



Entropic contribution to enhanced thermal stability in the thermostable P450 CYP119

Zhuo Liu^{a,1}, Sara Lemmonds^{b,1}, Juan Huang^{c,d}, Madhusudan Tyagi^{e,f}, Liang Hong^{a,g,2}, and Nitin Jain^{b,2}

^aSchool of Physics and Astronomy, Shanghai Jiao Tong University, 200240 Shanghai, People's Republic of China; ^bDepartment of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996; ^cState Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University, 200240 Shanghai, People's Republic of China; ^dSchool of Life Sciences and Biotechnology, Shanghai Jiao Tong University, 200240 Shanghai, People's Republic of China; ^eCenter for Neutron Research, National Institute of Standards and Technology, Gaithersburg, MD 20899; ^fDepartment of Materials Science and Engineering, University of Maryland, College Park, MD 20742; and ^gInstitute of Natural Sciences, Shanghai Jiao Tong University, 200240 Shanghai, People's Republic of China

Edited by Gregory A. Petsko, Weill Cornell Medical College, New York, NY, and approved September 12, 2018 (received for review May 1, 2018)

The enhanced thermostability of thermophilic proteins with respect to their mesophilic counterparts is often attributed to the enthalpy effect, arising from strong interactions between protein residues. Intuitively, these strong interresidue interactions will rigidify the biomolecules. However, the present work utilizing neutron scattering and solution NMR spectroscopy measurements demonstrates a contrary example that the thermophilic cytochrome P450, CYP119, is much more flexible than its mesophilic counterpart, CYP101A1, something which is not apparent just from structural comparison of the two proteins. A mechanism to explain this apparent contradiction is that higher flexibility in the folded state of CYP119 increases its conformational entropy and thereby reduces the entropy gain during denaturation, which will increase the free energy needed for unfolding and thus stabilize the protein. This scenario is supported by thermodynamic data on the temperature dependence of unfolding free energy, which shows a significant entropic contribution to the thermostability of CYP119 and lends an added dimension to enhanced stability, previously attributed only to presence of aromatic stacking interactions and salt bridge networks. Our experimental data also support the notion that highly thermophilic P450s such as CYP119 may use a mechanism that partitions flexibility differently from mesophilic P450s between ligand binding and thermal stability.

thermophilic protein | cytochrome p450 | entropy-driven | thermostability | flexibility

Cytochrome P450s are ubiquitous heme protein monooxygenases that catalyze various biochemical reactions constituting drug metabolism, lipid and steroid biosynthesis, carcinogenesis, and degradation of pollutants in higher organisms (1–5). The functional activity of these enzymes relies on conformational flexibility of the binding regions to recognize substrates with different physicochemical properties and convert them to monooxygenated products with high regioselectivity and stereo selectivity. Further development of P450 enzymes as practical catalysts in medical and industrial applications would benefit from this flexibility, but would also require them to be highly stable to enable efficient usage in a bioreactor-type setting for such applications. While most regular (mesophilic) P450 enzymes denature around 50 °C to 60 °C, thermally stable (thermophilic) P450s are adapted to meet this requirement, as they are structurally stable and function optimally at temperatures of 80 °C or above (6), and thus can be particularly advantageous in catalyzing reactions such as oxidation of unactivated C–H bonds, for example in lipophilic substrates, often feasible only at very high temperatures. In addition, the enhanced structural stability has great potential utility in carrying out mechanistic investigations of the P450 catalytic cycle that are difficult to carry out with the mesophilic enzymes (7, 8). Thus, understanding the microscopic factors governing the thermostability of these enzymes is of significance in designing new enzymes that can work at extreme environment for even a wider variety of defined substrates (9–13).

Discerning the basis of thermostability, in principle, can potentially be accomplished from a detailed structural and thermodynamic study of a large number of thermophilic P450s and then comparison with mesophilic homologs within a superfamily to identify the relationships between structural elements/interactions and thermodynamic factors that impart thermostability. Elucidation of the specific determinants of thermostability in thermophilic P450s so far have relied mainly on structural analysis of four thermophilic P450s with available crystal structures: CYP119, CYP175A1, CYP231A2, and P450st (14–17). Structural comparison with mesophilic P450 analogs suggests varying structural features such as shorter loops, more compact hydrophobic core, aromatic side-chain stacking, and/or salt bridge networks that may primarily contribute to the enhanced stability of thermophilic P450s, with no single universal feature believed to be responsible for thermostability (15, 18–21). For example, by comparing the crystal structures of CYP119 with CYP101A1, considered to be its mesophilic counterpart based on closest sequence and structural homology (22), it is suggested that the extensive clustering of aromatic residues and salt bridge interaction networks enhances the thermal stability of CYP119. On the other hand, CYP175A1 lacks such aromatic clustering and other factors such as salt bridge

Significance

Understanding the thermodynamic factors responsible for enhanced stability of thermophilic cytochrome P450 enzymes is significant in their development as efficient catalysts for high-temperature monooxygenation reactions. These factors are usually inferred from structural comparison with mesophilic P450s and invoke rigidifying enthalpic interactions as primary thermodynamic determinants for thermostability. Using the thermophilic P450 CYP119 as an example, the present work, however, questions this enthalpy-driven notion of P450 thermostability and provides strong experimental evidence that their thermostability may actually be entropy-driven due to increased flexibility in the folded state relative to mesophilic P450s and that this flexibility is partitioned effectively between functional activity and thermal stability. This represents a major paradigm shift in understanding the basis of thermostability in thermophilic P450s.

Author contributions: L.H. and N.J. designed research; Z.L., S.L., J.H., M.T., L.H., and N.J. performed research; Z.L., L.H., and N.J. analyzed data; and L.H. and N.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

¹Z.L. and S.L. contributed equally to this work.

²To whom correspondence may be addressed. Email: njain@utk.edu or hong3liang@sjtu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1807473115/-DCSupplemental.

Published online October 8, 2018.

interactions, hydrogen bonds, and shorter loops have been invoked to account for its thermostability (15, 22). Such comparisons have also been extended to other thermophilic P450s identified so far and used to interpret the thermodynamics basis of their thermal stability (14–17).

Based on such comparisons, a prevailing view is that the thermostability of thermophilic proteins is primarily enthalpic in origin and the entropic contribution to the thermal stability [$\Delta G(T)$] of a folded protein (Eq. 1) is generally deemed to be negligible in such cases.

$$\Delta G(T) = \Delta H_T - T \cdot \Delta S_T. \quad [1]$$

Intuitively, the strong attractive forces inside the protein molecule, such as hydrogen bonds, salt bridges, aromatic stacking interactions, and nonpolar van der Waals interactions that are enthalpic in nature, will not only stabilize the structure but also rigidify the macromolecule in the folded state to protect it from denaturing at high temperature (23–25). However, a wealth of experimental data accumulated on flexibility of various mesophilic P450s suggest that these proteins are highly dynamic entities which use their conformational dynamics as an integral feature to facilitate substrate recognition and catalysis (26). Structures of thermophilic P450s with various ligands indicate a similar trend, with the flexibility especially high within the active site regions involving the F/G helices and the F/G loop region, where distinct changes allow each thermophilic P450 to adapt to specific ligands with appropriate physicochemical properties, similar to what is observed in mesophilic P450s (27–29). Recent experimental data on various mesophilic P450s by several groups, including ours, indicate that the protein dynamics is not only restricted to the active site regions but also involves regions remote from the catalytic site (27, 30–34). For example, in P450-BM3, mutations done at remote sites affected ligand binding, likely due to change in active site dynamics (35). Also, in CYP101A1, collective dynamic modes on the picosecond to nanosecond timescales were detected to be prevalent throughout the protein, and not just the active site, that create access to the active site for the substrate (26, 36). Even in the case of other thermophilic proteins such as rubredoxin and glutamate dehydrogenase, increased fluctuations were detected throughout the protein by hydrogen/deuterium exchange NMR spectroscopy and molecular dynamics (MD) simulations (37–39). These motions were correlated to a broadened energy landscape exhibiting more conformational states than their mesophilic counterparts, which was then related to an increased entropic reservoir to explain the observed thermodynamic properties of the proteins. Results from these studies therefore raise an inevitable question as to whether the thermal resilience of thermophilic P450s arises primarily from the rigidity resulting from enthalpic interactions or whether there is an important previously unobserved role for the inherent flexibility of these proteins that would provide a thermodynamic explanation for the increased thermal stability.

In the present study, we reconcile this flexibility/rigidity notional discrepancy by providing evidence from solution NMR and neutron scattering measurements on CYP119 and its mesophilic counterpart CYP101A1, which shows that the thermophilic CYP119 exhibits increased global flexibility and conformational disorder in the folded state relative to its mesophilic counterpart CYP101A1, while, at the same time, retaining its enthalpic interactions and thermostability. CYP119 is an ideal candidate for such dynamic and thermodynamic characterization since it was the first thermophilic P450 discovered from the thermophilic organism *Sulfolobus acidocaldarius* with a melting temperature of $\sim 90^\circ\text{C}$ compared with a value of $\sim 50^\circ\text{C}$ for its mesophilic counterpart CYP101A1 and has been the most experimentally characterized from the group of thermophilic P450s in terms of mechanistic and structural features (22, 40). By performing detailed measurements of thermodynamic parameters for CYP119 using a fluorescence-based technique and

comparing them to similar parameters in the mesophilic counterpart CYP101A1, we show that increased dynamics of CYP119 in the folded state can lower the entropy gain and thus increase the free energy of unfolding, making an important entropic contribution to the protein's thermal stability.

