Received: 9 May 2020 | Revised: 16 September 2020 | Accepted: 17 September 2020

DOI: 10.1111/mmi.14614

RESEARCH ARTICLE

WILEY

A recently isolated human commensal *Escherichia coli* ST10 clone member mediates enhanced thermotolerance and tetrathionate respiration on a P1 phage-derived IncY plasmid

Shady Mansour Kamal [1,2] | Annika Cimdins-Ahne [3] | Changhan Lee [1] | Fengyang Li [1] | Alberto J. Martín-Rodríguez [1] | Zaira Seferbekova [4,5] | Robert Afasizhev [4] | Haleluya Tesfaye Wami [3] | Panagiotis Katikaridis [6] | Lena Meins [6] | Heinrich Lünsdorf [7] | Ulrich Dobrindt [3] | Axel Mogk [6] | Ute Römling [1]

[1]Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden

[2]Department of Microbiology and Immunology, Faculty of Pharmaceutical Sciences & Pharmaceutical Industries, Future University in Egypt, Cairo, Egypt

[3]Institute of Hygiene, University of Münster, Münster, Germany

[4]Kharkevich Institute for Information Transmission Problems, RAS, Moscow, Russia

[5]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, Russia

[6]Center for Molecular Biology, University of Heidelberg (ZMBH), German Cancer Research Center (DKFZ), DKFZ-ZMBH Alliance, Heidelberg, Germany

[7]Helmholtz Centre for Infection Research, Braunschweig, Germany

Correspondence
Ute Römling, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden.
Email: Ute.Romling@ki.se

Present address
Changhan Lee, Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, MI, USA

Funding information
RSF, Grant/Award Number: 18-14-00358; Swedish Research Council for Natural Sciences and Engineering, Grant/Award Number: 621-2013-4809 and 2017-04465; Deutsche Forschungsgemeinschaft, Grant/Award Number: MO970/4-3

Abstract

The ubiquitous human commensal *Escherichia coli* has been well investigated through its model representative *E. coli* K-12. In this work, we initially characterized *E. coli* Fec10, a recently isolated human commensal strain of phylogroup A/sequence type ST10. Compared to *E. coli* K-12, the 4.88 Mbp Fec10 genome is characterized by distinct single-nucleotide polymorphisms and acquisition of genomic islands. In addition, *E. coli* Fec10 possesses a 155.86 kbp IncY plasmid, a composite element based on phage P1. pFec10 harbours multiple cargo genes such as coding for a tetrathionate reductase and its corresponding regulatory two-component system. Among the cargo genes is also the Transmissible Locus of Protein Quality Control (TLPQC), which mediates tolerance to lethal temperatures in bacteria. The disaggregase $ClpG_{GI}$ of TLPQC constitutes a major determinant of the thermotolerance of *E. coli* Fec10. We confirmed stand-alone disaggregation activity, but observed distinct biochemical characteristics of $ClpG_{GI\text{-}Fec10}$ compared to the nearly identical *Pseudomonas aeruginosa* $ClpG_{GI\text{-}SG17M.}$ Furthermore, we noted a unique contribution of $ClpG_{GI\text{-}Fec10}$ to the exquisite thermotolerance of *E. coli* Fec10, suggesting functional differences between both disaggregases in vivo. Detection of thermotolerance in 10% of human commensal *E. coli* isolates hints to the successful establishment of food-borne heat-resistant strains in the human gut.

KEYWORDS
disaggregase ClpG, *Escherichia coli*, IncY plasmid, phylogenetic analysis, tetrathionate respiration, thermotolerance

1 | INTRODUCTION

Escherichia coli colonizes the gastrointestinal tract of almost every human in low numbers as the most predominant commensal facultative anaerobic bacterium. *E. coli* conquers the gastrointestinal tract soon after birth (Bettelheim and Lennox-King, 1976), which contributes to the early stimulation of the immune system, strengthening of the epithelial barrier function, production of vitamin B12 and K, and provides the physiological basis for the growth of anaerobes (Blount, 2015). However, the human gastrointestinal tract is only one of the habitats of *E. coli*. Members of the highly diverse species *E. coli*, which consists of *E. coli* sensu stricto and several cryptic *E. coli* clades are found in diverse ecological niches including the gastrointestinal tract of wild and domestic animals, sewage, slurry, plants, and the environment (Jorgensen et al., 2017; Meric et al., 2013; Tenaillon et al., 2010).

Present epidemiological approaches classify *E. coli* strains into seven phylogroups (Beghain et al., 2018; Clermont et al., 2000) and/or sequence types (ST) based on sequence variability in a restricted number of conserved loci that allow discrimination on the subspecies level (Manges et al., 2019). Epidemiological analyses indicate that commensal *E. coli* can be mainly classified into phylogroup A and B2, with the predominant phylogroups to widely shift temporally and among different populations (Lescat et al., 2013; Massot et al., 2016). Classification according to sequence types unraveled the ubiquitous occurrence of certain sequence types such as the multidrug-resistant ST131 in fecal and clinical samples, but also ST10 is found in the gastrointestinal tract of humans, animals, in association with plants and in the environment (Day et al., 2019; Dublan Mde et al., 2014; Freitag et al., 2018; Reid et al., 2017). Other sequence types are more restricted (Manges et al., 2019). Genome sequencing revealed that intestinal and extraintestinal *E. coli* pathogens can be closely related to innocuous commensal *E. coli* strains (Cimdins et al., 2017b). Indeed, enterotoxigenic *E. coli* can readily arise from commensal strains by the acquisition of few genetic elements including the transfer of characteristic virulence plasmids (von Mentzer et al., 2014), indicating that plasmids can readily alter the characteristics of strains.

Plasmids can evolve from phages (Sternberg and Hoess, 1983; Venturini et al., 2019), some of which exist as circular plasmids in the lysogenic state. The contribution of plasmids toward genomic plasticity reaches beyond the provision of pathogenicity factors, antimicrobial resistance determinants, and degradation pathways (Koraimann, 2018; Pilla and Tang, 2018; Shintani et al., 2010). One plasmid-encoded feature is the up to 19 kbp long Transmissible Locus of Protein Quality Control (TLPQC) which mediates elevated tolerance against lethal temperature, pressure, and oxidatives (Boll et al., 2017; Lee et al., 2016; Li and Ganzle, 2016; Li et al., 2020; Wang et al., 2020). The TLPQC locus, originally derived from an environmental species, has also been obtained by clinically relevant bacteria (Bojer et al., 2012). The three core genes *dna-shsp20*$_{GI}$*-clpG*$_{GI}$ encoding a transcription factor, a holdase chaperone and a disaggregase are the most conserved loci of TLPQC. ClpG disaggregases with the horizontally transferred ClpG$_{GI}$ as a distinct subgroup constitute stand-alone disaggregases with high intrinsic ATPase activity determinative for the tolerance phenotype to lethal temperatures in *Klebsiella pneumoniae*, *P. aeruginosa*, and *E. coli* (Bojer et al., 2010; Lee et al., 2018). ClpG members surpass the disaggregation activity of the canonical Hsp70 (DnaK)/ClpB bi-chaperone disaggregase, which mediates basal thermotolerance in bacteria (Katikaridis et al., 2019).

The commensal *E. coli* K-12 strain, isolated in 1921, has been the genetic reference strain since then (Bachmann, 1972). Sequence type 10 *E. coli* K-12 is closely related to NCTC86, the *E. coli* strain originally identified by Theodor Escherich in 1886 and has been considered as a typical commensal strain (Khetrapal et al., 2017). In this work, we initially characterized the biofilm-forming commensal strain Fec10, a member of the ST10 clonal complex, closely related to *E. coli* K-12 and NCTC86, recently isolated from a healthy human being (Bokranz et al., 2005). Fec10 contains the phage P1-derived IncY plasmid pFec10 bearing distinct cargo gene sequences, which indicate the adaptation of Fec10 to adverse growth conditions. One of the distinct features of pFec10 is the acquisition of an operon for tetrathionate respiration, which is involved in the regulation of rdar biofilm expression in the presence of the electron acceptor. Another cargo gene cluster is the TLPQC locus that provides enhanced tolerance to severe lethal temperature stress. We confirm autonomous disaggregation activity of thermotolerance mediating ClpG$_{GI}$ and show that ClpG$_{GI}$ withstands higher temperatures than the widespread ClpB/DnaK bi-chaperone system in agreement with the overall exquisite temperature-tolerance phenotype of *E. coli* Fec10 and other commensal *E. coli* isolates.

2 | RESULTS

2.1 | The commensal isolate Fec10 is a modern member of the *E. coli* K-12 clonal group

Recently, we isolated and characterized commensal *E. coli* strains from the gastrointestinal tract of healthy humans (Bokranz et al., 2005). Thereby, initial phylogenetic analysis identified our internal reference strain ST10 *E. coli* Fec10 as closely related to *E. coli* K-12 (Cimdins et al., 2017b). As a first step toward a more detailed phylogenetic analysis of Fec10, we determined its genome sequence by PacBio RSII sequencing combined with Illumina sequencing. Subsequently, single-nucleotide polymorphism (SNP) analysis of

the core genomes of 521 fully sequenced and representative *E. coli* strains from different phylotypes including *E. coli* K-12 derivatives such as the MG1655 reference strain; the Theodor Escherich isolate NCTC86; a recently identified Swine-derived ETEC strain closely related to K-12 (Shepard et al., 2012) and strains most closely related to *E. coli* K-12 and Fec10 in the NCBI and NCTC database including strains from the ECOR collection demonstrated that Fec10, indeed, is closely related to *E. coli* K-12 (Figure 1; Figure S1a,b; see below). Unlike *E. coli* K-12, Fec10 displays a full red, dry, and rough (rdar) morphotype, a distinct biofilm colony phenotype characterized by the production of the exopolysaccharide cellulose and amyloid curli fimbriae as extracellular matrix components at ambient temperature (Figure S2; Cimdins et al., 2017b). Fec10 is a member of the ST10 strain complex, which opens up the opportunity to document adaptation of a member of the abundant phylogroup A/ST10 *E. coli* subgroup to post-industrial human beings.

2.2 | Description of Fec10 genome features

The genome of strain Fec10 is composed of one circular chromosome of 4.88 Mbp in size and at least two circular entities, a plasmid (see below), and a pro483-like *E. coli* phage (NC_028943). The chromosome of *E. coli* Fec10 is colinear with the *E. coli* K-12 MG1655 reference genome (Figure S1b). With in total 4,828 (and 4,598 chromosomal) annotated coding sequences, Fec10 has 3,926 (85.4%) coding sequences (CDS) in synteny with *E. coli* K-12 MG1655. Comparing Fec10 with six *E. coli* K-12 derivatives (whereby each isolate differs in gene composition), Fec10 has 3,853 core, 975 variable, and 636 strain-specific coding sequences. Common regions of *E. coli* Fec10 and closely related NCTC86 (Figure 1, S1 a,b) comprise ~ 80% of the genomes, the level of sequence identity being 95%, and the vertically inherited fraction 86%. These numbers are similar when Fec10 is compared to other closely related strains such as VR50, FORC 064, NCTC9102, RR1, AG100, D4, D9, TO73, 26,561 with the common genome parts ranging from 77% to 84%, the sequence identity from 95% to 96%, and the vertically inherited the fraction of the common genome from 86% to 90%. With respect to the *E. coli* population (1,080 genomes), Fec10 comprises a core genome of 1,321 (27.4%) CDSs and 3,567 (72.6%) variable CDSs (Vallenet et al., 2020). Among the acquired distinct genomic islands and regions of plasticity are several phage-like genomic islands, but also the yersiniabactin operon, designated "high pathogenicity island", which is a virulence determinant in several enterobacterial genera (Figure S1b).

