Protein oxidation is known to compromise vital cellular functions. Therefore, invading pathogenic bacteria must resist damage inflicted by host defenses via reactive oxygen species. Using comparative genomics and experimental approaches, we provide multiple lines of evidence that proteins from pathogenic bacteria have acquired resistance to oxidative stress by an increased conformational stability. Representative pathogens exhibited higher survival upon HSP90 inhibition and a less-oxidation-prone proteome. A proteome signature of the 46 pathogenic bacteria encompasses 14 physicochemical features related to increasing protein conformational stability. By purifying ten representative proteins, we demonstrate in vitro that proteins with a pathogen-like signature are more resistant to oxidative stress as a consequence of their increased conformational stability. A compositional signature of the pathogens’ proteomes allowed the design of protein fragments more resilient to both unfolding and carbonylation, validating the relationship between conformational stability and oxidability with implications for synthetic biology and antimicrobial strategies.
proteins are indeed more resistant to oxidation as a direct consequence of their increased conformational stability.

RESULTS

Pathogen Proteomes Are More Resistant to Oxidation In Vitro and to HSP90 Inhibition In Vivo

We selected six representative pathogenic and six nonpathogenic bacteria and tested their total protein extracts for susceptibility to oxidation. We found that protein extracts from representative pathogens were on average 1.9-fold more resistant to carbonylation than the extracts from nonpathogens (p = 0.0002) (Figure 1A; species names are listed in Table S1).

Furthermore, we measured the resistance of the 12 bacteria to geldanamycin, a known inhibitor of the HSP90 chaperone. An intrinsically stable proteome should be able to withstand a HSP90 deficiency in normal conditions, reflected in high cell survival and a low level of geldanamycin-induced protein damage. Indeed, geldanamycin exposure resulted in 19% average survival for the nonpathogens (p < 0.0001) (Figure 1B; species names are listed in Table S1).

Furthermore, we measured the resistance of the 12 bacteria to geldanamycin, a known inhibitor of the HSP90 chaperone. An intrinsically stable proteome should be able to withstand a HSP90 deficiency in normal conditions, reflected in high cell survival and a low level of geldanamycin-induced protein damage. Indeed, geldanamycin exposure resulted in 19% average survival for the nonpathogens (p < 0.0001) (Figure 1B; species names are listed in Table S1).

Furthermore, we evaluated the protection against geldanamycin granted by the antioxidant enzymes KatE, KatG, and SodD and by the small-molecule antioxidant N-acetylcysteine (NAC) in E. coli. Regardless of the source of the antioxidant activity, it can account for, on average, only 18% of the difference in geldanamycin resistance observed between the pathogens and nonpathogens reported in Figure 1B (Supplemental Results; Figure S1D). Namely, while the four sources of antioxidant activity yielded an average 30% increase in survival in the presence of geldanamycin, the average survival difference between the six pathogenic and six nonpathogenic bacteria was 3.0-fold (Figure 1B). Consistent trends were observed in the level of protein carbonylation (Supplemental Results; Figure S1E). In other words, the extent of proteome oxidation protection that can be obtained by increased amounts of antioxidants appears to be small in comparison to the differences we observed between pathogenic and nonpathogenic bacteria. Thus, such differences likely also stem from other sources, such as intrinsic properties of the proteome.

Pathogenicity Is Reflected in the General Trends of Variation between Proteome Physicochemical Properties

Next, we investigated whether the resilience of the six chosen pathogens to HSP90 inhibition and proteome oxidation is representative of a broader trend across more microbes. We thus
compared proteomes of 46 known pathogenic to 40 matched nonpathogenic bacteria according to their overall physicochemical properties, calculated from the genome-encoded amino acid composition (Supplemental Experimental Procedures). In a principal component (PC) analysis, which captures the salient differences between bacterial proteome composition (73% variability in first two PCs; Figure S2), the pathogen proteomes were significantly shifted from the nonpathogens, and the difference was observed separately in two bacterial phyla (Figures 2A and 2B; p = 0.0035 and 0.0009, two-tailed Kolmogorov-Smirnov test). Crucially, we observed a strong agreement between the direction of the pathogenicity-related trend in Firmicutes (Figure S2A) and in γ-proteobacteria (Figure S2B), which is suggestive of a convergent evolution toward a common pathogen-like proteome composition.

