This is a repository copy of *Differential effects of hydrophobic core packing residues for thermodynamic and mechanical stability of a hyperthermophilic protein.*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/102761/

Version: Accepted Version

Differential effects of hydrophobic core packing residues for thermodynamic and mechanical stability of a hyperthermo-philic protein

Katarzyna M. Tych†,‡, Matthew Batchelor†,‡,‖, Toni Hoffmann†,‡,‖, Michael C. Wilson‡,‖, Megan L. Hughes†,‡, Emanuele Paci‡,‖ David J. Brockwell‡,‖ and Lorna Dougan†,‡*

†School of Physics and Astronomy, University of Leeds, Leeds, LS2 9JT, ‡Astbury Centre for Structural and Molecular Biology, University of Leeds, Leeds, LS2 9JT, ‖School of Molecular and Cellular Biology, University of Leeds, Leeds, LS2 9JT

ABSTRACT Proteins from organisms which have adapted to environmental extremes provide attractive systems to explore and determine the origins of protein stability. Improved hydrophobic core packing and decreased loop-length flexibility can increase the thermodynamic stability of proteins from hyperthermophilic organisms. However, their impact on hyperthermophilic protein mechanical stability is not known. Here, we use protein engineering, biophysical characterization, single molecule force spectroscopy (SMFS) and molecular dynamics (MD) simulations to measure the effect of altering hydrophobic core packing on the stability of the cold shock protein TmCSP from the hyperthermophilic bacterium Thermotoga maritima. We make two variants of TmCSP in which a mutation is made to reduce the size of aliphatic groups from buried hydrophobic side chains. In the first, a mutation is introduced in a long loop (TmCSP L40A); in the other, the mutation is introduced on the C-terminal β-strand (TmCSP V62A). We use MD simulations to confirm that the mutant TmCSP L40A shows the most significant increase in loop flexibility, and mutant TmCSP V62A shows greater disruption to the core packing. We measure the thermodynamic stability (ΔG_{D-N}) of the mutated proteins and show there is a more significant reduction for TmCSP L40A ($\Delta\Delta G$ = 63%) than TmCSP V62A ($\Delta\Delta G$ = 47%) as might be expected, based on the relative reduction in the size of the side chain. By contrast SMFS measures the mechanical stability (ΔG^*) and shows a greater reduction for TmCSP V62A ($\Delta\Delta G^*$ = 8.4%) than TmCSP L40A ($\Delta\Delta G^*$ = 2.5%). While the impact on mechanical stability is subtle, the results demonstrate the power of tuning non-covalent interactions to modulate both the thermodynamic and mechanical stability of a protein. Such understanding and control provides the opportunity to design proteins with optimized thermodynamic and mechanical properties.

INTRODUCTION

Proteins from organisms adapted to high-temperatures have evolved to retain their native, folded structure and function in the challenging environments in which the organisms grow.[1, 2] These heat-adapted organisms are classified according to the optimal growth temperature, T_{OPT}, of the organism, with thermophiles having a T_{OPT} between ~ 45 and 80 °C and hyperthermophiles a T_{OPT} above ~ 80 °C.[3, 4] The most thermostable proteins are found within these latter organisms, at or near the upper temperature limits of life.[5] Many of these proteins show melting temperatures (T_M) in excess of 100 °C, and as high as 200 °C, and retain activity for periods of hours at these temperatures.[6] The discovery of thermophiles and hyperthermophiles has stimulated a wealth of fundamental and applied research into their proteins, particularly their enzymes.[7-10] This research has included efforts to understand the molecular basis of protein thermostability at high temperatures and the potential applications of these proteins.[7, 9]

Like their mesophilic counterparts, proteins from thermophiles and hyperthermophiles are only marginally stable at the physiological temperature of the organisms in which they are