2.3 | *E. coli* Fec10 harbors a phage P1-derived IncY plasmid with the TLPQC locus

E. coli Fec10 harbors a 155.9 kbp plasmid (termed pFec10) with an IncY origin of replication (Figure 2; Table S1). The rare IncY replicon originates from the P1 phage (Sternberg and Hoess, 1983). Several IncY/P1 phage derivatives have recently been described (Venturini et al., 2019). Indeed, BLAST search with standard parameters indicated 25 and 15% query cover by *E. coli* phage P1 (NC_005856.1) and *Salmonella* phage SJ46 (NC_031129.1), respectively, which

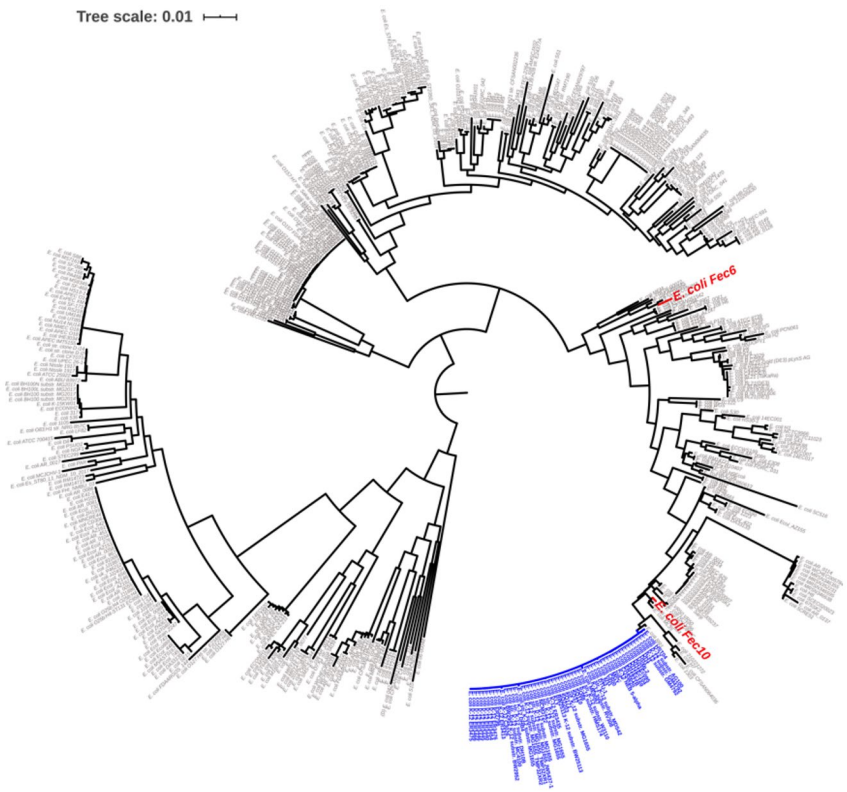

FIGURE 1 Unrooted phylogenetic tree of the species *E. coli*. The tree is constructed with 521 complete *E. coli* genomes using 238 orthologous genes. The two commensal *E. coli* strains (Fec10 and Fec6) are shown in red, and the cluster of *E. coli* K-12 strains is shown in blue. Tree visualized with iTOL https://itol.embl.de/ (Letunic and Bork 2016)

thus cover a significant part of the plasmid backbone (Figure S1c; Table S1; Altschul et al., 1990; Arndt et al., 2016). However, major head and sheat proteins are missing, making the assembly of a functional phage unlikely.

The closest plasmid homolog of pFec10 is pSSE which is a multireplicon IncHI2/IncHI2A plasmid from the thermo-resistant *Salmonella enterica* subsp. *enterica* serovar Senftenberg strain 775W (ATCC42845) isolated from Chinese egg powder in 1941. pFec10 is 99.96% identical over 50% coverage to pSSE, partly due to the presence of the highly conserved TLPQC locus present on pFec10 and pSEE, which harbors two distinct TLPQC loci (Figure S1d, S3; Table S1; (Nguyen et al., 2017)).

Assessment of characteristic elements for plasmid maintenance showed that the pFec10 plasmid does not encode genetic elements that mediate antibiotic resistance (Hendriksen et al., 2019). Plasmid stability is most likely maintained by the type II toxin-anti-toxin system YdeE-YdeD (Prevent host death-death on curing (Phd-Doc)), whereby *ydeD* encodes a "death on curing" toxin (Lehnherr et al., 1993; Liu et al., 2008). Death on curing toxins arrest elongation factor EF-Tu in a non-functional state by phosphorylation. Within the pFec10 sequence, there is also a region representing an origin of transfer (Figure 2).

2.4 | pFec10 is an unconventional plasmid

Distinct features of pFec10 include a multitude of distinct IS elements (Figure 2; Table S1, S2). In addition, a variety of cargo gene products are encoded on pFec10. With 11,586 bp, the longest open reading frame encodes a 3,862 amino acid long tandem-95-repeat protein with the closest homolog (50.8% identity) in *Butticuxella* spp. Another distinct feature is the presence of a type I restriction-modification system with the modification, sensitivity, and restriction subunit. Besides the TLPQC locus involved in stress resistance, cargo gene modules, for example, for detoxification, nutrient acquisition and redox balance are encoded on pFec10. For example, the plasmid codes for genes involved in arsenate resistance, a DyP-type decolorizing peroxidase and its encapsulating protein, a lactate utilization gene cluster, and a tetrathionate reductase operon and its respective sensing/regulatory two-component system (Figure 2; Table S1). Those elements are embedded into the phage-derived plasmid scaffold (Figure 2, Table S1; (Arndt et al., 2016)).

The plasmid encodes a DyP-decolorizing peroxidase unit involved in detoxification. DyP decolorizing peroxidases of fungi and bacteria have an unknown physiological function, but wide biotechnological applicability as they degrade lignin and oxidize a variety of aromatic compounds, dyes, and other small molecules (Lin et al., 2019; Singh et al., 2013). A 32% identical DyP homolog is encoded on the Fec10 and *E. coli* K-12 chromosome. Furthermore, an encapsulin, which encapsulates peroxidases via a recognized C-terminus in nanocompartments, is encoded downstream of the peroxidase (Tracey et al., 2019) and a NAD-dependent formate dehydrogenase is encoded upstream.

As another element of detoxification, the plasmid encodes two gene clusters involved in arsenate reduction. These newly arranged gene clusters encode gene products closely related to gene products from the Fec10 and *E. coli* K-12 core genomes, but also gene products with no close homologs. The four-gene cluster encodes a protein with partial homology to the ArsA efflux transporter ATPase subunit, an ArsC reductase, an ArsF efflux pump membrane protein, and an ArsA efflux transporter ATPase subunit. ArsC and ArsF have close homologs with >90% identity on the chromosome of Fec10 and *E. coli* K-12. The three-gene cluster encodes an ArsR transcriptional regulator, an ArsD transacting repressor, and another ArsA

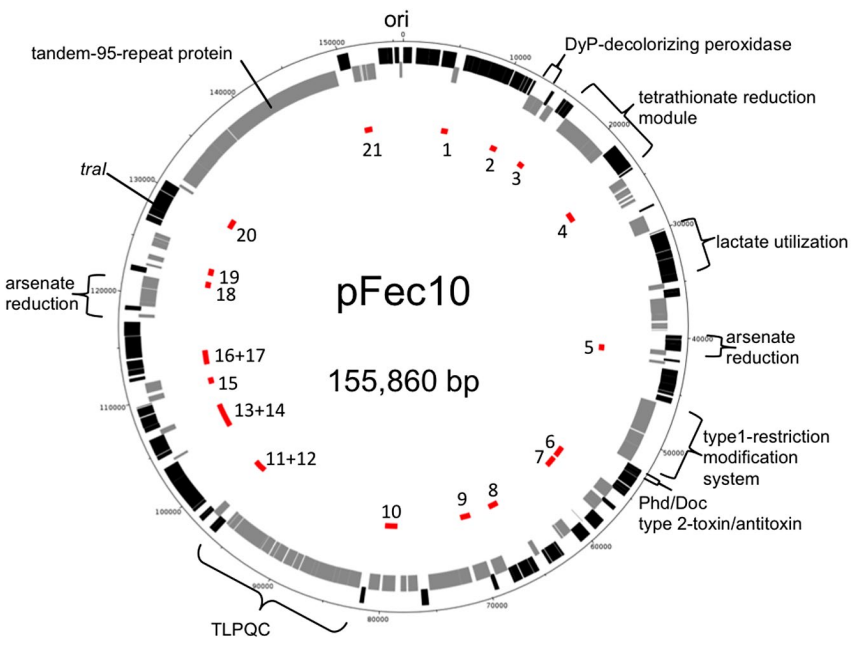

FIGURE 2 Plasmid map of pFec10. Coding regions are indicated in black (sense strand) or grey (anti-sense strand). Origin of replication (ori) and major cargo genes mentioned in the text are highlighted. TLPQC, transmissible locus of protein quality control. IS element as identified by ISfinder (http://www-is-biotoul.fr; (Siguier et al., 2006)) is highlighted in red. Numbers refer to IS elements as listed in Table S1 and Table S2. The figure was created by DNAPlotter and manually modified

efflux transporter ATPase subunit. Only the ArsR regulator possesses a 74.1% identical homolog on the core genome.

Furthermore, several components required for D- and L-Lactate catabolism are encoded on pFec10. The plasmid encodes an LctP L-lactate permease, which has no counterpart on the chromosome of Fec10 and *E. coli* K-12 (MG1655 strain). Moreover, a gene of the D-lactate dehydrogenase Dld is present with a >85% identity counterpart in Fec10 and K-12. Adjacently located is the *ykgEFG* gene cluster with a corresponding highly similar *ykgEFG* operon in Fec10 and K-12. For example, the iron-sulphur oxidoreductase subunit YkgE of the L-lactate dehydrogenase shows 94% identity. *YkgEFG* is involved in L-lactate catabolism and biofilm formation in *Bacillus subtilis* (Chai et al., 2009), but has no L-lactate catabolism phenotype in *E. coli* K-12, as an alternative operon for lactate catabolism is present (Chai et al., 2009). Therefore, the evolutionary driving force for the acquisition of a highly homologous second copy of *ykgEFG* remains to be determined.

The species *E. coli* usually does not encode the genetic information to respire tetrathionate. However, the BLAST search indicates that an almost identical tetrathionate operon is present in selected *E. coli* and *Klebsiella pneumoniae* strains (data not shown), suggesting that the acquisition of the tetrathionate operon by a commensal *E. coli* strain aids adaptation to altered conditions in the gastrointestinal tract. The tetrathionate operon might have originated from *Citrobacter* spp. (homology can be as high as >99%), as tetrathionate reduction is a characteristic feature of, for example, the species *Citrobacter freundii* (Kapralek, 1972). In order to get the first insight into the biological impact and physiological contribution of the plasmid under certain environmental conditions, we chose to initially investigate the biological consequences of two cargo operons on the plasmid, the tetrathionate operon and the core *dna-shsp20*$_{GI}$-*clpG*$_{GI}$ operon of the TLPQC locus for elevated temperature tolerance.

2.5 | A Biofilm phenotype is associated with tetrathionate sensing and respiration

Tetrathionate is used as an alternative electron acceptor by *Salmonella typhimurium* (Hensel et al., 1999; Winter and Baumler, 2011), but not *E. coli*. A combination of tetrathionate versus thiosulfate, the product of tetrathionate respiration, is even toxic for *E. coli* (Palumbo and Alford, 1970). We have recently shown that the reduction of alternative electron acceptors triggers alternations in the amount of biofilm formation in bacteria (Martín-Rodríguez et al., submitted; Martín-Rodríguez et al., 2020). To this end, we deleted the tetrathionate operon genes on the pFec10 plasmid. We constructed three different mutants; (1) a deletion mutant of *ttrSR* encoding the two-component system histidine kinase and response regulator required for tetrathionate sensing and activation of the reductase operon; (2) a deletion mutant of *ttrBCA* encoding the tetrathionate reductase in combination with a short uncharacterized upstream open reading frame of 144 bps encoding a hypothetical protein; (3) a mutant with the deletion of the two divergently transcribed operons; to test whether exposure to tetrathionate has any effect on the rdar biofilm formation (Figure S2e; (Römling et al., 1998b)). When grown on Congo Red (CR) agar plate to display biofilm formation, the mutants did not show a change in the rdar biofilm morphotype. We observed, however, a mutant and temperature-dependent effect of tetrathionate on colony morphology appearance when cells were grown on a CR agar plate with 1.5 and 15 mM tetrathionate (Figure 3). At 28°C, the rdar morphotype was diminished strongest upon the deletion of the *ttrBCA* operon. At 37°C, the colony appearance of the wild type turned pale at 15 mM tetrathionate, while the deletion of the *ttrBCA* operon led to an upregulation of the rdar morphotype. Thus, tetrathionate has a differential temperature-dependent effect on colony morphology.

