**Regulatory Evolution of TEM β-lactamases through Large-Scale Genomic Analysis Reveals Promoter Diversity and Its Impact on β-lactam Resistance in *Escherichia coli***

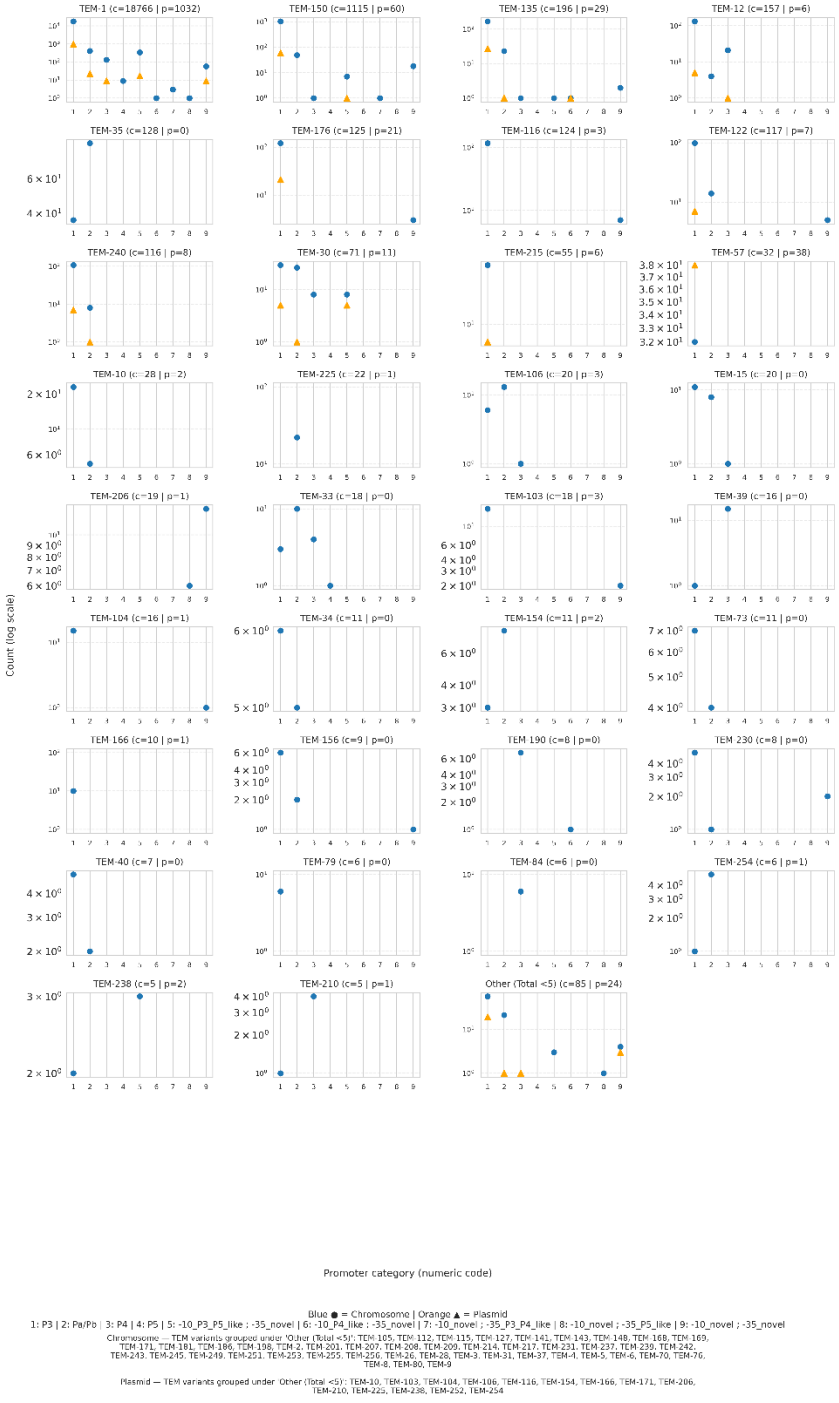
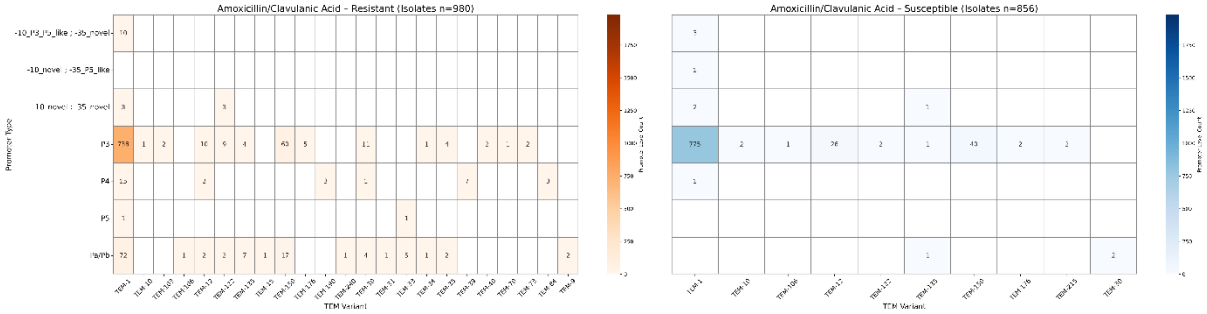
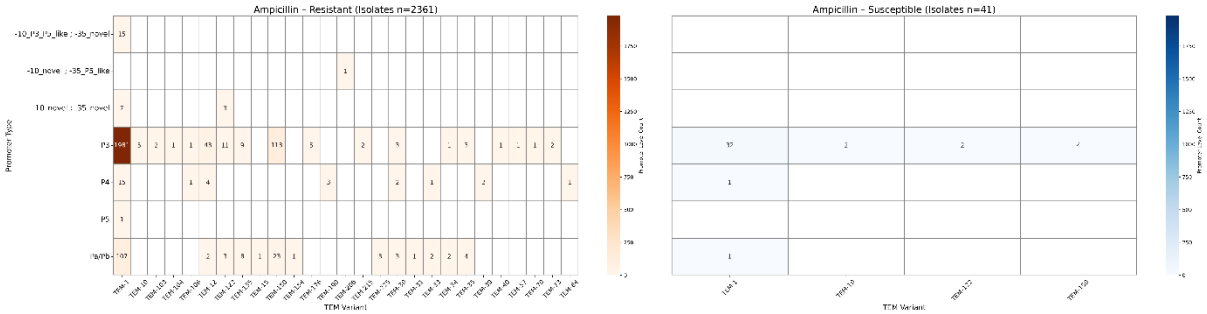