Results

Far-UV Circular Dichroism Spectra and Thermal Melting Curves of CYP101A1 and CYP119. Far-UV circular dichroism (CD) spectra of CYP101A1 and CYP119 were acquired at 25°C (Fig. 1A) and indicate that both proteins are well folded and contain a mix of alpha-helices and beta-sheets in terms of secondary structure, similar to what is observed in the ligand-free crystal structures of the two proteins (14, 40), as shown in Fig. 1C. Thermal melting curves which monitor the unfolding of protein as a function of temperature were also obtained for the two proteins. Unfolding profile for the secondary structure in each protein was measured at a wavelength of 222 nm as a function of increasing temperature until the protein unfolded. The unfolded fraction, f_u , was calculated using

$$f_u = \frac{C_T - C_{min}}{C_{max} - C_{min}} \times 100\%, \quad [2]$$

where C_T is the observed CD value (millidegrees) of CYP101A1 or CYP119 at a given temperature T , and C_{max} and C_{min} are the maximum and minimum CD values (millidegrees) of CYP101A1 or CYP119 during the thermally induced unfolding process. By fitting the melting curves presented in Fig. 1B with a Boltzmann function, the melting temperature (T_m) for each protein was calculated. From the resulting values, it is evident that CYP119 used in our studies is much more thermostable than its mesophilic counterpart CYP101A1, since its T_m value ($89.0 \pm 1.2^\circ\text{C}$) is significantly higher than that of CYP101A1 ($50 \pm 1.0^\circ\text{C}$), in agreement with previously reported results (41).

Comparison of Thermodynamic Parameters Between CYP101A1 and CYP119 Based on Free Energy Stability Curves. Free energy stability curves were determined for both CYP101A1 and CYP119 via the chemical denaturation method using guanidinium hydrochloride (GdnHCl). The temperature dependence of GdnHCl-induced unfolding of each protein was measured as described in the experimental section to determine the free energy of unfolding ΔG and also the free energy stability curve (Fig. 2A). A fluorescence-based technique was utilized to monitor the unfolding of the protein which relies on use of the fluorescent dye, SYPRO Orange. SYPRO Orange is known to bind to hydrophobic regions of a protein with high affinity and has been used previously to study ligand binding, protein stability, and unfolding of proteins to determine T_m values using differential scanning fluorimetry or thermal shift assays (42, 43). The fluorescent intensity of the dye increases dramatically when it binds to exposed hydrophobic regions of the protein as the protein unfolds, making this method a very useful quantitative monitor of protein unfolding. We have adapted this fluorescence-based method in a manner to determine protein free energy stability curves for thermodynamic parameter measurements by systematically measuring the protein unfolding process as a function of increasing amounts of chemical denaturant, which, to our knowledge, has not been demonstrated previously for any protein. A major advantage of using fluorescence as a means of detection rather than traditional CD spectroscopic methods for such thermodynamic measurements is that it can be done with greater speed and requires lower amounts of samples. This is because the method is amenable to setup in a 96-well or 384-well format in a real-time PCR instrument, where, due to availability of a temperature-gradient feature and compatible fluorescence filter sets (λ_{ex} 470 nm/ λ_{em} 570 nm), the entire stability curve comprising free energy measurements at different

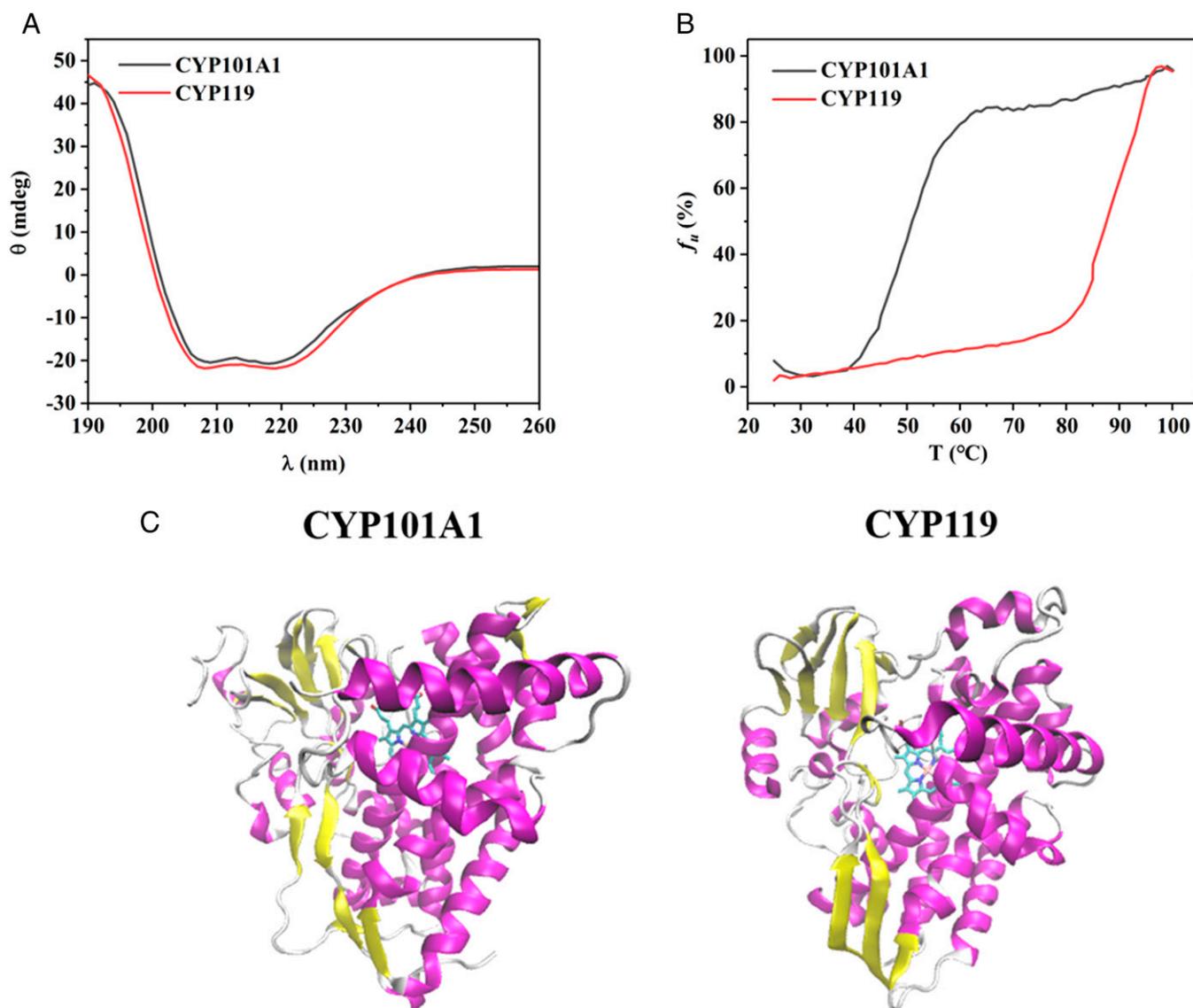


Fig. 1. CD spectra and thermal melting curves of ligand-free CYP101A1 and ligand-free CYP119. (A) Comparison of far-UV CD spectra of CYP101A1 (black) and CYP119 (red). Spectra were acquired at 25 °C on a sample of each protein with a concentration of 5 μ M in 50 mM potassium phosphate, 50 mM potassium chloride buffer, pH 7.4. (B) Comparison of thermal unfolding curves of CYP101A1 (black) and CYP119 (red). Unfolding of the secondary structure of each protein was monitored at a fixed wavelength of 222 nm as a function of increasing temperature, where f_u is the unfolding fraction. Sample conditions were similar to those used in measurement of far-UV CD spectra in A. Experimentally measured ellipticity values were converted to percentage unfolded fraction f_u , as described in the text. (C) Crystal structures of (Left) ligand-free CYP101A1 (PDB ID code 3L61) and (Right) ligand-free CYP119 (PDB ID code 11O7).

temperatures can be acquired in a single run, reducing the overall measurement time from days to a few hours.

For comparison, the free energy stability curve of CYP119 obtained by the traditional CD spectroscopy method is presented in *SI Appendix, Fig. S2* and agrees quite well with the fluorescence measurements. We envision that this method will find general applicability in similar thermodynamic measurements for other thermophilic P450s, although due diligence must be performed for factors such as solubility of the protein, exposed hydrophobic patches, and background fluorescence from intrinsically fluorescent groups that may complicate the accuracy of the measurement and interpretation of results. In our case, these factors were not an issue for either protein based on structural analysis and control measurements performed (see experimental section). The free energy stability curve for CYP119 is much higher than that for CYP101A1 at all temperatures below

the melting temperature of CYP101A1 (Fig. 2A), demonstrating that CYP119 is thermally more stable. It is also found that CYP119 is maximally stable at 35 °C with a maximum stabilization free energy of \sim 25 kJ/mol, which is considerably higher than that for CYP101A1, which is \sim 12 kJ/mol at its maximal stability temperature of 15 °C.