1

found. There is little evidence that there is one single molecular-level adaptation for increasing the stability of proteins from thermophiles and hyperthermophiles compared with those from mesophiles (organisms with moderate optimal growth temperatures, between 20 and 45 °C).[11] The thermal stabilization of a protein can be achieved in a number of ways including: increased numbers of ionic interactions and ionic networks,[3, 12-19] increased packing density,[20] increased hydrophobicity[3, 21-25] and a reduction in the length of flexible, loop regions in the protein.[26, 27] However, there are examples of particular strategies being found in structurally related proteins. For example, a study of 373 protein families examined the importance of different non-covalent interactions through comparisons between mesophilic and thermophilic homologues.[21] This study predicted that the optimization of the hydrophobic core is the most significant contribution to the enhanced stability of thermophilic proteins, where the 'average surrounding hydrophobicity' of a protein was defined as the sum of hydrophobic indices obtained from thermodynamic transfer experiments. Using this criteria they found that 80% of the thermophilic proteins studied displayed higher hydrophobicity than their mesophilic equivalents. In another study, a protein structure dataset was constructed from one specific organism, the hyperthermophilic bacterium Thermotoga maritima, and compared with those of close homologs from mesophilic bacteria.[20] The study found that the proteins from Thermotoga maritima had an increased number of salt-bridges and were more compact than the mesophilic proteins. .

Homologous proteins from hyperthermophilic and mesophilic organisms therefore offer important model systems with which to investigate the importance of specific non-covalent interactions on protein stability.[28-30] We recently took this approach to examine the role of salt bridges in the stability of a homologous pair, the cold shock protein B (TmCSP) from the hyperthermophilic organism Thermotoga maritima, also used in this study, and the cold shock protein (BsCSP) from the mesophilic organism Bacillus subtilis. We determined that TmCSP, which has many more salt bridges than BsCSP, is thermodynamically more stable. Using single molecule force spectroscopy (SMFS) we also showed that TmCSP has increased mechanical softness.

Given the insight gained by this approach, we now use this model system to examine the role of hydrophobicity and loop flexibility on the stability of the hyperthermophilic cold stock protein TmCSP. Improved hydrophobic packing and therefore reduced solvent accessibility of hydrophobic residues will be entropically favorable. This is in part due to the lower degree of ordering of water molecules that occurs when these hydrophobic side chains are removed from the solvent. An additional contribution may come from enthalpically favorable interactions if there is a concomitant increase in the van der Waals contributions to the hydrophobicity. Previous studies have demonstrated that an increase in the number of hydrophobic interactions per amino acid residue can increase the thermostability of a protein[20, 21, 31] and have successfully demonstrated the impact of increased hydrophobic core packing in enhancing protein thermostability.[31] Through their inherent flexibility, loops are also considered to be potential initiation sites for thermal denaturation.[32] A reduction in loop length or loop flexibility may therefore contribute to enhancing protein kinetic stability.[33] In addition to thermodynamic stability, previous SMFS studies have examined the impact of hydrophobic core packing on the mechanical stability of mesophilic proteins including; the I27 domain of titin,[34, 35] TNfn3 from tenascin-C,[36] protein L,[37] and protein GB1.[38] These studies have shown that hydrophobic core packing, modulated by changing the length of amino acid side chains, plays a potentially important role in the mechanical unfolding of some proteins. SMFS is therefore an attractive tool with which to measure the contribution of hydrophobic interactions on the mechanical stability.

We elected to use the cold shock protein TmCSP as a model system. We made two TmCSP mutants in which side chain deletions allowed us to probe the effects of subtle changes in hydrophobic core packing and loop flexibility. Using a combination of SMFS, fluorescence spectroscopy and molecular dynamics (MD) simulations we examine and compare the two TmCSP mutants with benchmarked TmCSP.[28, 30] We obtain information about their thermodynamic and mechanical stability and identify, as well as quantify, the non-covalent interactions in the proteins using MD. Our experiments reveal insight into the importance of hydrophobic interactions in determining protein stability, and highlight the advantages of using proteins from extremophilic organisms as model systems.