**Fig. S1.** TEM variant-wise distribution of promoter types identified in chromosome- and plasmid-associated *bla*$_{\text{TEM}}$ sequences. Each panel represents an individual TEM variant, with the x-axis showing promoter categories (numeric codes) and the y-axis showing the number of sequences on a logarithmic scale. Chromosome-derived sequences are indicated by blue circles and plasmid-derived sequences by orange triangles. Low-frequency TEM variants with fewer than five observations were grouped under “Other.” The numeric promoter category codes correspond to the promoter definitions provided below the figure.

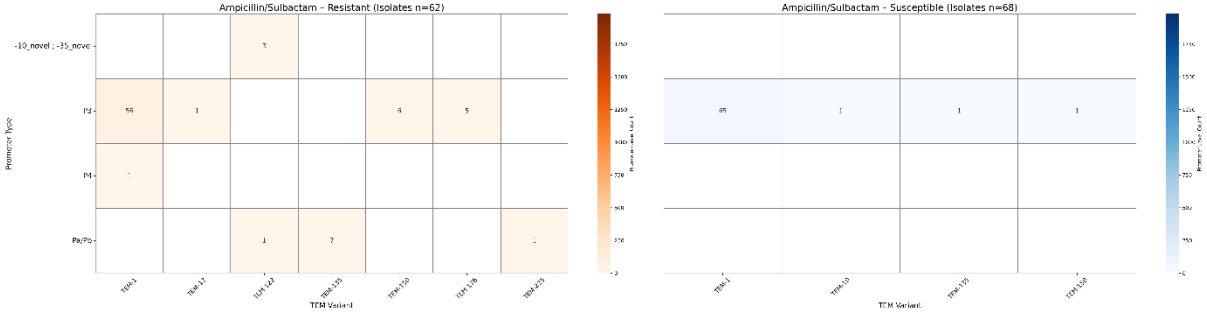
## Amoxicillin/Clavulanic Acid – Promoter × TEM Heatmaps



