THE MOLECULAR BASIS FOR LIFE IN EXTREME ENVIRONMENTS

Nozomi Ando[1,2], Blanca Barquera[3], Doug Bartlett[4], Eric Boyd[5], Audrey A. Burnim[1], Amanda S. Byer[1], Daniel Colman[5], Richard E. Gillilan[6], Martin Gruebele[7], George Makhatadze[3], Catherine A. Royer[3,*], Everett Shock[8], A. Joshua Wand[9], Maxwell B. Watkins[2]

[1]Department of Chemistry & Chemical Biology, Cornell University, Ithaca, NY 14853
[2]Department of Chemistry, Princeton University, Princeton, NJ 08544
[3]Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy NY 12185
[4]Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA 92093-0202
[5]Department of Microbiology and Immunology, Montana State University, Bozeman, MT 59717
[6]Center for High Energy X-ray Sciences (CHEXS), Ithaca, NY 14853
[7]Departments of Chemistry, Physics, and Center for Biophysics and Quantitative Biology, University of Illinois, Urbana-Champaign, Il 61801
[8]Department of …
[9]9Departments of Biochemistry & Biophysics, Chemistry, and Molecular & Cellular Medicine, Texas A&M University, College Station, TX 77845

[*]To whom correspondence should be addressed

Extreme environments and extremophiles – What are they?

Extreme environments and the organisms that inhabit them – extremophiles – serve as model platforms for probing biomolecular adaptations that allow life to thrive under environmental conditions that are intolerable for the life forms with which most of us are familiar. Extreme environments are host to a large fraction of the microbial biomass on Earth (1, 2). But what exactly do we mean by extreme environments and extremophiles? There are several ways to approach this question, including from the perspective of the environment or from the perspective of the organism. Here, we integrate aspects from both perspectives to arrive at a working definition of what adaptive features unify extremophilic organisms and how environmental extremes help to illuminate such features. The key concept in this context is habitability.

The concept of habitability is often considered binary, whereby some environments are habitable while other environments are not. What then defines whether an environment is habitable?

Given that non-water-based life has yet to be detected, a first order constraint on an environment's habitability is the presence of water. An extreme environment with respect to water availability could thus be defined as one that has extremely low water activity, such as in a brine or in a water inclusion in a sodium chloride crystal (3). Such environments have been widely shown to host thriving populations of halophilic Archaea that can tolerate salinities approaching that of halite precipitation (4). Alternatively, the habitability of an environment with respect to water availability could be defined with respect to the periodicity and extent by which water becomes available, such as in the Atacama desert in Chile where precipitation is low and infrequent. Numerous environments in the Atacama, nonetheless, have been shown to host abundant microbial life (5).

The habitability of an environment must also be defined with respect to its physical and chemical properties. A major constraint on life is temperature, with the upper limit for life currently standing at 122 °C for a methanogenic archaeon (6). Habitable limits also include extremes of pH. The current lower limit is held by the archaeon *Picrophilus torridus* that can replicate at a pH as low as -0.06 (7), equivalent to a solution of 1.2 M sulfuric acid. At the other end of the pH spectrum, *Alkaliphilus transvaalensis* isolated from a South African gold mine has been shown to grow at a pH as high as 12.5 (8).

Deep-ocean and deep subsurface environments harbor organisms that exist at elevated hydrostatic pressure over a range of temperatures. These include psychropiezophilic and obligately psychropiezophilic bacteria, such as *Colwellia marinimaniae*, obtained from the Challenger Deep in the Mariana Trench which can grow at pressures up to 140 MPa (9), and hyperthermophilic archaea, such as the obligate piezophile *Pyrococcus yayanosii* obtained from a deep-sea hydrothermal vent on the Mid-Atlantic ridge which is capable of growth up to 120 MPa (10). Microorganisms can survive pressures more than an order of magnitude above those permitting growth (11, 12), despite dramatic changes in membrane viscosity (13) and decreases in the diffusion of water, ions and metabolites within the cytoplasm (14).

A habitable environment must also be defined as one that has all of the pre-requisites for growth, including major and trace elements to support cellular metabolism. This includes sources of carbon, nitrogen, phosphorous, sulfur and oxygen, as well as metal ions for processes such as metabolic catalysis. And finally, a habitable environment must provide a source of energy. As cells encounter conditions that destabilize their biomolecules, they require an input of energy to drive *de novo* biomolecule synthesis and to operate ATP-dependent active transporters to maintain gradients across the cellular membrane (15–17). If an environment's supply of energy exceeds a cell's demand for energy, life can persist (18), otherwise it becomes uninhabitable. In many subsurface ecosystems, the availability of electron donors or acceptors is limited, leading to conditions of energy limitation and extremely slow growth with long turnover times (19–21). These conditions can culminate in very low biodiversity, as has been observed in fracture fluids obtained from continental deep (2.8 km) subsurface environments (22).

In summary, habitability means a solvent to carry out metabolism and information processing, a physical environment in which constituents are at least metastable, a source of building blocks, and a source of energy to maintain homeostasis and replication. This definition of habitability provides a framework to unify extreme environments as those that impose temporary or chronic stress on cells, either through stress from physical or chemical characteristics of an environment (e.g., water availability, temperature, pressure, pH, trace elements), or stress from lack of energy (e.g. lack of electron acceptors/donors, lack of high energy photons to create electron donors). As such, extremophiles can be defined as organisms that have evolved biomolecular traits or physiological strategies to not just survive but to thrive under conditions that impose chronic energy limitation (23). In such a scenario, selection should act to evolve mechanisms and biomolecules that allow cells to diversify into environments that impose energetic stress while balancing or minimizing the energetic costs of doing so.

While life is thought to have originated at hydrothermal vents (24), organisms that eventually populated more moderate conditions are thought to have diversified to inhabit available ecological niches and in doing so, have continually expanded the habitable limits for life. This appears to have been the case for adaptations to chronic energy stress imposed by pH extremes. In the case of acidic high temperature habitats, which are characteristically dominated by Archaea (23), it appears that several lineages diversified from moderately acidic environments into progressively more acidic environments (25). Intriguingly, data indicates that these diversification events were enabled by acquisition of metabolic pathways and enzymes that allowed cells to integrate oxygen into their energy metabolism. It has been suggested that acidophiles and their aerobic sulfur oxidizing activity were responsible for generating acidic hot spring environments (25), allowing these cells to evolve in concert with the progressive acidification of their habitats in what has been termed niche construction (26). Comparative genomic analyses also indicate that diversification of archaeal life into acidic hot spring habitats was facilitated by acquisition of numerous proteins and pathways that enable cells to maintain cytoplasmic osmotic balance (15, 25) and an ability to synthesize more structurally rigid isoprenoid lipid structures that decrease the potential for proton permeation into the cytoplasm (27).

In the case of alkaliphiles, the pattern is a bit more nuanced. Nonetheless, a recent study of microbial life in a subsurface habitat undergoing the geological process of serpentinization suggests that organisms may have also diversified into these habitats (28). The serpentinization process begins when ultramafic rocks become hydrated, initiating a series of chemical reactions that ultimately generate fluids with hyperalkaline pH and that are enriched in lithogenic hydrogen, formate, and methane (29). The increased pH leads to inorganic carbon limitation a characteristic that presents energetic challenges to methanogens and acetogens since CO_2 serves as both the electron acceptor and carbon source. A comparative phylogenomic analysis of two closely related methanogen (genus *Methanobacterium*; phylum *Euryarchaeota*) populations from the subsurface of the Samail Ophiolite, Oman, showed that the population from less reacted, near-neutral waters (pH 7.6) is ancestral to the population recovered from hyperalkaline waters (pH 11.3), suggesting diversification of this lineage into higher pH waters (Fones et al, submitted). Intriguingly while the genome of the *Methanobacterium* population from pH-neutral waters encodes proteins to couple hydrogen oxidation to CO_2 reduction, typical of hydrogenotrophic and autotrophic methanogens, the genome of the more recently evolved *Methanobacterium* population from hyperalkaline waters lacks homologs of two key oxidative [NiFe]-hydrogenases and these functionalities are replaced by formate dehydrogenases. This elegant adaptation allows cells to oxidize formate to yield reductant and cytoplasmic CO_2, allowing cells to overcome CO_2/oxidant limitation in hyperalkaline waters.

To date, it is not firmly established whether life diversified into or out of high pressure environments. In any case, piezophiles are able to grow at high pressure in part through modulation of membrane fluidity. Elevated pressure orders lipid membranes in an analogous

fashion that accompanying a decreased temperature. Piezophilic bacteria compensate for these effects by further fluidizing their membranes via increasing the proportions of unsaturated fatty acids within their membranes (30). The story with piezophilic archaeal membranes is still developing but can include pressure-induced changes in the amount of membrane spanning lipids and perhaps also changes in lipid unsaturation (31). It is also likely that piezophilic microbes adapt to high pressure via the incorporation of specific osmolytes, although evidence is lacking. Deep-sea animals increase their levels of the compatible solute trimethylamine-N-oxide with increasing capture depth all the way to solution saturation, presumably as a mechanism for compensating for pressure effects on cellular macromolecules (32).

These case studies highlight the potential for extremophiles to illuminate the diversity of life on Earth, its generation and maintenance, feedbacks between environmental and biological change, and the physiological and biomolecular adaptations in proteins, pathways, and other cellular structures that allow life to diversify and push the limits of environmental habitability. The diversity of extremophilic organisms and the biomolecules that they synthesize represent a database that can be exploited for biotechnological and industrial applications that require enzymes or other biomolecules with specialized properties (17, 33). This includes applications in pharmacology, agriculture, textile production, and biofuels production, among others. Application of biophysical tools to better understand the basic tenets underpinning their structure and function allows for such enzyme platforms to be engineered for optimal target activity and/or to be tailored and integrated into suitable production platforms.

Genomic insights into extremophiles

The development and widespread application of molecular biology tools, in particular DNA sequencing technologies, to natural environments over the past several decades have revolutionized our understanding of the diversity, distribution, and evolution of microorganisms (Fig. 1). DNA sequencing tools are especially useful for predicting the metabolisms of and uncovering unique adaptations among uncultured microorganisms, which comprise the overwhelming majority of microbial taxa. Up to ~87% of genera and 64% of phyla lack cultured representatives, and the extent of uncultured phyla is environment-specific (34). The development of modern sequencing techniques, including sequencing-by-synthesis methods (e.g., Illumina sequencing), and concomitant computational improvements and software developments that enable the analysis of large sequence datasets have allowed for the recovery of high quality genomes from environmental DNA (i.e., metagenomes) *via* metagenome-assembled-genomes (MAGs). These techniques are now also commonly employed to sequence and analyze the genomes of single cells (single amplified genomes; SAGs).

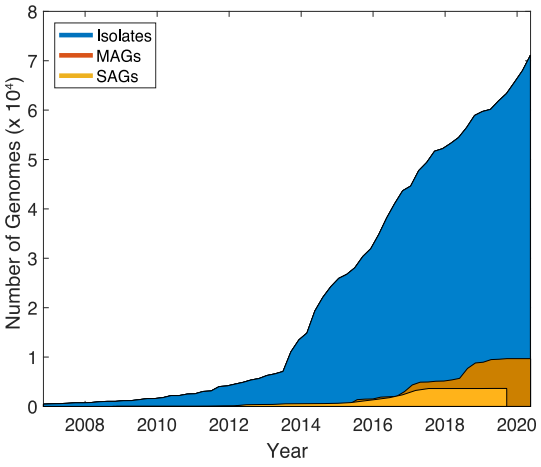

Figure 1. The cumulative number of genomes in the Department of Energy Integrated Microbial Genomes (IMG) database. Genomes are separated into those generated from cultivars, genomes assembled from metagenome data (MAGs), and single cell amplified genomes (SAGs) and ordered by their 'addition date'. The availability of SAG data for 2020 has yet to be reported and is thus not included on this plot.

To date, sequencing efforts (in particular of cultivar genomes) have skewed towards organisms from host-associated environments (e.g., from human-derived isolates; Figure 2), consistent with the general bias of culture collections towards organisms recovered from such environments (34). Similarly, most metagenomics studies have focused on non-extreme surface environments (e.g., soils, marine, and freshwaters). This is in spite of evidence that extreme hydrothermal and subsurface environments harbor a greater extent of phylum-level archaeal and bacterial diversity (Figure 2). Below, we describe recent insights

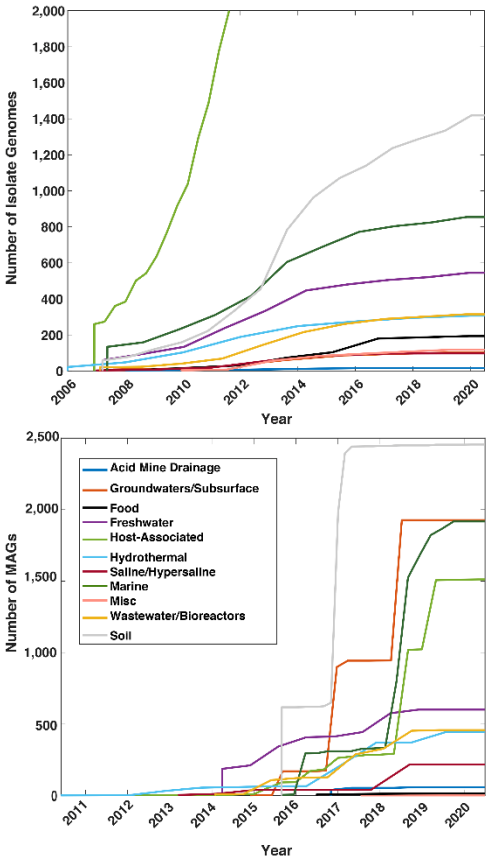

Figure 2. The distribution of environment types where isolate genomes (top) and MAGs (bottom) have been generated. Metadata for each genome was used to group genomes into environment types from where they originated. The curves show accumulation of genomes from each environment type over time and are colored by the legend at the bottom. The host-associated isolate genomes curve is not shown for clarity, given the high number of genomes recovered from this environment type.

provided by metagenomics studies of two model extreme environments that capture environmental attributes that are most likely to drive chronic energy stress in microbial inhabitants: high temperature hydrothermal environments and high pressure and/or nutrient limited subsurface environments.