Next, we determined the enthalpic and entropic contribution to the free energy of unfolding (Eq. 1) at any given temperature via the modified Gibbs–Helmholtz equation,

$$\Delta G(T) = \Delta H_m \left(1 - \frac{T}{T_m} \right) + \Delta C_p (T - T_m - T \ln \left(\frac{T}{T_m} \right)), \quad [3]$$

where T_m is the melting temperature, ΔG is the unfolding free energy, ΔC_p is the change in heat capacity, and ΔH_m and ΔS_m are the enthalpy penalty and entropy gain of unfolding at the

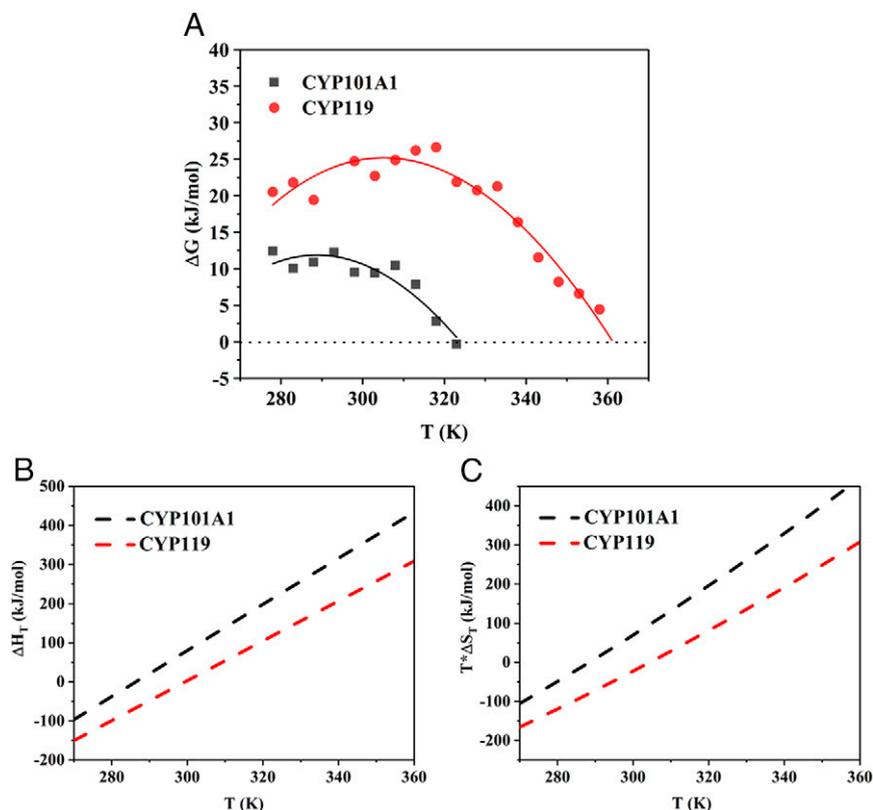


Fig. 2. (A) Free energy stability curves for ligand-free CYP101A1 (black) and ligand-free CYP119 (red) determined by fluorescence dye method as described in *Methods*. The solid lines were fitted to the $\Delta G(T)$ points obtained at various temperatures by using Eq. 3, where $\Delta G(T)$ is the free energy of unfolding determined via chemical denaturation using GdnHCl as described in *Methods* and *SI Appendix*. (B) Enthalpy change (ΔH_T) and (C) entropy change ($T^*\Delta S_T$) during unfolding versus temperature for CYP101A1 (black) and CYP119 (red). Values of ΔH_T and $T^*\Delta S_T$ were derived from Eqs. 4 and 5, respectively. The dashed lines in the plot are illustrative lines for the enthalpy change and entropy change parameters for each protein as a function of temperature.

melting temperature. All of these thermal parameters can be determined by fitting Eq. 3 to the free energy stability curves (Fig. 2A), and the results are presented in Table 1. From the fit, the values of T_m obtained are ~ 361 K (CYP119) and ~ 323 K (CYP101A1), in general agreement with the values derived from the thermal unfolding experiments obtained by circular dichroism (Fig. 1B). The change in heat capacity ΔC_p was also obtained from the fit and shows that CYP119 has smaller ΔC_p (~ 5.1 $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$) compared with CYP101A1 (~ 5.9 $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$), which is reflected in the broadening of the CYP119 stability curve relative to that of CYP101A1 (Fig. 2A). It is also observed that the ΔH_m value for CYP119 (314 $\text{kJ}\cdot\text{mol}^{-1}$) is about 45% higher than that of CYP101A1 (216 $\text{kJ}\cdot\text{mol}^{-1}$), which would usually imply that the enthalpic interactions play a major role in stabilizing CYP119 relative to CYP101A1. Additionally, the ΔS_m value for CYP119 (869 $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$) is found to be about 30% higher than the ΔS_m value for CYP101A1 (668 $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$), suggesting that there is a larger entropy change for CYP119 at its unfolding temperature. Thus, it would seem that the favorable enthalpic change for CYP119 is solely responsible for determining the thermostability of CYP119 while the entropic contribution would not be of any significance, as has also been suggested by previous thermodynamic studies with other P450s such as CYP175A1 (15). However, this conclusion changes if the comparison is done at similar temperatures rather than the respective unfolding temperatures for the two proteins which are significantly different.

Such a comparison can be made by determining the individual enthalpic and entropic contributions to unfolding free energy at any given temperature other than T_m using the relationships

$$\Delta H_T = \Delta H_m + \Delta C_p(T - T_m) \quad [4]$$

$$\Delta S_T = \Delta S_m + \Delta C_p \ln\left(\frac{T}{T_m}\right). \quad [5]$$

The resulting plots of ΔH_T and ΔS_T as a function of temperature are presented in Fig. 2B and C. A striking trend observed from these plots is that both the enthalpy and entropy changes for CYP119 are actually much lower than those for CYP101A1 at all temperatures. Thus, even though CYP119 seems thermally more stable than CYP101A1 at their respective unfolding temperatures due to higher enthalpic contribution, when the comparison is made at the same temperature, the relative enthalpic contribution is actually lower, while the entropic contribution to thermal stability is much more significant for CYP119 due to a smaller entropy gain during unfolding, which reduces the free energy of unfolding and thus thermally stabilizes CYP119. In

Table 1. Thermodynamic parameters of CYP101A1 and CYP119 determined from the thermal stability curves

Protein	T_m , K	ΔH_m , $\text{kJ}\cdot\text{mol}^{-1}$	ΔC_p , $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$	ΔS_m , $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$
CYP101A1	323 ± 1.7	216 ± 25	5.9 ± 0.6	668 ± 73
CYP119	361 ± 1.4	314 ± 22	5.1 ± 0.6	869 ± 60

T_m , ΔH_m , and ΔC_p were determined by fitting the stability curve for each protein to the Gibbs–Helmholtz equation (Eq. 3). ΔS_m was calculated from ΔH_m and T_m values. Errors represent SEM.

other words, at the same temperature, CYP119 is actually less stable compared with CYP101 if only the enthalpic term is taken into account, and the reason to render CYP119 thermally more stable is that it has much less entropy gain during unfolding. Therefore, assessment of thermostability solely based on crystal structures could sometimes be misleading, as structural analysis normally addresses the enthalpic contribution while ignoring the entropic contribution.

Comparison of Inherent Protein Flexibility Between CYP101A1 and CYP119. To compare the internal dynamics in CYP101A1 and CYP119, we performed neutron scattering experiments on protein powders hydrated with D₂O, and derived the mean squared atomic displacements (MSDs), $\langle x^2(\Delta t) \rangle$, of CYP101A1 and CYP119 as a function of temperature as outlined in the experimental section. Use of neutron scattering is inherently advantageous in this case due to its ability to readily observe change in protein flexibility on a global scale over a wide temperature range and, in particular, at very high temperatures that are typically not accessible to crystallographic methods or solution methods such as NMR. Analysis of the MSD results (Fig. 3A) from neutron scattering reveals a clear dynamic transition, i.e., change in the slope of MSD at ~ 200 K, for both proteins. This dynamic transition is coupled to the thermal activation of higher-order, large-scale collective protein motions and is thus thought to be important for specific protein function in proteins (36, 44). Below this transition temperature, protein dynamics consists of mostly harmonic vibrations and methyl rotations, which are essentially the same for both CYP101A1 and CYP119, as expected. Above the transition temperature, anharmonic large-scale collective dynamic motions set in, and the MSD comparison between the two proteins demonstrates that these dynamical processes are clearly more abundant in CYP119 than in CYP101A1 at all temperatures up to their melting temperatures. Hence, it can be concluded that CYP119 is, overall, more dynamic than CYP101A1 at any given temperature, especially in terms of anharmonic, collective atomic motions.