Next, we combined the first two PCs into the summary statistic PaPS (the pathogen-like proteome signature) that maximizes the separation between the pathogen and nonpathogen proteomes (Experimental Procedures); pathogens have significantly lower PaPS values (Figures 2A and 2B). Upon examining the contribution of individual amino acid frequencies to PaPS by multiple regression, we found strongly significant contributions of the secondary-structure breakers Gly and Pro (Figure 2C; p < 0.0001 for both) followed by Arg and Ser (p < 0.0001 for both), all of which are known to have disorder-promoting effects (Uversky, 2013). The avoidance of secondary-structure-disfavoring amino acids in the pathogens’ proteomes suggests a general trend toward more structured proteins. An examination of the physicochemical properties strongly correlated with PaPS (Figure 2D) indicated that pathogens favor bulkier, polar, and charged residues (generally having high β sheet propensities and stabilizing effects) but disfavor residues with high propensity for extended structures. In addition, pathogen proteomes tend toward a lower free energy of α helices. A direct comparison of amino acid propensities between the
pathogens and nonpathogens is broadly consistent with the above, with an avoidance of Gly and Pro in pathogens (Figures S2C–S2E; p ≤ 0.05 and ≤ 0.016, respectively, Mann-Whitney test) and an enrichment for the order-promoting Phe and Tyr (Figures S2C–S2E; p ≤ 0.045 and p ≤ 0.006, respectively).

While the PaPS scores were, overall, highly significantly different between the pathogens and nonpathogens (Figures 2A and 2B), there were also exceptions to this trend. Such cases may reflect the subtleties in the definition of pathogenicity that are difficult to capture using a binary classification. For instance, two lowly virulent pathogens clustered with nonpathogens, and two insect-only pathogens, clustered with the vertebrate pathogens (Supplemental Discussion; organisms highlighted in Figure S2). Moreover, two nonpathogenic Listeria cluster tightly with a close pathogenic relative that they have recently diverged from (Supplemental Discussion).

The observed proteome signature of elevated conformational stability in pathogens could theoretically result from other adaptations (for instance, to the hosts’ body temperatures). However, a regression of the organisms’ PaPS scores against seven bacterial phenotypes indicates specifically the pathogenicity as the strongest contributor to PaPS, as opposed to, e.g., the optimal growth temperature (Figures S3A and S3B; phenotypes listed in Table S2). Moreover, a known thermophilic proteome signature (Zeldovich et al., 2007) does not strongly relate to PaPS or to the composition of pathogen proteomes (Figures S3C and S3D). Furthermore, we turn to experimentally evaluate whether differences exist between the representative pathogens and nonpathogens in their ability to withstand temperature increases. Although slight differences in survival can be observed between the studied species, they cannot be attributed to their pathogenicity (Supplemental Results; Figures S3E and S3F).

To experimentally validate the trends across proteomes captured by the PaPS statistic, we investigated a series of ten purified proteins selected from both pathogenicity classes and both phylogenetic groups of bacteria (Supplemental Experimental Procedures). We chose the thymidylate synthase (ThyA) and the nicotinic acid mononucleotide adenylyltransferase protein (NadD) families, which exhibited highly significant differences of PaPS scores between the pathogens and nonpathogens (Figures 2E–2H; for NadD, p < 0.0001 and p = 0.0027 for Firmicutes and γ-proteobacteria; for ThyA, p = 0.0013 and 0.0001). Thus, these two protein families are representative of the broader across-proteome trends resulting from the PC analysis.