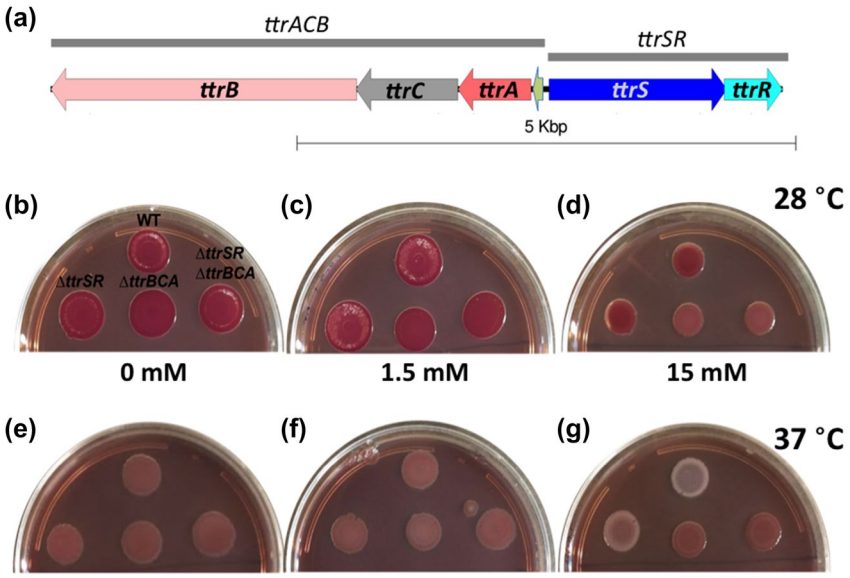

FIGURE 3 Effect of tetrathionate supplementation on the rdar biofilm colony morphology of strain *E. coli* Fec10 on CR agar plates. (A) Arrangement of the tetrathionate gene cluster with the constructed mutants indicated. *E. coli* Fec10 WT, Δ*ttrSR*, Δ*ttrBCA*, and Δ*ttrSR* Δ*ttrBCA* colony biofilms after incubation on CR agar plates supplemented with sodium tetrathionate (1.5 or 15 mM) for 48 h at 28°C (B-D) or 37°C (E-G) [Colour figure can be viewed at wileyonlinelibrary.com]

2.6 | *ClpG$_{GI}$* and *dna-hsp20$_{GI}$-clpG$_{GI}$* mediate elevated tolerance to lethal temperature

High tolerance to lethal heat treatment has been associated with the TLPQC locus (alternatively called LHR locus) in various species with the three-gene operon *dna-shsp20$_{GI}$-clpG$_{GI}$* to play a determinative role in extended heat tolerance (Bojer et al., 2010; Lee et al., 2015). In particular, horizontally transferred ClpG$_{GI}$ (ClpK) has been shown to mediate increased tolerance to lethal temperature in *K. pneumoniae*, *P. aeruginosa*, and food-derived and ESBL clinical *E. coli* isolates (Bojer et al., 2011; Boll et al., 2017; Lee et al., 2018). In order to investigate whether and to which extent *clpG$_{GI}$* and/or the entire *dna-shsp20$_{GI}$-clpG$_{GI}$* operon contribute to tolerance against lethal temperatures, we deleted solely *clpG$_{GI}$* and the entire *dna-shsp20$_{GI}$-clpG$_{GI}$* operon in the commensal *E. coli* strain Fec10. Deletion of *clpG$_{GI}$* and *dna-shsp20$_{GI}$-clpG$_{GI}$* dramatically reduced the tolerance to lethal heat treatment >3-fold $\log_{10}$ upon exposure to 60°C for 15 min with a major contribution derived from *clpG$_{GI}$*, whereas the wild-type Fec10 remained almost unaffected (Figures 4a, S4). Similarly, exposure to 65°C for 2 min reduced survival of both deletion mutants >4-fold $\log_{10}$ with Fec10 wild type not affected (Figure S4). The loss of tolerance to lethal heat treatment could, however, be partially restored by the provision of *clpG$_{GI}$* and *dna-shsp20$_{GI}$-clpG$_{GI}$* on plasmid pBAD30 under the control of the L-arabinose inducible P_{BAD} promoter (Figure 4). These results show that in commensal isolates of *E. coli* provision of tolerance to elevated heat shock is largely mediated by the *dna-shsp20$_{GI}$-clpG$_{GI}$* operon with *clpG$_{GI}$* constituting the major determining factor.

In *P. aeruginosa* clone C, gene products of the chromosomal *dna-shsp20$_{GI}$-clpG$_{GI}$* locus are predominantly expressed in the stationary phase of growth (Lee et al., 2015, 2018). Investigation of the ClpG$_{GI}$ production of *E. coli* Fec10 in the logarithmic and stationary phase of growth at 37°C revealed production of the protein in log phase cells with ClpG$_{GI}$ levels to increase during the stationary phase (Figure 4b). Notably, the canonical disaggregase ClpB is expressed at detectable levels in Fec10 at this temperature mainly in the stationary phase of growth.

2.7 | *ClpG$_{GI\text{-}Fec10}$* can complement *clpB* and *dnaK103* mutants in *E. coli* MC4100

We have previously shown that *P. aeruginosa* ClpG$_{GI\text{-}SG17M}$ expressed from a plasmid restores thermotolerance to *E. coli* MC4100 Δ*clpB* and *dnaK103* mutants, lacking individual components of the canonical bi-chaperone disaggregase (Lee et al., 2018). ClpG$_{GI\text{-}Fec10}$ is 95.8% identical to ClpG$_{GI\text{-}SG17M}$ with a conserved N-terminal domain and two highly conserved AAA + ATPase domains containing conserved catalytic motifs such as the Walker A/Walker B ATP binding motif, while amino acid changes accumulate also at the C-terminal end of the N2 domain and the highly charged C-terminus (Figures S5, S6; Lee et al., 2018). ClpG$_{GI\text{-}Fec10}$ restored heat tolerance at 50°C to the same extent as ClpG$_{GI\text{-}SG17M}$ in MC4100 *clpB* and *dnaK103* mutants. Both ClpG$_{GI}$ variants were superior as compared to plasmid-expressed *clpB*, confirming that enhanced heat resistance can be transferred to other bacteria by the sole expression of *clpG$_{GI}$* (Figure 4c). Expression of *clpG$_{GI\text{-}Fec10}$* also effectively rescued *dnaK103* null mutants, though lack of the central Hsp70 chaperone causes massive protein aggregation upon heat stress (Figure 4c) (Mogk et al., 1999). This indicates that *clpG$_{GI}$* covers and expands the functionality of *dnaK* and *clpB* in tolerance to lethal temperatures.

2.8 | ClpG$_{GI\text{-}Fec10}$ is a stand-alone disaggregase with high intrinsic ATPase activity

The complementation experiments indicated that ClpG$_{GI\text{-}Fec10}$ has similar functionality as ClpG$_{GI\text{-}SG17M}$. To assess the activity of ClpG$_{GI\text{-}Fec10}$ against model substrates, we purified ClpG$_{GI\text{-}Fec10}$ according to previously described protocols (Figure S7; Lee et al., 2018). To this end, the end-point assessment of the restoration of the catalytic activity of heat-aggregated luciferase indicated that both, ClpG$_{GI\text{-}Fec10}$ and ClpG$_{GI\text{-}SG17M}$, recovered >30% of luciferase activity after 120 min, while the 15% recovery rate for ClpB was more than 50% lower (Figure 5). Residual disaggregase activity was observed for the DnaK chaperone system (KJE: including the DnaJ, GrpE cochaperones), which also competes with ClpG$_{GI\text{-}Fec10}$ for substrate binding, as 1:1 co-incubation of DnaKJE with ClpG$_{GI\text{-}Fec10}$ abolished luciferase recovery above residual levels (Figures 5, S8). As expected, ClpB alone did not show any activity.

Similarly, we monitored the recovery of catalytic activity of the model substrate malate dehydrogenase (MDH) (Figure S8). In this case, the chaperonin complex GroES/EL was added to aid the refolding of the disaggregated amino acid chain (Lee et al., 2018). As previously reported, we observed that the end-point recovery rate was highest for the ubiquitous ClpB/DnaKJE system with >80% MDH recovered activity after 120 min. ClpG$_{GI\text{-}Fec10}$, and to a lesser extent ClpG$_{GI\text{-}SG17M}$, showed a trend to display a reduced recovery rate (Figure S8). Lack of disaggregase activity was observed for ClpB and the DnaKJE system, which competed with ClpG$_{GI\text{-}Fec10}$ for substrate binding, as 1:1 co-incubation of DnaKJE with ClpG$_{GI\text{-}Fec10}$ abolished recovery of MDH activity (Figure S8).

Interaction of ClpB with its aggregate-loaded DnaK co-chaperon does not only deliver aggregates to ClpB, but also activates the ATPase activity of ClpB dedicated to disaggregation (Zolkiewski, 1999). We have recently shown that stand-alone ClpG$_{GI\text{-}SG17M}$ has high intrinsic basal ATPase activity, >11-fold higher than the basal activity of ClpB (Figure 5C; (Lee et al., 2018)). The estimated basal ATPase activity of ClpG$_{GI\text{-}Fec10}$ was >15-fold higher than ClpB basal activity and even significantly higher as compared to ClpG$_{GI\text{-}SG17M}$. Addition of the ClpB substrate casein stimulated ATPase activity >6.5-fold, while the ATPase activity of ClpG$_{GI\text{-}Fec10}$ and ClpG$_{GI\text{-}SG17M}$ was inhibited 2.4-fold and >1.2-fold, respectively. This suggests differences in ATPase regulation and substrate specificities of ClpB and ClpG$_{GI}$ disaggregases, but also between different ClpG$_{GI}$ disaggregases. Of note, ClpG$_{GI}$ protein