The inherent protein dynamics in CYP101A1 and CYP119 was also studied by solution NMR spectroscopy to obtain residue-level comparison of protein dynamics throughout the protein. The 2D ¹H-¹⁵N transverse relaxation-optimized spectroscopy (TROSY) NMR spectra of CYP101A1 and CYP119 were acquired under identical conditions at 30 °C, and peak patterns in the two spectra were compared. As can be seen from the NMR spectra of ligand-free CYP101A1 and CYP119 in Fig. 4A, C, and E, extensive line-broadening and splitting of peaks into multiplets was observed for CYP119 resonances throughout the spectra relative to the resonances for CYP101A1. The increased NMR line-broadening in CYP119 relative to CYP101A1 is unexpected since it is a smaller protein than CYP101A1 (368 vs. 414 residues), which indicates that CYP119 is sampling multiple conformations that may be in intermediate or slow exchange, unlike CYP101A1, which exhibits a smaller amount of such characteristics. The presence of peak clusters instead of individual peaks combined with extensive line-broadening clearly implies the presence of higher conformational disorder as well as increased dynamics in CYP119 compared with CYP101A1, consistent with the results obtained from neutron scattering experiments (Fig. 3A). Interestingly, the conformational disorder still persists to a significant extent in CYP119 even in the presence of a high-affinity ligand such as 4-phenylimidazole, although the overall conformational dynamics is reduced (Fig. 4D and F). Similar reduction in dynamics is also observed for 4-phenylimidazole-bound CYP101A1 (Fig. 4B) compared with 4-phenylimidazole-bound CYP119. This observation supports our hypothesis that the inherent flexibility of CYP119 thus may not be entirely utilized for binding ligands as is generally deduced from analysis of ligand-bound and ligand-free crystal structures of CYP119, but may also likely make an important entropic contribution to its enhanced thermostability. Presence of large confor-

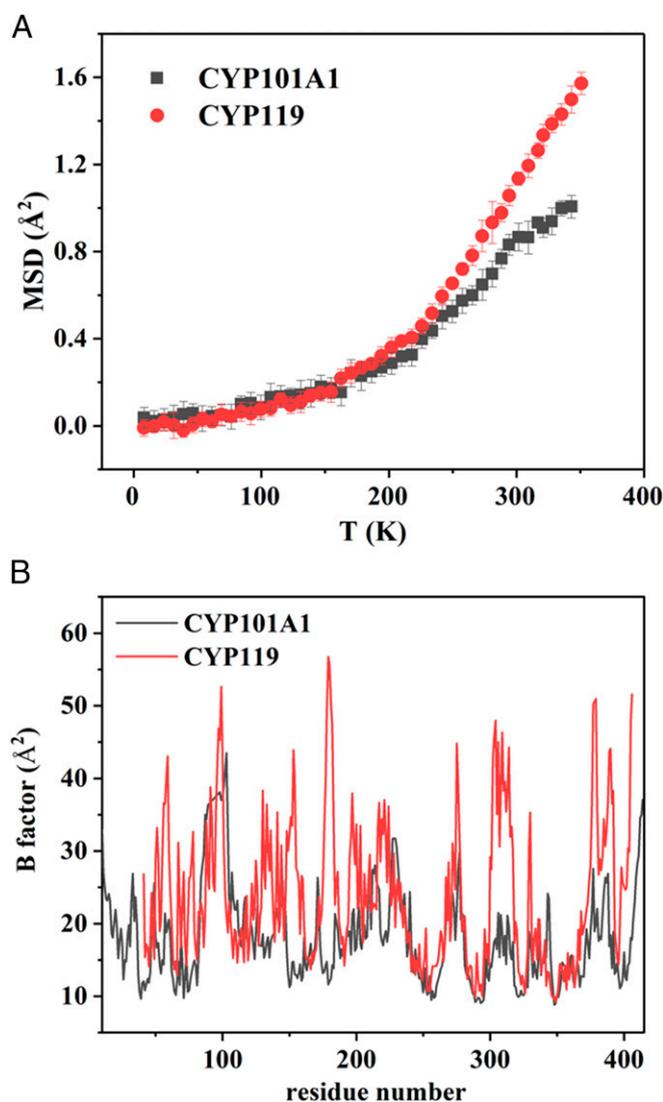


Fig. 3. (A) Comparison of MSD of ligand-free CYP101A1 (black) and ligand-free CYP119 (red) as a function of temperature from 4 K to 353 K determined by neutron scattering spectroscopy. Error bars represent one standard deviation. The detailed procedures for preparing the samples for neutron scattering, data acquisition, and analysis are provided in *Methods*. (B) Comparison of crystallography B factors of ligand-free CYP101A1 (PDB ID code 3L61) and ligand-free CYP119 (PDB ID code 11O7) at 100 K.

mational flexibility in CYP119 in the F/G loop region has been previously inferred from crystal structures of the ligand-protein complexes (14) in the form of conformational changes, from previous MD simulations, and also via 2D NMR spectroscopy of ¹³C-*p*-methoxyphenylalanine insertions or ¹⁵N selectively labeled Phe residues in CYP119 (45, 46), which all suggest that CYP119 samples a range of predetermined conformational states in solution and ligand binding selects for the appropriate conformation (s). However, the observations from these studies have been restricted primarily to the ligand binding regions that require such flexibility for diverse substrate recognition, similar to several mesophilic P450s (47). None of these previous studies have been able to detect the extent of temperature-dependent dynamic disorder observed from neutron scattering and NMR spectra in our study. Given that a protein encounters a hierarchy of states in its conformational landscape, dynamic fluctuations can occur on several timescales and length scales. Thus, it may not be possible to assess various forms of protein flexibility from MD simulations on the

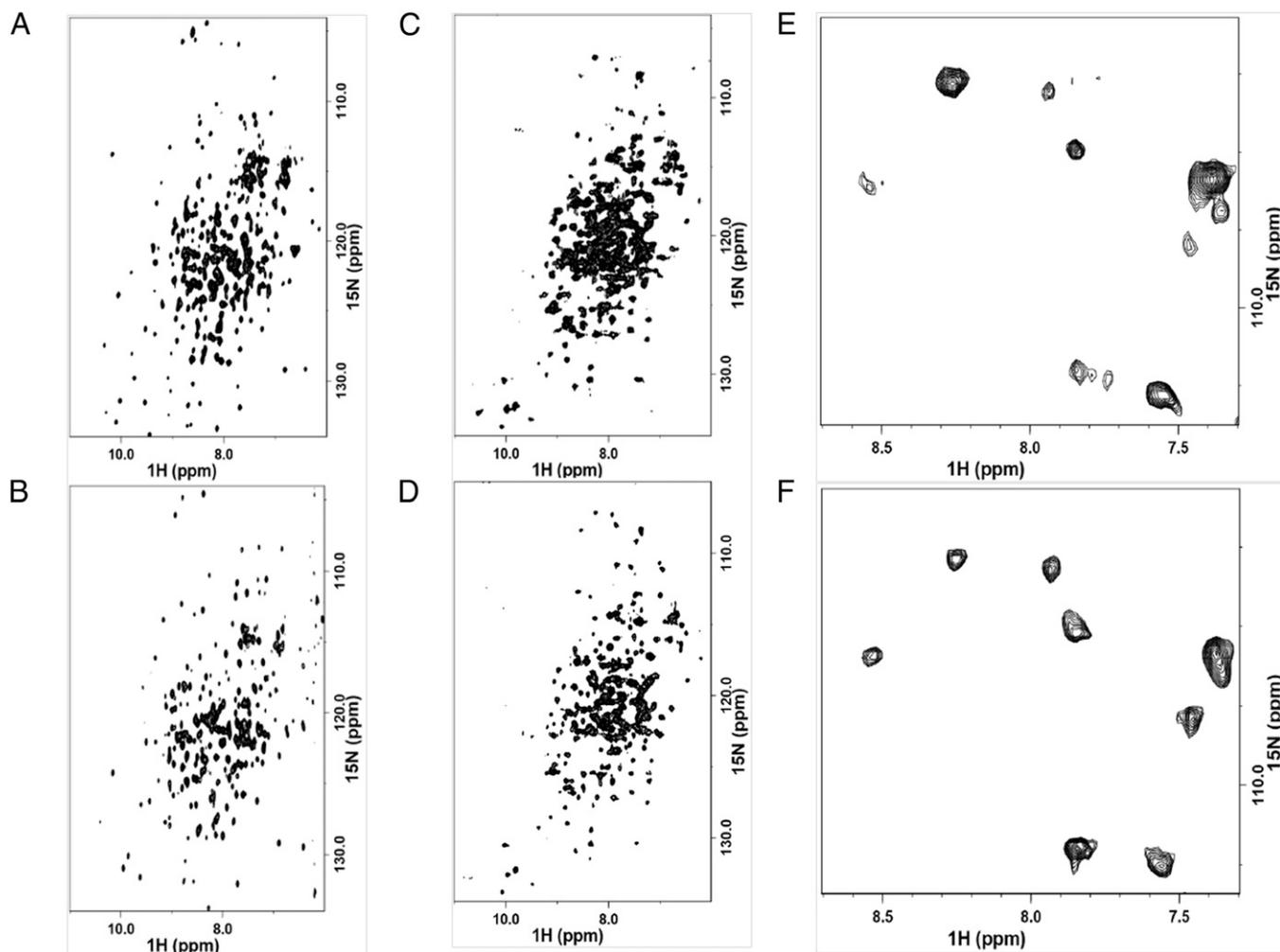


Fig. 4. Comparison of 2D ^1H - ^{15}N TROSY-HSQC (heteronuclear single-quantum coherence) NMR spectra of (A) ligand-free CYP101A1, (B) 4-phenylimidazole-bound CYP101A1, (C) ligand-free CYP119, and (D) 4-phenylimidazole-bound CYP119 at 30 °C. (E and F) Comparison of zoomed-in example regions of ligand-free CYP119 and 4-phenylimidazole-bound CYP119 full spectra in C and D to better show increased conformational disorder relative to CYP101A1 and dynamic changes upon ligand binding in CYP119. Spectra were acquired using samples of ^{15}N uniformly labeled protein with and without the ligands at a concentration of 150 μM in 50 mM potassium phosphate and 50 mM potassium chloride, pH 7.4.

small range of timescales employed in these studies or from limited NMR snapshots of the protein with selectively labeled amino acids, as the contributing fluctuation modes may be partitioned in different proportions over the protein matrix and do not represent the entire protein dynamics.