An Increase in Conformational Stability Is Coupled to an Increase in Oxidation Resistance

Within the NadD and ThyA protein families, we selected six and four proteins, respectively, with highly pathogen-like or highly non-pathogen-like PaPS values (highlighted in Figures 2E–H). These proteins were produced in E. coli to experimentally corroborate the link of the PaPS signature to conformational stability and oxidation resistance. By measuring protein carbonylation kinetics, we found that ThyA and NadD proteins from pathogens were significantly more resistant to oxidation relative to their homologs from nonpathogens (Figures 3A and 3B). In particular, the ThyA and NadD with non-pathogen-like PaPS were 68% and 40% more carbonylated (p = 0.0004 and 0.0005), respectively, after Fenton-reaction-induced oxidation relative to the pathogen-like ThyA and NadD. Consistently, ThyA from control pathogens (atypical pathogen proteins that have non-pathogen-like PaPS) oxidized similarly to the ones from nonpathogenic bacteria. This demonstrates that it is specifically the PaPS signature and not some other feature of the pathogens’ proteome composition that leads to the oxidation resistance.

Furthermore, to determine their conformational stability and its potential correlation to oxidation resistance, we measured the absorbance at 280 nm during thermal denaturation of ThyA and NadD. Purified ThyA proteins from the two pathogenic bacteria had a melting temperature (T_m) on average 23°C higher than the nonpathogen ThyA (Figures 3C and S4A). Similarly, NadD showed a 21°C average difference (Figures 3C and S4B) (p = 0.0001). On the contrary, ThyA proteins from the negative control pathogens (pathogen proteins with a non-pathogen-like PaPS; see Experimental Procedures) had a T_m on average 26°C lower than their homologs from pathogens, thus resembling ThyA from the nonpathogenic bacteria. The large differences in the melting temperature indicate a highly increased conformational stability of the proteins with a pathogen-like PaPS signature. We further tested the link between the two variables by directly modifying the conformational stability using trimethylamine N-oxide (TMAO), an osmolyte known to stabilize protein structures.

Exposure of ThyA and NadD from nonpathogenic bacteria to 1 M TMAO increased the T_m of these proteins by 15°C on average (p = 0.011), indicating an increase in conformational stability (Figure 3C; Table S3). Such TMAO-stabilized ThyA and NadD also displayed an increased resistance to oxidation in vitro (Figure 3D); the levels of protein carbonylation were on average 37% lower for ThyA and 23% lower for NadD (p < 0.0001 for both; Figure 3D). The control ThyA proteins (proteins from pathogens characterized by non-pathogen-like PaPS) showed similar trends (Figure 3D), with an average decrease of protein carbonylation of 24% following TMAO stabilization (p < 0.0001). Furthermore, while stabilization of the proteins makes them more oxidation resistant, the converse also holds; denaturation of all tested proteins by guanidinium hydrochloride (GuHCl) rendered them all extremely sensitive to oxidation (Figure 3E; 2.7-fold average carbonylation increase, p < 0.0001). The experiments with external protein stabilization (via TMAO) or destabilization (via GuHCl) corroborate the hypothesis that the protective role of native protein structure against oxidation stems specifically from conformational stability and not from an unrelated property of structure and/or sequence composition.

The PaPS Signature Is Sufficient to Design Conformationally Stable and Oxidation-Resistant Peptides

