THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 27, pp. 19606–19618, July 6, 2007

Molecular and Structural Characterization of the PezAT Chromosomal Toxin-Antitoxin System of the Human Pathogen *Streptococcus pneumoniae*[*] [S]

Received for publication, February 27, 2007, and in revised form, April 23, 2007 Published, JBC Papers in Press, May 8, 2007, DOI 10.1074/jbc.M701703200

Seok Kooi Khoo[‡1], Bernhard Loll[§1], Wai Ting Chan[‡], Robert L. Shoeman[§], Lena Ngoo[‡], Chew Chieng Yeo[‡2], and Anton Meinhart[§3]

From the [§]Department of Biomolecular Mechanisms, Max-Planck-Institute for Medical Research, 69120 Heidelberg, Germany and the [‡]Department of Biotechnology, Malaysia University of Science and Technology, 47301 Petaling Jaya, Malaysia

The chromosomal *pezT* gene of the Gram-positive pathogen *Streptococcus pneumoniae* encodes a protein that is homologous to the zeta toxin of the *Streptococcus pyogenes* plasmid pSM19035-encoded epsilon-zeta toxin-antitoxin system. Overexpression of *pezT* in *Escherichia coli* led to severe growth inhibition from which the bacteria recovered ~3 h after induction of expression. The toxicity of PezT was counteracted by PezA, which is encoded immediately upstream of *pezT* and shares weak sequence similarities in the C-terminal region with the epsilon antitoxin. The *pezAT* genes form a bicistronic operon that is co-transcribed from a σ^{70}-like promoter upstream of *pezA* and is negatively autoregulated with PezA functioning as a transcriptional repressor and PezT as a co-repressor. Both PezA and the non-toxic $PezA_2PezT_2$ protein complex bind to a palindrome sequence that overlaps the promoter. This differs from the epsilon-zeta system in which epsilon functions solely as the antitoxin and transcriptional regulation is carried out by another protein designated omega. Results from site-directed mutagenesis experiments demonstrated that the toxicity of PezT is dependent on a highly conserved phosphoryltransferase active site and an ATP/GTP nucleotide binding site. In the $PezA_2PezT_2$ complex, PezA neutralizes the toxicity of PezT by blocking the nucleotide binding site through steric hindrance.

Toxin-antitoxin (TA)[4] systems were initially discovered on low copy number plasmids where they function as post-segregational killing systems that help ensure the segregational stability of plasmids. TA systems usually consist of two genes: the toxin gene encodes a stable protein whereas the antitoxin gene encodes either a labile protein or an untranslated, antisense RNA species. The toxic effect is neutralized by inhibition of toxin translation when the antitoxin is an RNA (type I), or by strong binding of the cognate antitoxin when the antitoxin is a protein (type II). When encoded on plasmids, TA systems were known as "addiction modules," because cells which lose these plasmids would be killed, thus causing the cells to be "addicted" to the short-lived antitoxin product because its *de novo* synthesis is essential for cell survival (1, 2).

Over the past few years, homologs of these plasmid-borne TA systems have been identified on the chromosomes of *Escherichia coli* and various other bacteria. The discovery of the *E. coli*-encoded *mazEF* TA system led to the postulation that TA modules may function as mediators of a type of programmed cell death in bacteria, because the transcription of *mazEF* was found to be inhibited by various environmental stresses such as nutrient starvation (3). The MazE antitoxin is relatively unstable compared with the MazF toxin and any inhibition of *mazEF* transcription would "activate" MazF and lead to cell death. *E. coli* was therefore hypothesized to undergo altruistic cell death, thus helping to ensure the survival of the population during adverse conditions (3). The MazF toxin was found to exert its lethality by functioning as a sequence-specific endoribonuclease (4). Since then, several homologs of MazF from *Mycobacterium tuberculosis* (5) and *Bacillus subtilis* (6) have been demonstrated to be functional endoribonucleases and are lethal when overexpressed in *E. coli*.

Another toxin that was found to act as an endoribonuclease is the RelE toxin of the *relBE* TA system encoded on the *E. coli* chromosome. However, RelE was reported to cleave mRNA molecules within ribosomes but not as free transcripts like for MazF, and the RelE-mediated cleavage of mRNA is dependent on translation (7). Experiments conducted with RelE showed that while the toxin-afflicted cells cannot form colonies, they do not lyse and can be rescued from this "dormant" state by subsequent expression of the cognate antitoxin (8). TA systems were therefore proposed to be stress response elements rather than mediators of programmed cell death, as these toxins induce a reversible stasis that enable cells to survive episodes of extreme nutritional stress (2).

* This research was supported by a SAGA/MUST/MZA-YCC/1 (M18) Grant from the Malaysian Academy of Sciences and the Ministry of Science, Technology and Innovation (to C. C. Y.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1 and Tables S1 and S2.

The atomic coordinates and structure factors (code 2P5T) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] These authors contributed equally to this work.

[2] To whom correspondence may be addressed. Tel.: 60-3-7880-1777; Fax: 60-3-7880-1762; E-mail: ccyeo@must.edu.my.

[3] To whom correspondence may be addressed. Tel.: 49-6221-486-505; Fax: 49-6221-486-585; E-mail: Anton.Meinhart@mpimf-heidelberg.mpg.de.

[4] The abbreviations used are: TA, toxin-antitoxin; IPTG, isopropyl-1-thio-β-D-galactopyranoside; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; MES, 4-morpholineethanesulfonic acid; RACE, rapid amplification of cDNA ends; r.m.s.d., root mean square deviation; HTH, helix-turn-helix; Pez, pneumococcal epsilon-zeta; ITC, isothermal titration calorimetry.

A survey of sequenced genomes indicated the widespread occurrence of TA systems among free living prokaryotes and these chromosomally encoded TA systems could be divided into eight families: RelBE, ParDE, HigBA, VapBC, MazEF, Phd/Doc, CcdAB, and epsilon-zeta (2, 9). To date only TA systems that belong to the MazEF and RelBE families have been extensively studied. The epsilon-zeta system was initially identified as an addiction module encoded on plasmid pSM19035 of *S. pyogenes* with epsilon functioning as the cognate antitoxin to the zeta toxin. The epsilon, zeta, and omega genes form an operon on pSM19035 with the omega gene product acting as a repressor of transcription for omega-epsilon-zeta along with other transcription units on pSM19035 (10, 11). The products of the epsilon and zeta genes were shown to form a stable epsilon$_2$zeta$_2$ heterotetramer (12, 13). The toxic protein zeta contains a highly conserved Walker-A motif for ATP/GTP binding, as well as a phosphoryltransferase active site. Epsilon neutralizes the toxicity of zeta by blocking access to the phosphoryltransferase active site through formation of a tight complex with zeta (12). However, the mode of action of the zeta toxin remained elusive.

A homolog of zeta was found to be encoded in the genome of the Gram-positive pathogen *S. pneumoniae* which harbors at least five other TA loci (12). Two of these pneumococcal TA loci, namely *relBE-2Spn* and *yefM-yoeBSpn*, were recently shown to be functional TA systems (14, 15). In this report, we present the crystal structure of the pneumococcal epsilon-zeta homolog, designated PezAT (for pneumococcal epsilon-zeta) and show that it is a functional TA system, with PezA as the cognate antitoxin to the PezT toxin. We also show that *pezAT* is negatively autoregulated and, unlike the omega-epsilon-zeta system, PezA functions as a transcriptional repressor with PezT as a co-repressor.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions—E. coli DH5α was used as a host strain for most of the cloning experiments and transcriptional analyses using *lacZ* fusions. Cloning of recombinant bait and target vectors for bacterial two hybrid assays was performed in *E. coli* XL-Blue MRF′ Km (Stratagene). BacterioMatch II *E. coli* validation reporter (Stratagene) was co-transformed with both bait and target pairs to enable detection of protein-protein interactions through transcriptional activation of the *HIS* reporter gene. *E. coli* BL21(DE3)/pLysS was used as the host strain to investigate the effect of the expression of the *pezAT* genes on cellular growth and *E. coli* BL21(DE3)/Codon Plus-RIL strain (Stratagene) was used for overexpression of the PezA and PezT proteins for purification. All cultures were grown in LB medium (Difco) at 37 °C unless otherwise stated. When necessary, the growth media was supplemented with antibiotics at the following concentrations: ampicillin (Ap), 100 μg/ml; kanamycin (Km), 50 μg/ml; chloramphenicol (Cm), 34 μg/ml for maintenance of *E. coli* BL21 strains, and 10 μg/ml for the bacterial two-hybrid assay strains.

Construction of Plasmids for Overexpression of PezA and PezT in E. coli—The *pezT* open reading frame was PCR-amplified from the *S. pneumoniae* genome using the primer pair PezT-F and PezT-R (for these and all other primer sequences see supplemental Table 1). The 762-bp PCR fragment obtained was initially cloned into the pGEM-T Easy PCR cloning vector (Promega). The resulting recombinant pGEM-PezT plasmid was digested to release the cloned *pezT* fragment, which was then ligated into pET11a to generate a pET11a-PezT plasmid. Similarly, a DNA fragment spanning both the *pezA* and *pezT* genes was PCR-amplified using the primers PezA-F and PezT-R, producing a single 1236-bp PCR product, which was then digested and cloned into pET11a resulting in a construct designated pET11a-PezAT. For separate expression of PezA, the open reading frame was cloned into pET28a by amplification with the primers PezA-EcoRI-F and PezA-XhoI-R and the restriction digested 477-bp fragment was ligated into pET28a resulting in a pET28a-PezA recombinant plasmid. For protein purification, both the *pezAT* and *pezA* open reading frames were each cloned into the pET28a vector such that their expression would result in an N-terminal hexahistidine fusion protein. The construct expressing the *pezAT* open reading frame, pET28a-(His)$_6$-PezAT, was obtained using the primers PezA-F and PezT-R for amplification, whereas the *pezA* construct, designated pET28a-(His)$_6$-PezA, was generated with the primers PezA-F and PezA-R for amplification. Site-directed mutagenesis was performed as described in Ref. 16.