The presence of increased overall flexibility of CYP119 relative to CYP101A1 is also evident from comparison of B factors of crystal structures of ligand-free CYP101A1 [Protein Data Bank (PDB) ID code 3L61] and ligand-free CYP119 (PDB ID code 1IO7), which shows that CYP119 is globally more dynamic than CYP101A1, as shown in Fig. 3B, and not just the active site regions, which is in agreement with the neutron scattering and NMR experimental results.

Discussion

Increased Protein Flexibility as a Source of Entropic Contribution to Thermostability of CYP119. Various enthalpic factors that rely on interactions between individual amino acid residues have been proposed to account for the thermostability of thermophilic proteins (48). In the case of thermophilic P450s such as CYP119, based on comparison of crystal structures, the enhanced thermostability of the protein relative to the mesophilic counterparts such as CYP101A1 has been suggested to arise primarily from an

increased network of salt bridge interactions and extensive clustering of aromatic rings in CYP119, both of which are deemed to contribute to the enthalpy penalty toward unfolding (14, 22). Mutations involving these interactions are known to lower T_m in the range of ~ 1 °C to 10 °C (49, 50). This indicates that, while the aromatic stacking interactions and salt bridge interaction networks are important in contributing to the thermostability of CYP119, they, by themselves, do not account entirely for the observed thermostability of CYP119, the melting temperature of which is more than 35 °C higher than that of CYP101A1. Recently, non-polar interactions within the Cys ligand loop in CYP119 were also implicated in stabilization of the CYP119 fold based on lowering of T_m by up to ~ 16 °C for mutations made in this region (27). Other factors suggested to augment thermostability of CYP119 include stronger internal hydrogen bonding, a smaller, more compact structure with shorter β -turns compared with CYP101A1, and an increased number of branched amino acids such as isoleucine leading to greater hydrophobic interactions in the core of the protein. However, the contribution of these factors to thermostability of CYP119 was not deemed to be significant, based on statistical analysis of different P450 structures that found no large deviations for these factors between CYP119 and mesophilic P450s (14, 22).

A striking characteristic feature of the above factors, especially the aromatic stacking and salt bridge interactions, is that they all mostly contribute to the enthalpy term of the free energy (Eq. 1) (51). Intuitively, these stronger attractive interactions between residues would rigidify the protein molecule, and therefore one would expect CYP119 to be more rigid compared with CYP101A1. However, this is contrary to the results obtained here using neutron scattering, NMR, and comparison of crystallography B factors, which all clearly indicate that CYP119 is globally more dynamic than CYP101A1 (Figs. 3 and 4). To explain this discrepancy, it is important to consider the entropy term in understanding the thermostability of CYP119. As is observed from the plot of enthalpy and entropy terms as a function of temperature, while the enthalpy of unfolding for CYP119 clearly makes a negative contribution to the free energy stability curve compared with CYP101A1 at every temperature, the enhanced thermal stability, i.e., higher unfolding free energy of the thermophilic enzyme, results instead from its much smaller entropy gain during unfolding. The reduction in entropy gain will definitely benefit from the protein being more flexible and disordered in the folded state, as shown in Figs. 3 and 4. Such entropic driving forces are not unusual for stability of proteins, and it has been argued that the enhanced thermostability of many thermophilic proteins are indeed entropy-driven, with the thermostable proteins, on average, exhibiting a lower value of conformational entropy change (per amino acid) of unfolding (52, 53). In fact, enhancement of protein stability by site-directed mutations to reduce conformational entropy gain of unfolding has been demonstrated quite early on in bacteriophage T4 lysozyme, where Gly and Pro substitutions at key positions in the lysozyme structure enabled entropic stabilization of the protein (54). A recent study of thermal stability of Taq DNA polymerase similarly proposed a reduced entropic folding penalty as the basis for enhanced thermal stabilization for this enzyme upon analysis of thermodynamic data on thermophilic Taq and mesophilic *Escherichia coli* DNA polymerases (55). When the analysis was extended to available thermodynamic data for 17 other thermophilic–mesophilic protein pairs, analogous entropic patterns were observed in seven of these systems, which suggests that this mechanism of stabilizing thermophilic proteins may be more common than previously thought (11, 55). More recently, thermodynamic studies carried out on another thermophilic P450, CYP175A1, also support the same conclusion that the thermostability is entropy-driven, when their thermodynamic data are reanalyzed in a manner similar to ours (*SI Appendix*, Fig. S3) (15). Since the original analysis of thermodynamic data on CYP175A1 and mesophilic CYP101A1 in that publication was carried out at their respective unfolding temperatures, the entropic effects were found to be less prominent for the thermostability of the protein. However, if the comparison is carried out at the same temperature to assess entropy values, the results are actually opposite and more in line with that for CYP119 (*SI Appendix*, Fig. S3). This would also explain why CYP175A1 is highly thermostable, similar to CYP119, even though its structure is missing the characteristic aromatic stacking interactions observed in CYP119. Accounting for entropic effects is an important consideration that is often missed when the analysis is solely based on invoking enthalpic interactions by simply comparing the static crystal structures.

It is interesting to trace the origin of this entropic contribution toward thermostability in CYP119. The neutron scattering data clearly indicate significantly higher MSDs for CYP119 compared with CYP101A1 at all temperatures up to the melting temperature of CYP119, pointing to the presence of greater dynamic and conformational heterogeneity in CYP119. However, it is not readily obvious from this data what the nature of this dynamic difference is and whether this occurs throughout the protein. Analysis of NMR spectra of the two proteins offers further clues in this regard, since the NMR resonances can be analyzed in a residue-wise manner and give a dynamic assessment of the entire

protein on all timescales. Upon comparison of Fig. 4 *A* and *C*, considerable conformational disorder and dynamic differences are observed throughout the spectrum of CYP119, whereas only a small percentage of residues (<30%) are affected in CYP101A1 spectrum to the extent that they are in CYP119. This indicates that multiple conformations on a variety of timescales are being sampled by CYP119 across most of the protein, and thus we deduce that the entire structure of CYP119 protein is flexible and this flexibility is not completely abolished in CYP119 even when a strong affinity ligand such as 4-phenylimidazole binds to the active site, indicating that the purpose of the high inherent flexibility in CYP119 is not solely to accomplish the function of ligand binding. Additionally, the increased global flexibility of CYP119 is reconfirmed by the comparison of the crystallography B factors of these two proteins (Fig. 3*B*).

Such increased global flexibility is not unprecedented for thermophilic proteins. Enhanced flexibility for the thermophilic variants over the mesophilic counterparts on the millisecond to second timescale were detected for rubredoxin and α -amylase by hydrogen/deuterium exchange experiments (20, 37). Also, a study of backbone dynamics on multiple timescales of thermophilic and mesophilic ribonuclease enzymes by ^{15}N NMR relaxation measurements found that the extent of flexibility for thermophilic ribonuclease is significantly more than the mesophilic protein on certain timescales (56). Larger dynamic fluctuations on the picosecond timescale at a global level have also been reported for thermophilic malate dehydrogenase and dihydrofolate reductase by elastic incoherent neutron scattering (57, 58). The above studies clearly dispel the notion that thermophilic proteins are more rigid than their mesophilic counterparts, at least on certain timescales. These dynamic fluctuations throughout the protein may cause the low-level conformational disorder observed throughout the protein, creating a large number of accessible conformational microstates that can be occupied as the temperature increases and contribute to the conformational entropy of the folded state, thereby increasing its thermostability. In contrast, the mesophilic counterpart can initiate unfolding at earlier temperatures if it has fewer of these accessible states, especially conformations corresponding to the native folded state. It is interesting to note that another previous MD simulation on CYP119 predicted a coupled outer shell which had different dynamic characteristics than the active site that is not observed in CYP101A1 (29). Also, another recent simulation on CYP119 found hotspots remote from the active site which formed a coupled network of residues to the Cys ligand loop that, when disrupted, leads to rapid unfolding of the protein (27). Mutation of certain aromatic residues in these hotspots was found to lower the thermostability up to 16 °C (27). It is therefore plausible that a dynamic network is pervasive throughout the entire protein which has distinct microstates modulated on various timescales to impart additional thermal stability via an entropic contribution, evidence of which is visible in the differential linewidths and/or peak splittings observed throughout the protein backbone in our NMR spectra (Fig. 4*B*). This change in backbone flexibility can be small and lead to entropic effects without causing major changes in structure that would disrupt enthalpic interactions on a global scale, and would therefore mask the entropic contribution when analyzing crystal structures of thermophilic proteins.

