Substrate Recognition and Activity Regulation of the *Escherichia coli* mRNA Endonuclease MazF*

Received for publication, January 15, 2016, and in revised form, March 9, 2016 Published, JBC Papers in Press, March 29, 2016, DOI 10.1074/jbc.M116.715912

Valentina Zorzini‡§1,2, Andrej Mernik‡, Jurij Lah‡, Yann G. J. Sterckx‡, Natalie De Jonge‡, Abel Garcia-Pino§11, Henri De Greve15, Wim Versées15, and Remy Loris15

From the ‡Structural Biology Brussels, Department of Biotechnology, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium, the §Structural Biology Research Center, VIB, Pleinlaan 2, B-1050 Brussels, Belgium, the ¶Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, 1000 Ljubljana, Slovenia, and the ¶¶Biologie Structurale et Biophysique, Université Libre de Bruxelles, Rue des Professeurs Jeener et Brachet 12, 6041 B-Gosselies, Belgium

*MazF (EcMazF) is the archetypal of a large family of ribonucleases involved in bacterial stress response. The crystal structure of EcMazF in complex with a 7-nucleotide substrate mimic explains the relaxed substrate specificity of the *E. coli* enzyme relative to its *Bacillus subtilis* counterpart and provides a framework for rationalizing specificity in this enzyme family. In contrast to a conserved mode of substrate recognition and a conserved active site, regulation of enzymatic activity by the antitoxin *EcMazE* diverges from its *B. subtilis* homolog. Central in this regulation is an *EcMazE*-induced double conformational change as follows: a rearrangement of a crucial active site loop and a relative rotation of the two monomers in the EcMazF dimer. Both are induced by the C-terminal residues Asp-78–Trp-82 of *EcMazE*, which are also responsible for strong negative cooperativity in EcMazE-EcMazF binding. This situation shows unexpected parallels to the regulation of the F-plasmid CcdB activity by CcdA and further supports a common ancestor despite the different activities of the MazF and CcdB toxins. In addition, we pinpoint the origin of the lack of activity of the E24A point mutant of EcMazF in its inability to support the substrate binding-competent conformation of EcMazF.

Bacterial populations are capable of overcoming periods of harsh conditions that are normally lethal for metabolically active cells through the stochastic generation of persisters (1, 2). These phenotypically distinct cells encompass only a very small fraction of the total population and are in a metabolically dormant state that makes them tolerant to a wide variety of antibiotics without having acquired resistance. Discovered almost 70 years ago, the molecular mechanisms behind persistence are becoming clear only recently (3–6). The activities of toxin-antitoxin (TA)5 modules have been tightly linked with the establishment of the persister phenotype (7–9). TA modules form a diverse family of two-component systems that can block distinct components of basic metabolism. In all cases, the “toxin” acts upon a specific basic physiological activity such as translation, cell wall synthesis, or transcription and replication, while the corresponding “antitoxin” harnesses this activity (10, 11). Very often, the antitoxin is modular and contains a folded DNA-binding/dimerization domain linked to an intrinsically disordered toxin-neutralization segment (12). The nature of the DNA-binding domain can vary and is not strictly correlated with the nature of the toxin (13). In contrast, the toxin-neutralizing domain folds upon binding to the toxin, and the latter interaction seems much better linked to the corresponding toxin family (12).

Based on the nature and activity of the toxin, TA modules are classified in a growing number of families (13). One of the most prevalent TA modules is the *mazEF* family, of which the toxin MazF is an mRNA interferase. Activation of MazF shifts translation toward a specific subset of genes during times of stress (14–16). It cleaves mRNA and in some cases rRNA at specific sequences (17–21). In *Escherichia coli*, it degrades certain mRNAs and truncates others to generate leaderless mRNAs. The latter are specifically transcribed by MazF-modified ribosomes (20).

MazF proteins belong to a large family that is ubiquitous among bacteria and are well studied with respect to their physiological effects and their cutting specificity. *E. coli* MazF was shown to specifically cut at the 5′ end of A in UAC sequences (19). Other family members typically have their own specific RNA recognition sequence that can vary in length, but often it contains the ACA consensus sequence. For example, *Bacillus subtilis* MazF (BsMazF) recognizes the penta-nucleotide sequences UACAU (22). The specificity of ChpBK (the second MazF protein encoded on the *E. coli* chromosome) is broadened to XACY where X is preferentially U, but also A or G, and
Y is U, A, or G (23). In addition, not all MazF recognition sites are cleaved in each mRNA, indicating that RNA secondary structure also plays a major role in directing MazF activity (17, 23). The MazF ribonuclease activity occurs in vitro in the absence of any co-factor (24, 25). In vivo, MazF is able to operate independently of translation, but its cleavage efficiency for specific RNAs is enhanced dramatically during translation, probably by destabilization of the mRNA secondary structure (26).

Currently, a structure of MazF in complex with a substrate mimic is only available for a homolog from B. subtilis (BsMazF, also referred to as YdcE− (27)). However, sequence variation within the MazF family of mRNA interferase is high (e.g. only 23% sequence identity between EcMazF and BsMazF), and neither a structural basis for the substrate specificity of the archetypal E. coli enzyme nor a unifying enzymatic mechanism for the MazF family has yet been proposed. In this study, we describe the structures of wild type and E24A mutant of E. coli MazF in their free forms and in complex with a 7-nucleotide substrate mimic or a peptide corresponding to the 15 C-terminal residues of E. coli MazE, respectively. These structures together with activity and ITC data provide novel insights into substrate recognition and regulation of E. coli MazF and further support an evolutionary relationship and a common regulatory mechanism for ccdAB and mazEF.