Hydrothermal systems Hydrothermal systems present collectively the widest variety of ecological niche space capable of supporting microbial life on Earth. This includes niches that span temperature gradients (ambient to 122°C), pH gradients (<1 to 12), and gradients in the availability of soluble and precipitated (mineral) nutrient substrates (18). Given the plethora of potentially habitable niche space in hydrothermal systems, it is not surprising that they are also host to among the most diverse microbial communities studied, both at the taxonomic and functional levels (35). Despite being considerably under-sampled relative to less extreme environments (Figure 2), a plot of the phylum level taxonomic diversity of genomes recovered from hydrothermal systems reveals them to comprise the second highest level of diversity among habitat types in the IMG genome database (Figure 3). For example, a metagenomics study of a single Yellowstone National Park (YNP) spring revealed the presence of populations that belong to nearly 50% of the then-known bacterial and archaeal taxonomic higher order clades (e.g., phyla), including several previously undocumented microbial phyla (35). Intriguingly, MAGs recovered from this single spring expanded the known distribution of key microbial metabolisms (e.g., methane/alkane metabolism and sulfate/sulfite reduction) among phyla where these metabolisms had previously not been documented (35). One would expect other hydrothermal systems with similar geochemical gradients to those in YNP to harbor comparable taxonomic and functional diversity.

Among the most intriguing discoveries made over the past decade is the identification of protein homologs in novel phyla involved in global biogeochemical cycles. For example, the discovery of Methyl coenzyme reductase (Mcr) proteins, which catalyze the last step of methanogenesis or the first step of anaerobic alkane oxidation in the genomes of uncultured '*Bathyarchaeota*' lineage provided evidence that these Mcr proteins in *Bathyarchaeota* may be involved in the oxidation of longer chain alkanes, such as butane (36). Metagenomics studies of marine and terrestrial hydrothermal systems have also greatly expanded the diversity of Mcr in taxonomic groups not previously known to conduct methanogenesis or alkanotrophy (35–37). Cross-environment comparisons indicate that the unique geochemistry of these environments, in particular their sourcing with volcanic gases enriched in alkanes, may act as a selective pressure for inclusion of organisms with the ability to metabolize short chain hydrocarbons during community assembly (38). Notably, the structural determinants of catalytic bias in Mcr are not known but could have tremendous importance in biotechnological applications. Additional biochemical and biophysical characterization is needed to better understand the determinants of

Mcr enzymatic directionality, the enzymatic basis for activating specific short chain alkanes, and whether additional novel functionalities may be conducted by divergent Mcr.

Subsurface Environments Our understanding of the presence, extent, and nature of subsurface microbial ecosystems has only recently come into focus. Drilling and subterranean exploration have provided access to the subsurface for study, and recent evidence indicates that marine and continental subsurface environments may be habitable to greater than 5 km below the surface (1). Further, this vast habitable environment may host a significant fraction of all microbial biomass on Earth (1, 2). Given the shear enormity of these ecosystems, their potential to inform on the origin and diversity of life on Earth, and their role in global biogeochemical cycles, they have received considerable attention in recent years and metagenomics datasets from them are increasingly becoming available (Figure 2).

Subsurface life is exposed to numerous extreme stresses including higher temperatures at greater depths, and a limited flux of nutrients or limited gradients in electron donors/acceptors to fuel metabolism and replication and, notably, increased pressure (39, 40). Consequently, adaptations to overcome these stressors are critical for organisms inhabiting subsurface systems. Organisms that live in deep ocean environments are likewise exposed to high pressure. Processes that appear to be modified in piezophiles based on comparative genomics, omics measurements, and genetic experiments are energy metabolism and initiation of DNA replication and translation (41–43). For example, deep-sea archaea of the *Thaumarchaeota* phylum, one of the most abundant microbes on Earth, have acquired vacuolar-type ATPase genes via horizontal gene transfer that facilitates growth at low pH. Its adaptive value to growth in the deepest ocean trenches may relate to pressure effects on the membrane environment in which this complex molecular machine operates (44).

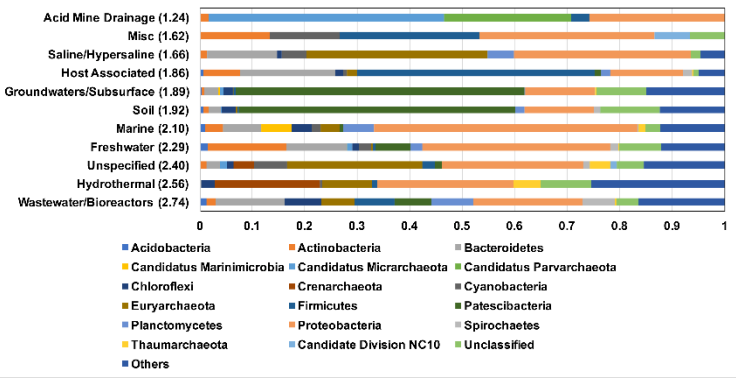

Figure 3. Phylum-level taxonomic composition of MAGs within environment types identified in Figure 2. Environment types are arranged with those with highest inferred phylum diversity (based on the Shannon diversity index) at the bottom. The numbers in parentheses show the Shannon diversity index values for phylum-level composition in each environment type. Phyla are shown for those that comprised over 5% of the phyla for each environment type, while the rest are pooled as 'others'. Phylum designations are those provided by IMG metadata, with the exception of Candidate Phylum Radiation lineage, which are collapsed as 'Patescibacteria' based on recent taxonomic revisions ref 46.

One of the most well-studied subsurface systems from a genomics perspective is that of the Rifle aquifer of Colorado, which exhibits low-level mine waste contamination. Several metagenomic studies of this system have revealed an incredibly high diversity of uncultivated lineages, primarily affiliated with the proposed 'Candidate Phyla Radiation' (CPR; later subsumed by taxonomic classification into the Patescibacteria group ;) (45, 46), which generally also exhibit streamlined, small genomes. Indeed, streamlined genomes have also been observed in other subsurface systems where nutrient limitation is prevalent (47, 48), suggesting that minimizing genome size to reduce the cost of replication may be a common adaptation in subsurface life. Additional insights into physiological adaptations that allow for subsurface life have been obtained from metagenomic investigations of fracture fluids from deep continental mines. In particular, subsurface mine waters from South Africa identified *Candidatus Desulforudis audaxviator* as the near-only inhabitant of fracture fluids (22) and this organism has since been shown to be cosmopolitan member of subsurface communities (49). Like many other subsurface environments, including those undergoing the geological process of serpentinization (47, 48, 50), many organisms in subsurface environments appear to be supported by products of water-rock interactions, including hydrogen, formate, carbon monoxide and methane. A study of a single hot spring in Yellowstone sourced by H_2 rich volcanic gas

revealed that presence of a diverse array of hydrogenases in >72% of the organisms that inhabit the spring (51). Whether the diverse enzymes involved in transformation of substrates like H2 discovered in metagenomic datasets harbor features of interest (e.g., acid or alkali tolerance, thermal or pressure tolerance) for industrial or biotechnological applications awaits further biochemical and biophysical characterization. Such questions can be extended to any protein homolog whose distribution spans natural geochemical gradients, providing a robust dataset and a natural platform to interrogate and understand the molecular basis for extreme life.

The Molecular Basis for Extreme Life

As noted above, our understanding of extreme environments and the extremophiles living in them has expanded tremendously in recent years. We now have access to the sequences of literally millions of genes, and in principle, metadata describing the physical and chemical characteristics of their environments. While the linking of meta-data to environmental genomes and metagenomes remains to be accomplished holistically and constitutes a significant roadblock to progress in understanding the molecular signatures of adaptation, we can already envision mining the genomic data for biochemical and biophysical insights. The fact that oftentimes the function of less than 50% of putative genes can be deduced in the genomes of extremophiles using homology-based methods presents another significant roadblock, but also an opportunity for identifying novel bio-catalytic systems of potential technological importance.

At the same time that information about extreme environments is becoming more readily available, considerable strides are being made in understanding of the mechanisms by which protein sequence underpins functional properties. Proteins are highly complex molecules with tens to hundreds of thousands of atoms. They populate dynamic conformational ensembles that can be modified significantly by single amino acid substitutions. Enzyme activity, or the lack thereof, hangs on just a few kT units of energy ($1\ kT \sim 2.5$ kJ/mol), and function is often dependent on the transient population of excited conformational states. Such order-disorder transitions are central to biomolecular recognition and signaling. Relatively modest amino acid substitutions can have profound effects on these dynamic equilibria. In many cases, 'folding intermediates' of proteins may turn out to be signals of altered flexibility and dynamics of excited conformations under temperature, pressure, or other stresses.

We now know how and why extreme physical and chemical conditions modulate the functionally important conformational dynamics of proteins. However, conceptually connecting the systemic information available in the genomic databases with the subtlety of protein conformational landscapes presents a major challenge to identifying the molecular signatures of adaptation to extreme environments. It is the interplay between biochemistry and biophysics with the geochemistry and physical conditions of these environments that defines, ultimately, what biological functions are present and how they operate. An understanding of the underlying molecular basis for extreme life lies in both the biophysical aspects of any given protein function and the environmental properties that have shaped the evolutionary history of that protein. Efforts to reconstruct ancestral proteins and correlate them climate- and geochemical changes over geological time scales are still in their infancy (52).

Biomolecular Adaptation – What do we know?

Fortunately our current knowledge of extremophile proteins is rapidly growing. General properties of biomolecules from extremophiles have been reported, and specific types of amino acids appear to be preferred in protein sequences from specific environments (53, 54). Proteins from thermophilic organisms tend to exhibit a higher proportion of hydrophobic residues or ion pairs. Halophilic organisms, which must maintain high levels of intracellular ions such as potassium, have proteins whose sequences exhibit a strong preference for surface-exposed acidic residues, particularly aspartate. Proteins from psychrophilic organisms have evolved to maintain

appropriate levels of dynamic motion even in the cold. They tend to have fewer stabilizing interactions, more hydrophobic residues on their surfaces, more overall negative charge, smaller amino acids and additional polar surface loops. Here, we discuss examples of biomolecular adaptations to what can be argued are the most stressful conditions imposed by extreme environments on living organisms (and their associated biomolecules): temperature, pressure, and composition (e.g., pH and salinity).

Temperature adaptations Biomolecules obtained from thermophiles are by far the best characterized (*e.g.* (55)). Thermophilic enzymes exhibit increased global stability of their native, active states relative to their unfolded, inactive states, ΔG_u, making them useful in biotechnological applications (56, 57). This increased stability can arise from one of three thermodynamic mechanisms (58) (Figure 4): Model I) a decreased entropy change of unfolding, ΔS_u, with no change in either the enthalpy of unfolding, ΔH_u, at the melting temperature, Tm, or in the heat capacity of unfolding ΔC_{pu}, Model II) a smaller change in ΔCp_u or Model III) constant ΔC_{pu}, with a decrease in both ΔH_u and ΔS_u, with that in ΔS_u being more pronounced. Theoretical studies suggest that entropic stabilization may be the most common mechanism for stabilization at high temperature (59).

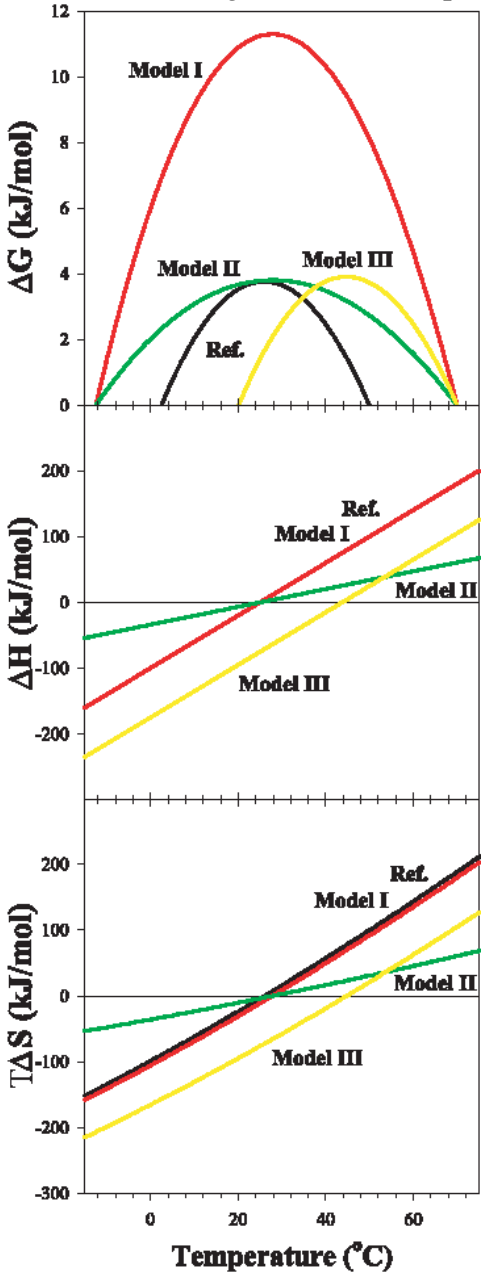

Figure 4. Thermodynamic basis for the temperature stabilization of proteins. Taken from (58). A, the dependence of Gibbs energy of unfolding on temperature; B, the dependence of enthalpy on temperature; C, the dependence of entropic contribution on temperature. Models correspond to those discussed in the text. Ref refers to mesophilic protein thermodynamic stability.