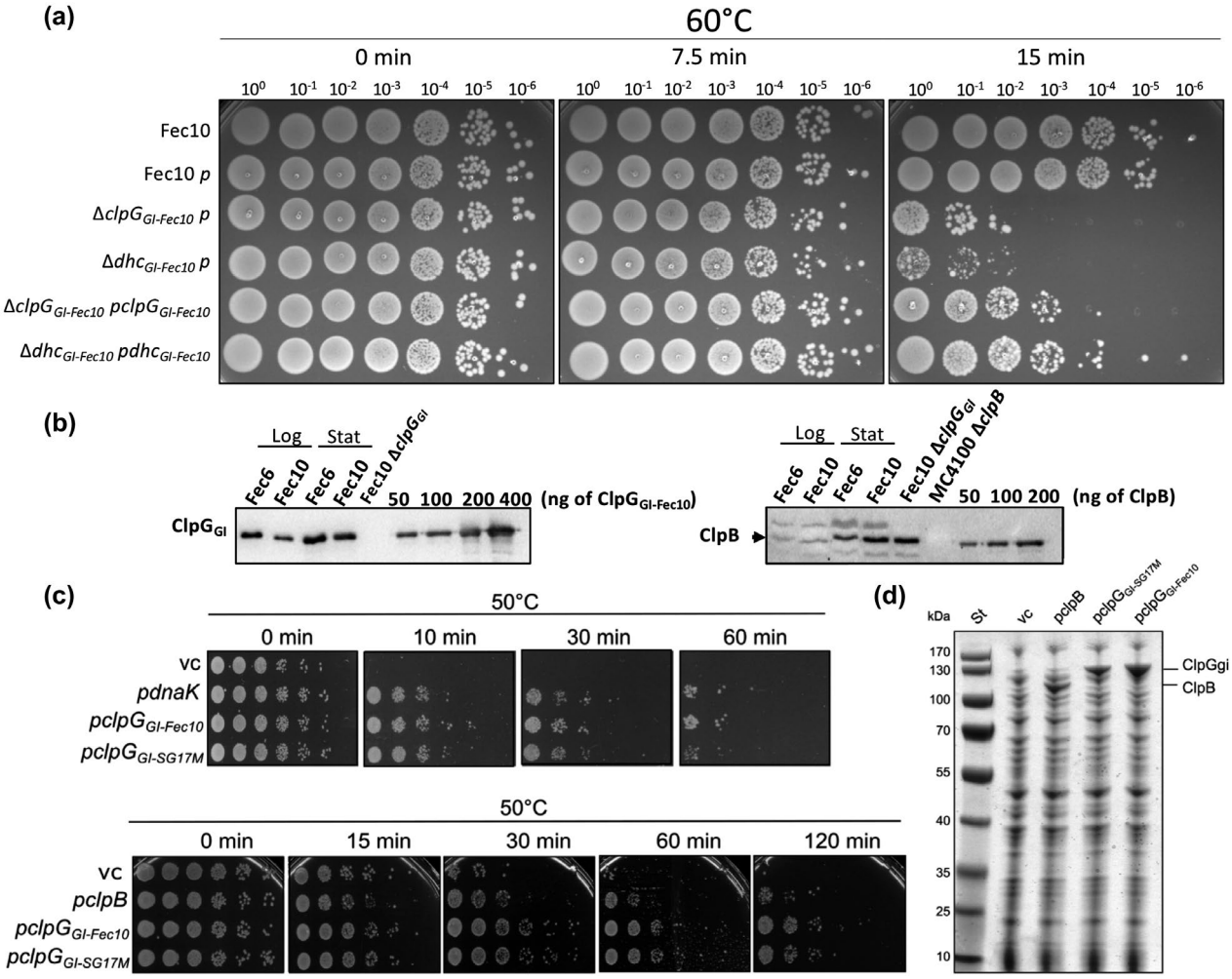

FIGURE 4 $ClpG_{GI\text{-}Fec10}$ contributes to heat tolerance in *E. coli* Fec10 and restores heat tolerance in *E. coli* MC4100 in the absence of *dnaK* and *clpB*. (A) Assessment of heat shock tolerance in the *E. coli* Fec10 $clpG_{GI\text{-}Fec10}$ and $dhc_{GI\text{-}Fec10}$ deletion mutants and respective complementation. Bacterial cells were exposed to 60°C for 7.5 and 15 min. p = plasmid pBAD30; p$clpG_{GI\text{-}Fec10}$ = $clpG_{GI\text{-}Fec10}$ cloned in plasmid pBAD30. p$dhc_{GI\text{-}Fec10}$ = *dna-shsp20*$_{GI}$-$clpG_{GI}$ cloned in plasmid pBAD30. (B) Western blot analysis of the production level of $ClpG_{GI}$ and ClpB in *E. coli* Fec10 and Fec6 in the logarithmic and stationary phase of growth at 37°C. Detection of protein production was performed using antisera generated against $ClpG_{GI\text{-}SG17M}$ and $ClpB_{K\text{-}12}$. (C) Complementation of the heat shock tolerance of the *E. coli* MC4100 deletion mutant of *dnaK* (upper panel) and *clpB* (lower panel) in MC4100 with plasmids expressing *dnaK*, *clpB*, $clpG_{GI\text{-}SG17M}$, and $clpG_{GI\text{-}Fec10}$. Cells were exposed to 50°C for the indicated amount of time. Vc = vector control pUHE21; p*dnaK* = *dnaK* cloned in pUHE21; p*clpB* = *clpB* cloned in pUHE21; p$clpG_{GI\text{-}Fec10}$ = $clpG_{GI\text{-}Fec10}$ cloned in pUHE21; p$clpG_{GI\text{-}SG17M}$ = $clpG_{GI\text{-}SG17M}$ cloned in pUHE21. (D) SDS-PAGE gel of the expression of ClpB, $ClpG_{GI\text{-}SG17M}$, and $ClpG_{GI\text{-}Fec10}$ from the complementation experiment in panel (C). The positions of the proteins on the gel are indicated

variability is grossly correlated with the phylogenetic position of the host organism (Figure S5). We assessed the differential binding of the disaggregases to fluorescein-labeled FITC-casein by anisotropy measurements (Figure 5D). In the presence of ATPγS ClpB bound with high affinity (K_d: 0.56 ± 0.14 μM) to the substrate, while binding of both $ClpG_{GI}$ disaggregases was much weaker and K_d values could not be calculated as the binding curves did not saturate.

2.9 | Thermal stability of $ClpG_{GI}$ and ClpB

The extreme heat tolerance phenotype of strain Fec10 (Figures 4; S4b), which is comparable to the highest heat-tolerant *E. coli* strain isolated after the artificial decontamination of meat from slaughterhouse strains (Dlusskaya et al., 2011) also raised the question about the thermodynamic stability of $ClpG_{GI}$ proteins. Circular dichroism (CD) is a straightforward approach to assess changes in the secondary structure of proteins upon temperature upshift as the CD spectra of α-helix, β-sheet and random coil differ substantially (Figure S9a). We, therefore, recorded the thermal transition of the protein structure by monitoring CD spectra of $ClpG_{GI\text{-}Fec10}$, $ClpG_{GI\text{-}SG17M}$, and ClpB, allowing to compare stabilities of stand-alone $ClpG_{GI}$ and canonical ClpB disaggregases. First, we determined the CD spectrum of $ClpG_{GI\text{-}Fec10}$, $ClpG_{GI\text{-}SG17M}$, and ClpB without and with ATPγS and Mg^{2+}. The CD spectrum of all three proteins without ATPγS and Mg^{2+} resembled an α-helical spectrum (Figure S9a,b). Measurement of the melting temperature of the secondary structure by monitoring the

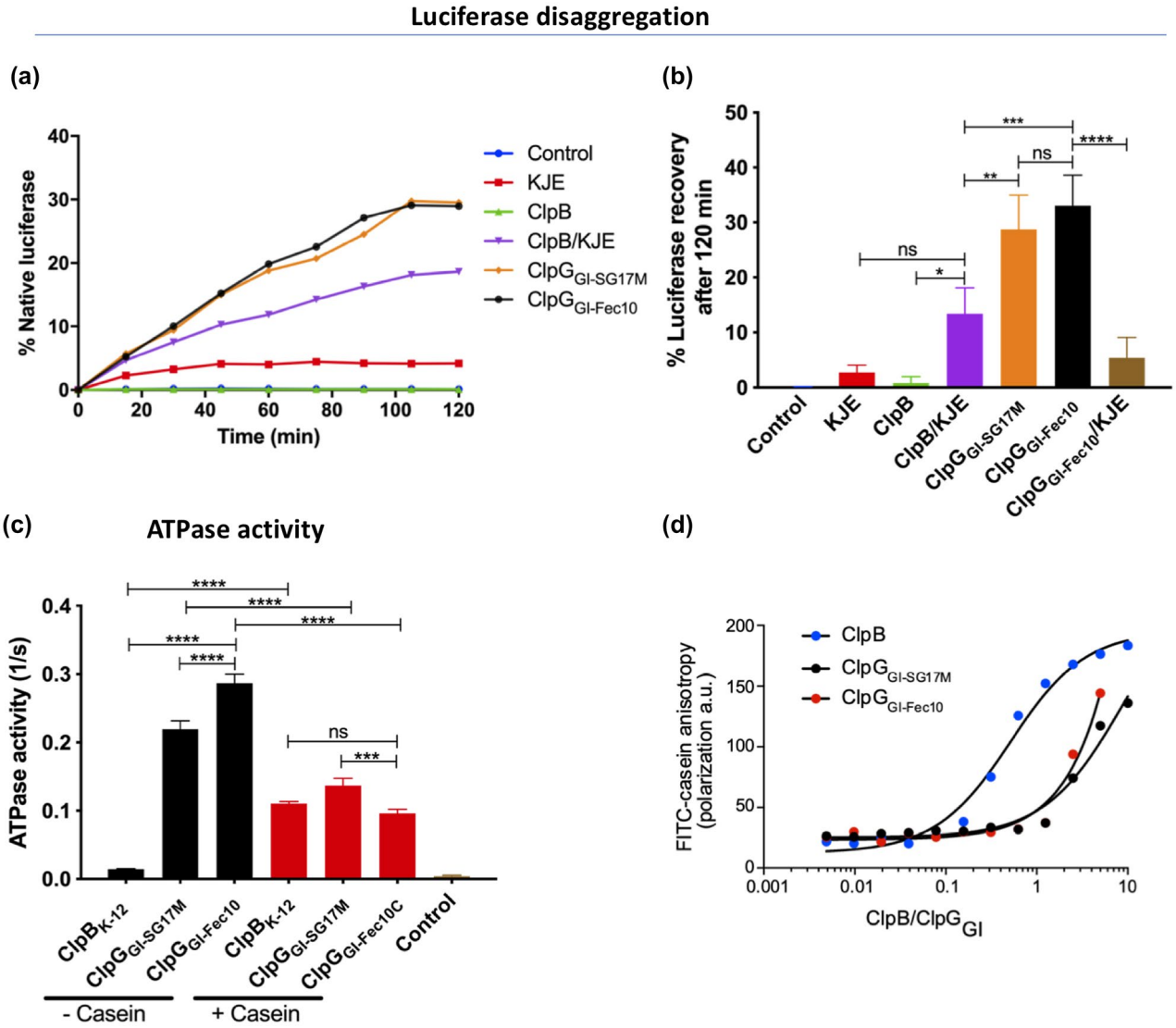

FIGURE 5 $ClpG_{GI\text{-}Fec10}$ disaggregates heat-folded luciferase independent of accessory chaperons and demonstrates high basal ATPase activity. (A) Refolding of aggregated luciferase was monitored over 120 min. The activity of the native luciferase was set to 100%. One representative experiment is shown. (B) Recovery of luciferase after 120 min incubation with the indicated disaggregase. The mean value was calculated from three independent experiments with nine technical replicates. Error bars indicate SD (****$p < .0001$). (C) ATPase rates were calculated with and without casein, the stimulator substrate for ClpB ATPase activity. The mean value was calculated from three independent experiments with nine technical replicates. Error bars indicate SD (****$p < .0001$). (D) Binding of ClpB, $ClpG_{GI\text{-}Fec10}$, and $ClpG_{GI\text{-}SG17M}$ to fluorescein-labeled FITC-casein was determined in the presence of 2 mM ATPγS by anisotropy measurements

change in the absorption of left and right circular polarized light at 222 nm showed that ClpB displayed a far lower melting temperature of 53.2°C compared to $ClpG_{GI\text{-}Fec10}$ and $ClpG_{GI\text{-}SG17M}$, which showed structural transition at 64.5 and 62.9°C, respectively (Figures 6, S9b).

We next determined the CD spectra and transition curves in the presence of 1 mM ATPγS and 5 mM Mg^{2+}, which increased melting temperatures for all proteins (e.g., ClpB: 61.9°C, $ClpG_{GI\text{-}Fec10}$: 69.0°C and $ClpG_{GI\text{-}SG17M}$: 69.8°C) (Figures 6, S9d,e). An increase in protein stability is gained by the energy derived from ATPγS binding and ATP-driven oligomerization. T_M values of $ClpG_{GI\text{-}Fec10}$ and $ClpG_{GI\text{-}SG17M}$ determined in the presence of nucleotide remained >7°C higher as compared to ClpB, demonstrating higher thermal stability under all conditions tested (Figure 6). Notably, the melting temperatures of ATP-bound $ClpG_{GI\text{-}Fec10}$ and $ClpG_{GI\text{-}SG17M}$ are close to 70°C which are in congruence with the viability of temperature tolerant *E. coli* upon exposure to lethal temperatures (Figure 4a, S4; Dlusskaya et al., 2011).

As both $ClpG_{GI\text{-}Fec10}$ and $ClpG_{GI\text{-}SG17M}$ show increased, but not identical thermostability in vitro compared to the ClpB disaggregase, we were wondering whether $ClpG_{GI\text{-}SG17M}$ and $ClpG_{GI\text{-}Fec10}$ show comparable performance in vivo upon extreme temperature exposure. To this end, we expressed $clpG_{GI\text{-}Fec10}$ and $clpG_{GI\text{-}SG17M}$ from plasmid pBAD30 in *E. coli* Fec10 $clpG_{GI}$ and $dna\text{-}shsp20_{GI}\text{-}clpG_{GI}$ mutants (Figure S10). Surprisingly, in contrast to complementation with $clpG_{GI\text{-}Fec10}$, we observed no complementation with $clpG_{GI\text{-}SG17M}$

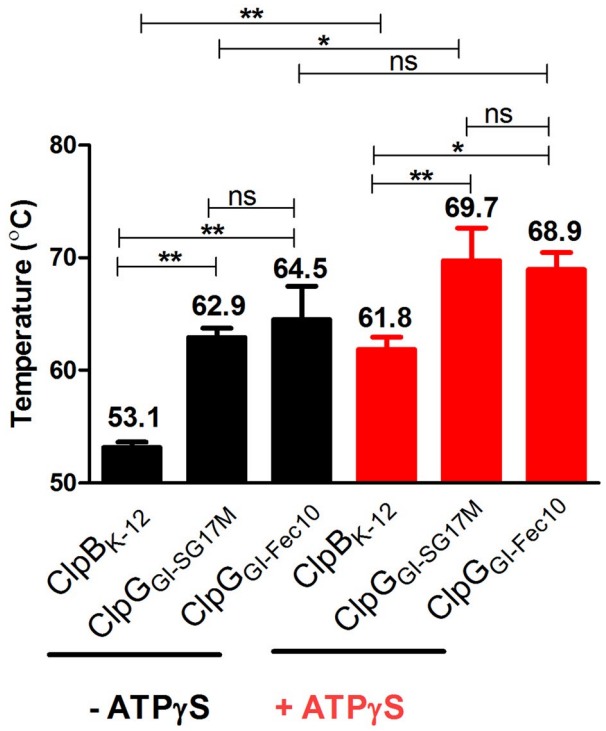

FIGURE 6 Determination of the melting temperature for $ClpB_{K-12}$, $ClpG_{GI-SG17M}$, and $ClpG_{GI-Fec10}$. Circular Dichroism melting temperature was recorded in the absence (black bars) and presence (red bars) of Mg^{2+}/ATPγS. The change in the differential absorption of the circular polarized light upon secondary structure alteration to random coil was recorded at 222 nm. The melting temperatures of the indicated enzymes are shown above the bars

upon exposure to a lethal temperature of 60°C. Lowering the temperature to 55°C, though, allowed the complementation of the *E. coli* Fec10 $clpG_{GI}$ mutant partially also by $ClpG_{GI-SG17M}$. Thus, how in vivo complementation capacity of both proteins is reflected by the in vitro biochemical characteristics needs to be sorted out. Of note, tolerance to lethal temperatures of *E. coli* Fec10 $clpG_{GI}$ and *dna-shsp-20*$_{GI}$-$clpG_{GI}$ is superior compared to *E. coli* K-12 and several investigated wild-type strains (Figure S4a, S10), suggesting the presence of additional factors that contribute to this phenotype.