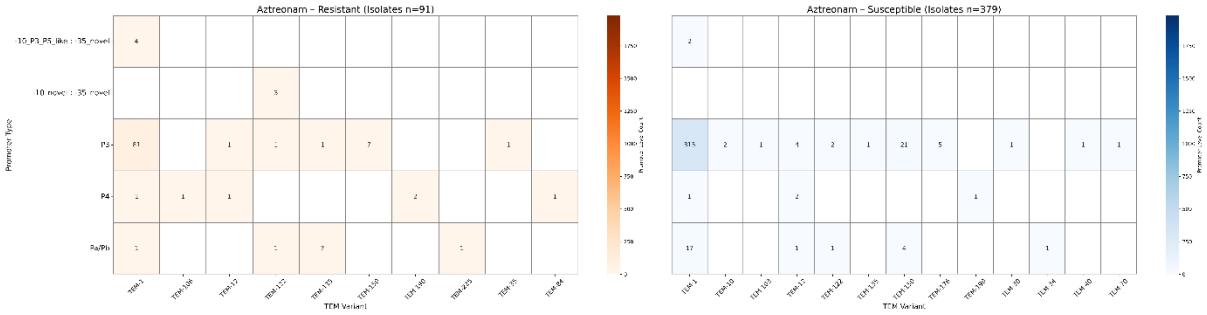

## Ampicillin – Promoter × TEM Heatmaps



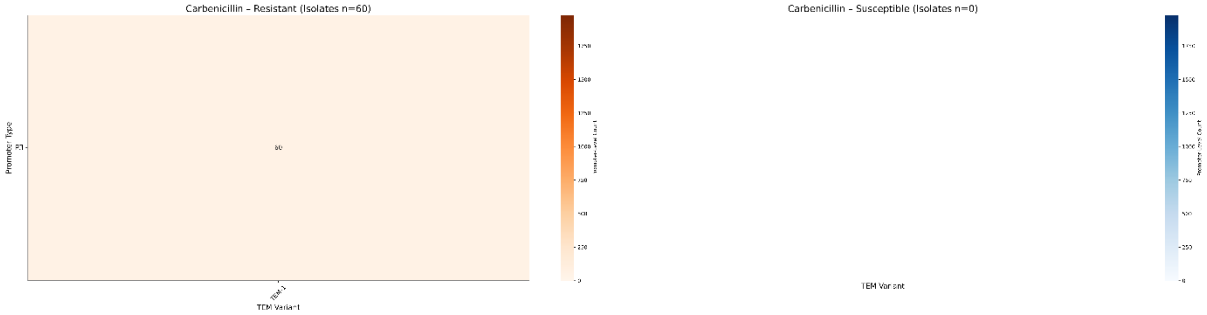

## Ampicillin/Sulbactam – Promoter × TEM Heatmaps



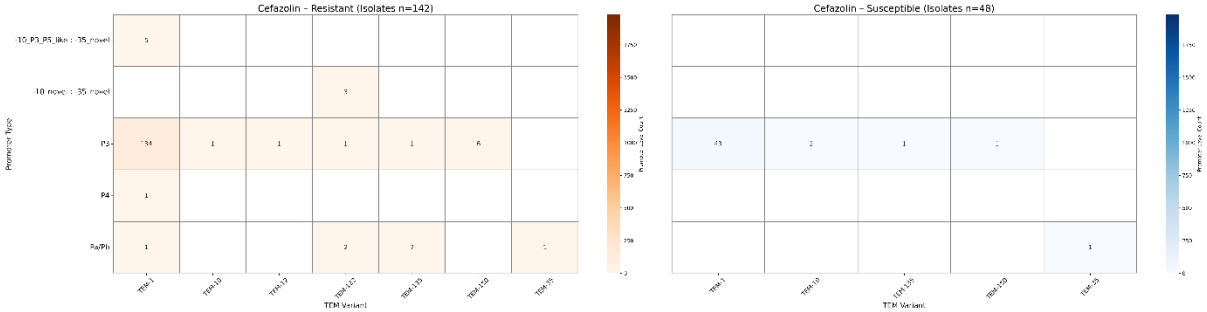

## Aztreonam – Promoter × TEM Heatmaps



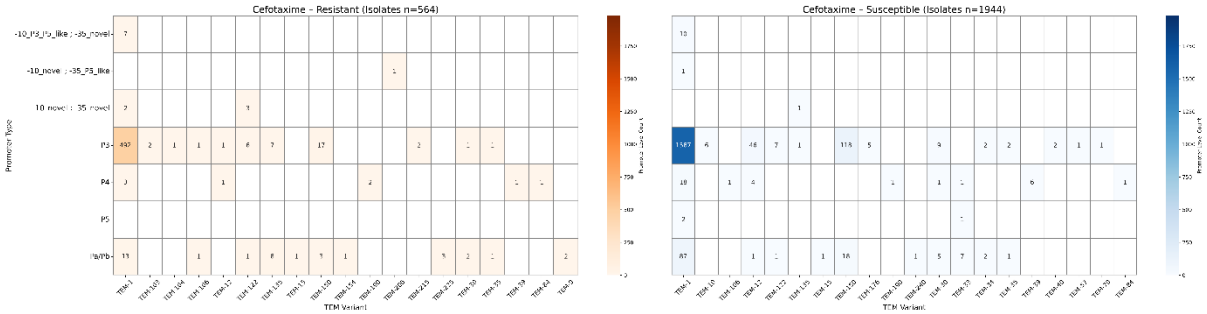

## Carbenicillin – Promoter × TEM Heatmaps



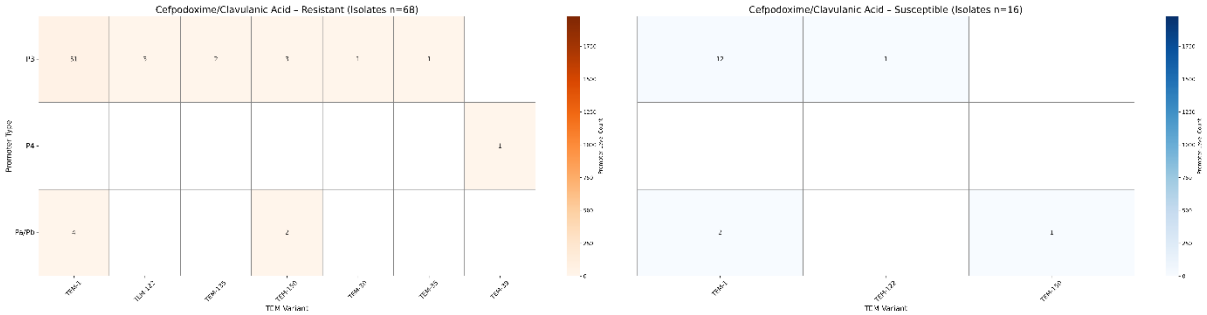
## Cefazolin – Promoter × TEM Heatmaps



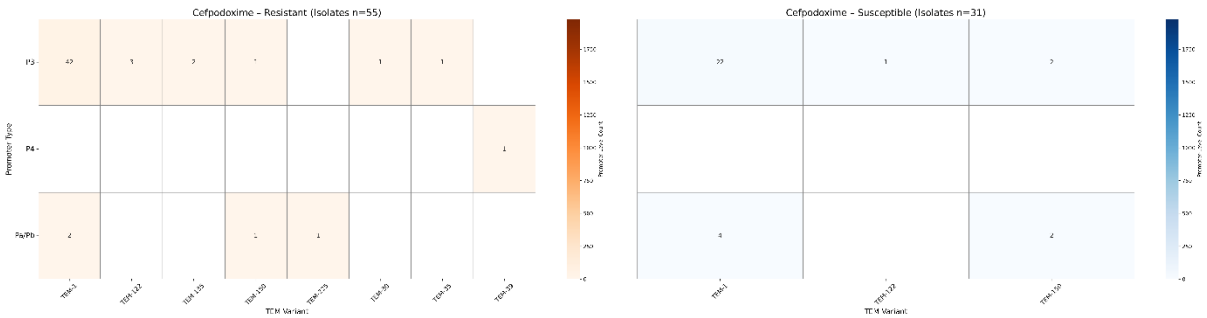