Experimental Procedures

Protein Production—Peptides corresponding to residues 54–77 (EcMazE(54–77)), 70–82 (EcMazE(70–82)), 68–82 (EcMazE(68–82)), and 50–82 (EcMazE(50–82)) were obtained from Bio-Synthesis (Lewisville, TX). Expression and purification of wild type EcMazF were performed as described (28). The E24A point mutant of E. coli MazF (EcMazEF24A) was generated by amplification and mutation of the wild type E. coli mazF gene. This gene is flanked by NdeI and XhoI restriction sites, which allowed its insertion into the pET22b vector. The pET22b-mazF E24A plasmid encoding a His tag at the C terminus of EcMazEF24A was transformed into E. coli BL21(DE3) competent cells. Expression in LB medium was induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside when the absorbance at 600 nm reached 0.6, and the culture was incubated overnight at 37 °C, 200 rpm. The cells were harvested by centrifugation (30 min, 4,000 × g, 4 °C) and resuspended in lysis buffer, 20 mM Tris-HCl, 150 mM NaCl, 10 mM imidazole, pH 7.0, 0.1 mg/ml 4-(2-aminoethyl)benzenesulfonyl fluoride, and 1 µg/ml leupeptin. After breaking the cells and passing them twice through a French press (1,000–1,200 bars, 12,000 p.s.i.), the suspension was again centrifuged (30 min, 15,000 g, 4 °C), and the supernatant was filtered through a 0.45-µm filter. The recombinant protein was affinity-purified using nickel-NTA agarose. The purity of the fractions was controlled on an SDS-PAGE, and protein concentrations were determined using ultraviolet spectrophotometry at 280 nm. The protein was flash-frozen using liquid nitrogen for storage at −80 °C.

Crystal Structure Determination—Protein solutions were checked for monodispersity and lack of aggregation by small angle x-ray scatter on a Rigaku BioSAXS 2000. Crystallization conditions for wild type EcMazF and its complex with the substrate analog d(AUACAUA) were screened at 293 K by the sitting-drop vapor diffusion method using a Phoenix robot (Art Robbins Instruments, Asbach, Germany) and manually using the hanging-drop vapor diffusion method. Drops containing 0.1 µl of complex (molar ratio protein/DNA equal to 1:6) or 1 µl of protein and 0.1 or 1 µl of precipitant solution were equilibrated against 70 or 200 µl of precipitant solution. Various commercial screens from Hampton Research I and II, Morpheus I and II, JCSG-plus, and ProPlex were used for screening. The conditions yielding diffraction-quality crystals are reported in Table 1 and supplemental Table S1. All the datasets were collected at SOLEIL beamline PROXIMA-2a (Saint-Aubin, France) using an ADSC Quantum 315 CCD detector, and data for the ligand-free enzyme were collected at SOLEIL beamline PROXIMA-1 using a Dectris PILATUS 6M detector. All data were indexed, integrated, and scaled using XDS (Table 1 and supplemental Table S1) (27).

Crystallization conditions for EcMazEF24A free form and its complex with EcMazE(68–82) were screened at 293 K by the sitting-drop vapor diffusion method using a Phoenix robot (Art Robbins Instruments) and manually using the hanging-drop vapor diffusion method using the Hampton Research Crystal Screens I and II, Morpheus I and II, JCSG-plus, and ProPlex. Drops containing 1 or 0.1 µl of complex (molar ratio protein/peptide equal to 1:6) and 1 or 0.1 µl precipitant solution were equilibrated against 200 or 70 µl of precipitant solution, respectively. The conditions yielding diffraction-quality crystals are reported in Table 1 and supplemental Table S1. Data for the E.cMazE(68–82) complex were collected at SOLEIL beamline PROXIMA-2a (Saint-Aubin, France) using a using a Dectris PILATUS 6M detector and indexed, integrated, and scaled using XDS (Table 1) (29). Data for the free mutant were collected at ESRF beamline ID29 (Grenoble, France) using an ADSC Q315 CCD detector. These data were indexed, integrated, and scaled using the HKL suite (30) (supplemental Table S1).

Intensities from the merged and scaled diffraction data were converted to structure factor amplitudes using the CCP4 program TRUNCATE (31). All structures were determined by molecular replacement with Phaser (32) using the structure of the EcMazF dimer present in the crystal structure of the EcMaz-EcMazF heterohexameric complex (PDB entry 1UB4 (33)) as search model. All structures were refined against an intensity-based maximum likelihood target using Phoenix (34) without σ cutoff. After an initial rigid body refinement, a Cartesian simulated annealing protocol (starting at a Boltzmann temperature of 5,000 K) was applied to uncouple Rwork and Rfree. Rounds of positional and isotropic B-factor refinements interspersed by manual rebuilding using Coot (35) were performed. Multiple refinement cycles combined with manual inspection of the stereochemistry of the models via Ramachan-
RNA Binding and Activity Regulation of MazF

dran plot significantly decreased the R-factors for all the structures reported. The R-factors were further decreased by including water molecules where relevant and by refining TLS parameters (one TLS group per chain was used).

**RNase Activity Assay**—Bacteriophage MS2 genomic RNA was obtained from Roche Applied Science. Reaction mixtures contained 0.25 μl of RNA (0.8 μg/μl in 10 mM Tris-HCl, pH 7.0, 1 mM EDTA) and 1.25, 2.5, or 5 μl of EcMazF (either wild type or the E24A mutant 4 μM in 20 mM Tris-HCl, pH 7.0, 150 mM NaCl). The volumes of the reaction mixtures were adjusted with buffer (20 mM Tris-HCl, pH 7.0, 150 mM NaCl). Prior to the titrations, the samples were filtrated with 0.22-μm filters. Solutions of EcMazF (85 μM) were titrated into 6 μM solution of RNA/DNA fragments at 305 K.

