving ‘mysterious’ computer problems and his cheerful mood.

… Melanie Müller and Marion Gradl for MS support; and Marion Gradl for help in cloning constructs.

… the employees of the institute’s media kitchen and glass washing facilities for enduring several large-scale microbiology experiments.

… the X-ray data collection teams of MPIs in Heidelberg and Dortmund for data collection.

… the project students Roman Sakson and Annika Grimmer for their great assistance.

… Robert Lindner for writing scripts in ‘just a minute’.

… the current and former members of the AG Meinhart for the nice working atmosphere and many more or less scientific conversations: Hannes Mutschler, Brad Lunde, Iris Magler, Aytaç Dikfidan, Julian Kellner, Juliane Benz and Stefano DaVela. And especially Maike Gebhard, Christina Sendlmeier and Florence Jungblut for keeping the lab running and the great assistance.

… the BMM community for sharing materials and ideas, jolly lunch discussions, parties and a nice atmosphere.

… Juliane Benz for scientific discussions and critically reading my manuscripts, for her supply of snacks, emotional support and lengthy discussions.

… the Heidees for many intensive practices, long weekends and the friendly community.

… Arne for insisting that there is a life outside the lab and his tremendous support in all non-scientific issues.

… my parents and my sister for believing in me and my decisions, their patience and their affection.

6 References

1. **Orcutt, B.N., J.B. Sylvan, N.J. Knab, and K.J. Edwards**, Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiol Mol Biol Rev*, 2011. **75**(2): p. 361-422.
2. **Stetter, K.O.**, Extremophiles and their adaptation to hot environments. *FEBS Lett*, 1999. **452**(1-2): p. 22-5.
3. **Tindall, B.J.**, Prokaryotic diversity in the Antarctic: the tip of the iceberg. *Microb Ecol*, 2004. **47**(3): p. 271-83.
4. **Wommack, K.E. and R.R. Colwell**, Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev*, 2000. **64**(1): p. 69-114.
5. **Strom, S.L.**, Microbial ecology of ocean biogeochemistry: a community perspective. *Science*, 2008. **320**(5879): p. 1043-5.
6. **Hibbing, M.E., C. Fuqua, M.R. Parsek, and S.B. Peterson**, Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol*, 2010. **8**(1): p. 15-25.
7. **Giovannoni, S.J. and K.L. Vergin**, Seasonality in ocean microbial communities. *Science*, 2012. **335**(6069): p. 671-6.
8. **Finlay, B.B. and G. McFadden**, Anti-immunology: evasion of the host immune system by bacterial and viral pathogens. *Cell*, 2006. **124**(4): p. 767-82.
9. **Cases, I., V. de Lorenzo, and C.A. Ouzounis**, Transcription regulation and environmental adaptation in bacteria. *Trends Microbiol*, 2003. **11**(6): p. 248-53.
10. **Boto, L.**, Horizontal gene transfer in evolution: facts and challenges. *Proc Biol Sci*, 2010. **277**(1683): p. 819-27.
11. **Hall-Stoodley, L., J.W. Costerton, and P. Stoodley**, Bacterial biofilms: from the natural environment to infectious diseases. *Nat Rev Microbiol*, 2004. **2**(2): p. 95-108.
12. **Lewis, K.**, Persister cells. *Annu Rev Microbiol*, 2010. **64**: p. 357-72.
13. **Claessen, D., D.E. Rozen, O.P. Kuipers, L. Sogaard-Andersen, and G.P. van Wezel**, Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol*, 2014. **12**(2): p. 115-24.
14. **Shapiro, J.A.**, Thinking about bacterial populations as multicellular organisms. *Annu Rev Microbiol*, 1998. **52**: p. 81-104.
15. **Fineran, P.C., T.R. Blower, I.J. Foulds, D.P. Humphreys, K.S. Lilley, and G.P. Salmond**, The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc Natl Acad Sci U S A*, 2009. **106**(3): p. 894-9.
16. **Webb, J.S., L.S. Thompson, S. James, T. Charlton, T. Tolker-Nielsen, B. Koch, M. Givskov, and S. Kjelleberg**, Cell death in *Pseudomonas aeruginosa* biofilm development. *J Bacteriol*, 2003. **185**(15): p. 4585-92.
17. **Bayles, K.W.**, Bacterial programmed cell death: making sense of a paradox. *Nat Rev Microbiol*, 2014. **12**(1): p. 63-9.
18. **Yamaguchi, Y., J.H. Park, and M. Inouye**, Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet*, 2011. **45**: p. 61-79.
19. **Gerdes, K., S.K. Christensen, and A. Lobner-Olesen**, Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol*, 2005. **3**(5): p. 371-82.
20. **Magnuson, R.D.**, Hypothetical functions of toxin-antitoxin systems. *J Bacteriol*, 2007. **189**(17): p. 6089-92.
21. **Schuster, C.F. and R. Bertram**, Toxin-antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS Microbiol Lett*, 2013. **340**(2): p. 73-85.
22. **Blower, T.R., F.L. Short, F. Rao, K. Mizuguchi, X.Y. Pei, P.C. Fineran, B.F. Luisi, and G.P. Salmond**, Identification and classification of bacterial Type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res*, 2012. **40**(13): p. 6158-73.

23. **Masuda, H., Q. Tan, N. Awano, K.P. Wu, and M. Inouye**, YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*. *Mol Microbiol*, 2012. **84**(5): p. 979-89.
24. **Aakre, C.D., T.N. Phung, D. Huang, and M.T. Laub**, A Bacterial Toxin Inhibits DNA Replication Elongation through a Direct Interaction with the beta Sliding Clamp. *Mol Cell*, 2013. **52**(5): p. 617-28.
25. **Wang, X., D.M. Lord, *et al.***, A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nat Chem Biol*, 2012.
26. **Gerdes, K. and E. Maisonneuve**, Bacterial persistence and toxin-antitoxin loci. *Annu Rev Microbiol*, 2012. **66**: p. 103-23.
27. **Ogura, T. and S. Hiraga**, Mini-F plasmid genes that couple host cell division to plasmid proliferation. *Proc Natl Acad Sci U S A*, 1983. **80**(15): p. 4784-8.
28. **Gerdes, K., P.B. Rasmussen, and S. Molin**, Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci U S A*, 1986. **83**(10): p. 3116-20.
29. **Jensen, R.B. and K. Gerdes**, Programmed cell death in bacteria: proteic plasmid stabilization systems. *Mol Microbiol*, 1995. **17**(2): p. 205-10.
30. **Hayes, F.**, Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science*, 2003. **301**(5639): p. 1496-9.
31. **Leplae, R., D. Geeraerts, R. Hallez, J. Guglielmini, P. Dreze, and L. Van Melderen**, Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Res*, 2011. **39**(13): p. 5513-25.
32. **Makarova, K.S., Y.I. Wolf, and E.V. Koonin**, Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct*, 2009. **4**: p. 19.
33. **Sberro, H., A. Leavitt, R. Kiro, E. Koh, Y. Peleg, U. Qimron, and R. Sorek**, Discovery of functional toxin/antitoxin systems in bacteria by shotgun cloning. *Mol Cell*, 2013. **50**(1): p. 136-48.
34. **Sevin, E.W. and F. Barloy-Hubler**, RASTA-Bacteria: a web-based tool for identifying toxin-antitoxin loci in prokaryotes. *Genome Biol*, 2007. **8**(8): p. R155.
35. **Yamaguchi, Y. and M. Inouye**, Regulation of growth and death in *Escherichia coli* by toxin-antitoxin systems. *Nat Rev Microbiol*, 2011. **9**(11): p. 779-90.
36. **Sala, A., P. Bordes, and P. Genevaux**, Multiple toxin-antitoxin systems in *Mycobacterium tuberculosis*. *Toxins (Basel)*, 2014. **6**(3): p. 1002-20.
37. **Hayes, F. and L. Van Melderen**, Toxins-antitoxins: diversity, evolution and function. *Crit Rev Biochem Mol Biol*, 2011. **46**(5): p. 386-408.
38. **Szekeres, S., M. Dauti, C. Wilde, D. Mazel, and D.A. Rowe-Magnus**, Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. *Mol Microbiol*, 2007. **63**(6): p. 1588-605.
39. **Unterholzner, S.J., B. Poppenberger, and W. Rozhon**, Toxin-antitoxin systems: Biology, identification, and application. *Mob Genet Elements*, 2013. **3**(5): p. e26219.
40. **Saavedra De Bast, M., N. Mine, and L. Van Melderen**, Chromosomal toxin-antitoxin systems may act as antiaddiction modules. *J Bacteriol*, 2008. **190**(13): p. 4603-9.
41. **Christensen-Dalsgaard, M., M.G. Jorgensen, and K. Gerdes**, Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol Microbiol*, 2010. **75**(2): p. 333-48.
42. **Kaspy, I., E. Rotem, N. Weiss, I. Ronin, N.Q. Balaban, and G. Glaser**, HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nat Commun*, 2013. **4**: p. 3001.
43. **Keren, I., D. Shah, A. Spoering, N. Kaldalu, and K. Lewis**, Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J Bacteriol*, 2004. **186**(24): p. 8172-80.
44. **Kasari, V., K. Kurg, T. Margus, T. Tenson, and N. Kaldalu**, The *Escherichia coli* mqsR and ygiT genes encode a new toxin-antitoxin pair. *J Bacteriol*, 2010. **192**(11): p. 2908-19.