Cell Viability Assay—E. coli BL21(DE3)/pLysS cells were individually transformed with either pET11a-PezT or pET11a-PezAT and co-transformed with pET11a-PezT together with pET28a-PezA. Overnight cultures of *E. coli* BL21(DE3)/pLysS carrying the recombinant pET11a plasmids were diluted to an A_{600} of ~0.02 in fresh LB broth supplemented with Ap and Cm and cells were allowed to grow at 37 °C to an A_{600} of ~0.45. Cells carrying pET28a recombinant plasmids were grown in the presence of Km instead of Ap. Overexpression of the cloned genes was induced at mid-log phase (A_{600} ~ 0.45) by addition of 1 mM IPTG and A_{600} values were monitored at hourly intervals for up to 7 h following induction. At the same time intervals, aliquots of the cultures were taken to determine the number of viable cells. Identical overexpression experiments were also conducted with strains grown in M9 minimal medium (Fermentas).

Overexpression and Purification of Recombinant (His)$_6$-PezAT and (His)$_6$-PezA—E. coli BL21(DE3)/Codon Plus-RIL were transformed with plasmids pET28a-(His)$_6$-PezAT or pET28a-(His)$_6$-PezA. Cells were grown in LB-medium to an A_{600} of 0.5 at 37 °C, cooled to 20 °C and protein overproduction was induced with 1 mM IPTG overnight. Selenomethionine labeling was performed according to Ref. 17, and induction was performed similarly to wild-type protein. For purification of the PezAT protein complex, cells were harvested by centrifugation and the cell pellet was resuspended in buffer A (see supplemental Table S2 for buffer compositions). Cell walls were broken by sonication, and the slurry was cleared by centrifugation. Affinity chromatography was performed using a Ni-NTA material (Qiagen) equilibrated with buffer A. Unbound sample was removed by washing with buffer A. The protein-complex was eluted with buffer A with additional 250 mM imidazole. Fractions containing the PezAT protein complex were pooled and the N-terminal hexahistidine-tag of PezA was removed by

thrombin (GE Healthcare) cleavage in a dialysis against cleavage buffer B at 4 °C for 12 h. The protein-complex was further purified over a MonoQ column (GE Healthcare). Briefly, the dialysed PezAT protein complex was diluted with buffer C and loaded onto the column. The PezAT proteins were eluted in a gradient of up to 600 mM NaCl and were subjected to a final clean-up by size exclusion chromatography using Superose 6 (GE Healthcare) equilibrated with buffer D. Selenomethionine-labeled PezAT proteins were purified similarly to the wild type protein with the addition of 5 mM DTE throughout all purification steps.

For purification of PezA, harvested cells were resuspended in buffer I, and the cells were lysed by sonication. The cell debris was cleared by ultracentrifugation and the crude lysate was subsequently subjected to ammonium sulfate fractionation to a final concentration of 40% saturation. After centrifugation, the pellet was dissolved in and dialyzed against buffer J. The proteins were further applied onto a HisTrap HP affinity column (GE Healthcare) equilibrated with buffer J. Following extensive washing with buffer J, the bound proteins were eluted from the column using buffer J containing 500 mM imidazole.

For purification of the mutated variant PezT(K45A), harvested cells were resuspended in buffer E. After cell lysis by sonication, the cleared supernatant was loaded over a HisTrap FF column (GE Healthcare) equilibrated with buffer E. Non-specifically bound proteins were removed in a high salt wash with buffer E containing additional 500 mM NaCl, and bound proteins were eluted in a gradient to 500 mM imidazole. The pooled protein fractions were diluted with buffer F prior to loading on a heparin column (GE Healthcare), equilibrated with buffer G. PezT(K45A) was then eluted in a gradient to 600 mM NaCl. Fractions containing PezT(K45A) were further purified over Superose 12 (GE Healthcare) size exclusion column equilibrated with buffer H.

For crystallization experiments, the protein was concentrated to 17 mg/ml. Protein concentrations were determined by the Bradford assay (Bio-Rad). Protein purity, complete cleavage of the hexahistidine tag and incorporation of selenomethionine were monitored by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF MS).

Mass Spectrometry—MALDI-TOF mass spectrometry analysis was performed in the linear, positive ion mode with blanking ($<$700 m/z) and pulsed (time-delayed) extraction using a Shimadzu Biotech Axima TOF^2 instrument (Shimadzu Biotech Germany). All reagents and protein standards were purchased from Sigma-Aldrich. Sinapinic acid (10 mg/ml in 50% acetonitrile, 50% of 0.1% trifluoroacetic acid (all % v/v)) was used as a matrix. Sample positions on the steel 384-position sample plate were washed once with matrix solution. A 1-μl drop of the protein preparation or 1 μl of water into which the crystal was transferred with the cryoprotectant was allowed to dry onto the sample plate at room temperature and the preparation was then washed with a small drop of 0.1% trifluoroacetic acid applied for ~10 s. Immediately after removal of the trifluoroacetic acid, 1 μl of matrix solution was added to the sample and it was allowed to dry at room temperature. The standard proteins (about 1 pmol of insulin, ubiquitin, cytochrome *c*, myoglobin, chymotrypsinogen A, aldolase, bovine serum albumin, conalbumin, or β-galactosidase) were spotted onto the washed plate and an equal volume (usually 1 μl) of matrix was immediately added to the protein drop. Each protein standard was analyzed separately and a combined calibration of the near external standards was employed to determine the mass/charge (m/z) values.

Isothermal Titration Calorimetry—ITC was performed using a VP-ITC microcalorimeter (MicroCal Inc.). Data were analyzed using the Origin-based software provided by the manufacturer. A buffer exchange to buffer D was performed by size exclusion chromatography and protein concentrations were determined according to their absorbance at 280 nm. Titrations were carried out at 15 °C and, in a typical experiment, 3 μM PezT(K45A) were loaded into the sample cell and titrated against 30 μM PezA in the injection syringe.

Protein Crystallization and Data Collection—Crystals of the PezAT protein complex were obtained from drops made of 1.5 μl of protein solution and 1.5 μl of precipitant solution (12.5 to 15% (v/v) isopropyl alcohol, 100 mM MES-NaOH, pH 6.0) in a hanging-drop vapor diffusion setup at 293 K. Crystal size was optimized by adding 1 to 6% (v/v) of a 30% (v/v) dioxane stock solution to the mixed crystallization drop. Single individual crystals grew to a maximum size of about $300 \times 100 \times 50$ μm^3 within 2 weeks. Before flash-cooling in liquid nitrogen, crystals were harvested into precipitant solution with additional 10% (v/v) glycerol, and the glycerol concentration was increased stepwise to 30% (v/v) for cryo protection. Native and anomalous diffraction data were collected at beamline PXII at the Swiss Light Source with the crystals kept at 90 K. The datasets were processed and scaled using XDS (18). The selenium sites were located with the program SHELXD (19). Initial phase improvement and density modification was performed with SHELXE (20). Model building was performed with the program "O" (21) and refined using CNS (22) in the beginning and REFMAC (23) in the later stage of refinement. The overall quality of the structure was evaluated by PROCHECK (24).

In Vivo Detection of Protein-Protein Interaction Using a Bacterial Two-hybrid System—The BacterioMatch II Two-Hybrid System Vector Kit (Stratagene) was employed according to the manufacturer's instructions to investigate *in vivo* interactions between PezA and PezT. Both *pezA* and *pezT* were separately cloned as fusion proteins with the λcI repressor into pBT as bait constructs and in frame with the RNAPα reading frame into pTRG as target vectors. In the case of *pezA*, DNA fragments were amplified using the primer pair BPezA-F and BPezA-R and cloned into the pBT bait vector for the bait construct pBT-PezA. The recombinant target of *pezA* was generated by PCR using primer pair TPezA-F and BPezA-R and cloned into pTRG for the target construct pTRG-PezA. For generation of bait and target constructs of *pezT*, an intermediate cloning step into the pGEM-T Easy PCR cloning vector (Promega) after amplification with the primer pairs PezT1-F and PezT1-R was introduced. Following restriction digest, the isolated *pezT* DNA fragment was ligated into pBT for the bait construct pBT-PezT and into pTRG for the target vector construct pTRG-PezT. The *E. coli* reporter strain (Stratagene) was co-transformed with each of the recombinant bait and target pairs (*i.e.* pBT-PezA and pTRG-PezT pair, or pBT-PezT and pTRG-PezA pair). Pos-

itive interaction between the λcI-fusion protein produced by the recombinant pBT and the RNAPα-fusion protein produced by the recombinant pTRG was monitored by growth of the reporter strain co-transformants on the selective medium containing 5 mM 3-amino-1,2,4-triazole (3-AT).

RNA Preparation and Reverse Transcriptase-PCR (RT-PCR)—Total RNA samples were prepared from *E. coli* cultures using an RNeasy Mini kit (Qiagen) according to the manufacturer's protocol. The total RNA concentrations were determined by UV spectrophotometry, and the sample integrity was checked by agarose gel electrophoresis. RT-PCR was carried out using 400 ng of total RNA with the Access RT-PCR System (Promega) according to the manufacturer's procedure. The gene specific primers 1050-F, 1050-R, 1051-F, and 1051-R were used in the RT-PCR analysis.

Construction of lacZ Transcriptional Fusions and β-Galactosidase Assays to Detect Promoter Activity—Four constructs spanning different regions of the *pezAT* operon were amplified and subsequently ligated into the vector pQF52 (25) to obtain different *lacZ* transcriptional fusions for the detection of promoter activity. First, the putative promoter region upstream from *pezA* was amplified using the primer pair upPezA-F and upPezA-R yielding a 209-bp fragment, which was cloned into pQF52 resulting in the pQF-PezAup construct. Second, the promoter region together with *pezA* was amplified (using the upPezA-F and lacZ_PezA-R primers) as a 712-bp fragment, which was then cloned to generate pQF-PezA. Third, the entire *pezAT* operon (1388-bp DNA fragment) along with its promoter region was obtained with the primer pair upPezA-F and lacZ_PezT-R for the recombinant pQF-PezAT. Finally, pQFPezTup was constructed by ligation of pQF52 with a 233-bp fragment that contained a truncated *pezA* reading frame that was amplified using the primer pair PezA283-F and PezA-R. *E. coli* DH5α was transformed with the recombinant pQF52 plasmids and transformants were assayed for β-galactosidase activity (when A_{600} reached ~0.5) according to the method of Miller (26), using SDS and chloroform to permeabilize the cells.

