Featured in Physics

Proteome-Wide Assessment of Protein Structural Perturbations under High Pressure

Haley M. Moran,[1] Edgar Manriquez-Sandoval,[2] Piyoosh Sharma,[1] Stephen D. Fried,[1,2,*] and Richard E. Gillilan[3,†]
[1]Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, USA
[2]Department of Biophysics, Johns Hopkins University, Baltimore, Maryland 21218, USA
[3]Center for High Energy X-ray Sciences (CHEXS), Ithaca, New York 14853, USA

(Received 27 February 2024; accepted 28 June 2024; published 9 September 2024)

One of the planet's more understudied ecosystems is the deep biosphere, where organisms can experience high hydrostatic pressures (30–110 MPa); yet, by current estimates, these subsurface and deep ocean zones host the majority of the Earth's microbial and animal life. The extent to which terrestrially relevant pressures up to 100 MPa deform most globular proteins—and which kinds—has not been established. Here, we report the invention of an experimental apparatus that enables structural proteomic methods to be carried out at high pressures for the first time. The method, called high-pressure limited proteolysis (Hi-P LiP), involves performing pulse proteolysis on whole cell extracts brought to high pressure. The resulting sites of proteolytic susceptibility induced by pressure are subsequently read out by sequencing the peptide fragments with tandem liquid chromatography–mass spectrometry. The method sensitively detects pressure-induced structural changes with residue resolution and on whole proteomes, providing a deep and broad view of the effect of pressure on protein structure. When applied to a piezosensitive thermophilic bacterium, *Thermus thermophilus*, we find that approximately 40% of its soluble proteome is structurally perturbed at 100 MPa. Proteins with lower charge density are more resistant to pressure-induced deformation, as expected; however, contrary to expectations, proteins with lower packing density (i.e., more voids) are also more resistant to deformation. Furthermore, high pressure has previously been shown to preferentially alter conformations around active sites. Here, we show this is also observed in Hi-P LiP, suggesting that the method could provide a generic and unbiased modality to detect binding sites on a proteome scale. Hence, data sets of this kind could prove useful for training emerging artificial intelligence models to predict cryptic binding sites with greater accuracy.

DOI: 10.1103/PRXLife.2.033011

I. INTRODUCTION

Extremophiles are organisms uniquely adapted to environments deeply inhospitable to mesophiles such as humans [1]. Earth is replete with numerous "extreme" ecosystems, including those associated with very high temperature, acidity, or salt concentration (such as in hot springs, acid lakes, or hypersaline lakes) [2]. As most of the Earth's surface is covered by water (71%), much of the planet's biomass is subjected to significant hydrostatic pressure. With an average depth of 3800 m, pressures experienced by organisms residing on the ocean floor are typically close to 38 MPa. Though the deepest parts of the ocean extend below 10 000 m (with pressures above 100 MPa), life exists in ocean sediment and subsurface realms kilometers deeper (Fig. 1) [3]. Accordingly, biophysicists, microbiologists, and geologists alike have realized a newfound appreciation for organisms in deep sea and subsurface ecosystems. The pressure limits of life and the behavior of biomolecules under pressure have also attracted attention in food science due to the advent and widespread use of high-pressure processing (HPP) to sterilize pathogens [4]. At the same time, pressure itself has become an important biophysical tool for probing biomolecule structure and function [5].

Precisely which parts of the cellular machinery are most sensitive to pressure and why remains an open area of research. Polynucleotides, lipids, and proteins all have pressure sensitivity to varying extents [6]; however, existing studies tend to focus on large structural changes at pressure levels well beyond those known to support life. On first examination at the gross structural level, many biomolecules appear relatively pressure resistant. Double-stranded DNA, for example, is known to be intrinsically resistant to high pressure, undergoing little compression up to 200 MPa [7], roughly double the hydrostatic pressure of the deepest point in the ocean [8,9].

Much less is known about the structure of RNA under pressure [10], but it does appear that terrestrially relevant pressures can alter riboswitch function [11]. Phospholipids have been more extensively studied than polynucleotides and are known to be sensitive to terrestrially relevant pressures, with some mesophase transitions occurring well below 100 MPa [5,6,12,13]. Mesophilic membranes can transition to a liquid-ordered (gel) phase in a pressure-dependent manner that some piezophiles resist by possessing an altered lipid profile featuring greater unsaturation and plasmalogens. Furthermore,

*Contact author: sdfried@jhu.edu
†Contact author: reg8@cornell.edu

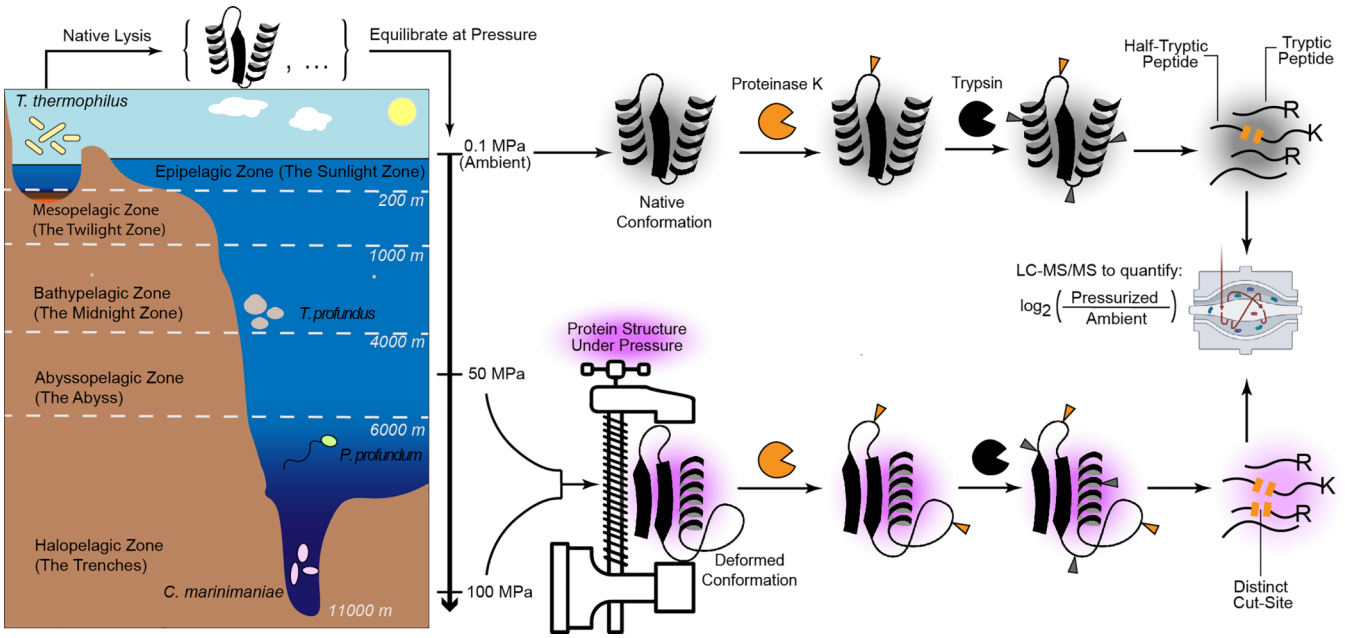

FIG. 1. High-pressure limited proteolysis (Hi-P LiP) deeply interrogates protein structure under pressure. Left: Marine microorganisms inhabit aqueous environments at all depths, from the surface to the pelagic zones to deep trenches. Habitats for *Thermus thermophilus*, *Thermococcus profundus, Photobacterium profundum,* and *Colwellia marinimaniae* are depicted. Right: Proteins are extracted from piezosensitive thermophile *T. thermophilus* and subjected to limited proteolysis (LiP) with Proteinase K (PK), either at ambient pressure or at high pressure (Hi-P). To assess structural deformations induced by Hi-P, proteins are fully trypsinized, and the profile of PK cut-sites is determined by sequencing and quantifying tryptic and half-tryptic peptides by tandem liquid chromatography-mass spectrometry (LC-MS/MS).

it appears that nonlamellar phases are less efficiently packed and are more sensitive to pressure; consequently, organisms alter their lipid chemistry to maintain access to these essential states [14].

A recurring theme in high-pressure structural biology is that functionally important states can be less well packed at the molecular level, therefore exhibiting greater pressure sensitivity—this extends to proteins. The effect of high hydrostatic pressure (Hi-P) on proteins has been investigated using an array of biophysical methods including NMR, fluorescence, and x-ray scattering, as well as biochemical methods such as enzymatic activity assays [5,15–18]. Most model proteins show little large-scale change at terrestrially relevant pressures but begin to unfold at pressures above 300 MPa [5,19,20]. The effect has been described with a two-state thermodynamic model accounting for the *change* in hydrostatic volume upon unfolding (typically small and negative: $\Delta V_{N\to U} \sim -10$ to -100 Å^3 representing $\Delta V/V \sim -0.1\%$ to -1%) [21,22]. Hi-P favors unfolded states because they occupy *less* volume, which is because they eliminate *voids*, small unoccupied volumes trapped within the cores of folded proteins. While contraction of water volume can occur as a result of high charge density (electrostriction), the prevailing model is that proteins with more voids would be more pressure sensitive because they yield a greater thermodynamic weight ($P\Delta V$), favoring the unfolded form [22].

Most recent literature on protein structure under pressure has focused on the range 200–400 MPa, likely because those pressures are at the upper end of what can be reached by biophysical instrumentation, and the effects in that range tend to be more prominent and easily observed. This lack of evidence for significant structural changes to proteins under terrestrially relevant pressures below 100 MPa (and the ease with which lipids transform under pressure) has given the impression that the proteome may not be a bottleneck for adaptation to high pressure.

In the following, we present the first proteome-wide assessment of the effect of high pressure on protein structure. The study couples a purpose-built apparatus that enables biochemical manipulations at high pressure (Fig. 2) with limited-proteolysis mass spectrometry (LiP-MS; Fig. 1) [23,24]. LiP-MS is an emerging structural proteomics approach in which complex samples (such as total cell extracts) are subjected to pulse proteolysis with Proteinase K (PK). This enzyme cleaves selectively at surface-exposed and unstructured regions of proteins but with little sequence specificity. Hence, structural information about the conformational ensemble of each protein is encoded into cleavage sites, which can be read out by sequencing the resulting fragments with tandem liquid chromatography–mass spectrometry (LC-MS/MS) (Fig. 1). In our experiment—high-pressure limited proteolysis (Hi-P LiP)—we perform pulse proteolysis on cell extracts at high pressures, quench PK, and then retrieve the resulting peptides for subsequent mass spectrometry sample preparation at ambient pressure. The experiment provides deep coverage of most soluble proteins in a target organism's proteome and a sensitive readout of local changes in protein structure.