45. **Nariya, H. and M. Inouye**, MazF, an mRNA interferase, mediates programmed cell death during multicellular *Myxococcus development. Cell*, 2008. **132**(1): p. 55-66.
46. **De la Cruz, M.A., W. Zhao, C. Farenc, G. Gimenez, D. Raoult, C. Cambillau, J.P. Gorvel, and S. Meresse**, A toxin-antitoxin module of *Salmonella* promotes virulence in mice. *PLoS Pathog*, 2013. **9**(12): p. e1003827.
47. **Harvey, R.M., U.H. Stroeher, A.D. Ogunniyi, H.C. Smith-Vaughan, A.J. Leach, and J.C. Paton**, A variable region within the genome of *Streptococcus pneumoniae* contributes to strain-strain variation in virulence. *PLoS One*, 2011. **6**(5): p. e19650.
48. **Goeders, N. and L. Van Melderen**, Toxin-antitoxin systems as multilevel interaction systems. *Toxins (Basel)*, 2014. **6**(1): p. 304-24.
49. **Unoson, C. and E.G. Wagner**, A small SOS-induced toxin is targeted against the inner membrane in *Escherichia coli. Mol Microbiol*, 2008. **70**(1): p. 258-70.
50. **Tan, Q., N. Awano, and M. Inouye**, YeeV is an *Escherichia coli* toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB. *Mol Microbiol*, 2011. **79**(1): p. 109-18.
51. **Couturier, M., M. Bahassi el, and L. Van Melderen**, Bacterial death by DNA gyrase poisoning. *Trends Microbiol*, 1998. **6**(7): p. 269-75.
52. **Castro-Roa, D., A. Garcia-Pino, S. De Gieter, N.A. van Nuland, R. Loris, and N. Zenkin**, The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nat Chem Biol*, 2013. **9**(12): p. 811-7.
53. **Mutschler, H., M. Gebhardt, R.L. Shoeman, and A. Meinhart**, A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biol*, 2011. **9**(3): p. e1001033.
54. **Barreteau, H., A. Kovac, A. Boniface, M. Sova, S. Gobec, and D. Blanot**, Cytoplasmic steps of peptidoglycan biosynthesis. *FEMS Microbiol Rev*, 2008. **32**(2): p. 168-207.
55. **Swoboda, J.G., J. Campbell, T.C. Meredith, and S. Walker**, Wall teichoic acid function, biosynthesis, and inhibition. *Chembiochem*, 2010. **11**(1): p. 35-45.
56. **Wang, X. and P.J. Quinn**, Lipopolysaccharide: Biosynthetic pathway and structure modification. *Prog Lipid Res*, 2010. **49**(2): p. 97-107.
57. **Zielenkiewicz, U., M. Kowalewska, C. Kaczor, and P. Ceglowski**, In vivo interactions between toxin-antitoxin proteins epsilon and zeta of streptococcal plasmid pSM19035 in *Saccharomyces cerevisiae. J Bacteriol*, 2009. **191**(11): p. 3677-84.
58. **Ceglowski, P., A. Boitsov, N. Karamyan, S. Chai, and J.C. Alonso**, Characterization of the effectors required for stable inheritance of *Streptococcus pyogenes* pSM19035-derived plasmids in Bacillus subtilis. *Mol Gen Genet*, 1993. **241**(5-6): p. 579-85.
59. **Brown, J.S., S.M. Gilliland, and D.W. Holden**, A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. *Mol Microbiol*, 2001. **40**(3): p. 572-85.
60. **Croucher, N.J., D. Walker, *et al.***, Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone *Streptococcus pneumoniae*Spain23F ST81. *J Bacteriol*, 2009. **191**(5): p. 1480-9.
61. **Mutschler, H. and A. Meinhart**, epsilon/zeta systems: their role in resistance, virulence, and their potential for antibiotic development. *J Mol Med (Berl)*, 2011. **12**(89): p. 1183-94.
62. **Khoo, S.K., B. Loll, W.T. Chan, R.L. Shoeman, L. Ngoo, C.C. Yeo, and A. Meinhart**, Molecular and structural characterization of the PezAT chromosomal toxin-antitoxin system of the human pathogen *Streptococcus pneumoniae. J Biol Chem*, 2007. **282**(27): p. 19606-18.
63. **Meinhart, A., J.C. Alonso, N. Strater, and W. Saenger**, Crystal structure of the plasmid maintenance system epsilon/zeta: functional mechanism of toxin zeta and inactivation by epsilon 2 zeta 2 complex formation. *Proc Natl Acad Sci U S A*, 2003. **100**(4): p. 1661-6.
64. **Saraste, M., P.R. Sibbald, and A. Wittinghofer**, The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci*, 1990. **15**(11): p. 430-4.
65. **Walker, J.E., M. Saraste, M.J. Runswick, and N.J. Gay**, Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J*, 1982. **1**(8): p. 945-51.