To exactly match buffer composition, the single strand RNA/DNA fragments and EcMazF (wild type or and E24A mutant) were dialyzed overnight against 2 liters of 50 mM phosphate buffer at pH 7.0 and 150 mM NaCl. Prior to the titrations, the samples were filtrated with 0.22-μm filters. Solutions of EcMazF (85 μM) were titrated into 6 μM solution of RNA/DNA fragments at 305 K.

**Isothermal Titration Calorimetry**—ITC experiments were carried out on a MicroCal iTC200 system (GE Healthcare). Investigation of the wild type and EcMazF*E24A*. RNA binding activity was performed using modified single strand RNAs (RNA1, 5′-AGAUAdUACAUAUGAA-3′; RNA2, 5′-GCUCC-dUACAUUGUCAG-3′; RNA3, 5′-UUGGcdAAUUCAUAAUC-AAU-3′) with the (potential) scissile nucleotide being replaced by its 3′-deoxy analog and obtained from GenScript.

To exactly match buffer composition, the single strand RNA/DNA fragments and EcMazF (wild type or and E24A mutant) were dialyzed overnight against 2 liters of 50 mM phosphate buffer at pH 7.0 and 150 mM NaCl. Prior to the titrations, the samples were filtrated with 0.22-μm filters. Solutions of EcMazF (85 μM) were titrated into 6 μM solution of RNA/DNA fragments at 305 K.

Toxin-antitoxin binding was assessed at 305 K by titrating MazE(50 – 82) (100 μM) or MazE(54 – 77) (640 μM) into wild type MazF and MazF*E24A* mutant (4 or 16 μM). All the samples were previously dialyzed against 2 liters of 50 mM phosphate buffer, pH 7.0, 150 mM NaCl, and 1 mM EDTA and filtrated with 0.22-μm filters.

Data analysis was performed with MicroCal Origin software accompanying the ITC instrument. The binding affinity (Kp) and change in enthalpy (ΔH) associated with the binding events were calculated after fitting the dataset assuming two sequential binding sites. The ITC curve for MazF*E24A* titrated with MazE(68 – 82) was fitted assuming an n = 1.

**Results**

**Substrate Binding Folds an Otherwise Disordered Substrate-binding Loop of EcMazF**—The structure of wild type *E. coli* MazF (EcMazF) was determined in three crystal forms, encompassing five crystallographically independent dimers (*supplemental Table S1*). This structural ensemble shows a well conserved monomeric fold consisting of a five-stranded β-sheet packed upon a C-terminal α-helix (Fig. 1). Most loop regions are well structured, with the exception of loops β1-β2 (residues Asp-16 – Arg-29) and β4-β5 (residues Leu-64 – Gly-71) (Fig. 1C). Loop β1-β2 is partially unstructured in 8 out of 10 crystallographic independent chains and adopts two conformations in the remaining two chains (*supplemental Fig. S1*). Loop β4-β5 is observed in two distinct conformations in chains A and B of crystal form II, whereas the conformations observed in the remaining eight monomers correspond to an ensemble of closely related conformations.

The crystal structure of EcMazF in complex with the substrate-mimicking DNA sequence d(AUACAUA) (PDB entry 5CR2) shows clear electron density for the d(U2A3C4A5U6A7) pentanucleotide sequence, whereas in both other monomers, only a 5′pdU3p′ unit can be identified (Fig. 2A). Loops β1-β2 and
β4-β5 are fully ordered and adopt identical conformations in all three monomers. The conformation adopted by β4-β5 resembles most but is not identical to the conformation of this loop adopted by molecule A in crystal for II of the ligand-free structures. In contrast, the conformation adopted by β1-β2 in the substrate-bound structure does not match any of the conformations seen in the ligand-free structures (Fig. 1C and supplemental Fig. S1).

Loops β1-β2 and β4-β5 together with loop β3-β4 (Thr-52–Phe-60) and the short α-helix α1 (Phe-37–Thr-43) from the adjacent EcMazF monomer constitute the substrate recognition site (Fig. 2D). Contacts between the substrate mimic and both MazF proteins are represented schematically in Fig. 2F. With the exception of U2, all 2'-OHs, if they would have been present (as in the real RNA substrate), would point outward to the solvent and therefore do not seem to be required for interaction with the protein. The recognition site can be divided into two distinct regions as follows: the one where dU2 is located, which we will call the upstream-binding site, and the one that accommodates d(A3C4A5U6), which will be referred to as the downstream-binding groove (Fig. 2D).

**Upstream Subsite in MazF Helps Define Substrate Specificity in This Family of RNases**—Several of the MazF proteins studied so far preferably cleave at the 3' side of a uridine. In our structure, the corresponding uracil base (dU2) is buried in a crevice on the surface of the protein similar to the equivalent uridine in the BsMazF substrate complex, the only other family member for which a structure with a substrate analog is available (Fig. 2, B and E) (27). The uracil base is sandwiched between mostly aliphatic groups from side chains (Fig. 3A). Of these, the most prominent is the side chain of Pro-30, which is conserved within the MazF family. In contrast, other contacting residues are not conserved within the MazF family. The width and hydrophobic nature of the crevice holding the uracil base are similar in both EcMazF and BsMazF.