One of the most well-known of such enzymes is the DNA polymerase from *Thermus aquaticus* (Taq), a thermophilic bacterium isolated from a hot spring (60, 61). Taq polymerase, as it is nicknamed, is the enzyme used In the Polymerase Chain Reaction (PCR). The activity profile for Taq exhibits a maximum at 80 °C, allowing it to withstand the multiple heating/annealing steps used in PCR. The increased stability of Taq polymerase has been attributed to a much smaller entropic penalty for folding (62). Thus, although the stabilizing interactions in Taq are decreased (smaller ΔH_u) with respect to its mesophilic homologue from *Escherichia coli*, the entropic penalty is reduced even further, leading to increased stability at all temperatures, and a shift in the Tm (Model III in Figure Y). The difference in the entropic cost of folding likely arises from differences in either the configurational entropy or exposed surface area of the unfolded state.

Pressure adaptations Pressure is known to modulate protein hydration (63), destabilize tertiary interactions (64), decrease dynamics (65) and dissociate oligomers (66, 67) because the more disrupted states exhibit a smaller molar volume, largely due to the elimination of internal void volume (68, 69). Global signatures of pressure adaptation have been more difficult to identify, in part due to the temperature differences across high-pressure environments. Certain classes of proteins appear to undergo stronger adaptive changes in piezophilic organisms. Campanaro and co-workers (70) found that genes whose products were

involved in motility, transport, and DNA synthesis were preferentially modified in two piezophiles, compared to their shallow water counterparts. Interestingly, for those proteins for which 3D structures were available, the substitutions mapped primarily to the surfaces of the proteins. The prevalence of adaptive substitutions on protein surfaces suggests that selection may act to adapt protein interactions given the hydration, ionic ad metabolite environment. Pressure adaptation of a few specific enzymes from deep sea organisms, dihydrofolate reductase (DHFR) (71, 72) and lactate and malate dehydrogenases (LDH and MDH, respectively) (73), revealed changes in enzymatic mechanisms, enhanced enzyme activity at moderate pressures and the over-representation of pressure-adaptive sites on the proteins' surfaces, flexible regions and at subunit interfaces. Machine-learning algorithms identified differences in amino acid composition between piezophilic and non-piezophilic organisms when divided into three different temperature classes, psychrophilic, mesophilic and thermophilic organisms (74). Polar, hydrophilic and small amino acids were over-represented in psychrophilic piezophiles, whereas in mesophilic and thermophilic piezophiles hydrophobic, non-polar and aliphatic amino acids were more prevalent. Substitutions toward smaller amino acids in psychrophiles have been interpreted as providing for

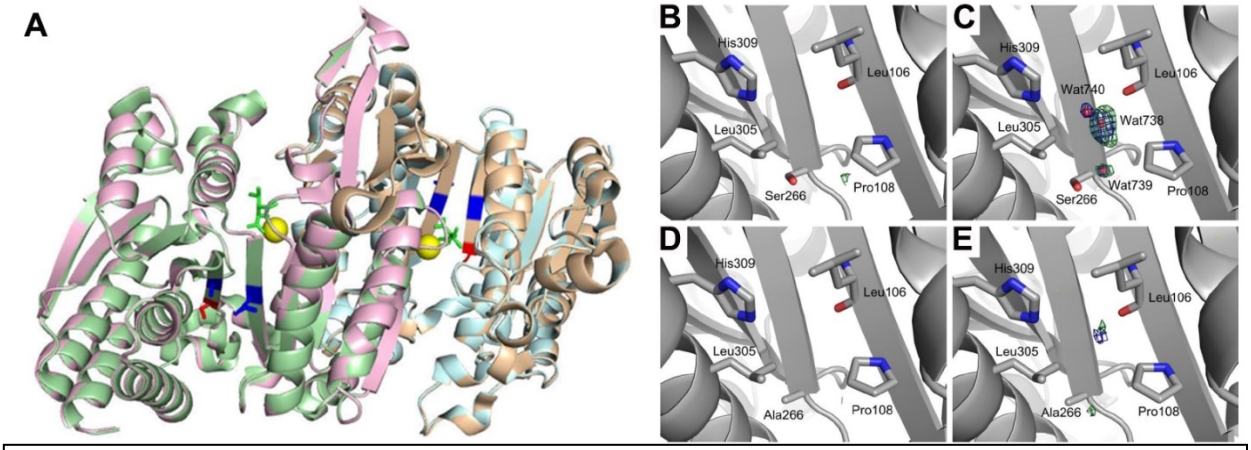

Figure 5. **Proposed structural basis for piezophilic character of IMPDH from *Shewanella benthica* DB21MT-2.** Taken from (70). A) Structures of DHFR from *Shewanella oneidensis* (SoIMPDH) and *Shewanella benthica* DB21MT-2 overlaid with the a and b subunits of soIPMDH in pink and brown and those of SpIPMDH in green and cyan. The IPM molecules are shown in green stick and the magnesium ions in yellow spheres. The single amino acid substitution that confers high-pressure tolerance is S266 in SoIPMDH, and is shown in red stick. B-E) Zoom of the vicinity of residue 266 for soIPMDH at atmospheric pressure (B) and 580 MPa (C) and for soIPMDH at atmospheric pressure (D) and at 580 MPa. Several water molecules are visible in the high pressure structure of SoIPMDH which are not present in that of SpIPMDH.

the required flexibility at low temperature by adding voids.

The subtlety of pressure adaptation is well-illustrated by comparison of the pressure-dependent enzymatic activity, sequence and structure of two isopropyl malate dehydrogenases (IMPDH), one from the atmospheric pressure-adapted *Shewanella oneidensis* (SoIMPDH) and the other from *Shewanella benthica* DB21MT-2 isolated from the Challenger Deep in the Mariana Trench at a depth of nearly 11 km (SbIPMDH)(75–77). Pressure adaptation was attributed to a single amino acid substitution at position 266, an alanine in the piezophile and a serine in the mesophile (Figure 5A). The rate-limiting step of the reaction for an IPMPDH corresponds to a hinge motion, domain-opening conformational change on the backside of the active site required for product release. No water molecules were present near position 266 in structures of either enzyme at 0.1 MPa (Figure 5B, D). In contrast, water was present in the structure of the low-pressure adapted enzyme at 580 MPa (Figure 5C), whereas this is not the case in the high pressure adapted SbIPMDH (Figure 5E). The authors postulate that the serine hydroxyl group in the mesophilic enzyme can form a hydrogen bond with the water molecules present at high pressure, stabilizing the closed form of the enzyme and inhibiting the rate-limiting step. Since the alanine residue in the enzyme from the extreme piezophile cannot make such stabilizing interactions with

the solvent, pressure does not inhibit of the dynamically dependent activity. Indeed, a smaller residue such as alanine at this position could enhance the conformational dynamics as well.

Chemical adaptations Many organisms are also able to live under extreme conditions of pH or salinity. In the case of pH, the cell interior generally remains near neutral pH (78), requiring proton pumps to maintain homeostasis. For acidophiles, this provides an easy access to a proton gradient for energy production (ATP synthesis), whereas alkaliphiles have developed mechanisms to enrich protons near their surface, such teichuronopeptides in cell walls or negatively charged polyglutamic acid surface peptides to lower the free energy for protons near the cell surface (79). Instead of proton pumps, antiporters that correlate ion pumping with reverse proton pumping also enhance H^+ concentration (80) in alkaliphiles.

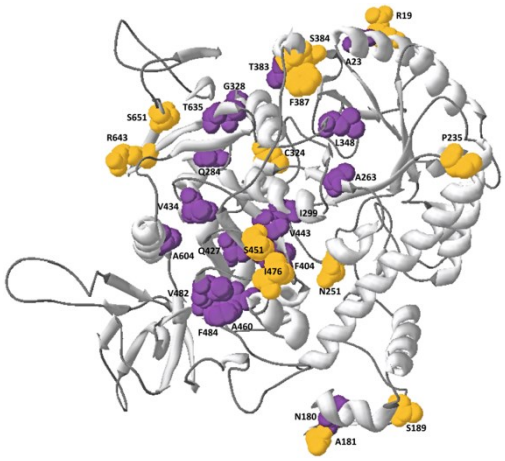

Figure 6. Substitutions for the β-galactosidase from a halophilic organism. Surface residues (orange) and interior residues (purple) substituted in *H. lactusprofundi* relative to non-halophilic organisms, highlighting a complex network of substitutions responsible for the wide range of temperature and salinity and solvents in which this enzyme remains functional.

In addition to extremes in the concentrations of protons and hydroxide, high concentrations of ions such as sodium, potassium, and magnesium can also represent a source of stress for cells. Organisms that thrive under such saline conditions are termed halophiles. The halophilic archaeon, *Halorhabdus utahensis*, from the Great Salt Lake can grow at up to 30% NaCl, and grows optimally at 27% (81). Unlike pH resistance microorganisms, halophiles have adapted to either pump out ions, or by 'letting it all in' and thereby adapting their biomolecules to be tolerant of high ionic strength inside the cell. Proteins from organisms that 'let it all in' tend to have acidic surface residues (negative charge) that enhance hydration and reduce aggregation propensity. Notably most mesophilic proteins will salt out under such conditions. In addition, hydrophobic patches in proteins may be reduced (82). Organisms with high salinity cytoplasms are generally not tolerant to fluctuations in salinity, whereas those in unstable environments (e.g. estuaries) tend to maintain moderate cytoplasmic salinities closer to 200 mM. As in the case of pressure, the devil is in the details for these enzymes. For example, β-galactosidase from the Antarctic archaeon *Halorubrum lacusprofundi*, an enzyme that breaks milk sugar into monosaccharides, is functional over a wide temperature range of - 5 to 70 °C as well as in up to 5 M chloride and aqueous solutions of ethanol. It turns out a set of specific interior and surface mutations is responsible for this enhanced stability and function (83) (Figure 6).

The above commentary is largely framed in the context of protein stability. It is important to note that the free energy governing complex functions such as allostery are also inherently connected to the entire ensemble of states (84). Thus, the impact of extreme environments on protein stability, folding, function and evolution is likely to be complex and remains largely unexplored. Indeed, while the above examples of the structure-function relationships of enzymes from extremophilic organisms provide interesting biophysical insights, a truly systematic biochemical/biophysical picture of enzymes across broad temperature, pressure and compositional gradients is missing. Ideally, properly annotated metagenomics datasets, rather than cultivar genomes, provide a more global view of organisms that are best adapted to given environments and what proteins are most abundant, and hence, important to their growth and survival. Properly exploited, the sequences in these datasets harbor the clues to understanding the molecular basis for extreme life. Such systematic studies of these gene products would allow examination of the biophysical consequences of expansion beyond the temperature, pressure, pH or salt concentration limits to their function.

Advances in Instrumentation for Extreme Biophysics

Extreme biophysics requires sophisticated instrumentation capable of handling high temperature, pressure or other conditions that impose engineering challenges, and in some cases rapidly adjusting those variables over a wide range. Here we discuss some of the technology developed in the areas of magnetic resonance techniques, fluorescence and microscopy techniques, thermodynamic measurements and scattering techniques. The focus is on recent advances with application to complex biomolecules.

Magnetic resonance Owing to its prominence as a fundamental variable in thermodynamics, adaptations of high pressure in NMR spectroscopy began shortly after the demonstration of magnetic resonance itself. The "autoclave" design of the NMR probe, which entails placing the sample and the probe electronics under pressure within a self-contained metal housing, arose as the dominant design strategy (85). Nearly 40 years later, this approach would be used to study the pressure-induced unfolding of proteins (86). For modern multi-dimensional NMR, the viability of the autoclave probe design made little headway in the context of proteins and other biopolymers (86).

An obvious alternate strategy was to limit the pressurized region to the sample itself through the construction of a pressure-tolerant NMR cell to avoid subjecting delicate components of the modern NMR probe to pressure. Early approaches employed quartz capillaries that take advantage of the scaling of pressure tolerance with a reduced inner diameter (87). Impressive pressures could be stably attained. However, the inherent insensitivity of nuclear spin spectroscopy combined with the small active volume of capillary cells results in limited signal-to-noise that severely restricts the experiments that can be undertaken. Joining the NMR sample cell to a high-pressure generator was a major challenge (88), but a novel joining valve to mate to a multi-stage sealing pressure flange permitted the first high resolution triple-resonance protein NMR experiments at 100 MPa (89). Single crystal sapphire tubes are brittle and therefore risky inside expensive NMR instrumentation, whereas advanced zirconia ceramics were subsequently developed to meet the requirements of high-resolution NMR spectroscopy (90). Commercially available alumina-toughened zirconia tubes now provide routine, stable and safe access to pressures up to 300 MPa and over a temperature range of -15 to 115 °C at sample volumes comparable to conventional tubes.

The development of robust, reliable and efficient apparatus for variable high-pressure state-of-the-art NMR spectroscopy of biopolymers has enabled a host of applications in biophysics (91). For example, pressure-induced chemical shift perturbation has permitted detailed analysis of the internal cooperativity of protein structure (92). The pressure dependence of fast internal motion of the protein backbone and both methyl- and aromatic-bearing amino acid side chains have revealed novel insights into the local compressibility (93) and the liquid-like nature of the protein interior. Thus, the full palette of modern solution NMR spectroscopy of biopolymers can be employed in a variable pressure context up to 300 MPa (94). Most recently, rapid pressure transition strategies have enabled pressure-jump experiments that further expand entry into the arena of extreme biophysics of biopolymers (95, 96).

Electron paramagnetic resonance has also been adapted to high pressure samples. Owing to its inherently high sensitivity, the diameter of the EPR sample cell can be kept small to avoid excessive dielectric heating. The small sample cell diameter combined with the relatively open structure of the EPR instrument allows the capillary approach to be employed to great advantage. A "winding" capillary strategy developed for NMR by Yonker and single capillary joined to a metal housing similar to the NMR high pressure described above have been adapted for EPR (97). A broad temperature range is accessible with this apparatus up to 240 MPa and potentially beyond.