66. **Mutschler, H. and A. Meinhart**, Type II Toxin-Antitoxin Loci: The Epsilon/zeta Family, in *Prokaryotic Toxin-Antitoxins*, K. Gerdes, Editor. 2013, Springer Berlin Heidelberg. p. 205-223.
67. **Brown, J.S., S.M. Gilliland, B.G. Spratt, and D.W. Holden**, A locus contained within a variable region of pneumococcal pathogenicity island 1 contributes to virulence in mice. *Infect Immun*, 2004. **72**(3): p. 1587-93.
68. **de la Hoz, A.B., S. Ayora, I. Sitkiewicz, S. Fernandez, R. Pankiewicz, J.C. Alonso, and P. Ceglowski**, Plasmid copy-number control and better-than-random segregation genes of pSM19035 share a common regulator. *Proc Natl Acad Sci U S A*, 2000. **97**(2): p. 728-33.
69. **Murayama, K., P. Orth, A.B. de la Hoz, J.C. Alonso, and W. Saenger**, Crystal structure of omega transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 A resolution. *J Mol Biol*, 2001. **314**(4): p. 789-96.
70. **Camacho, A.G., R. Misselwitz, *et al.***, In vitro and in vivo stability of the epsilon2zeta2 protein complex of the broad host-range *Streptococcus pyogenes* pSM19035 addiction system. *Biol Chem*, 2002. **383**(11): p. 1701-13.
71. **Brzozowska, I. and U. Zielenkiewicz**, The ClpXP protease is responsible for the degradation of the Epsilon antidote to the Zeta toxin of the streptococcal pSM19035 plasmid. *J Biol Chem*, 2014. **289**(11): p. 7514-23.
72. **Mutschler, H., J. Reinstein, and A. Meinhart**, Assembly dynamics and stability of the pneumococcal epsilon zeta antitoxin toxin (PezAT) system from *Streptococcus pneumoniae*. *J Biol Chem*, 2010. **285**(28): p. 21797-806.
73. **Brzozowska, I., K. Brzozowska, and U. Zielenkiewicz**, Functioning of the TA cassette of streptococcal plasmid pSM19035 in various Gram-positive bacteria. *Plasmid*, 2012. **68**(1): p. 51-60.
74. **Zielenkiewicz, U. and P. Ceglowski**, The toxin-antitoxin system of the streptococcal plasmid pSM19035. *J Bacteriol*, 2005. **187**(17): p. 6094-105.
75. **Stewart-Tull, D.E.**, The immunological activities of bacterial peptidoglycans. *Annu Rev Microbiol*, 1980. **34**: p. 311-40.
76. **Typas, A., M. Banzhaf, C.A. Gross, and W. Vollmer**, From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol*, 2012. **10**(2): p. 123-36.
77. **Vollmer, W., D. Blanot, and M.A. de Pedro**, Peptidoglycan structure and architecture. *FEMS Microbiol Rev*, 2008. **32**(2): p. 149-67.
78. **Du, W., J.R. Brown, D.R. Sylvester, J. Huang, A.F. Chalker, C.Y. So, D.J. Holmes, D.J. Payne, and N.G. Wallis**, Two active forms of UDP-N-acetylglucosamine enolpyruvyl transferase in gram-positive bacteria. *J Bacteriol*, 2000. **182**(15): p. 4146-52.
79. **Lioy, V.S., M.T. Martin, *et al.***, pSM19035-encoded zeta toxin induces stasis followed by death in a subpopulation of cells. *Microbiology*, 2006. **152**(Pt 8): p. 2365-79.
80. **Morse, S.A., S.R. Johnson, J.W. Biddle, and M.C. Roberts**, High-level tetracycline resistance in *Neisseria gonorrhoeae* is result of acquisition of streptococcal tetM determinant. *Antimicrob Agents Chemother*, 1986. **30**(5): p. 664-70.
81. **Pachulec, E. and C. van der Does**, Conjugative plasmids of *Neisseria gonorrhoeae*. *PLoS One*, 2010. **5**(4): p. e9962.
82. **Bertani, G.**, Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol*, 1951. **62**(3): p. 293-300.
83. **Van Duyne, G.D., R.F. Standaert, P.A. Karplus, S.L. Schreiber, and J. Clardy**, Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol*, 1993. **229**(1): p. 105-24.
84. **Pedelacq, J.D., S. Cabantous, T. Tran, T.C. Terwilliger, and G.S. Waldo**, Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*, 2006. **24**(1): p. 79-88.
85. **Luft, J.H.**, Improvements in epoxy resin embedding methods. *J Biophys Biochem Cytol*, 1961. **9**: p. 409-14.

86. **Jensen, R.B., E. Grohmann, H. Schwab, R. Diaz-Orejas, and K. Gerdes**, Comparison of ccd of F, parDE of RP4, and parD of R1 using a novel conditional replication control system of plasmid R1. *Mol Microbiol*, 1995. **17**(2): p. 211-20.
87. **Laemmli, U.K.**, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 1970. **227**(5259): p. 680-5.
88. **Switzer, R.C., 3rd, C.R. Merril, and S. Shifrin**, A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal Biochem*, 1979. **98**(1): p. 231-7.
89. **Gharahdaghi, F., C.R. Weinberg, D.A. Meagher, B.S. Imai, and S.M. Mische**, Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis*, 1999. **20**(3): p. 601-5.
90. **Pace, C.N., F. Vajdos, L. Fee, G. Grimsley, and T. Gray**, How to measure and predict the molar absorption coefficient of a protein. *Protein Sci*, 1995. **4**(11): p. 2411-23.
91. **Gill, S.C. and P.H. von Hippel**, Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*, 1989. **182**(2): p. 319-26.
92. **Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, and A. Bairoch**, Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor. 2005, Humana Press. p. 571-607
93. **Pappin, D.J., P. Hojrup, and A.J. Bleasby**, Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol*, 1993. **3**(6): p. 327-32.
94. **Consalvi, V., R. Chiaraluce, L. Giangiacomo, R. Scandurra, P. Christova, A. Karshikoff, S. Knapp, and R. Ladenstein**, Thermal unfolding and conformational stability of the recombinant domain II of glutamate dehydrogenase from the hyperthermophile *Thermotoga maritima*. *Protein Eng*, 2000. **13**(7): p. 501-7.
95. **Berghauser, J.**, A reactive arginine in adenylate kinase. *Biochim Biophys Acta*, 1975. **397**(2): p. 370-6.
96. **Reinstein, J., I.R. Vetter, I. Schlichting, P. Rosch, A. Wittinghofer, and R.S. Goody**, Fluorescence and NMR investigations on the ligand binding properties of adenylate kinases. *Biochemistry*, 1990. **29**(32): p. 7440-50.
97. **Thrall, S.H., J. Reinstein, B.M. Wohrl, and R.S. Goody**, Evaluation of human immunodeficiency virus type 1 reverse transcriptase primer tRNA binding by fluorescence spectroscopy: specificity and comparison to primer/template binding. *Biochemistry*, 1996. **35**(14): p. 4609-18.
98. **Kabsch, W., :**, Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr.*, 1993. **26**: p. 795-800.
99. **Hendrickson, W.A., J.L. Smith, and S. Sheriff**, Direct phase determination based on anomalous scattering. *Methods Enzymol*, 1985. **115**: p. 41-55.
100. **Sheldrick, G.M.**, A short history of SHELX. *Acta Crystallogr A*, 2008. **64**: p. 112-122.
101. **Pape, T. and T.R. Schneider**, HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Cryst.*, 2004(37): p. 843-844.
102. **McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read**, Phaser crystallographic software. *J Appl Crystallogr*, 2007. **40**(Pt 4): p. 658-674.
103. **Cowtan, K.**, Recent developments in classical density modification. *Acta Cryst D*, 2010(66): p. 470-478.
104. **Cowtan, K.**, dm: An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, 1994. **31**: p. 34-38.
105. **Winn, M.D., C.C. Ballard, *et al.***, Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*, 2011. **67**(Pt 4): p. 235-42.
106. **Murshudov, G.N., A.A. Vagin, and E.J. Dodson**, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr*, 1997. **53**(Pt 3): p. 240-55.
107. **Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan**, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, 2010. **66**(Pt 4): p. 486-501.
108. **Brunger, A.T.**, Version 1.2 of the Crystallography and NMR system. *Nat Protoc*, 2007. **2**(11): p. 2728-33.