Determination of lacZ Transcript Levels using Quantitative Real-time RT PCR—The LightCycler RNA master SYBR Green I kit (Roche Applied Science) was used to carry out quantitative real-time RT-PCR on the LightCycler 2.0 (Roche Applied Science) for the detection of *lacZ* transcript levels using the primer pair lacZ-F and lacZ-R. Total RNA from *E. coli* DH5α harboring the various pQF-derived recombinants were extracted using the RNeasy mini kit. Reverse transcription was carried out at 61 °C for 20 min, and the cDNA/RNA hybrid was then denatured at 95 °C for 30 s. After that, 45 cycles of real-time PCR amplification were carried out using the following parameters: denaturation for 1 s at 95 °C, annealing for 5 s at 5 °C below the predicted melting temperature of the primer pair, and elongation for 13 s at 72 °C. The *E. coli*-encoded 16S rRNA gene was used as a housekeeping control (using the primer pair 16S-F and 16S-R). The relative expression ratio of the target gene, *lacZ*, was calculated based on its real-time PCR efficiencies (*E*) and the crossing point difference (ΔCP) for an unknown sample *versus* a control using the mathematical model published by Pfaffl (27). For each gene (*i.e. lacZ* or 16S rRNA), cDNA dilution curves were generated and used to calculate the individual real-time PCR efficiencies with the values for the slopes obtained using the LightCycler 4.0 software (Roche Applied Science). For each sample, at least three different real-time RT PCR reactions were carried out from which the mean CP value was obtained utilizing the LightCycler 4.0 software (Roche Applied Science).

Determination of the Transcriptional Start Site using 5′-RACE—The transcriptional start site of the *pezAT* operon was determined using 5′-RACE (Version 2.0; Invitrogen), following the manufacturer's protocol. Total RNA was extracted from *E. coli* DH5α harboring the pQF-PezAT recombinant plasmid and was subjected to first strand cDNA synthesis using 20 ng of GSP1 primer, which was designed to anneal 399-bp downstream of the *pezA* start codon. A homopolymeric tail was then added to the 3′-end of the synthesized cDNA using terminal deoxytransferase and dCTP. The dC-tailed cDNA was then PCR-amplified using the provided anchor primer, which was complementary to the homopolymeric tail, and a nested primer, GSP2, located 121-bp downstream of the *pezA* start codon. The amplified product was purified and sequenced to determine the transcriptional start site, which would be the nucleotide that is immediately after the poly(G) tail.

Electrophoretic Mobility Shift Assays—A 203-bp DNA fragment containing palindrome sequence PS, which is the putative regulatory protein binding site, was PCR-amplified using the primer pair PS-F and PS-R, purified and labeled with biotin using the Biotin 3′-End DNA Labeling kit (Pierce). Subsequent electrophoretic mobility shift assays were carried out using the LightShift chemiluminescent EMSA kit (Pierce). Biotinylated DNA (20 fmol) was incubated at room temperature for 20 min with 50 ng/μl polydeoxyinosinic-polydeoxycytydilic acid [poly(dI-dC)] and different amounts of purified protein (either PezA or the PezAT protein complex) in binding buffer K. Nucleoprotein complexes were separated from free DNA by electrophoresis on a 6% polyacrylamide gel in 0.5× Tris borate/EDTA buffer and subsequently transferred onto a nylon membrane. The biotin end-labeled DNA was detected using a streptavidin-horseradish peroxidase conjugate and the supplied chemiluminescent substrate.

RESULTS

Sequence Analysis—A homolog of the *S. pyogenes* plasmid pSM19035-encoded zeta toxin was found in the genome sequence of *S. pneumoniae* TIGR4. The PezT protein encoded by the *sp1051* reading frame shares 42% sequence identity with zeta (Fig. 1*A*) suggesting that both proteins have similar functions in different Streptococcal species. However, a BLAST search for the corresponding antitoxin homolog to the epsilon protein in the genome sequence of *S. pneumoniae* did not yield any positive results. As TA systems are usually organized as operons, the reading frames that flank *sp1051* were further analyzed. Interestingly, pairwise amino acid sequence alignment of the upstream open reading frame *sp1050* gene product PezA with epsilon indicated that the C-terminal residues 70–158 of PezA share 21% sequence identity with epsilon (Fig. 1*A*). Apparently arranged in a bicistronic operon, the *sp1050* (*pezA*) reading frame overlaps with *sp1051* (*pezT*) by one nucleotide, where the last A nucleotide of the TAA stop codon of *pezA* is

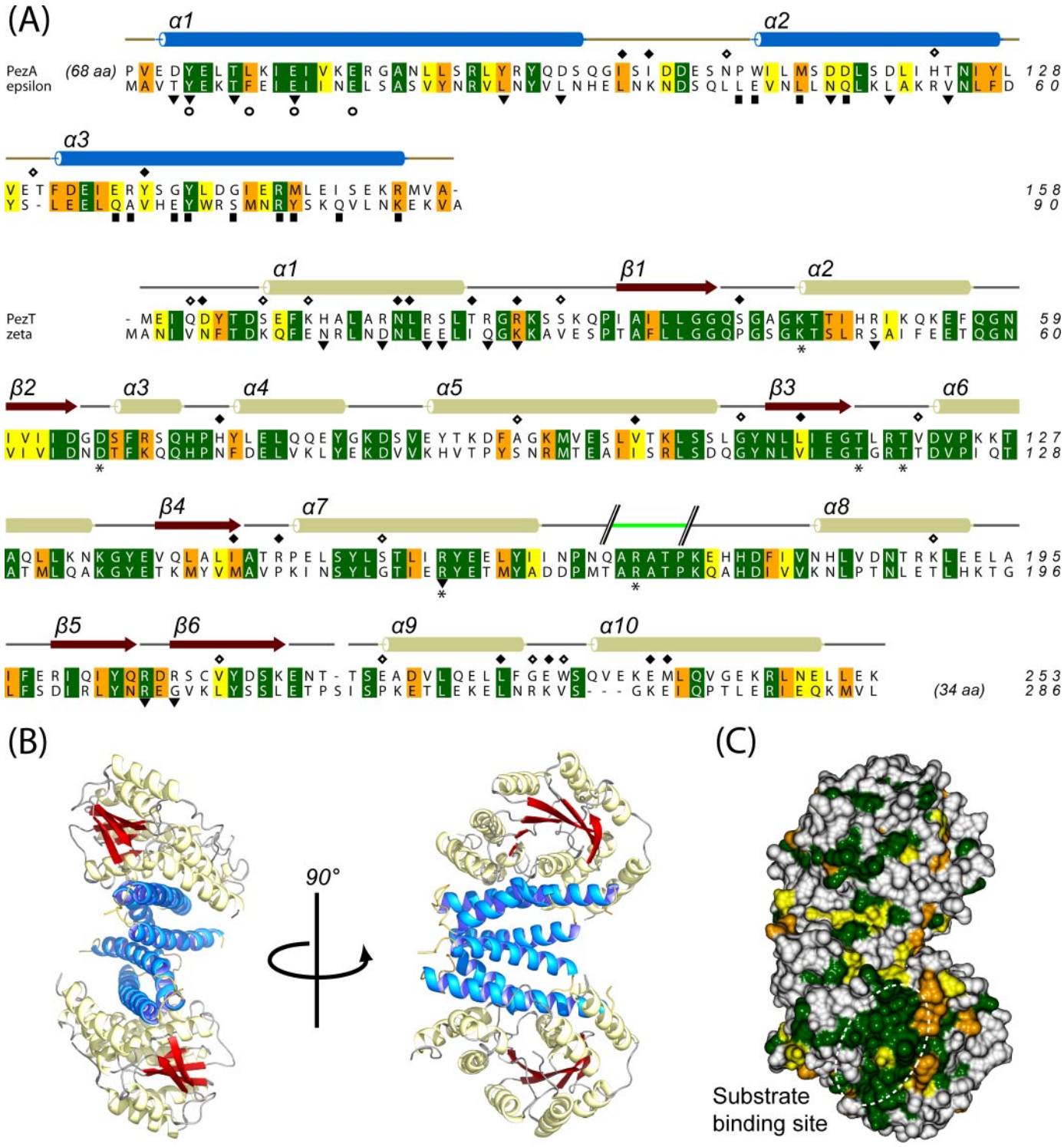

FIGURE 1. *A*, sequence alignment of the C-terminal domain of PezA and PezT with the epsilon/zeta system. Identical residues are colored *green*, highly homologous residues *orange*, and homologous residues *yellow*, respectively. Secondary structures are shown above the sequence alignment with α-helices illustrated as *cylinders* and β-strands as *arrows*. The region in PezT that could not be modeled into the electron density is indicated by a *green line*. Residues that are involved in PezA-PezA interaction are labeled with *filled squares beneath* the sequence alignment, and those that are involved in PezA-PezT interaction with *filled triangles*. *Open circles* label residues of PezA that inhibit the toxin PezT. *Asterisks* indicate those residues that were changed by site-directed mutagenesis in PezT and were shown to abolish PezT toxicity. *Filled diamonds above* the sequence alignment indicate residues of different PezT variants where mutations to homologous residues were observed, *open diamonds* mark mutations of a non-conservative nature. *B*, ribbon representation of the heterotetrameric PezA/PezT arrangement. Helices of PezA are colored in *blue* and helices of PezT in *beige*, respectively. β-Strands in PezT are colored in *red*. *C*, surface representation of the complex colored according to the color scheme of *A*.

also the start of the ATG initiation codon of *pezT*. This arrangement is similar to that found in the other two *S. pneumoniae* genomes that have been sequenced, strains R6 (designated spr0951–spr0952; accession no. NC_003098) and D39 (designated SPD_0930-SPD_0931; accession no. CP000410). Furthermore, we found that the N-terminal 70 – 80 amino acids of