2.10 | Commensal *E. coli* strains can show high tolerance to lethal temperatures

We were wondering whether gastrointestinal commensal strains commonly harbor the TLPQC locus and show resistance to lethal temperatures. We screened additional 31 representatives of human commensal *E. coli* strains, genetically unrelated as judged by pulsed-field gel electrophoresis (Bokranz et al., 2005), for the presence of the small heat shock protein 20 gene ($shsp20_{GI}$) as an indication for the TLPQC locus (Lee et al., 2015). In total, four out of 32 strains (12.5%), Fec6, Fec10, Fec32, and Fec89, contained the TLPQC locus (data not shown). Subsequently, we exposed those strains and four non-TLPQC bearing strains, Fec41, Fec55, Fec75, and Fec113, to a lethal temperature treatment at 55°C. Indeed, three out of four $shsp20_{GI}$-positive strains were more thermotolerant with less than 100-fold reduction in viability count after exposure to 55°C for 45 min (Figure S4a). Fec6 and Fec10 were the two most thermotolerant strains that substantially survived lethal heat at 60°C for 15 min (Figure 4a) and exposure to 65°C for 5 min (Figure S4b).

E. coli strains have previously been classified into seven major phylotypes (Beghain et al., 2018; Clermont et al., 2000). Thereby, commensal strains represent most commonly phylotype A or B2 depending on the investigated human population (Lescat et al., 2013; Massot et al., 2016). All *E. coli* strains, which showed an elevated lethal-heat-tolerance phenotype were phylotype A (Bokranz et al., 2005), an observation that has previously also been reported for food-derived heat-tolerant *E. coli* isolates.

2.11 | Two TLPQC loci are present in the high heat-tolerant strain *E. coli* Fec6

We determined the whole genome sequence of *E. coli* Fec6, which displayed an extraordinarily high tolerance to lethal temperature surviving 65°C for 5 min with no reduction in viable counts (Figure S4b). The Fec6 genome encodes two TLPQC loci integrated at different locations on the chromosome. TLPQC-1 of Fec6 and TLPQC of Fec10 are grossly identical, align to well-characterized TLPQC loci, and might represent the original form of the island (Figure S3). TLPQC-2 of Fec6 has acquired a number of cargo gene products that are inserted between a gene of unknown function and the *trx* thioredoxin gene (Figure S3). Truncated inserted genes are found downstream of *kefC* encoding a K^+/H^+ anti-transporter.

The TLPQC locus of *E. coli* Fec10 has 100% query coverage with >99.8% identity to the TLPQC locus of plasmid A of *E. coli* H15. TLPQC-1 of Fec6 has >99% query coverage and >99% identity in several *E. coli* genomes including P12b, NCTC11121, NCTC9966, H1, NCTC9040, and *E. coli* C (NCBI BLAST accessed October 7, 2019). TLPQC-2 of Fec6 has >90% query coverage and >98.8% identity with sequences from *E. coli* VREC0864, S43, and VREC0761. Rapid evolution of this locus is indicated as the island is found with >97% query coverage and >97.5% sequence identity in *E. coli* MEM (BLAST accessed October 7, 2019).

The $ClpG_{GI-Fec6}$ disaggregases of TLPQC-1 and TLPQC-2 are highly similar, but not identical to $ClpG_{GI-Fec10}$ and $ClpG_{GI}$s from other TLPQC islands (Figures S5, S6). We speculate differential substrate recognition and processing or temperature-dependent activity by $ClpG_{GI-Fec6-1}$ and $ClpG_{GI-Fec6-2}$ as preliminary indicated for $ClpG_{GI-Fec10}$ and $ClpG_{GI-SG17M}$ (Figure 5c), which might contribute to the extended temperature tolerance of strain Fec6. The diversified N- and C-termini eventually provide broadened substrate specificity or direct the disaggregases to specific cellular sites. As the total level of disaggregase is increased in *E. coli* Fec6 compared to Fec10 (Figure 4b), the presence of two $ClpG_{GI}$ copies in strain Fec6 can alternatively explain enhanced stress protection.

3 | DISCUSSION

Commensal *E. coli* strains, albeit present at low frequency, are found in the gastrointestinal tract of almost every human being as well as in animals (Escobar-Paramo et al., 2006; Rossi et al., 2018). Despite the ubiquitous association of commensal *E. coli* with humans and other higher organisms, the breadth and characteristics of the *E. coli* strain population in the ecological niche of the gut has not been covered in detail.

In this work, we initially characterized a commensal representative of the ST10 clonal complex closely related to *E. coli* K-12. This strain produces a rdar biofilm morphotype at 28°C (Figure S2; (Cimdins et al., 2017b)) and harbors an IncY plasmid with the previously identified TLPQC/LHR locus that mediates tolerance toward temporal exposure to lethal temperature and additional cargo genes involved in detoxification.

Among the *E. coli* strains, the historical strain *E. coli* K-12 isolated from a reconvalescent patient is the most well-investigated. However, it has been questioned to which extent *E. coli* K-12 has adapted to laboratory conditions during years of laboratory maintenance and subject to extensive manipulation (Browning et al., 2013). ST10 *E. coli* strains, which belong to phylogroup A, are not only found at high frequency in the gastrointestinal tract of human beings, but also in the environment, plants, animals, and the clinical habitat, suggesting a superior capacity for survival (Freitag et al., 2018; Jorgensen et al., 2017; Manges et al., 2015; Reid et al., 2017; Richter et al., 2018; Shepard et al., 2012). Whether and, if so, which distinct subgroups exist within the ST10 type, which horizontally transferred genetic elements and single-nucleotide polymorphisms contribute to persistence and survival needs to be further explored, but O-antigen and flagella types seem to be highly variable among closely related ST10 strains (Day et al., 2019).

IncY plasmids derive from phage P1, which is present as an episomal copy in its lysogenic cycle, stably maintained by a toxin/antitoxin system (Lobocka et al., 2004). pFec10 clearly is a composite plasmid with a phage gene scaffold harboring a variety of unconventional cargo genes that mediate, for example, stress resistance and nutrient acquisition. We assume that we observe, with pFec10, the early development of a prophage into a plasmid, which subsequently acquired a high number of IS elements indicative for rapid evolution and cargo genes including a TLPQC locus to provide fitness advantages for strain Fec10. Furthermore, as judged by the majority of cargo genes, we assume that Fec10 may experience inflammatory and/or toxic conditions in the gut. For example, tetrathionate reduction is utilized by the gastrointestinal pathogen *S. typhimurium*, a close relative, to successfully thrive in the inflamed gut (Winter and Baumler, 2011). The ability to reduce tetrathionate could thus contribute to the in vivo fitness of commensal Fec10 upon inflammation whereby the produced thiosulfate might be reduced by other microorganisms. Furthermore, the TLPQC locus mediates not only thermotolerance, but also tolerance against the inflammatory gut oxidants chlorine and hydrogen peroxide (Wang et al., 2020).

Interestingly, a collection of minimally processed commensal *E. coli* strains of independent genetic background, previously assessed for their biofilm formation capability (Bokranz et al., 2005), harbored >10% thermotolerant strains, while overall, 2% of sequenced *E. coli* genomes are thermotolerant (Li and Ganzle, 2016). Three out of four of those TLPQC bearing strains, including Fec10, were of low abundance in feces of healthy human beings (Bokranz et al., 2005). Processes that select for thermotolerant strains in the environment and the medical setting include manipulations with temporal temperature upshift such as the processing of raw milk cheese, natural fermentation processes, steam treatment of meat, and mild thermosterilization of endoscopes (Bojer et al., 2011, Dlusskaya et al., 2011; Peng et al., 2012; Wang et al., 2018). We conclude from this work that the consumption of raw milk cheese or other food products does eventually establish *E. coli* strains with enhanced thermotolerance in the human gastrointestinal tract albeit at low frequency. Elevation of this strain population might occur, though, upon inflammation. Thermotolerance, however, is not a recently acquired phenotype of *E. coli* as a highly similar TLPQC locus is already present in *E. coli* C, isolated prior to 1920 (unpublished observation) and encoded on a plasmid of *S. enterica* serovar Senftenberg strain isolated in 1941 (Nguyen et al., 2017). *S. enterica* probably readily survived in dried egg powder that was a frequent source of outbreaks of salmonellosis in the 1940s. In this line, the TLPQC locus has recently also been shown to mediate pressure tolerance and tolerance against oxidizing agents used in the food industry (Li et al., 2020; Wang et al., 2020).

The only other TLPQC bearing plasmid in a food-borne *E. coli* belongs to the IncFII incompatibility group (Boll et al., 2017). The presence of the TLPQC locus on IncFII and IncY plasmids indicates that this locus can be readily transferred to other *E. coli* strains or other bacterial species within or outside of the *Enterobacteriaceae* family. Plasmid transfer is especially effective in the gastrointestinal tract. In combination with a bloom of *E. coli* upon pathogen-driven inflammation, horizontal gene transfer can reach extremely high numbers even in the absence of antibiotic pressure (Stecher et al., 2012).

As observed for *E. coli* strains from different origins (Dlusskaya et al., 2011; Li and Ganzle, 2016), the TLPQC locus can mediate thermotolerance to extremely high temperatures in commensal *E. coli* strains. This physiology has its foundation in the biochemical characteristics of the isolated disaggregase $ClpG_{GI}$, a major determinant of tolerance toward lethal temperatures (this work; Lee et al., 2018). $ClpG_{GI}$ exhibits enhanced disaggregation activity as compared to the canonical ClpB disaggregase, enabling the protein to process tight protein aggregates formed at more extreme temperature or higher protein concentration (Katikaridis et al., 2019). Of note, although highly similar, $ClpG_{GI}$ proteins from different bacteria differ in their biochemical characteristics and their ability to complement exposure to the extreme lethal temperature. As such, $ClpG_{GI\text{-}Fec10}$ has a significantly higher ATPase activity than $ClpG_{GI\text{-}SG17M}$. The N2 domain, recently shown to suppress ATPase activity of $ClpG_{GI\text{-}SG17M}$ (Lee et al., 2018), shows extended sequence variability (Figure S6b) and could contribute to the elevated ATPase activity of $ClpG_{GI\text{-}Fec10}$. Our data indicated that $ClpG_{GI}$ proteins are particularly heat stable thus be protected from unfolding and aggregation at more extreme temperatures. Indeed, both $ClpG_{GI}$ proteins showed increased in vitro thermostability compared to the ClpB/DnaK bi-chaperon system, the core genome disaggregase

involved in temperature tolerance in *E. coli* and other organisms. The high melting temperature upon binding to ATPγS (T_M value of 69-70°C) of $ClpG_{GI}$ represents another distinguishing feature from canonical ClpB and rationalizes why $ClpG_{GI}$, but not ClpB (T_M: 62°C) and its essential partner DnaK (T_M: 59°C) (Palleros et al., 1992) protects bacteria from severe heat shock. Loss of canonical ClpB and DnaK at extreme temperatures due to unfolding (and likely aggregation) will eradicate the disaggregation potential of *E. coli* cells that do not encode $ClpG_{GI}$ as an alternative disaggregation system. Nevertheless, only $ClpG_{GI\text{-}Fec10}$ significantly complemented *E. coli* Fec10 $clpG_{GI}$ and $dna\text{-}shsp20_{GI}\text{-}clpG_{GI}$ mutants in vivo upon exposure to 60°C in contrast to $ClpG_{GI\text{-}SG17M}$. This implies that the slightly different in vitro biochemical characteristics are enhanced in vivo or that species-specific factors or substrates impact on $ClpG_{GI}$ activity in vivo. In any case, our observations provide the biochemical basis that supports the $clpG_{GI}$ phenotype of increased thermotolerance in Fec10 and other commensal *E. coli* strains, but also suggests species-specific adaptation of $ClpG_{GI}$ proteins.