109. **Brunger, A.T., P.D. Adams*, et al.***, Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr*, 1998. **54**(Pt 5): p. 905-21.
110. **Headd, J.J., N. Echols, P.V. Afonine, R.W. Grosse-Kunstleve, V.B. Chen, N.W. Moriarty, D.C. Richardson, J.S. Richardson, and P.D. Adams**, Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr D Biol Crystallogr*, 2012. **68**(Pt 4): p. 381-90.
111. **Headd, J.J., N. Echols, P.V. Afonine, N.W. Moriarty, R.J. Gildea, and P.D. Adams**, Flexible torsion-angle noncrystallographic symmetry restraints for improved macromolecular structure refinement. *Acta Crystallogr D Biol Crystallogr*, 2014. **70**(Pt 5): p. 1346-56.
112. **Brunger, A.T.**, Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, 1992. **355**(6359): p. 472-5.
113. **Chen, V.B., W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, and D.C. Richardson**, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, 2010. **66**(Pt 1): p. 12-21.
114. **National Library of Medicine, N.C.f.B.I.**, The NCBI handbook. 2002, Internet: Bethesda.
115. **Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman**, Basic local alignment search tool. *J Mol Biol*, 1990. **215**(3): p. 403-10.
116. **Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman**, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 1997. **25**(17): p. 3389-402.
117. **Stajich, J.E., D. Block*, et al.***, The Bioperl toolkit: Perl modules for the life sciences. *Genome Res*, 2002. **12**(10): p. 1611-8.
118. **Sievers, F. and D.G. Higgins**, Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol*, 2014. **1079**: p. 105-16.
119. **Sievers, F., A. Wilm*, et al.***, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*, 2011. **7**: p. 539.
120. **Livingstone, C.D. and G.J. Barton**, Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Comput Appl Biosci*, 1993. **9**(6): p. 745-56.
121. **Jones, D.T.**, Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol*, 1999. **292**(2): p. 195-202.
122. **Barton, G.J.**, ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng*, 1993. **6**(1): p. 37-40.
123. **Crooks, G.E., G. Hon, J.M. Chandonia, and S.E. Brenner**, WebLogo: a sequence logo generator. *Genome Res*, 2004. **14**(6): p. 1188-90.
124. **Holm, L. and P. Rosenstrom**, Dali server: conservation mapping in 3D. *Nucleic Acids Res*, 2010. **38**(Web Server issue): p. W545-9.
125. **Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne**, The Protein Data Bank. *Nucleic Acids Res*, 2000. **28**(1): p. 235-42.
126. **Kabsch, W.**, A solution for the best rotation to relate two sets of vectors. *Acta Cryst D*, 1976. **A32**: p. 922-923.
127. **Krissinel, E. and K. Henrick**, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr*, 2004. **60**(Pt 12 Pt 1): p. 2256-68.
128. **Krissinel, E. and K. Henrick**, Inference of macromolecular assemblies from crystalline state. *J Mol Biol*, 2007. **372**(3): p. 774-97.
129. **Baker, N.A., D. Sept, S. Joseph, M.J. Holst, and J.A. McCammon**, Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*, 2001. **98**(18): p. 10037-41.
130. **Shultzaberger, R.K., R.E. Bucheimer, K.E. Rudd, and T.D. Schneider**, Anatomy of *Escherichia coli* ribosome binding sites. *J Mol Biol*, 2001. **313**(1): p. 215-28.

131. **Tavernarakis, N., M. Driscoll, and N.C. Kyrpides**, The SPFH domain: implicated in regulating targeted protein turnover in stomatins and other membrane-associated proteins. *Trends Biochem Sci*, 1999. **24**(11): p. 425-7.
132. **Domingue, G.J., Sr. and H.B. Woody**, Bacterial persistence and expression of disease. *Clin Microbiol Rev*, 1997. **10**(2): p. 320-44.
133. **Martin, H.H.**, Bacterial protoplasts--a review. *J Theor Biol*, 1963. **5**(1): p. 1-34.
134. **Amitai, S., Y. Yassin, and H. Engelberg-Kulka**, MazF-mediated cell death in *Escherichia coli*: a point of no return. *J Bacteriol*, 2004. **186**(24): p. 8295-300.
135. **Lioy, V.S., C. Machon, M. Tabone, J.E. Gonzalez-Pastor, R. Daugelavicius, S. Ayora, and J.C. Alonso**, The zeta toxin induces a set of protective responses and dormancy. *PLoS One*, 2012. **7**(1): p. e30282.
136. **Pribnow, D.**, Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proc Natl Acad Sci U S A*, 1975. **72**(3): p. 784-8.
137. **Schaller, H., C. Gray, and K. Herrmann**, Nucleotide sequence of an RNA polymerase binding site from the DNA of bacteriophage fd. *Proc Natl Acad Sci U S A*, 1975. **72**(2): p. 737-41.
138. **Gaal, T., W. Ross, S.T. Estrem, L.H. Nguyen, R.R. Burgess, and R.L. Gourse**, Promoter recognition and discrimination by EsigmaS RNA polymerase. *Mol Microbiol*, 2001. **42**(4): p. 939-54.
139. **Typas, A., G. Becker, and R. Hengge**, The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. *Mol Microbiol*, 2007. **63**(5): p. 1296-306.
140. **Marianovsky, I., E. Aizenman, H. Engelberg-Kulka, and G. Glaser**, The regulation of the *Escherichia coli* mazEF promoter involves an unusual alternating palindrome. *J Biol Chem*, 2001. **276**(8): p. 5975-84.
141. **Li, G.Y., Y. Zhang, M. Inouye, and M. Ikura**, Structural mechanism of transcriptional autorepression of the *Escherichia coli* RelB/RelE antitoxin/toxin module. *J Mol Biol*, 2008. **380**(1): p. 107-19.
142. **Ishihama, Y., T. Schmidt, J. Rappsilber, M. Mann, F.U. Hartl, M.J. Kerner, and D. Frishman**, Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics*, 2008. **9**: p. 102.
143. **Mutschler, H.**, Functional characterization of the chromosomally encoded toxin-antitoxin system PezAT from *Streptococcus pneumoniae*, Dissertation, in *Functional characterization of the chromosomally encoded toxin-antitoxin system PezAT from Streptococcus pneumoniae*. 2011, Ruperto-Carola University of Heidelberg: Heidelberg.
144. **de Boer, P.A., R.E. Crossley, A.R. Hand, and L.I. Rothfield**, The MinD protein is a membrane ATPase required for the correct placement of the *Escherichia coli* division site. *EMBO J*, 1991. **10**(13): p. 4371-80.
145. **Umbarger, H.E.**, Regulation of amino acid metabolism. *Annu Rev Biochem*, 1969. **38**: p. 323-70.
146. **Wilbaux, M., N. Mine, A.M. Guerout, D. Mazel, and L. Van Melderen**, Functional interactions between coexisting toxin-antitoxin systems of the ccd family in *Escherichia coli* O157:H7. *J Bacteriol*, 2007. **189**(7): p. 2712-9.
147. **Wang, X. and T.K. Wood**, Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Appl Environ Microbiol*, 2011. **77**(16): p. 5577-83.
148. **Dörr, T., M. Vulic, and K. Lewis**, Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. *PLoS Biol*, 2010. **8**(2): p. e1000317.
149. **Johnson, S.R. and S.A. Morse**, Antibiotic resistance in *Neisseria gonorrhoeae*: genetics and mechanisms of resistance. *Sex Transm Dis*, 1988. **15**(4): p. 217-24.
150. **Flett, F., G.O. Humphreys, and J.R. Saunders**, Intraspecific and intergeneric mobilization of non-conjugative resistance plasmids by a 24.5 megadalton conjugative plasmid of *Neisseria gonorrhoeae*. *J Gen Microbiol*, 1981. **125**(1): p. 123-9.
151. **Murzin, A.G.**, OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J*, 1993. **12**(3): p. 861-7.