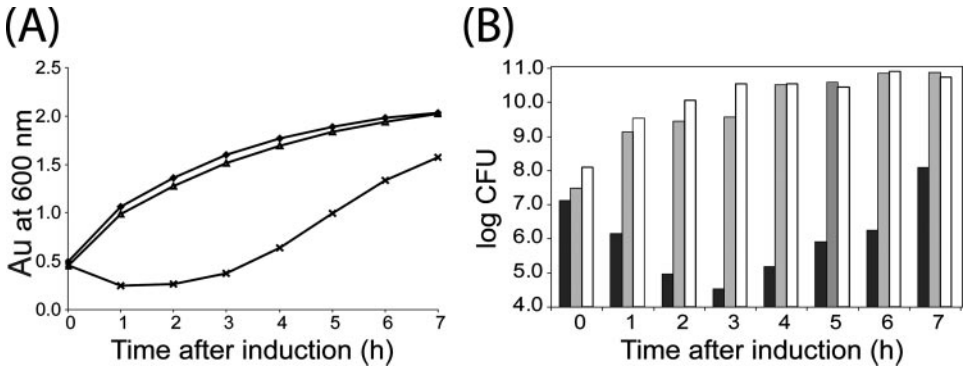

FIGURE 2. **Effect of PezT and PezAT overexpression on the growth of *E. coli*.** *A*, *E. coli* BL21(DE3)/pLysS cells harboring plasmid pET11a (vector alone, ♦), or pET11a-pezT (expressing the PezT toxin, ×), or pET11a-pezAT (expressing the PezA antitoxin and the PezT toxin, Δ) were grown in LB broth at 37 °C until A_{600} of ~0.45. IPTG was added to the medium at a final concentration of 1 mM to induce the expression of the gene(s) carried on the pET11a vector. Cellular growth was then monitored at hourly intervals by measuring A_{600} and by determining the number of colony-forming units (*CFU*) (*B*) by plating out appropriate dilutions onto LB agar and enumerating them after incubation at 37 °C overnight. *Bars* representing log CFU are for *E. coli* harboring pET11a vector (*white bars*), pET11a-pezT (*black bars*), and pET11a-pezAT (*gray bars*). *AU*, absorbance units.

PezA share 30 to 40% sequence identity with transcriptional regulators of the Xre and Cro/CI families. Thus *pezAT* was investigated to determine if the two genes encode for a functional toxin-antitoxin pair that is autoregulated by PezA, which also functions as the cognate antitoxin to the PezT toxin.

Overexpression of pezAT in E. coli—To determine if *pezAT* is indeed a functional TA system and whether *pezT* encodes for the toxic protein, the *pezAT* genes were cloned together as well as separately into the pET11a expression vector. The *pezAT* genes were obtained by PCR amplification using genomic DNA from clinical isolates of *S. pneumoniae* as template. The recombinant plasmids were sequenced on both strands for verification. We found that the cloned *pezAT* sequences obtained from six *S. pneumoniae* clinical isolates differed by between 4–6% when compared with the data base sequence for the TIGR4 strain. Although most of these sequence differences led to silent mutations (no frameshift mutations were identified), the translated PezA and PezT proteins do differ by up to 3% with the TIGR4 strain (Fig. 1*A*). A similar degree of variation was found when the *pezAT* sequences from the R6 and D351 strains in the data base were compared with the TIGR4 strain, thus indicating a natural degree of inherent sequence variability within this locus. The *pezAT* fragment showing the least variation from the TIGR4 sequence was chosen for this study. Here, the PezA amino acid sequence is identical with the TIGR4 sequence whereas the PezT sequence showed five differences: K26R, R109G, V121I, K146R, and A220E. Two of these differences, R109G and V121I, were also found in the PezT sequences of the R6 and D351 strains. Except for R109G, all non-homologous sequence variations observed within the PezT family align with non-conserved amino acids between PezT and zeta (Fig. 1*A*). Interestingly, all isolated variants had a glycine replacement for residue Arg^{109}. Whereas cloning of both *pezA* and *pezT* on a single DNA fragment proceeded without any problems, cloning of the toxin-encoding *pezT* was difficult and was successful only after several attempts. Similar problems had also been reported for the epsilon/zeta locus (28).

In cell viability assays, overexpression of PezT resulted in growth inhibition with a reduction of $\sim 10^3$ in cell viability for the first hour post-induction. The reduction in cell growth was even more drastic for the subsequent 2 h where the viable counts dropped by $\sim 10^5$ compared with the uninduced state (Fig. 2). However, restoration of cell growth was observed ~3-h post-induction with a gradual increase in both the A_{600} value as well as viable cell counts (Fig. 2). Following overnight growth, there were no discernable differences in both the A_{600} as well as the viable cell counts of the *E. coli* BL21 cells overexpressing PezT when compared with its uninduced counterpart and the negative control (data not shown). This growth profile is remarkably similar to what has been reported for the overexpression of zeta in *E. coli* (28). The similar experiment was repeated by growing the cells in M9 minimal medium instead of LB broth which revealed an almost identical growth profile (data not shown). Sequencing of plasmid DNA isolated from cells that have recovered from stasis did not reveal any mutations in the *pezT* gene nor in the T7 promoter region of the pET11a expression vector.

This inhibition in cell growth was not observed when *pezA* was co-expressed along with *pezT* in *E. coli* BL21 harboring the pET11a-pezAT construct. To investigate if neutralization of PezT toxicity by PezA could also occur *in trans*, a pET28a-PezA plasmid was co-transformed with pET11a-PezT into *E. coli* BL21(DE3)/pLysS and the transformants obtained were subjected to IPTG induction for protein expression. Cellular growth was not inhibited, showing that PezA was able to neutralize PezT toxicity even when the *pezA* reading frame was provided *in trans*. Taken together, these results indicate that PezT functions as a toxin that inhibits the growth of *E. coli* cells for the first 3 h of overexpression and this toxicity is abolished by co-expression of PezA, which apparently functions as its cognate antitoxin.

PezA Interacts with PezT and Forms a Tight Complex—The pET28a-pezAT recombinant plasmid was constructed such that the PezA protein is expressed as a hexahistidine fusion protein from the T7 promoter of the pET28a vector. In pull-down experiments, PezT was successfully co-purified with $(His)_6$-PezA immobilized on Ni-NTA agarose beads. Following elution, two distinct protein bands were observed on SDS-PAGE corresponding to the expected mobilities of the $(His)_6$-PezA fusion protein (19 kDa) and PezT (29 kDa) indicating that under native conditions, both PezA and PezT were co-purified (Fig. 3*A*). Mass spectrometry demonstrated that both PezA and PezT had the expected mass.

To characterize the binding properties of PezA and PezT in detail, ITC experiments were performed. The proteins bind to each other with an apparent K_D below 1 nM (Fig. 3*B*). This low K_D is at the borderline of measurable affinities by ITC, because the signal to noise ratio decreases at the required low concentrations. When PezA was titrated with PezT, the signal slowly

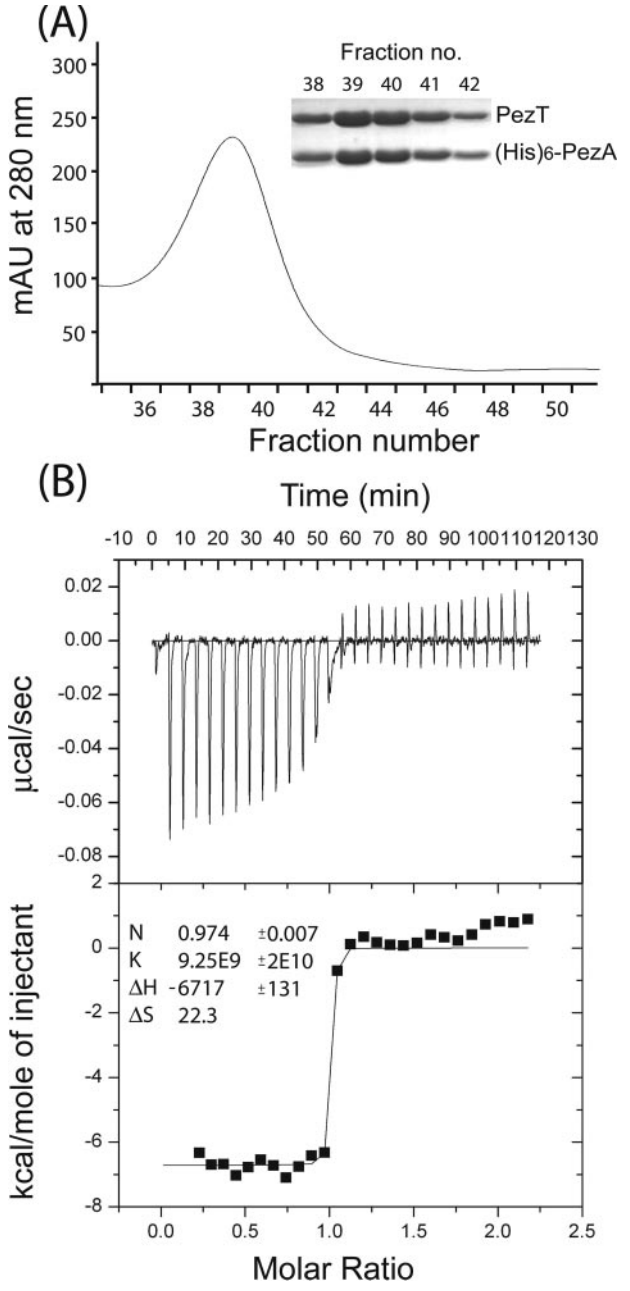

FIGURE 3. *A*, pull-down experiment of $(His)_6$-PezA and PezT. Chromatogram showing elution profile of proteins from the nickel affinity column on which crude extract of *E. coli* BL21 Codon Plus RIL strain harboring pET28-pezAT was applied. SDS-PAGE analysis of the eluted fractions is shown in the *inset*. *B*, ITC experiment with PezA and PezT(K45A). *Top*, thermogram resulting from a titration of 3 μM PezT in the sample cell with 30 μM PezA in the injection syringe. *Bottom*, best fit (*solid line*) to the experimental data (*filled squares*). Given results were obtained from an analysis of the data assuming a single binding site model.

decreased to background level, possibly due to disassembly of PezA oligomers before PezA is able to bind to PezT. This finding is supported by size exclusion chromatography experiments, where PezA has a shorter retention time than expected, suggesting that PezA alone tends to forms higher oligomeric aggregates. The obtained stoichiometric relation for the complex is around 1 suggesting that *in vitro* both proteins form oligomers with a stoichiometry of $PezA_nPezT_n$.