4 | EXPERIMENTAL PROCEDURES

4.1 | Strains and growth conditions

The commensal *E. coli* strains used in this study are described in Table 1. Strains were routinely cultivated aerobically at 30°C or 37°C with shaking at 120-200 rpm in Luria-Bertani (LB) broth (BD Difco), unless otherwise indicated. If required, 100 μg ml^{-1} of ampicillin (Amp), 30 μg ml^{-1} of kanamycin (Km), 50 μg ml^{-1} of spectinomycin (Spec) or 25 μg ml^{-1} of chloramphenicol (Cm) were added. *E. coli* TOP10 was used to propagate recombinant plasmids. Genotypically distinct Fec6, Fec9, Fec10, Fec12, Fec17, Fec23, Fec27, Fec32, Fec34, Fec35, Fec41, Fec51, Fec55, Fec56, Fec59, Fec60, Fec61, Fec65, Fec75, Fec81, Fec89, Fec93, Fec95, Fec97, Fec98, Fec99, Fec100, Fec101, Fec108, Fec110, Fec112, Fec113 (Bokranz et al., 2005) were screened for the presence of $shsp20_{GI}$ as a TLPQC marker by PCR (primers in Table S3).

4.2 | Determination of the genome sequence and annotation

To sequence the genomes of *E. coli* Fec6 and Fec10, the genomic DNA was isolated using the Qiagen Genomic DNA isolation Kit (500/G binding capacity) according to the user's instructions. Briefly, cells were cultured in 30 ml of LB broth to OD_{600} = 0.4. Cell lysates were obtained by incubation in bacterial lysis buffer containing RNases A (200 μg/ml) and proteinase K (250 μg/ml). The cell lysate was applied onto a Genomic-tip 500/G by gravity flow. After elution with TE buffer pH 8, the genomic DNA was analyzed on a 0.5% agarose gel, and the concentration and purity (260/280 nm ratio) were measured by NanoDrop.

TABLE 1 *E. coli* strains used in this study

Strain	Genotype/Source/Characteristic	Reference
Commensal *E. coli* isolates and derivatives		
Fec6	commensal strain/human feces	Bokranz et al. (2005)
Fec10	commensal strain/human feces, Culture Collection University of Gothenburg CCUG 74881	Bokranz et al. (2005)
Fec10 Δ$clpG_{GI}$	deletion of disaggregase $ClpG_{GI}$	This study
Fec10 Δdhc_{GI}	deletion of $dna\text{-}shsp20_{GI}\text{-}clpG_{GI}$ operon	This study
Fec10 Δ*ttrSR*	deletion of two component system TtrSR	This study
Fec10 Δ*ttrBCA*[1]	deletion of tetrathionate reductase TtrBCA	This study
Fec10 Δ*ttrSR* Δ*ttrBCA*[1]	double deletion of tetrathionate reductase regulatory and structural genes	This study
Fec32	commensal strain/human feces	Bokranz et al. (2005)
Fec41	commensal strain/human feces	Bokranz et al. (2005)
Fec55	commensal strain/human feces	Bokranz et al. (2005)
Fec75	commensal strain/human feces	Bokranz et al. (2005)
Fec89	commensal strain/human feces	Bokranz et al. (2005)
Fec113	commensal strain/human feces	Bokranz et al. (2005)
TOB1	commensal strain/human feces	Bokranz et al. (2005)
E. coli K-12 derivatives		
XL1 blue	XL1 blue placIq	Weibezahn et al. (2004)
MC4100 Δ*clpB*	MC4100 △*clpB*::Km placIq	Weibezahn et al. (2004)
MC4100 Δ*dnaK103*	MC4100 Δ*dnaK103* placIq	Winkler et al. (2010)
TOP10	cloning and propagating plasmids; F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7,697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen

Note: Deletion of the *ttrBAC* operon included a hypothetical coding sequence located immediately upstream *ttrB*.

TABLE 2 Plasmids used in this study

Plasmid	Description	Source/Reference
pKD3	template plasmid for λ Red mediated recombination, Cm^r, Amp^r	Datsenko and Wanner (2000)
pKD46	P_{BAD} promoter, encodes λ Red recombinase, Amp^r	Datsenko and Wanner (2000)
pBAD30	cloning vector, pACYC origin, L-arabinose inducible *araBAD* promoter, Amp^r	Guzman et al. (1995)
pJN105	broad-host range vector with L-arabinose inducible *araBAD* promoter; pBBR1ori, Gm^r	Newman and Fuqua (1999)
pBAD30 $clpG_{GI\text{-}Fec10}$	Amp^r; $clpG_{GI\text{-}Fec10}$ cloned with C-terminal 6xHis tag	This study
pBAD30 $dhc_{GI\text{-}Fec10}$	Amp^r; $dhc_{GI\text{-}Fec10}$ cloned with C-terminal 6xHis tag	This study
pJN105 $clpG_{GI\text{-}Fec10}$	Gm^r; $clpG_{GI\text{-}Fec10}$ cloned with C-terminal 6xHis tag	This study
pBAD30 $clpG_{GI\text{-}SG17M}$	Amp^r; $clpG_{GI\text{-}SG17M}$ cloned with C-terminal 6xHis tag	This study
placIq	$Spec^r$, harbors lacIq to allow for IPTG controlled gene expression, p15A ori	Lee et al. (2018)
pUHE21	Amp^r, IPTG-inducible Ptac promoter, pBR322 ori	Lee et al. (2018)
pUHE21 *clpB*	Amp^r, *clpB* cloned into BamHI/XbaI restriction sites	Lee et al. (2018)
pUHE21 *dnaK*	Amp^r, *dnaK* cloned into BamHI/HindIII restriction sites	Lee et al. (2018)
pUHE21 $clpG_{GI\text{-}Fec10}$	Amp^r, $clpG_{GI\text{-}Fec10}$ cloned into BamHI/XbaI restriction sites	This study
pUHE21 $clpG_{GI\text{-}SG17M}$	Amp^r, $clpG_{GI\text{-}SG17M}$ cloned into BamHI/XbaI restriction sites	Lee et al. (2018)
pET24a $clpG_{GI\text{-}SG17M}$	Kan^r, $clpG_{GI\text{-}SG17M}$ with C-terminal 6x His tag	Lee et al. (2018)

Genomic DNA sequencing and assembly were performed as described by PacBio RSII and Illumina sequencing (Cimdins et al., 2017a). Illumina raw reads were aligned to the PacBio assembly to correct sequencing errors based on PacBio sequencing inaccuracy. Annotation was performed using Prokka (Seemann, 2014) and RASTtk (Brettin et al., 2015).

The correct plasmid sequence was identified by matching Illumina reads to the plasmid contig and confirmed by the conventional DNA Sanger sequencing of isolated plasmid DNA (primers in Table S3). Plasmid circularization was performed using the circulator minimus2 option (Hunt et al., 2015). The presence of direct repeats was checked with Gepard (Krumsiek et al., 2007). IS elements were identified using ISfinder (https://isfinder.biotoul.fr) with multiple hits and hits with identities below 95% omitted. DNAPlotter (Carver et al., 2009) was used to create the plasmid map. Genome sequences of *E. coli* Fec10 and Fec6 were deposited in the NCBI database under GenBank accession number MDLJ00000000.2 and WIPD00000000.2.

4.3 | Genome characterization and bioinformatic analyses

Genome annotation and search for genome characteristics were performed using RAST and MicroScope (Aziz et al., 2008; Brettin et al., 2015; Vallenet et al., 2020). Initial strain typing was performed using the ANItools (Han et al., 2016) and the Center for Genomic Epidemiology webservices (http://www.genomicepidemiology.org/). ST-type was determined by MLST 2.0 (Larsen et al., 2012). In silico analysis of plasmid replicons was performed by PlasmidFinder 2.1 (https://cge.cbs.dtu.dk/services/PlasmidFinder/; Carattoli et al., 2014). Phylogenetic groups were identified by Clermont typing (http://clermontyping.iame-research.center/).

Search for phage sequences was performed with PHASTER (Arndt et al., 2016). Protein sequences were aligned with Clustal W and MEGA X was used to construct the maximum likelihood (ML) and neighborhood joining (NJ) phylogenetic tree for $ClpG_{GI}$ proteins (Kumar et al., 2018). Jalview was used to visualize the protein alignment and DNAPlotter was used to visualize the plasmid map (Carver et al., 2009).

Additional bioinformatic procedures are found in Supporting Information.

4.4 | Construction of gene deletion mutants

Construction of the deletion mutants (Δ$clpG_{GI}$, Δdhc_{GI} (*dna-shsp20*$_{GI\text{-}ClpG_{GI}}$), Δ*ttrSR* and Δ*ttrBCA*) in *E. coli* Fec10 was performed via λ Red recombination (Datsenko and Wanner, 2000). The chloramphenicol

and kanamycin-resistant cassette from the template plasmid pKD3 and pKD4, respectively, flanked by 40 bps (50 bps for *ttr* deletion mutants) upstream of the start codon and 40 bp downstream of the stop codon of the respective genes was amplified with primers as stated in Table S3. After purification and DpnI digestion, the PCR products were transformed into Fec10 harboring the pKD46 plasmid. Mutants were selected on 12 μg $ml^{-1}$of chloramphenicol and subsequently verified by PCR for ORF replacement using primers outside the homologous recombination regions. To cure the pKD46 plasmid, mutants were incubated at 42°C overnight. Colonies were tested by streaking on LB agar containing Cm and Amp and Cm^R and Amp^S colonies were selected. Deletion mutants were verified by PCR.

4.5 | Plasmid construction

Plasmids used in this study are described in Table 2. To express the protein in *E. coli* TOP10 for subsequent purification, $clpG_{GI\text{-}Fec10}$ was cloned into the broad-host-range vector pJN105 (Newman and Fuqua, 1999) with the L-arabinose inducible P_{BAD} promoter. Primers $clpG_{GI}$ _Fec10_NheI_pJN_F/R (Table S3) were used to amplify $clpG_{GI\text{-}Fec10}$ with a C-terminal 6xHis tag from *E. coli* Fec10 DNA. After restriction digestion, the amplified fragment was cloned between NheI/XbaI restriction sites of the multiple cloning site using standard procedures. For complementation, $clpG_{GI\text{-}Fec10}$, $dhc_{GI\text{-}Fec10}$, and $clpG_{GI\text{-}SG17M}$ were cloned with a C-terminal 6xHis tag between EcoRI/SalI restriction sites into the pBAD30 vector under the L-arabinose inducible P_{BAD} promoter (Guzman et al., 1995). $ClpG_{GI\text{-}Fec10}$ was cloned via BamHI/XbaI restriction sites into the pUHE21 vector for complementation studies in *E. coli* Δ*clpB* and *dnaK103* chaperone mutants. Primers are listed in Table S3.

4.6 | Rdar biofilm assay

Ten microliters of overnight bacterial culture were spotted onto a 25 ml of LB without NaCl agar plate supplemented with CR (40 μg ml^{-1}) and Coomassie brilliant blue (20 μg ml^{-1}) (CR agar plate) or 50 μg ml^{-1} of Calcoflour white (fluorescence brightener 28) (Römling et al., 1998a). When indicated, sodium tetrathionate (1.5 or 15 mM) was added. Plates were incubated at 28 and 37°C for 48 hr and the colony morphotype was documented.

4.7 | Lethal-temperature-tolerance assay

Cells grown in LB broth containing 0.1% L-arabinose and 100 μg ml^{-1} of ampicillin with shaking at 200 rpm at 37°C were harvested after 18 hr. OD_{600} was adjusted to one in the same medium and 500 μl of the cell suspension was incubated at 55, 60 and 65°C for the indicated amount of time, while a control cell suspension was kept on ice. After the preparation of 10-fold serial dilutions in LB medium, cell viability was assessed by the spot assay. Ten microliters of cell suspension of each dilution were spotted onto LB agar and plates were incubated at 37°C for 14–15 hr.

Viability assays for *E. coli* K-12 Δ*clpB* and *dnaK103* cells harboring the plasmid placIq and pUHE21-derivatives after IPTG-controlled expression of *clpB*, *dnaK*, and $clpG_{GI}$ were performed as described previously (Lee et al., 2018).

