A novel protein quality control mechanism contributes to heat shock resistance of world-wide distributed *Pseudomonas aeruginosa* clone C strains

Changhan Lee¹, Edvard Wigren², Janja Trček^{1†}, Verena Peters¹, Jihong Kim³, Muhammad Sharif Hasni¹, Manfred Nimtz⁵, Ylva Lindqvist², Chankyu Park³, Ute Curth⁴, Heinrich Lünsdorf⁵ and Ute Römling^{1*}

¹ Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, 171 77 Stockholm, Sweden

² Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden

³ Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

⁴ Institute for Biophysical Chemistry, Hannover Medical School, 30625 Hannover, Germany

⁵ The Helmholtz Center for Infection Research, 38124 Braunschweig, Germany

[†] Present address: Faculty of Natural Sciences and Mathematics, Department of Biology, University of Maribor, Maribor, Slovenia

Running title

Role of Hsp20 in *P. aeruginosa* clone C strains

Correspondence

Ute Römling

Department of Microbiology, Tumor and Cell Biology, Karolinska Institute, 171 77 Stockholm, Sweden

E-mail: Ute.Romling@ki.se; Tel. +46-8-524 87319; Fax: +46-8-31 11 01.

Summary

Pseudomonas aeruginosa is a highly successful nosocomial pathogen capable to cause a wide variety of infections with clone C strains most prevalent worldwide. In this study, we initially characterize a molecular mechanism of survival unique to clone C strains. We identified a *P. aeruginosa* clone C-specific genomic island (PACGI-1) which contains the highly expressed small heat shock protein sHsp20c, the founding member of a novel subclass of class B bacterial small heat shock proteins. sHsp20c and adjacent gene products are involved in resistance against heat shock. Heat stable sHsp20c is unconventionally expressed in stationary phase in a wide temperature range from 20-42°C. Purified sHsp20c has characteristic features of small heat shock protein class B as it is monodisperse, forms sphere-like 24-meric oligomers and exhibits significant chaperone activity. As the *P. aeruginosa* clone C population is significantly more heat shock resistant than genetically unrelated *P. aeruginosa* strains without sHsp20c, the horizontally acquired *shsp20c* operon might contribute to survival of world-wide distributed clone C strains.

Introduction

Pseudomonas aeruginosa is a highly successful nosocomial pathogen, which causes a wide variety of infections such as wound, ear, lung, eye and urinary tract infections (Kerr and Snelling, 2009). Decades-long chronic lung infections in cystic fibrosis (CF) caused by *P. aeruginosa* significantly affects morbidity and mortality (Lyczak *et al.*, 2002). On the other hand, *P. aeruginosa* is widely distributed in diverse environmental niches including the aquatic habitat covering natural and processed water and the clinical environment (Grobe *et al.*, 1995; Pirnay *et al.*, 2005). Despite the ecological and clinical importance of *P. aeruginosa*, genetic studies have been mainly limited to strains PAO1 and PA14. Strain PAO1, though, belongs to an uncommon clonal group. While PA14 belongs to the second most common clonal group, the isolate contains a truncated *ladS* gene, which affects biofilm formation and virulence (Bodilis *et al.*, 2009; Mikkelsen *et al.*, 2011).

To date, the most abundant group of closely related *P. aeruginosa* strains is the clone C strain cluster (De Soyza *et al.*, 2013). Clone C strains infect patients worldwide, covering a broad spectrum of diseases and are found in the clinical as well as aquatic environments (Romling *et al.*, 1994a; Romling *et al.*, 2005; Rakhimova *et al.*, 2009; Kidd *et al.*, 2011; Tielen *et al.*, 2011; Selezska *et al.*, 2012; Petit *et al.*, 2013). For example, clone C strains are highly prone to infect CF patients over decades (Romling *et al.*, 1994b), cause chronic obstructive bronchitis (Rakhimova *et al.*, 2009) and urinary tract infections (Tielen *et al.*, 2011). The genome structure of clone C strains has been extensively investigated by low resolution typing techniques (Romling *et al.*, 1997). In contrast to clonal groups of limited occurrence, extensive chromosomal plasticity has been observed among clone C strains (Romling *et al.*, 1997). However, the mechanisms how *P. aeruginosa* clone C strains successfully survive in the environment and infect patients have not been addressed. Acquisition of accessory genetic

elements by individual strains, specific features encoded on the core genome, or both factors may play a role.

The accessory genome is a flexible gene pool which accelerates adaptive evolution (Dobrindt *et al.*, 2004) to provide individual isolates with specific fitness traits (He *et al.*, 2004) (Finck-Barbancon *et al.*, 1997; Roy-Burman *et al.*, 2001; Kulasekara *et al.*, 2006). In *P. aeruginosa* clone C, strain and clone specific genetic elements have been identified. For example, genomic islands PAGI-2 and PAGI-3 are present in two distinct clone C strains (Larbig *et al.*, 2002). pKLC102 is a 103.5-kb plasmid found ubiquitously in *P. aeruginosa* clone C strains, occasionally integrated into the chromosome (Klockgether *et al.*, 2004; Klockgether *et al.*, 2007).

Recently, we determined the whole-genome sequence of *P. aeruginosa* SG17M, a clone C isolate from a river in Germany (Lee *et al.*, 2014). In this work, we characterize a small heat shock protein (sHsp20), sHsp20c, one of the most highly expressed gene products of *P. aeruginosa* clone C strains. sHsp20c is encoded on a novel *P. aeruginosa* clone C-specific genomic island (PACGI-1) of 86 kbp (Fig. 1). sHsp20s are molecular chaperones of molecular mass between 12-43 kDa widespread throughout all kingdoms of life from bacteria over archaea to humans (Narberhaus, 2002; Haslbeck *et al.*, 2005). In humans, sHsps are involved in prevention of neurological disorders and apoptosis (Stege *et al.*, 1999; Clark and Muchowski, 2000) and a major component of the eye lens is sHsp alpha-crystallin giving name to the conserved α -crystallin domain of around 90 amino acids and which is flanked by variable amino- and carboxyl-terminal regions in sHsps (Kriehuber *et al.*, 2010; Haslbeck and Vierling, 2015). Generally, sHsps act as ATP-independent molecular chaperones to prevent aggregation of proteins upon stress conditions (Haslbeck *et al.*, 2005). Our study reveals a role of the sHsp20c protein and adjacent operon genes in stress resistance thus contributing to

the survival advantage of clone C strains.

Results

Detection of sHsp20c as a clone C-specific protein

Comparing the whole-genome sequence of the *P. aeruginosa* clone C strain SG17M (Romling *et al.*, 2005; Lee *et al.*, 2014) with PA14 and PAO1, a SG17M-specific genomic island of 86 kbp, PACGI-1, was identified. Previously, comparative proteome analysis of clone C strain C with PAO1 identified differentially expressed proteins associated with adaptation to the CF lung habitat (Sriramulu *et al.*, 2005). A 21.4 kDa protein with pI of 5.33 was solely expressed at high level in strain C (Fig. 2A). MALDI-TOF analysis identified the protein as a sHsp, termed sHsp20c (Fig. S1, Table S6). *shsp20c* is located on PACGI-1 as the 94th gene (Table S1).

Established bacterial sHsp20 proteins can be classified into class A and B (Fu *et al.*, 2006). Phylogenetic analysis of sHsp20 representatives from animals, plants, fungi and bacterial class A and B (Fig. 2B, Fig. S2; (Fu *et al.*, 2006)) revealed that sHsp20c is with 24.3% identity clearly distinct from IbpA, encoded on the core genome of *P. aeruginosa* as well as from IbpA2, another sHsp20 of certain *P. aeruginosa* strains ((Fig. 2BC; (Chenna *et al.*, 2003)). Proteins highly similar to sHsp20c are, however, found in individual isolates of *Klebsiella pneumoniae* (Kpn), *Burkholderia multivorans* (Bmu), *Pseudomonas stutzeri* (Pst) and other species, which form a subgroup of sHsp20c-like proteins within the bacterial class B (Fig. 2B).

The common structural feature of sHsps is the α -crystallin domain. The alignment of amino acid sequences between sHsp20s shows that sHsp20c harbors the conserved α -crystallin

domain (NCBI/blastp; pfam00011; (Kriehuber *et al.*, 2010)) from amino acid 85 to 173 and sHsp-conserved region I and II, which are involved in higher oligomerization and dimerization, respectively (Fig. 2C; (Waters, 1995; Kappe *et al.*, 2002; Fu *et al.*, 2006)). However, in contrast to core genome IbpA, sHsp20c has a longer N-terminus. From secondary structure prediction and homology modeling the conserved α -crystallin domain consisting of a β -sandwich comprising two antiparallel β -sheets can be predicted to be structurally intact in sHsp20c. However, due to low sequence identity no accurate structural information could be obtained for the flanking amino- and carboxyl-terminal regions. As a representative 3D-structure of monomeric sHsp20c we generated a structural model using the closest homologue in PDB as a template, the 30.2% identical heat shock protein 16.0 of *Schizosaccharomyces pombe* (PDB: 3w1z), Fig. 2D).