In vivo interactions between the PezA and PezT proteins were investigated using a bacterial two-hybrid system (see "Experimental Procedures") where growth of the *E. coli* indicator strain on the 5 mM 3-AT selective medium was observed for strains that were co-transformed with pBT-pezA and pTRG-pezT (data not shown). Interestingly, when the interaction pair was switched, *i.e.* using pBT-pezT and pTRG-pezA, no growth of the *E. coli* indicator strain was detected on the selective medium.

Transcriptional Analyses of pezAT—Analyses of sequences upstream of *pezA* indicated the presence of a characteristic −10 promoter sequence (TATAAT) 33 nucleotides upstream of the *pezA* start codon as well as an imperfect −35 sequence (TTGTGT) that is separated by 18 nucleotides from the −10 hexamer (Fig. 4*A*). A 56-bp imperfect palindrome sequence, designated PS, overlaps both the −10 and −35 region and is centered 59 nucleotides upstream of *pezA*. A putative ribosome-binding site (AGGAG) could also be found 6 nucleotides upstream of the ATG start codon (Fig. 4*A*).

To obtain information on the activity of the putative promoter designated P_{pezAT}, transcriptional fusions with a promoterless *lacZ* gene as a reporter in the *E. coli* plasmid pQF52 were carried out. The functionality of P_{pezAT} was validated when *E. coli* DH5α cells harboring pQF-pezAup (191 bp of the *pezA* upstream region cloned upstream of *lacZ* in pQF52) showed mean β-galactosidase activity of 130 ($\pm$25) Miller units compared with a value of less than 5 for cells harboring the pQF52 vector alone (Fig. 4*D*). No β-galactosidase activity could be detected in cells harboring pQF-pezTup (which contained sequences in the 193-bp region upstream of the *pezT* start codon) indicating no detectable promoter activity within the *pezA* reading frame immediately upstream of *pezT*.

The *pezA* and *pezT* reading frames overlap by 1 nucleotide and would appear to constitute a bicistronic operon, as in the case of other TA systems (1, 2). Total RNA was isolated from *E. coli* DH5α harboring the pQF-PezAT recombinant constructs, which contained the entire *pezAT* reading frames as well as 191 bp of sequences upstream of *pezA* encompassing the P_{pezAT} promoter. Reverse transcriptase was used to synthesize cDNAs that were complementary to *pezAT* mRNA and the resulting cDNAs were PCR-amplified using two oligonucleotide primers (1050-F and 1051-R) that would anneal to the 5′-end of *pezA* and the 3′-region of *pezT*, respectively (Fig. 4*B*). A single band was observed corresponding to the expected size if a single continuous RNA transcript containing *pezAT* were present (Fig. 4*C*). Sequencing of this PCR product confirmed its identity. The individual *pezA* and *pezT* genes were also amplified using primers that anneal to regions within the respective reading frames as controls (Fig. 4*C*). Possible DNA contamination can be excluded since no PCR products were evident in the negative controls without a reverse transcriptase extension step but still applying the same RNA samples.

Total RNA isolated from *E. coli* harboring pQF-PezAT was subjected to 5′-RACE to determine the transcriptional start site. Sequencing of the 5′-RACE products obtained from three different preparations of RNA showed that the *pezAT* transcript initiated at the A residue 21 bp upstream of the *pezA* start codon (designated +1 in Fig. 4*A*). Taken together, these results

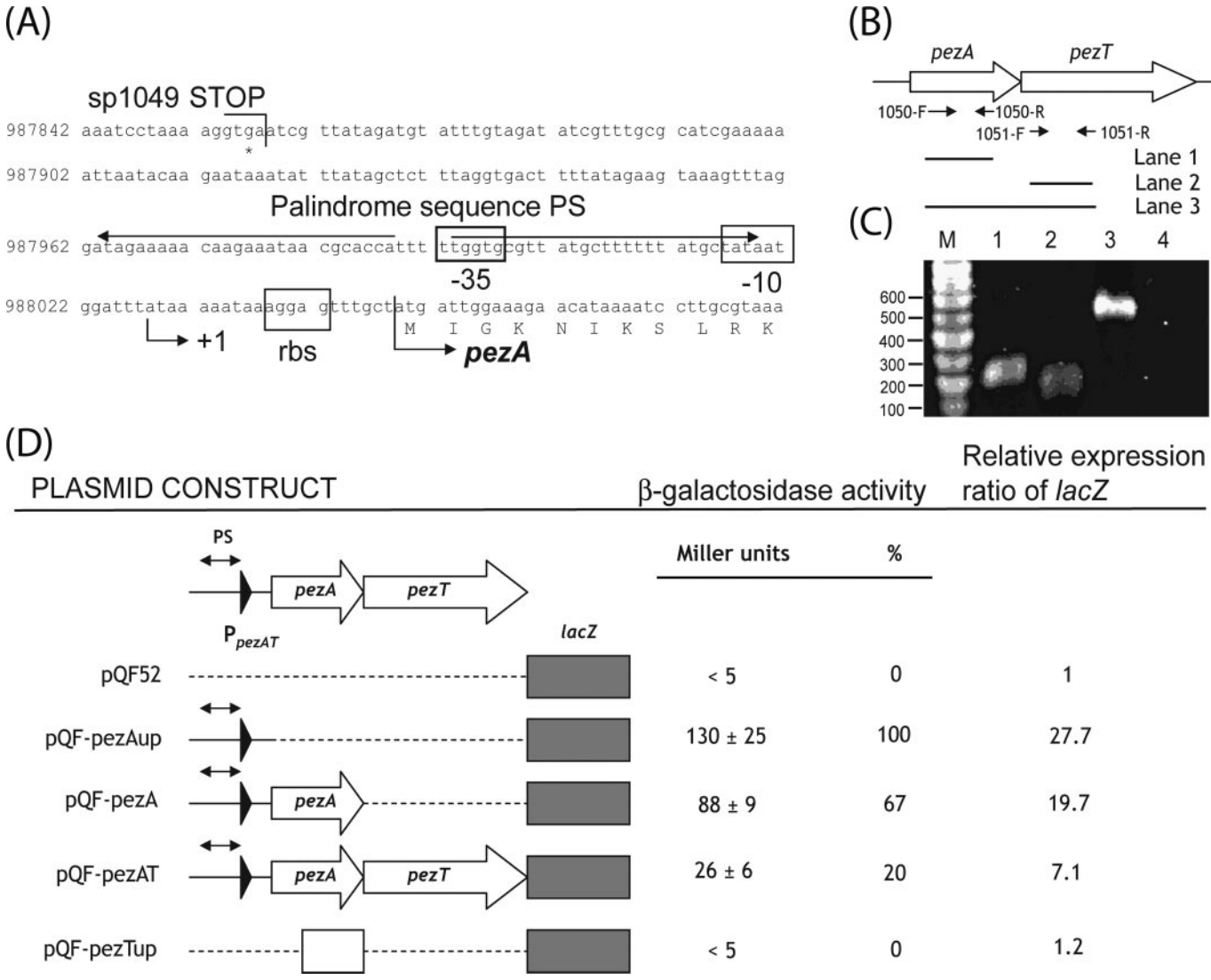

FIGURE 4. **Regulation of the *pezAT* operon.** *A*, nucleotide sequence of the *pezA* upstream region along with the first 33 nt of the *pezA* coding sequence. The −10 and −35 elements of the P_{pezAT} promoter are *boxed* and labeled as is the putative ribosome binding site (designated *rbs*). The palindrome sequence PS is depicted with two divergent *arrows*. The transcript start site as determined by 5′-RACE is labeled as +1. The sequences are numbered as in the data base entry for the *S. pneumoniae* TIGR4 genome sequence (accession no. AE005672). *B*, genetic structure of the *pezAT* operon. *Arrows* depict primers which were used for RT-PCR analyses. *C*, results of RT-PCR. A cDNA fragment spanning the *pezAT* genes was synthesized using as template total RNA prepared from *E. coli* DH5α harboring recombinant plasmid pQF-pezAT. Subsequent PCR reactions utilized oligonucleotide pairs as indicated by the *horizontal bars*; *lane 4* contains the negative control, *i.e.* PCR reaction similar to *lane 3* but without the reverse transcriptase reaction. *Lane M*, DNA molecular weight standard with sizes as indicated in bp. *D*, results of the *lacZ* transcriptional fusions. Schematic diagram of the various *pezAT-lacZ* transcriptional fusion constructs in pQF52 is shown in the *table* with the organization of the *pezAT* genes, its P_{pezAT} promoter and the palindrome sequence PS shown at the *top* of the table. DNA fragments containing combinations of the *pezAT* locus and its upstream region were cloned upstream of the promoterless *lacZ* (depicted as a *gray box*) in plasmid pQF52 resulting in various pQF52-derived recombinant constructs as depicted. Complete open reading frames are indicated as *block arrows*. The β-galactosidase activity levels presented are the average results of at least six independent experiments and are also presented as relative levels (in %). The relative ratios of *lacZ* transcript levels to 16 S rRNA as determined by real-time RT-PCR are also presented.

show that *pezAT* is a bicistronic operon that is transcribed from a σ^{70}-like promoter, P_{pezAT}, the Pribnow box starting 33 bp upstream of the *pezA* start codon.

The PezA Antitoxin Represses Transcription from P_{pezAT} and the PezT Toxin Acts as a Co-repressor—When comparing β-galactosidase activities of *E. coli* DH5α cells harboring pQF-pezAup and pQF-pezA (which encompassed the *pezA* reading frame along with the upstream region), pQF-pezA cells showed an ~33% decrease in activity (Fig. 4*D*) indicating that PezA possibly functions as a repressor. In the pQF-PezAT clone (which spans the upstream fragment and both *pezA* and *pezT* reading frames), β-galactosidase activity levels were down by nearly 80% when compared with levels observed in the pQF-pezAup clone (Fig. 4*D*) implying a much stronger repression in the presence of PezT.