## Cefotaxime – Promoter × TEM Heatmaps



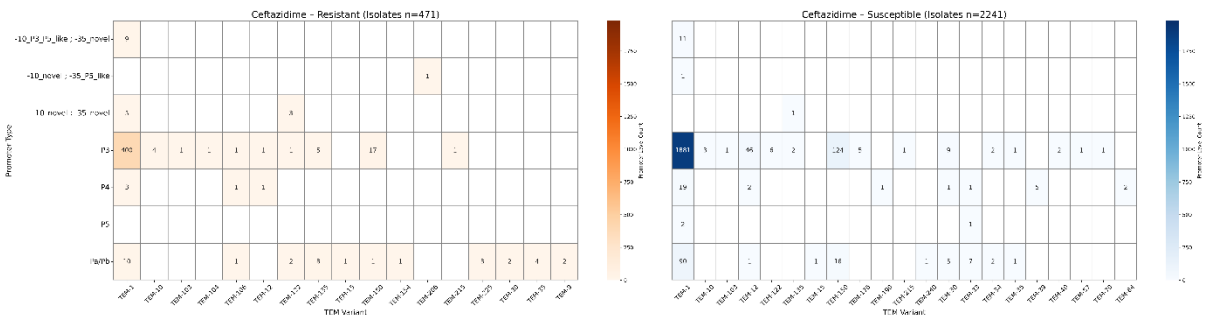

## Cefpodoxime/Clavulanic Acid – Promoter × TEM Heatmaps



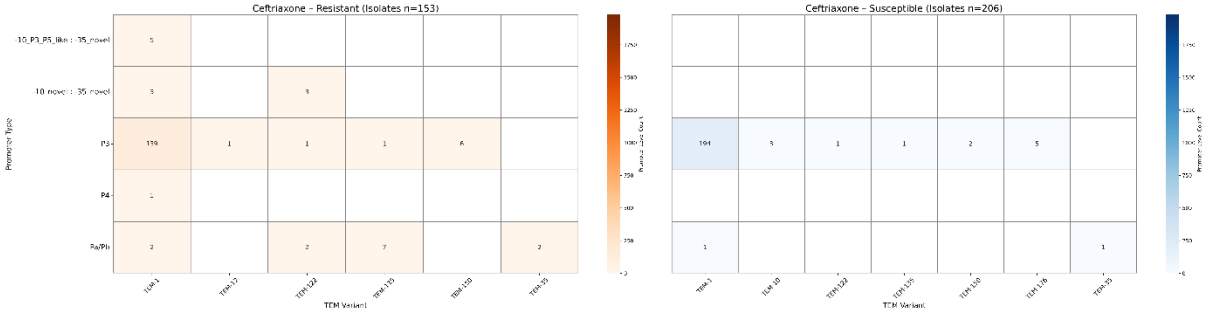

## Cefpodoxime – Promoter × TEM Heatmaps



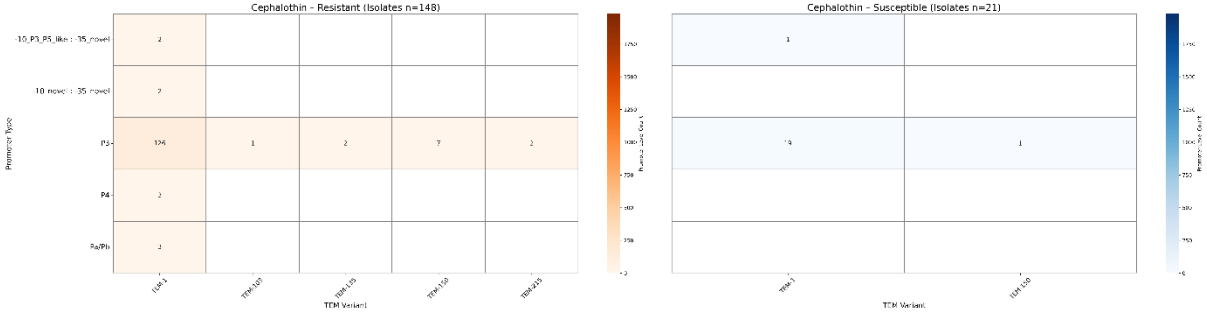

## Ceftazidime – Promoter × TEM Heatmaps



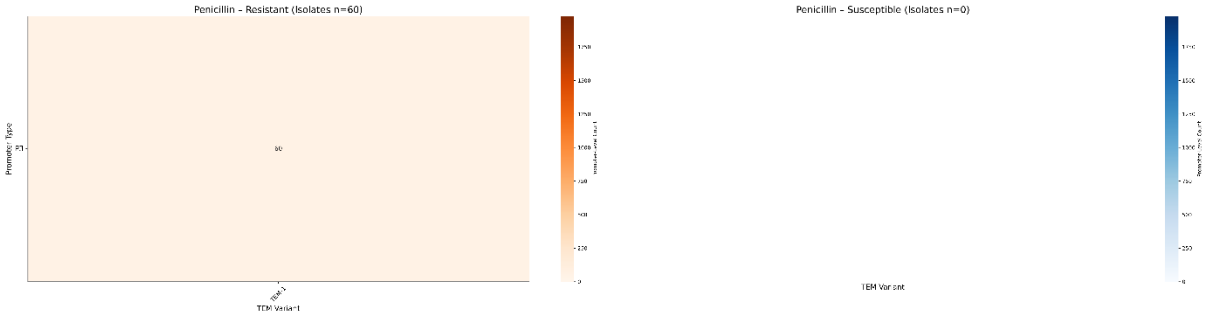
## Ceftriaxone – Promoter × TEM Heatmaps