4.8 | Protein purification

$ClpG_{GI\text{-}Fec10}$ with a C-terminal 6xHis-tag was expressed from the pJN105 plasmid in *E. coli* TOP10 (Table 2). Cultures were incubated at 37°C, expression induced at $OD_{600} = 0.5$ by 0.1% L-arabinose and cells subsequently harvested 6 hr postinduction. Cells were broken with a French press and the cell debris removed by centrifugation at 15 000 *g* for 45 min. The protein was purified according to the standard procedure using Ni-NTA (Qiagen) or Ni-IDA (Protino). The purified protein was subsequently subjected to Superdex S200 size exclusion chromatography in MDH assay buffer supplemented with 5% (v/v) glycerol. Afterward, dialysis was performed in 50 mM Na_2HPO_4, 300 mM NaCl and 5% glycerol, pH 8 overnight at 4°C. The protein purity was assessed by SDS-PAGE and the protein concentration was determined by the Bradford assay using a BSA calibration curve. Protein concentrations refer to monomers. $ClpG_{GI\text{-}SG17M}$, ClpB, DnaK, DnaJ, GrpE, and firefly Luciferase were purified as described before (Katikaridis et al., 2019; Lee et al., 2018).

4.9 | In vitro disaggregation activity assays

Disaggregation assays were performed as described previously with modifications (Mogk et al., 2003). The model substrates used to test the disaggregation activity were malate dehydrogenase (MDH; Roche) and firefly luciferase (Roche). In brief, 2 μM MDH and 0.2 μM luciferase were heat aggregated at 47°C for 30 min and 45°C for 15 min, respectively, in MDH assay buffer (50 mM Tris pH 7.5, 20 mM $MgCl_2$, 150 mM KCl, and 2 mM DTT) and subsequently incubated at room temperature for 5 min. Fifty microliters of aggregated MDH and luciferase (final concentration 1 and 0.1 μM, respectively) were mixed with 54 μl of assay buffer containing 2 μM disaggregase. In the MDH assay, 1 μM each GroEL/GroES was added to aid the refolding of the enzyme. The KJE mix consisted of 1 μM DnaK, 0.2 μM DnaJ, 0.1 μM GrpE. The reaction was started by adding 6 μl of ATP regeneration system (2 mM ATP, 3 mM phosphoenolpyruvate, and 20 ng/μl of pyruvate kinase) and samples were incubated at 30°C for disaggregation and refolding. To determine the recovered MDH activity, consumption of NADH was measured by monitoring the decrease in absorbance at 340 nm using a UV spectrophotometer (Novaspec Plus, Amersham). For each MDH activity assay, 10 μl of sample was mixed with 690 μl of measurement buffer (150 mM potassium phosphate pH 7.6, 2 mM DTT, 0.5 mM oxaloacetate, and 0.28 mM NADH). To assess luciferase activity, 2 μl of sample was mixed with 125 μl of 2X luciferase assay

buffer (25 mM glycylglycine pH 7.4, 12.5 mM $MgSO_4$, 5 mM ATP) and 125 µl of 25 mM luciferin (Gold Biotech.), and activity was measured by Berthold Lumat LB 9,507. Similarly, 10 and 2 µl of native MDH and luciferase, respectively, were mixed with measurement buffer. The value corresponds to 100% substrate activity before heat denaturation.

4.10 | ATPase assay

The ATPase assay is based on the coupled enzymatic activity of pyruvate kinase (PK) and lactate dehydrogenase (LDH). The disaggregase hydrolyzes ATP to ADP that is recycled by PK to convert phosphoenolpyruvate (PEP) to pyruvate. LDH in turn converts pyruvate into lactate thereby oxidizing NADH to NAD+. The decrease in NADH corresponds to the ATPase rate of the disaggregases and is recorded by the decrease in absorbance at 340 nm in a 96 well plate by FLUOstar-Omega (BMG-Labtech). One µM disaggregase was assessed in 100 µl of MDH assay buffer containing 10 µl of 10X reaction mix (50 µl of 5 mM NADH, 50 µl of 10 mM PEP (Sigma) and 50 µl of PK/LDH (Sigma)). Optionally, casein was added to a final concentration of 0.1 µg/µl to stimulate the ATPase activity of ClpB. The reaction was started by adding 100 µl of 4 mM ATP in the MDH assay buffer.

4.11 | Circular Dichroism (CD) spectroscopy

Circular dichroism spectroscopy was used to assess the thermal stability of the ClpB and $ClpG_{GI}$ proteins. The disaggregases were dialyzed against 10 mM potassium phosphate buffer pH 7.4 and diluted to a final concentration of 3 µM. Optionally, $MgAc_2$ and ATPγS were added to a final concentration of 5 and 1 mM, respectively. CD spectra were recorded from 190 to 250 nm for the different disaggregases using a Jasco J750 spectropolarimeter. The melting curves of disaggregases were obtained by increasing the temperature at a rate of 0.5°C per minute from 20 to 85°C with the circular dichroism signal recorded at 222 nm. T_M values were calculated using Prism software.

4.12 | Anisotropy measurements

Binding of ClpB, $ClpG_{GI\text{-}SG17M}$, and $ClpG_{GI\text{-}Fec10}$ to 100 nM FITC-casein was monitored by fluorescence anisotropy measurements using a BMG Biotech CLARIOstar plate reader. Samples were incubated in MDH assay buffer for 5 min at 30°C in the presence of 2 mM ATPγS and polarization of FITC-casein was determined in black 384 well plates (excitation: 482 nm; emission: 530 nm, Target mP: 35). A sample containing FITC-casein only served as reference. K_d values were determined using nonlinear regression curve fitting (Prism software).

4.13 | Transmission electron microscopy

A detailed description of this experimental procedure is found in Supporting Information.

ACKNOWLEDGMENTS

Mikhail S. Gelfand supervised Zaira Seferbekova and Robert Afasizhev under RSF grant 18-14-00358. Inge Kristen (Helmholtz Center for Infection Research, Braunschweig, Germany) is gratefully thanked for excellent electron microscopic sample preparation. Shady Mansour Kamal received a travel grant from the Swedish Society to conduct work at the University of Heidelberg, Germany. Panagiotis Katikaridis was supported by the Heidelberg Biosciences International Graduate School (HBIGS). This work was supported by the Swedish Research Council for Natural Sciences and Engineering (project number 621-2013-4809 and 2017-04465) to U.R. and by a grant of the Deutsche Forschungsgemeinschaft (MO970/4-3) to A.M. The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council, and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with massively parallel sequencing and access to the UPPMAX computational infrastructure.

AUTHOR CONTRIBUTIONS

Conception of the study: UR and AM; acquisition of data: SMK, HL, CL, ACA, FL, AJM-R, ZS, RA, HTW, LM, PK, AM, UR; analysis of data: SMK, HL, ACA, FL, AJM-R, HTW, ZS, RA, PK, UD, AM, UR; interpretation of the data: SMK, HL, ACA, AJM-R, UD, AM, UR; writing of the manuscript: UR and SMK. All authors contributed to the revision and commented on the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Material and strains are available upon request.

ORCID

Shady Mansour Kamal https://orcid.org/0000-0002-7565-811X
Annika Cimdins-Ahne https://orcid.org/0000-0001-7414-9914
Changhan Lee https://orcid.org/0000-0003-0327-4712
Fengyang Li https://orcid.org/0000-0001-5102-6284
Alberto J. Martín-Rodríguez https://orcid.org/0000-0003-2422-129X
Zaira Seferbekova https://orcid.org/0000-0002-3131-3697
Robert Afasizhev https://orcid.org/0000-0002-3513-8873
Haleluya Tesfaye Wami https://orcid.org/0000-0002-9929-2570
Panagiotis Katikaridis https://orcid.org/0000-0001-5783-1487
Lena Meins https://orcid.org/0000-0001-6214-477X
Heinrich Lünsdorf https://orcid.org/0000-0003-2778-8611
Ulrich Dobrindt https://orcid.org/0000-0001-9949-1898
Axel Mogk https://orcid.org/0000-0003-3674-5410
Ute Römling https://orcid.org/0000-0003-3812-6621

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*, W16–W21.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., et al. (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, *9*, 75.

Bachmann, B.J. (1972) Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriological Reviews*, *36*, 525–557.

Beghain, J., Bridier-Nahmias, A., Le Nagard, H., Denamur, E. and Clermont, O. (2018) ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microbial Genomics*, *4*, e000192.

Bettelheim, K.A. and Lennox-King, S.M. (1976) The acquisition of *Escherichia coli* by new-born babies. *Infection*, *4*, 174–179.

Blount, Z.D. (2015) The unexhausted potential of *E. coli*. *Elife*, *4*, e05826.

Bojer, M.S., Hammerum, A.M., Jorgensen, S.L., Hansen, F., Olsen, S.S., Krogfelt, K.A., et al. (2012) Concurrent emergence of multidrug resistance and heat resistance by CTX-M-15-encoding conjugative plasmids in *Klebsiella pneumoniae*. *APMIS*, *120*, 699–705.

Bojer, M.S., Krogfelt, K.A. and Struve, C. (2011) The newly discovered ClpK protein strongly promotes survival of *Klebsiella pneumoniae* biofilm subjected to heat shock. *Journal of Medical Microbiology*, *60*, 1559–1561.

Bojer, M.S., Struve, C., Ingmer, H., Hansen, D.S. and Krogfelt, K.A. (2010) Heat resistance mediated by a new plasmid encoded Clp ATPase, ClpK, as a possible novel mechanism for nosocomial persistence of *Klebsiella pneumoniae*. *PLoS One*, *5*, e15467.

Bokranz, W., Wang, X., Tschäpe, H. and Römling, U. (2005) Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. *Journal of Medical Microbiology*, *54*, 1171–1182.

Boll, E.J., Marti, R., Hasman, H., Overballe-Petersen, S., Stegger, M., Ng, K., et al. (2017) Turn up the heat-food and clinical *Escherichia coli* isolates feature two transferrable loci of heat resistance. *Frontiers in Microbiology*, *8*, 579.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., et al. (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, *5*, 8365.

Browning, D.F., Wells, T.J., Franca, F.L., Morris, F.C., Sevastsyanovich, Y.R., Bryant, J.A., et al. (2013) Laboratory adapted *Escherichia coli* K-12 becomes a pathogen of C*aenorhabditis elegans* upon restoration of O antigen biosynthesis. *Molecular Microbiology*, *87*, 939–950.

Carattoli, A., Zankari, E., Garcia-Fernandez, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014) In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, *58*, 3895–3903.

Carver, T., Thomson, N., Bleasby, A., Berriman, M. and Parkhill, J. (2009) DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*, *25*, 119–120.

Chai, Y., Kolter, R. and Losick, R. (2009) A widely conserved gene cluster required for lactate utilization in *Bacillus subtilis* and its involvement in biofilm formation. *Journal of Bacteriology*, *191*, 2423–2430.

Cimdins, A., Lüthje, P., Li, F., Ahmad, I., Brauner, A. & Römling, U. (2017a) Draft genome sequences of semiconstitutive red, dry, and rough biofilm-forming commensal and uropathogenic *Escherichia coli* isolates. *Genome Announcements*, *5*, e01249–e1316.

Cimdins, A., Simm, R., Li, F., Lüthje, P., Thorell, K., Sjöling, A., et al. (2017b) Alterations of c-di-GMP turnover proteins modulate semi-constitutive rdar biofilm formation in commensal and uropathogenic *Escherichia coli*. *Microbiologyopen*, *6*, e00508.

Clermont, O., Bonacorsi, S. and Bingen, E. (2000) Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Applied and Environment Microbiology*, *66*, 4555–4558.

Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 6640–6645.

Day, M.J., Hopkins, K.L., Wareham, D.W., Toleman, M.A., Elviss, N., Randall, L., et al. (2019) Extended-spectrum beta-lactamase-producing *Escherichia coli* in human-derived and foodchain-derived samples from England, Wales, and Scotland: an epidemiological surveillance and typing study. *The Lancet Infectious Diseases*, *19*, 1325–1335.

Dlusskaya, E.A., McMullen, L.M. and Gänzle, M.G. (2011) Characterization of an extremely heat-resistant *Escherichia coli* obtained from a beef processing facility. *Journal of Applied Microbiology*, *110*, 840–849.

Dublan Mde, L., Ortiz-Marquez, J.C., Lett, L. and Curatti, L. (2014) Plant-adapted *Escherichia coli* show increased lettuce colonizing ability, resistance to oxidative stress and chemotactic response. *PLoS One*, *9*, e110416.

Escobar-Paramo, P., Le Menac'h, A., Le Gall, T., Amorin, C., Gouriou, S., Picard, B., et al. (2006) Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environmental Microbiology*, *8*, 1975–1984.

Freitag, C., Michael, G.B., Li, J., Kadlec, K., Wang, Y., Hassel, M., et al. (2018) Occurrence and characterisation of ESBL-encoding plasmids among *Escherichia coli* isolates from fresh vegetables. *Veterinary Microbiology*, *219*, 63–69.

Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, *177*, 4121–4130.