To corroborate the results of the β-galactosidase activity assay and to preclude any possible reduced β-galactosidase activity caused by cloning artifacts, quantitative real-time RT-PCR was carried out to determine the *lacZ* transcript levels in the recombinant *E. coli* DH5α clones. Levels of *lacZ* mRNA were measured relative to 16 S rRNA and presented as ratios of *lacZ* to 16 S rRNA levels as described under "Experimental Procedures." Confirming the results from the β-galactosidase assay, the presence of the *pezA* reading frame in pQF-pezA

resulted in a 29% decrease in *lacZ* transcript levels compared with the levels in pQF-pezAup and this decreased further by nearly 74% in the presence of both the *pezAT* reading frames (Fig. 4*D*). Thus, the PezA antitoxin functions as a transcriptional repressor with the PezT toxin acting as a co-repressor and both proteins are therefore involved in the autoregulation of the *pezAT* operon, like in most other TA systems (1, 2).

To investigate whether the palindromic sequence PS upstream of *pezAT* is indeed the DNA binding site for the PezA and PezT proteins, electrophoretic mobility shift assays (EMSA) were performed. PezA and PezT proteins were expressed and purified as a recombinant N-terminal $(His)_6$-PezA fusion and a $(His)_6$-PezAT protein complex from *E. coli* as described under "Experimental Procedures." The purified PezA protein and the PezAT protein complex were separately incubated with a biotin-labeled 203 bp DNA fragment encompassing PS. Binding reactions were carried out in the presence of 50 ng of poly(dI-dC) as a nonspecific competitor DNA. Results indicated that the majority of labeled DNA was retarded when 750 ng of PezA was included in the binding reaction (Fig. 5*A*). Essentially similar results were obtained with the PezAT complex where the majority of the retardation occurred with 750 ng of the protein complex but initial retardation was detected already at lower PezAT concentrations than with PezA alone (Fig. 5*B*).

The $PezA_2PezT_2$ Architecture—To understand the PezAT system in detail, we solved the crystal structure of the PezAT protein complex at a resolution of 3.2 Å by the single anomalous diffraction (SAD) method. Details describing x-ray diffraction data quality and refinement statistics are given in Table 1, and the Ramachandran statistics are given in Table 2.

The apparent biological unit of the PezAT TA system is a $PezA_2PezT_2$ heterotetramer, which has a dumbbell-shaped arrangement with an internal 2-fold symmetry (Fig. 1*B*). The central core is formed by a stack of two PezA antitoxin polypeptide chains that are sandwiched between two PezT toxin molecules whose dyad axis lies parallel to the PezA homodimer interface. Overall, this molecular arrangement is very similar to the architecture observed in the crystal structure of the epsilon-zeta system (12). Despite the significant sequence homology between the two different antitoxin and toxin polypeptide chains, conserved residues on the surface of the $PezA_2PezT_2$ heterotetramer map only to one region (Fig. 1*C*), which was recently proposed to be the substrate binding side of the putative phosphoryltransferase zeta (12). Further, larger conserved patches that would suggest common modes of interaction of these two different TA systems could not be identified on the surface of the antitoxin-toxin protein complex. The majority of conserved residues is either shielded within the protein complex or is important for establishing a similar fold.

Two heterotetrameric protein complex molecules were located within one asymmetric unit in the crystal. Although the individual PezA and PezT polypeptide chains have a very similar geometry and their r.m.s.d. on Cα positions is at most 0.924 Å for all toxin and 1.225 Å for the antitoxin polypeptide chains, several regions within the polypeptide chain show distinct differences. Consequently, a non-crystallographic restraint could not be used during refinement, because such restraint caused

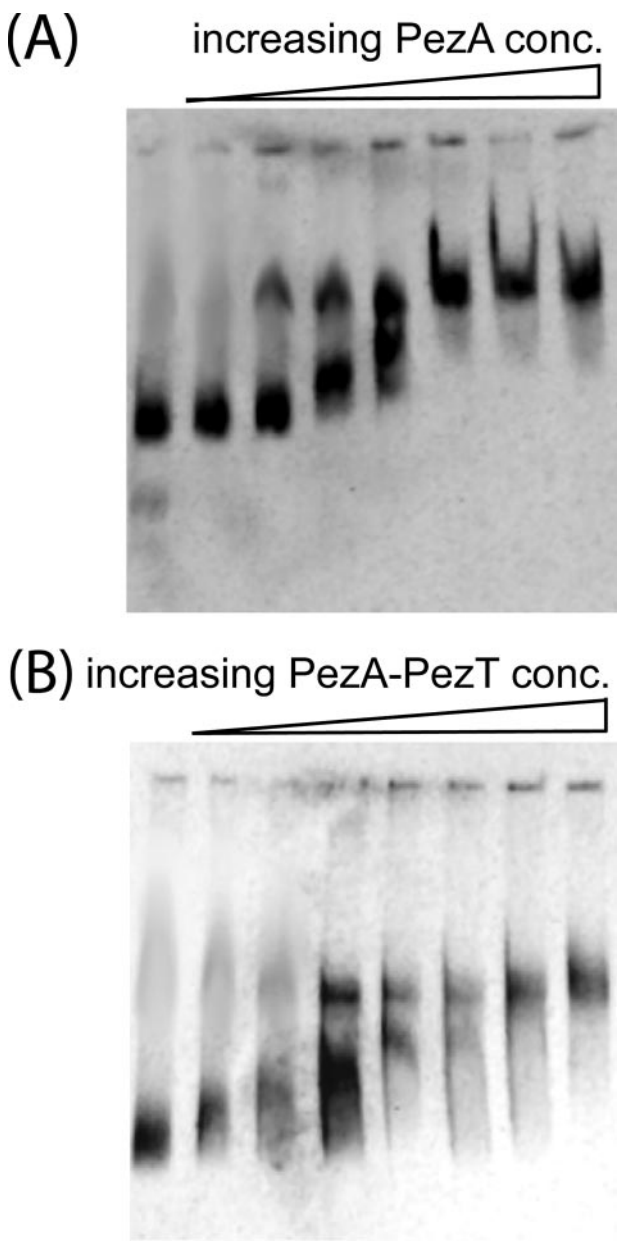

FIGURE 5. **Electrophoretic mobility shift assays demonstrating *in vitro* binding of the purified $(His)_6$-PezA and $(His)_6$-PezA-PezT protein complex with 20 fmol of a 203 bp DNA fragment encompassing the 56-bp imperfect palindrome sequence PS.** Increasing amounts of PezA (*A*) and the PezA-PezT protein complex (*B*) (from *left* to *right*: 0, 50, 150, 300, 500, 750, 850, and 1000 ng) were incubated with the 203-bp DNA fragment that was labeled with biotin. Unbound DNA fragments were separated from nucleoprotein complexes by electrophoresis in 6% polyacrylamide gels.

the R-factors to increase. These differences between individual PezT polypeptide chains can be attributed to crystal contacts to neighboring molecules and probably do not reflect biological relevant conformational states of the proteins. Furthermore, some regions within the polypeptide chain of PezT did not show up in the electron density and the polypeptide chain could not be modeled for the corresponding residues (see Table 2). The major portion of the polypeptide chains that lacks electron density interpretable at atomic resolution resides within the helix-turn-helix (HTH) motif within the N-terminal region of the PezA polypeptide chain. Although we found residual elec-

TABLE 1
X-ray diffraction data and refinement statistics

	Selenomethionyl	Native
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell (*a,b,c* (Å))	81.3, 103.8, 255.0	80.5, 102.9, 254.4
Wavelength (Å)	0.97908	1.007466
Resolution (Å)[a]	20-3.8 (3.9-3.8)	10-3.2 (3.3-3.2)
Unique reflections[a]	40,716 (3075)	34,352 (4708)
Completeness (%)[a]	99.1 (99.4)	95.3 (85.5)
$<I/\sigma(I)>^a$	19.6 (8.1)	20.8 (3.1)
$R_{meas}^{a,b}$	6.7 (22.1)	6.0 (38.7)
Redundancy	6.7	5.7
Refinement		
Non-hydrogen atoms		11093
$R_{cryst}^{a,c}$		21.4 (27.7)
$R_{free}^{a,d}$		27.7 (34.7)
Overall B factor (Å^2)		91.27
r.m.s.d.[e] from ideal geometry		
Bond length (Å)		0.008
Bond angles (°)		1.059

[a] Values in parentheses refer to the highest resolution shell.
[b] $R_{meas} = \Sigma_h[n/(n-1)]^{1/2}\,\Sigma_i|\,I_h - I_{h,i}|/\Sigma_h\Sigma_i\,I_{h,i}$.where I_h is the mean intensity of symmetry-equivalent reflections, and n is the redundancy.
[c] $R_{cryst} = \Sigma_h|F_o - F_c|/\Sigma F_o$ (working set, no σ cut-off applied).
[d] R_{free} is the same as R_{cryst}, but calculated on 5% of the data excluded from refinement.
[e] r.m.s.d. from target geometries.

TABLE 2
Model completeness of the individual polypeptide chains and Ramachandran statistics

PezA						
Chain A	G1[a]	–	G33[a]	L67	–	A158
Chain C				M66	–	A158
Chain E				D64	–	A158
Chain G				M66	–	A158
PezT						
Chain B	I3	–	Q168	H176	–	K253
Chain D	M1	–	Q168	P173	–	L251
Chain F	E2	–	N165	D178	–	K253
Chain H	M1	–	P166	D178	–	K253
Ramachandran statistics						
Most favored region				91.0%		
Additional allowed region				9.0%		
Generously allowed region				0.0%		
Disallowed region				0.0%		

[a] Because of weakly defined electron density, this region was modeled as a polyglycine chain.

tron density that matches a common fold of classical HTH motifs for one PezA polypeptide chain, we were unable to clearly trace the amino acid sequence. In general, for other PezA polypeptide chains, these regions were ill defined and did not allow for unambiguous model building (for detailed information, see Table 2). Because corrections for solvent content failed within these regions, it is likely that this particular disorder problem causes the increased free R-factor (see Table 1). However, we could show that all molecules within the crystals indeed harbor full-length PezA protein. MALDI-TOF MS experiments of dissolved crystals clearly show a peak corresponding to the mass of full-length PezA, but lack any peak that would correspond to a proteolytically truncated protein. Structure factor amplitudes and model coordinates have been deposited at the Protein Data Bank under entry 2P5T.