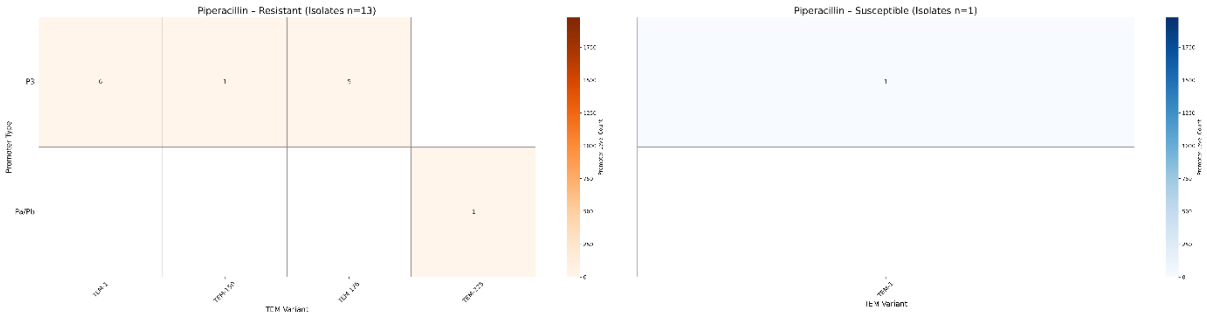

## Cephalothin – Promoter × TEM Heatmaps




## Penicillin – Promoter × TEM Heatmaps




## Piperacillin – Promoter × TEM Heatmaps



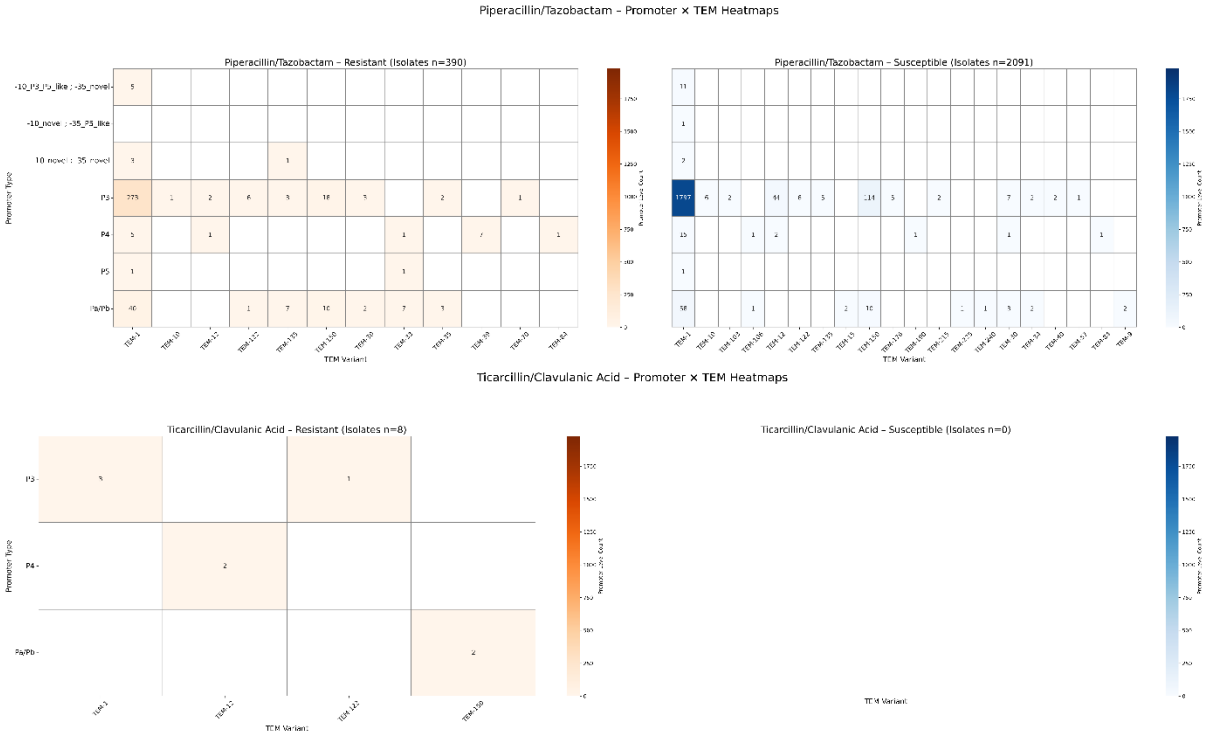



**Fig. S2**. Distribution of promoter types across resistant and susceptible chromosome-derived isolates for different TEM-relevant β-lactam antibiotics and β-lactam/β-lactamase inhibitor combinations. For each antibiotic, the heatmaps show the number of isolates corresponding to each promoter type-TEM variant combination, separately for resistant and susceptible phenotypes. This phenotype-level analysis was restricted to chromosome-derived isolates because corresponding phenotype data were unavailable for plasmid-associated isolates.