The crevice in EcMazF is markedly deeper than in BsMazF, and a prominent cavity is observed in the EcMazF complex that provides sufficient space for this subsite to accommodate a purine (Fig. 2, D and E). The latter would not be possible for BsMazF, and this feature thus explains the less strict specificity of EcMazF compared with BsMazF. Indeed, although BsMazF cuts strictly after uracil, EcMazF can also cut after guanine in GACA as well as after the first adenine in UACA and GACA indicating that the upstream subsite can accommodate adenine and guanine as well as uracil (18, 24). Cleavage after cytosine is not observed and can be explained in our crystal structure by the hydrogen bond between O4 of uracil and the backbone nitrogen of Arg-69. Binding of a cytosine base in an analogous manner would be hindered by the presence of its NH2 group. Additional contacts with the ribose-phosphate moiety of dU2 include a hydrogen bond from the guanidinium group of Arg-69 to the 5'-phosphate group of dU2 (absent in BsMazF) and van der Waals contacts between Trp-14 and the ribose moiety of dU2.

**ACA Motif of the Substrate Mimic Stacks into a Broad Surface Groove Formed by Residues from Both Monomers**—The bases of dA4 up to dU6 form a planar stacking arrangement and dock into a broad groove on the protein surface (Figs. 2D and 3B). A series of four main chain and three side chain hydrogen bonds with the bases of the d(A3C4A5) sequence are observed that are common in the EcMazF and BsMazF complexes and define substrate specificity. In the complex of BsMazF, two additional hydrogen bonds are observed between dA3 and the side chains of Glu-78 and Gln-79 that are not present in the

---

**TABLE 1**

| Table 1: Crystallization, data collection, and refinement |
|----------------------------------------------------------|
| **MazF** | **Maz** | **Maz** |
| d(AUACAUA) | d(AUACAUA) | d(AUACAUA) |
| Crystallization | Protein solution | 18 mg/ml (0.7 mM) EcMazF + 4.2 mM oligonucleotide in 50 mM phosphate, pH 7.0, 150 mM NaCl | 10 mg/ml (0.4 mM) EcMazE24A + 2.5 mM MazE (68–82) in 20 mM Tris-HCl, pH 7.0, 150 mM NaCl | 10 mg/ml (0.4 mM) EcMazE24A + 2.5 mM MazE (68–82) in 20 mM Tris-HCl, pH 7.0, 150 mM NaCl |
| Reservoir solution | 0.2 M CH3COONa, 0.1 M Tris-HCl, pH 8.5, 30% PEG 4000 | 0.2 M MgCl2, 0.1 M (CH3)2AsO2Na, pH 6.5, 50% PEG6000 | 24% PEG1500, 20% glycerol |
| Drop contents | 0.1 + 0.1 μl | 0.1 + 0.1 μl | 0.1 + 0.1 μl |

**Data collection**

| Space group | P3_1, 21 | C2 | P6_1 |
| Cell dimensions | 114.14, 47.5 | 68.8, 32.2, 86.3 | 52.2, 52.2, 198 |
| Resolution range | 4.75–2.90 (3.07–2.90) | 42.74–1.63 (1.73–1.63) | 44.04–2.48 (2.63–2.48) |
| Rmerge | 0.224 (0.778) | 0.105 (0.372) | 0.087 (0.883) |
| Completeness (%) | 92.2 (92.4) | 95.9 (90.9) | 98.7 (98.4) |
| Unique reflections | 7,461 (1,130) | 22,847 (3,469) | 10,603 (1,681) |
| Total no. of reflections | 40,771 (10,292) | 82,883 (12,312) | 39,042 (6,268) |
| Wavelength (Å) | 0.9801 | 0.9801 | 0.9801 |

**Refinement**

| Rfree | 0.223/0.290 | 0.178/0.236 | 0.197/0.248 |
| Root mean square deviation Bond angles (°) | 0.73 | 1.05 | 1.23 |
| Root mean square deviation Bond lengths (Å) | 0.004 | 0.020 | 0.010 |
| Content asymmetric unit | 1.5 MazF dimers + 3 DNA fragments | 1 MazF dimer + 1 MazE peptide | 1 MazF dimer + 1 MazE peptide |
| No. of atoms | Protein 2,558 | Solvent 663 | 33 |
| Other 152 | 25.31 | 68.25 |
| Average B-factor (Å²) | All atoms 38.32 | Protein 38.87 | 68.29 |
| Solvent 38.17 | Solvent 25.95 | 45.19 | 61.75 |
| Ramachandran profile (%) Core region 97.0 | Additional allowed disallowed 1.3 | 98.1 | 1.9 |
| PDB code | 5CR2 | 5CQX | 5CQY |
E. coli complex (Fig. 2, F and G). In EcMazF, Asp-76 is the equivalent residue of Glu-78, and the shorter side chain can no longer contact the base. BsMazF Gln-79 is substituted to Gln-77 in EcMazF, where this side chain is located further away from dA1, preventing a direct hydrogen bond. BsMazF further discriminates at dU6, for which an additional hydrogen bond is observed between this base and the side chain of Glu-78. Again, this hydrogen bond is absent in the EcMazF complex. As in A, the exposed surfaces of Leu-64, Gln-67, Val-72, Leu-74, Asp-76, and Gln-77, which form the “bottom” of the upstream subsite and downstream-binding groove and do not contact the substrate directly, are colored red. The substrate-mimicking DNA 5’-AUACAUA-3’ is represented as sticks. Upstream and downstream binding sites are indicated. E, equivalent representation of 5’-UUDUAUCAUA-3’ bound to BsMazF. Here, additional specificity-determining contacts are made between the oligonucleotide and the upstream and downstream binding sites, leading to a better complementarity of the surfaces of both macromolecules. F, schematic representation of nucleotide-specific recognition of d(AUACAU) substrate by EcMazF. The substrate-mimicking DNA 5’-AUACAUA-3’ is represented as sticks. Black lines indicate hydrogen bonds, and red lines indicate hydrophobic contacts. Asterisks designate interactions with main chain atoms of the given EcMazF amino acid. No electron density is seen for dA1 and dA7, and hence no interactions with the protein can be deduced. The extended conformation shown for the oligonucleotide does not represent the conformation of the molecule in the crystal, but is intended as a schematic. G, equivalent representation for the 5’-dUACAU-3’ moiety in the BsMazF complex (PDB entry 4MDX).
RNA Binding and Activity Regulation of MazF