152. **Bycroft, M., T.J. Hubbard, M. Proctor, S.M. Freund, and A.G. Murzin**, The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, 1997. **88**(2): p. 235-42.
153. **Leipe, D.D., E.V. Koonin, and L. Aravind**, Evolution and classification of P-loop kinases and related proteins. *J Mol Biol*, 2003. **333**(4): p. 781-815.
154. **Vetter, I.R. and A. Wittinghofer**, Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Q Rev Biophys*, 1999. **32**(1): p. 1-56.
155. **Dyguda, E., W. Szefczyk, and A.S. Sokalski**, The Mechanism of Phosphoryl Transfer Reaction and the Role of Active Site Residues on the Basis of Ribokinase-Like Kinases. *Int. J. Mol. Sci.*, 2004. **5**: p. 141-153.
156. **Schulz, G.E., C.W. Muller, and K. Diederichs**, Induced-fit movements in adenylate kinases. *J Mol Biol*, 1990. **213**(4): p. 627-30.
157. **Lager, G.A. and G.V. Gibbs**, Effect of Variations in O-P-O and P-O-P Angles on P-O Bond Overlap Populations for Some Selected Ortho- and Pyrophosphates. *American Mineralogist*, 1973. **58**: p. 756-764.
158. **Michaelis, L., M.L. Menten, K.A. Johnson, and R.S. Goody**, The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry*, 2011. **50**(39): p. 8264-9.
159. **Brzozowska, I. and U. Zielenkiewicz**, Regulation of toxin-antitoxin systems by proteolysis. *Plasmid*, 2013. **70**(1): p. 33-41.
160. **Aizenman, E., H. Engelberg-Kulka, and G. Glaser**, An *Escherichia coli* chromosomal "addiction module" regulated by guanosine [corrected] 3',5'-bispyrophosphate: a model for programmed bacterial cell death. *Proc Natl Acad Sci U S A*, 1996. **93**(12): p. 6059-63.
161. **Christensen, S.K., M. Mikkelsen, K. Pedersen, and K. Gerdes**, RelE, a global inhibitor of translation, is activated during nutritional stress. *Proc Natl Acad Sci U S A*, 2001. **98**(25): p. 14328-33.
162. **Jørgensen, M.G., D.P. Pandey, M. Jaskolska, and K. Gerdes**, HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *J Bacteriol*, 2009. **191**(4): p. 1191-9.
163. **Belitsky, M., H. Avshalom, A. Erental, I. Yelin, S. Kumar, N. London, M. Sperber, O. Schueler-Furman, and H. Engelberg-Kulka**, The *Escherichia coli* extracellular death factor EDF induces the endoribonucleolytic activities of the toxins MazF and ChpBK. *Mol Cell*, 2011. **41**(6): p. 625-35.
164. **Khan, A.R. and M.N. James**, Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci*, 1998. **7**(4): p. 815-36.
165. **Demidyuk, I.V., A.V. Shubin, E.V. Gasanov, and S.V. Kostrov**, Propeptides as modulators of functional activity of proteases. *BioMol Concepts*, 2010. **1**: p. 305–322.
166. **Park, J.H., Y. Yamaguchi, and M. Inouye**, Intramolecular regulation of the sequence-specific mRNA interferase activity of MazF fused to a MazE fragment with a linker cleavable by specific proteases. *Appl Environ Microbiol*, 2012. **78**(11): p. 3794-9.
167. **Shapira, A., S. Shapira, M. Gal-Tanamy, R. Zemel, R. Tur-Kaspa, and I. Benhar**, Removal of hepatitis C virus-infected cells by a zymogenized bacterial toxin. *PLoS One*, 2012. **7**(2): p. e32320.
168. **Maisonneuve, E., L.J. Shakespeare, M.G. Jorgensen, and K. Gerdes**, Bacterial persistence by RNA endonucleases. *Proc Natl Acad Sci U S A*, 2011. **108**(32): p. 13206-11.
169. **Gerdes, K.**, Introduction, in *Prokaryotic Toxin-Antitoxins*, K. Gerdes, Editor. 2013, Springer Berlin Heidelberg. p. 1-8.
170. **Garcia-Pino, A., S. Balasubramanian, L. Wyns, E. Gazit, H. De Greve, R.D. Magnuson, D. Charlier, N.A. van Nuland, and R. Loris**, Allostery and intrinsic disorder mediate transcription regulation by conditional cooperativity. *Cell*, 2010. **142**(1): p. 101-11.
171. **Donegan, N.P. and A.L. Cheung**, Regulation of the mazEF toxin-antitoxin module in *Staphylococcus aureus* and its impact on sigB expression. *J Bacteriol*, 2009. **191**(8): p. 2795-805.