To detect production of sHsp20c, we developed a peptide antibody against a highly conserved region of sHsp20c, which specifically recognizes sHsp20c in SG17M (Fig. S3). sHsp20c production was detected in 15 environmental and 26 clinical isolates of clone C and J, with the exception of one isolate (Fig. 2E, Fig. S3, data not shown). Clone J is the 4th most frequent epidemic clone in aquatic habitats and patients (Hilker *et al.*, 2014). No significant difference in sHsp20c production between environmental and clinical isolates was observed (Fig. S4). However, screening of 117 genotypically independent *P. aeruginosa* isolates, 56 environmental and 61 clinical, for the presence of *shsp20c* indicated that significantly more environmental (55%) than clinical (26%) isolates harbored *shsp20c* (Fig. 2F).

Involvement of sHsp20c in stress resistance

In bacteria, sHsps are involved in stress response, such as heat shock and oxidative stress (Thomas and Baneyx, 1998; Kuczyńska-Wisnik *et al.*, 2010). To identify the physiological

function of sHsp20c, we constructed an in-frame deletion mutant of *shsp20c* in SG17M (Fig. S3). Assessment of the *shsp20c* mutant for resistance against heat shock and hydrogen peroxide, a generator of extrinsic oxidative stress, showed modestly increased sensitivity of the *shsp20c* mutant (Fig. 3AB). Sensitivity was complemented by single copy *shsp20c* inserted at the Tn7 insertion site downstream of essential *glmS* (Choi *et al.*, 2005) and expressed from the 275 bp of promoter region in conjunction with upstream *dna* (Fig. 3), but not by sHsp20c cloned on various plasmids (data not shown). Although sHsp20c is one of the most highly expressed proteins in *P. aeruginosa* SG17M, we reasoned that uncontrolled production of sHsp20c from plasmids is responsible for non-complementation.

Additive effect of adjacent genes

The *shsp20c* gene is located downstream of *dna* encoding a predicted DNA-binding protein and *clpBc* encoding a ClpB-like ATP dependent chaperone (Fig. 1). This genetic organization is conserved in strains, which harbor sHsp20c homologues (Bojer *et al.*, 2010; Gajdosova *et al.*, 2011; Bojer *et al.*, 2013), indicating functional relationship. The 93 aa Dna protein contains an N-terminal helix-turn-helix domain and is an AlpA-like transcriptional regulator (Fig. S5). The ClpBc protein is 949 aa long and contains two AAA+ ATPases domains flanking a CHD5-like coiled-coil domain. In addition, ClpB specific domains are located at the N- and C-terminus (Fig. S5). The triple deletion mutant of *dna*, *shsp20c* and *clpBc* significantly increased susceptibility to heat shock and, partially, oxidative stress compared to the *shsp20c* single mutant (Fig. 3AB). Sensitivity of the triple mutant was complemented by single copy insertion of the three genes at the Tn7 insertion site. We also investigated the role of *dna* in sHsp20c production and heat resistance (Fig. S6). Of note, *dna* is a suppressor of sHsp20c production and decreased heat resistance when expressed from the pJN105 plasmid

in the *dna shsp20c clpBc* triple mutant strain of SG17M.

Functional interplay between sHsp20c and IbpA

As all *P. aeruginosa* strains, clone C strains including SG17M code for the sHsp IbpA in the core genome raising the question whether the two heat shock proteins sHsp20c and IbpA could act additively. To our knowledge, the phenotype of an *ibpA* mutant of *P. aeruginosa* has not been investigated. An *ibpA* mutant of SG17M has increased sensitivity towards heat shock (Fig. 3A), in agreement with the phenotype in *Pseudomonas putida* (Krajewski *et al.*, 2013). A slight additive effect on heat shock sensitivity was detected in the double mutant of *ibpA* and *shsp20c* (Fig. 3A). Sensitivity against oxidative stress induced by treatment of hydrogen peroxide was, however, not altered (Fig. 3B).

Clone C strains are more resistant against heat stress

PACGI-1 contains additional genes potentially involved in stress resistance, such as superoxide dismutases and a thioredoxin (*trx*) (Table S1). We hypothesized the island to confer increased survival to stress to clone C strains. Therefore, we tested stress resistance between two groups of *P. aeruginosa* strains. One group consisted of five non-sHsp20c harboring non-clone C strains including PA14 and the other group contained sHsp20c positive clone-C strains of environmental and clinical origin (Fig. 3C). PT1M, an aquatic clone J strain was also included in the sHsp20c-positive group. Interestingly, the sHsp20c-harboring strain group was significantly more resistant against heat shock than the group of non-sHsp20c harboring strains with an average survival of 81.7% and 26.5%, respectively, whereby survival of SG17M was normalized to 100%. We could not detect any significant difference in susceptibility to hydrogen peroxide between the two strain groups (Fig. 3D).

The *shsp20c* operon confers heat shock resistance to non-clone C strains

The heat shock resistance phenotype conferred by sHsp20c and enhanced by downstream encoded ClpBc is significant, but modest, in SG17M. To generalize the phenotype, we assessed whether the *dna shsp20c clpBc* gene products provide heat resistance also in other strain backgrounds. Introduction of *dna shsp20c* and *dna sHsp20c clpBc* in single copy at the Tn7 site in strains PA14 and SG9M, which had shown the greatest sensitivity to heat shock (Fig. 3C), mediated a gradually enhanced heat resistance to both strains (Fig. 3F). All together, heat resistance, which was 3 and 2% of SG17M was enhanced to 32 and 10% in PA14 and SG9M, respectively, upon introduction of the three genes. Of note, coexpression of *clpBc* significantly enhanced expression of the sHsp20c protein (Fig. 3E).

Unconventional expression of sHsp20c

Expression of heat shock protein has been well characterized in mesophilic bacteria, such as *E. coli*. Conventionally, in bacteria, sHsps are coordinately regulated with high molecular mass heat shock proteins and are specifically induced at high temperature with expression controlled by sigma factor 32 (Allen *et al.*, 1992). Elements positioned in the 5'-untranslated region of small heat shock genes, called ROSE (Repression of heat Shock gene Expression), act as repressor of expression at temperatures below 30°C (Kortmann and Narberhaus, 2012). As we originally identified sHsp20c highly expressed at 37°C (Fig. 2E), we systematically investigated sHsp20c expression. Interestingly, sHsp20c is expressed within a wide temperature range from 20 to 42°C. Further, sHsp20c is expressed on solid and in liquid medium as well as in rich Luria–Bertani (LB) medium and minimal M63 medium (Fig. 4ABC). sHsp20c expression starts in mid-log phase with high induction in stationary phase

(Fig. 4B). This unique expression pattern of sHsp20c was also observed in the unrelated PT1M clone J strain (Fig. 4C).

Although sHsp20c is expressed independently of temperature, we observed induction of sHsp20c by heat shock as well as oxidative stress (Fig. 4DE). When cell were shifted from 20 to 37 and 42°C, respectively, expression of sHsp20c was induced at the same growth phase (Fig. 4E). In addition, exposure to hydrogen peroxide for an hour also induced expression of sHsp20c (Fig. 4D). Consistent with a role of sHsp20c in stress response, we observed more than 50% of the entire amount of sHsp20 in aggregates (Fig. 4F), which corresponds to a more than 150-fold enrichment of sHsp20c relative to the total protein content in aggregates compared to soluble protein.

Next, we wondered whether sHsp20c is differentially expressed at different life-style stages of *P. aeruginosa* SG17M. By separating cell aggregates resembling biofilm formation from single planktonic cells in liquid culture, elevated expression of sHsp20c in planktonic cells compared to cell aggregates was detected at 20 and 37°C (Fig. 4G,H). Thus, sHsp20c might preferentially protect planktonic cells from stress.

Chaperone activity of sHsp20c

Chaperone activity has been demonstrated for bacterial and eukaryotic sHsps (Jakob *et al.*, 1993; Studer and Narberhaus, 2000). As sHsp20c is a member of a novel subgroup of sHsp20c proteins, we wondered whether sHsp20c has chaperone activity. To study this, sHsp20c containing an N-terminal 6XHis-tag was expressed in *P. aeruginosa* SG17M $\Delta shsp20c$ and purified to homogeneity (Fig. 5A, Experimental Procedures). Size exclusion chromatography (SEC) performed on 6XHis-sHsp20c resulted in a single symmetrical elution

peak running at an apparent molecular mass of 400-500 kDa, indicative of a homogenous high molecular mass 6XHis-sHsp20c oligomer (Fig. 5A). Importantly, a similar SEC elution profile as for 6XHis-sHsp20c was observed for non-tagged sHsp20c, indicating that the N-terminal 6XHis-tag does not change the oligomeric properties of sHsp20c. Over-digestion of 6XHis-sHsp20c with thrombin, in order to remove the N-terminal 6XHis-tag, resulted in a truncated sHsp20c form missing 5-6 amino acids, which only formed dimers (data not shown). These findings are consistent with previous reports on sHsps forming higher order oligomers, with bacterial group B sHsps being monodisperse, and assembly of oligomers to be disrupted by deletion of the N- or C-terminal tail of the proteins (Haslbeck *et al.*, 2005).