An additional factor that can add to this entropic contribution from the backbone is side-chain configurational entropy from the numerous salt bridge networks identified in CYP119. CYP119 has a smaller number of two-residue salt bridges but is known to have a higher number of three or more residue salt bridge networks (~10) compared with CYP101A1 (~6) (41). The multitude of favorable electrostatic interactions in these networks is believed to be an important source of thermostability for CYP119 (14, 22). Previous structural and computational

studies suggest that salt bridge networks in thermophilic proteins generally are dynamic rather than static, with multiple configurations observed for salt bridge residue side chains (12). Incorporation of salt-bridges in networks overcomes the energetic cost of desolvation at high temperatures, due to favorable geometry and better electrostatic interactions resulting from shorter pair distances (an enthalpic effect), while, at the same time, the stochastic disorder due to increased thermal motions of the side chains at higher temperatures is easily accommodated in a network by increasing salt bridge occupation frequencies among the various charged side chains in the network (59). This makes the side-chain interactions long-lived, even though it reduces specific charged residue pair correlations and tends to increase the configurational entropy of both backbone and side-chain atoms in the folded form. This entropic effect would augment the enthalpic component of thermal stability imparted by the electrostatic effects. This was very elegantly demonstrated for a homotrimeric coiled coil LPP56 containing a well-packed hydrophobic core and a number of salt bridge networks on the surface (59). Analysis of salt bridge networks in ligand-free CYP119 by comparing them between two molecules present in the same asymmetric unit of the crystal structure similarly reveals presence of multiple configurations for several side chains in these networks (14). For example, in the salt-bridge network formed by Arg188/Glu184/Asp85, the side chain of Arg188 is strongly hydrogen-bonded to the side chains of both Glu184 and Asp85 in one of the molecules, but is only strongly hydrogen-bonded to the side chain of Glu184, since it is flipped in orientation and exhibits a much weaker hydrogen bonding to Asp85 in the second molecule. Another example is the side-chain network formed by Arg100/Glu96/Arg328/Glu332/Glu333, which shows different configurations in the two molecules, with differences in hydrogen bonding between Arg100 and Glu333 due to change in distance and orientation of the side chains and also in the hydrogen bonding between Arg100 and Glu332 side chains due to change in their orientation. Moreover, differences in side-chain configurations are also observed when comparing crystal structures of ligand-free CYP119 solved at two different temperatures (100 K and 270 K), with some of the largest differences in B factors between the two structures corresponding to residues incorporating the above side chains (14). The resulting side-chain configurational entropy seen in these salt bridge networks therefore sheds additional light on the role of salt bridge networks in enhancing thermal stability in thermophilic P450s, which is quite distinct from their typically enthalpic role. Since the overall entropy is distributed throughout the protein, it would also explain why individual single-site mutations in salt bridge networks only cause minor changes in T_m values, as they would result only in small amounts of entropic destabilization, but multisite mutations would likely cause larger changes due to cumulative entropic effects (50). Another worthwhile possibility to consider with regards to an alternate entropic role for side chains is aromatic residues, especially those involved in aromatic clusters in thermophilic proteins such as CYP119 that have been invoked as contributing to increased stability by virtue of their characteristic enthalpic interactions, such as aromatic stacking (41). Two prominent aromatic clusters, not found in other thermophilic P450s, are present in CYP119: one cluster formed by five residues—Tyr2, Trp4, Phe5, Phe24, and Trp281—and another cluster formed by seven residues—Phe225, Phe228, Trp231, Tyr250, Phe298, Phe334, and Phe338 (41). Apart from increasing thermal stability in CYP119 via aromatic interactions, such clusters could, in part, contribute favorably to the entropic term in these interactions via a hydrophobic effect that entails release of ordered water molecules from the surfaces of these hydrophobic aromatic side chains when they come together to form such clusters in the folded structure. This contribution, while important, is unlikely to be a major factor for enhanced stability, since these clusters are not found in other thermophilic P450s.

Also, in CYP119, they are present close to the surface and exposed to water, which has the potential to form similar ordered shells around the exposed part of aromatic side chains, albeit to a reduced extent compared with the unfolded state. This would again explain the small changes in T_m values ($<10^\circ\text{C}$) observed upon site-specific mutations carried out in these clusters (49, 50).

The presence of an increased number of accessible dynamic microstates would not only contribute toward the conformational entropy but also in increasing the heat capacity of the folded state toward that of the denatured form. As a result, ΔC_p of unfolding is reduced in CYP119 similar to what is observed in other thermophilic proteins. For example, by measuring the backbone conformational dynamics of B1 domain of streptococcal protein G, reduced conformational heat capacity was calculated from backbone NMR order parameters and correlated to increased thermal stability (60). Assuming that the unfolded states for both CYP101A1 and CYP119 are similar in nature, this would explain the lower heat capacity change for CYP119 relative to CYP101A1 reflected in the broadening of the free energy stability curve for CYP119 with a shift of T_m to higher temperatures. However, it is important to note that other factors may also contribute to the smaller ΔC_p and ΔS , such as more rigid or compact structure of the denatured protein molecules, different nature of interactions, and/or smaller entropy gain of water molecules during denaturation. These factors are very difficult to quantify in the present work and would require further extensive investigation, and thus no firm conclusions for these factors can be drawn from experiment at this time. We note here that a high-temperature (500 K) MD simulation to study unfolding of CYP119 and CYP101A1 found only a marginal increase in the radius of gyration for CYP101A1 versus CYP119 upon unfolding (27), making it less likely that differences in denatured state of the two proteins are a major source of entropic differences between the two proteins.

Partitioning of Protein Flexibility Between Ligand Binding and Thermal Stability. Finally, one may also consider the relative contribution of protein flexibility toward functional aspects such as ligand binding and catalytic activity vis-à-vis its role in thermostability. Our NMR data on the dynamic changes in CYP119 with and without ligand indicate preservation of a residual level of dynamics throughout the protein even in the presence of ligand, which points to a notion of flexibility partitioning where part of the flexibility is used for ligand binding/catalysis and part for imparting thermal stability, and that this partitioning may operate differently between thermophilic and mesophilic enzymes. This represents a previously unobserved paradigm for use of protein dynamics in two distinct biological roles within the same protein and, to our knowledge, has not been reported previously for any other thermophilic enzyme. A similar notion in context of thermal stability (but not ligand binding) has been put forward from rigidity analysis and H/D exchange studies on rubredoxin that is based upon a different way of flexibility partitioning between the mesophilic and thermophilic forms (37, 61, 62). It is possible that increased conformational fluctuations may be utilized at lower temperatures to reduce the entropic gain for thermal stability in thermophilic enzymes such as CYP119 which is not necessary in mesophilic enzymes, whereas flexibility related to activity is restricted at lower temperature and becomes more prominent at higher temperatures in these enzymes. It has been generally observed that the functional activity of thermophilic enzymes in terms of substrate turnover are several-fold lower than mesophilic enzymes around room temperature; however, activities of both types of enzymes become optimum much closer to their respective T_m values, which would support this conjecture. Further detailed experimental studies in terms of thermodynamic and dynamic measurements on multiple timescales (e.g., by NMR) are needed on CYP119 in various ligand-bound forms to characterize the extent of this partitioning as a function of temperature and

delineate the entropic portion. It remains to be seen whether the significance of entropic contribution toward enhanced thermal stability manifested in CYP119 and CYP175A1 is a general trend in thermophilic P450s. If so, further characterization could potentially open up new avenues for engineering thermostability in P450 enzymes using entropic considerations.

Conclusion

In conclusion, the thermophilic cytochrome P450, CYP119, shows higher flexibility throughout the protein than its mesophilic counterpart CYP101A1 as found by our experimental solution NMR and neutron scattering studies, which is not readily apparent only by their structural comparison. This higher flexibility in the folded state is not necessarily contradictory to enhanced thermal stability, but rather positively contributes to the latter through an entropic effect which can be reconciled by the fact that less entropy gain during denaturation will increase the unfolding free energy and thus stabilize the protein. This scenario is supported by thermodynamic measurements of the temperature dependence of unfolding free energy in the present work, which shows that the enhanced thermal stability of CYP119 relative to mesophilic CYP101A1 is indeed due to less entropy gain of the enzyme during unfolding. It is likely to result from an increased conformational entropy in the folded state that can arise from multiple sources, including backbone fluctuations on multiple timescales and dynamic disorder of side chains in salt bridge networks. As such, we propose that the higher flexibility of CYP119 may be partitioned effectively between functional activity and thermal stability and in a way that is different from mesophilic P450s.