EcMazF<sup>E24A</sup> Is Defective in Substrate Recognition and Prevents EcMazF to Adopt an Active Conformation—Previous in vitro interaction studies between EcMazF and EcMazE using biophysical techniques involved the catalytically inactive MazF<sup>E24A</sup> mutant for practical reasons (40). The reason for its lack of activity, defective in catalysis or substrate recognition, remains unclear. The supplemental Fig. S3 shows that although EcMazF degrades bacteriophage MS2 genomic RNA and is inhibited by stoichiometric amounts of EcMazE, EcMazF<sup>E24A</sup> is incapable of RNA cleavage. We then performed ITC experiments with three substrate-mimicking RNAs where the potential scissile nucleotide is replaced by its 2'-deoxy analog, thus preventing cleavage but allowing binding (see under “Experimental Procedures” for the full sequences). RNA1 corresponds to a <i>bona fide</i> substrate for EcMazF (24) and indeed binds to the wild type enzyme with low micromolar affinity, in agreement for what is expected for an enzyme-substrate interaction (Table 2 and supplemental Fig. S4). However no binding was observed in EcMazF<sup>E24A</sup>. RNA2 corresponds to an RNA sequence that is not cleaved by the enzyme despite containing an UACAU recognition sequence (24). It was assumed that the lack of cleavage would be due to secondary structure. In agreement with this, neither wild type nor mutant shows binding in the conditions used. RNA3 does not contain a UACAU recognition sequence and is also not bound by either wild type or mutant enzyme as expected.

The crystal structures of EcMazF<sup>E24A</sup> again show conformational flexibility for loops β1-β2 and β4-β5. The conformations of loop β1-β2 appear to be distributed differently from the wild type ensemble and to adopt a more “open” conformation, pointing away from the body of the EcMazF monomer (Fig. 1C and supplemental Fig. S1).

In the EcMazF-d(AUCAUA) complex, the side chain of Glu-24 does not interact with the ligand. Rather, this side chain is buried under loop β1-β2, its negative charge being neutralized by salt bridges with the side chains of 6 and Lys-79 (Fig. 5A). Absence of the Glu-24 side chain in the d(AUCAUA)-bound state would create a cavity where the positive charges of Arg-86 and Lys-79 are buried and repel each other, likely a highly destabilizing situation.

E24A Mutation Does Not Affect Antitoxin Recognition—We next compared the binding of the intrinsically disordered domain of EcMazE (EcMazE<sup>(50–82)</sup>) to EcMazF and EcMazF<sup>E24A</sup> using ITC (Table 2 and Fig. 6). We find that both EcMazF and EcMazF<sup>E24A</sup> possess two binding sites for EcMazE<sup>(50–82)</sup>. Binding is sequential and with strong negative cooperativity. The first EcMazE<sup>(50–82)</sup> molecule binds with an affinity around 10 nM, although subsequent binding of the second EcMazE<sup>(50–82)</sup> is 3 orders of magnitude weaker. This cooperativity is largely due to the C-terminal EcMazE residues Asp-78–Trp-82 (Table 2 and Fig. 6).

To investigate the interaction between EcMazF and residues Asp-78–Trp-82 of EcMazE, we determined the crystal structure of EcMazF<sup>E24A</sup> in complex with EcMazE<sup>(68–82)</sup>. In this structure, a single EcMazE<sup>(68–82)</sup> peptide is bound to the EcMazF<sup>E24A</sup> dimer, in agreement with the stoichiometry genes of the phosphate (Fig. 4A), indicating that these residues might contribute to catalysis by stabilizing the build up of negative charge on this atom in the bipyramidal transition state. This is in agreement with the observation that substitution of BsMazF Thr-48 (the equivalent of <i>E. coli</i> Thr-52) leads to an inactive enzyme (27).

The guanidinium group of Arg-29 is located such that it can interact with both the 2’-OH and the 5’-O’ group of the two nucleosides adjacent to the cleavage position (Fig. 4A), suggesting a key role in catalysis rather than charge stabilization as suggested earlier (27). The peculiar position of the guanidinium group of Arg-29, bridging the nucleophile and the leaving group, would allow for a dual general base/general acid role. In such a scenario we hypothesize that Arg-29 would act as a proton relay by concertedly abstracting a proton from the 2’-OH and donating a proton to the 5’-O’ leaving group. This finally results in the reshuffling of a proton from the 2’-OH to the 5’-O’, with the arginine acting as a relay in a Grotthuss-like mechanism (Fig. 4B).