The chaperone activity of sHsp is well documented (Narberhaus, 2002). Inhibition of thermal aggregation of the model substrate citrate synthase (CS) was used to assess the chaperone activity of 6XHis-tagged sHsp20c (Fig. 5B). A significant chaperone activity was detected with correlation to the increment of molar ratio, although the chaperone activity of sHsp20c was lower than that of the positive control α -crystallin. To monitor the thermal stability of sHsp20c, its secondary structure was examined by far-UV circular dichroism (CD) (Fig. 5C). Spectra containing a single broad negative ellipticity minimum between 210 and 218 nm, characteristic of a β sheet- β turn rich secondary structure, were monitored within the temperature range from 20–50°C. Further temperature elevation resulted in a spectrum profile indicative of loss of β -sheet structure, beginning noticeably at 60°C. Therefore, sHsp20c is most likely physiologically stable under the heat induced condition up to 50°C to function as a chaperone in the cell (Fig. 5C).

Oligomeric structure of sHsp20c

In order to more precisely define the oligomerization state of sHsp20c, purified 6XHis-tagged sHsp20c oligomer was subjected to sedimentation velocity (SV) and equilibrium (SE) analysis in an analytical ultracentrifuge. SV revealed a single oligomeric species of sHsp20c independent of protein concentration with a sedimentation coefficient of $s_{20,w}^0 = 18.1$ S (Fig. 6A). SE analysis with the program SEDPHAT using a single species model (Vistica *et al.*, 2004) resulted in a molecular mass of 547 kDa (Fig. 6B; details in Supporting Information, see below). Since the molecular mass of a 6XHis-tagged sHsp20c monomer as calculated from amino acid composition is 22.3 kDa, the protein most probably forms 24- or 25-mers in solution.

To determine the shape of the sHsp20 oligomer, transmission electron microscopy (TEM) was performed with purified 6XHis-sHsp20c particles, negatively stained with uranylacetate (Fig. 7). As is shown in the survey view, molecules were monodisperse (Fig. 7A, circles), in general, but doublets (Fig. 7A; elliptic circles) and even linear triplets could be found. Doublets were regularly found in distinct contact, assuming some kind of affine coupling or interactions. Larger aggregates obviously were caused by crowding effects upon EM-sample preparation. Occasionally, small particles were seen that can be interpreted as debris or polypeptide dimers (Fig. 7A, small circles). Interestingly, individual particles, beside the general circular outline, often showed a hexagonal top-view projection with fine details of substructure.

With the information from SE and symmetry constraints from electron microscopy (Fig. 7BC), sHsp20c most probably forms 24-mers in solution with a most probable molecular arrangement of 24 subunits organized as finally 2 x 6 polypeptide dimers. With a sedimentation coefficient of 18.1 S, as determined from SV analysis, a frictional ratio of 1.3

relative to a spherical, unhydrated 24-mer can be calculated. Since the frictional ratio of a spherical hydrated protein is expected to be 1.1 – 1.2 (Lebowitz *et al.*, 2002), the shape of the sHsp20c 24-mer deviates only moderately from that of a sphere.

Discussion

P. aeruginosa has the remarkable ability to inhabit diverse environments and infect a range of organisms including humans (Grobe *et al.*, 1995; Rahme *et al.*, 2000; Kerr and Snelling, 2009; Papaioannou *et al.*, 2013). Although the population structure of *P. aeruginosa* is mainly epidemic, few clonal clusters, but often locally or spatially restricted, can dominate (Mathee *et al.*, 2008). *P. aeruginosa* clone C strain cluster is the most common clonal group of closely related strains world-wide in environmental and clinical niches (Romling *et al.*, 2005; Selezska *et al.*, 2012; De Soyza *et al.*, 2013). In this study, we identified *P. aeruginosa* clone C-specific genomic island PACGI-1. So far, *P. aeruginosa* genomic islands were functionally characterized in terms of pathogenicity (Mathee *et al.*, 2008; Kung *et al.*, 2010). Interestingly, PACGI-1 codes for diverse stress response related genes such as a several superoxide dismutases (SOD), a thioredoxin and glutathione-regulated potassium-efflux protein KefC, DNA-damage repair related RadC as well as heat shock and protein stability related sHsp20c, ClpBc and HtpX (Table S1) suggesting that PACGI-1 is an island contributing to quality control of macromolecules.

The proximal region of a tRNA gene is a common integration site of genomic island (Kung *et al.*, 2010). Remarkably, however, PACGI-1 is located within spacer 7 of the CRISPR 2 adaptive bacterial immune system against plasmid and phages (Sorek *et al.*, 2008). Currently, we do not have a rational explanation why the CRISPR locus was targeted by PACGI-1.

The PACGI-1 encoded sHsp20c is one of the highest expressed proteins of clone C strains (Fig. 2A and preliminary data not shown). sHsps of the sHsp20 family are holding chaperones involved in preventing protein aggregation under heat shock, oxidative stress and desiccation in bacteria (Thomas and Baneyx, 1998; Bojer *et al.*, 2010; Kuczynska-Wisnik *et al.*, 2010; Cocotl-Yanez *et al.*, 2014). Despite ubiquitous occurrence, few physiological functions have been assigned to sHsps in prokaryotes. Here, we show that the additional copy of sHsp20c in the PACGI-1 genomic island contributes to resistance to heat shock. In addition, we examined the virulence of *shsp20c* mutant strain in the larvae of *Galleria mellonella* as a model host. Although we could not observe a significant difference in early and end-point 48 h survival compared with SG17M wild-type strain, the mutant showed slightly attenuated virulence after 36 hours of infection (data not shown). The contribution of island genes to virulence is in agreement with clone C strains being highly virulent in this environmentally relevant model host (Hilker *et al.*, 2014).

Environmental strains of *P. aeruginosa* have a higher prevalence of encoding sHsp20c and presumably the PACGI-1 island than clinical strains (Fig. 2F). It is commonly considered that environmental strains of *P. aeruginosa* are indistinguishable from clinical isolates (Wolfgang *et al.*, 2003). The presence of sHsp20c in environmental isolates points to a predominant role of sHsp20c in the environment. As, on the other hand, the first (clone C) and forth (clone J) most prevalent clone cluster in patients and aquatic habitats harbors sHsp20c, a specific genetic background might favor a beneficial role of sHsp20c and the PACGI-1 island. Despite high expression in *P. aeruginosa* clone C strains, a balanced protein quality control homeostasis seems to be essential since expression of plasmid encoded sHsp20c could not complement the phenotype of the *shsp20c* mutant or was even detrimental (data not shown). However, *shsp20c* and *shsp20c clpBc* in single copy, in combination with the upstream

repressor *dna*, significantly contributed to heat resistance in the unrelated strains PA14 and SG9M, showing the generality of the heat resistance phenotype (Fig. 3F).

sHsp20c has chaperone function and prevents the model substrate citrate synthase from thermal aggregation. In *E. coli*, sHsps IbpA and IbpB work co-operatively with heat shock proteins ClpB and DnaK (Mogk *et al.*, 2003; Carroni *et al.*, 2014) with task distribution of sHsps as holding chaperone to prevent protein aggregation, while ClpB and DnaK act as disaggregating and folding chaperones, form a complex and cooperate to reactivate aggregated proteins (Arifuzzaman *et al.*, 2006; Seyffer *et al.*, 2012; Rosenzweig *et al.*, 2013). Interestingly, a combination of sHsp20c with the downstream located gene encoding a ClpB-like disaggregating chaperone is conserved in all strains with sHsp20c homologues (Bojer *et al.*, 2010; Gajdosova *et al.*, 2011; Bojer *et al.*, 2013). Functional relatedness is indicated as enhanced susceptibility to heat shock was observed compared to single deletion mutant of sHsp20c.

sHsp20c is unconventionally expressed in *P. aeruginosa* strains. In contrast to other bacterial sHsps expressed at elevated temperature (Laksanalamai *et al.*, 2001; Usui *et al.*, 2001; Krajewski *et al.*, 2013), sHsp20c is expressed in late logarithmic and stationary phase over a wide temperature range with 20°C as the lowest examined temperature. In most Gram-negative bacteria, transcription of sHsp and other heat shock responsive genes is coordinately regulated by the alternative sigma factor σ^{32} (alternatively called RpoH) (Allen *et al.*, 1992). The sHsp20c encoding operon might be regulated by σ^S an alternative sigma factor in stationary phases in response to general stress or nutrient starvation (Battesti *et al.*, 2011). Recently, σ^S -dependent sHsp20 has been reported in *Azotobacter vinelandii* to be involved in desiccation resistance (Cocotl-Yanez *et al.*, 2014). The distinct expression pattern of sHsp20c thus points to a protein quality control function in stationary phase clearly distinct from the

sHsp20 core genome protein IbpA. In this context, it is interesting to note that we observed life-style biased expression of sHsp20c as increased expression of sHsp20c in planktonic compared to cell aggregates. As protein aggregation is associated with aging, which is defined as the loss of reproductive ability (Lindner *et al.*, 2008), this mechanism might contribute to rejuvenation of bacterial cells with increased survival rates.