MATERIALS AND METHODS

Protein Engineering and Expression

Polyproteins were constructed using a method which makes use of Gibson Assembly cloning[39] and purified using a method described previously,[28] including an additional stage to remove any bound nucleic acid.[29] Three $(His)_6$-tagged chimeric polyprotein constructs, each containing four domains of I27 interdigitated with three domains of a CSP were produced: (i) a polyprotein containing the wild-type CSP from the hyperthermophilic organism Thermotoga maritima (TmCSP), (I27-TmCSP)$_3$-I27 (ii) a polyprotein containing the L40A variant of CSP which we refer to as TmCSP L40A, (I27-TmCSP L40A)$_3$-I27 and (iii) a polyprotein containing the V62A variant of the CSP which we refer to as TmCSP V62A, (I27-TmCSP V62A)$_3$-I27. The $(His)_6$ tag, inter-domain linker sequences, I27 domains and two cysteine residues at the C-terminus are identical in all three polyprotein constructs.

Protein Thermodynamic Stability

Chemically and thermally induced unfolding transitions of the cold shock proteins TmCSP, TmCSP L40A and TmCSP V62A were measured using a PTI fluorimeter (Photon Technology International, UK) with a Peltier temperature controller and an LPS-100 lamp. Protein samples (0.1 mg ml^{-1} in 63 mM sodium phosphate buffer pH 7.4 containing different concentrations of GdnHCl) were equilibrated at 7 °C (thermal unfolding) and 23 °C (chemical denaturation) overnight before measurements were recorded. Fluorescence spectra were measured in a 1 cm pathlength quartz cuvette using an excitation wavelength of 280 nm and emission range of 320–380 nm with a 1 nm step size. Unfolding transitions were followed by a change in the barycentric median (BCM) as described previously. {Tych, 2016 #196} The BCM 'center of mass' of each spectrum between 320 nm and 380 nm was calculated where $I(\lambda)$ is the fluorescence value at a respective wavelength.

The BCM value for each spectrum was plotted against temperature (T) or denaturant concentration ([GdnHCl]) and the

unfolding transition followed by an increase in BCM due to a shift to a higher wavelength of the unfolded peak.

Chemical equilibrium curves were fitted to a two-state unfolding model as described previously {Tych, 2016 #196}.

For the thermal denaturation curves, 0.1 mg ml^{-1} protein samples were prepared in 63 mM sodium phosphate buffer pH 7.4 at a range of GdnHCl concentrations below the mid-point of protein unfolding. Samples were heated using a 3 °C stepped gradient with four minutes equilibration before each spectrum was recorded, equating to an average temperature (T) increase rate of 0.4 °C/min. Thermal unfolding curves were fitted to an integrated van't Hoff equation (equation 1) where a_F and a_U represent the signal at the start and end of the run and b_F and b_U represent the rate of change of signal with temperature ([T]) in the pre-transitional and post-transitional baselines.[41] Q is the quantum yield, R is the ideal gas constant, ΔH is the change in enthalpy and T_M is the temperature at which 50% of protein is folded.

$$Eqn\ (1): f(T) = \frac{(a_F + b_F T)(\frac{1}{Q})e^{\frac{-\Delta H}{R}(\frac{1}{T_M} - \frac{1}{T})} + (a_U + b_U T)}{1 + (\frac{1}{Q})e^{\frac{-\Delta H}{R}(\frac{1}{T_M} - \frac{1}{T})}}$$

The equation yields values for T_M and ΔH of the transition from each unfolding experiment. The pairs of values for ΔH and T_M of the unfolding transition for each concentration of GdnHCl were plotted and a weighted linear fit to these data used to determine the change in specific heat capacity between folded and unfolded states, ΔC_p. The values of ΔC_p and the average values of ΔH and T_M across three thermal denaturations in the absence of GdnHCl were inserted into the Gibbs–Helmholtz equation (equation 2) using values of T in 3 °C intervals, to produce a thermal stability curve.[5, 42, 43]

$$Eqn\ (2): \Delta G = \Delta H \left(1 - \frac{T}{T_M}\right) - \Delta C_p [(T_M - T) + T ln\left(\frac{T}{T_M}\right)]$$