172. **Vogel, J., L. Argaman, E.G. Wagner, and S. Altuvia**, The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol*, 2004. **14**(24): p. 2271-6.
173. **Kawano, M., L. Aravind, and G. Storz**, An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol*, 2007. **64**(3): p. 738-54.
174. **Domingue, G.J.**, Demystifying pleomorphic forms in persistence and expression of disease: Are they bacteria, and is peptidoglycan the solution? *Discov Med*, 2010. **10**(52): p. 234-46.
175. **Elena, S.F., T.S. Whittam, C.L. Winkworth, M.A. Riley, and R.E. Lenski**, Genomic divergence of *Escherichia coli* strains: evidence for horizontal transfer and variation in mutation rates. *Int Microbiol*, 2005. **8**(4): p. 271-8.
176. **Walk, S.T., E.W. Alm, L.M. Calhoun, J.M. Mladonicky, and T.S. Whittam**, Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environ Microbiol*, 2007. **9**(9): p. 2274-88.
177. **Meric, G., E.K. Kemsley, D. Falush, E.J. Saggers, and S. Lucchini**, Phylogenetic distribution of traits associated with plant colonization in *Escherichia coli*. *Environ Microbiol*, 2013. **15**(2): p. 487-501.
178. **Gordon, D.M. and A. Cowling**, The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology*, 2003. **149**(Pt 12): p. 3575-86.
179. **Berthe, T., M. Ratajczak, O. Clermont, E. Denamur, and F. Petit**, Evidence for coexistence of distinct *Escherichia coli* populations in various aquatic environments and their survival in estuary water. *Appl Environ Microbiol*, 2013. **79**(15): p. 4684-93.
180. **Dixit, S.M., D.M. Gordon, X.Y. Wu, T. Chapman, K. Kailasapathy, and J.J. Chin**, Diversity analysis of commensal porcine *Escherichia coli* - associations between genotypes and habitat in the porcine gastrointestinal tract. *Microbiology*, 2004. **150**(Pt 6): p. 1735-40.
181. **Savageau, M.A.**, *Escherichia coli* Habitats, Cell Types, and Molecular Mechanisms of Gene Control. *The American Naturalist*, 1983. **122**(6): p. 732-744.
182. **Archer, C.T., J.F. Kim, H. Jeong, J.H. Park, C.E. Vickers, S.Y. Lee, and L.K. Nielsen**, The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC Genomics*, 2011. **12**: p. 9.
183. **Tamman, H., A. Ainelo, K. Ainsaar, and R. Horak**, A moderate toxin, GraT, modulates growth rate and stress tolerance of *Pseudomonas putida*. *J Bacteriol*, 2014. **196**(1): p. 157-69.
184. **Pandey, D.P. and K. Gerdes**, Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res*, 2005. **33**(3): p. 966-76.
185. **Tabone, M., S. Ayora, and J.C. Alonso**, Toxin zeta reversible induces dormancy and reduces the UDP-N-acetylglucosamine pool as one of the protective responses to cope with stress. *Toxins (Basel)*, 2014. **6**(9): p. 2787-803.
186. **Mengin-Lecreulx, D., B. Flouret, and J. van Heijenoort**, Pool levels of UDP N-acetylglucosamine and UDP N-acetylglucosamine-enolpyruvate in *Escherichia coli* and correlation with peptidoglycan synthesis. *J Bacteriol*, 1983. **154**(3): p. 1284-90.
187. **Mengin-Lecreulx, D., E. Siegel, and J. van Heijenoort**, Variations in UDP-N-acetylglucosamine and UDP-N-acetylmuramyl-pentapeptide pools in *Escherichia coli* after inhibition of protein synthesis. *J Bacteriol*, 1989. **171**(6): p. 3282-7.
188. **Namboori, S.C. and D.E. Graham**, Enzymatic analysis of uridine diphosphate N-acetyl-D-glucosamine. *Anal Biochem*, 2008. **381**(1): p. 94-100.
189. **Bochner, B.R. and B.N. Ames**, Complete analysis of cellular nucleotides by two-dimensional thin layer chromatography. *J Biol Chem*, 1982. **257**(16): p. 9759-69.
190. **Cheek, S., H. Zhang, and N.V. Grishin**, Sequence and structure classification of kinases. *J Mol Biol*, 2002. **320**(4): p. 855-81.
191. **Peisajovich, S.G., L. Rockah, and D.S. Tawfik**, Evolution of new protein topologies through multistep gene rearrangements. *Nat Genet*, 2006. **38**(2): p. 168-74.
192. **Zhang, D., L.M. Iyer, A.M. Burroughs, and L. Aravind**, Resilience of biochemical activity in protein domains in the face of structural divergence. *Curr Opin Struct Biol*, 2014. **26**: p. 92-103.

193. **Grishin, N.V.**, KH domain: one motif, two folds. *Nucleic Acids Res*, 2001. **29**(3): p. 638-43.
194. **Takagi, H., Y. Kakuta, T. Okada, M. Yao, I. Tanaka, and M. Kimura**, Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nat Struct Mol Biol*, 2005. **12**(4): p. 327-31.
195. **Anantharaman, V. and L. Aravind**, New connections in the prokaryotic toxin-antitoxin network: relationship with the eukaryotic nonsense-mediated RNA decay system. *Genome Biol*, 2003. **4**(12): p. R81.
196. **Grady, R. and F. Hayes**, Axe-Txe, a broad-spectrum proteic toxin-antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Mol Microbiol*, 2003. **47**(5): p. 1419-32.
197. **Schmidt, O., V.J. Schuenemann, N.J. Hand, T.J. Silhavy, J. Martin, A.N. Lupas, and S. Djuranovic**, prlF and yhaV encode a new toxin-antitoxin system in *Escherichia coli*. *J Mol Biol*, 2007. **372**(4): p. 894-905.
198. **Vonrhein, C., G.J. Schlauderer, and G.E. Schulz**, Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure*, 1995. **3**(5): p. 483-90.
199. **Cornish-Bowden, A.**, A simple graphical method for determining the inhibition constants of mixed, uncompetitive and non-competitive inhibitors. *Biochem J*, 1974. **137**(1): p. 143-4.
200. **Dixon, M.**, The determination of enzyme inhibitor constants. *Biochem J*, 1953. **55**(1): p. 170-1.
201. **Arcus, V.**, OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol*, 2002. **12**(6): p. 794-801.
202. **Teplyakov, A., G. Obmolova, K.S. Wilson, K. Ishii, H. Kaji, T. Samejima, and I. Kuranova**, Crystal structure of inorganic pyrophosphatase from *Thermus thermophilus*. *Protein Sci*, 1994. **3**(7): p. 1098-107.
203. **Norman, A., L.H. Hansen, and S.J. Sorensen**, Conjugative plasmids: vessels of the communal gene pool. *Philos Trans R Soc Lond B Biol Sci*, 2009. **364**(1527): p. 2275-89.
204. **Eberhard, W.G.**, Evolution in bacterial plasmids and levels of selection. *Q Rev Biol*, 1990. **65**(1): p. 3-22.
205. **Van Melderen, L.**, Toxin-antitoxin systems: why so many, what for? *Curr Opin Microbiol*, 2010. **13**(6): p. 781-5.
206. **Wen, Y., E. Behiels, and B. Devreese**, Toxin-Antitoxin systems: their role in persistence, biofilm formation, and pathogenicity. *Pathog Dis*, 2014. **70**(3): p. 240-9.
207. **Whitfield, C.**, Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem*, 2006. **75**: p. 39-68.
208. **Domenech, M., E. Garcia, and M. Moscoso**, Versatility of the capsular genes during biofilm formation by *Streptococcus pneumoniae*. *Environ Microbiol*, 2009. **11**(10): p. 2542-55.
209. **Munoz-Elias, E.J., J. Marcano, and A. Camilli**, Isolation of *Streptococcus pneumoniae* biofilm mutants and their characterization during nasopharyngeal colonization. *Infect Immun*, 2008. **76**(11): p. 5049-61.
210. **Albers, S.V. and B.H. Meyer**, The archaeal cell envelope. *Nat Rev Microbiol*, 2011. **9**(6): p. 414-26.
211. **Chen, S.C., C.H. Huang, C.S. Yang, J.S. Liu, S.M. Kuan, and Y. Chen**, Crystal structures of the archaeal UDP-GlcNAc 2-epimerase from *Methanocaldococcus jannaschii* reveal a conformational change induced by UDP-GlcNAc. *Proteins*, 2014. **82**(7): p. 1519-26.
212. **Severi, E., D.W. Hood, and G.H. Thomas**, Sialic acid utilization by bacterial pathogens. *Microbiology*, 2007. **153**(Pt 9): p. 2817-22.