Phylogenetically, sHsp20c-like proteins constitute a distinct subfamily among bacterial class B sHsp20 proteins (Fig. 2B) (Fu *et al.*, 2006). In addition to sequence homology, the higher order protein structure is distinct between different classes of sHsps. Bacterial class A sHsps form polydisperse oligomers, whereas class B sHsps exist as monodisperse oligomers (Fu *et al.*, 2006). Consistent with the classification as a class B sHsp, sHsp20c forms a sphere-like mono-oligomeric 24-mer. Though from direct observation of negatively stained particles detailed substructural features can be recognized, views on class-averages did not reveal a distinct regular surface-pattern which reflects the subunit arrangement within the molecules. Beside the low-resolution of the 3D-reconstruction as such the reason for insufficient surface-patterning may be caused by irregular, noisy positioning of the outside oriented 6XHis-tag appendages that may distort visualization of the regular sHsp20c-subunit arrangement.

In summary, in this study, we identified clone C-specific genomic island PACGI-1. The island harbors a novel small heat shock protein sHsp20c, highly expressed in clone C strains, involved in heat shock response. This discovery contributes to the beginning elucidation of the physiological features of highly successful strains.

Experimental Procedures

Strains and growth conditions

P. aeruginosa SG17M (Romling *et al.*, 2005) was used as the representative clone C strain and was the parent of all constructed mutants. The construction of in-frame deletion mutant was performed using established protocols (Choi and Schweizer, 2005). All strains, plasmids and primers used in this study are listed in the table S3-5. If not otherwise specified, cells were grown in Luria–Bertani (LB) medium at 37°C containing appropriate antibiotics. For *P. aeruginosa* SG17M, gentamycin (Gm) was used at 30 mg/ml, and tetracycline (Tc) at 70 mg/ml. *E. coli* TOP10 was used for cloning and to retain the plasmids. For *E. coli*, Gm at 30 mg/ml, Tc at 20 mg/ml, kanamycin (Km) at 25 mg/ml and ampicillin (Ap) at 100 mg/ml was used.

Mutant construction

Briefly, deletion mutants were constructed by cloning flanking regions of the target genes into pSG001 (Table S4) vector, which is a derivative of pEX18Tc (Shanks *et al.*, 2006) harboring an FLP-excisable gentamycin cassette between *Bam*HI and *Sal*I. To amplify the gentamycin cassette, pUC18T mini-Tn7T-Gm (Choi and Schweizer, 2006) was used as a template of PCR with primers are described in table S5. Besides, all the primers which amplified the flanking regions of to be deleted genes are described in table S5. The resulting vectors were introduced into SG17M by triparental mating with *E. coli* HB101 harboring pRK2013 and *E. coli* TOP10 harboring the constructed plasmid. Selection was on citrated-based VBMM medium containing gentamycin (Choi and Schweizer, 2005). To confirm the double crossover, the conjugants were tested for Tc and Gm resistance. Tc^S and Gm^R conjugants were selected and streaked to obtain a single colony. pFLP3 plasmid was used for

Flp-mediated excision of the Gm marker, whereby established procedures were followed (Choi and Schweizer, 2005). In all cases, deletion mutant were confirmed by PCR using primers outside the entire region of homologous recombination.

Complementation

pUC18T-mini-Tn7T-Gm was used to complement mutants (Choi and Schweizer, 2006). The genes including their own promoter were amplified by using the primers (*shsp20c/dna shsp20c clpBc* mini F, *shsp20c* mini R and *dna shsp20c clpBc* mini R, Table S5) and cloned into the *Bam*HI and *Kpn*I site of multiple cloning site of the vector. For chromosomal integration, the cloned vector and pTNS2 helper plasmid encoding the TnsABCD transposase subunits were co-conjugated into the *dna shsp20c* or *dna shsp20c clpBc* deletion mutant of SG17M with *E. coli* HB101 containing pRK2013 conjugation-helper strain (Table 3). Gm^R conjugants were selected and chromosomal integration was confirmed by PCR using primers PglmS-up and PglmS-down (Table 5). For the excision of the Gm marker, pLFP3 plasmid and same procedure was used as described above for mutant construction (Choi and Schweizer, 2005).

DNA sequencing, annotation and bioinformatics methods

The whole genome sequence of SG17M determined by whole-genome shotgun sequencing is available under GenBank accession number JALF000000000 (Lee *et al.*, 2014). The sequence was compared to the annotated genomes of PAO1 and PA14 (www.pseudomonas.com; (Winsor *et al.*, 2011)). The location of the island corresponds to 2,317,400 to 2,403,499 bp of scaffold 1 of the SG17M genome sequence. Open reading frames (ORFs) of SG17M were annotated by BASys (www.basys.ca; (Van Domselaar *et al.*, 2005)) and RAST

(<http://rast.nmpdr.org/>; (Overbeek *et al.*, 2014)). The ORFs were analyzed for homologies by using BLASTN, BLASTP and BLASTX (www.ncbi.nlm.nih.gov/blast/). ClustalX2 and W2 were used for the sequence alignments of sHsps and phylogenetic analysis (Chenna *et al.*, 2003). Phylogenetic tree was visualized with Treeview (Page, 2002) and structure-based sequence alignment was visualized using the program ESPript (Robert and Gouet, 2014). The structure of sHsp20c was modeled from the coordinates of the heat shock protein 16.0 from *Schizosaccharomyces pombe* (PDB: 3w1z) using the SWISS-MODEL Workspace (Arnold *et al.*, 2006). Figure for molecular structure description was prepared with the program PyMOL (WL, 2002).

Susceptibility assays

Cell viability after exposure to heat shock and hydrogen peroxide (Sigma) was determined by counting colony forming unit (CFU). To assess susceptibility to heat shock, bacteria were grown for 48 h in LB broth at 20°C (stationary phase conditions with sHsp20c expression). 500 µl of a cell suspension with O.D.₆₀₀ 2.5 was exposed to 50°C for 30 and 60 min and a control cell suspension was kept at room temperature. For comparison cell viability of sHsp20c-harboring and non-harboring strains, cells were grown for 24 h in LB broth at 37°C before subjection to heat shock.

To assess susceptibility to hydrogen peroxide, cells were grown for 24 h in LB broth at 37°C. 500 µl bacterial suspension of O.D.₆₀₀ 2.5 was treated with 30 and 50 mM hydrogen peroxide for an hour at 37°C and a control suspension was kept at the same temperature without treatment. After exposure, 100 µl cell suspension diluted 10⁻⁷ was spread on LB-agar medium. Colonies were counted after 12–14 h incubation at 37°C and % of survival was calculated as the ratio of CFU of stress-treated and non-treated cells.

To compare susceptibility between sHsp20c-harboring and non-harboring strains, the survival rate of *P. aeruginosa* SG17M was arbitrarily set to 100%. At least two biological replicates were performed with three technical replicates on each occasion.

Development of an antiserum against sHsp20c

The sequence of sHsp20c was assessed for antigenicity and sHsp20c proteins available at the beginning of the study were aligned to identify conserved regions. Based on this evaluation, the amino acid sequences R₁₁₉GEKRQEETKDGGFHR₁₃₅ was selected for production of a polyclonal peptide antibody (PolyPeptide Laboratories, France).

For antibody purification, the serum was loaded onto a CnBr-activated sepharose coupled with peptide containing a N-terminal cysteine residue. The column was washed with 100 mM acetate buffer, fractions were eluted with 200 mM acetate buffer and immediately neutralized with 1.5 M Tris-HCl, pH 8.8. In a second step, fractions were eluted with 100 mM Triethylamine buffer, pH 11.5 and neutralized immediately. The fractions containing eluted antibody were combined, buffer exchanged to PBS and the antibody concentrated to 2 mg/ml.

Western blot

Cell lysates with equal protein content from cultures growing under specific conditions were subjected to SDS-PAGE (12% separating and 4% stacking gel) protein gel and blotted onto polyvinylidene difluoride (PVDF) membrane (Millipore). After 1 h incubation in blocking solution (1 x TBS pH 7.4, 0.01% Tween 20, 0.5% skim milk powder), sHsp20c was detected with 1/2000 anti-sHsp20c antiserum and 1/10,000 rabbit anti-mouse secondary antibody (Jackson Immuno Research, UK). Bound antibody was visualized with ECL chemiluminescent detection reagent (Roche). Detection of protein production was at least

performed twice in independent experiments.

Expression of sHsp20c in aggregated and planktonic cells

SG17M was grown in LB medium with vigorous agitation (200 rpm) at 20 and 37°C for 24 and 48 h, respectively. Cell aggregates were separated from planktonic cells by a short centrifugation at 500 rpm for 10 sec. The supernatant was collected as planktonic cells. Aggregated and single cell status was confirmed by light transmission microscopy (Deltavision). sHsp20c was detected by Western blot analysis using sHsp20c-specific antibody as described above. Experiments were repeated two times independently.