Force Spectroscopy
SMFS experiments were completed using an Asylum MFP-3D AFM (Asylum Research, Santa Barbara, CA, USA). Silicon nitride cantilevers (MLCT) were obtained from Bruker (Billerica, MA, USA). Before each experimnets the spring constant of the cantilever was calibrated in buffer, using the equipartition theorem method,[44] and was found to be within the range of 38 (± 3) pN nm^{-1}. Lyophilized protein (0.1 mg) was reconstituted to a concentration of 0.2 mg ml^{-1} in sterile sodium phosphate buffer (63 mM, pH 7.4) and incubated on a gold substrate for 10 min. Mechanical unfolding experiments were completed at constant at pulling velocities of 100, 200, 600 and 2000 nm s^{-1} at room temperature (23 °C) over a distance of 400 nm. At each pulling velocity three different datasets, each using a different calibrated cantilever and sample, were obtained. Each dataset contained more than 34 total unfolding events. For the subsequent data analysis traces were only included which contained one polyprotein chain unfolding, characterized by there being seven or fewer unfolding events. Of these, only traces with a minimum of two I27 unfolding events, without non-specific unbinding events at high force, or other sources of noise were used. Given the interdigitated nature of the polyprotein, the presence of two I27 unfolding events ensured that force would be applied to at least one CSP domain.[28, 30] The force-extension data were subsequently analyzed using in-house software written for Igor Pro.

Kinetics of protein mechanical unfolding
The Zhurkov–Bell model was used to model mechanical unfolding, by assuming that each protein unfolds via a two-state all-or-none process governed by a rate constant, k_U, and the distance from the native state to the transition state along the measured reaction co-ordinate, Δx_U.[45]

$$Eqn\ (3): k_U(F) = A\exp\frac{-(\Delta G - F\Delta x_U)}{k_B T}$$

where $k_u(F)$ is the force-dependent rate constant, F is the applied force, A is the attempt frequency, Δx_U is the distance from the folded state to the transition state, k_B is Boltzmann's constant and T the temperature. The value of Δx_U is the distance between the folded state and the transition barrier on the unfolding pathway. ΔG_U is the height of the activation energy barrier to unfolding and can be related to the unfolding rate at zero force by,

$$Eqn\ (4): k_U = A\exp\frac{-(\Delta G_U)}{k_B T}$$

Monte Carlo (MC) simulations were completed to produce unfolding forces at each pulling velocity. These were used to create simulated unfolding force which were then compared to those generated experimentally. A straight line was fitted to the simulated F_U–ln(pulling speed) dependence to compare with the experimentally determined line of best fit. The pair of k_U and Δx_U values that provided the best global fit to the experimental data over all pulling velocities was obtained. The range of k_U and Δx_U values that provided a fit to the data within the experimental uncertainty was used to calculate the standard deviation for each parameter. In the simulations, due to the similarity in slope of the F_U–pulling speed dependence for both I27 and the CSP for each of the three different constructs, it was assumed that the Δx_U for the proteins were unchanged. These values were set to $\Delta x_U = 0.70$ nm for the CSPs and $\Delta x_U = 0.32$ nm for I27 based on a recent study.[11b]

MD Simulations
The behavior of TmCSP, TmCSP L40A and TmCSP V62A was simulated using the CHARMM param36 force field and explicit solvent. Initial structures from which to start simulations of TmCSP used the PDB structure 1G6P (model #1). TmCSP L40A and TmCSP V62A starting structures were created by manual deletion of side chain atoms from residues L40 or V62, respectively, from the 1G6P structure, followed by incorporation of the missing hydrogen atom using CHARMM. [46] A steepest descent minimization (1000 steps) was then performed for each of the three proteins. Proteins were solvated in a water box containing a 1.2 nm surround of water molecules (~4400) and NaCl ions were added at a concentration of ~50 mM using VMD.[47] NAMD[48] was then used to run simulations: a short heating protocol (0–300 K) followed by 0.2 ns of equilibration preceded a 400 ns simulation for each protein at 300 K. For analysis, the coordinates of all atoms were recorded every 1 ps. The simulation of TmCSP is an extension of the explicit solvent simulation previously published by us.[11b]