Methods

Protein Expression and Purification. CYP101A1 protein was expressed and purified in ligand-free form using methods described previously (26). The native CYP119 gene with a C-terminal His₆-tag sequence was synthesized and subcloned into a pET29a vector (Genscript Inc). The resulting plasmid DNA was transformed into BL-21 pLysS (DE3) *E. coli* cells and placed on an LB/Agar plate containing 50 mg/mL of kanamycin for selection, followed by incubation overnight at 37 °C. A single isolated colony with kanamycin resistance from the resulting colonies was used to inoculate 50 mL of LB media containing 50 µg/mL of kanamycin and grown at 37 °C for 4 h to 5 h with shaking at 250 rpm until the culture reached an OD₆₀₀ of roughly 0.6. The cells were then spun down, and the pellet was transferred to 1 L of defined M9 medium containing 50 µg/mL of kanamycin, trace metals, 4 g of glucose as the carbon source, and 1 g of ammonium sulfate as the nitrogen source, followed by growth for an additional 6 h to 7 h at 37 °C to a density of OD₆₀₀ > 1. The heme precursor *l*-amino levulinic acid was added to the culture at this point for a final concentration of 300 µM, followed by induction with 1 mM isopropyl-D thio-galactopyranoside. Protein expression was then allowed overnight for 10 h to 12 h at 37 °C. Cell pellets were harvested in the morning by centrifugation and resuspended in buffer containing 50 mM potassium phosphate and 50 mM KCl at pH 7.4 (buffer A). The resulting solution was then subjected to sonication to disrupt the cell membranes, followed by high-speed centrifugation for 30 min to remove cell debris. The supernatant was loaded onto a nickel-charged agarose affinity resin column equilibrated with buffer A to allow binding of CYP119. The column was then washed with five column volumes of buffer A, and CYP119 was eluted with buffer A containing 150 mM imidazole. Next, the elute was loaded onto an anion exchange column, washed with five column volumes of buffer A, and the protein was again eluted with buffer A containing 250 mM KCl. The eluted protein was then concentrated and further purified by size exclusion chromatography on an ÄKTA-FPLC system (GE Healthcare Life Sciences), using a S-200 column (BioRad) equilibrated with buffer A. Protein fractions coming off from the size exclusion column were collected and checked for purity by measuring the (417/280)_{nm} wavelength ratio using UV-visible spectroscopy. Fractions with the (417/280)_{nm} ratio greater than 1.4 were judged to be pure and pooled together for storage at -80 °C until needed for further experiments.

Protein Secondary Structure Characterization and Thermal Melting Curves Using CD Spectroscopy. CD measurements on protein secondary structure were conducted at 25 °C in a Jasco J-815 spectropolarimeter using a 1-mm-pathlength cell. For determination of the melting curves, signals in the far-UV CD region (222 nm) were monitored as a function of temperature. The ligand-free CYP101A1 and CYP119 protein concentrations were 5 µM in

buffer A (pH 7.4) for the various measurements. Both proteins were heated in the range of 25 °C to 100 °C with a heating rate of 0.5 °C/min, equilibrated for at least 2 min at each temperature, and the CD data were collected at 0.5 °C intervals. Reversibility of thermal unfolding was confirmed by changing the temperature by a few degrees in either direction at each acquired data point and was observed to present no hysteresis.

Protein Chemical Denaturation Studies Using Fluorescence Spectroscopy. For determination of the stability curves, samples of ligand-free CYP101A1 and ligand-free CYP119 were prepared with different guanidinium hydrochloride (GdnHCl) concentrations ranging from 0 M to 6 M. Chemical denaturation curves were collected corresponding to the various GdnHCl concentrations at temperatures ranging from 5 °C to 50 °C for CYP101A1 and 5 °C to 85 °C for CYP119. Samples containing 2 µM of each protein in buffer A were prepared in triplicate and added to PCR tubes (USA Scientific). Each sample was incubated at the set temperature for up to 24 h before the fluorescence measurement to achieve equilibrium unfolding conditions. SYPRO Orange dye available as 5,000× stock (Sigma-Aldrich) was added just before the measurement to the protein samples in an appropriate amount to achieve a final concentration of the dye of 5×. Chemical denaturation of the protein was monitored by exciting the SYPRO Orange dye at 470 nm and monitoring its fluorescence emission at 570 nm using a Chromo4 DNAEngine real-time 96-well PCR instrument. Controls included samples containing buffer only, protein plus buffer only, and dye plus buffer only, measurements from which were used for baseline correction using the Opticon Monitor software available on the PCR instrument. Normalized intensity values I_0 used in the various chemical denaturation curves were obtained by normalizing the experimental fluorescence intensity values from 0 to 1 and plotted as a function of increasing GdnHCl concentration in the range 0 M to ~6 M. The normalized fluorescence intensity values were plotted as a function of GdnHCl concentration and nonlinearly fitted to Eq. 6 to extract the unfolding free energy value $\Delta G(T)$ of CYP101A1 and CYP119 at zero denaturant concentration and a given temperature (63).

$$I_0 = (I_f + s_f \times c) + (I_u + s_u \times c) \times \frac{\exp\left[\frac{s \times c - \Delta G(T)}{RT}\right]}{1 + \exp\left[\frac{s \times c - \Delta G(T)}{RT}\right]}, \quad [6]$$

where I_0 is the normalized fluorescence intensity of the SYPRO Orange dye bound to CYP101A1 or CYP119 at a given GdnHCl concentration, c . I_f and I_u represent the intercepts, while s_f and s_u are the slopes of folded and unfolded baselines, respectively. R is the gas constant, T is the temperature, and s is the slope in the transition state. The unfolding free energy values $\Delta G(T)$ were then plotted versus temperature in the respective temperature range for each protein to obtain the protein stability curve for both CYP101A1 and CYP119 (Fig. 2A).

Elastic Incoherent Neutron Scattering Studies of CYP101A1 and CYP119. Elastic incoherent neutron scattering spectra were collected on ligand-free CYP101A1 and CYP119 samples at $h \approx 0.4$ (0.4 g water per gram protein) using the NG2 high-flux backscattering spectrometer at the Center for Neutron Research at National Institute of Standard and Technology with a fixed energy resolution of ~0.8 µeV (corresponding to a time resolution of ~1 ns). Elastic scans were performed for both protein samples in the temperature range of 4 K to 353 K at a heating rate of 1.0 K/min. No correction for multiple scattering is needed since the neutron transmission was over 0.9, and multiple scattering is thus negligible (64). The experimentally measured quantity is the so-called elastic intensity, i.e., the intensity of the peak in the dynamic structure factor, $S(q, \Delta t)$, as a function of temperature, where Δt is the instrument resolution, which is 1 ns here. MSD, $\langle x^2(\Delta t) \rangle$, is derived using the widely used Gaussian approximation that $S(q, \Delta t) = \exp(-1/6q^2 \langle x^2(\Delta t) \rangle)$ in the q range from 0.36 Å⁻¹ to 1.32 Å⁻¹ (65). The obtained MSDs are plotted in Fig. 3A.

NMR Studies of CYP101A1 and CYP119. For solution NMR spectroscopy experiments on CYP101A1 and CYP119, ¹⁵N labeled samples of both proteins were prepared in a similar fashion to that described earlier in *Protein Expression and Purification*, except that ¹⁵N-labeled ammonium sulfate was used as a nitrogen source during cell growth. The final protein concentration used for the NMR experiments on both proteins was ~0.15 mM. For the 4-phenylimidazole-bound form, 4-phenylimidazole was prepared as a 20-mM stock solution in methanol and added to the protein in small increments to achieve a final 1:1 relative molar concentration. The ¹H-¹⁵N TROSY 2D spectra (Fig. 4) were acquired using these samples in ligand-free and 4-phenylimidazole-bound forms for both CYP101A1 and CYP119 in buffer containing 50 mM potassium phosphate and 50 mM KCl at pH 7.4 at 30 °C. The resulting data were processed and analyzed with Felix2000 software (Accelrys).

ACKNOWLEDGMENTS. We thank Miranda Pratt for help with thermodynamic measurements using CD spectroscopy and Qin Xu for helpful discussions. The Shanghai Jiao Tong University (SJTU) group acknowledges support for Circular Dichroism Spectroscopy from Instrumental Analysis Center of SJTU. This work was supported by the National Science

Foundation of China Grants 11504231 and 3163002. Access to High Flux Backscattering Source was provided by the Center for High Resolution Neutron Scattering, a partnership between the National Institute of Standards and Technology and the National Science Foundation under Agreement DMR-1508249.