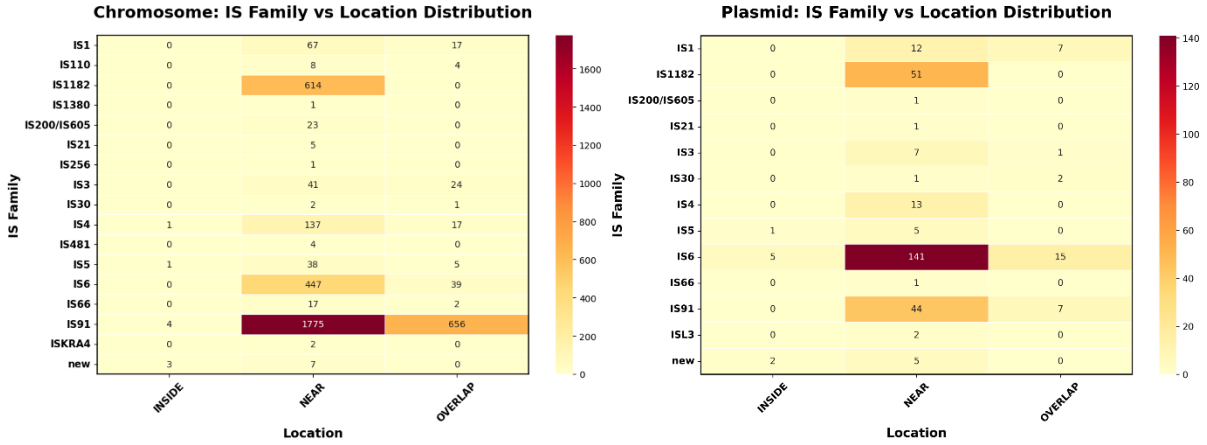
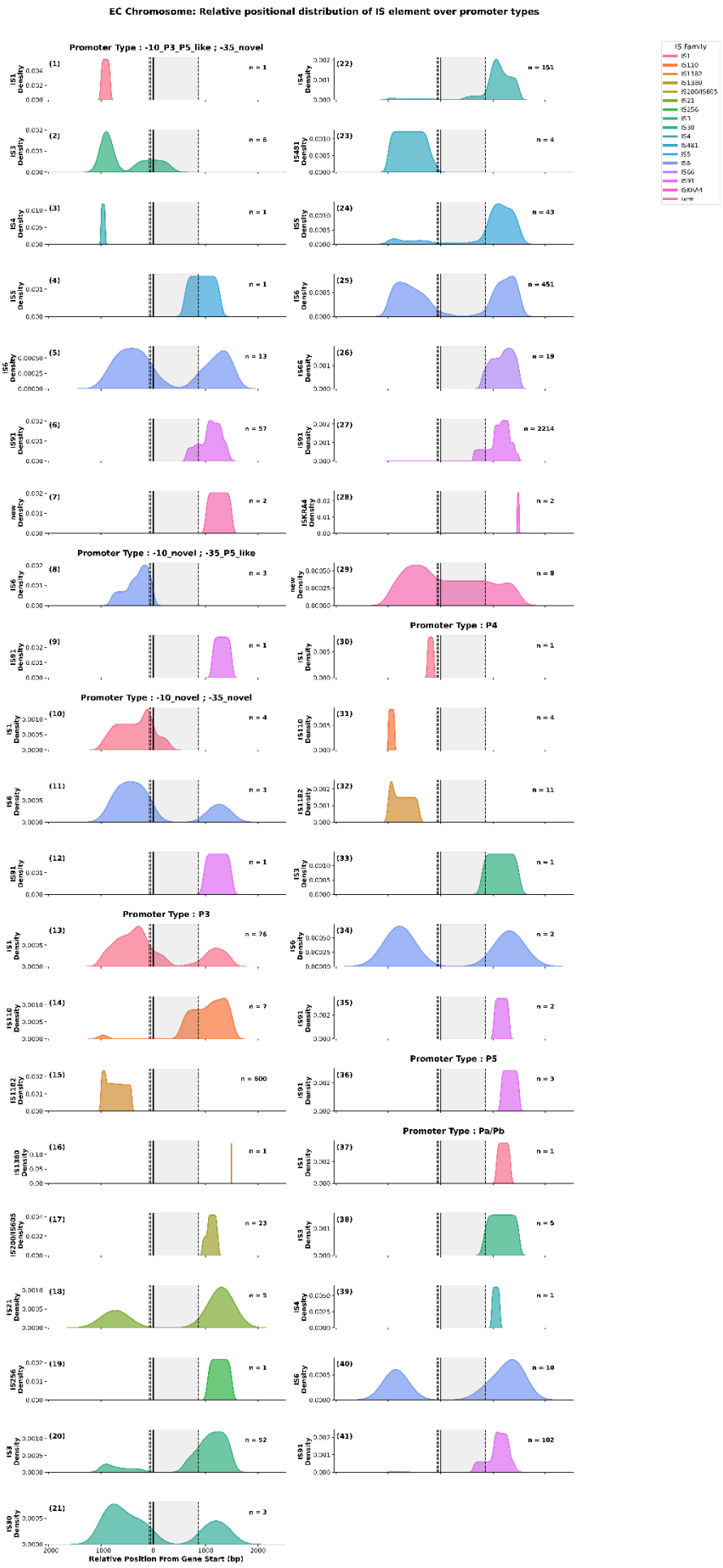



**Fig. S3.** Distribution of insertion sequence (IS) families identified within the analyzed ±1000-bp region surrounding the *bla*$_{\text{TEM}}$ locus in chromosome- and plasmid-associated genes. The y-axis lists the different IS families, while the x-axis represents the positional categories INSIDE, NEAR, and OVERLAP relative to the *bla*$_{\text{TEM}}$ region. Heatmap values indicate the number of IS occurrences corresponding to each IS family and positional category. The definitions of the positional categories are provided in the methods.

a)


EC Chromosome: Relative positional distribution of IS element over promoter types
Promoter Type : -10_P3_P5_like ; -35_novel
Promoter Type : -10_novel ; -35_P5_like
Promoter Type : -10_novel ; -35_novel
Promoter Type : P3
Promoter Type : P4
Promoter Type : P5
Promoter Type : Pa/Pb
IS Family
IS1
IS110
IS1182
IS1380
IS200/IS605
IS21
IS256
IS3
IS30
IS4
IS481
IS5
IS6
IS66
IS91
ISKRA4
new
Relative Position From Gene Start (bp)