The PezT Toxin Structure—PezT adopts the fold of NMP-kinase and the chloramphenicol phosphotransferase as had been reported for the zeta toxin (12) (Fig. 1*B*). Furthermore, PezT also contains sequence motifs that are characteristic for ATP/GTP binding proteins such as a Walker A motif or P-loop. In order to verify whether PezT and zeta indeed have the same toxic activity, we mutated residues that were previously shown to be involved in zeta toxicity (12) at equivalent positions in PezT and tested these constructs in cell viability assays. For instance, individual mutations of K45A, R157A, or R170A (Fig. 6) abolished PezT lethality. All three residues are known from related structures to be involved in ATP/GTP binding and hydrolysis (12, 29). Furthermore, an aspartate residue was shown in the latter report to deprotonate the substrate before phosphoryltransfer and mutation of the equivalent residue D66T indeed resulted in a non-toxic variant of PezT. Similar observation of the loss of toxicity was also observed in the zeta toxin (12). Consequently, both toxic proteins seem to use a similar mechanism of ATP/GTP binding, hydrolysis, and phosphoryltransfer.

The effects of the D66T mutation in PezT and the corresponding residue in zeta corroborate that a phosphoryltransfer reaction from ATP/GTP to an unknown substrate is much more likely responsible for toxicity than a simple nucleoside triphosphate hydrolysis. However, based on the epsilon-zeta structure alone, neither the substrate binding site could be experimentally verified nor could any conclusion be derived concerning the true chemical nature of the toxin substrate. Using a combination of structural data analysis with bioinformatics tools, we can now identify residues that are located adjacent to the ATP/GTP binding site and structurally conserved between PezT and zeta (Fig. 6). These residues cluster in a groove that is formed by helix α5 and helix α6. To show that this region is indeed involved in substrate binding, we mutated two conserved threonine residues, T117V and T120V. These two threonine residues project into the groove and are in position to form hydrogen bonds with the substrate. Both mutations impaired PezT toxicity. Whereas T117V resulted in a non-toxic variant that showed a normal growth profile when expressed in *E. coli* BL21, the T120V mutation led to a delay in the growth inhibition, which occurred ~1 h following IPTG induction for overexpression (Fig. 6*B*). The duration in which cellular growth was inhibited in the T120V mutant was similar to that of the wild-type PezT, *i.e.* about 3 h.

The PezA Three Helix Bundle—On a primary structural level, the major difference between the PezAT and the epsilon/zeta TA systems is found between the two different antitoxin polypeptide chains. Interestingly, the C-terminal region of the PezA polypeptide chain folds into a three helix bundle similar to that of epsilon. Strikingly, residues that form the hydrophobic core of the three helix bundle are conserved, whereas residues involved in PezA homodimer formation are only moderately conserved (Fig. 1*A*). Surface-exposed residues of PezA are much less conserved (Fig. 1*D*). The lack of conservation of the antitoxin at the grip of the dumbbell-shaped molecule suggests that this region might not be important for PezA functionality.

Inactivation of PezT by PezA—In the inactive $PezA_2PezT_2$ heterotetrameric protein complex, the nucleotide binding site of PezT is covered by two helices of PezA (helix α1 and α2, Fig. 1). Strikingly, three evolutionarily invariant residues of the antitoxin PezA (Tyr^{73}, Leu^{77}, and Glu^{80}, see supplemental Fig. S1) protrude with their side chains into the putative nucleotide

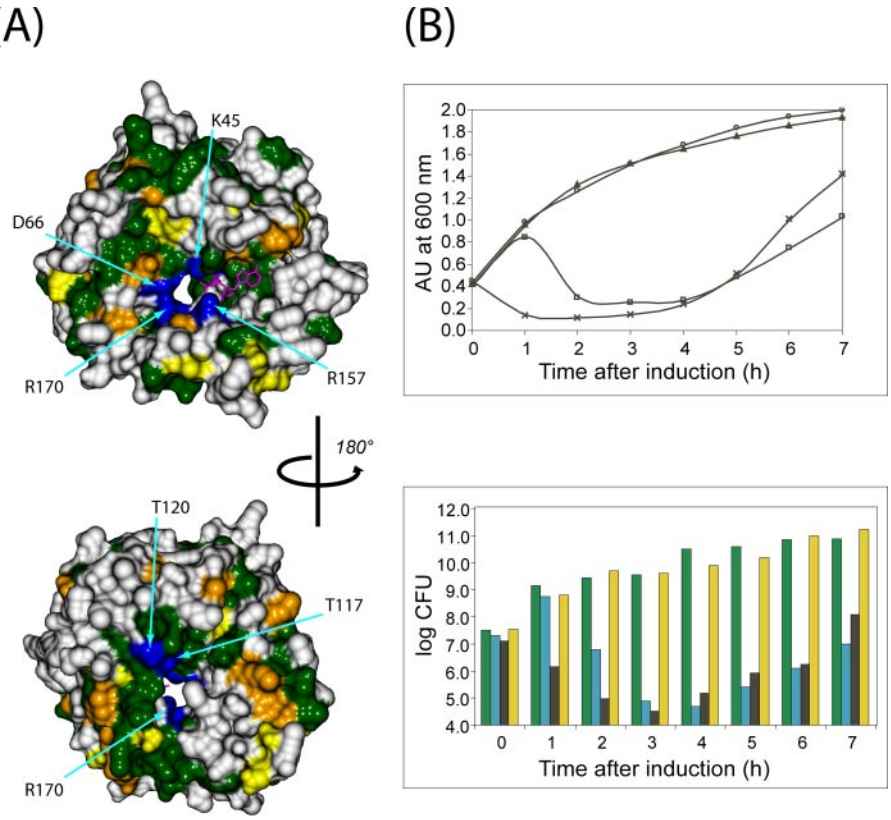

FIGURE 6. *A*, surface representation of PezT alone, colored according to Fig. 1*A*. Residues that are identical with zeta and showed to impair PezT toxicity are colored in *blue*. An ATP molecule colored in *magenta* was modeled according to the position found in the chloramphenicol-phosphotransferase (29). *B*, effect of PezT mutations on the growth of *E. coli* BL21(DE3)/pLysS carrying recombinant pET11a plasmids as determined by A_{600} and log CFU. Growth profile of wild-type PezT is shown as (×) for A_{600} values (on *top*) and *black bars* in log CFU (*bottom*). Growth profiles of the following mutants following IPTG induction are shown: T117V (▲; *green bars*), T120V (□; *blue bars*), and K45A (○; *yellow bars*). The R157A, R170A, and D66T mutations yielded nontoxic growth profiles similar to those of K45A and T117V and are not presented in the figure.

binding site of PezT and thereby prevent nucleotide binding in a manner similar to the epsilon/zeta system. Tyr^{73} and Leu^{77} stick their hydrophobic moieties into the active site of PezT and block this by steric hindrance. The adenine/guanine moiety in the nucleotide-bound state is mimicked by the aromatic ring of residue Tyr^{73} in PezA that is stabilized by π-stacking interaction with the guanidinium group of Arg^{205} in PezT. Residue Arg^{157} in PezT, which we showed to be involved in toxicity, is fixed by a salt bridge established to the side chain of Glu^{80} of PezA (see supplemental Fig. S1). Upon dissociation of PezA from PezT, the putative nucleotide binding site becomes accessible. A detailed analysis of the PezA-PezT interaction revealed that the majority of residues involved are not conserved (Fig. 1*A*). Only residues that play an essential role in toxin inactivation are identical between PezT and zeta. This suggests that the only evolutionary constraint is to abolish the harmful activity of the toxin, but residues involved in protein-protein interaction can, however, diverge.

The Helix-Turn-Helix Motif in PezA—Although we do not have an atomic model for the HTH-motif in the N terminus of PezA, the presented structure allows us to derive a hypothesis about the orientation of this domain. The HTH motif identified in the electron density is located in proximity of the dyad axis of the $PezA_2PezT_2$ heterotetramer. In the absence of operator DNA, the HTH domains apparently do not form strong interactions with each other or PezT and are therefore flexible. Assuming that the overall architecture of the $PezA_2PezT_2$ complex remains preserved upon binding to the operator site, the two HTH domains must obey the overall 2-fold symmetry of the complex and should, in this fashion, bind to a palindromic sequence such as found for the P_{pezAT} promoter.

DISCUSSION

In this study, we showed that the chromosomally encoded *pezAT* bicistronic operon of the human pathogen *S. pneumoniae* is a functional TA locus with *pezA* encoding the cognate antitoxin to the toxic product of *pezT*. Like most TA systems, the *pezAT* locus is autoregulated with the PezA antitoxin functioning as a transcriptional repressor by binding to a palindromic sequence that overlaps the P_{pezAT} promoter and the PezT toxin acts as a co-repressor. Our *in vivo* and *in vitro* data were further corroborated by the crystal structure, where the N-terminal domain of PezA adopts the fold of a DNA binding domain with a HTH motif belonging to the Xre and Cro/CI family. Thereby, the PezAT system differs significantly from the omega-epsilon-zeta locus of the *S. pyogenes* plasmid pSM19035, where omega, as a third component, is the transcriptional regulator for the epsilon-zeta bicistronic operon. Epsilon functions solely as the antitoxin for zeta (11, 28, 30) and omega has been shown to bind to operator DNA via a ribbon-helix-helix motif that belongs to the MetJ/Arc superfamily of transcriptional repressors (31, 32). Apparently, the regulatory domains for transcription of the PezAT system and the omega-epsilon-zeta systems are evolutionarily not related and both systems have different transcriptional regulatory mechanisms. Similar to most TA systems, epsilon was shown to have a shorter half-life than zeta (13). However, the regulatory mechanism via which omega senses a decreasing concentration of epsilon for proper functioning of the TA system remains elusive. In case of the PezAT system, this regulation becomes immediately evident, because the transcriptional repressor and the toxin-neutralizing activity can be found within a single polypeptide chain in PezA. A decreasing amount of antitoxin would be compensated by its transcriptional repressor function and thereby PezA ensures that the toxin PezT does not become free and active.