7 Appendix

7.1 Primer sequences

primer name	sequence[1]
	EzeT
T7_for	TAATACGACTCACTATAGGG
EzeT_3'NotI_His_r	GCATCGCGGCCGCTTGTGTATTCCCGCTG
pET_bone_for	ATGCGTCCGGCGTAGAGGATC
EzeT_3'HindIII_r	CAGCATAAGCTTCTTTGTTAGCAGCCGGATCTCAG
EzeT_83stop_HindIII_r	GCATAAGC**TTA**CGTAGTATTGTTG
EzeT_5'NcoI_start82_f	GCATCC**ATGGAT**TCAGAAAATAAAG
EzeT_5'NcoI_start62_f	CATGCC**ATG**GGTAAGATTCTCCTTCAGC
EzeT_5'NcoI_start36_f	CATGCC**ATG**GGCGATGCTCTCAAGCG
EzeT_3'SalI_r	CGTCGACCGGATCTCAGTGGTGGTGGTGGTG
EzeT_K121A_f	GGGCAACCCGGAGCAGGT**GCA**TCATCTCTGATTG
EzeT_K121A_r	CAATCAGAGATGA**TGC**ACCTGCTCCGGGTTGCCC
EzeT_M5A_D8A_f	CCATGGCTACTTTA**GCG**GAAAAA**GCT**GCACTTTTAGCC
EzeT_M5A_D8A_r	GGCTAAAAGTGC**AGC**TTTTTC**CGC**TAAAGTAGCCATGG
EzeT_N12A_Q16A_f	GATGCACTTTTA**GCC**GGAGCCAGT**GCG**TGCATCACG
EzeT_N12A_Q16A_r	CGTGATGCA**CGC**ACTGGCTCC**GGC**TAAAAGTGCATC
	ε1/ζ1
E1_5'NdeI_f	GGCCATATGAATAAAGTTGAGCCCCAAGAAAGTAA
Z1_3'HindIII_r	TTTTAAGCTTGATAATCAGCGAAATTTCCGTGATC
Z1_EcoRI_K115A_f	GGGAATTCCGGGTTCGGGA**GCA**ACTTCGACAG
E1_3'HindIII_r	GCCAAGCTTATTGCTCCTTATTTGCCGCCCACAAC
Z1_5'NcoI_f	CCGCCATGGTAAAACTGTCTAGCGATATTAATTTG
Z1_3'XhoI_His_r	GGCCCTCGAGTCTGCTGATGGATTTTTTGGCGG
T7_rev	TATGCTAGTTATTGCTCAGCG
Z1_A115K_f	GGAATTCCGGGTTCGGGA**AAA**ACTTCGACAGTAAAAACAATG
Z1_A115K_r	CATTGTTTTTACTGTCGAAGT**TTT**TCCCGAACCCGGAATTCC
	ε2/ζ2
E2_5'NdeI_f	AGCCCATATGCAAATGCAAGAAACTATTCC
Z2_K46S_r	GACT**GGA**TCCCGCTCCGGGTTGTGC
Z2_K46S_f	AGCGGGA**TCC**AGTAAGGCAGCTTCTTTAGC
Z2_3'XhoI_r	GCTACTCGAGCTCATCATCGCTTCATCG
E2_3'HindIII_r	CGTAAAGCTTCATTTTACACGTCCAACTG
Z2_5'NcoI_f	TGTACCATGGTTGAGTACACCCC
Z2_S46K_f	CCCGGAGCGGGT**AAG**AGTAAGGCAGC
Z2_S46K_r	GCTGCCTTACT**CTT**ACCCGCTCCGGG

	ε3/ζ3
Z3_K46A_f	TCAACCGGGAGCGGGA**GCA**AGTAAGACCGCCGACATC
Z3_K46A_r	GATGTCGGCGGTCTTACT**TGC**TCCCGCTCCCGGTTG
E3_5'NdeI_f	CCGGAGATATCATATGGGCATGAAAG
E3_3'HindIII_r	GGCCTTGAAGCTTACTCATACAGAACCCCCAAC
Z3_5'NcoI_f	TGGGGGTTCTCCATGGGTGATTACCAAAAG
Z3_404stop_HindIII_r	TGGTGCTCGAGTAAGCTTCATCGCTTCATC
	ezeT promoter
EzeT_5'NdeI_His_f	CATTTAACAGGAGTGTTAATG**CATATG**CACCATCACCATCACCCC
EzeT_5'NdeI_His_r	GGGGTGATGGTGATGGTG**CATATG**CATTAACACTCCTGTTAAATG
GFP_3'EcoRI_r	GACTGAATTCGGTTATGCTAGTTATTGCTC
GFP_delSD_f	CGAAAAATATCATTTAAC**GTAC**GTGTTAATGCATATGAGC
GFP_delSD_r	GCTCATATGCATTAACAC**GTAC**GTTAAATGATATTTTTCG
GFP_delPSdown_f	CCTGTAGAACAGTATTATAAT**TGTT**GAAAAATATCATTTAACAGG
GFP_delPSdown_r	CCTGTTAAATGATATTTTTC**AACA**ATTATAATACTGTTCTACAGG
GFP_delPS_f	CTCCTGTAGAACAGTAT**AGGTCG**CACCGAAAAATATC
GFP_delPS_r	GATATTTTTCGGTG**CGACCT**ATACTGTTCTACAGGAG
GFP_s70_f	GATTAAATAATATCTGAAAAGC**TTGACA**ACTCCTGTAGAACAG
GFP_s70_f	CTGTTCTACAGGAGT**TGTCAA**GCTTTTCAGATATTATTTAATC
GFP_del35_f	GATTAAATAATATCTGAAAAGCA**GCTAG**ACTCCTGTAGAACAG
GFP_del35_r	CTGTTCTACAGGAGT**CTAGC**TGCTTTTCAGATATTATTTAATC
GFP_delUPE_1_f	CCATTCGCCATTCAGGC**AAGCTT**GCCATGATTTGTTACG
GFP_delUPE_1_r	CGTAACAAATCATGGC**AAGCTT**GCCTGAATGGCGAATGG
GFP_delUPE_2_f	CATAGATTAAATAA**AAGCTT**AAAAGCAAAATCACTCC
GFP_delUPE_2_r	GGAGTGATTTTGCTTTT**AAGCTT**TTATTTAATCTATG
GFP_delprom_f	CACTCCTGTAGAACA**AGC**TTATAATCACCGAAAAATATC
GFP_delprom_r	GATATTTTTCGGTGATTATAA**GCT**TGTTCTACAGGAGTG
	Integration
HA-tag_1_f	GGAATACACAAGCGGCC**TACCCATACGATGTT**CACCACCACCAC TGAG
HA-tag_1_r	CTCAGTGGTGGTGGTG**AACATCGTATGGGTA**GGCCGCTTGTGTA TTCC
HA-tag_2_f	CCTACCCATACGATGTT**CCAGATTACGCT**TGAGATCCGGCTGC
HA-tag_2_r	GCAGCCGGATCTCA**AGCGTAATCTGG**AACATCGTATGGGTAGG
EzeT_5'NdeI_f	CTAGCATATGGCTACTTTAATGG
HA-tag_3'BamHI_r	GACTGGATCCTCAAGCGTAATCTGGAAC
Neo_5'MfeI_f	CATGCAATTGAATTAACCCTCACTAAAGGGCG
Neo_3'MfeI_r	CATGCAATTGTAATACGACTCACTATAGGGCTC
Int_f	GAGCAGGTGGTGCAGAAGCTGC
M13_rev	CAGGAAACAGCTATGAC
Int_control_f	CGTTGATTGGCAATGACCTGTC
Int_control_r	GCAGAGAAGGGGACGTTAAGG

[1] Restriction sites are underlined, mutated codons/promoter regions are in bold.