FIGURE 3. Subsite details. (A) details of the interactions in the upstream U binding cleft. Residues of EcMazF interacting with the base are shown in stick representation and colored as in Fig. 1A. The uridine base is surrounded by the aliphatic/aromatic parts of the side chains of Trp-14, Arg-29, Pro-30, Thr-52, Thr-53, and Arg-69. The specificity-determining hydrogen bond from the main chain NH of Arg-69 to O4 of the uridine base is shown as a gray dotted line. Amino acid side chains that are part of the bottom of the binding site but do not touch the substrate are colored red. The equivalent residues of MazF Thr-48 (the equivalent of <i>E. coli</i> Thr-52) are superimposed as black sticks and labeled in black. B, details of the interactions in the downstream ACA-binding groove. Residues of EcMazF interacting with d(A<sup>2</sup>C<sup>3</sup>A<sup>4</sup>U<sup>5</sup>) are shown in stick representation and colored as in Fig. 1A. Corresponding secondary structure elements are shown as a schematic. Hydrogen bonds are represented as dotted lines. Amino acid side chains that are part of the bottom of the binding groove but do not touch the substrate are colored red.
RNA Binding and Activity Regulation of MazF

Thermodynamic parameters for wild type MazF and RNA1, wild type MazF, and MazF/E24A mutant with MazE peptides

Table 2

| Interaction          | n  |  | kcal/mol | kcal/mol | kcal/mol |
|----------------------|----|---|----------|----------|----------|
| EcMazF-RNA1          | 1  |  | 5.2 ± 0.3 10^-7 | -8.6 ± 0.2 | -1.0 ± 0.1 | -9.6 ± 0.2 |
| EcMazF-EcMazE(50–82) | 2  |  | 1.1 ± 0.3 10^-8 | -2.1 ± 0.5 | 11.5 ± 0.2 | -10.6 ± 0.2 |
| EcMazF/E24A-EcMazE(50–82) | 2  |  | 5.9 ± 0.1 10^-5 | -5.7 ± 0.4 | -3.7 ± 0.1 | -9.4 ± 0.2 |
| EcMazF-EcMazE(54–77) | 2  |  | 1.6 ± 0.4 10^-8 | -19.9 ± 0.3 | 9.2 ± 0.2 | -10.6 ± 0.2 |
| EcMazF/E24A-EcMazE(54–77) | 2  |  | 1.9 ± 0.2 10^-5 | -45 ± 0.1 | -3.2 ± 0.1 | -7.7 ± 0.1 |
| EcMazF-EcMazE(54–77) | 2  |  | 1.1 ± 0.1 10^-8 | -10.3 ± 0.1 | -0.5 ± 0.1 | -10.8 ± 0.1 |
| EcMazF/E24A-EcMazE(54–77) | 2  |  | 2.2 ± 0.5 10^-7 | -4.0 ± 0.2 | -5.0 ± 0.1 | -9.0 ± 0.1 |
| EcMazF/E24A-EcMazE(68–82) | 2  |  | 1.8 ± 0.1 10^-8 | -13.4 ± 0.2 | 2.8 ± 0.1 | -10.5 ± 0.1 |
| EcMazF(68–82)        | 1  |  | 2.0 ± 0.1 10^-7 | -7.6 ± 0.2 | -1.4 ± 0.3 | -9.1 ± 0.1 |
| EcMazF(68–82)        | 1  |  | 4.7 ± 0.1 10^-6 | -3.2 ± 0.4 | -0.4 ± 0.1 | -3.6 ± 0.2 |

obtained from ITC. Electron density is visible for the complete EcMazE(68–82) peptide (Fig. 5B).

In our structures, Pro-76–Trp-82 forms a β-hairpin, the presence of which prevents loop β1-β2 of both EcMazF monomers to adopt the catalytically competent conformation and simultaneously blocks both substrate-binding sites on the EcMazF dimer (Fig. 5C). Segment Glu-69–Glu-75 would clash with the β1-β2 loop from one monomer, and segment Glu-75–Trp-82 would clash with the β1-β2 loop from the second monomer.

Interestingly, EcMazE(68–82) forms a partial mimic for loop β1-β2. Not only does the path of the backbone of EcMazE residues Pro-76–Val-81 coincide with that of EcMazF residues Thr-20–Gln-25, the side chain of EcMazE Glu80 takes over the role of EcMazF Glu-24 (Fig. 5C). EcMazE Glu-80 makes identical salt bridges to EcMazF Arg-86 and Lys-79, neutralizing the positive charges that otherwise would become buried upon MazE binding. Furthermore, Trp-73 and Trp-82 insert into identical hydrophobic pockets (one in each subunit of the MazF dimer) bordered by the side chains of Val-15, Ala-31, Cys-48, Pro-50, and Ile-81 and the aliphatic parts of the side chains of Arg-29 and Lys-79 (Fig. 7, A and B). In the d(AUACUAU)-bound conformation as well as in some of the conformations in the ensemble of wild type EcMazF structures (but not in the EcMazF/E24A ensemble), these tryptophans are mimicked by EcMazF Phe-17.
RNA Binding and Activity Regulation of MazF

Discussion

Molecular Framework for MazF Substrate Recognition and Catalysis—Knowledge of the structural basis of substrate recognition by MazF proteins has been limited to the crystal structure of a substrate mimic complex of BsMazF, which shares only 24% sequence identity with the E. coli enzyme. Comparison of both proteins reveals a common mode of substrate recognition on a large interaction surface (around 400 Å² of protein surface gets buried). The EcMazF substrate-binding site is characterized by fewer specificity-determining hydrogen bonds to the RNA bases and poorer surface complementarity compared with BsMazF, thus explaining the broader specificity of EcMazF. The differences between EcMazF and BsMazF thus provide a framework for rationalizing the differences in specificity observed within the MazF family.