Isolation of aggregated proteins

In brief, for detection of aggregated proteins (Tomoyasu *et al.*, 2001), cells of an overnight culture grown in LB medium at 37°C were disrupted by sonication and intact cells were separated by mild centrifugation. Soluble protein was separated by centrifugation at 13 000 g and aggregated protein in the pellet was separated from the membrane associated proteins by addition of 10% NP40. Soluble and aggregated protein was analyzed by Coomassie staining after SDS-PAGE.

Purification of 6XHis-sHsp20c

Two versions of sHsp20c encoding genes, one encoding sHsp20c with an N-terminal 6XHis-tag and an alternative variant with a thrombin cleavage site after the N-terminal 6XHis-tag were inserted into pJN105 (Newman and Fuqua, 1999) using primers containing sequences for 6XHis and 6XHis with a thrombin cleavage site (aa sequence of thrombin cleavage site: LVPRGS) in *E. coli* TOP10 (Table S3-5). The construct was transformed into the *shsp20c*

deletion mutant of SG17M. SG17M containing the plasmid was grown to the mid-exponential phase (optical density at 600 nm of 0.5) in LB broth with gentamycin (30 µg/ml) at 37°C. Protein expression was induced by adding 0.1 % of arabinose. Cells were grown for an additional 4 h, harvested, and resuspended in binding buffer (20 mM Tris-HCl [pH 7.9], 5 mM imidazole, 500 mM NaCl). After disruption by sonication, cell debris was removed by centrifugation at $15.000 \times g$ for 30 min and protein was purified according to standard procedures for Ni-NTA columns (Qiagen). The protein concentration was determined by using the Bradford reagent and purity was assessed by SDS-PAGE. Thrombin kit (Novagen) was used to cleave off the His-tag; 10 µg of purified 6XHis-tagged sHsp20c protein was digested with 1 U in buffer (200 mM Tris-HCl, 1.5 M NaCl, 25 mM CaCl₂, pH 8.4) at 20°C for 16 h. The cleavage was monitored by SDS-PAGE analysis.

Size-exclusion chromatography

As a final purification step, all sHsp20c preparations were subjected to size-exclusion chromatography (SEC) using a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) equilibrated with a buffer consisting of 25 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 5% glycerol, at a flow rate of 1 ml/min. After SEC, the protein was concentrated, aliquoted, flash frozen in liquid nitrogen and stored at -80°C until further use. The molecular mass of purified non- and His-tagged sHsp20c was analyzed using the same settings. The column was calibrated with Ferretin, Catalase, Aldolase, Albumin, Ovalbumin, Chymotrypsinogen A, Ribonuclease A as molecular mass standards.

Circular dichroism (CD) spectroscopy

CD measurements were conducted on a JASCO J-810 spectropolarimeter using a cuvette of

0.1 cm path length with sHsp20c at a concentration of 0.1 mg/ml in a buffer containing 10 mM Tris-HCl (pH 7.5) and 10 mM NaCl. The spectra were recorded from 195 to 260 nm. Five scans were averaged and corrected for the buffer spectrum. Thermal unfolding were studied using a peltier apparatus in the range 20–70°C, with a scanning rate of 2 °C/min and at 10°C intervals.

Analytic ultracentrifugation

Sedimentation velocity (SV) and sedimentation equilibrium experiments (SE) were carried out in a Beckman Coulter ProteomeLab XL I analytical ultracentrifuge at 20°C in a buffer containing 150 mM NaCl, 25 mM Tris pH 7.5 using an An-50 Ti rotor. Concentration profiles were measured with the manufacturer's data acquisition software ProteomeLab XL-I Version 6.0 (Firmware 5.7) using either the absorption scanning optics at 280 nm only, or in combination with the Rayleigh interference optics.

Partial specific volume, extinction coefficient at 280 nm, buffer density and viscosity were calculated from amino acid and buffer composition, respectively, by the program SEDNTERP (Laue, 1995). Protein concentrations were determined spectrophotometrically and are given in monomers throughout the text. Figures were prepared using the program GUSSI (<http://biophysics.swmed.edu/MBR/software.html>, kindly provided by Dr. C. Brautigam).

SV analysis was performed at 25000 rpm in a concentration range from 4 to 72 µM 6XHis-tagged sHsp20c in 3 or 12 mm double sector centerpieces filled with 100 µl or 400 µl sample, respectively. For data analysis a model for diffusion-deconvoluted differential sedimentation coefficient distributions [continuous $c(s)$ distributions] implemented in the program SEDFIT (Schuck, 2000) was used, subsequently experimental s -values were corrected to $s_{20, w}$.

SE of 6, 12 and 20 μ M 6XHis-tagged sHsp20c were performed with 150 μ l sample in standard 12 mm double sector centerpieces at 3500, 5000, 7000 and 9000 rpm. Before SE, centrifugation cells were mechanically stabilized and water blanks were collected using the interference optics as described by Vistica *et al.*, 2004. Samples were spun until no further change in concentration gradients could be observed, as judged by the program GUSSE. After subtraction of water blanks, interference data acquired at multiple loading concentrations and rotor speeds were modelled globally together with the corresponding absorbance data in the program SEDPHAT (Vistica *et al.*, 2004) using a single species model.

Chaperone activity assay

The chaperone activity of 6XHis-sHsp20c was measured by monitoring changes in aggregation of the thermolabile model substrate citrate synthase from porcine heart (CS; Sigma) (Sastry *et al.*, 2002; Malki *et al.*, 2003). The thermal aggregation experiments with CS were conducted at 43°C in 50 mM NaPO₄ pH 6.8 buffer. Indicated molar ratios of 6XHis-sHsp20c were transferred to a cuvette and preincubated for 5 min with stirring, prior to mixing with CS at a final concentration of 0.1 μ M. After preincubation, right-angle light scattering of aggregated CS was measured using a Shimadzu RF-5000 equipped with a stirrer, enabling constant temperature control with an EYELA model CTP-101 circulating water bath. Both excitation and emission wavelengths were set to 360 nm with a slit width of 3.0 nm.

Electron microscopy

Sample preparation of 6XHis-sHsp20c for TEM analysis followed mainly the procedures described in Chen *et al.* (2010) and references therein. Negatively stained (1% (w/v) uranyl acetate, pH 4.0) samples were analyzed by an in-column energy-filter transmission electron

microscopy (EF-TEM; Zeiss Libra 120plus, Zeiss, Oberkochen, Germany) doing image acquisition close to Scherzer and Gaussian focus at medium to low dose irradiation in the elastic bright-field mode at nominal x40000 magnification. The energy-slit width was set to 10 eV, using a 60 μ m objective aperture and a beam current of 1.0 μ A. Images were recorded with a 2048 x 2048 cooled bottom-mounted CCD camera (SharpEye, Tröndle, Moorenwies, Germany). Single particle analysis was performed with the EMAN software package (Ludtke *et al.*, 1999; Tang *et al.*, 2007), reference-free 2D analysis and class-averaging of a 4292 particle dataset was performed.

Acknowledgments

The authors would like to thank Lidia Litz for excellent technical assistance of AUC analysis, I. Kristen (HZI, Braunschweig) for admirable technical support in electron microscopic preparation and BILS (Bioinformatics Infrastructure for Life Sciences) for bioinformatics support.

The project was funded by the Swedish Research Council for Medicine and Health (K2012-56X-22039-01-3), the Carl Trygger Foundation (grant CTS 10:324) and the Karolinska Institutet to UR; Basic Science Research Program from the National Research Foundation of Korea (NRF) [2012R1A1A2041712] and a grant from KRIBB Research Initiative Program [KGM4611411] to CP. Changan Lee receives a personal scholarship through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2014R1A6A3A03057742).

References

582 Allen, S.P., Polazzi, J.O., Gierse, J.K., and Easton, A.M. (1992) Two novel heat shock genes
583 encoding proteins produced in response to heterologous protein expression in *Escherichia*
584 *coli*. *J Bacteriol* **174**: 6938-6947.

585 Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R. *et al.* (2006) Large-
586 scale identification of protein-protein interaction of *Escherichia coli* K-12. *Genome Res* **16**:
587 686-691.

588 Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a
589 web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195-
590 201.

591 Battesti, A., Majdalani, N., and Gottesman, S. (2011) The RpoS-mediated general stress
592 response in *Escherichia coli*. *Annu Rev Microbiol* **65**: 189-213.

593 Bodilis, J., Ghysels, B., Osayande, J., Matthijs, S., Pirnay, J.P., Denayer, S. *et al.* (2009)
594 Distribution and evolution of ferripyoverdine receptors in *Pseudomonas aeruginosa*. *Environ*
595 *Microbiol* **11**: 2123-2135.

596 Bojer, M.S., Struve, C., Ingmer, H., and Krogfelt, K.A. (2013) ClpP-dependent and -
597 independent activities encoded by the polycistronic *clpK*-encoding locus contribute to heat
598 shock survival in *Klebsiella pneumoniae*. *Res Microbiol* **164**: 205-210.

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

602 Carroni, M., Kummer, E., Oguchi, Y., Wendler, P., Clare, D.K., Sinning, I. *et al.* (2014) Head-
603 to-tail interactions of the coiled-coil domains regulate ClpB activity and cooperation with
604 Hsp70 in protein disaggregation. *Elife* **3**: e02481.