To introduce this new "omics" generation of high-pressure protein biophysics, we performed Hi-P LiP on the proteome of a piezosensitive thermophilic bacterium, *Thermus thermophilus*. We presumed its proteome would be perturbed

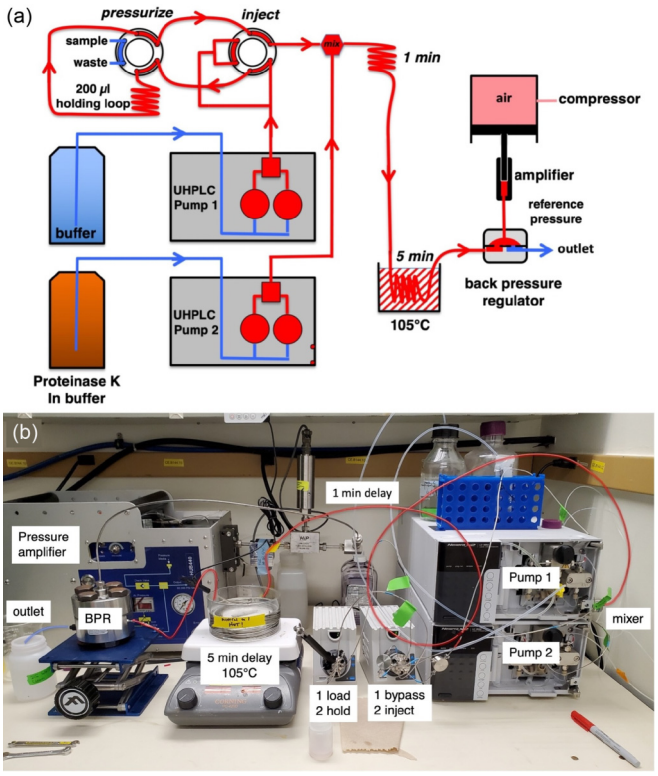

FIG. 2. Apparatus to perform Hi-P LiP. (a) Schematic of apparatus that can perform precisely timed pulse proteolysis of cellular extracts under high pressure followed by thermal denaturation and recovery. Blue lines represent fluid supplied at ambient pressure; red lines are fluid under pressure. (b) Photograph of apparatus *in operando*. BPR, back-pressure regulator. The two six-port valves shown, each with setting 1 or 2, define the sample injection complex (see Figs. S1(a)–S1(c) in the SM [23]).

by Hi-P, given that it is not adapted to deep subsurface environments. The structural changes we observe in the *T. thermophilus* proteome under pressure are surprisingly widespread: 39% of the 1125 assessed proteins. We mapped out the locations of the PK cut-sites on several structurally characterized *T. thermophilus* proteins and found that these locations map to functionally important cavities. These findings serve as a positive control that Hi-P LiP can recapitulate known effects that other biophysical methods have observed associated with Hi-P and can do so with much greater throughput. Consequently, our method appears well suited for unbiased detection of binding sites in complex samples.

We discover that isoelectric point, cofactors, and packing density are significant determinants for pressure sensitivity in the 50–100 MPa range. Moreover, we find that, in *T. thermophilus,* the most pressure-sensitive proteins possess *higher* packing density, suggesting that the very feature (voids) that promotes unfolding at high pressures (>300 MPa) provides protection from structural deformation at terrestrially relevant pressures. Based on these studies, we propose an explanation for why specific proteins in this thermophile are more sensitive to high pressures than others.

II. APPROACH

LiP-MS experiments typically begin by preparing replicates of two closely related protein extracts that differ by a specific condition—pressure in this case. The extracts are then treated with PK for a brief period (1 min) with a low loading of the enzyme (1:100 ratio) so that constituent proteins preferentially receive a single PK cut. This condition ensures that only the most solvent-accessible or unstructured portions of the constituent proteins within the extract are cleaved by the nonspecific protease (cf. Fig. 1) [25]. PK is rapidly quenched through thermal denaturation at $T \geqslant 100\,^{\circ}\mathrm{C}$, and then the samples are subjected to complete trypsinolysis under semi-denaturing conditions overnight, resulting in universal cleavage at lysines (K) and arginines (R). These peptides are then separated by nano–liquid chromatography and analyzed by mass spectrometry to be sequenced and quantified. The complex mixture of peptides in these samples can be sorted into those that are *tryptic* (meaning both ends of the peptide arose from trypsin cleavage) and *half tryptic* (meaning that one end of the peptide arose from PK cleavage), which can be inferred by mapping the sequenced peptides back to their parent proteins. Half-tryptic peptides represent solvent-accessible sites (residues) within proteins; their abundance measures those accessibilities. If a treatment causes local unfolding or distortion at a particular location in a particular protein, then one would expect a higher abundance of the half-tryptic peptide (with a PK cut at that location) in the test condition (i.e., high pressure) relative to a control condition (i.e., ambient pressure). To date, LiP-MS has been applied to various biological problems and has uncovered conformational changes on the proteome scale in response to nutrients [26] and aging [27]. It has also been used to interrogate biophysical problems (in which treatments are applied to cell extracts *ex vivo* as opposed to varying a biological condition), such as protein thermostability [28], protein folding [24,29], and osmolyte-based stabilization [30].

The technical challenge associated with performing LiP-MS to study the structural consequences of high pressure is that such an experiment requires quick mixing (to initiate limited proteolysis) and temperature changes (to quench PK) at high pressure on relatively large sample volumes (0.1–1 mL, ruling out diamond anvils), as well as means to retrieve sample to ambient pressure without incurring a large dilution. Facilitated by recent advances in ultrahigh-performance liquid chromatography (UHPLC) for high-pressure x-ray experiments [Fig. 2(a)] [31], we designed and built an experimental apparatus capable of meeting these technical specifications [Fig. 2(b)]. In brief, clarified extracts (0.2 mL, 0.116 mg/mL) from bacteria lysed by cryogenic pulverization are loaded into an injection valve complex designed to equilibrate samples at the desired pressure prior to processing (see Methods for details). An injection complex allows the sample to be loaded at atmospheric pressure into the 200 µL holding loop, pressurized, and held at 100 MPa without interrupting the overall buffer flow through the system (Figs. S1(a)–S1(c) in the Supplemental Material (SM) [23]). A dilute solution of PK (0.0116 mg/mL) in an identical buffer is brought up to pressure and then mixed in a 1:10 ratio with the flowing sample. The 1 min enzyme incubation and

5 min quench steps (at 105 °C) are accomplished by running the mixed sample through delay lines of appropriate length; to digest, the delay line is kept at room temperature, and to quench, the delay line is submerged in a mineral oil heat bath (Fig. 2). An active back-pressure regulator (BPR) enables the quenched sample to return to ambient pressure at an approximately 1/3 dilution from the original pressurized bolus (Fig. S1(d) in the SM [23]). The heat-denatured depressurized liquid is eluted into a container prefilled with solid urea until the full width at half maximum (FWHM) portion of the sample peak (0.674 mL; 2 min elapsed time) has passed (73% of the total sample as measured) to ensure PK activity remains quenched. The digested extracts are stored in the 8 M urea at 4 °C until all samples are ready to proceed to processing for mass spectrometry sample preparation (see Methods).

For this study, samples consisted of clarified extracts of *T. thermophilus* (substrain HB27) that were grown at 60 °C. Extracts were pressurized to 50 MPa (500 bars) or 100 MPa (1 kbar) in technical triplicates. Control samples were prepared by running the identical extracts through the apparatus but without activating the pressure cell (thereby accounting for any dilution factors associated with running through the setup). Spectra were searched, and label-free quantification was conducted in FragPipe [32] using IonQuant [33]. Our newly developed FLiPPR pipeline [34] (FragPipe LiP Processor) was employed to analyze the raw ion intensities, calculate effect sizes (reported as $\log_2$ of the ratio of abundances of a given peptide in the pressurized samples relative to control samples) and FDR-corrected p-values, and to collate metadata to facilitate bioinformatic analysis and discovery of trends.

III. RESULTS AND DISCUSSION

A. High pressure induces structural deformation across the *T. thermophilus* proteome

In experiments quantifying changes in peptide abundances at 50 MPa (100 MPa) relative to ambient pressure, we confidently identified and quantified 18 791 (20 964) peptides. These peptides belonged to 1197 (1226) distinct proteins, representing an acceptable coverage of the soluble *T. thermophilus* HB27 genome (∼66%), which comprises 2200 protein-coding genes, of which 390 are predicted to be membrane proteins. Peptide-level volcano plots [Figs. 3(a) and 3(b)] show how the abundances of these peptides change in response to pressure along with p-values (adjusted for multiple hypothesis testing [34]) against the null hypothesis that pressure treatment does not alter a protein's structure. Pressure treatment perturbs many proteins' conformations.

At 50 MPa, 5.5% of the sites within proteins were perturbed (using cutoffs of $|\log_2(\text{pressurized}/\text{ambient})| > 1$, adjusted $p < 0.05$); this fraction increases dramatically to 20.9% of sites at 100 MPa [Figs. 3(c) and 3(d)]. These volcano plots exhibit an unusual asymmetry in that many peptides *only* appear in the pressurized samples (the rightward lobes; 737 at 50 MPa and 2178 at 100 MPa), and these peptides are disproportionately half tryptic (shown as blue dots, 71% at 50 MPa and 74% at 100 MPa). Hence, the data unambiguously show that high pressure, particularly 100 MPa, renders many PK-inaccessible locations within proteins susceptible to PK cleavage. We define a protein as structurally perturbed if we can detect two or more sites with a significant change in PK susceptibility. By this metric, 12% of the *T. thermophilus* proteins are structurally perturbed at 50 MPa, and 39% are structurally perturbed at 100 MPa [Figs. 3(e) and 3(f)]. It is worth pointing out that this observation differs from other high-pressure biophysical studies on proteins, in which unfolding transitions typically are found for $\geqslant$300 MPa [5,19,20]. The two findings can be reconciled with a model that, as pressure increases, local structural deformations precede global unfolding and that the site-specific limited-proteolysis approach employed here is a very sensitive reporter to local deformation.

Though 50 MPa does not alter an immense fraction of *T. thermophilus*'s proteins, the 131 affected proteins include many essential ones, including 15 ribosomal proteins [35], the division protein FtsZ [36], and the essential enzyme peptide deformylase [37], consistent with these pressures still being too high to support viability [Fig. 3(g)]. Curiously, only one of the 20 aminoacyl-tRNA synthetases (threonyl-tRNA synthetase) is perturbed at 50 MPa despite very high coverage of this group of proteins (only 7 altered sites across 617 quantified).

B. High-pressure limited proteolysis is a robust experiment

Given that enzymatic cleavage at high pressure has not been extensively used as a structural reporter, we wanted to examine this approach's reliability and reproducibility critically. The median coefficient of variation (CV) for peptide abundance across three replicate digests at elevated pressures was 17% (at both 50 and 100 MPa), implying that limited proteolysis can be conducted reproducibly at high pressure (Figs. S2(e) and S2(f) in the SM [23]). A potential confounding variable in this approach is that high pressures have been shown to enhance enzymatic activity [15], including proteases. Though we cannot rule out that pressure-induced PK activation contributes to the signal in these experiments, several features within our data sets argue against it being the primary explanatory variable of our data. First, we find that the fraction of sequenced peptides that are half tryptic is virtually the same between pressurized and ambient samples (46% ± 0.6% at ambient, 44% ± 1.9% at 100 MPa; see Figs. S2(c) and S2(d) in the SM [23]). Second, we find a large dispersion in the effect of pressure on half-tryptic peptide abundance [Figs. 3(a) and 3(b)], with the majority (approximately 80% at 100 MPa) being invariant to pressure. These observations are inconsistent with pressure-induced enzyme activation, for which a uniform increase in the production of half-tryptic fragments would be expected.