Microscopy and optical techniques Pressure (speed of sound) and heat (thermal diffusion) equilibrate quickly over micrometer length scales, and thus the interior of extremophile cells is generally directly subject to these environmental variables. On the other hand, chemical variables such as pH are controlled by molecular diffusion, and the interior even of extremophile cells is

usually near neutral pH and at sub-molar ionic strength thanks to the activity of proton and ion pumps (98). Nonetheless, studying the latter parameters inside extreme organisms reveals that they have evolved strong mechanisms to maintain homeostasis. For example, acidophiles have a net positive charge in the cytoplasm, contributing to a very different interaction environment for proteins than in mammalian cells.

High pressure spectroscopy and microscopy experiments have made significant progress in both *in vitro* and *in vivo* applications over the last decade (99). Commercially available equipment such as computer-controlled high pressure pumps (e.g. Pressure BioSciences) and high pressure cells (e.g. ISI), connected by small diameter tubing allows implementation of a variety of optical and microscopy geometries under computer control. Recent high pressure chamber designs allow microscopy to be carried out near (~0.5 μm) the diffraction limit in dark field or epifluorescence modes up to 300 MPa (100, 101).

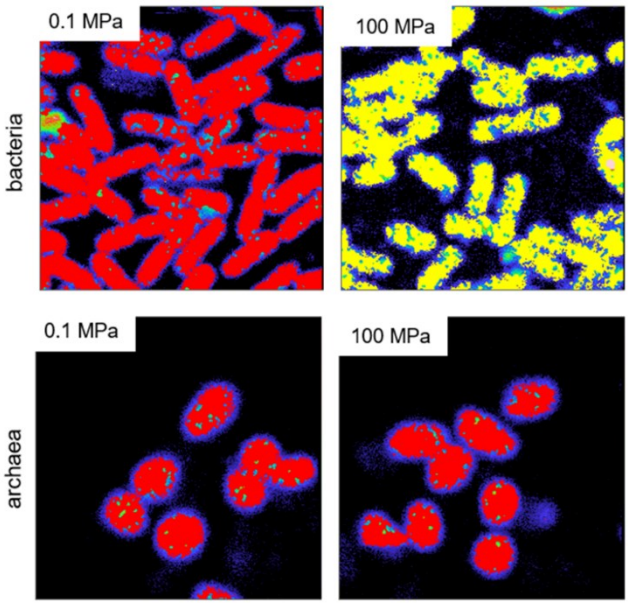

Figure 7. High resolution time-resolved microscopy of live microbial cells at high pressure. High pressure measurements in bacteria (A) vs. archaea (B) reveal that the NADPH bound-unbound ratio (color coded pink to yellow) is very sensitive to pressure in bacteria, but not in archaea. Adapted from ref. (105) with permission.

A variety of fast pressure-jump techniques have been developed to monitor relaxation kinetics optically. Modern stopped flow implementations allow both a wide temperature range (e.g. -40 to 100 °C) and pressure range (up to 200 MPa), without the need for probe light for fluorescence or absorption detection to pass through the pressurization fluid (102). The fastest possible pressure drop of up to 250 MPa in under 1 microsecond is currently implemented by an instrument that bursts a miniature membrane with high current (103). The method has been applied to monitor very fast-folding proteins (104, 105).

Recently measurements in live bacterial cells have become possible. Bourges *et al.* reported cell imaging up to 100 MPa by immobilizing bacteria in chitosan in a square cross-section capillary to reduce optical aberration on the microscope (106). They were able to implement fluorescence lifetime imaging as well as fluctuation spectroscopies in a single system, and measure metabolic response via bound and unbound fluorescent enzyme co-factors. Chen *et al.* developed FRET (Förster Resonant Energy Transfer) detection of protein thermodynamics in live bacteria as a function of temperature and pressure (98). They found that destabilizing surface interactions in the bacterial cytoplasm overcome crowding under thermal denaturation, but not under pressure denaturation. This suggests that many of these destabilizing interactions may be hydrophobic in nature because hydrophobic interactions increase strongly with temperature, but not with pressure. Lower resolution fluorescence microscopy on live cells is possible in the range up to 700 MPa now, and most recently, such instrumentation has been extended to monitor single dye molecule binding to DNA, indicating that even single molecule studies will be possible in harsh environments (107, 108). In-cell laser temperature jumps have also become possible, and can easily be used to induce temperature changes of 50 °C (109).

As for spectroscopic approaches, Circular dichroism (CD) under pressure, useful for measuring protein secondary structure, deserves special mention because of the added challenge of optical distortion perturbing the very small (~1/1000 of absorption) CD signal. Synthetic diamond windows provide distortion-free signals down to 230 nm into the UV (110), and YAG

(Yttrium-Aluminum-Garnet) windows have enabled measurements up to 400 MPa and in presence of harsh solvents (111). CD has also been combined with spin-labels (EPR) to follow pressure denaturation in proteins (112).

Although pH is maintained near neutral even in extremophiles, studying its variation between organisms and different compartments offers clues as to how extremophiles manage proton pumping. Recent approaches have enabled ratiometric (for reliability) (113) and genetically encoded (for *in situ* simplicity) pH probes, such as using a chimeric protein (114). The relaxation time of pH upon sudden external pH jump was found to be about 20 minutes for the fungus, *Aspergillus niger*. Likewise, other chemical stresses, such as salt fluctuations around *Halobacterium salinarium* can be studied by time-lapse microscopy, revealing changes in cell shape (115). Ion-sensitive dye probes have been developed to monitor relative concentrations of multiple ions inside extremophiles (116). Such experiments highlight perhaps the most extreme of extreme: organisms that not only live in extreme conditions, but that can survive large fluctuations in such conditions.

Single particle techniques such as electron microscopy remain to be adapted to extreme conditions such as high pressure. It has been shown recently that samples encapsulated in carbon nanotubes can be subject to high pressure during electron microscopy (117). Similar technology adapted to graphene sheets can hold water pockets, and could be implemented to study extreme conditions by electron microscopy.

Thermodynamic measurements Any spectroscopic method that allows determination of populations of different states can be used to extract thermodynamic information by using standard thermodynamic relations between populations of states and the equilibrium constant (see e.g., (118)). The equilibrium constant allows estimates of Gibbs free energy, ΔG, changes that describe the thermodynamic stability under given set of conditions: temperature, pressure, pH, ionic strength, concentration of co-solute. The dependence of ΔG on temperature (see Fig. 4) leads to estimates of enthalpy, ΔH, and entropy, ΔS, while the dependence of ΔG on pressure defines the volume changes in a system, ΔV. Two experimental methods stand apart because they can provide direct measurements of these parameters.

Differential scanning calorimetry (DSC) allows direct measurements of enthalpy of conformational transitions and has been used widely for comparative studies of thermodynamic properties of proteins from non-extremophilic and extremophilic organisms. In particular, the comparison between mesophiles, psychrophiles and thermophiles have revealed the thermodynamic basis for higher stabilities of thermophilic proteins (59, 119, 120) (see Figure 4). Some of these rules have been further validated experimentally through mutagenesis and protein engineering approaches (119, 121, 122), including for a protein involved in copper tolerance, CutA1, from *Pyrococcus horikoshii* that denatures near 150°C (123). Furthermore, comparative bioinformatics analysis of sequences from the genomes of mesophilic and thermophilic organisms (124, 125) and comparison of structures of mesophilic-thermophilic protein pairs provided further support to these experimentally derived principles (121, 122, 126).

Pressure perturbation calorimetry (PPC) is relatively new method (127–129), which allows direct measurements of volume changes upon conformational transitions at the transition temperature. The sign and magnitude of ΔV dictates the response of the system to the changes in hydrostatic pressure: systems with negative ΔV are destabilized by an increase in pressure, while positive ΔV will result in an increase in stability as the pressure increases. PPC has been successfully applied to study proteins (69, 128–130), nucleic acids (131, 132) and even lipid bilayers (133). It was found that ΔV values are relatively small and can be both positive and negative at ambient temperature (128, 129, 134, 135). As temperature increases, the ΔV values increase due to the effects of hydration on the net volume changes in the system (135). Interestingly, addition of co-solvent does not have a direct effect on the values of ΔV, but can modulate pressure tolerance by increasing the net (thermo)stability (130). Computational analysis of the structures of proteomes inferred for piezophilic and non-piezophilic organisms

showed no difference in ΔV values for proteins from piezophilic and non-piezophilic organisms, an observation that is supported by the direct experimental measurements of ΔV of unfolding using PPC (69).

Scattering techniques Protein structure and oligomeric state are sensitive to extreme conditions. Two principal types of X-ray scattering are widely used in structural characterizations: macromolecular crystallography (MX) and small-angle X-ray solution scattering (SAXS). The corresponding techniques for neutrons are known as neutron crystallography (NC) and small-angle neutron solution scattering (SANS). Macromolecular crystallography remains the gold-standard for obtaining high-resolution biomolecular structure determination, but application to extreme conditions, in particular high pressure, has been technically challenging (136, 137). Sample enclosures with sufficiently thick walls to withstand high pressures also tend to absorb X-rays and produce high background scattering levels that obscure weak sample signals.

The vast majority of ambient-pressure MX today is performed with flash-cooled samples. At temperatures of 100 K, X-ray damage on biological samples is greatly reduced and complete rotation datasets can be obtained from single crystals as small as a few microns in diameter. High pressure cryo-crystallography is an important tool for high resolution structures of enzyme-substrate interactions in extremophilic proteins. Pressure cryo-cooling equipment for crystallography has been commercialized (Advanced Design Consulting, Lansing, NY) and devices exist at several user facilities (e.g. CHESS, UNIST and ESRF). Room temperature crystallography is experiencing a renaissance in the structural biology community. Increasing evidence suggests that cryo-cooling alters the observed conformational heterogeneity of proteins (138, 139). Synchrotron facilities worldwide have developed a strategy, referred to as serial crystallography, for determining non-frozen structures by rapidly collecting data on multiple crystals. Given that pressure-driven conformational changes are similar in magnitude to temperature effects, the tools and technologies emerging from multi-temperature/serial crystallography will likely prove useful to high-pressure structural biology.

Most recent high-pressure structures (HP-MX) have been solved at room temperature using the diamond anvil cell (DAC) (140). While biomolecules in HP-MX are confined to a crystal lattice, the numerous small-scale (e.g., sub-Ångstrom) conformational shifts observed can be derived from structures at 2 Å resolution, providing atomic-level insight into the pressure sensitivity of enzymes (141, 142). Much work remains to be done in field of HP-MX/NC. DACs are generally designed for pressures far beyond those of interest to biologists, at the expense of precision. However, the recent innovation of pneumatic control allows more precise regulation of biologically relevant pressures.

High-resolution structures from crystalline and frozen states give, at best, only snapshots of what happens to molecules under realistic physiological conditions. Small angle solution scattering of X-rays (SAXS) and neutrons (SANS), yields valuable structural information by sampling the conformational ensembles present in solution (143). While not as information-rich as crystallography, these techniques are applicable to a wide range of extreme conditions, including temperature, ionic strength, pH, and pressure. SAX(N)S is widely used to quantify flexibility, disorder and unfolding in biomolecules, to determine stability limits, to identify oligomeric states and conformational changes and to validate atomistic models. While X-ray and neutron scattering techniques have revolutionized structural biology, structural information on biomolecules under pressure has been scarce due to experimental difficulty and lack of access to specialized equipment and expertise. This situation is changing with recent innovations, commercial availability of equipment, and support at national X-ray and neutron facilities.

HP-SAXS/SANS cell designs depart from the beryllium rod and DAC designs of crystallography in favor of flat windows of various materials (144–149). The challenge with HP-SAXS systems is the cumbersome process of sample change, which is required for the essential buffer background subtraction. Even small movements of the sample cell during sample change must be minimized to achieve reproducibility (148). Building upon an earlier design by Ando

(144), Rai et al. present a next-generation HP-SAXS cell designed for rapid, reproducible sample change at pressures of up to 700 MPa (Rai et al, submitted for publication). HP-SAXS is particularly well-suited to detecting the complete or partial unfolding of proteins that can result from high pressure (150, 151). Pressure-induced changes in protein-protein interaction potentials have been quantified in concentrated protein solutions using this technique. The method has been applied to the study of osmolytes related to high-pressure adaptation of organisms (152, 153), as well as dissociation of oligomers (146).

Future directions and roadblocks to progress

Massive and growing genomic information represents intrinsically molecular information for proteins that allow life to persist across gradients in temperature, pH, salinity, and pressure. Within these sequences lie the clues to the molecular limits of life. Pursuing these clues, however, is not straightforward due to the massive number of individual proteins they represent. In other words, it is not feasible to define the usually subtle sequence-structure-function-environmental relationships for a representative protein homolog from even a significant fraction of these functionally distinct protein groups. How then can we prioritize proteins for biophysical characterization?

As outlined above, a working definition for extremophiles that inhabit environments with extreme temperature, pressure, pH or composition is that they have evolved under chronic energy stress. It seems appropriate then to focus attention on the enzymes implicated in the key processes that sustain life in these environments in which energy demands are significant compared to energy sources, as these proteins would presumably be under strong evolutionary pressure, and hence hold the most easily identifiable clues to the molecular limits of life. While many biochemical processes are important, we can define two major classes of biochemical processes that are central to sustaining life: 1) Metabolism and bioenergetics, and 2) Genome function and repair.

Metabolism and bioenergetics The mechanisms by which energy is obtained from the environment share strong functional and structural similarities across all of life. This indicates that the sequences of metabolic and bio-energetic enzymes are highly constrained by their function. Differences between the structure-function determinants of such enzymes from mesophiles and various classes of extremophiles are likely to reveal important adaptations to the extreme environments.