605 Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson,

606 J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids*
607 *Res* **31**: 3497-3500.

608 Choi, K.H., and Schweizer, H.P. (2005) An improved method for rapid generation of
609 unmarked *Pseudomonas aeruginosa* deletion mutants. *BMC Microbiol* **5**: 30.

610 Choi, K.H., and Schweizer, H.P. (2006) mini-Tn7 insertion in bacteria with single *attTn7*
611 sites: example *Pseudomonas aeruginosa*. *Nat Protoc* **1**: 153-161.

612 Choi, K.H., Gaynor, J.B., White, K.G., Lopez, C., Bosio, C.M., Karkhoff-Schweizer, R.R.,
613 and Schweizer, H.P. (2005) A Tn7-based broad-range bacterial cloning and expression system.
614 *Nat Methods* **2**: 443-448.

615 Clark, J.I., and Muchowski, P.J. (2000) Small heat-shock proteins and their potential role in
616 human disease. *Curr Opin Struct Biol* **10**: 52-59.

617 Cocotl-Yanez, M., Moreno, S., Encarnacion, S., Lopez-Pliego, L., Castaneda, M., and Espin,
618 G. (2014) A small heat-shock protein (Hsp20) regulated by RpoS is essential for cyst
619 desiccation resistance in *Azotobacter vinelandii*. *Microbiology* **160**: 479-487.

620 De Soyza, A., Hall, A.J., Mahenthiralingam, E., Drevinek, P., Kaca, W., Drulis-Kawa, Z. *et al.*
621 (2013) Developing an international *Pseudomonas aeruginosa* reference panel.
622 *Microbiologyopen* **2**: 1010-1023.

623 Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004) Genomic islands in
624 pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**: 414-424.

625 Finck-Barbancon, V., Goranson, J., Zhu, L., Sawa, T., Wiener-Kronish, J.P., Fleiszig, S.M. *et*
626 *al.* (1997) ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity
627 and epithelial injury. *Mol Microbiol* **25**: 547-557.

628 Fu, X., Jiao, W., and Chang, Z. (2006) Phylogenetic and biochemical studies reveal a
629 potential evolutionary origin of small heat shock proteins of animals from bacterial class A. *J*

630 *Mol Evol* **62**: 257-266.

631 Gajdosova, J., Benedikovicova, K., Kamodyova, N., Tothova, L., Kaclikova, E., Stuchlik, S.
632 *et al.* (2011) Analysis of the DNA region mediating increased thermotolerance at 58 degrees
633 C in *Cronobacter sp.* and other enterobacterial strains. *Antonie Van Leeuwenhoek* **100**: 279-
634 289.

635 Grobe, S., Wingender, J., and Truper, H.G. (1995) Characterization of mucoid *Pseudomonas*
636 *aeruginosa* strains isolated from technical water systems. *J Appl Bacteriol* **79**: 94-102.

637 Haslbeck, M., and Vierling, E. (2015) A first line of stress defense: small heat shock proteins
638 and their function in protein homeostasis. *J Mol Biol* **427**: 1537-1548.

639 Haslbeck, M., Franzmann, T., Weinfurtner, D., and Buchner, J. (2005) Some like it hot: the
640 structure and function of small heat-shock proteins. *Nat Struct Mol Biol* **12**: 842-846.

641 He, J., Baldini, R.L., Deziel, E., Saucier, M., Zhang, Q., Liberati, N.T. *et al.* (2004) The broad
642 host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands
643 harboring plant and animal virulence genes. *Proc Natl Acad Sci U S A* **101**: 2530-2535.

644 Hilker, R., Munder, A., Klockgether, J., Moran Losada, P., Chouvarine, P., Cramer, N. *et al.*
645 (2014) Interclonal gradient of virulence in the *Pseudomonas aeruginosa* pangenome from
646 disease and environment. *Environ Microbiol.*

647 Jakob, U., Gaestel, M., Engel, K., and Buchner, J. (1993) Small heat shock proteins are
648 molecular chaperones. *J Biol Chem* **268**: 1517-1520.

649 Kappe, G., Leunissen, J.A., and de Jong, W.W. (2002) Evolution and diversity of prokaryotic
650 small heat shock proteins. *Prog Mol Subcell Biol* **28**: 1-17.

651 Kerr, K.G., and Snelling, A.M. (2009) *Pseudomonas aeruginosa*: a formidable and ever-
652 present adversary. *J Hosp Infect* **73**: 338-344.

653 Kidd, T.J., Grimwood, K., Ramsay, K.A., Rainey, P.B., and Bell, S.C. (2011) Comparison of

three molecular techniques for typing *Pseudomonas aeruginosa* isolates in sputum samples from patients with cystic fibrosis. *J Clin Microbiol* **49**: 263-268.

Klockgether, J., Reva, O., Larbig, K., and Tummler, B. (2004) Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J Bacteriol* **186**: 518-534.

Klockgether, J., Wurdemann, D., Reva, O., Wiehlmann, L., and Tummler, B. (2007) Diversity of the abundant pKLC102/PAGI-2 family of genomic islands in *Pseudomonas aeruginosa*. *J Bacteriol* **189**: 2443-2459.

Kortmann, J., and Narberhaus, F. (2012) Bacterial RNA thermometers: molecular zippers and switches. *Nat Rev Microbiol* **10**: 255-265.

Krajewski, S.S., Nagel, M., and Narberhaus, F. (2013) Short ROSE-like RNA thermometers control IbpA synthesis in *Pseudomonas* species. *PLoS One* **8**: e65168.

Kriehuber, T., Rattei, T., Weinmaier, T., Bepperling, A., Haslbeck, M., and Buchner, J. (2010) Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *FASEB J* **24**: 3633-3642.

Kuczynska-Wisnik, D., Matuszewska, E., and Laskowska, E. (2010) *Escherichia coli* heat-shock proteins IbpA and IbpB affect biofilm formation by influencing the level of extracellular indole. *Microbiology* **156**: 148-157.

Kulasekara, B.R., Kulasekara, H.D., Wolfgang, M.C., Stevens, L., Frank, D.W., and Lory, S. (2006) Acquisition and evolution of the *exoU* locus in *Pseudomonas aeruginosa*. *J Bacteriol* **188**: 4037-4050.

Kung, V.L., Ozer, E.A., and Hauser, A.R. (2010) The accessory genome of *Pseudomonas aeruginosa*. *Microbiol Mol Biol Rev* **74**: 621-641.

Laksanalamai, P., Maeder, D.L., and Robb, F.T. (2001) Regulation and mechanism of action of the small heat shock protein from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J*

678 *Bacteriol* **183**: 5198-5202.

679 Larbig, K.D., Christmann, A., Johann, A., Klockgether, J., Hartsch, T., Merkl, R. *et al.* (2002)

680 Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas*

681 *aeruginosa* clone. *J Bacteriol* **184**: 6665-6680.

682 Laue, T.M. (1995) Sedimentation equilibrium as thermodynamic tool. *Methods Enzymol* **259**:

683 427-452.

684 Lebowitz, J., Lewis, M.S., and Schuck, P. (2002) Modern analytical ultracentrifugation in

685 protein science: a tutorial review. *Protein Sci* **11**: 2067-2079.

686 Lee, C., Peters, V., Melefors, O., and Romling, U. (2014) Draft Genome Sequence of

687 *Pseudomonas aeruginosa* SG17M, an Environmental Isolate Belonging to Clone C, Prevalent

688 in Patients and Aquatic Habitats. *Genome Announc* **2**.

689 Lindner, A.B., Madden, R., Demarez, A., Stewart, E.J., and Taddei, F. (2008) Asymmetric

690 segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc*

691 *Natl Acad Sci U S A* **105**: 3076-3081.

692 Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999) EMAN: semiautomated software for high-

693 resolution single-particle reconstructions. *J Struct Biol* **128**: 82-97.

694 Lyczak, J.B., Cannon, C.L., and Pier, G.B. (2002) Lung infections associated with cystic

695 fibrosis. *Clin Microbiol Rev* **15**: 194-222.

696 Malki, A., Kern, R., Abdallah, J., and Richarme, G. (2003) Characterization of the

697 *Escherichia coli* YedU protein as a molecular chaperone. *Biochem Biophys Res Commun* **301**:

698 430-436.

699 Mathee, K., Narasimhan, G., Valdes, C., Qiu, X., Matewish, J.M., Koehrsen, M. *et al.* (2008)

700 Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proc Natl Acad Sci U S A* **105**:

701 3100-3105.

702 Mikkelsen, H., McMullan, R., and Filloux, A. (2011) The *Pseudomonas aeruginosa* reference
703 strain PA14 displays increased virulence due to a mutation in *ladS*. *PLoS One* **6**: e29113.

704 Mogk, A., Deuerling, E., Vorderwulbecke, S., Vierling, E., and Bukau, B. (2003) Small heat
705 shock proteins, ClpB and the DnaK system form a functional triade in reversing protein
706 aggregation. *Mol Microbiol* **50**: 585-595.