We conducted two separate fluorescence-based proteolysis assays under pressure to more directly assess whether pressure-induced activation occurs during Hi-P LiP experiments (Fig. 4). As a positive control, we found that pressure does indeed activate PK to degrade casein micelles [as measured by loss of fluorescein isothiocyanate (FITC) self-quenching; Figs. 4(a)–4(c)]; this is a well-known example of a substrate-level effect as pressure is known to disrupt casein micelles, which would increase FITC-casein's proteolytic susceptibility [38]. On the other hand, pressures up to 100 MPa

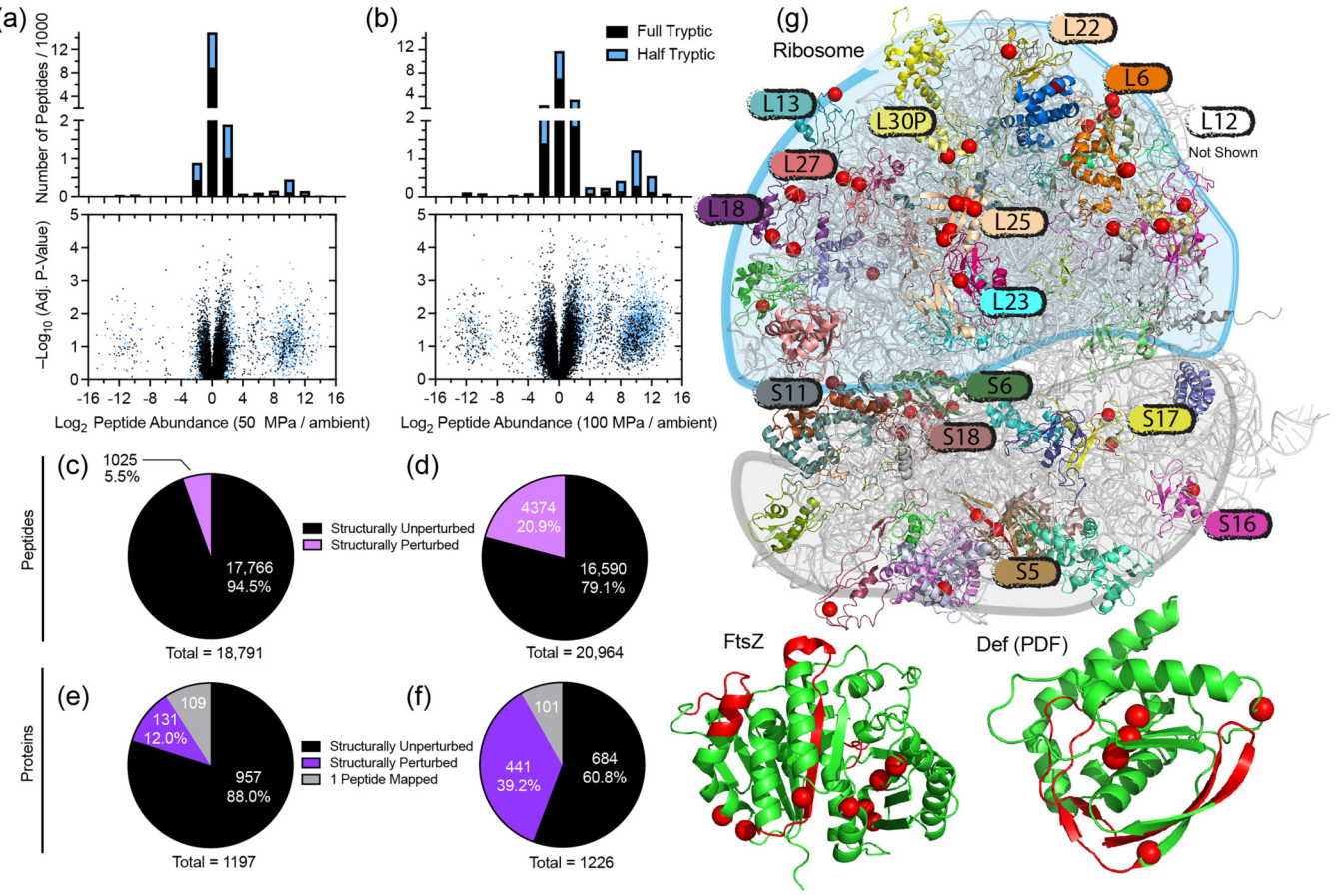

FIG. 3. A proteome-wide view of protein structural deformation under pressure. [(a), (b)] Peptide-based volcano plots depicting the relative abundance of confidently quantified peptides in pressurized samples relative to those retained at ambient pressure for pressures of (a) 50 MPa and (b) 100 MPa. Blue dots represent half-tryptic peptides; black dots, tryptic peptides. Histograms show the number of peptides in each abundance ratio category. Adjusted p-values were calculated by applying Benjami-Hochberg correction per protein following Welch's t-test. [(c), (d)] Pie charts indicating the total number of confidently quantified peptides at (c) 50 MPa and (d) 100 MPa, and the percent that are present at significantly different levels under pressure (effect-size >2, adjusted $p < 0.05$) representing structurally perturbed regions. [(e), (f)] Pie charts indicating the total number of confidently identified proteins and the percent structurally perturbed under pressure at (e) 50 MPa and (f) 100 MPa. Proteins with only one quantified peptide are discounted; proteins with two or more peptides with significant changes are labeled structurally perturbed. (g) The ribosome and two other essential proteins (cell-division protein, FtsZ, and peptide deformylase, Def) are structurally altered at 50 MPa. Each half-tryptic peptide with a significant change in abundance is demarcated with a red sphere at the Proteinase K cleavage site; significant tryptic peptides are colored red.

had only a modest enhancing effect (1.5 ± 0.4 fold) on PK activity on a disordered peptide reporter [Figs. 4(d)–4(f)]. Altogether, these data argue that the intrinsic PK activity is not strongly modulated by pressure (though this is not true for some other proteases [15]). Instead, apparent increases in cleavage in a pressure-dependent manner are substrate-level effects in which pressure locally deforms substrate structure, thereby altering its susceptibility to PK.

C. Proteins with particular characteristics are susceptible to pressure-induced deformation

We next collected a wide range of metadata about each of the *T. thermophilus* proteins to assess whether any biophysical, biochemical, or topological criteria could explain differences in pressure sensitivity or resistance. Several of these criteria can be calculated directly from protein sequences in BioPython, such as molecular weight and isoelectric point (pI). We also calculated the disorder content of each protein (using Metapredict [39]) and the domain structure (using DomainMapper [40]). Manual annotation is not generally available for all of *T. thermophilus*'s proteins. Hence, in order to infer proteins' subcellular locations, subunit composition, and cofactors, we first assessed whether a likely unique ortholog existed within the *Escherichia coli* proteome (based on a reciprocal best hit criterion [41]). We transferred this information to the corresponding *T. thermophilus* protein (using the comprehensive curated data in EcoCyc [42]; see Sec. V G for bioinformatic methods). Several key trends emerged (Fig. 5).

1. Isoelectric point

We found a robust relationship between pressure sensitivity and predicted isoelectric point in *T. thermophilus* proteins. Specifically, sites within very acidic (protein pI < 5) and

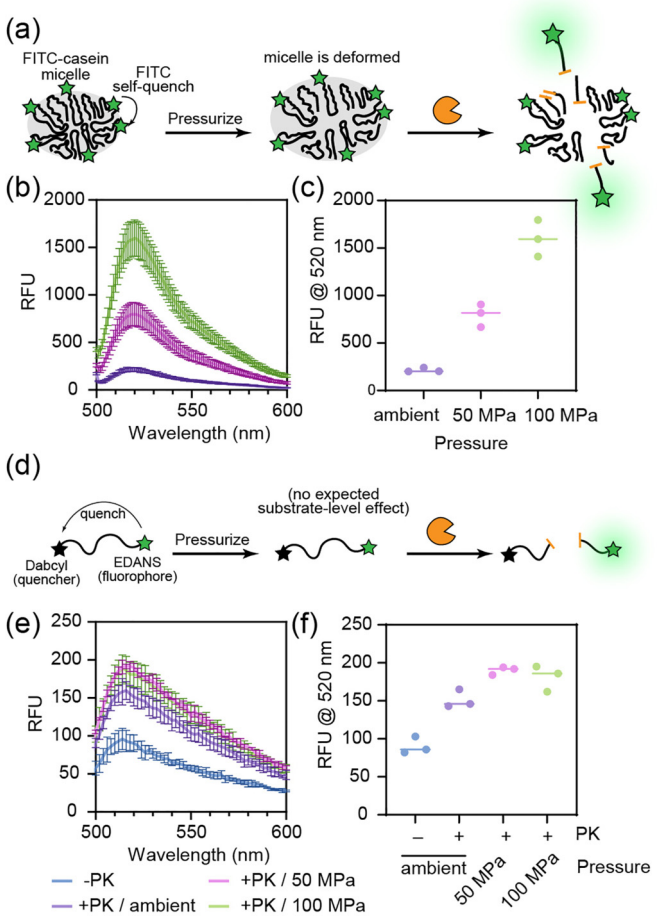

FIG. 4. Fluorescence assays to measure Proteinase K (PK) activity under pressure. (a) Scheme for casein cleavage assay. FITC-labeled casein forms micron-sized micelles with low fluorescence due to self-quenching. Pressure disturbs the micelle, increasing casein susceptibility to PK. Cleavage of FITC-casein liberates FITC into the bulk solvent, activating fluorescence. (b) Fluorescence spectra collected on 7.2 μg/mL FITC-casein (excited at 485 nm), proteolyzed with PK (1:2.5 w/w) for 1 min under ambient conditions or at pressure (symbols represent means, error bars represent standard deviations, $n = 3$). (c) Fluorescence levels at 520 nm following pulse proteolysis with PK at three conditions interrogated. (d) Scheme for peptide cleavage assay. A fluorophore (EDANS)-quencher (Dabcyl) paper on a short peptide produces little fluorescence unless the peptide is cleaved by PK, liberating the fluorophore. Pressure is expected to exert minimal structural effect on the disordered peptide. (e) Fluorescence spectra collected on 40 μg/mL Dabcyl-peptide-EDANS (excited at 355 nm), proteolyzed with PK (1:100 w/w) for 1 min under ambient conditions or at pressure (symbols represent means, error bars represent standard deviations, $n = 3$). Fluorescence without proteolysis (–PK) was recorded as well. (f) Fluorescence levels at 520 nm following pulse proteolysis with PK at three conditions interrogated (or without PK).

basic (protein pI > 10) proteins had a high propensity to be structurally altered at 100 MPa (34% and 38%, respectively). In contrast, sites within proteins that are near neutral at physiological pH (protein pI 6–9) were less likely to be structurally altered at 100 MPa [15%; $p < 10^{-92}$ by chi-square test; Fig. 5(e)]. This trend is recapitulated at 50 MPa as well [Fig. 5(b)]. For context, the distribution of protein isoelectric points for the *T. thermophilus* proteome is provided in Fig. 5(g). This finding is consistent with physical intuition: high pressures will provide a driving force to reduce molecular hydrostatic volume. Polyanions and polycations, when so compressed, would experience greater electrostatic repulsive forces. Alternative hypotheses may explain these results in terms of pressure-induced changes to the hydration layer [48]. We point out that the trend between pI and pressure sensitivity is nearly the *opposite* of our previously reported trend between pI and *refoldability*, in which proteins at the extremities of the pI spectrum typically refold more efficiently [24,29].

2. Cofactors

Another connection was found between the cofactors a protein hosts and pressure sensitivity (Figs. 5(a) and 5(d) and Fig. S3 in the SM [23]). Overall, holoproteins that contain either organic cofactors (TPP, PLP, FAD, FMN) or covalently coordinated metals (iron-sulfur clusters and heme) are resistant to deformation (with only 1–2% of sites altered at 50 MPa and 10% at 100 MPa). Proteins that coordinate divalent metals without covalent bonds (Fe^{2+}, Mg^{2+}, Mn^{2+}, Zn^{2+}) are much more susceptible to deformation (5.5% at 50 MPa and 18% at 100 MPa), and proteins predicted to carry no cofactors are the most susceptible (5.8% at 50 MPa and 23% at 100 MPa). Figure S3 in the SM [23] presents a more granular view of these data, separating data for each individual cofactor.