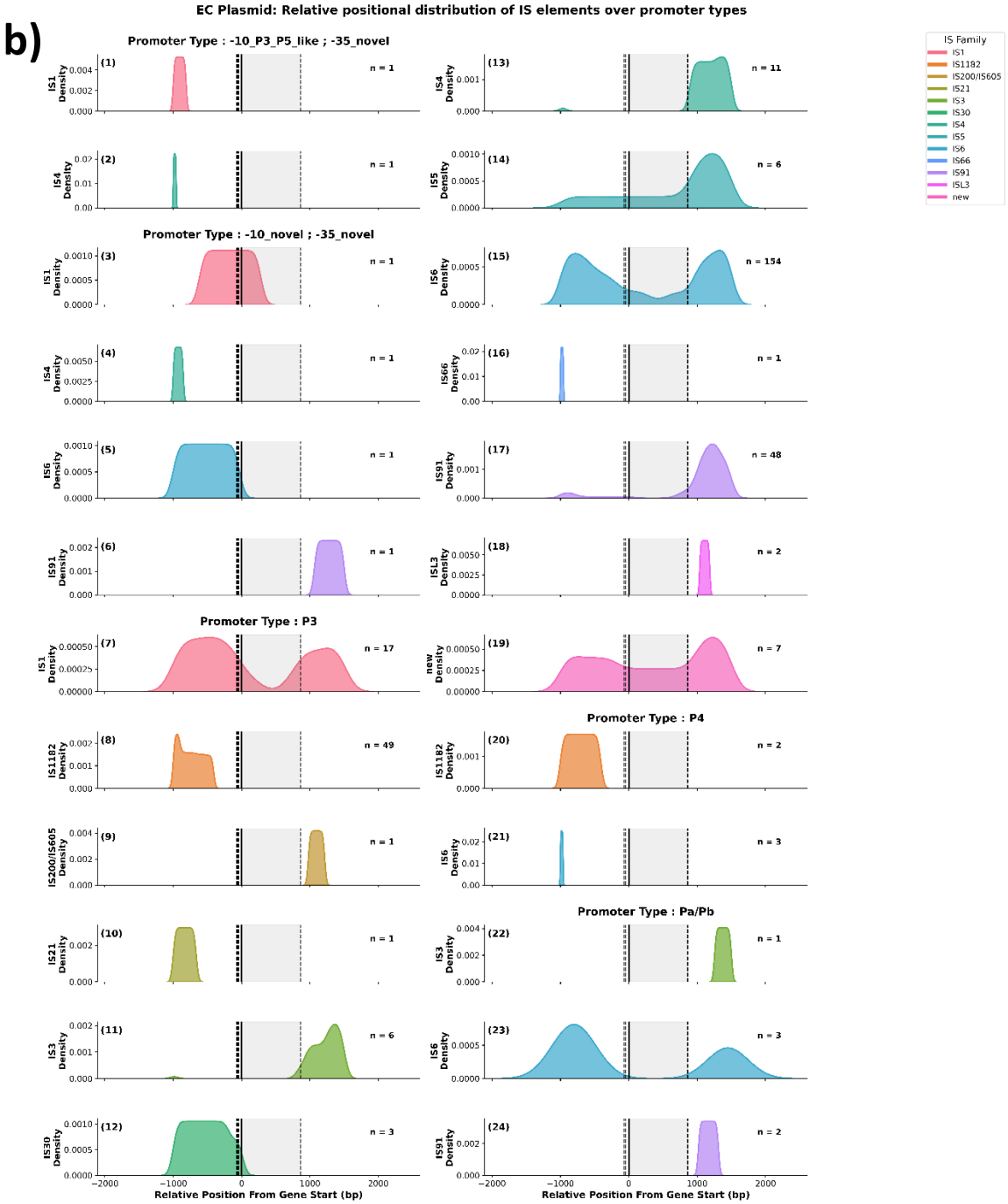



**Fig. S4.** (a) Chromosome-associated and (b) plasmid-associated IS families are represented using kernel density estimation (KDE) plots to show their positional distribution relative to *bla*$_{\text{TEM}}$ genes across different promoter types. Only IS elements identified within the predefined 1-kb upstream and downstream regions surrounding the *bla*$_{\text{TEM}}$ locus were included. Positions were normalized according to *bla*$_{\text{TEM}}$ gene orientation, with negative and positive positions representing upstream and downstream locations, respectively. The vertical line at 0 bp marks the gene start, the grey-shaded region represents the *bla*$_{\text{TEM}}$ gene body (0-861 bp), and the dashed line at 861 bp marks the gene end. Vertical reference lines at −70 bp and −47 bp indicate the approximate positions of the -35 and -10 promoter motifs, respectively. Each panel represents a promoter type-IS family combination, and n indicates the number of IS-element occurrences contributing to that combination.