707 Narberhaus, F. (2002) Alpha-crystallin-type heat shock proteins: socializing minichaperones
708 in the context of a multichaperone network. *Microbiol Mol Biol Rev* **66**: 64-93; table of
709 contents.

710 Newman, J.R., and Fuqua, C. (1999) Broad-host-range expression vectors that carry the L-
711 arabinose-inducible *Escherichia coli* *araBAD* promoter and the *araC* regulator. *Gene* **227**:
712 197-203.

713 Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T. *et al.* (2014) The
714 SEED and the Rapid Annotation of microbial genomes using Subsystems Technology
715 (RAST). *Nucleic Acids Res* **42**: D206-214.

716 Page, R.D. (2002) Visualizing phylogenetic trees using TreeView. *Curr Protoc*
717 *Bioinformatics* **Chapter 6**: Unit 6 2.

718 Papaioannou, E., Utari, P.D., and Quax, W.J. (2013) Choosing an appropriate infection model
719 to study quorum sensing inhibition in *Pseudomonas* infections. *Int J Mol Sci* **14**: 19309-
720 19340.

721 Petit, S.M., Lavenir, R., Colinon-Dupuich, C., Boukerb, A.M., Cholley, P., Bertrand, X. *et al.*
722 (2013) Lagooning of wastewaters favors dissemination of clinically relevant *Pseudomonas*
723 *aeruginosa*. *Res Microbiol* **164**: 856-866.

724 Pirnay, J.P., Matthijs, S., Colak, H., Chablain, P., Bilocq, F., Van Eldere, J. *et al.* (2005)
725 Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. *Environ*

726 *Microbiol* **7**: 969-980.

727 Rahme, L.G., Ausubel, F.M., Cao, H., Drenkard, E., Goumnerov, B.C., Lau, G.W. *et al.* (2000)

728 Plants and animals share functionally common bacterial virulence factors. *Proc Natl Acad Sci*

729 *U S A* **97**: 8815-8821.

730 Rakhimova, E., Wiehlmann, L., Brauer, A.L., Sethi, S., Murphy, T.F., and Tummeler, B. (2009)

731 *Pseudomonas aeruginosa* population biology in chronic obstructive pulmonary disease. *J*

732 *Infect Dis* **200**: 1928-1935.

733 Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new

734 ENDscript server. *Nucleic Acids Res* **42**: W320-324.

735 Romling, U., Schmidt, K.D., and Tummeler, B. (1997) Large genome rearrangements

736 discovered by the detailed analysis of 21 *Pseudomonas aeruginosa* clone C isolates found in

737 environment and disease habitats. *J Mol Biol* **271**: 386-404.

738 Romling, U., Wingender, J., Muller, H., and Tummeler, B. (1994a) A major *Pseudomonas*

739 *aeruginosa* clone common to patients and aquatic habitats. *Appl Environ Microbiol* **60**: 1734-

740 1738.

741 Romling, U., Kader, A., Sriramulu, D.D., Simm, R., and Kronvall, G. (2005) Worldwide

742 distribution of *Pseudomonas aeruginosa* clone C strains in the aquatic environment and

743 cystic fibrosis patients. *Environ Microbiol* **7**: 1029-1038.

744 Romling, U., Fiedler, B., Bosshammer, J., Grothues, D., Greipel, J., von der Hardt, H., and

745 Tummeler, B. (1994b) Epidemiology of chronic *Pseudomonas aeruginosa* infections in cystic

746 fibrosis. *J Infect Dis* **170**: 1616-1621.

747 Rosenzweig, R., Moradi, S., Zarrine-Afsar, A., Glover, J.R., and Kay, L.E. (2013) Unraveling

748 the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science* **339**:

749 1080-1083.

750 Roy-Burman, A., Savel, R.H., Racine, S., Swanson, B.L., Revadigar, N.S., Fujimoto, J. *et al.*
 751 (2001) Type III protein secretion is associated with death in lower respiratory and systemic
 752 *Pseudomonas aeruginosa* infections. *J Infect Dis* **183**: 1767-1774.
 753 Sastry, M.S., Korotkov, K., Brodsky, Y., and Baneyx, F. (2002) Hsp31, the *Escherichia coli*
 754 *yedU* gene product, is a molecular chaperone whose activity is inhibited by ATP at high
 755 temperatures. *J Biol Chem* **277**: 46026-46034.
 756 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity
 757 ultracentrifugation and lamm equation modeling. *Biophys J* **78**: 1606-1619.
 758 Selezska, K., Kazmierczak, M., Musken, M., Garbe, J., Schobert, M., Haussler, S. *et al.*
 759 (2012) *Pseudomonas aeruginosa* population structure revisited under environmental focus:
 760 impact of water quality and phage pressure. *Environ Microbiol* **14**: 1952-1967.
 761 Seyfffer, F., Kummer, E., Oguchi, Y., Winkler, J., Kumar, M., Zahn, R. *et al.* (2012) Hsp70
 762 proteins bind Hsp100 regulatory M domains to activate AAA+ disaggregase at aggregate
 763 surfaces. *Nat Struct Mol Biol* **19**: 1347-1355.
 764 Shanks, R.M., Caiazza, N.C., Hinsa, S.M., Toutain, C.M., and O'Toole, G.A. (2006)
 765 *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from gram-
 766 negative bacteria. *Appl Environ Microbiol* **72**: 5027-5036.
 767 Sorek, R., Kunin, V., and Hugenholtz, P. (2008) CRISPR--a widespread system that provides
 768 acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* **6**: 181-186.
 769 Sriramulu, D.D., Nimtz, M., and Romling, U. (2005) Proteome analysis reveals adaptation of
 770 *Pseudomonas aeruginosa* to the cystic fibrosis lung environment. *Proteomics* **5**: 3712-3721.
 771 Stege, G.J., Renkawek, K., Overkamp, P.S., Verschuure, P., van Rijk, A.F., Reijnen-Aalbers,
 772 A. *et al.* (1999) The molecular chaperone alphaB-crystallin enhances amyloid beta
 773 neurotoxicity. *Biochem Biophys Res Commun* **262**: 152-156.

774 Studer, S., and Narberhaus, F. (2000) Chaperone activity and homo- and hetero-oligomer
775 formation of bacterial small heat shock proteins. *J Biol Chem* **275**: 37212-37218.

776 Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007)
777 EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* **157**: 38-
778 46.

779 Thomas, J.G., and Baneyx, F. (1998) Roles of the *Escherichia coli* small heat shock proteins
780 IbpA and IbpB in thermal stress management: comparison with ClpA, ClpB, and HtpG In
781 vivo. *J Bacteriol* **180**: 5165-5172.

782 Tielen, P., Narten, M., Rosin, N., Biegler, I., Haddad, I., Hogardt, M. *et al.* (2011) Genotypic
783 and phenotypic characterization of *Pseudomonas aeruginosa* isolates from urinary tract
784 infections. *Int J Med Microbiol* **301**: 282-292.

785 Tomoyasu, T., Mogk, A., Langen, H., Goloubinoff, P., and Bukau, B. (2001) Genetic
786 dissection of the roles of chaperones and proteases in protein folding and degradation in the
787 *Escherichia coli* cytosol. *Mol Microbiol* **40**: 397-413.

788 Usui, K., Yoshida, T., Maruyama, T., and Yohda, M. (2001) Small heat shock protein of a
789 hyperthermophilic archaeum, *Thermococcus* sp. strain KS-1, exists as a spherical 24 mer and
790 its expression is highly induced under heat-stress conditions. *J Biosci Bioeng* **92**: 161-166.

791 Van Domselaar, G.H., Stothard, P., Shrivastava, S., Cruz, J.A., Guo, A., Dong, X. *et al.* (2005)
792 BASys: a web server for automated bacterial genome annotation. *Nucleic Acids Res* **33**:
793 W455-459.

794 Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R.A., Rouault, T.A., and Schuck, P.
795 (2004) Sedimentation equilibrium analysis of protein interactions with global implicit mass
796 conservation constraints and systematic noise decomposition. *Anal Biochem* **326**: 234-256.

797 Waters, E.R. (1995) The molecular evolution of the small heat-shock proteins in plants.

798 *Genetics* **141**: 785-795.

799 Winsor, G.L., Lam, D.K., Fleming, L., Lo, R., Whiteside, M.D., Yu, N.Y. *et al.* (2011)

800 *Pseudomonas* Genome Database: improved comparative analysis and population genomics

801 capability for *Pseudomonas* genomes. *Nucleic Acids Res* **39**: D596-600.

802 WL, D. (2002) The PyMOL Molecular Graphics System. In. San Carlos, CA: DeLano

803 Scientific.

804 Wolfgang, M.C., Kulasekara, B.R., Liang, X., Boyd, D., Wu, K., Yang, Q. *et al.* (2003)

805 Conservation of genome content and virulence determinants among clinical and

806 environmental isolates of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **100**: 8484-

807 8489.