These results are consistent with studies of individual protein-ligand complexes at high (>300 MPa) pressures: holoproteins tend to be more resistant to unfolding than apoproteins, presumably because the ligand fills a cavity and lends stability to the structure [49]. Divalent cations, on the other hand, have a high electrostriction effect on bulk water; that is, their electric fields reduce the molar volume of water by $>18\,\mathrm{cm}^3/\mathrm{mole}$ [50]. A divalent cation liberated from its binding pocket in a protein would, therefore, be more likely to result in a net volume decrease favored by pressure.

3. Packing density

Perhaps the most striking trend we observed concerned pressure sensitivity and packing density, a bioinformatic metric quantifying how closely spaced the atoms in a protein are under standard conditions, defined as $1–V_{\mathrm{void}}/V_{\mathrm{total}}$ [Figs. 5(c) and 5(f)]. Previous work has established strategies to calculate the hydrostatic volume of proteins and the volumes of "voids" within the folded region from structure [44,51,52]. We applied one such method to all *T. thermophilus* HB27 proteins (in their monomeric form) using the refined AlphaFold2 structure predictions [46] on the European Bioinformatics Institute (EBI) database. As AlphaFold2 models do not possess hydrogen atoms, we used the pdb2gmx tool from GROMACS [47] to hydrogenate all the proteins to better estimate their true packing densities. Particularly at 100 MPa, we found a striking correlation in which sites within *T. thermophilus* proteins whose packing densities are less than 0.74 are protected from pressure-induced deformation (16% altered), with a dramatic rise for sites within proteins whose packing densities are greater than 0.74 (28% altered). A statistical test

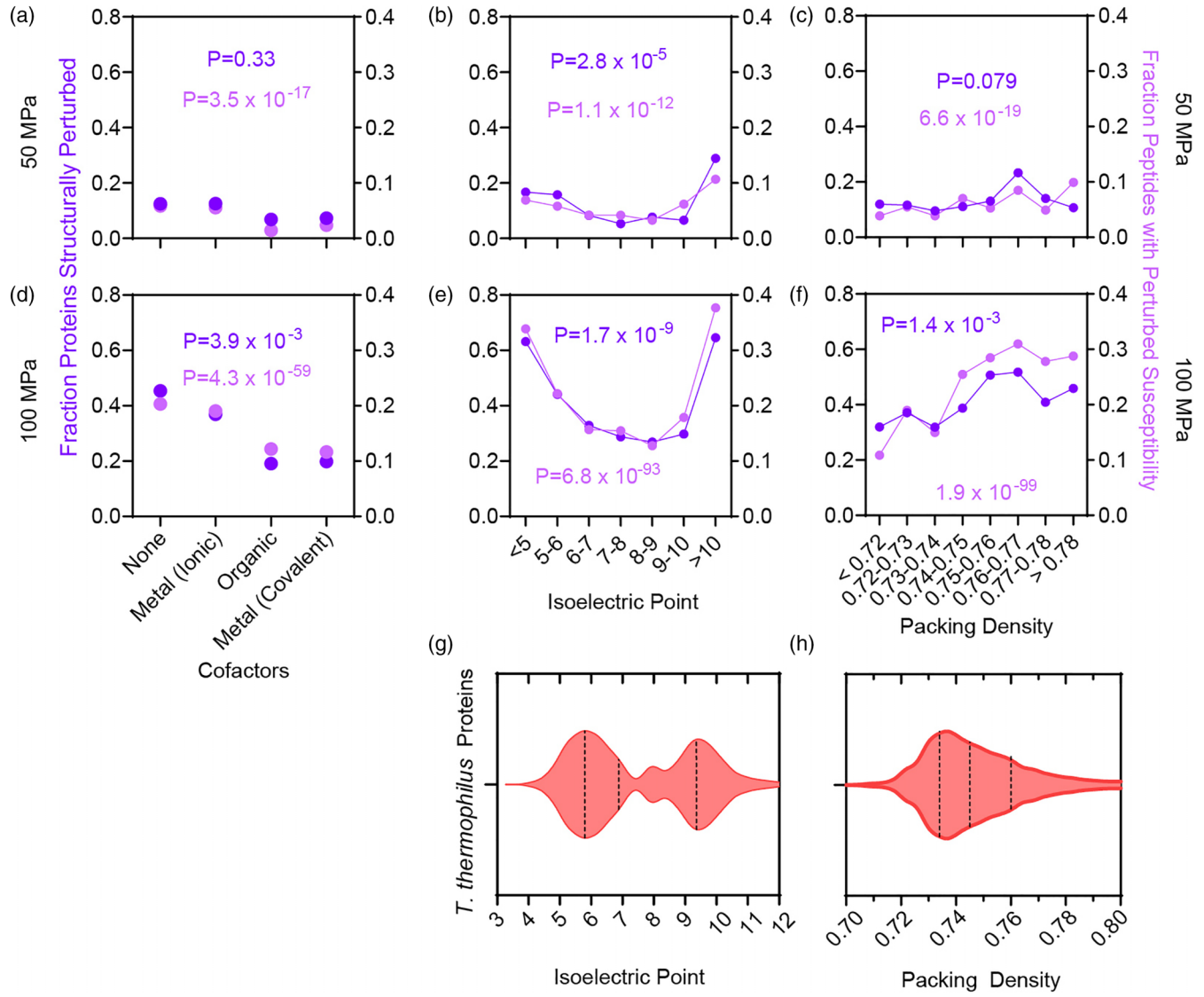

FIG. 5. Biophysical features explain susceptibility to structural deformation under pressure. [(a)–(c)] Fraction of proteins that are structurally perturbed (dark violet symbols, left *y* axis) and fraction of peptides with perturbed susceptibility to proteinase K (light pink symbols, right *y* axis) at 50 MPa pressure, grouped by the protein's (a) cofactor binding (inferred from reciprocal best-hit orthology), (b) isoelectric point (pI), and (c) packing density (calculated by ProteinVolume v1.3 [43,44] using AlphaFold2-v4 structural models [45,46] hydrogenated with the pdb2gmx tool [47]). Each trend line is associated with a *p*-value according to the chi-square test against the null hypothesis that the categorical variable does *not* explain differences in susceptibility to pressure. [(d)–(f)] Analogous to panels (a)–(c), except at 100 MPa pressure. (g) pI distribution of the *T. thermophilus* proteome. (h) Packing density distribution of the *T. thermophilus* proteome.

against the null hypothesis that packing density does not describe these differences in pressure sensitivity yields a *p*-value of 10^{-99} [Fig. 5(f)]. For context, the distribution of protein packing density for the *T. thermophilus* proteome is provided in Fig. 5(h). Packing density is less relevant for explaining effects at 50 MPa, though a slight uptick in pressure sensitivity was noted for the absolute highest densities above 0.78 [Fig. 4(e)]. This finding is intriguing because it runs counter to the theory that proteins with greater void volume (lower packing density) would be more prone to unfold because of a greater thermodynamic contribution of $P\Delta V_{\mathrm{N}\to\mathrm{U}}$ to $\Delta G_{\mathrm{N}\to\mathrm{U}}$. For a typical value for ΔV ($\sim$30 $\mathrm{cm^3\,mol^{-1}}$ or $30\times10^{-6}\,\mathrm{m^3\,mol^{-1}}$) [53], a pressure of 100 MPa ($100\times10^{6}\,\mathrm{N\,m^{-2}}$) would contribute 0.72 $\mathrm{kcal\,mol^{-1}}$ ($P\Delta V_{\mathrm{N}\to\mathrm{U}} = 3\times10^{3}\,\mathrm{J\,mol^{-1}}\times2.39\times10^{-4}\,\mathrm{kcal/J}$) of destabilization to the native state (relative to the unfolded state), too low to unfold all but the most unstable proteins (typical $\Delta G_{\mathrm{N}\to\mathrm{U}}$ are 4–8 $\mathrm{kcal\,mol^{-1}}$ [54]). Hence, terrestrially relevant pressures are not high enough to unfold most proteins but are likely high enough to deform or remodel them locally. According to these data, tiny voids may protect *T. thermophilus* proteins from pressure-based deformation, possibly by allowing the chain to compress without incurring major steric clashes.

To ensure that these findings were not biased by systematic error in the structural predictions of AlphaFold2 or an artifact of crystal packing effects in Protein Data Bank (PDB) structures, we performed two computational controls (Fig. S4 in the SM [23]). First, for the 109 *T. thermophilus* HB27 proteins

for which x-ray structures were available in the PDB, we calculated their packing densities and compared them to their corresponding values based on AlphaFold2 structural models. Excepting the large ribosomal subunit proteins (based on a low 3.8-Å-resolution structure), the correlation was very good ($R^2 = 0.77$; $m = 0.98 \pm 0.06$; Fig. S4(a) in the SM [23]). We also ran short 50–100 ps equilibrations with molecular dynamics simulations to solvate and relax 14 *T. thermophilus* protein structures at standard temperature and pressure; this resulted in nearly negligible changes to packing density ($R^2 = 0.99$; $m = 1.01 \pm 0.01$; Fig. S4(b) in the SM [23]). Hence, we conclude that AlphaFold2 structural models are decent proxies for estimating packing densities.

We are mindful that computed packing densities are sensitive to the atomic radii parameters, and the analyses shown in Fig. 5 utilize the van der Waals radii from Bondi, as implemented in ProteinVolume v1.3. To assess the robustness of these conclusions, we calculated all protein packing densities using the alternative atomic radii parameters proposed by Gaines *et al.* [55]. Using Bondi's radii parameters, an average protein packing density is predicted to be approximately 0.74, a value similar to that of hexagonal close packing of hard spheres [56]. In contrast, Gaines *et al.*'s parameters select atom sizes designed to minimize side-chain overlap in crystal structures. These parameters result in a mean packing density closer to 0.56 [57]. Although these parameters result in a marked global decrease in packing densities (Fig. S5(a) in the SM [23]), the trends between packing density and pressure sensitivity remain unaltered (Figs. 5(c) and 5(f), and Figs. S5(b) and S5(c) in the SM [23]).

4. Subunit composition

Multiple previous studies [20,21,58,59] have linked pressure sensitivity to protein oligomeric state, and a common view is that pressure-induced subunit dissociation precedes pressure-induced unfolding [60]. While easy-to-dissociate homo-oligomeric complexes have been convenient targets for Hi-P studies, it is unclear to what extent biologically relevant pressures disrupt oligomers within the proteome. If this trend were general, monomeric proteins would exhibit fewer pressure-sensitive changes in their PK susceptibility compared to oligomers since the dissociation of a complex would expose new protein surfaces to PK. However, we could not find evidence supporting this expectation in *T. thermophilus* (Fig. S6 in the SM [23]); monomeric proteins were not more resistant to pressure-induced changes, and there was no discernible trend as the predicted number of subunits in the complex increased.

5. Disorder

Intrinsically disordered regions have been posited to play several roles in biology, and one of relevance is their capacity to respond rapidly to environmental changes [61]. One study has pointed out that their ensemble properties can acutely shift in response to changes in osmolytes [62]. Thus, their potential reaction to hydrostatic pressure is of interest. Here, we must temper our findings because *T. thermophilus* possesses very little disorder in its proteome; the median protein has a predicted disorder content of 2.6% (Fig. S7(c) in the SM [23]), and in our 100 MPa study, we only recorded data for 57 proteins with $\geqslant$15% disorder. Nevertheless, our findings tentatively support the view that disordered proteins are more sensitive to pressure (Fig. S6(a) in the SM [23]). Specifically, we find a sharp uptick in the degree of structural perturbation for proteins with $\geqslant$15% disorder in both the 50 and 100 MPa data sets.