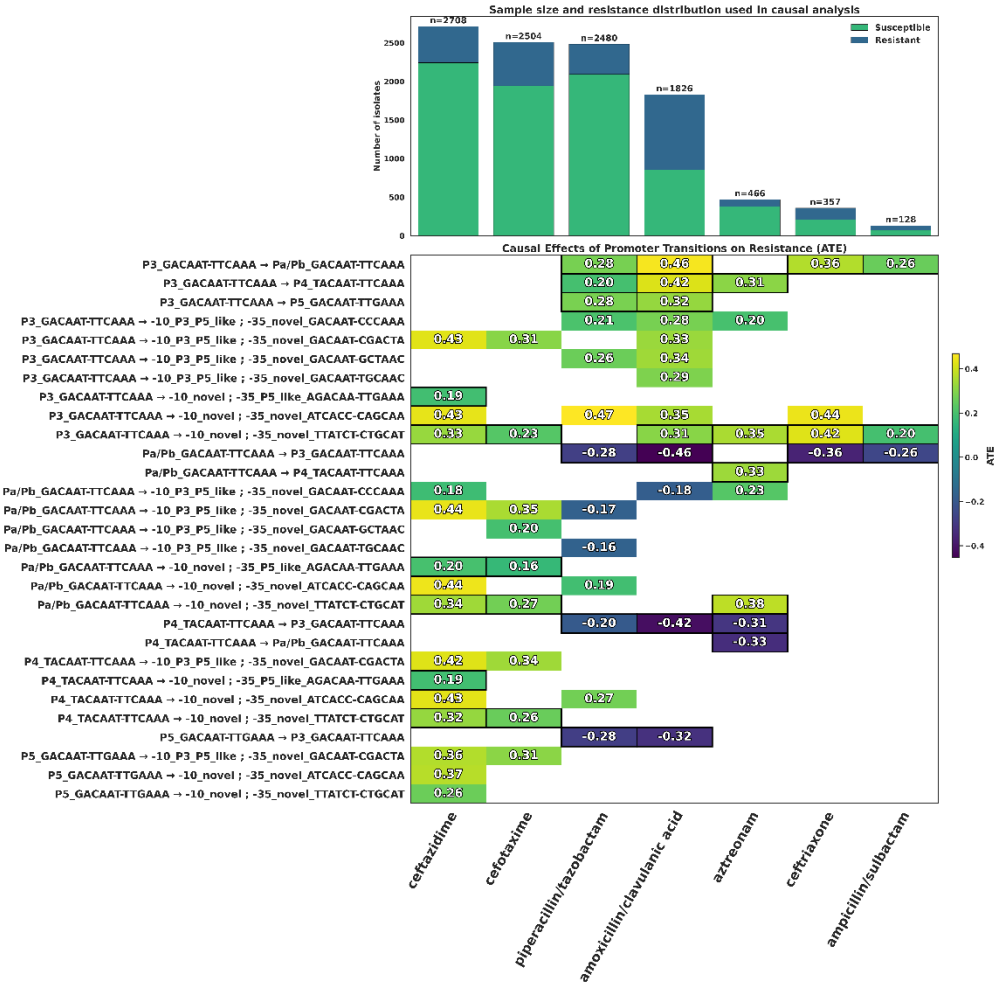



**Fig. S5.** Counterfactual effects of *bla*$_{\text{TEM}}$ promoter-motif transitions on antimicrobial resistance estimated using average treatment effects (ATEs). The lower heatmap shows the ATEs for ordered transitions between individual *bla*$_{\text{TEM}}$ promoter motifs across the indicated β-lactam antibiotics, estimated using the counterfactual framework while holding the observed TEM variant profile constant. To specifically examine the resistance-associated effects of alternative promoter motifs relative to established canonical promoter backgrounds, the heatmap focuses on transitions originating from canonical promoter motifs and includes both canonical-to-canonical and canonical-to-novel transitions. Each row represents a transition from the promoter motif shown before the arrow to the promoter motif shown after the arrow, with the corresponding -10 and -35 motif sequences indicated in the labels. Higher positive ATE values, represented toward green–yellow, indicate an increase in model-predicted resistance probability under the second promoter-motif condition, whereas negative ATE values, represented toward blue–purple, indicate a decrease in predicted resistance probability. Only promoter motif–antibiotic transitions meeting the predefined criteria of ATE stability $\geq 0.80$ and permutation ($p \leq 0.05$) are displayed; all other combinations are left blank. Bold black borders indicate motif-antibiotic transitions with a non-zero pairwise overlap fraction, whereas displayed cells without borders lack empirical overlap between the compared promoter conditions and therefore represent model-based counterfactual extrapolations. The upper stacked bar chart shows the total number of isolates with available resistance phenotype information for each antibiotic and their phenotype distribution, with susceptible isolates shown in green and resistant isolates shown in blue.

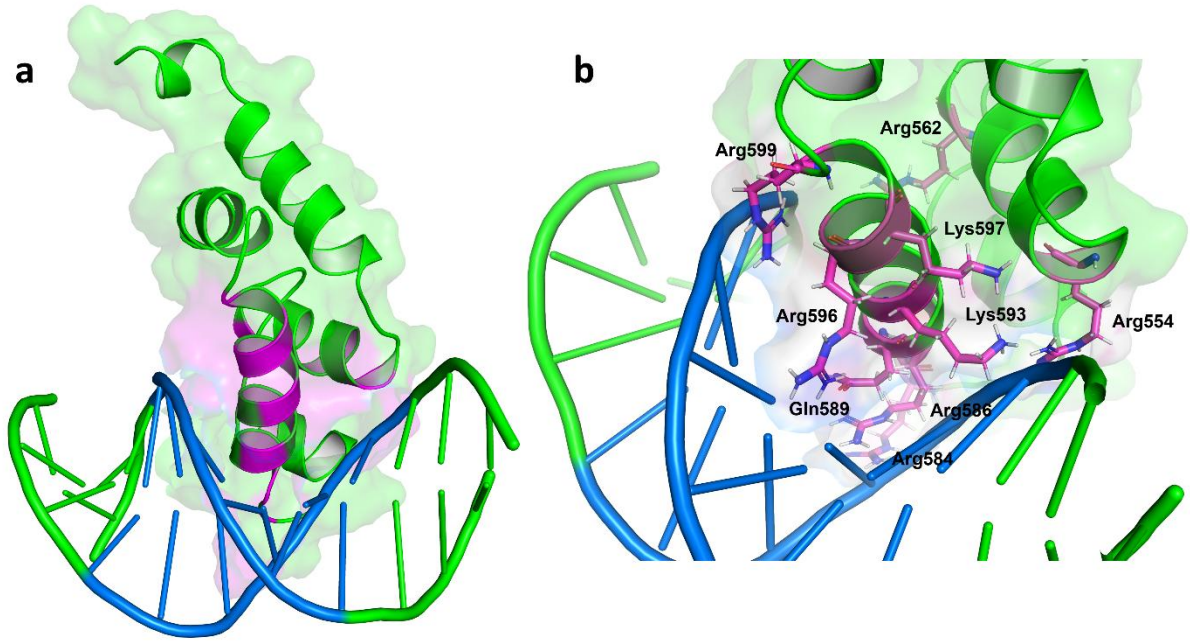



**Fig. S6.** Structural representation of the putative novel −35 motif M2 (TGCAAC) and its interactions with $\sigma^{70}$ Domain4. (a) Structural representation of $\sigma^{70}$ Domain 4 bound to the putative novel M2 motif (TGCAAC). The six nucleotides corresponding to the M2 motif are highlighted in blue, while $\sigma^{70}$ residues forming protein-DNA interactions with a trajectory occupancy of ≥10% during the 100-ns molecular dynamics simulation are highlighted in magenta. (b) Enlarged view of the interaction interface showing the nine $\sigma^{70}$ residues involved in interactions with the M2-containing DNA region over the simulation trajectory.