- Sligar SG (2010) Chemistry. Glimpsing the critical intermediate in cytochrome P450 oxidations. *Science* 330:924–925.
- Coon MJ (2005) Cytochrome P450: Nature's most versatile biological catalyst. *Annu Rev Pharmacol Toxicol* 45:1–25.
- Schuler MA, Sligar SG (2007) Diversities and similarities in P450 systems: An introduction. *The Ubiquitous Roles of Cytochrome P450 Proteins* (Wiley, New York), pp 1–26.
- Guengerich FP (2001) Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem Res Toxicol* 14:611–650.
- Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* 138:103–141.
- Cowan DA (1997) Thermophilic proteins: Stability and function in aqueous and organic solvents. *Comp Biochem Physiol A Physiol* 118:429–438.
- Kellner DG, Hung S-C, Weiss KE, Sligar SG (2002) Kinetic characterization of compound I formation in the thermostable cytochrome P450 CYP119. *J Biol Chem* 277:9641–9644.
- Sivaramakrishnan S, et al. (2011) A novel intermediate in the reaction of seleno CYP119 with m-chloroperbenzoic acid. *Biochemistry* 50:3014–3024.
- Fágán CÖ (1995) Understanding and increasing protein stability. *Biochim Biophys Acta* 1252:1–14.
- Taylor TJ, Vaisman II (2010) Discrimination of thermophilic and mesophilic proteins. *BMC Struct Biol* 10:55.
- Razvi A, Scholtz JM (2006) Lessons in stability from thermophilic proteins. *Protein Sci* 15:1569–1578.
- Kumar S, Nussinov R (2001) How do thermophilic proteins deal with heat? *Cell Mol Life Sci* 58:1216–1233.
- Urlacher VB, Girhard M (2012) Cytochrome P450 monooxygenases: An update on perspectives for synthetic application. *Trends Biotechnol* 30:26–36.
- Park S-Y, et al. (2002) Thermophilic cytochrome P450 (CYP119) from *Sulfolobus solfataricus*: High resolution structure and functional properties. *J Inorg Biochem* 91:491–501.
- Behera RK, Mazumdar S (2010) Thermodynamic basis of the thermostability of CYP175A1 from *Thermus thermophilus*. *Int J Biol Macromol* 46:412–418.
- Ho WW, Li H, Nishida CR, de Montellano PR, Poulos TL (2008) Crystal structure and properties of CYP231A2 from the thermoacidophilic archaeon *Picrophilus torridus*. *Biochemistry* 47:2071–2079.
- Oku Y, et al. (2004) Structure and direct electrochemistry of cytochrome P450 from the thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain 7. *J Inorg Biochem* 98:1194–1199.
- Colombo G, Merz KM (1999) Stability and activity of mesophilic subtilisin E and its thermophilic homolog: Insights from molecular dynamics simulations. *J Am Chem Soc* 121:6895–6903.
- Dalhus B, et al. (2002) Structural basis for thermophilic protein stability: Structures of thermophilic and mesophilic malate dehydrogenases. *J Mol Biol* 318:707–721.
- Fitter J, Heberle J (2000) Structural equilibrium fluctuations in mesophilic and thermophilic α -amylase. *Biophys J* 79:1629–1636.
- Beadle BM, Baase WA, Wilson DB, Gilkes NR, Shoichet BK (1999) Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme. *Biochemistry* 38:2570–2576.
- Yano JK, et al. (2000) Crystal structure of a thermophilic cytochrome P450 from the archaeon *Sulfolobus solfataricus*. *J Biol Chem* 275:31086–31092.
- Vihinen M (1987) Relationship of protein flexibility to thermostability. *Protein Eng* 1:477–480.
- Radestock S, Gohlke H (2008) Exploiting the link between protein rigidity and thermostability for data-driven protein engineering. *Eng Life Sci* 8:507–522.
- Chan C-H, et al. (2004) Relationship between local structural entropy and protein thermostability. *Proteins* 57:684–691.
- Miao Y, et al. (2012) Coupled flexibility change in cytochrome P450cam substrate binding determined by neutron scattering, NMR, and molecular dynamics simulation. *Biophys J* 103:2167–2176.
- Mehareenna YT, Poulos TL (2010) Using molecular dynamics to probe the structural basis for enhanced stability in thermal stable cytochromes P450. *Biochemistry* 49:6680–6686.
- Nishida CR, Ortiz de Montellano PR (2005) Thermophilic cytochrome P450 enzymes. *Biochem Biophys Res Commun* 338:437–445.
- Brandman R, Lampe JN, Brandman Y, de Montellano PRO (2011) Active-site residues move independently from the rest of the protein in a 200 ns molecular dynamics simulation of cytochrome P450 CYP119. *Arch Biochem Biophys* 509:127–132.
- Colthart AM, et al. (2016) Detection of substrate-dependent conformational changes in the P450 fold by nuclear magnetic resonance. *Sci Rep* 6:22035.
- Navrátilová V, et al. (2017) Molecular insights into the role of a distal F240A mutation that alters CYP1A1 activity towards persistent organic pollutants. *Biochim Biophys Acta* 1861:2852–2860.
- Tietz DR, Podust LM, Sherman DH, Pochapsky TC (2017) Solution conformations and dynamics of substrate-bound cytochrome P450 MycG. *Biochemistry* 56:2701–2714.
- Yasutake Y, Kameda T, Tamura T (2017) Structural insights into the mechanism of the drastic changes in enzymatic activity of the cytochrome P450 vitamin D₃ hydroxylase (CYP107BR1) caused by a mutation distant from the active site. *Acta Crystallogr F Struct Biol Commun* 73:266–275.
- Asciutto EK, Pochapsky TC (2018) Some surprising implications of NMR-directed simulations of substrate recognition and binding by cytochrome P450_{cam} (CYP101A1). *J Mol Biol* 430:1295–1310.
- Hu S, et al. (2010) Altering the regioselectivity of cytochrome P450 BM-3 by saturation mutagenesis for the biosynthesis of indirubin. *J Mol Catal B Enzym* 67:29–35.
- Hong L, et al. (2016) Determination of functional collective motions in a protein at atomic resolution using coherent neutron scattering. *Sci Adv* 2:e1600886.
- Hernández G, Jenney FE, Jr, Adams MWWW, LeMaster DM (2000) Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proc Natl Acad Sci USA* 97:3166–3170.
- Jaenicke R (2000) Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *Proc Natl Acad Sci USA* 97:2962–2964.
- Danculescu C, Ladenstein R, Nilsson L (2007) Dynamic arrangement of ion pairs and individual contributions to the thermal stability of the cofactor-binding domain of glutamate dehydrogenase from *Thermotoga maritima*. *Biochemistry* 46:8537–8549.
- Lee Y-T, Wilson RF, Rupniewski I, Goodin DB (2010) P450cam visits an open conformation in the absence of substrate. *Biochemistry* 49:3412–3419.
- Yano JK, et al. (2003) Preliminary characterization and crystal structure of a thermostable cytochrome P450 from *Thermus thermophilus*. *J Biol Chem* 278:608–616.
- Krishna SN, et al. (2013) A fluorescence-based thermal shift assay identifies inhibitors of mitogen activated protein kinase kinase 4. *PLoS One* 8:e81504.
- Niesen FH, Berglund H, Vedadi M (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* 2:2212–2221.
- Rasmussen BF, Stock AM, Ringe D, Petsko GA (1992) Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* 357:423–424.
- Basudhar D, et al. (2015) Analysis of cytochrome P450 CYP119 ligand-dependent conformational dynamics by two-dimensional NMR and X-ray crystallography. *J Biol Chem* 290:10000–10017.
- Lampe JN, et al. (2008) Ligand-induced conformational heterogeneity of cytochrome P450 CYP119 identified by 2D NMR spectroscopy with the unnatural amino acid (13)C-p-methoxyphenylalanine. *J Am Chem Soc* 130:16168–16169.
- Pochapsky TC, Kazanis S, Dang M (2010) Conformational plasticity and structure/function relationships in cytochromes P450. *Antioxid Redox Signal* 13:1273–1296.
- Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. *Proc Natl Acad Sci USA* 103:5869–5874.
- Puchkaev AV, Koo LS, Ortiz de Montellano PR (2003) Aromatic stacking as a determinant of the thermal stability of CYP119 from *Sulfolobus solfataricus*. *Arch Biochem Biophys* 409:52–58.
- Maves SA, Sligar SG (2001) Understanding thermostability in cytochrome P450 by combinatorial mutagenesis. *Protein Sci* 10:161–168.
- Lazaridis T, Archontis G, Karplus M (1995) Enthalpic contribution to protein stability: Insights from atom-based calculations and statistical mechanics. *Advances in Protein Chemistry*, eds Anfinsen CB, Richards FM, Edsall JT, Eisenberg DS (Academic, New York), Vol 47, pp 231–306.
- Sawle L, Ghosh K (2011) How do thermophilic proteins and proteomes withstand high temperature? *Biophys J* 101:217–227.
- Pica A, Graziano G (2016) Shedding light on the extra thermal stability of thermophilic proteins. *Biopolymers* 105:856–863.
- Matthews BW, Nicholson H, Becktel WJ (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci USA* 84:6663–6667.
- Liu C-C, LiCata VJ (2014) The stability of Taq DNA polymerase results from a reduced entropic folding penalty; identification of other thermophilic proteins with similar folding thermodynamics. *Proteins* 82:785–793.
- Butterwick JA, et al. (2004) Multiple time scale backbone dynamics of homologous thermophilic and mesophilic ribonuclease HI enzymes. *J Mol Biol* 339:855–871.
- Tehei M, Madern D, Franzetti B, Zaccai G (2005) Neutron scattering reveals the dynamic basis of protein adaptation to extreme temperature. *J Biol Chem* 280:40974–40979.
- Meinhold L, et al. (2008) Protein dynamics and stability: The distribution of atomic fluctuations in thermophilic and mesophilic dihydrofolate reductase derived using elastic incoherent neutron scattering. *Biophys J* 94:4812–4818.
- Missimer JH, et al. (2007) Configurational entropy elucidates the role of salt-bridge networks in protein thermostability. *Protein Sci* 16:1349–1359.
- Seewald MJ, et al. (2000) The role of backbone conformational heat capacity in protein stability: Temperature dependent dynamics of the B1 domain of streptococcal protein G. *Protein Sci* 9:1177–1193.
- Rader AJ (2002) Protein rigidity and flexibility: Application to folding and thermostability. Doctor of Philosophy Doctoral thesis (Michigan State Univ, Ann Arbor, MI).
- Rader AJ (2009) Thermostability in rubredoxin and its relationship to mechanical rigidity. *Phys Biol* 7:16002.
- Santoro MM, Bolen DW (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27:8063–8068.
- Khodadadi S, et al. (2010) Dynamics of biological macromolecules: Not a simple slaving by hydration water. *Biophys J* 98:1321–1326.
- Gabel F, et al. (2002) Protein dynamics studied by neutron scattering. *Q Rev Biophys* 35:327–367.