We aimed to further establish a causal link between the specific proteome signature recognized herein, PaPS, and the detected oxidation resistance and conformational stability. To this end, we have designed a series of six helix-turn-helix (HTH) and six β-hairpin (BH) peptides with amino acid composition selected according to PaPS (see Experimental Procedures; peptides are listed in Table S4).
Following Fenton-reaction-induced oxidation, the HTH (Figure 4A) and BH (Figure 4B) peptides with non-pathogen-like PaPS values exhibited an average 30% increase in carbonylation (p < 0.0001), unlike the peptides with pathogen-like PaPS, which did not exhibit an increase in their carbonylation level. The negative control peptides (which had a different overall composition while retaining non-pathogen-like PaPS values; see Experimental Procedures) again displayed oxidation resistance resembling the non-pathogen-like peptides. Furthermore, the peptides with pathogen-like PaPS displayed an increased conformational stability, with a Tm 13°C higher for HTH and 20°C higher for the BH than the corresponding non-pathogen-like peptides (p < 0.0001) (Figures 4C and S4D–S4F). The negative control peptides resembled the non-pathogen-like peptides; their Tm is 18°C higher than the pathogen-like peptides (Figures 4C and S4D–S4F).

Conformational stability of the synthetic peptides correlates with their oxidation resistance, thus establishing that the relationship between oxidation resistance and conformational stability is not limited to natural proteins.

**DISCUSSION**

Even though pathogenic bacteria are characterized by different mechanisms of how they cause disease, common hallmarks of infectivity exist across evolutionarily distant pathogenic species.
In our study, we investigated the characteristics of overall proteome composition shared between pathogenic bacteria and searched for a hypothesized link to tolerance to oxidative stress, a challenge faced by pathogens during host invasion (Rosenberger and Finlay, 2003).

As a consequence of oxidative damage, proteins are unable to maintain their native structure and thus perform their function with a decreased efficiency (Krisko and Radman, 2010), leading to elevated mutation rates and fueling a vicious cycle of protein and genetic damage (Krisko and Radman, 2013). Although proteins can undergo different oxidative modifications in a number of ways, carbonylation is a common irreversible type of damage elevated during oxidative stress (Nystro¨ m, 2005). Importantly, even in conditions of physiological ROS levels, misfolding of proteins increases their susceptibility to carbonylation in vivo, for instance during heat shock or streptomycin or puromycin treatments (Dukan et al., 2000). However, the natively folded proteins can also be carbonylated, and the individual proteins display extreme inequalities in their resistance to carbonylation, with ~10% of the E. coli proteome being highly susceptible (Maisonneuve et al., 2009). The sequence or structure determinants of these differences are not clear, but the strong tendency of carbonylation-prone sites (Maisonneuve et al., 2009) intrinsically more resistant to misfolding might also be more resistant to carbonylation. Finally, we demonstrate in E. coli that solely the level of ROS-scavenging enzymes or the presence of small-molecule scavengers (NAC) cannot explain the full extent of ROS resistance differences among pathogen proteomes (Figure S1). Rather, we propose that an intrinsic property of proteins is responsible for the extra ROS resistance not explained by other mechanisms we tested.

Our work provides multiple lines of experimental evidence that demonstrate a causal relationship between the conformational stability of a folded protein structure and its resilience to carbonylation. Furthermore, the PaPS score that estimates structural stability was derived from proteome composition of pathogenic bacteria, suggesting a link between PaPS and pathogenicity, although alternative (or complementary) explanations for PaPS are possible. Our data suggest that optimal growth temperatures, doubling times, or aerobicity does not fully explain PaPS (Figure S3). In addition, PaPS could be indirectly linked to pathogenicity via the genomic G+C content. Host-associated bacteria tend toward lower G+C than free-living ones, and thus pathogens also have somewhat lower G+C overall. Genomic G+C is known to influence the amino acid composition of the proteome (Moura et al., 2013), where low G+C corresponds to a proteome more stable against unfolding (Mendez et al.,...
Thus, the observed signature of conformational stability in pathogenic proteomes might be, in part, an indirect effect of selection (or relaxation thereof) on G+C content (reviewed in Rocha and Fell, 2010). Nevertheless, genomic G+C does not appear to fully explain our experimental observations; three of the six representative pathogens had moderately high (S. typhimurium [52%] and K. pneumoniae [57%]) or very high G+C (P. aeruginosa [67%]) but displayed similar geldanamycin resistance as the three lower G+C pathogens (Figure 1B) as well as similar proteome resistance to oxidation (Figures 1A and 1C). Pathogenicity is a multifaceted phenotype, and the G+C content and the associated proteome composition appear to be additions to the set of factors already known to influence pathogenicity.