The three-dimensional structure of the PezT toxin is very similar to zeta and mutations in equivalent residues of PezT

which have previously been shown to be involved in zeta toxicity also abolished PezT activity. It is therefore very likely that PezT and zeta have similar mechanisms for toxicity. Apparently, both use ATP/GTP for phosphoryltransfer to an unknown substrate. However, we could now clearly demonstrate by site-directed mutagenesis that a recently proposed substrate binding site of zeta (12) is indeed involved in toxicity. Mutation of two threonine residues that are located distal from the nucleotide binding site to valine abolished toxicity. A common mechanism is further supported by the growth profile of *E. coli* cells overexpressing PezT. Inhibition of growth was only temporary and restoration of growth occurred about 3 h after induction of overexpression, similar to the phenomenon observed in cells that express zeta. Microscopic examination of cultures during this growth inhibition period revealed massive filament formation, but no induction of the SOS system was detected (28). Preliminary examination of PezT-inhibited cultures also indicated similar morphological traits (data not shown). Recent studies showed that induction of zeta expression in *E. coli* CC118 cells led to the inhibition of DNA, RNA and protein synthesis. However, it was hypothesized that zeta exerts a pleiotropic effect on the physiological state of the cells (33). The target and mechanism of PezT/zeta lethality remains to be elucidated and is the subject of our on-going investigations.

Although TA systems have been found to be widely distributed especially among free-living prokaryotes (9), the targets of many toxins have yet to be identified. The toxins of three well-studied chromosomal TA loci, namely MazEF, RelBE, and YefM-YoeB, were lately shown to be site-specific endoribonucleases (4, 7, 34). The recently elucidated structure of the FitAB complex from *Neisseria gonorrhoeae* indicated that the FitB toxin is composed of a PIN (PilT N terminus) domain, which is different from that of the RelE and YoeB toxins. However, FitB did not display nuclease activity against tested PIN substrates (35). The structures of PezT and zeta and subsequent mutagenesis experiments strongly infer the role of a phosphoryltransferase reaction in their toxicity. Notably, a recent report on the HipBA TA locus in *E. coli* indicated that the HipA toxin is a serine kinase that autophosphorylates itself in the presence of ATP (36). However, HipA is a member of the phosphotidylinositol 3/4 kinase superfamily whose structure is not related to the NMP kinase fold of PezT and zeta. A common mechanism can therefore be excluded.

Expression of PezT and zeta in *E. coli* led to temporary inhibition in cell growth and is therefore seen as triggering cell stasis rather than cell death. However, expression of zeta in the Gram-positive *B. subtilis* was reported to be bactericidal (28). Consequently, there is a need to investigate the effects of PezT expression in Gram-positives such as *B. subtilis* and its native host, *S. pneumoniae*, to find out if PezT has a bactericidal effect in these organisms as described for zeta.

An earlier report had shown that the *pezAT* locus is part of a 27-kb pathogenicity island termed pneumococcal pathogenicity island 1 (PPI1), which also harbors the iron uptake locus *piaABC* found to be required for full *S. pneumoniae* virulence in mice (37). A similar location of TA loci on mobile genetic elements has been previously reported (2, 9) for example, the presence of 13 TA loci in a 125-kb superintegron in *Vibrio cholerae* (9, 38). It was postulated that the presence of TA loci may contribute to the stability of these mobile elements similar to how plasmid-encoded TA loci function to stabilize plasmid replicons (2). In contrast to this, deletions were discovered in the genomes of many *S. pneumoniae* strains containing variable number of genes within PPI1 and 33% of the isolates examined did not contain *sp1051* (*i.e. pezT*) (39). Intriguingly, it was demonstrated that disruption of *pezT* impaired virulence in mouse models but showed no growth defect in laboratory broth, serum or blood. Although the function of the *pezAT* locus was not clear, the authors concluded that since *pezT* is missing in a proportion of strains, it is unlikely to be an essential gene for virulence, but rather may modulate the virulence of strains that carry it (39). A similar linkage between toxins of TA systems and virulence was postulated for the FitAB system of *N. gonorrhoeae*. *N. gonorrhoeae* that lacks *fitAB* transit a polarized epithelial monolayer more quickly than the wild-type parent and has an accelerated rate of intracellular replication (35). The potential link between TA systems and virulence highlights the increasing importance for the study of TA systems in pathogenic bacteria where they have also been implicated in multidrug tolerance through their ability to mediate persistence.

Acknowledgments—We thank W. Saenger, F. Alings, Y. Groemping, E. Hartmann, and J. Reinstein for helpful discussions and encouragement and I. Schlichting for support. We are grateful to J. Geib, K. F. Ng, M. K. Tan, and C. K. Lim for excellent technical support, J. H. Siong for help with real-time PCR analyses, I. Vetter for support of the crystallographic software, and the staff at the beamline X10SA, Swiss Light Source (Villigen, Switzerland) for excellent support. We also thank F. Jamal and Y. F. Ngeow for providing clinical isolates of S. pneumoniae.

REFERENCES

1. Hayes, F. (2003) *Science* **301,** 1496–1499
2. Gerdes, K., Christensen, S. K., and Lobner-Olesen, A. (2005) *Nat. Rev. Microbiol.* **3,** 371–382
3. Aizenman, E., Engelberg-Kulka, H., and Glaser, G. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 6059–6063
4. Zhang, Y., Zhang, J., Hoeflich, K. P., Ikura, M., Qing, G., and Inouye, M. (2003) *Mol. Cell* **12,** 913–923
5. Zhu, L., Zhang, Y., Teh, J. S., Zhang, J., Connell, N., Rubin, H., and Inouye, M. (2006) *J. Biol. Chem.* **281,** 18638–18643
6. Pellegrini, O., Mathy, N., Gogos, A., Shapiro, L., and Condon, C. (2005) *Mol. Microbiol.* **56,** 1139–1148
7. Christensen, S. K., and Gerdes, K. (2003) *Mol. Microbiol.* **48,** 1389–1400
8. Pedersen, K., Christensen, S. K., and Gerdes, K. (2002) *Mol. Microbiol.* **45,** 501–510
9. Pandey, D. P., and Gerdes, K. (2005) *Nucleic Acids Res.* **33,** 966–976
10. Ceglowski, P., Boitsov, A., Chai, S., and Alonso, J. C. (1993) *Gene* (*Amst.*) **136,** 1–12
11. Ceglowski, P., Boitsov, A., Karamyan, N., Chai, S., and Alonso, J. C. (1993) *Mol. Gen. Genet.* **241,** 579–585
12. Meinhart, A., Alonso, J. C., Strater, N., and Saenger, W. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 1661–1666
13. Camacho, A. G., Misselwitz, R., Behlke, J., Ayora, S., Welfle, K., Meinhart, A., Lara, B., Saenger, W., Welfle, H., and Alonso, J. C. (2002) *Biol. Chem.* **383,** 1701–1713
14. Nieto, C., Pellicer, T., Balsa, D., Christensen, S. K., Gerdes, K., and Espinosa, M. (2006) *Mol. Microbiol.* **59,** 1280–1296
15. Nieto, C., Cherny, I., Khoo, S. K., de Lacoba, M. G., Chan, W. T., Yeo, C. C.,

Gazit, E., and Espinosa, M. (2007) *J. Bacteriol.* **189,** 1266–1278
16. Meinhart, A., Silberzahn, T., and Cramer, P. (2003) *J. Biol. Chem.* **278,** 15917–15921
17. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) *J. Mol. Biol.* **229,** 105–124
18. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26,** 795–800
19. Schneider, T. R., and Sheldrick, G. M. (2002) *Acta Crystallogr. Sect. D* **58,** 1772–1779
20. Sheldrick, G. M. (2002) *Z. Kristallog.* **217,** 644–650
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
22. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D* **54,** 905–921
23. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr., Sect. D* **53,** 240–255
24. Laskowski, R. A., Mac Arthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
25. McLean, B. W., Wiseman, S. L., and Kropinski, A. M. (1997) *Can. J. Microbiol.* **43,** 981–985
26. Miller, J. H. (1972) *Experiments in Molecular Genetics,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
27. Pfaffl, M. W. (2001) *Nucleic Acids Res.* **29,** e45
28. Zielenkiewicz, U., and Ceglowski, P. (2005) *J. Bacteriol.* **187,** 6094–6105
29. Izard, T., and Ellis, J. (2000) *EMBO J.* **19,** 2690–2700
30. de la Hoz, A. B., Ayora, S., Sitkiewicz, I., Fernandez, S., Pankiewicz, R., Alonso, J. C., and Ceglowski, P. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 728–733
31. Murayama, K., Orth, P., de la Hoz, A. B., Alonso, J. C., and Saenger, W. (2001) *J. Mol. Biol.* **314,** 789–796
32. Weihofen, W. A., Cicek, A., Pratto, F., Alonso, J. C., and Saenger, W. (2006) *Nucleic Acids Res.* **34,** 1450–1458
33. Lioy, V. S., Martin, M. T., Camacho, A. G., Lurz, R., Antelmann, H., Hecker, M., Hitchin, E., Ridge, Y., Wells, J. M., and Alonso, J. C. (2006) *Microbiology* **152,** 2365–2379
34. Kamada, K., and Hanaoka, F. (2005) *Mol. Cell* **19,** 497–509
35. Mattison, K., Wilbur, J. S., So, M., and Brennan, R. G. (2006) *J. Biol. Chem.* **281,** 37942–37951
36. Correia, F. F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B. L., Makarova, K., Koonin, E. V., and Lewis, K. (2006) *J. Bacteriol.* **188,** 8360–8367
37. Brown, J. S., Gilliland, S. M., and Holden, D. W. (2001) *Mol. Microbiol.* **40,** 572–585
38. Christensen-Dalsgaard, M., and Gerdes, K. (2006) *Mol. Microbiol.* **62,** 397–411
39. Brown, J. S., Gilliland, S. M., Spratt, B. G., and Holden, D. W. (2004) *Infect. Immun.* **72,** 1587–1593