Common with BsMazF, and likely a general feature in the MazF family, is induced fit substrate binding. Both EcMazF and BsMazF only adapt a catalytically competent conformation for loop β1-β2 when a substrate or substrate mimic is bound (27). In both enzymes, the antitoxin also prevents loop β1-β2 from adopting this conformation. However, loop β1-β2 of BsMazF adopts the active conformation in the unbound state as well. In contrast, substrate binding to BsMazF refolds loop β4-β5 (27). The latter is in agreement with both NMR and x-ray data that are available for the closely related Staphylococcus aureus SaMazF (64% sequence identity with BsMazF). It was observed that loop β4-β5 of SaMazF shows the largest conformational variability in the unbound state (47).

The catalytic site encompasses only two residues that can potentially fulfill a catalytic role and are also conserved within the MazF family. Arg-29 is most likely the key catalytic residue by reshuffling a proton from the 2’-OH to the 5’O-leaving group (Fig. 4B). Because of its high pK_a in water (−12), arginine is usually not considered as a very likely general acid/base. However, such a catalytic role for arginine would not be unprecedented, and arginine residues with perturbed pK_a values are found to act as general acid/base in a growing number of enzymes, including IMP dehydrogenase, pectate/pectin lyases, fumarate reductase, l-aspartate oxidase, and tyrosine-phenol lyase (41). An arginine residue has also been suggested to act as a general acid for the 3’-O leaving group in RNA cleavage by E. coli RelE where a lysine side chain is presumed to act as a catalytic base (42).

Involvement of an arginine residue in a proton wire has also been observed in Shewanella frigidimarina fumarate reductase, where the proton wire finally feeds protons to another arginine residue that acts as a general acid in the enzyme-catalyzed reaction (43). Our proposed mechanism would lead to a 2’,3’-cyclic phosphate end product, in agreement with previous observations for EcMazF and Kid (a MazF homolog encoded on plasmid R1 – (36)).

Structural Similarities Suggest That the Functional Divergent MazF and CcdB Families Are of a Common Ancestral Origin—Although CcdB and MazF proteins are structurally related (44, 45), their activities are divergent; CcdB inhibits gyrase through binding to its A subunit, although MazF is a ribonuclease. Still key features of how the antitoxin regulates activity of the toxin are paralleled. Two symmetrically placed pockets on the EcMazF dimer that are filled by distinct aromatic residues from the C terminus of MazE are reminiscent of the interaction between CcdB and CcdA (46). Indeed, superposition of the F-plasmid CcdB-CcdA(37–72) complex (PDB entry 3HPW) on our EcMazF-MazE(68–82) complex reveals that Phe-65 and Trp-72 of CcdA appear to be the structural and functional equivalents of Trp-73 and Trp-82 of EcMazE (Fig. 7C).
Furthermore, locking of CcdA residues Phe-65 and Trp-72 into their recognition pockets on CcdB results in a significant 12° relative rotation of both subunits in the CcdB dimer compared with the gyrase-bound conformation of CcdB (46). A similar 10° relative rotation is observed for the EcMazF dimer upon binding EcMazE(68–82) (Fig. 7D). Thus, both F-plasmid CcdB and EcMazF employ the relative rotation of the two monomers as an allosteric signal to alter their target binding surface. The allosteric effector in both cases consists of two aromatic residues at the C terminus of the antitoxin. Next to this allosteric component, both EcMazF and F-plasmid CcdB proteins also use steric hindrance from the N-terminal side of the intrinsically disordered domain of the antitoxin to prevent target/substrate binding. In addition, like in MazF proteins, the loop β1-β2 of CcdB is highly flexible in the free state of the protein (45, 47, 48) and becomes ordered upon antitoxin binding. The latter loop in CcdB does not, however, seem to fold in a unique conformation when gyrase is bound (49), although correct folding of the equivalent loop in EcMazF is required for substrate binding and catalysis. Together, these strong parallels

**FIGURE 6. Isothermal titration calorimetry of EcMazE fragments binding to EcMazF.** A, titration of EcMazE(50–82) into EcMazFE24A. B, titration of EcMazE(54–77) into EcMazFE24A. C, titration of EcMazE(68–82) into EcMazFE24A. D, titration of EcMazE(50–82) into wild type EcMazF. E, titration of EcMazE(54–77) into wild type EcMazF. All the experiments were done at 305 K in 50 mM phosphate buffer, pH 7.0, 150 mM NaCl, 1 mM EDTA. Each time, the raw ITC data are shown in the upper panel, and the corresponding binding isotherm obtained from integrating the area under the peaks is shown in the lower panel.
strongly favor common ancestor origins for the MazF and CcdB proteins, and possibly also for the neutralization domains of their corresponding antitoxins.

**Author Contributions**—V. Z. purified proteins, carried out and interpreted ITC measurements, performed the crystallization, determined the crystal structures, and co-wrote the manuscript. A. M. contributed to the ITC experiments and their interpretation. J. L. designed ITC experiments and interpreted the thermodynamic data. W. V. contributed to the interpretation of the MazF mechanism and wrote part of the manuscript. Y. G. J. S. performed the MazFE24A mutation, crystallization, and data collection of two MazFE24A crystal forms. N. D. J. participated with formulating the hypotheses and designed part of the study. A. G. P. contributed to x-ray data processing and interpretation. H. D. G. designed and supervised cloning, site-specific mutagenesis, as well as purification of the two MazFE24A crys-

**Acknowledgments**—We thank the beamline scientists from BM29 (ESRF, Grenoble, France) and PROXIMA-1 and PROXIMA-2a (SOLEIL, Gif-sur-Yvette, France) for their support.