On the other hand, when we analyzed the individual peptides and mapped them to either disordered regions or folded regions, we did *not* find that locations within disordered regions were more likely to be altered by pressure (Fig. S7(b) in the SM [23]). In other words, *T. thermophilus* proteins with disorder are more prone to be altered by pressure, but it is not necessarily the disordered regions that are altered. The latter finding can be partially rationalized on the basis that disordered regions—which are typically fully solvated—do not generally have voids and are, hence, volumetrically well packed. Nevertheless, disordered regions appear to modulate how proteins sense pressure, the basis of which will require further investigation.

D. Sites perturbed by pressure cluster around functional regions

We curated our data set to identify a smaller number of well-characterized proteins—ideally which possess high-resolution x-ray structures in the PDB—to identify whether sites that are structurally altered by pressure cluster into functional hubs or other notable locations within protein architectures. Toward this end, we filtered our LiP-MS data to the highest-confidence subset by first performing a newly described data-merging scheme that pools together distinct peptides that report on a common PK cut-site [34] and then focusing only on those cut-sites whose Benjami-Hochberg adjusted p-values are below a stringent threshold of 0.02. Manual inspection of this set of cut-sites revealed that locations within *T. thermophilus* proteins that undergo structural deformation under pressure cluster near active site pockets.

GroEL is a bacterial chaperonin responsible for facilitating the folding of hundreds of protein clients in *E. coli*, and x-ray structures have been solved for both the *E. coli* and *T. thermophilus* (PDB: 1AON, 4V4O [63,64]) homologs. The chaperone functions as a tetradecamer, consisting of two heptameric rings that interface at their equatorial domains (Fig. S8(a) in the SM [23]). Central to its functional cycle is the binding of ATP to its equatorial domain, which stimulates a large conformational change [65,66]. At 50 MPa, only 2 PK cleavage sites meet the adjusted p-value criterion, but they occur directly adjacent to the nucleotide binding pocket [Fig. 6(a)]. At 100 MPa, pressure-sensitive sites near the nucleotide binding pocket have proliferated significantly in the equatorial and intermediate domains, with additional sites appearing in the apical domain between helices. The nucleotide binding pocket is capped on the exterior by a solvent-exposed loop that contains multiple sites with significant structural perturbation (Fig. 6(a), and Fig. S8(b) in the SM [23]). When visualized on the heptamer, these sites arrange into a belt-like structure around the equatorial domains, suggestive of the allosteric communication responsible for cooperative nucleotide binding within the heptameric ring (Figs. S8(c) and S8(d) in the SM [23]).

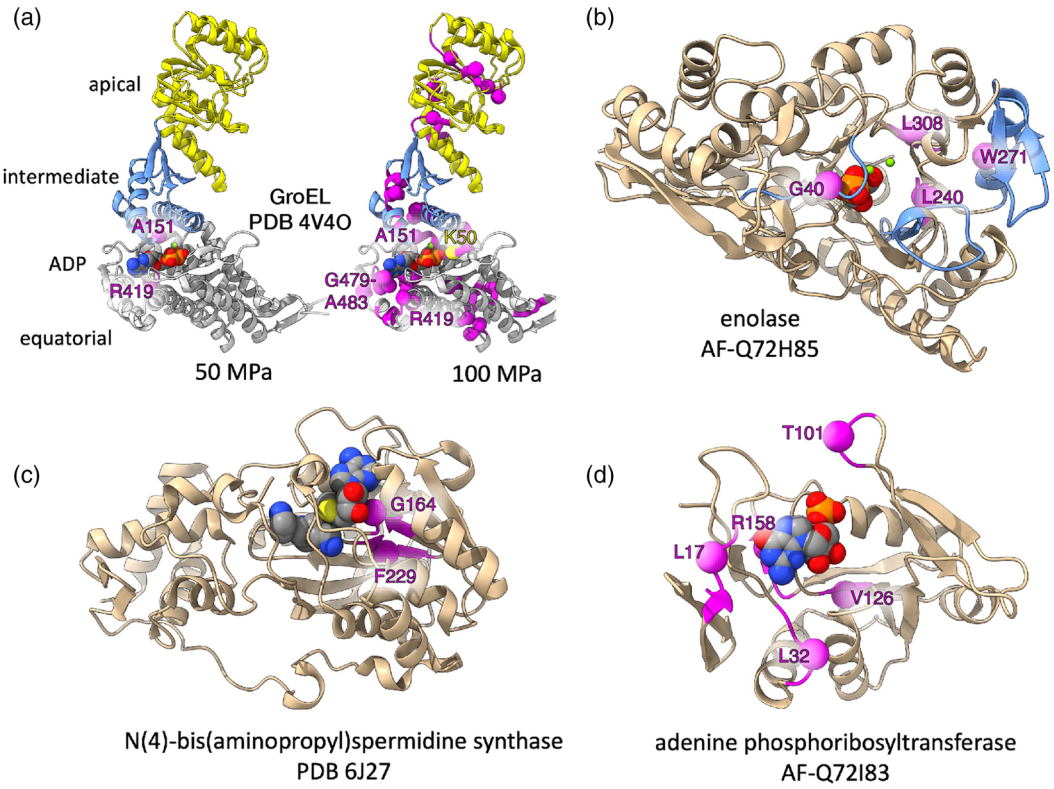

FIG. 6. Pressure-induced structural alteration clusters at active sites. Significant cut-sites (greater than twofold change in PK susceptibility, adjusted $p < 0.02$) are represented as spheres on the residues' α carbons, with shading extending two additional residues in both directions. Purple spheres represent sites of increased susceptibility at elevated pressure; yellow spheres represent sites of decreased susceptibility at elevated pressure. Cut-sites illustrated represent data from the 100 MPa condition, except for GroEL [panel (a)], wherein data from both the 50 MPa and 100 MPa conditions are shown. Panels (a) and (c) use x-ray structures from the PDB (codes given); panels (b) and (d) use AlphaFold2-v4 models, aligned onto orthologous proteins (see text) to place ligands. (a) Apical, intermediate, and equatorial domains of GroEL shown in yellow, blue, and gray, respectively. (b) Blue represents known dynamic loops in enolase.

In a model for the glycolytic enzyme enolase [created by aligning the AlphaFold2 *T. thermophilus* model to the x-ray structure of the *E. coli* ortholog bearing ligands (PDB: 6BFY [67])], the four pressure-sensitive PK sites (L240, W271, L308, G40) cluster around the 2-phosphoglycerate binding site [Fig. 6(b)]. Two of the four sites (G40, W271) are located on loops identified as conformationally variable in the literature [blue in Fig. 6(b)]; significantly, G40 is on the flexible lid to the active site [68]. Likewise, the two structurally perturbed PK sites in N4-aminopropylspermidine synthase (PDB 6J27; G164, F229) [69] occur directly in the joint 5'-methylthioadenosine N4-aminopropylspermidine binding pocket [Fig. 6(c)]. Aligning the AlphaFold2 model of adenine phosphoribosyltransferase (PRTase) [70] to its human ortholog (PDB 6HGS) [71,72] shows that the pressure-sensitive positions envelop the binding site [Fig. 6(d)]. In particular for PRTase, the perturbed sites primarily localize to loops that close around the substrate upon binding.

A final case that we found notable in our manual inspection was (homo)isocitrate dehydrogenase (Fig. 7) (PDB 3AH3) [73], where a large cluster of pressure-sensitive sites are found near the dimer interface lining the active site where homoisocitrate, NADH, and a divalent Ca ion bind. This active site is located in "front" of the enzyme's central β sheet. Intriguingly, we also detect a second cluster of pressure-sensitive sites that line a noncanonical binding site on the backside of the central β sheet. In the x-ray structure, several ethylene diol molecules from the crystallization buffer bind in this cavity, though we are not aware of an x-ray structure of (homo)isocitrate dehydrogenase with biological ligands in this cavity. These observations suggest that this could be an allosteric site for homoisocitrate dehydrogenase for an as-of-yet unknown effector molecule.

In high-pressure NMR studies, Williamson observed that pressure-induced structural changes in four model proteins (BPTI, lysozyme, protein G, and barnase) cluster near known active sites [74]. Our manual inspection of a few examples suggests that Hi-P LiP recapitulates this property but can detect these changes with much higher throughput and much lower experimental cost. As artificial intelligence (AI) methods in protein science continue to develop, global predictions for ligand binding sites will likely be readily available, facilitating a more comprehensive comparison. Alternatively, Hi-P LiP data sets could be used to train AIs to predict binding sites.

IV. CONCLUSION

High-pressure limited proteolysis (Hi-P LiP) is a novel structural proteomics technique that can detect conformational shifts at the residue level, on the proteome scale, for

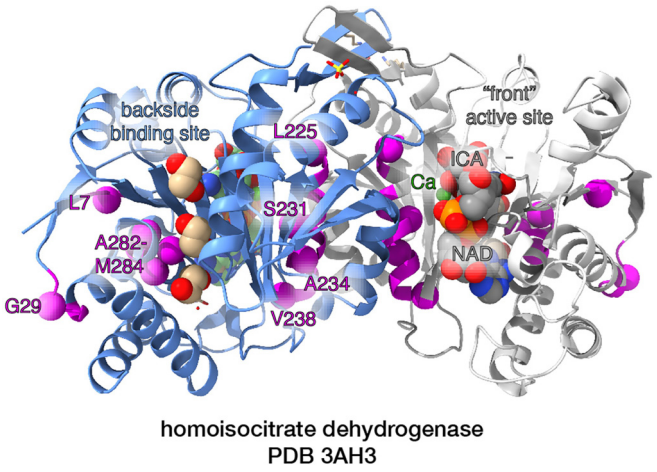

FIG. 7. Pressure-induced structural alterations in (homo)isocitrate dehydrogenase. Significant cut-sites (greater than twofold change in PK susceptibility, adjusted $p < 0.02$) are represented as purple spheres on the residues' α carbons, with shading extending two additional residues in both directions. Cut-sites illustrated represent data from the 100 MPa condition. NAD, isocitrate (ICA), and calcium (Ca) are bound within the active site in front of the central β sheet. Ethylene diols are bound within a second pocket on the backside of the central β sheet. Pressure-sensitive sites line both binding pockets. Model generated by aligning x-ray structure for *T. thermophilus* HB27 homoisocitrate dehydrogenase (PDB 3AH3 [73]) with an ortholog (PDB 6M3S) with NAD and ICA bound.

terrestrially relevant hydrostatic pressures. We have found that this approach is robust and sensitive, reporting on modest structural deformations induced by pressure that were probably undetectable through other experimental approaches. The action of modest terrestrial pressures on the proteome of *T. thermophilus*, a piezosensitive bacterium, is much more widespread than anticipated based on previous high-pressure structural studies conducted on individual proteins. Indeed, we find that 39% of identified proteins were structurally perturbed at 100 MPa, as based on their possessing at least two sites with altered susceptibility to Proteinase K. The unexpectedly large fraction of pressure-sensitive proteins in this organism rejuvenates the hypothesis that proteins in piezosensitive species can be quite sensitive to pressure, even at comparatively modest magnitudes. Moreover, we demonstrate that Proteinase K is fortuitously suitable for these studies in that it is not substantively activated by pressures up to 100 MPa itself.