In many extreme environments, particularly in energy-limited subsurface ecosystems, microbial inhabitants are supported by chemolithotrophic metabolisms (i.e., chemical and inorganic source of energy). Examples include H_2-dependent and autotrophic methanogens and hydrogenotropic sulfate reducing microbes (SRMs), which can use small organic compounds, if available, in heterotrophic growth. As such, the energy metabolism (electron donating half) of these cells is dependent on one of several hydrogenase enzymes that reversibly activate H_2, including [NiFe]- and [FeFe]-hydrogenases (154). In the case of SRMs, the electron accepting half of the cells metabolism is dependent on dissimilatory sulfite reductase (Dsr), which catalyzes the reduction of (bi)sulfite that is available in the cells environment or that is produced by reduction of sulfate (155). For autotrophic methanogens, the electron acceptor is CO_2 and the final step in the reduction of CO_2 to CH_4 is catalyzed by methyl-coenzyme M reductase (Mcr) (156). Ion gradients are established across the cytoplasmic membrane by energy coupled membrane pumps (156, 157), that can then be exploited for ATP production by ATP synthase. Often times, electron transfer reactions, including those used to generate ion gradients in SRMs and methanogens, are linked by iron sulfur cluster containing ferredoxin proteins. While the enzymes implicated in sulfate reduction and methanogenesis can be multi-subunit complexes, several tend to be soluble thereby facilitating biochemical and biophysical study. Indeed, structures are available for three Dsrs (158–160), a variety of hydrogenases (*e.g.,* (161)), several Mcrs (*e.g.,* (162)) and Fd (36

entries in the Protein Data Bank). Structures are also available for the ATP synthase from *T. thermophilus* (*e.g.,* (163)). These proteins represent amenable and fundamentally important targets to systematically probe the relationships between sequence and environment and the biophysical and molecular basis that allows for diversification to the extremes of habitable space.

Another family of key metabolic proteins is the nitrogenase-like family of enzymes. Nitrogenases are responsible for the production of much of the fixed nitrogen on Earth ((164) and references therein). They are members of a much larger family of nitrogenase-like enzymes that are thought to have originated via gene duplication from one proto-nitrogenase-like ancestor. At least three major biochemical processes with evolutionary significance are associated with this family: nitrogen fixation, bacteriochlorophyll biosynthesis, and Ni-containing co-factor, F430, biosynthesis (165). The catalytic subunits, called NifDK, have homologs among nitrogenase metallocofactor-maturation proteins, catalytic subunits of dark-operative protochlorophyllide oxidoreductase/chlorophyllide oxidoreductases (DPOR - enzymes with important roles in anoxic and oxygenic photosynthesis), and enzymes involved in biosynthesis of F430.

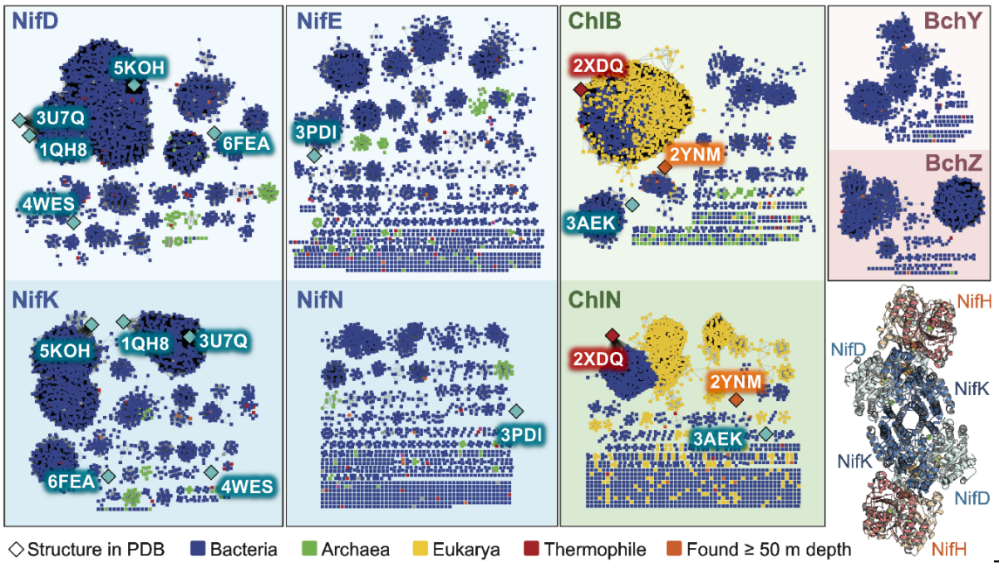

Figure 8. Sequence similarity network (SSN) for nitrogenase component 1 type oxidoreductase (pfam00148, May 2020) (165). Each square corresponds to one enzyme, and connecting edges represent sequence similarity above or equal to a threshold cutoff of an E-value of 1 x 10^{-230} (corresponding to 62% identity). Individual enzyme nodes are colored by super-kingdom, with enzymes from thermophilic organisms highlighted in red and enzymes from organisms found at greater than or equal to 50 m depth highlighted in orange. Enlarged diamonds are enzymes with deposited structures in the PDB, labeled by accession code of the highest-resolution structure available. The more distantly related CfbD subfamily, although a nitrogenase-like oxidoreductase, is not shown. Nitrogenase crystal structure (PDB: 1M34, *Av*) depicts catalytic NifDK heterotetramer with two ATPase NifH homodimers coordinated at opposing ends. NifDK is homologouse to NifEN, which is responsible for nitrogenase cofactor maturation. The catalytic subunits for DPOR and COR are ChlBN and BchYZ.

Sequence similarity networks (SSNs) annotated with environmental metadata provide useful roadmaps for identifying sequences from extremophile enzymes of interest (166). In Figure 8, the SSN for nitrogenase component 1 type oxidoreductase subunits reveal a cluster of similar sequences for the DPOR subunit ChlB from organisms found at depth (orange squares clustering with PDB: 2YNM), whereas no significant clusters of thermophilic sequences (red squares) are observed. Moreover, although the nitrogenase catalytic subunits (NifDK) are homologous to the FeMo cofactor maturation protein (NifEN), greater sequence conservation is observed within the catalytic subunits. In addition to the DPOR complex from the pressure-adapted organism, *Prochlorococcus marinus* (PDB 2YNM), a crystal structure is available for the DPOR catalytic component from *Thermosynechococcus elongatus*, a thermophile (PDB 2XDQ). However, large swaths of the nitrogenase-like enzyme family are completely uncharacterized structurally, including the entire COR family. Within the nitrogenase system, analysis of mutations across SSN clusters reveals that a highly conserved glycine in the Mo-Fe nitrogenases (residue 69 in

Azotobacter vinelandii NifD numbering) is mutated in SSN clusters with archaeal and thermophilic content, including those associated with V-bearing nitrogenases. In all nitrogenase-like enzymes, catalysis requires long-range electron transfer through transient complex formation with a separate ATP-dependent Fe protein. Detailed examination of sequences reveals that in *A. vinelandii* NifD, all of the residues that form this interface are charged, whereas in SSN clusters with high thermophile content, a handful of mutations to uncharged or hydrophobic residues are observed, in addition to deletions. With recent advances in both computation and structural techniques, a bioinformatics-guided structural approach can provide insight into the sequence modifications that enable adaptation to extreme environments as well as fundamental relationships between primary sequence, tertiary structure, and catalytic function.

Genome function and repair Enzymes implicated in genome function are also considered as key targets for understanding the molecular limits of life. The largest fraction of genes associated with pressure sensitivity in randomly generated mutants of the piezophile, *P. profundum* SS9, were associated with chromosome structure and function, as well as ribosome assembly (42). A strain of *E. coli* adapted to grow at 62 MPa, AN62 (167), bears very few single site mutations (Bartlett and co-workers, unpublished), one of which is in the β subunit of RNA polymerase (RNAP); two are in transcriptional repressors and one is in the transcription anti-terminator, Rho. An unprecedented number of viruses are present in the extreme environments, along with a large number of sequences for DNA modifying and repair enzymes in the genomes of the microorganisms in these ecosystems. In addition to performing central functions for sustaining life, they also participate in the evolution of extremophiles, and represent interesting avenues for biotechnological applications. RNA and DNA polymerases, gyrases, nucleases, transcription and translation factors have long been targets of biophysical, biochemical and structural investigations of their molecular mechanisms. Building on this wealth of knowledge with comparative studies across temperature, pressure, pH or compositional gradients would yield significant insight into molecular adaptation of these important enzymes.

Extreme Dark Matter Finally, a major roadblock to interpreting the available and ever-increasing genomic data from extreme environments, and hence to defining and understanding the molecular limits of life, is that at least at the level of simple analyses, the functions of well over half of the extremophile gene sequences produced by (meta)genomic sequencing of cultivars or environmental samples cannot be deduced (168). The abundance of Dark Matter is particularly pronounced among extremophiles (169). Among the sequences that make up this biological Dark Matter are some of the most abundant proteins across all of biology. As such, these functionally unannotated proteins are likely to play essential roles in sustaining life under extreme conditions. Tackling this massive gap in our understanding of the functional biodiversity on Earth represents a truly grand and largely multidisciplinary challenge. Addressing it will require discovery-based research that integrates across advanced structural and functional bioinformatics and modeling, database development with clear and uniformly reported metadata, and the development of high throughput targeted functional screening approaches both *in vitro* and *in vivo*. Making headway in addressing this grand challenge will require additional advances in computational infrastructure, structural modeling approaches, and biophysical/biochemical instrumentation. Nonetheless, assigning function to this Dark Matter and studying such proteins from the structure-function perspective may provide just the insights that are needed to understand the molecular basis for extreme life, and to harness it for applications ranging from human health to energy production.

Acknowledgments

This work was supported by the Extreme Biophysics Research Coordination Network grant MCB 1817845 from the National Science Foundation to CAR.

Bibliography

1. McMahon, S., and J. Parnell. 2014. Weighing the deep continental biosphere. FEMS Microbiol Ecol. 87: 113–120.
2. Bar-On, Y.M., R. Phillips, and R. Milo. 2018. The biomass distribution on Earth. Proc. Natl. Acad. Sci. 115: 6506–6511.
3. Norton, C., and W. Grant. 1988. Survival of Halobacteria within Fluid Inclusions in Salt Crystals. J. Gen. Microbiol. 134: 1365–1373.
4. Oren, A. 2013. Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. Front. Microbiol. 4: 1–6.
5. Maier, R.M., K.P. Drees, J.W. Neilson, D.A. Henderson, J. Quade, and J.L. Betancourt. 2004. Microbial Life in the Atacama Desert. Science (80-.). 306: 1289–1290.
6. Takai, K., K. Nakamura, T. Toki, U. Tsunogai, M. Miyazaki, J. Miyazaki, H. Hirayama, S. Nakagawa, T. Nunoura, and K. Horikoshi. 2008. Cell proliferation at 122 C and isotopically heavy CH4 production by a hyperthermophillic methanogen under high-presure cultivation. 105: 10949–10954.
7. Schleper, C., G. Puehler, I. Holz, A. Gambacorta, D. Janekovic, U. Santarius, and E. Al. 1995. Picrophilus gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0. J Bacteriol. 117: 7050–7059.
8. Takai, K., D.P. Moser, T.C. Onstott, N. Spoelstra, S.M. Pfiffner, A. Dohnalkova, and J.K. Fredrickson. 2001. Alkaliphilus transvaalensis gen. nov., sp. nov., an extremely alkaliphilic bacterium isolated from a deep South African gold mine. Int. J. Syst. Evol. Microbiol. 51: 1245–1256.
9. Kusube, M., T.S. Kyaw, K. Tanikawa, R.A. Chastain, K.M. Hardy, J. Cameron, and D.H. Bartlett. 2017. Colwellia marinimaniae sp. nov., a hyperpiezophilic species isolated from an amphipod within the challenger deep, Mariana Trench. Int. J. Syst. Evol. Microbiol. 67: 824–831.
10. Zeng, X., J.L. Birrien, Y. Fouquet, G. Cherkashov, M. Jebbar, J. Querellou, P. Oger, M.A. Cambon-Bonavita, X. Xiao, and D. Prieur. 2009. Pyrococcus CH1, an obligate piezophilic hyperthermophile: Extending the upper pressure-temperature limits for life. ISME J. 3: 873–876.
11. Sharma A, Scott JH, Cody GD, Fogel ML, Hazen RM, Hemley RJ, H.W. 2002. Microbial activity at gigapascal pressures. Science (80-.). 295: 1514–6.
12. Vanlint, D., R. Mitchell, E. Bailey, F. Meersman, P.F. McMillan, C.W. Michiels, and A. Aertsen. 2011. Rapid acquisition of gigapascal-high-pressure resistance by Escherichia coli. MBio. 2: 1–3.
13. Winter, R., and C. Jeworrek. 2009. Effect of pressure on membranes. Soft Matter. 5: 3157–3173.
14. Foglia, F., R. Hazael, F. Meersman, M.C. Wilding, V.G. Sakai, S. Rogers, L.E. Bove, M.M. Koza, M. Moulin, M. Haertlein, V.T. Forsyth, and P.F. McMillan. 2019. In Vivo Water Dynamics in Shewanella oneidensis Bacteria at High Pressure. Sci. Rep. 9: 1–11.
15. Baker-Austin, C., and M. Dopson. 2007. Life in acid: pH homeostasis in acidophiles. Trends Microbiol. 15: 165–171.
16. Reed, C.J., H. Lewis, E. Trejo, V. Winston, and C. Evilia. 2013. Protein Adaptations in Archaeal Extremophiles. Archaea. 2013: 373275.
17. Pennisi, E. 1997. Biotechnology: In Industry, Extremophiles Begin to Make Their Mark. Science (80-.). 276: 705–706.
18. Shock, E.L., and M.E. Holland. 2007. Quantitative habitability. Astrobiology. 7: 839–851.
19. D'Hondt, S., S. Rutherford, and A.J. Spivack. 2002. Metabolic activity of subsurface life in deep-sea sediments. Science (80-.). 295: 2067–70.