7.2 Amino acid sequence alignment of EzeT proteins from different *Escherichia coli* strains

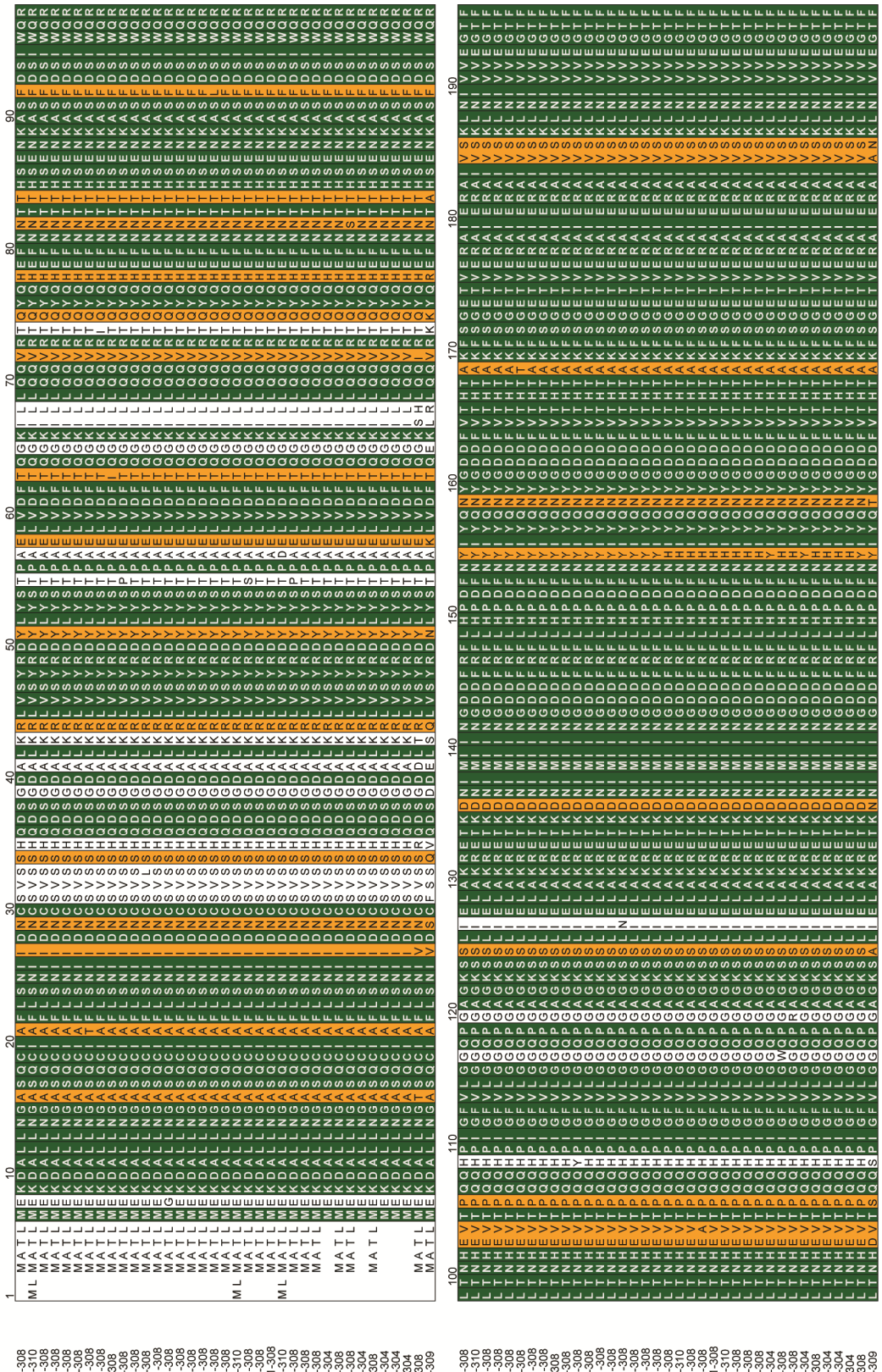


Figure 7.2-1: Amino acid sequence alignment of EzeT proteins from different E. coli strains. 37 groups of identical EzeT sequences with an overall amino acid sequence identity of 88 % and pairwise sequence identities above 96 % have been identified in 364 different E. coli strains. Most residues are either identical (green) or highly conserved (orange) between the different proteins.

7.3 Abbreviations

(*v*/*v*)	Volume per volume
(*w*/*v*)	Weight per volume
Å	Ångström
aa	Amino acid
a.u.	Arbitrary units
APS	Ammonium persulfate
AppNHp	Adenosine-5'-[(β,γ)-imido]triphosphate
ATPγS	Adenosine-5'-(γ-thiol)-triphosphate
bp	Base pairs
C-terminus	Carboxy terminus
CD	Circular dichroism
CFU	Colony forming unit
CV	Column volume
DTE	Dithioerythritol
EDTA	Ethylenediaminetetraacetic acid
GFP	Green fluorescent protein
HEPES	2-[4-(2-Hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
His_6	Hexahistidine
IPTG	Isopropyl β-*D*-1-thiogalactopyranoside
kb	Kilobase
K_d	Binding constant
kDa	Kilodalton
K_i	Inhibitory constant
K_m	Michaelis-Menten constant
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
LB	Lysogeny broth medium
LMW	Low molecular weight marker
MALDI-TOF	Matrix-assisted laser desorption/ionization time-of-flight
Mant	2'-/3'-O-(N'-Methylanthraniloyl)
MES	2-(N-Morpholino)ethanesulfonic acid
mS	Millisiemens
MW	Molecular weight
MWCO	Molecular weight cut-off
N-terminus	Amino terminus

NADH/NAD	Nicotinamide adenine dinucleotide, reduced/oxidized form
OD_{600}	Optical density at a wavelength of 600 nm
ORF	Open reading frame
PCD	Programmed cell death
PCR	Polymerase chain reaction
PDB	Protein data bank
PEP	Phosphoenol pyruvate
PMF	Peptide mass fingerprinting
RMSD	Root mean square deviation
Rpm	Revolutions per minute
PSK	Postsegregational killing
SAD	Single wavelength anomalous diffraction
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SeMet	Selenomethionine
T	Temperature
TA	Toxin-antitoxin
TEMED	N,N,N′,N′-Tetramethylethane-1,2-diamine
TEV	Tobacco etch virus
T_m	Melting temperature
Tris	Tris(hydroxymethyl)-aminomethan
U	Unit
UNAG	Uridine diphosphate-N-acetylglucosamine
UNAG-3P	Uridine diphosphate-N-acetylglucosamine-3'-phosphate
UV	Ultraviolet light
v_{max}	Maximal velocity
wt	Wild-type

Amino acids in one and three letter code

A	Ala	Alanine	M	Met	Methionine
C	Cys	Cysteine	N	Asn	Asparagine
D	Asp	Aspartic acid	P	Pro	Proline
E	Glu	Glutamic acid	Q	Gln	Glutamine
F	Phe	Phenylalanine	R	Arg	Arginine
G	Gly	Glycine	S	Ser	Serine
H	His	Histidine	T	Thr	Threonine
I	Ile	Isoleucine	V	Val	Valine
K	Lys	Lysine	W	Trp	Tryptophan
L	Leu	Leucine	Y	Tyr	Tyrosine