In *T. thermophilus*, pressure sensitivity correlates strongly with isoelectric point, the presence and chemical nature of cofactors, and the packing density of atoms. The relationship between pressure sensitivity and high atomic packing density runs counter to conventional thinking based on pressure-induced unfolding studies that have highlighted the role of volumetric changes in a two-state model. The structural changes we observe in the 50–100 MPa regime are likely local, are more subtle, and may represent a regime not previously studied in which tiny voids provide a "cushion," enabling proteins to compress without undergoing large-scale conformation changes. Interestingly, this view is consistent with x-ray crystallography studies on 3-isopropylmalate dehydrogenase, in which the authors detected more voids in the piezophilic protein (compared to the mesophilic homolog), which shrink in response to pressure without altering the overall shape of the protein [75,76].

It is important to point out that our research at this stage can only confidently identify the properties of thermophilic proteins that make them more piezosensitive. It would be imprudent to conclude that piezophile proteins are adapted to have lower packing densities or closer-to-neutral isoelectric points, as they may possess unidentified mitigation strategies at both the protein and cellular level. Further Hi-P LiP studies, comparing mesophile to piezophile proteins, both *in vitro* and *in vivo*, will be addressed in ongoing studies.

The examples of pressure-sensitive sites that cluster around binding cavities—particularly those involved in allosteric networks such as in GroEL—are consistent with the previous high-pressure NMR literature and the theory that allosteric communication requires densely packed regions to facilitate mechanical coupling [77]. Hence, these early observations suggest that Hi-P LiP may provide a means to detect functional sites and allosteric networks in an unbiased proteome-wide manner. To summarize, Hi-P LiP is an important new tool which will help us better understand how life adapts to high pressure and move high-pressure biophysics to the "omics" level.

V. METHODS

A. Culture and lysis

Thermus thermophilus HB27 cells (ATCC BAA-163) were grown overnight to saturation and then inoculated into 100 mL of Castenholz TYE Media (ATCC Medium 461) in 250 mL flasks at a starting OD_{600} of 0.05. Cells were grown to a final OD_{600} of ~0.7 at 60 °C with shaking (220 rpm) and harvested via centrifugation at 7200*g* for 5 min at room temperature. Cell pellets were resuspended in 2.0 mL of lysis buffer (20 mM Tris pH 8.2, 100 mM NaCl, 2 mM $MgCl_2$) and flash frozen by slow drip over liquid nitrogen. Frozen cell suspensions were then precooled for 4 min and cryogenically milled via SPEX SamplePrep freezer mill for nine cycles alternating between 1 min grinding and 1 min cooling steps. The pulverized lysate was thawed at room temperature for 30 min, transferred to a fresh 1.5 mL microfuge tube, and clarified at 16000*g* for 15 min at room temperature to remove cell debris. The lysate protein concentration was measured via bicinchoninic acid assay (Rapid Gold BCA Assay, Pierce) using a plate reader (Molecular Devices iD3). Using these clarified extracts, nine 0.2 mL samples of 0.116 mg/mL protein in 20 mM Tris pH 8.2, 100 mM NaCl, 2 mM $MgCl_2$, and 1 mM DTT were prepared for HiP-LiP, allowing for three technical replicates for each pressure condition (ambient, 50 MPa, and 100 MPa). We estimate that at pressures up to 100 MPa, the pH of this buffer would be perturbed by <0.2 pH units [78].

B. High-pressure limited proteolysis

High-pressure buffers were supplied using a pair of Shimadzu Nexera LC-30 AD chromatography pumps (Shimadzu North America, Somerset, NJ) capable of generating constant

flow up to 130 MPa (Fig. 2). All tubing prior to the mixer was stainless steel [outside diameter (o.d.) 1.58 mm, inside diameter (i.d.) 0.25 mm] with a pressure rating of 138 MPa (IDEX Health & Science, LLC, Oak Harbor, WA). The *injection complex* is a pair of MX Series II 2-position/6-port Ultra−life^{TM} switching valves with maximum pressure rating 103 MPa (IDEX Health & Science, LLC, Oak Harbor, WA). One valve (left side in Fig. 2) was fitted with a sample injection port (INV-907 Cytiva Life Sciences, Marlborough, MA) and a 200 μL stainless steel holding loop (o.d. 1.58 mm, i.d. 0.76 mm; IDEX Health & Science, LLC, Oak Harbor, WA). The second valve (right side in Fig. 2) was plumbed with zero dead volume internal stainless steel TEE connectors having internal diameter 0.25 mm (Vici Valco Instruments, Houston, TX). Sample and PK-spiked buffer streams were mixed inline using a 100-μL high-efficiency mixer (MiRC Mixer MR100, Shimadzu Scientific Instruments, Columbia, MD). The 1-min and 5-min delay loops were made from a single contiguous 10-m length of stainless-steel tubing (Avantor Hichrom, o.d. 1.58 mm, i.d. 0.50 mm, Hichrom Limited, Reading, United Kingdom). At a flow rate of 0.337 mL/min, sample traversed the 10-m tubing length in 6 min, spending 1 min at 23 ℃ (1.7 m) and 5 min at 105 ℃ (8.3 m). Temperature in the 8.3-m loop was maintained by a heated oil bath. In operation, pump 1 delivered pure buffer at 0.3 mL/min while pump 2 delivered buffer containing (11.6 μg/mL) proteinase K (PK) at 0.037 mL/min resulting in a 1:10 dilution after mixing.

To simultaneously maintain constant flow and constant pressure in these experiments, it was necessary to utilize a system of active back-pressure regulation. The back-pressure regulator (BPR) here has been described elsewhere [31]. The reference water pressure required by the BPR was supplied by an HUB 440 high-pressure generator (Pressure BioSciences Inc., Medford, MA). Unlike the previous application, sample in this setup passes directly through the BPR and is collected into a tube containing sufficient urea to make up an 8 M solution; then the critical fraction is received. To evaluate the effect of sample injection, mixing, and pressure regulation on the sample elution profile at the outlet, we conducted an acetone pulse experiment. An injection of 5% v/v acetone in water was passed through the system and measured by UV absorption at 280 nm. Absorption spectra were recorded using an AveSpec ULS2048 × 64 fiber optic spectrometer and AvaLight Deep UV source (Avantes, Lafayette, CO) connected to an FIA flow cell (FIAlab, Seattle, WA) at the outlet. Figure S1(d) of the SM [23] shows the 280-nm absorption as a function of time normalized to pure 5% acetone solution. A 200 μL injection of 5% acetone running in buffer at 0.337 mL/min thus has a full width at half maximum (FWHM) of 2 min. A 2-min fraction collected at peak is 0.674 mL with a dilution factor of 0.3.

C. Mass spectrometry sample preparation

All protein samples were deposited directly from the HiP-LiP apparatus to falcon tubes containing solid urea (512 mg) and vortexed to create a solution with a final urea concentration of 8 M. Freshly prepared dithiothreitol (DTT) (15.2 μL of a 700 mM stock) was added to the falcon tube for a final concentration of 10 mM, and the mixture was incubated at 37 °C for 30 min at 700 rpm on a thermomixer to reduce cysteine residues. A 60.6 μL portion of a freshly prepared 700 mM stock of iodoacetamide (IAA) was then added to a final concentration of 40 mM, and the mixture was incubated at room temperature in the dark for 45 min to alkylate cysteines. Subsequently, 3.174 mL of 100 mM ammonium bicarbonate solution was added to dilute urea to a concentration of 2 M. After thorough mixing of samples via vortexing, a 1 μL portion of a 1 mg/mL stock of trypsin (Pierce) was added to the samples. Mixtures were incubated on thermomixers overnight at 25 °C and 700 rpm.

Digests were then desalted with SepPak C18 1cc Vacuum Cartridges (Waters) over vacuum manifold. First, tryptic digests were acidifed with 42 μL of trifluoroacetic acid (TFA) to a final concentration of 1% (v/v). Cartridges were conditioned with 2 × 1 mL of 80% acetonitrile (ACN) with 0.5% TFA and equilibrated with 4 × 1 mL of 0.5% TFA in Optima Water. Peptides were slowly loaded under diminished vacuum, then columns were washed with 4 × 1 mL of 0.5% TFA before elution with 1 mL of 80% ACN with 0.5% TFA. During elution, cartridges were suspended above 15 mL falcon tubes in a swing-bucket rotor and spun for 4 min at 350*g*. Eluted peptides were transferred to 1.5 mL microfuge tubes and dried via vacuum centrifuge (Eppendorf Vacufuge) before being stored at −80 °C until analysis. For analysis, samples were resuspended via sonication and vortexing in 0.1% formic acid (FA) in Optima Water to a final concentration of 1 mg/mL.

D. Mass spectrometry acquisition

Chromatographic separation of digests was carried out on a Thermo UltiMate3000 UHPLC system with an Acclaim Pepmap RSLC, C18, 75 μm × 25 cm, 2 μm, 100 Å column. Approximately 1 μg of protein was injected onto the column. The column temperature was maintained at 40 °C, and the flow rate was set to 0.300 μL min^{-1} for the duration of the run. Solvent A (0.1% FA) and Solvent B (0.1% FA in ACN) were used as the chromatography solvents. The samples were run through the UHPLC system as follows: peptides were allowed to accumulate onto the trap column (Acclaim PepMap 100, C18, 75 μm × 2 cm, 3 μm, 100 Å column) for 10 min (during which the column was held at 2% Solvent B). The peptides were resolved by switching the trap column to be in line with the separating column, quickly increasing the gradient to 5% B over 5 min and then applying a 95-min linear gradient from 5% B to 40% B. Subsequently, the gradient held at 40% B for 5 min and then increased again from 40% B to 90% B over 5 min. The column was then cleaned with a sawtooth gradient to purge residual peptides between runs in a sequence.

A Thermo Q-Exactive HF-X Orbitrap mass spectrometer was used to analyze protein digests. A full MS scan in positive ion mode was followed by 20 data-dependent MS scans. The full MS scan was collected using a resolution of 120000 (at m/z 200), an AGC target of 3E6, a maximum injection time of 64 ms, and a scan range from 350 to 1500 m/z. The data-dependent scans were collected with a resolution of 15000 (at m/z 200), an AGC target of 1E5, a minimum AGC target of 8E3, a maximum injection time of 55 ms, and an isolation window of 1.4 m/z units. To dissociate precursors

prior to their reanalysis by MS2, peptides were subjected to an HCD of 28% normalized collision energies. Fragments with charges of 1, 6, 7, or higher and unassigned were excluded from analysis, and a dynamic exclusion window of 30.0 s was used for the data-dependent scans.

E. MS data analysis

The FragPipe v20.0 proteomics pipeline with IonQuant v1.9.8 with a match between runs (MBR) false discovery rate (FDR) of 5% was used to analyze spectra and perform label-free quantification (LFQ) of detected peptides. Using MSFragger v3.8 and Philosopher v5.0, a semitryptic search allowing up to two missed cleavages was conducted against the *T. thermophilus* HB27 (UP000000592, UniProt) reference proteome database, and identifications were filtered to a 1% FDR. An MS1 precursor mass tolerance of 10 ppm and an MS2 fragmentation tolerance of 20 ppm were used. Methionine oxidation and N-terminus acetylation were allowed as dynamic modifications, while carbamidomethylation on cysteines was defined as a static modification. Raw ion intensity data for identified peptides were exported and processed utilizing FLiPPR.[33] Data were merged either from the ions to the peptide level (for the proteome-wide analyses in this study) or from ions to the cut-site level (for the analysis of individual proteins in Sec. III D)—these are compiled in supporting data sets 2 and 3, respectively. In all cases, missing data imputation, filtering, and Benjami-Hochberg FDR correction were carried out per protein, as implemented in FLiPPR. In proteome-wide analyses, peptides were labeled significantly perturbed by pressure if their abundance changed by more than twofold and adjusted *p*-values were less than 0.05. In analyses of individual proteins, cut-sites were only considered if adjusted *p*-values were less than 0.02. Metadata were compiled from various sources (EcoCyc [42,79]; ECOD [40,80]; BioPython [81]), and assembled together along with the FLiPPR outputs using custom PYTHON scripts. To test whether these features discriminate for greater or lesser sensitivity to pressure, the numbers of (un)perturbed proteins and (un)perturbed peptides in various categories were counted in Excel spreadsheets and compared to a null hypothesis with the chi-square test. These counts and statistical analyses can be found in supporting data set 1.