20. Jorgensen, B.B. 2011. Deep subseafloor microbial cells on physiological standby. Proc Natl Acad Sci U S A. 108: 18193–18194.
21. Starnawski, P., T. Bataillon, T.J.G. Ettema, L.M. Jochum, L. Schreiber, X. Chen, and E. Al. 2017. Microbial community assembly and evolution in subseafloor sediment. Proc Natl Acad Sci U S A. 114: 2940–2945.
22. Chivian, D., E.L. Brodie, E.J. Alm, D.E. Culley, P.S. Dehal, T.Z. DeSantis, T.M. Gihring, A. Lapidus, L.-H. Lin, S.R. Lowry, D.P. Moser, P.M. Richardson, G. Southam, G. Wanger, L.M. Pratt, G.L. Andersen, T.C. Hazen, F.J. Brockman, A.P. Arkin, and T.C. Onstott. 2008. Environmental genomics reveals a single-species ecosystem deep within Earth. Science (80-.). 322: 275–8.
23. Valentine, D. 2007. Adaptations to energy stress dictate the ecology and evolution of the Archaea. Nat Rev Microbiol. 5: 316–323.
24. Martin, W., J. Baross, D. Kelley, and M.J. Russell. 2008. Hydrothermal vents and the origin of life. Nat. Rev. Microbiol. 6: 805–814.
25. Colman, D.R., S. Poudel, T.L. Hamilton, J.R. Havig, M.J. Selensky, E.L. Shock, and E.S. Boyd. 2018. Geobiological feedbacks and the evolution of thermoacidophiles. ISME J. 12: 225–236.
26. Odling-Smee, F.J., K.N. Laland, and M.W. Feldman. 1996. Niche Construction. Am. Nat. 147: 641–648.
27. Macalady, J.L., M.M. Vestling, D. Baumler, N. Boekelheide, C.W. Kaspar, and J.F. Banfield. 2004. Tetraether-linked membrane monolayers in Ferroplasma spp: a key to survival in acid. Extremophiles. 8: 411–419.
28. Fones, E.M., D.R. Colman, E.A. Kraus, A.S. Stepanauskas, R., Templeton, J.R. Spear, and E.S. Boyd. 2020. Diversification of methanogens into hyperalkaline serpentinizing environments through adaptations to minimize oxidant limitation. Under Rev. .
29. Moody, J. 1976. Serpentinization: a review. Lithos. 9: 125–138.
30. Allen, E.E., D. Facciotti, and D.H. Bartlett. 1999. Monounsaturated but not polyunsaturated fatty acids are required for growth of the deep-sea bacterium Photobacterium profundum SS9 at high pressure and low temperature. Appl. Environ. Microbiol. 65: 1710–1720.
31. Cario, A., V. Grossi, P. Schaeffer, and P.M. Oger. 2015. Membrane homeoviscous adaptation in the piezo-hyperthermophilic archaeon Thermococcus barophilus. Front. Microbiol. 6: 1–12.
32. Yancey, P.H. 2020. Cellular responses in marine animals to hydrostatic pressure. J. Exp. Zool. Part A Ecol. Integr. Physiol. 333: 398–420.
33. Coker, J. 2016. Extremophiles and biotechnology: current uses and prospects. F100Res. 5.
34. Lloyd, K.G.G., A.D.D. Steen, J. Ladau, J. Yin, and L. Crosby. 2018. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. mSystems. 3: 00055–18.
35. Colman, D.R., M.R. Lindsay, and E.S. Boyd. 2019. Mixing of meteoric and geothermal fluids supports hyperdiverse chemosynthetic hydrothermal communities. Nat Commun. 10: 681.
36. Hahn, C.J., R. Laso-Perez, F. Vulcano, K.M. Vaziourakis, R. Stokke, I.H. Steen, A. Teske, A. Boetius, M. Liebeke, R. Amann, K. Knittel, and G. Wegener. 2020. “Candidatus Ethanoperedens,” a Thermophilic Genus of Archaea Mediating the Anaerobic Oxidation of Ethane. MBio. 11.
37. Borrel, G., P.S. Adam, L.J. McKay, L.X. Chen, I.N. Sierra-Garcia, C.M.K. Sieber, Q. Letourneur, A. Ghozlane, G.L. Andersen, W.J. Li, S.J. Hallam, G. Muyzer, V.M. de Oliveira, W.P. Inskeep, J.F. Banfield, and S. Gribaldo. 2019. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. Nat Microbiol. 4: 603–613.
38. Colman, D.R., M.R. Lindsay, M.J. Amenabar, and E.S. Boyd. 2019. The Intersection of

Geology, Geochemistry, and Microbiology in Continental Hydrothermal Systems. Astrobiology. 19: 1505–1522.
39. Colman, D.R., S. Poudel, B.W. Stamps, E.S. Boyd, and J.R. Spear. 2017. The deep, hot biosphere: Twenty-five years of retrospection. Proc. Natl. Acad. Sci. U. S. A. 114: 6895–6903.
40. Kieft, T.L. 2016. Microbiology of the Deep Continental Biosphere. In: Hurst CJ, editor. Their World: A Diversity of Microbial Environments. Cham: Springer International Publishing. pp. 225–249.
41. Michoud, G., and M. Jebbar. 2016. High hydrostatic pressure adaptive strategies in an obligate piezophile Pyrococcus yayanosii. Sci. Rep. 6: 1–10.
42. Lauro, F.M., K. Tran, A. Vezzi, N. Vitulo, G. Valle, and D.H. Bartlett. 2008. Large-scale transposon mutagenesis of Photobacterium profundum SS9 reveals new genetic loci important for growth at low temperature and high pressure. J. Bacteriol. 190: 1699–1709.
43. Vezzi, A.A., S. Campanaro, M.D. Angelo, F. Simonato, N. Vitulo, F.M. Lauro, G. Malacrida, B. Simionati, N. Cannata, C. Romualdi, D.H. Bartlett, G. Valle, S. Science, N. Series, and N. Mar. 2005. Life at Depth : Photobacterium profundum Genome Sequence and Expression Analysis. Science (80-.). 307: 1459–1461.
44. Wang, B., W. Qin, Y. Ren, X. Zhou, M.Y. Jung, P. Han, E.A. Eloe-Fadrosh, M. Li, Y. Zheng, L. Lu, X. Yan, J. Ji, Y. Liu, L. Liu, C. Heiner, R. Hall, W. Martens-Habbena, C.W. Herbold, S. keun Rhee, D.H. Bartlett, L. Huang, A.E. Ingalls, M. Wagner, D.A. Stahl, and Z. Jia. 2019. Expansion of Thaumarchaeota habitat range is correlated with horizontal transfer of ATPase operons. ISME J. 13: 3067–3079.
45. Hug, L.A., B.C. Thomas, C.T. Brown, K.R. Frischkorn, K.H. Williams, S.G. Tringe, and J.F. Banfield. 2015. Aquifer environment selects for microbial species cohorts in sediment and groundwater. ISME J. 9: 1846–1856.
46. Parks, D.H., M. Chuvochina, D.W. Waite, C. Rinke, A. Skarshewski, P.A. Chaumeil, and P. Hugenholtz. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 36: 996–1004.
47. Fones, E.M., D.R. Colman, E.A. Kraus, D.B. Nothaft, S. Poudel, K.R. Rempfert, J.R. Spear, A.S. Templeton, and E.S. Boyd. 2019. Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. ISME J. 13: 1750–1762.
48. Suzuki, S., S. Ishii, T. Hoshino, A. Rietze, A. Tenney, P.L. Morrill, F. Inagaki, J.G. Kuenen, and K.H. Nealson. 2017. Unusual metabolic diversity of hyperalkaliphilic microbial communities associated with subterranean serpentinization at The Cedars. ISME J. 11: 2584–2598.
49. Karnachuk, O. V, Y.A. Frank, A.P. Lukina, V. V Kadnikov, A. V Beletsky, A. V Mardanov, and N. V Ravin. 2019. Domestication of previously uncultivated Candidatus Desulforudis audaxviator from a deep aquifer in Siberia sheds light on its physiology and evolution. ISME J. 13: 1947–1959.
50. Brazelton, W.J., C.N. Thornton, A. Hyer, K.I. Twing, A.A. Longino, S.Q. Lang, M.D. Lilley, G.L. Fruh-Green, and M.O. Schrenk. 2017. Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. PeerJ. 5: e2945.
51. Lindsay, M.R., D.R. Colman, M.J. Amenabar, K.E. Fristad, K.M. Fecteau, R. V. Debes, J.R. Spear, E.L. Shock, T.M. Hoehler, and E.S. Boyd. 2019. Probing the geological source and biological fate of hydrogen in Yellowstone hot springs. Environ. Microbiol. 21: 3816–3830.
52. Lim, S.A., E.R. Bolin, and S. Marqusee. 2018. Tracing a protein's folding pathway over evolutionary time using ancestral sequence reconstruction and hydrogen exchange. Elife. 7: 1–19.
53. Brininger, C., S. Spradlin, L. Cobani, and C. Evilia. 2018. The more adaptive to change,

the more likely you are to survive: Protein adaptation in extremophiles. Semin. Cell Dev. Biol. 84: 158–169.
54. Coker, J. 2019. Recent advances in understanding extremophiles. F1000Res. 8: 1917.
55. Atalah, J., P. Cáceres-Moreno, G. Espina, and J.M. Blamey. 2019. Thermophiles and the applications of their enzymes as new biocatalysts. Bioresour. Technol. 280: 478–488.
56. Dumorné, K., D.C. Córdova, M. Astorga-Eló, and P. Renganathan. 2017. Extremozymes: A potential source for industrial applications. J. Microbiol. Biotechnol. 27: 649–659.
57. Graham, J.E., M.E. Clark, D.C. Nadler, S. Huffer, H.A. Chokhawala, S.E. Rowland, H.W. Blanch, D.S. Clark, and F.T. Robb. 2011. Identification and characterization of a multidomain hyperthermophilic cellulase from an archaeal enrichment. Nat. Commun. 2: 375–379.
58. Makhatadze, G.I., V. V Loladze, A. V Gribenko, and M.M. Lopez. 2004. Mechanism of thermostabilization in a designed cold shock protein with optimized surface electrostatic interactions. J Mol Biol. 336: 929–942.
59. Sawle, L., and K. Ghosh. 2011. How do thermophilic proteins and proteomes withstand high temperature? Biophys. J. 101: 217–227.
60. Chien, A., D.B. Edgar, and J.M. Trela. 1976. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. J. Bacteriol. 127: 1550–1557.
61. Brock, T.D., and H. Freeze. 1969. Thermus aquaticus gen. n. and sp. n., a Non-sporulating Extreme Thermophile. J. Bacteriol. 98: 289–297.
62. Liu, C.C., and V.J. LiCata. 2014. The stability of Taq DNA polymerase results from a reduced entropic folding penalty; identification of other thermophilic proteins with similar folding thermodynamics. Proteins Struct. Funct. Bioinforma. 82: 785–793.
63. Collins, M.D., G. Hummer, M.L. Quillin, B.W. Matthews, and S.M. Gruner. 2005. Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. Proc. Natl. Acad. Sci. U. S. A. 102: 16668–71.
64. Silva, J.L., D. Foguel, and C.A. Royer. 2001. Pressure provides new insights into protein folding, dynamics and structure. Trends Biochem. Sci. 26: 612–8.
65. Al-Ayoubi, S.R., P.H. Schummel, M. Golub, J. Peters, and R. Winter. 2017. Influence of cosolvents, self-crowding, temperature and pressure on the sub-nanosecond dynamics and folding stability of lysozyme. Phys. Chem. Chem. Phys. 19: 14230–14237.
66. Gross, M., and R. Jaenicke. 1994. Proteins under pressure. The influence of high hydrostatic pressure on structure, function and assembly of proteins and protein complexes. Eur. J. Biochem. 221: 617–30.
67. Silva, J., and G. Weber. 1993. Pressure Stability of Proteins. Annu. Rev. Phys. Chem. 44: 89–113.
68. Roche, J., J.A. Caro, D.R. Norberto, P. Barthe, C. Roumestand, J.L. Schlessman, A.E. Garcia, B. García-Moreno E., and C.A. Royer. 2012. Cavities determine the pressure unfolding of proteins. Proc. Natl. Acad. Sci. U. S. A. 109: 6945–50.
69. Avagyan, S., D. Vasilchuk, and G.I. Makhatadze. 2020. Protein adaptation to high hydrostatic pressure: Computational analysis of the structural proteome. Proteins Struct. Funct. Bioinforma. 88: 584–592.
70. Campanaro, S., L. Treu, and G. Valle. 2008. Protein evolution in deep sea bacteria: an analysis of amino acids substitution rates. BMC Evol. Biol. 8: 313.
71. Murakami, C., E. Ohmae, S. Tate, K. Gekko, K. Nakasone, and C. Kato. 2011. Comparative study on dihydrofolate reductases from Shewanella species living in deep-sea and ambient atmospheric-pressure environments. Extremophiles. 15: 165–175.
72. Ohmae, E., C. Murakami, S.I. Tate, K. Gekko, K. Hata, K. Akasaka, and C. Kato. 2012. Pressure dependence of activity and stability of dihydrofolate reductases of the deep-sea bacterium Moritella profunda and Escherichia coli. Biochim. Biophys. Acta - Proteins Proteomics. 1824: 511–519.