To conclude, increased conformational stability may represent an additional mechanism by which proteins of pathogenic bacteria resist oxidation, providing an intrinsic resilience against the host defenses. Moreover, it is possible to design proteins with an enhanced conformational stability and, consequently, with an increased resistance to oxidative stress by applying the protein properties encoded in the PaPS statistic derived from pathogen proteomes. In addition to implications for synthetic biology, these findings could also contribute to the design of novel therapeutic approaches through destabilization of bacterial proteins. Such strategies could prove to be an effective means of delivering a synergistic attack on pathogen proteomes by combining protein destabilization with the ROS-inducing host defenses.

**EXPERIMENTAL PROCEDURES**

Wild-type E. coli MG1655 was cultured in liquid Luria-Bertani (LB) medium at 37°C (always supplemented with the appropriate antibiotic). Overnight cultures were diluted 100-fold in LB supplemented with isopropyl-β-D-1-thiogalactopyranoside (0.5 mM) at 37°C to the exponential phase (optical density 600 nm = 0.2–0.3).

Genes encoding ThyA and NadD were synthesized and inserted into pJexpress401 vector after codon usage optimization for E. coli MG1655 (purchased from DNA 2.0). Histidine tag (6x) was fused at the protein C terminus. Table S1 contains the list of organisms from which ThyA and NadD originate as well as the list of species’ full names from which total protein extracts were analyzed. PSEAE, STAAF, ECD01, SALTI, and KLEP3 were grown in LB medium, PSEP1 in nutrient medium; LACC3 in MRS medium; SHEON in oxoid CM3 medium; and CLOAB, CLOPS, STRAG, and BACC1 on tryptic soy medium. All strains listed were grown at 37°C to 0.2 optical density at 600 nm. CLOAB and CLOPS were grown in an oxygen-free atmosphere.

Cell Reports 7, 1393–1400, June 12, 2014 ©2014 The Authors 1399

The genes’ translations to protein were described by amino acid relative frequencies and by 59 physicochemical properties (Smole et al., 2011) from the Amino Acid Index database (Kawashima and Kanehisa, 2000) and a PC analysis performed (using XLStat 2010.6, Addinsoft) across the unweighted averages of proteins for each proteome. PaPS was defined as a rotation of the first two PCs that maximizes the separation of the pathogen and nonpathogen proteomes and pathogen unrelated proteome signature (PUPS) as orthogonal to it; details are given in Supplemental Experimental Procedures and a visualization is provided in Figure S2. The ThyA and NadD gene families were selected among 853 widespread clusters of orthologous groups (COGs) (Tatusov et al., 2003) to maximize the area under the curve score for the separation of pathogens and nonpathogens by PaPS (but not PUPS) of proteins within that COG, consistently in Firmicutes and γ-proteobacteria (Supplemental Experimental Procedures).

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Results, Supplemental Discussion, Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.04.057.

**ACKNOWLEDGMENTS**

The authors thank Miroslav Radman for stimulating discussions and Tea Copic for technical support. This work was supported by the Mediterranean Institute for Life Sciences. A.V. and A.N. are supported by the City of Split Fellowship for Excellence. The work of F.S. was supported in part by Marie Curie Actions; by grants 09980501 and 098-0000000-3168 of the Ministry of Science, Education, and Sport of Croatia; and by EU grant ICT-2013-612944 (MAESTRA). A.K. would like to thank Maria Marta Radman for her help with the graphics.

Received: December 12, 2013
Revised: March 26, 2014
Accepted: April 30, 2014
Published: May 29, 2014

**REFERENCES**