F. Fluorescence assays

For the FITC-casein assay, 40 µL of a 1.15 mg/mL stock of FITC-casein (ThermoFisher) stored in 20 mM Tris pH 8.2, 100 mM NaCl, and 2 mM $MgCl_2$ was diluted to 7.2 µg/mL using 6.36 mL of dilution buffer (20 mM Tris pH 8.2, 100 mM NaCl, 2 mM $MgCl_2$, 64 mM $CaCl_2$) to achieve a final $CaCl_2$ concentration of 60 mM. Then, 400-µL portions were aliquoted into nine tubes. Aliquots of 200 µL from each portion were injected into the instrument, equilibrated for 10 min at 0 or 100 MPa, and then experienced limited proteolysis at a flow rate of 0.337 mL/min with sample mixing with buffer in a 1:10 ratio. Three control replicates were not exposed to pressure or Proteinase K, three replicates experienced LiP in instrument at ambient pressures, and three replicates experienced Hi-P LiP at 100 MPa. Running buffer in the Hi-P LiP instrument was 40 mM Tris pH 8.2, 200 mM NaCl, 4 mM $MgCl_2$, and 60 mM $CaCl_2$. The PK buffer was 20 mM Tris pH 8.2, 100 mM NaCl, 2 mM $MgCl_2$, 0.029 mg/mL Proteinase K for a 1:2.5 PK-to-protein mass ratio. Samples were collected for a 2-min interval containing the FWHM portion of the sample peak after quenching in the 105 °C oil bath. Collected sample was deposited directly into a tube containing 6.67 µL of Protease Inhibitor Set 3 (Calbiochem). Samples were then doubled in volume with 20% glycerol (for a final glycerol concentration of 10%) and flash frozen for transport on dry ice. Samples were later thawed and 200 µL of each were deposited into an opaque-bottom microtiter plate. Fluorimetry was conducted on a Tecan Spark microplate reader. Each well was excited at 485 nm, and emission was collected from 500 to 650 nm using bandwidth settings of 5.0 and the Dichroic 510 filter.

For the DABCYL-Edans Peptide Assay, nine 400-µL aliquots of the fluorogenic peptide (GenScript, catalog no. RP30232) at concentrations of 40 µg/mL experienced analogous treatment to the above using a running buffer composed of 20 mM Tris pH 8.2, 100 mM NaCl, 2 mM $MgCl_2$, as well as a PK buffer of 20 mM Tris pH 8.2, 100 mM NaCl, 2 mM $MgCl_2$ and 4 µg/mL PK (for a final 1:100 PK-to-peptide mass ratio). Samples were collected after quenching in the 105 °C oil bath and deposited directly into a tube containing 6.74 µL of 50 mM PMSF. 200 µL of each sample were deposited into an opaque-bottom microtiter plate. Fluorimetry was conducted on a Tecan Spark microplate reader. Each well was excited at 355 nm, and emission was collected from 450 to 600 nm using bandwidth settings of 5.0 and the Dichroic 510 filter.

G. Bioinformatics and computational methods

Sequence files for the proteomes of *Thermus thermophilus* HB27 (UP000000592) and *Escherichia coli* K12 MG1655 (UP000000625) were obtained from the UniProt database. Then, the proteomes were iteratively searched against each other using jackhmmer v3.4 [82] (with default parameter settings). Pairwise orthologs were defined using the reciprocal best hit (RBH) criterion [41]. *E. coli* metadata for protein complexes, cofactors, essentiality, and cellular location were then acquired from the EcoCyc application programming interface (API) [42,79]. These metadata were then transferred to *T. thermophilus* HB27 proteins using the RBH matches.

Structural predictions for all but five *T. thermophilus* HB27 proteins were acquired from AlphaFold2 using the publicly available API. Structural models were hydrogenated using pdb2gmx from GROMACS using default settings. Fourteen structures were then selected for relaxation simulations. The explicit-solvent all-atom molecular dynamic simulations were performed remotely on the WAXSiS server, and the relaxed protein coordinates were extracted from the final simulation frame [83]. Experimentally derived structures of *T. thermophilus* proteins were acquired from the Protein Data Bank, and all structure chains were saved individually. Last, the solvent-excluded total volume, van der Waals volume, void volume, and packing density were calculated locally for all relaxed and unrelaxed protein structures under ambient pressure conditions with ProteinVolume 1.3 using Bondi's van der Waals atomic radii and default parameter settings [43].

Supporting data set 1 provides summary data for all proteins in the 50 MPa and 100 MPa Hi-P LiP experiments. Quantifications for all peptides in the 50 MPa and 100 MPa Hi-P LiP experiments can be found in supporting data set 2. Quantifications for all cut-sites (in which peptides are merged together that correspond to the same cut-site) in the 50 MPa and 100 MPa Hi-P LiP experiments can be found in supporting data set 3. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD047671.

ACKNOWLEDGMENTS

We would like to thank the NSF Division of Molecular and Cellular Biology for a CAREER grant (Grant No. 2045844) and NIH/NIGMS for a New Innovator Award (No. DP2-GM140926) both to S.D.F. E.M.S. is thankful for the Program in Molecular Biophysics training grant (Grant No. NIH-T32GM135131). P.S. would like to acknowledge support from the Albstein Foundation for Brain Research. H.M.M. is thankful for the Chemistry-Biology Interface Program training grant (Grant No. NIH-T32GM080189-13). This research was conducted in part at the Center for High Energy X-ray Sciences (CHEXS), which is supported by the NSF (BIO, ENG and MPS Directorates) under Award No. DMR-1829070, and the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by Award No. 1-P30-GM124166-01A1 from NIH/NIGMS and by New York State's Empire State Development Corporation (NYSTAR). The authors thank Alfredo Caro for pointing out the references on 3-isopropylmalate dehydrogenase.

[1] R. Durvasula and D. V. S. Rao, *Extremophiles: From Biology to Biotechnology* (CRC Press, Boca Raton, FL, 2018).

[2] N. Merino, H. S. Aronson, D. P. Bojanova, J. Feyhl-Buska, M. L. Wong, S. Zhang, and D. Giovannelli, Living at the extremes: Extremophiles and the limits of life in a planetary context, Front. Microbiol. **10**, 780 (2019).

[3] D. R. Colman, S. Poudel, B. W. Stamps, E. S. Boyd, and J. R. Spear, The deep, hot biosphere: Twenty-five years of retrospection, Proc. Natl. Acad. Sci. USA **114**, 6895 (2017).

[4] V. Balasubramaniam, Process development of high pressure-based technologies for food: Research advances and future perspectives, Curr. Opin. Food Sci. **42**, 270 (2021).

[5] L. Smeller, Biomolecules under pressure: Phase diagrams, volume changes, and high pressure spectroscopic techniques, Int. J. Mol. Sci. **23**, 5761 (2022).

[6] R. Winter, D. Lopes, S. Grudzielanek, and K. Vogtt, Towards an understanding of the temperature/pressure configurational and free-energy landscape of biomolecules, J. Non-Equilib. Thermodyn. **32**, 41 (2007).

[7] D. J. Wilton, M. Ghosh, K. V. A. Chary, K. Akasaka, and M. P. Williamson, Structural change in a B-DNA helix with hydrostatic pressure, Nucleic Acids Res. **36**, 4032 (2008).

[8] J. Somkuti, O. R. Molnár, A. Grád, and L. Smeller, Pressure perturbation studies of noncanonical viral nucleic acid structures, Biology (Basel) **10**, 1173 (2021).

[9] J.-M. Knop, S. Patra, B. Harish, C. A. Royer, and R. Winter, The deep sea osmolyte trimethylamine n-oxide and macromolecular crowders rescue the antiparallel conformation of the human telomeric G-quadruplex from urea and pressure stress, Chem. Eur. J. **24**, 14346 (2018).

[10] J. Wang *et al.*, Pressure pushes $tRNA^{Lys3}$ into excited conformational states, Proc. Natl. Acad. Sci. USA **120**, e2215556120 (2023).

[11] H.-L. Sung and D. J. Nesbitt, High pressure single-molecule FRET studies of the lysine riboswitch: Cationic and osmolytic effects on pressure induced denaturation, Phys. Chem. Chem. Phys. **22**, 15853 (2020).

[12] G. N. Somero, Adaptations to high hydrostatic pressure, Annu. Rev. Physiol. **54**, 557 (1992).

[13] R. Winter, Pressure effects on artificial and cellular membranes, in *High Pressure Bioscience: Basic Concepts, Applications and Frontiers*, edited by K. Akasaka and H. Matsuki (Springer Netherlands, Dordrecht, 2015), pp. 345–370.

[14] J. R. Winnikoff *et al.*, Homeocurvature adaptation of phospholipids to pressure in deep-sea invertebrates, Science **384**, 1482 (2024).

[15] M. J. Eisenmenger and J. I. Reyes-De-Corcuera, High pressure enhancement of enzymes: A review, Enzyme Microb. Technol. **45**, 331 (2009).

[16] J. L. Silva, A. C. Oliveira, T. C. R. G. Vieira, G. A. P. de Oliveira, M. C. Suarez, and D. Foguel, High-pressure chemical biology and biotechnology, Chem. Rev. **114**, 7239 (2014).

[17] *High Pressure Bioscience Basic Concepts, Applications and Frontiers*, edited by K. Akasaka, Vol. 72 (Springer, New York, 2015).

[18] N. Ando *et al.*, The molecular basis for life in extreme environments, Annu. Rev. Biophys. **50**, 343 (2021).

[19] R. Winter, Interrogating the structural dynamics and energetics of biomolecular systems with pressure modulation, Annu. Rev. Biophys. **48**, 441 (2019).

[20] J. L. Silva and G. Weber, Pressure stability of proteins, Annu. Rev. Phys. Chem. **44**, 89 (1993).

[21] L. Smeller, Pressure–temperature phase diagrams of biomolecules, Biochim. Biophy. Acta. Prot. Struct. Mol. Enzymol. **1595**, 11 (2002).

[22] J. Roche and C. A. Royer, Lessons from pressure denaturation of proteins, J. R. Soc. Interface. **15**, 20180244 (2018).

[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PRXLife.2.033011 for valve settings on Hi-P LiP apparatus (Fig. S1), descriptive statistics (Fig. S2), statistics broken down by cofactor (Fig. S3), controls for packing density calculations (Fig. S4), comparison of Bondi and O'Hern parameters on packing density (Fig. S5), statistics by number of subunits (Fig. S6), statistics by degree of disorder (Fig. S7), and pressure-sensitive sites on GroEL surface (Fig. S8). Material includes Refs. [39,63,83].

[24] P. To, B. Whitehead, H. E. Tarbox, and S. D. Fried, Nonrefoldability is pervasive across the *E. coli* proteome, J. Am. Chem. Soc. **143**, 11435 (2021).

[25] Y. Feng, G. De Franceschi, A. Kahraman, M. Soste, A. Melnik, P. J. Boersema, P. P. de Laureto, Y. Nikolaev, A. P. Oliveira, and P. Picotti, Global analysis of protein structural changes in complex proteomes, Nat. Biotechnol. **32**, 1036 (2014).