Han, N., Qiang, Y.andZhang, W. (2016) ANItools web: a web tool for fast genome comparison within multiple bacterial strains. *Database*, *2016*, baw084. https://doi.org/10.1093/database/baw084

Hendriksen, R.S., Bortolaia, V., Tate, H., Tyson, G.H., Aarestrup, F.M. and McDermott, P.F. (2019) Using genomics to track global antimicrobial resistance. *Frontiers in Public Health*, *7*, 242.

Hensel, M., Hinsley, A.P., Nikolaus, T., Sawers, G. and Berks, B.C. (1999) The genetic basis of tetrathionate respiration in *Salmonella typhimurium*. *Molecular Microbiology*, *32*, 275–287.

Hunt, M., Silva, N.D., Otto, T.D., Parkhill, J., Keane, J.A. and Harris, S.R. (2015) Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biology*, *16*, 294.

Jorgensen, S.B., Soraas, A.V., Arnesen, L.S., Leegaard, T.M., Sundsfjord, A. and Jenum, P.A. (2017) A comparison of extended spectrum beta-lactamase producing *Escherichia coli* from clinical, recreational water and wastewater samples associated in time and location. *PLoS One*, *12*, e0186576.

Kapralek, F. (1972) The physiological role of tetrathionate respiration in growing citrobacter. *Journal of General Microbiology*, *71*, 133–139.

Katikaridis, P., Mein, L., Mansour Kamal, S., Römling, U. and Mogk, A. (2019) ClpG provides increased heat resistance by acting as superior disaggregase. *Biomolecules*, *9*, 815.

Khetrapal, V., Mehershahi, K.S. and Chen, S.L. (2017) Complete genome sequence of the original *Escherichia coli* isolate, strain NCTC86. *Genome Announc*, *5*, e00243-17.

Koraimann, G. (2018) Spread and persistence of virulence and antibiotic resistance genes: a ride on the F plasmid conjugation module. *EcoSal Plus*, *8*. https://doi.org/10.1128/ecosalplus.-ESP-0003-2018

Krumsiek, J., Arnold, R. and Rattei, T. (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics*, *23*, 1026–1028.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*, 1547–1549.

Larsen, M.V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., et al. (2012) Multilocus sequence typing of total-genome-sequenced bacteria. *Journal of Clinical Microbiology*, *50*, 1355–1361.

Lee, C., Franke, K.B., Kamal, S.M., Kim, H., Lünsdorf, H., Jager, J., et al. (2018) Stand-alone ClpG disaggregase confers superior heat tolerance to bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, E273–E282.

Lee, C., Wigren, E., Lünsdorf, H. and Römling, U. (2016) Protein homeostasis-more than resisting a hot bath. *Current Opinion in Microbiology*, *30*, 147–154.

Lee, C., Wigren, E., Trcek, J., Peters, V., Kim, J., Hasni, M.S., et al. (2015) A novel protein quality control mechanism contributes to heat shock resistance of worldwide-distributed *Pseudomonas aeruginosa* clone C strains. *Environmental Microbiology*, *17*, 4511–4526.

Lehnherr, H., Maguin, E., Jafri, S. and Yarmolinsky, M.B. (1993) Plasmid addiction genes of bacteriophage P1: doc, which causes cell death on curing of prophage, and phd, which prevents host death when prophage is retained. *Journal of Molecular Biology*, *233*, 414–428.

Lescat, M., Clermont, O., Woerther, P.L., Glodt, J., Dion, S., Skurnik, D., et al. (2013) Commensal *Escherichia coli* strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Environmental Microbiology Reports*, *5*, 49–57.

Li, H. and Gänzle, M. (2016) Some like it hot: Heat resistance of *Escherichia coli* in food. *Frontiers in Microbiology*, *7*, 1763.

Li, H., Mercer, R., Behr, J., Heinzlmeir, S., McMullen, L.M., Vogel, R.F., et al. (2020) Heat and pressure resistance in *Escherichia coli* relates to protein folding and aggregation. *Frontiers in Microbiology*, *11*, 111.

Lin, L., Wang, X., Cao, L. and Xu, M. (2019) Lignin catabolic pathways reveal unique characteristics of dye-decolorizing peroxidases in *Pseudomonas putida*. *Environmental Microbiology*, *21*, 1847–1863.

Liu, M., Zhang, Y., Inouye, M. and Woychik, N.A. (2008) Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 5885–5890.

Lobocka, M.B., Rose, D.J., Plunkett, G. 3rd, Rusin, M., Samojedny, A., Lehnherr, H., et al. (2004) Genome of bacteriophage P1. *Journal of Bacteriology*, *186*, 7032–7068.

Manges, A.R., Geum, H.M., Guo, A., Edens, T.J., Fibke, C.D. and Pitout, J.D.D. (2019) Global extraintestinal pathogenic *Escherichia coli* (ExPEC) lineages. *Clinical Microbiology Reviews*, *32*, e00135-18.

Manges, A.R., Harel, J., Masson, L., Edens, T.J., Portt, A., Reid-Smith, R.J., et al. (2015) Multilocus sequence typing and virulence gene profiles associated with *Escherichia coli* from human and animal sources. *Foodborne Pathogens and Disease*, *12*, 302–310.

Martín-Rodríguez, A.J., Reyes-Darias, J.A., Martín-Mora, D., González, J.M., Krell, T. and Römling, U. (submitted) Reduction of alternative electron acceptors drives biofilm formation in Shewanella algae.

Martín-Rodríguez, A.J., Rhen, M., Melican, K. and Richter-Dahlfors, A. (2020) Nitrate metabolism modulates biosynthesis of biofilm components in uropathogenic *Escherichia coli* and acts as a fitness factor during experimental urinary tract infection. *Frontiers in Microbiology*, *11*, 26.

Massot, M., Daubie, A.S., Clermont, O., Jaureguy, F., Couffignal, C., Dahbi, G., et al. (2016) Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of *Escherichia coli* from community subjects in the Paris area in 2010 and evolution over 30 years. *Microbiology*, *162*, 642–650.

Meric, G., Kemsley, E.K., Falush, D., Saggers, E.J. and Lucchini, S. (2013) Phylogenetic distribution of traits associated with plant colonization in *Escherichia coli*. *Environmental Microbiology*, *15*, 487–501.

Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E. and Bukau, B. (2003) Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Molecular Microbiology*, *50*, 585–595.

Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H., et al. (1999) Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. *EMBO Journal*, *18*, 6934–6949.

Newman, J.R. and Fuqua, C. (1999) Broad-host-range expression vectors that carry the l-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. *Gene*, *227*, 197–203.

Nguyen, S.V., Harhay, G.P., Bono, J.L., Smith, T.P. and Harhay, D.M. (2017) Genome sequence of the thermotolerant foodborne pathogen *Salmonella enterica* Serovar Senftenberg ATCC 43845 and Phylogenetic Analysis of Loci Encoding Increased Protein Quality Control Mechanisms. *mSystems*, *2*(1). https://doi.org/10.1128/mSystems.00190-16

Palleros, D.R., Reid, K.L., McCarty, J.S., Walker, G.C. and Fink, A.L. (1992) DnaK, hsp73, and their molten globules. Two different ways heat shock proteins respond to heat. *Journal of Biological Chemistry*, *267*, 5279–5285.

Palumbo, S.A. and Alford, J.A. (1970) Inhibitory action of tetrathionate enrichment broth. *Applied Microbiology*, *20*, 970–976.

Peng, S., Stephan, R., Hummerjohann, J., Bianco, J. and Zweifel, C. (2012) In vitro characterization of Shiga toxin-producing and generic *Escherichia coli* in respect of cheese-production relevant stresses. *Journal of Food Safety and Food Quality*, *63*, 136–141.

Pilla, G. and Tang, C.M. (2018) Going around in circles: virulence plasmids in enteric pathogens. *Nature Reviews Microbiology*, *16*, 484–495.

Reid, C.J., Wyrsch, E.R., Roy Chowdhury, P., Zingali, T., Liu, M., Darling, A.E., et al. (2017) Porcine commensal *Escherichia coli*: a reservoir for class 1 integrons associated with IS26. *Microbial Genomes*, *3*, e000143.

Richter, T.K.S., Hazen, T.H., Lam, D., Coles, C.L., Seidman, J.C., You, Y. et al. (2018) Temporal variability of *Escherichia coli* diversity in the gastrointestinal tracts of Tanzanian children with and without exposure to antibiotics. *Msphere 3*, e00558-18.

Römling, U., Bian, Z., Hammar, M., Sierralta, W.D. and Normark, S. (1998a) Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. *Journal of Bacteriology*, *180*, 722–731.

Römling, U., Sierralta, W.D., Eriksson, K. and Normark, S. (1998b) Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agfD* promoter. *Molecular Microbiology*, *28*, 249–264.

Rossi, E., Cimdins, A., Lüthje, P., Brauner, A., Sjöling, A., Landini, P., et al. (2018) "It's a gut feeling"—*Escherichia coli* biofilm formation in the gastrointestinal tract environment. *Critical Reviews in Microbiology*, *44*, 1–30.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*, 2068–2069.

Shepard, S.M., Danzeisen, J.L., Isaacson, R.E., Seemann, T., Achtman, M. and Johnson, T.J. (2012) Genome sequences and phylogenetic analysis of K88- and F18-positive porcine enterotoxigenic *Escherichia coli*. *Journal of Bacteriology*, *194*, 395–405.

Shintani, M., Takahashi, Y., Yamane, H. and Nojiri, H. (2010) The behavior and significance of degradative plasmids belonging to Inc groups in *Pseudomonas* within natural environments and microcosms. *Microbes and Environments*, *25*, 253–265.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. and Chandler, M. (2006) ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Research*, *34*, D32–D36.

Singh, R., Grigg, J.C., Qin, W., Kadla, J.F., Murphy, M.E. and Eltis, L.D. (2013) Improved manganese-oxidizing activity of DypB, a peroxidase from a lignolytic bacterium. *ACS Chemical Biology*, *8*, 700–706.

Stecher, B., Denzler, R., Maier, L., Bernet, F., Sanders, M.J., Pickard, D.J., et al. (2012) Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. *Proceedings*

of the National Academy of Sciences of the United States of America, *109*, 1269–1274.
Sternberg, N. and Hoess, R. (1983) The molecular genetics of bacteriophage P1. *Annual Review of Genetics*, *17*, 123–154.
Tenaillon, O., Skurnik, D., Picard, B. and Denamur, E. (2010) The population genetics of commensal *Escherichia coli*. *Nature Reviews Microbiology*, *8*, 207–217.
Tracey, J.C., Coronado, M., Giessen, T.W., Lau, M.C.Y., Silver, P.A. and Ward, B.B. (2019) The discovery of twenty-eight new encapsulin sequences, including three in Anammox bacteria. *Scientific Reports*, *9*, 20122.
Vallenet, D., Calteau, A., Dubois, M., Amours, P., Bazin, A., Beuvin, M., et al. (2020) MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Research*, *48*, D579–D589.
Venturini, C., Zingali, T., Wyrsch, E.R., Bowring, B., Iredell, J., Partridge, S.R., et al. (2019) Diversity of P1 phage-like elements in multidrug resistant *Escherichia coli*. *Scientific Reports*, *9*, 18861.
von Mentzer, A., Connor, T.R., Wieler, L.H., Semmler, T., Iguchi, A., Thomson, N.R., et al. (2014) Identification of enterotoxigenic *Escherichia coli* (ETEC) clades with long-term global distribution. *Nature Genetics*, *46*, 1321–1326.
Wang, Z., Fang, Y., Zhi, S., Simpson, D.J., Gill, A., McMullen, L.M., et al. (2020) The locus of heat resistance confers resistance to chlorine and other oxidizing chemicals in *Escherichia coli*. *Applied and Environment Microbiology*, *86*, e02123–e2219.
Wang, Z., Li, P., Luo, L., Simpson, D.J. and Gänzle, M.G. (2018) Daqu fermentation selects for heat-resistant Enterobacteriaceae and Bacilli. *Applied and Environment Microbiology*, *84*, e01483–e1518.
Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., et al. (2004) Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell*, *119*, 653–665.
Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., et al. (2010) Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *EMBO Journal*, *29*, 910–923.
Winter, S.E. and Bäumler, A.J. (2011) A breathtaking feat: to compete with the gut microbiota, *Salmonella* drives its host to provide a respiratory electron acceptor. *Gut Microbes*, *2*, 58–60.
Zolkiewski, M. (1999) ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from *Escherichia coli*. *Journal of Biological Chemistry*, *274*, 28083–28086.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Kamal SM, Cimdins-Ahne A, Lee C, et al. A recently isolated human commensal *Escherichia coli* ST10 clone member mediates enhanced thermotolerance and tetrathionate respiration on a P1 phage-derived IncY plasmid. *Mol Microbiol* 2021;115:255–271. https://doi.org/10.1111/mmi.14614