**References**

1. Lewis, K. (2007) Persister cells, dormancy and infectious disease. Nat. Rev. Microbiol. 5, 48–56
2. Lewis, K. (2010) Persister cells. Annu. Rev. Microbiol. 64, 357–372
3. Amato, S. M., Orman, M. A., and Brynildsen, M. P. (2013) Metabolic control of persister formation in Escherichia coli. Mol. Cell 50, 475–487
4. Maisonneuve, E., and Gerdes, K. (2014) Molecular mechanisms underlying bacterial persisters. Cell 157, 539–548
5. Germain, E., Roghanian, M., Gerdes, K., and Maisonneuve E (2015) Stochastic induction of persister cells by HipA through (p)pGpp-mediated activation of mRNA endonucleases. Proc. Natl. Acad. Sci. U.S.A. 112, 5171–5176
6. Verstraeten, N., Knapen, W. J., Kint, C. I., Liebens, V., Van den Bergh, B., Dewachter, L., Michiels, J. E., Fu, Q., David, C. C., Fierro, A. C., Marchal, K., Beirlant, J., Versées, W., Hofkens, J., Jansen, M., et al. (2015) Obg and membrane depolarization are part of a microbial bet-hedging strategy that leads to antibiotic tolerance. Mol. Cell 59, 9–21
7. Correia, F. F., D’Onofrio, A., Rejtar, T., Li, L., Karger, B. L., Makarova, K., Koonin, E. V., and Lewis, K. (2006) Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in Escherichia coli. J. Bacteriol. 188, 8360–8367
8. Maisonneuve, E., Shakespeare, L. J., Jørgensen, M. G., and Gerdes K. (2011) Bacterial persistence by RNA endonucleases. Proc. Natl. Acad. Sci. U.S.A. 108, 13206–13211
9. Tripathi, A., Dewan, P. C., Siddique, S. A., and Varadarajan, R. (2014) MazF-induced growth inhibition and persister generation in Escherichia coli. J. Biol. Chem. 289, 4191–4205
RNA Binding and Activity Regulation of MazF

10. Yamaguchi, Y., Park, J. H., and Inouye, M. (2011) Toxin-antitoxin systems in bacteria and archaea. *Annu. Rev. Genet.* 45, 61–79

11. Hayes, F., and Van Melderen, L. (2011) Toxins-antitoxins: diversity, evolution and function. *Crit. Rev. Biochem. Mol. Biol.* 46, 386–408

12. Loris, R., and Garcia-Pino, A. (2014) Disorder- and dynamics-based regulatory mechanisms in toxin-antitoxin modules. *Chem. Rev.* 114, 6933–6947

13. Christensen, S. K., Pedersen, K., Hansen, F. G., and Gerdes, K. (2003) Toxin-antitoxin loci as stress-response-elements: CphA/MazF and CphB cleave translated RNAs and are counteracted by tmRNA. *J. Mol. Biol.* 332, 809–819

14. Mukherjee, J., Heflick, K. P., Ikura, M., Qian, G., and Inouye, M. (2003) MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Mol. Cell.* 12, 913–923

15. Belitsky, M., Avshalom, H., Erental, A., Yelin, I., Kumar, S., London, N., and Loris, R. (2005) Molecular basis of gyrase poisoning by the addiction protein domain of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Acta Crystallogr. D Biol. Crystallogr.* 61, 875–884

16. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68, 352–367

17. Otwinowski, Z., and Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326

18. Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A. C., Engelberg-Kulka, H. (2009) *Escherichia coli* MazF toxin exhibits selective, not global, mRNA cleavage. *J. Biol. Chem.* 284, 15421–15428

19. Safer, G., Bismut, J., and Tjian, R. (2009) Multiple RNA species are cleaved by the E. coli toxin MazF. *FEBS Lett.* 583, 316–320

20. Loris, R., Dao-Thi, M. H., Bahassi, E. M., Van Melderen, L., Poortmans, F., and Wyns, L. (2014) Structural and functional analysis of the Kid toxin protein from *Escherichia coli*. *J. Biol. Chem.* 289, 2526–2532

21. Simanshu, D. K., Yamaguchi, Y., Park, J. H., Inouye, M., and Patel, D. J. (2013) Structural basis of mRNAs recognition and cleavage by toxin MazF and its regulation by antitoxin MazE in *Bacillus subtilis*. *Mol. Cell* 52, 447–458

22. Belitsky, M., Avshalom, H., Erental, A., Yelin, I., Kumar, S., London, N., Sperber, M., Schueler-Furman, O., and Engelberg-Kulka, H. (2011) The *Escherichia coli* MazE/MazF complex: molecular bases of antidote-toxin recognition. *Acta Crystallogr. D Biol. Crystallogr.* 67, 619–623

23. De Jonge, N., Hohlweg, W., Garcia-Pino, A., Respondek, M., Buts, L., Haesaerts, S., Lah, J., Zangger, K., and Loris, R. (2010) Structural and thermodynamic characterization of *Vibrio fischeri* CcdB. *J. Biol. Chem.* 285, 5606–5613

24. Zorzini, V., Buts, L., Sleutel, M., Garcia-Pino, A., Talavera, A., Haesaerts, S., De Greve, H., Cheung, A., van Nuland, N. A., and Loris, R. (2014) Structural and biophysical characterization of *Staphylococcus aureus* SaMazF shows conservation of functional dynamics. *Nucleic Acids Res.* 42, 6709–6725

25. De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., Wyns, L., De Greve, H., and Loris, R. (2009) Rejuvenation of CcdB-poisoned gyrase by an intrinsically disordered protein domain. *Mol. Cell* 35, 154–163

26. De Jonge, N., Garcia-Pino, A., Buts, L., Haesaerts, S., Charlier, D., Zangger, K., Wyns, L., De Greve, H., and Loris, R. (2009) Molecular basis of gyrase poisoning by the addiction toxin CcdB. *J. Biol. Chem.* 348, 1091–1102