73. Winnikoff, J.R., W.R. Francis, E. V. Thuesen, and S.H.D. Haddock. 2019. Combing Transcriptomes for Secrets of Deep-Sea Survival: Environmental Diversity Drives Patterns of Protein Evolution. Integr. Comp. Biol. 59: 1–13.
74. Nath, A., and K. Subbiah. 2016. Insights into the molecular basis of piezophilic adaptation: Extraction of piezophilic signatures. J. Theor. Biol. 390: 117–126.
75. Nagae, T., T. Kawamura, L.M.G. Chavas, K. Niwa, M. Hasegawa, C. Kato, and N. Watanabe. 2012. High-pressure-induced water penetration into 3-isopropylmalate dehydrogenase. Acta Crystallogr. Sect. D Biol. Crystallogr. 68: 300–309.
76. Hamajima, Y., T. Nagae, N. Watanabe, E. Ohmae, Y. Kato-Yamada, and C. Kato. 2016. Pressure adaptation of 3-isopropylmalate dehydrogenase from an extremely piezophilic bacterium is attributed to a single amino acid substitution. Extremophiles. 20: 177–186.
77. Kasahara, R., T. Sato, H. Tamegai, and C. Kato. 2009. Piezo-adapted 3-isopropylmalate dehydrogenase of the obligate piezophile Shewanella benthica DB21MT-2 isolated from the 11,000-m depth of the Mariana Trench. Biosci. Biotechnol. Biochem. 73: 2541–3.
78. Siliakus, M., J. van der Oost, and S. Kengen. 2017. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. Extremophiles. 21: 651–670.
79. Preiss, L., D.B. Hicks, S. Suzuki, T. Meier, and T.A. Krulwich. 2015. Alkaliphilic Bacteria with Impact on Industrial Applications, Concepts of Early Life Forms, and Bioenergetics of ATP Synthesis. Front. Bioeng. Biotechnol. 3.
80. Krulwich, T.A., G. Sachs, and E. Padan. 2011. Molecular aspects of bacterial pH sensing and homeostasis. Nat. Rev. Microbiol. 9: 330–343.
81. Wainø, M., B.J. Tindall, and K. Ingvorsen. 2000. Halorhabdus utahensis gen. nov., sp. nov., an aerobic, extremely halophilic member of the Archaea from Great Salt Lake, Utah. Int. J. Syst. Evol. Microbiol. 50: 183–190.
82. Siglioccolo, A., A. Paiardini, M. Piscitelli, and S. Pascarella. 2011. Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. BMC Struct. Biol. 11: 50.
83. DasSarma, S., M.D. Capes, R. Karan, and P. DasSarma. 2013. Amino Acid Substitutions in Cold Adapted Proteins from Halorubrum lacusprofundi, an Extremely Halophilic Microbe from Antarctica. PLoS One. 8: e58587.
84. Motlagh, H.N., J.O. Wrabl, J. Li, and V.J. Hilser. 2014. The ensemble nature of allostery. Nature. 508: 331–9.
85. Benedek, G., and E. Purcell. 1954. Nuclear Magnetic Resonance in Liquids Under High Pressure. J. Chem. Phys. 22: 2003–2012.
86. Jonas, J. 2002. High-resolution nuclear magnetic resonance studies of proteins. Biochem. Biophys. Acta-Protein Struct. Mol. Enzymol. 1595: 145–159.
87. Yamada, H., K. Nishikawa, M. Honda, T. Shimura, K. Akasaka, and K. Tabayashi. 2001. Pressure-resisting cell for high-pressure, high-resolution nuclear magnetic resonance measurements at very high magnetic fields. Rev. Sci. Instrum. 72: 1463.
88. Roe, D. 1985. Sapphire NMR tube for high-resolution studies at elevated pressure. J. Magn. Res. 63: 388–391.
89. Urbauer, J.L., M.R. Eherhardt, R.J. Bieber, P.F. Flynn, and A.J. Wand. 1996. High resolution triple resonance NMR spectroscopy of a novel calmodulin•peptide complex at kilobar pressures. J. Am. Chem. Soc. 118: 11329–11330.
90. Peterson, R.W., and A.J. Wand. 2005. Self-contained high-pressure cell, apparatus, and procedure for the preparation of encapsulated proteins dissolved in low viscosity fluids for nuclear magnetic resonance spectroscopy. Rev. Sci. Instrum. 76: 094101.
91. Caro, J.A., and A.J. Wand. 2018. Practical aspects of high-pressure NMR spectroscopy and its applications in protein biophysics and structural biology. Methods. 148: 67–80.
92. Roche, J., C.A. Royer, and C. Roumestand. 2019. Exploring Protein Conformational Landscapes Using High-Pressure NMR. 1st ed. Elsevier Inc.

93. Fu, Y., V. Kasinath, V.R. Moorman, N. V Nucci, V.J. Hilser, and A.J. Wand. 2012. Coupled motion in proteins revealed by pressure perturbation. J Am Chem Soc. 134: 8543–8550.
94. Kasinath, V., Y. Fu, K.A. Sharp, and A.J. Wand. 2015. A sharp thermal transition of fast Aromatic-ring dynamics in ubiquitin. Angew. Chemie - Int. Ed. 54: 102–107.
95. Kremer, W., M. Arnold, C.E. Munte, R. Hartl, M.B. Erlach, J. Koehler, A. Meier, and H.R. Kalbitzer. 2011. Pulsed Pressure Perturbations, an Extra Dimension in NMR Spectroscopy of Proteins. J. Am. Chem. Soc. 133: 13646–13651.
96. Charlier, C., T. Reid Alderson, J.M. Courtney, J. Ying, P. Anfinrud, and A. Bax. 2018. Study of protein folding under native conditions by rapidly switching the hydrostatic pressure inside an NMR sample cell. Proc. Natl. Acad. Sci. U. S. A. 115: E4169–E4178.
97. Lerch, M.T., Z. Yang, C. Altenbach, and W.L. Hubbell. 2015. High-Pressure EPR and Site-Directed Spin Labeling for Mapping Molecular Flexibility in Proteins. In: Qin, PZ and Warncke K, editor. Electron Paramagentic Resonance Investigations of Biological Systems. . pp. 29–57.
98. Chen, T., K. Dave, and M. Gruebele. 2018. Pressure- and heat-induced protein unfolding in bacterial cells: crowding vs. sticking. FEBS Lett. 592: 1357–1365.
99. Dellarole, M., and C.A. Royer. 2014. High-Pressure Fluorescence Applications, chapter 4. In: Fluorescence Spectroscopy and Microscopy. . pp. 53–74.
100. Frey, B., M. Hartmann, M. Herrmann, R. Meyer-Pittroff, K. Sommer, and G. Bluemelhuber. 2006. Microscopy under pressure - An optical chamber system for fluorescence microscopic analysis of living cells under high hydrostatic pressure. Microsc. Res. Tech. 69: 65–72.
101. Hartmann, M., M. Kreuss, and K. Sommer. 2004. High pressure microscopy--a powerful tool for monitoring cells and macromolecules under high hydrostatic pressure. Cell. Mol. Biol. (Noisy-le-grand). 50: 479–484.
102. Bugnon, P., G. Laurenczy, Y. Ducommun, P.Y. Sauvageat, A.E. Merbach, R. Ith, R. Tschanz, M. Doludda, R. Bergbauer, and E. Grell. 1996. High-pressure stopped-flow spectrometer for kinetic studies of fast reactions by absorbance and fluorescence detection. Anal Chem. 68: 3045–3049.
103. Dumont, C., T. Emilsson, and M. Gruebele. 2009. Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. Nat. Methods. 6: 515-U70.
104. Prigozhin, M.B., Y. Zhang, K. Schulten, M. Gruebele, and T. V. Pogorelov. 2019. Fast pressure-jump all-atom simulations and experiments reveal site-specific protein dehydration-folding dynamics. Proc. Natl. Acad. Sci. 116: 5356–5361.
105. Wirth, A.J., Y. Liu, M.B. Prigozhin, K. Schulten, and M. Gruebele. 2015. Comparing Fast Pressure Jump and Temperature Jump Protein Folding Experiments and Simulations. J. Am. Chem. Soc. 137: 7152–7159.
106. Bourges, A.C., A. Lazarev, N. Declerck, K.L. Rogers, and C.A. Royer. 2020. Quantitative High-Resolution Imaging of Live Microbial Cells at High Hydrostatic Pressure. Biophys. J. 118: 2670–2679.
107. Vass, H., S.L. Black, E.M. Herzig, F.B. Ward, P.S. Clegg, and R.J. Allen. 2010. A multipurpose modular system for high-resolution microscopy at high hydrostatic pressure. Rev. Sci. Instrum. 81: 053710.
108. Vass, H., S. Lucas Black, C. Flors, D. Lloyd, F. Bruce Ward, and R.J. Allen. 2013. Single-molecule imaging at high hydrostatic pressure. Appl. Phys. Lett. 102.
109. Ebbinghaus, S., A. Dhar, J.D. McDonald, and M. Gruebele. 2010. Protein folding stability and dynamics imaged in a living cell. Nat. Methods. 7: 319–323.
110. Hayashi, R., Y. Kakehi, M. Kato, N. Tanimizu, S. Ozawa, M. Matsumoto, S. Kawai, and P. Pudney. 2002. Circular dichroism under high pressure. In: Progress in Biotechnology. Elsevier. pp. 583–590.
111. Nagata, Y., R. Takeda, and M. Suginome. 2016. High-pressure circular dichroism

spectroscopy up to 400 MPa using polycrystalline yttrium aluminum garnet (YAG) as pressure-resistant optical windows. RSC Adv. 6: 109726–109729.
112. Lerch, M.T., J. Horwitz, J. McCoy, W.L. Hubbell, J. Horwitza, John McCoy, and Wayne L. Hubbell. 2013. Circular dichroism and site-directed spin labeling reveal structural and dynamical features of high-pressure states of myoglobin. Proc. Natl. Acad. Sci. U. S. A. 110: E4714–E4722.
113. Ma, L., Q. Ouyang, G.C. Werthmann, H.M. Thompson, and E.M. Morrow. 2017. Live-cell Microscopy and Fluorescence-based Measurement of Luminal pH in Intracellular Organelles. Front. Cell Dev. Biol. 5: 71.
114. Bagar, T., K. Altenbach, N.D. Read, and M. Benčina. 2009. Live-Cell Imaging and Measurement of Intracellular pH in Filamentous Fungi Using a Genetically Encoded Ratiometric Probe. Eukaryot. Cell. 8: 703–712.
115. Vauclare, P., F. Natali, J.P. Kleman, G. Zaccai, and B. Franzetti. 2020. Surviving salt fluctuations: stress and recovery in Halobacterium salinarum, an extreme halophilic Archaeon. Sci. Rep. 10: 3298.
116. Weinisch, L., S. Kühner, R. Roth, M. Grimm, T. Roth, D.J.A. Netz, A.J. Pierik, and S. Filker. 2018. Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate Schmidingerothrix salinarum. PLOS Biol. 16: e2003892.
117. Wang, Z. 2006. Materials Science: High-Pressure Microscopy. Science (80-.). 312: 1149–1150.
118. Makhatadze, G., and P. Privalov. 1995. Energetics of Protein Structure. In: Advances in Protein Chemistry. . pp. 307–425.
119. Darr, S.C., B. Pace, and N.R. Pace. 1990. Characterization of ribonuclease P from the archaebacterium Sulfolobus solfataricus. J. Biol. Chem. 265: 12927–12932.
120. Kumar, S., C.J. Tsai, and R. Nussinov. 2001. Thermodynamic differences among homologous thermophilic and mesophilic proteins. Biochemistry. 40: 14152–14165.
121. Hollien, J., and S. Marqusee. 1999. A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. Biochemistry. 38: 3831–3836.
122. Razvi, A., and J.M. Scholtz. 2006. A thermodynamic comparison of HPr proteins from extremophilic organisms. Biochemistry. 45: 4084–4092.
123. Matsuura, Y., M. Takehira, G.I. Makhatadze, Y. Joti, H. Naitow, N. Kunishima, and K. Yutani. 2018. Strategy for Stabilization of CutA1 Proteins Due to Ion-Ion Interactions at Temperatures of over 100°c. Biochemistry. 57: 2649–2656.
124. Suhre, K., and J.M. Claverie. 2003. Genomic correlates of hyperthermostability, an update. J. Biol. Chem. 278: 17198–17202.
125. Berezovsky, I.N., W.W. Chen, P.J. Choi, and E.I. Shakhnovich. 2005. Entropic stabilization of proteins and its proteomic consequences. PLoS Comput. Biol. 1: 0322–0332.
126. Shiraki, K., S. Nishikori, S. Fujiwara, H. Hashimoto, Y. Kai, M. Takagi, and T. Imanaka. 2001. Comparative analyses of the conformational stability of a hyperthermophilic protein and its mesophilic counterpart. Eur. J. Biochem. 268: 4144–4150.
127. Lin, L.-N., J.M.F. Brandts, and V. Plotnikov. 2002. Determination of the Volumetric Properties of Proteins and Other Solutes Using Pressure Perturbation Calorimetry. Anal. Biochem. 302: 144–60.
128. Mitra, L., N. Smolin, R. Ravindra, C. Royer, and R. Winter. 2006. Pressure perturbation calorimetric studies of the solvation properties and the thermal unfolding of proteins in solution--experiments and theoretical interpretation. Phys. Chem. Chem. Phys. 8: 1249–65.
129. Schweiker, K.L., and G.I. Makhatadze. 2009. Use of pressure perturbation calorimetry to characterize the volumetric properties of proteins. 1st ed. Elsevier Inc.
130. Papini, C.M., P.P. Pandharipande, C.A. Royer, and G.I. Makhatadze. 2017. Putting the