[26] I. Piazza, K. Kochanowski, V. Cappelletti, T. Fuhrer, E. Noor, U. Sauer, and P. Picotti, A map of protein-metabolite interactions reveals principles of chemical communication, Cell **172**, 358 (2018).

[27] J. Paukštytė *et al.*, Global analysis of aging-related protein structural changes uncovers enzyme-polymerization-based control of longevity, Mol. Cell **83**, 3360 (2023).

[28] P. Leuenberger, S. Ganscha, A. Kahraman, V. Cappelletti, P. J. Boersema, C. von Mering, M. Claassen, and P. Picotti, Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability, Science **355**, eaai7825 (2017).

[29] P. To, Y. Xia, S. O. Lee, T. Devlin, K. G. Fleming, and S. D. Fried, A proteome-wide map of chaperone-assisted protein refolding in a cytosol-like milieu, Proc. Natl. Acad. Sci. USA **119**, e2210536119 (2022).

[30] M. Pepelnjak *et al.*, In situ analysis of osmolyte mechanisms of proteome thermal stabilization, Nat. Chem. Biol. **20**, 1053 (2024).

[31] R. C. Miller, C. Cummings, Q. Huang, N. Ando, and R. E. Gillilan, Inline small-angle x-ray scattering-coupled chromatography under extreme hydrostatic pressure, Protein Sci. **31**, e4489 (2022).

[32] A. T. Kong, F. V. Leprevost, D. M. Avtonomov, D. Mellacheruvu, and A. I. Nesvizhskii, MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics, Nat. Methods **14**, 513 (2017).

[33] F. Yu, S. E. Haynes, and A. I. Nesvizhskii, IonQuant enables accurate and sensitive label-free quantification with FDR-controlled match-between-runs, Mol. Cell. Proteomics **20**, 100077 (2021).

[34] E. Manriquez-Sandoval, J. Brewer, G. Lule, S. Lopez, and S. D. Fried, FLiPPR: A processor for limited proteolysis (LiP) mass spectrometry datasets built on FragPipe, J. Prot. Res. **23**, 2332 (2024).

[35] D. N. Wilson and K. H. Nierhaus, Ribosomal proteins in the spotlight, Crit. Rev. Biochem. Mol. Biol. **40**, 243 (2005).

[36] W. Margolin, FtsZ and the division of prokaryotic cells and organelles, Nat. Rev. Mol. Cell Biol. **6**, 862 (2005).

[37] D. Mazel, S. Pochet, and P. Marlière, Genetic characterization of polypeptide deformylase, a distinctive enzyme of eubacterial translation, EMBO J. **13**, 914 (1994).

[38] R. Gebhardt, N. Takeda, U. Kulozik, and W. Doster, Structure and stabilizing interactions of casein micelles probed by high-pressure light scattering and FTIR, J. Phys. Chem. B **115**, 2349 (2011).

[39] R. J. Emenecker, D. Griffith, and A. S. Holehouse, Metapredict: A fast, accurate, and easy-to-use predictor of consensus disorder and structure, Biophys. J. **120**, 4312 (2021).

[40] E. Manriquez-Sandoval and S. D. Fried, DomainMapper: Accurate domain structure annotation including those with non-contiguous topologies, Protein Sci. **31**, e4465 (2022).

[41] R. L. Tatusov, E. V. Koonin, and D. J. Lipman, A genomic perspective on protein families, Science **278**, 631 (1997).

[42] I. M. Keseler *et al.*, The EcoCyc database in 2021, Front. Microbiol. **12**, 711077 (2021).

[43] C. R. Chen and G. I. Makhatadze, ProteinVolume: Calculating molecular van der Waals and void volumes in proteins, BMC Bioinf. **16**, 101 (2015).

[44] C. R. Chen and G. I. Makhatadze, Molecular determinant of the effects of hydrostatic pressure on protein folding stability, Nat. Commun. **8**, 14561 (2017).

[45] J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold, Nature (London) **596**, 583 (2021).

[46] M. Varadi *et al.*, Alpha Fold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models, Nucleic Acids Res. **50**, D439 (2022).

[47] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX **1–2**, 19 (2015).

[48] V. Pattni and M. Heyden, Pressure effects on protein hydration water thermodynamics, J. Phys. Chem. B **123**, 6014 (2019).

[49] A. Levin, S. Cinar, M. Paulus, J. Nase, R. Winter, and C. Czeslik, Analyzing protein-ligand and protein-interface interactions using high pressure, Biophys. Chem. **252**, 106194 (2019).

[50] A. M. Couture and K. J. Laidler, The partial molal volumes of ions in aqueous solution: I. Dependence on charge and radius, Can. J. Chem. **34**, 1209 (1956).

[51] A. T. Grigas, Z. Liu, L. Regan, and C. S. O'Hern, Core packing of well-defined x-ray and NMR structures is the same, Protein Sci. **31**, e4373 (2022).

[52] J. D. Treado, Z. Mei, L. Regan, and C. S. O'Hern, Void distributions reveal structural link between jammed packings and protein cores, Phys. Rev. E **99**, 022416 (2019).

[53] K. Akasaka, Probing conformational fluctuation of proteins by pressure perturbation, Chem. Rev. **106**, 1814 (2006).

[54] K. L. Maxwell *et al.*, Protein folding: Defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins, Protein Sci. **14**, 602 (2005).

[55] J. C. Gaines, A. H. Clark, L. Regan, and C. S. O'Hern, Packing in protein cores, J. Phys. Condens. Matter **29**, 293001 (2017).

[56] F. M. Richards, The interpretation of protein structures: Total volume, group volume distributions and packing density, J. Mol. Biol. **82**, 1 (1974).

[57] J. C. Gaines, S. Acebes, A. Virrueta, M. Butler, L. Regan, and C. S. O'Hern, Comparing side chain packing in soluble proteins, protein-protein interfaces, and transmembrane proteins, Proteins **86**, 581 (2018).

[58] M. Gross and R. Jaenicke, Proteins under pressure, Eur. J. Biochem. **221**, 617 (1994).

[59] G. Weber, *Protein Interactions* (Chapman and Hall, New York, 1992).

[60] R. E. Gillilan, High-pressure SAXS, deep life, and extreme biophysics, in *Methods in Enzymology* (Academic Press, New York, 2022).

[61] D. Moses, G. M. Ginell, A. S. Holehouse, and S. Sukenik, Intrinsically disordered regions are poised to act as sensors of cellular chemistry, Trends Biochem. Sci. **48**, 1019 (2023).

[62] C. L. Cuevas-Velazquez *et al.*, Intrinsically disordered protein biosensor tracks the physical-chemical effects of osmotic stress on cells, Nat. Commun. **12**, 5438 (2021).

[63] T. Shimamura, A. Koike-Takeshita, K. Yokoyama, R. Masui, N. Murai, M. Yoshida, H. Taguchi, and S. Iwata, Crystal structure of the native chaperonin complex from *Thermus thermophilus*

revealed unexpected asymmetry at the cis-cavity, Structure **12**, 1471 (2004).

[64] Z. Xu, A. L. Horwich, and P. B. Sigler, The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex, Nature (London) **388**, 741 (1997).

[65] D. K. Clare, D. Vasishtan, S. Stagg, J. Quispe, G. W. Farr, M. Topf, A. L. Horwich, and H. R. Saibil, ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin, Cell **149**, 113 (2012).

[66] O. Llorca, S. Marco, J. L. Carrascosa, and J. M. Valpuesta, Conformational changes in the GroEL oligomer during the functional cycle, J. Struct. Biol. **118**, 31 (1997).

[67] J. Krucinska, E. Falcone, H. Erlandsen, A. Hazeen, M. N. Lombardo, A. Estrada, V. L. Robinson, A. C. Anderson, and D. L. Wright, Structural and functional studies of bacterial enolase, a potential target against Gram-negative pathogens, Biochemistry **58**, 1188 (2019).

[68] O. A. Zadvornyy, E. S. Boyd, M. C. Posewitz, N. A. Zorin, and J. W. Peters, Biochemical and structural characterization of enolase from *Chloroflexus aurantiacus*: Evidence for a thermophilic origin, Front. Bioeng. Biotechnol. **3**, 74 (2015).

[69] R. Hidese *et al.*, The C-terminal flexible region of branched-chain polyamine synthase facilitates substrate specificity and catalysis, FEBS J. **286**, 3926 (2019).

[70] W. Shi, K. S. E. Tanaka, T. R. Crother, M. W. Taylor, S. C. Almo, and V. L. Schramm, Structural analysis of adenine phosphoribosyltransferase from *Saccharomyces cerevisiae*, Biochemistry **40**, 10800 (2001).

[71] M. Ozeir *et al.*, Structural basis for substrate selectivity and nucleophilic substitution mechanisms in human adenine phosphoribosyltransferase catalyzed reaction, J. Biol. Chem. **294**, 11980 (2019).

[72] C. H. T. P. Silva, M. Silva, J. Iulek, and O. H. Thiemann, Structural complexes of human adenine phosphoribosyltransferase reveal novel features of the APRT catalytic mechanism, J. Biomol. Struct. Dyn. **25**, 589 (2008).

[73] Y. Suzuki, K. Asada, J. Miyazaki, T. Tomita, T. Kuzuyama, and M. Nishiyama, Enhancement of the latent 3-isopropylmalate dehydrogenase activity of promiscuous homoisocitrate dehydrogenase by directed evolution, Biochem. J. **431**, 401 (2010).

[74] M. P. Williamson, Pressure-dependent conformation and fluctuation in folded protein molecules, in *High Pressure Bioscience: Basic Concepts, Applications and Frontiers*, edited by K. Akasaka and H. Matsuki, Vol. 72 (Springer, New York, 2015).

[75] T. Nagae, T. Kawamura, L. M. G. Chavas, K. Niwa, M. Hasegawa, C. Kato, and N. Watanabe, High-pressure-induced water penetration into 3-isopropylmalate dehydrogenase, Acta. Crystallogr. D Biol. Crystallogr. **68**, 300 (2012).

[76] T. Nagae, C. Kato, and N. Watanabe, Structural analysis of 3-isopropylmalate dehydrogenase from the obligate piezophile *Shewanella benthica* DB21MT-2 and the nonpiezophile *Shewanella oneidensis* MR-1, Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. **68**, 265 (2012).

[77] J. Wang, A. Jain, L. R. McDonald, C. Gambogi, A. L. Lee, and N. V. Dokholyan, Mapping allosteric communications within individual proteins, Nat. Commun. **11**, 3862 (2020).

[78] Y. Kitamura and T. Itoh, Reaction volume of protonic ionization for buffering agents. Prediction of pressure dependence of pH and pOH, J. Solution Chem. **16**, 715 (1987).

[79] P. D. Karp *et al.*, The EcoCyc database (2023), EcoSal Plus **11**, eesp (2023).

[80] H. Cheng, R. D. Schaeffer, Y. Liao, L. N. Kinch, J. Pei, S. Shi, B.-H. Kim, and N. V. Grishin, ECOD: An evolutionary classification of protein domains, PLoS Comput. Biol. **10**, e1003926 (2014).

[81] P. J. A. Cock *et al.*, Biopython: Freely available PYTHON tools for computational molecular biology and bioinformatics, Bioinformatics **25**, 1422 (2009).

[82] L. S. Johnson, S. R. Eddy, and E. Portugaly, Hidden Markov model speed heuristic and iterative HMM search procedure, BMC Bioinf. **11**, 431 (2010).

[83] C. J. Knight and J. S. Hub, WAXSiS: A web server for the calculation of SAXS/WAXS curves based on explicit-solvent molecular dynamics, Nucleic Acids Res. **43**, W225 (2015).