808

809

810

811

Figure legends

Figure 1. Insertion of clone C-specific genomic island PACGI-1 at the CRISPR locus in strain SG17M. Chromosomal organization of the CRISPR genomic region in SG17M, including the CRISPRs and CRISPR-associated (*cas*, *csy*) genes. The 86 kb clone C-specific genomic island PACGI-1 is integrated in the 7th spacer (black boxes) between the 7th and 8th repeat of CRISPR 2.

Figure 2. Identification of sHsp20c as a clone C-specific protein. (A) The total proteomic pattern was compared between PAO1 and clone C strain C by two-dimensional gel electrophoresis. Clone C-specific expression of protein was detected at 21.4 kDa and pI 5.33, which was subsequently identified as sHsp20c by MALDI-TOF analysis. 2D gels reproduced with permission from Wiley-VCH from (31). (B) Phylogenetic tree of sHsp20c with sHsps (protein identities as in Fig. S2). (C) Sequence alignment of sHsp20c with sHsps IbpA and IbpA2 from *P. aeruginosa* and selected other species with protein identities as in Fig. S2 (Chenna *et al.*, 2003). Secondary structure annotation corresponding to results from homology modelling shown on top with α -helices represented by spirals, 3_{10} -helices denoted by η , β -strands shown as arrows and strict β -turns as TT. Strictly conserved residues are shaded in red and conserved regions are boxed by blue frames. Consensus sequence at bottom is indicated in uppercase for residues with 100% identity and in lowercase for residues higher than 50%. Predicted domain organization is shown with the same color coding as in (D). Alignment was carried out using the program ClustalW2 and the result was modified for the figure using the program ESPript (Robert and Gouet, 2014) (D) Structural model of monomeric sHsp20c (Ser42-Gly189) shown as cartoon representation colored according to predicted domain organization with the N-terminal domain, the α -crystallin domain and the C-term extension colored yellow, blue and green, respectively, and the dotted

yellow line representing the N-terminal residues that could not be modelled. Model was prepared using the coordinates of the 30.2% identical heat shock protein 16.0 from *Schizosaccharomyces pombe* ((PDB: 3w1z). (E) Expression of sHsp20c in clone C and J isolates was detected by Western blot analysis with sHsp20c specific antibody. The strains include environmental clone J 1: PT1M, Clone C 2: PT31M, 3: W5Aug28, 4: W15Aug28, 5: 43A, 6: 10A, 7: SG17M; and clinical isolates clone J 1: P7-2, 2: P7-37, 3: P7-38, Clone C 4: P7-1, 5: P7-2, 6: P7-8, 7: SG17M. (F) Occurrence of *shsp20c* detected in 56 environmental and 61 clinical isolates belonging to different clonal groups ($\chi^2=0.037$).

Figure 3. Susceptibility of *shsp20c* mutants to heat and oxidative stress. Sensitivity of SG17M and respective mutants to heat shock (A) and hydrogen peroxide (B). To assess heat shock resistance, cells were kept at 50°C for 30 and 60 mins. Cells were treated with 30 and 60 mM hydrogen peroxide for 90 mins. Bars show the mean of three technical replicates of one representative experiment and error bars indicate standard deviation (*, ** $p<0.001$; *** $p<0.05$; * compared with SG17M). (C) sHsp20c harboring strains are less susceptible to heat shock at 50°C for 30 min than sHsp20c negative isolates. (D) sHsp20c harboring strains are equally sensitive to oxidative stress than sHsp20c negative isolates. The sHsp20c negative, non-clone C group consisted of strains: SG9M, PT4M, ATCC33818, A5787 and PA14. The sHsp20 positive group included P4-1, 8277, 8735, SG31M, PT1M (clone J) and SG17M. (* $p<0.002$). (E) Expression of sHsp20c in clone PA14, SG9M and derivatives as detected by Western blot analysis. Cells were grown in LB medium at 20°C for 48h. Samples: 1, SG17M; 2, SG17M $\Delta shsp20c$; 3, PA14; 4, PA14 *attTn7::dna shsp20c*⁺; 5, PA14 *attTn7::dna shsp20c clpBc*⁺; 6, SG9M; 7, SG9M *attTn7::dna shsp20c*⁺; 8, SG9M *attTn7::dna shsp20c clpBc*⁺ (F) Introduction of *dna shsp20c* and *dna shsp20c clpBc* at the Tn7 site enhances heat resistance in PA14 and SG9M. Bars show the mean of three technical replicates of one

representative experiment and error bars indicate standard deviation (*, ** $p < 0.005$; *compared to corresponding wild type). Samples: 1, PA14; 2, PA14 *attTn7::dna shsp20c*⁺; 3, PA14 *attTn7::dna shsp20c clpBc*⁺; 4, SG9M; 5, SG9M *attTn7::dna shsp20c*⁺; 6, SG9M *attTn7::dna shsp20c clpBc*⁺.

Figure 4. Unconventional expression of sHsp20c and inducibility by heat and oxidative stress. sHsp20c is highly expressed in stationary phase cells on solid (A) and in liquid medium (B). (A) SG17M cultured at 20°C in LB and M63-citrate medium were collected at different time points; LB medium: lane 1: 24, 2: 36, 3: 48, 4: 60, 5: 72 h and M63 medium: lane 1: 31, 2: 43, 3: 55, 4: 67, 5: 79 h. Lane 6, positive control of SG17M grown on LB-agar medium at 37°C for 18 h. (B) SG17M cells were harvested at different O.D.₆₀₀ in liquid medium. Lanes, 1: 1; 2: 2.5; 3: 4; 4: 4.8 and 1: 1; 2: 1.5; 3: 1.7; 4: 1.8 of O.D.₆₀₀ value for LB and M63-citrate medium respectively. (C) sHsp20c is expressed at 20, 27 and 37°C on M63-citrate solid medium. Strains PT1M and SG17M were incubated 72, 48 and 24 h at 20, 27 and 37°C respectively. (D) sHsp20c is induced by hydrogen peroxide treatment in SG17M strain. SG17M was treated with hydrogen peroxide for 1 h at 37°C in LB medium. (E) SG17M was grown at 20°C for 4 h until O.D.₆₀₀ 0.2, and then aliquots were incubated at 20, 37 and 42°C respectively. At certain time points, cells were collected, the O.D.₆₀₀ was measured and sHsp20c production was assessed by Western blot analysis. (F) sHsp20c is preferentially associated with the aggregated cell fraction. S, soluble protein; A, aggregated protein. Cells were grown in LB medium at 37°C for 24 h. (G) sHsp20c expression is higher in planktonic cells compared with cell aggregates at both 20 and 37°C. (H) Successful separation of the culture in LB medium at 37°C for 24 h with shaking (left) into aggregated and planktonic cells was assessed by light transmission microscopy. Scale bar: 5 µm.

Figure 5. Oligomeric complex of sHsp20c and its chaperone activity as well as thermal

stability. (A) Size-exclusion chromatography analysis of oligomeric complexes formed by sHsp20c performed on a Superdex 200 prep grade column monitoring the absorbance at 280 nm. The elution profiles of non- (red) and 6XHis-tagged (blue) sHsp20c are shown as solid lines with molecular mass standards from calibration of the column indicated above the chromatograms. Purified 6XHis-sHsp20c protein was analyzed by SDS-PAGE before performing SEC (inset). (B) The chaperone activity of 6XHis-sHsp20c was monitored by prevention of thermal aggregation of CS at different molar ratio. BSA and α -crystallin proteins were negative and positive controls for chaperone activity, respectively. (C) Far-UV circular dichroism (CD) spectra of 6XHis-sHsp20c recorded in the temperature range 20 to 70 °C.

Figure 6. Molecular mass and hydrodynamic properties of sHsp20c. (A) Exemplary $c(s)$ distributions of 4 μ M (red), 8 μ M (yellow), 16 μ M (cyan) and 32 μ M 6XHis-tagged sHsp20c (blue) are shown as obtained by SEDFIT analysis (Schuck, 2000) of a sedimentation velocity experiment at 25000 rpm with a detection wavelength of 280 nm. sHsp20c sediments as a single species at all protein concentrations. Extrapolation to zero concentration yielded $s_{20,w}^0 = 18.1$ S. (B) Sedimentation equilibrium experiment at a loading concentration of 12 μ M 6XHis-tagged sHsp20c at 3500 rpm (blue circles), 5000 rpm (cyan circles), 7000 rpm (yellow circles) and 9000 rpm (red circles) using interference detection. Global analysis of data sets measured at 6 μ M, 12 μ M and 20 μ M, all four rotor speeds and with detection of interference and absorbance at 280 nm; using a single species model yielded a molecular mass of 547 kDa (solid lines). In the lower panel differences between modelled and measured values as a function of radial position are given (note different scales of upper and lower panel).

Figure 7. Oligomeric structure and shape of sHsp20c by transmission electron microscopy.

907 (A) Survey view of negatively stained particles. Large circles: individual molecules, small
908 circles: subunit dimers/molecular debris, ellipses: doublets. (B) Representative filtered class-
909 averages. Arrows indicate substructural centers of mass. (C) Class-averages after 12
910 refinement cycles with d6-symmetry (Cutout).

911

Fig. 1

912

913

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6

922

923

Fig. 7

924