Piezolyte Hypothesis under Pressure. Biophys. J. 113: 974–977.
131. Rayan, G., and R.B. Macgregor Jr. 2009. Pressure-induced helix-coil transition of DNA copolymers is linked to water activity. Biophys. Chem. 144: 62–66.
132. Carr, C.E., I. Khutsishvili, and L.A. Marky. 2018. Energetics, Ion, and Water Binding of the Unfolding of AA/UU Base Pair Stacks and UAU/UAU Base Triplet Stacks in RNA. J. Phys. Chem. B. 122: 7057–7065.
133. Heerklotz, H., and J. Seelig. 2002. Application of pressure perturbation calorimetry to lipid bilayers. Biophys J. 82: 1445–1452.
134. Chalikian, T. V., J. Völker, G.E. Plum, and K.J. Breslauer. 1999. A more unified picture for the thermodynamics of nucleic acid duplex melting: A characterization by calorimetric and volumetric techniques. Proc. Natl. Acad. Sci. U. S. A. 96: 7853–7858.
135. Chen, C.R., and G.I. Makhatadze. 2017. Molecular determinant of the effects of hydrostatic pressure on protein folding stability. Nat. Commun. 8: 14561.
136. Collins, M.D., C.U. Kim, and S.M. Gruner. 2011. High-pressure protein crystallography and NMR to explore protein conformations. Ann. Rev. Biophys. 40: 81–98.
137. Fourme, R., E. Girard, R. Kahn, A.-C. Dhaussy, and I. Ascone. 2009. Advances in High-Pressure Biophysics: Status and Prospects of Macromolecular Crystallography. Annu. Rev. Biophys. 38: 153–171.
138. Fraser, J.S., H. van den Bedem, A.J. Samelson, P.T. Lang, J.M. Holton, N. Echols, and T. Alber. 2011. Accessing protein conformational ensembles using room-temperature X-ray crystallography. Proc. Natl. Acad. Sci. U. S. A. 108: 16247–16252.
139. Keedy, D.A., H. van den Bedem, D.A. Sivak, G.A. Petsko, D. Ringe, M.A. Wilson, and J.S. Fraser. 2014. Crystal Cryocooling Distorts Conformational Heterogeneity in a Model Michaelis Complex of DHFR. Structure. 22: 899–910.
140. Colloc'h, N., E. Girard, A.-C. Dhaussy, R. Kahn, I. Ascone, M. Mezouar, and R. Fourme. 2006. High pressure macromolecular crystallography: the 140-MPa crystal structure at 2.3 A resolution of urate oxidase, a 135-kDa tetrameric assembly. Biochim. Biophys. Acta Proteins Proteomics. 1764: 391–7.
141. Barstow, B., N. Ando, C.U. Kim, and S.M. Gruner. 2008. Alteration of citrine structure by hydrostatic pressure explains the accompanying spectral shift. Proc Natl Acad Sci U S A. 105: 13362–13366.
142. Urayama, P., G.N. Phillips, and S.M. Gruner. 2002. Probing substates in sperm whale myoglobin using high-pressure crystallography. Structure. 10: 51–60.
143. Tuukkanen, A.T., A. Spilotros, and D.I. Svergun. 2017. Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons. IUCrJ. 4: 518–528.
144. Ando, N., P. Chenevier, M. Novak, M.W. Tate, and S.M. Gruner. 2008. High hydrostatic pressure small-angle X-ray scattering cell for protein solution studies featuring diamond windows and disposable sample cells. J. Appl. Crystallogr. 41: 167–175.
145. Brooks, N.J., B.L. Gauthe, N.J. Terrill, S.E. Rogers, R.H. Templer, O. Ces, and J.M. Seddon. 2010. Automated high pressure cell for pressure jump x-ray diffraction. Rev. Sci. Instrum. 81: 64103.
146. Fujisawa, T., and M. Kato. 2002. The Small-Angle X-Ray Scattering from Proteins Under Pressure. In: Taniguchi Y, HE Stanley, H Ludwig, editors. Biological Systems Under Extreme Conditions: Structure and Function. Springer. pp. 121–138.
147. Krywka, C., C. Sternemann, M. Paulus, M. Tolan, C. Royer, and R. Winter. 2008. Effect of osmolytes on pressure-induced unfolding of proteins: A high-pressure SAXS study. ChemPhysChem. 9: 2809–2815.
148. Teixeira, S. c. m., J.B. Leão, C. Gagnon, and M.A. McHugh. 2018. High pressure cell for Bio-SANS studies under sub-zero temperatures or heat denaturing conditions. J. Neutron Res. 20: 11–21.
149. Woenckhaus, J., R. Köhling, R. Winter, P. Thiyagarajan, and S. Finet. 2000. High

pressure-jump apparatus for kinetic studies of protein folding reactions using the small-angle synchrotron x-ray scattering technique. AIP Rev. Sci. Instrum. 71: 3895–3899.
150. Jenkins, K.A., M.J. Fossat, S. Zhang, D.K. Rai, S. Klein, R. Gillilan, Z. White, G. Gerlich, S.A. McCallum, R. Winter, S.M. Gruner, D. Barrick, and C.A. Royer. 2018. The consequences of cavity creation on the folding landscape of a repeat protein depend upon context. Proc. Natl. Acad. Sci. 115: E8153–E8161.
151. Schroer, M.A., M. Paulus, C. Jeworrek, C. Krywka, S. Schmacke, Y. Zhai, D.C.F. Wieland, C.J. Sahle, M. Chimenti, C.A. Royer, B. Garcia-Moreno, M. Tolan, and R. Winter. 2010. High-Pressure SAXS Study of Folded and Unfolded Ensembles of Proteins. Biophys. J. 99: 3430–3437.
152. Möller, J., M. Schroer, M. Erlkamp, S. Grobelny, M. Paulus, S. Tiemeyer, F. Wirkert, M. Tolan, and R. Winter. 2012. The effect of ionic strength, temperature, and pressure on the interaction potential of dense protein solutions: from nonlinear pressure response to protein crystallization. Biophys J. 102: 2641–2648.
153. Schroer, M.A., Y. Zhai, D.C.F. Wieland, C.J. Sahle, J. Nase, M. Paulus, M. Tolan, and R. Winter. 2011. Exploring the Piezophilic Behavior of Natural Cosolvent Mixtures. Angew. Chem. Int. Ed. Engl. 50: 11413–11416.
154. Peters, J.W., G.J. Schut, E.S. Boyd, D.W. Mulder, E.M. Shepard, J.B. Broderick, P.W. King, and M.W.W. Adams. 2015. [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. Biochim. Biophys. Acta - Mol. Cell Res. .
155. Pereira, I.A.C., A.R. Ramos, F. Grein, M.C. Marques, S.M. da Silva, and S.S. Venceslau. 2011. A comparative genomic analysis of energy metabolism in sulfate reducing bacteria and archaea. Front. Microbiol. 2: 1–22.
156. Thauer, R.K. 1998. Biochemistry of methanogenesis : a tribute to Mary Stevenson. 144: 2377–2406.
157. Tremblay, P.-L., T. Zhang, S.A. Dar, C. Leang, and D.R. Lovley. 2013. The Rnf Complex of of Clostridium ljungdahlii Is a Proton-Translocating Ferredoxin:NAD$^+$ Oxidoreductase Essential for Autotrophic Growth Pier-Luc. Appl. Environ. Microbiol. 8: 1–8.
158. Hsieh, Y.C., M.Y. Liu, V.C.C. Wang, Y.L. Chiang, E.H. Liu, W.G. Wu, S.I. Chan, and C.J. Chen. 2010. Structural insights into the enzyme catalysis from comparison of three forms of dissimilatory sulphite reductase from Desulfovibrio gigas. Mol. Microbiol. 78: 1101–1116.
159. Oliveira, T.F., E. Franklin, J.P. Afonso, A.R. Khan, N.J. Oldham, I.A.C. Pereira, and M. Archer. 2011. Structural insights into dissimilatory sulfite reductases: Structure of desulforubidin from Desulfomicrobium norvegicum. Front. Microbiol. 2: 1–12.
160. Schiffer, A., K. Parey, E. Warkentin, K. Diederichs, H. Huber, K.O. Stetter, P.M.H. Kroneck, and U. Ermler. 2008. Structure of the Dissimilatory Sulfite Reductase from the Hyperthermophilic Archaeon Archaeoglobus fulgidus. J. Mol. Biol. 379: 1063–1074.
161. Ilina, Y., C. Lorent, S. Katz, J.H. Jeoung, S. Shima, M. Horch, I. Zebger, and H. Dobbek. 2019. X-ray Crystallography and Vibrational Spectroscopy Reveal the Key Determinants of Biocatalytic Dihydrogen Cycling by [NiFe] Hydrogenases. Angew. Chemie - Int. Ed. 58: 18710–18714.
162. Ermler, U., W. Grabarse, S. Shima, M. Goubeaud, and R.K. Thauer. 1997. Crystal structure of methyl-coenzyme M reductase: The key enzyme of biological methane formation. Science (80-.). 278: 1457–1462.
163. Zhou, L., and L.A. Sazanov. 2019. Structure and conformational plasticity of the intact Thermus thermophilus V/A-type ATPase. Science (80-.). 365.
164. Boyd, E.S., and J.W. Peters. 2013. New insights into the evolutionary history of biological nitrogen fixation. Front. Microbiol. 4: 1–12.
165. Mus, F., A.B. Alleman, N. Pence, L. Seefeldt, and J.W. Peters. 2018. Exploring the alternatives of biological nitrogen fixation. Metallomics. 10: 523–538.

166. El-Gebali, S., J. Mistry, A. Bateman, S.R. Eddy, A. Luciani, S.C. Petter, M. Qureshi, L.J. Richardson, G.A.. Salazar, A. Smart, E.L.L. Sonnhammer, L. Hirsh, L. Paladin, D. Piovesan, S.C.E. Tosatto, and R.D. Finn. 2019. The Pfam protein families database in 2019. Nucleic Acids Res. 47: 427–432.
167. Marietou, A., A.T.T. Nguyen, E.E. Allen, and D.H. Bartlett. 2014. Adaptive laboratory evolution of Escherichia coli K-12 MG1655 for growth at high hydrostatic pressure. Front. Microbiol. 5: 749.
168. Rinke, C., P. Schwientek, A. Sczyrba, N.N. Ivanova, I.J. Anderson, J.F. Cheng, A. Darling, S. Malfatti, B.K. Swan, E.A. Gies, J.A. Dodsworth, B.P. Hedlund, G. Tsiamis, S.M. Sievert, W.T. Liu, J.A. Eisen, S.J. Hallam, N.C. Kyrpides, R. Stepanauskas, E.M. Rubin, P. Hugenholtz, and T. Woyke. 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature. 499: 431–437.
169. Hedlund, B.P., J.A. Dodsworth, S.K. Murugapiran, C. Rinke, and T. Woyke. 2014. Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter." Extremophiles. 18: 865–875.

Figure Legends

Figure 1. The cumulative number of genomes in the Department of Energy Integrated Microbial Genomes (IMG) database. Genomes are separated into those generated from cultivars, genomes assembled from metagenome data (MAGs), and single cell amplified genomes (SAGs) and ordered by their 'addition date'. The availability of SAG data for 2020 has yet to be reported and is thus not included on this plot.

Figure 2. The distribution of environment types where isolate genomes (top) and MAGs (bottom) have been generated. Metadata for each genome was used to group genomes into environment types from where they originated. The curves show accumulation of genomes from each environment type over time and are colored by the legend at the bottom. The host-associated isolate genomes curve is not shown for clarity, given the high number of genomes recovered from this environment type.

Figure 3. Phylum-level taxonomic composition of MAGs within environment types identified in Figure 2. Environment types are arranged with those with highest inferred phylum diversity (based on the Shannon diversity index) at the bottom. The numbers in parentheses show the Shannon diversity index values for phylum-level composition in each environment type. Phyla are shown for those that comprised over 5% of the phyla for each environment type, while the rest are pooled as 'others'. Phylum designations are those provided by IMG metadata, with the exception of Candidate Phylum Radiation lineage, which are collapsed as 'Patescibacteria' based on recent taxonomic revisions ref 46.

Figure 4. Thermodynamic basis for the temperature stabilization of proteins. Taken from (58). A, the dependence of Gibbs energy of unfolding on temperature; B, the dependence of enthalpy on temperature; C, the dependence of entropic contribution on temperature. Models correspond to those discussed in the text.

Figure 5. Proposed structural basis for piezophilic character of IMPDH from *Shewanella benthica* DB21MT-2. Taken from (70). A) Structures of DHFR from *Shewanella oneidensis* (SoIMPDH) and *Shewanella benthica* DB21MT-2 overlaid with the a and b subunits of soIPMDH in pink and brown and those of SpIPMDH in green and cyan. The IPM molecules are shown in green stick and the magnesium ions in yellow spheres. The single amino acid substitution that confers high-pressure tolerance is S266 in SoIPMDH, and is shown in red stick. B-E) Zoom of the vicinity of residue 266 for soIPMDH at atmospheric pressure (B) and 580 MPa (C) and for soIPMDH at atmospheric pressure (D) and at 580 MPa. Several water molecules are visible in the high pressure structure of SoIPMDH which are not present in that of SpIPMDH.

Figure 6. Substitutions for the β-galactosidase from a halophilic organism. Surface residues (orange) and interior residues (purple) substituted in *H. lactusprofund*i relative to non-halophilic organisms, highlighting a complex network of substitutions responsible for the wide range of temperature and salinity and solvents in which this enzyme remains functional.

Figure 7. High resolution time-resolved microscopy of live microbial cells at high pressure. High pressure measurements in bacteria (A) vs. archaea (B) reveal that the NADPH bound-unbound ratio (color coded pink to yellow) is very sensitive to pressure in bacteria, but not in archaea. Adapted from ref. (105) with permission.

Figure 8. Sequence similarity network (SSN) for nitrogenase component 1 type oxidoreductase (pfam00148, May 2020) (165). Each square corresponds to one enzyme, and connecting edges represent sequence similarity above or equal to a threshold cutoff of an E-value of 1 x 10-230 (corresponding to 62% identity). Individual enzyme nodes are colored by super-kingdom, with enzymes from thermophilic organisms highlighted in red and enzymes from organisms found at greater than or equal to 50 m depth highlighted in orange. Enlarged diamonds are enzymes with deposited structures in the PDB, labeled by accession code of the highest-resolution structure available. The more distantly related CfbD subfamily, although a nitrogenase-like oxidoreductase, is not shown. Nitrogenase crystal structure (PDB: 1M34, Av) depicts catalytic NifDK heterotetramer with two ATPase NifH homodimers coordinated at opposing ends. NifDK is homologouse to NifEN, which is responsible for nitrogenase cofactor maturation. The catalytic subunits for DPOR and COR are ChlBN and BchXY.