Constant velocity protein unfolding simulations that mimic AFM experiments were performed using an implicit solvent model (FACTS) and the united-atom force field CHARMM19. [49] This simplified model was used to ensure that solvent relaxation was not a dominating factor during the

kinetically controlled extension of the protein. The constant velocity protein unfolding simulations were run using CHARMM, applying Langevin dynamics at a temperature of 300 K. [46] The friction coefficient was 3 ps-1. The surface tension-like parameter was 0.025 kcal mol-1 Å-2 (the recommended value for modeling globular proteins with FACTS[49]). In these simulations an external force is applied between the N-terminal N atom and C-terminal C atom. The attached cantilever is moved away at constant velocity (v = 108 nm s⁻¹). To mirror the AFM experiments, the cantilever spring constant, kc, was set to 30 pN nm⁻¹. For all three proteins, 50 pulling simulations were run, each starting from a different initial structure extracted from a short equilibrium simulation (with no applied load). For analysis, the coordinates of all atoms were recorded every 1 ps.

Wordom was used to analyze the trajectories.[51] 'DSSPcont' assignments[52] were used to calculate secondary structure content on a per residue basis. VMD was used to locate and quantify hydrogen bonds between the different parts of the protein, using the criteria: N–O distance < 4 Å, and N–H–O angle >150°. Salt bridges were also located and quantified using VMD, [47] using the criterion of whether the distance between the center-of-mass of side chain N or O atoms was <0.7 nm apart, allowing for salt bridges to be separated by a gap the size of one water molecule. [53]

The program 'NACCESS' was used to ascertain which residues make up the hydrophobic core of the protein.[54, 55] The solvent accessibility (probe size 0.14 nm) of side-chain atoms within each residue from 4000 snapshots in each simulation was calculated and averaged. Using the default setting, Cα atoms were included as part of the side chain, but glycine residues were excluded from analysis. The contact order and relative contact order were calculated using a script kindly provided by Dmitry Ivankov.[56] Contacts made by core residue side-chains were analyzed by outputting the distance between the side-chain of each core residue with the side-chain of all other

residues in the sequence. Side chain positions were described using the center-of-mass of all side chain atoms plus the Cα atom. A cut-off value of 0.55 nm gave a good range of values for fraction of time in contact across the peptide serving to highlight important residues.

RESULTS

Design of TmCSP variants to probe the effect of hydrophobic core packing and loop flexibility on the thermodynamic and mechanical stability of TmCSP. The TmCSP protein (Fig. 1(a)) is composed of five β -strands organised in an anti-parallel manner, forming two β -sheets connected by a loop region. The amino acid sequence of of TmCSP is shown in Fig. 1(b). We elected to make conservative mutations such that; the mutation would be unlikely to change the protein structure, would conserve the chemical nature of the residue and not introduce any new interactions. This involved the reduction in the size of aliphatic groups from residues which contained buried hydrophobic side chains. For this type of mutation, any change in energy of the transition state is related to the degree of native structure formation close to the mutated residue. We chose two positions at which to make conservative side chain deletions (Fig. 1(a, b)). At both positions the two residue side-chains are buried in the PDB structures of TmCSP (see Supplementary text) but were hypothesized (based on our previous studies[30]) to be disrupted at different points in the mechanical unfolding pathway. Both mutations are located in different regions of TmCSP and were selected in order to determine their impact on the thermodynamic and mechanical unfolding of TmCSP. In the protein variant TmCSP L40A, the leucine at position 40, (located in a loop), is mutated to an alanine. This results in a reduction in the side-chain length, replacing -CH(CH₃)₂ with -H. In the second variant TmCSP V62A the valine at position 62 (located in the fifth β-strand) is mutated to an alanine (Fig. 1(a, b)). This results in a less dramatic reduction in side-chain length, replacing two -CH₃ groups with two hydrogens.
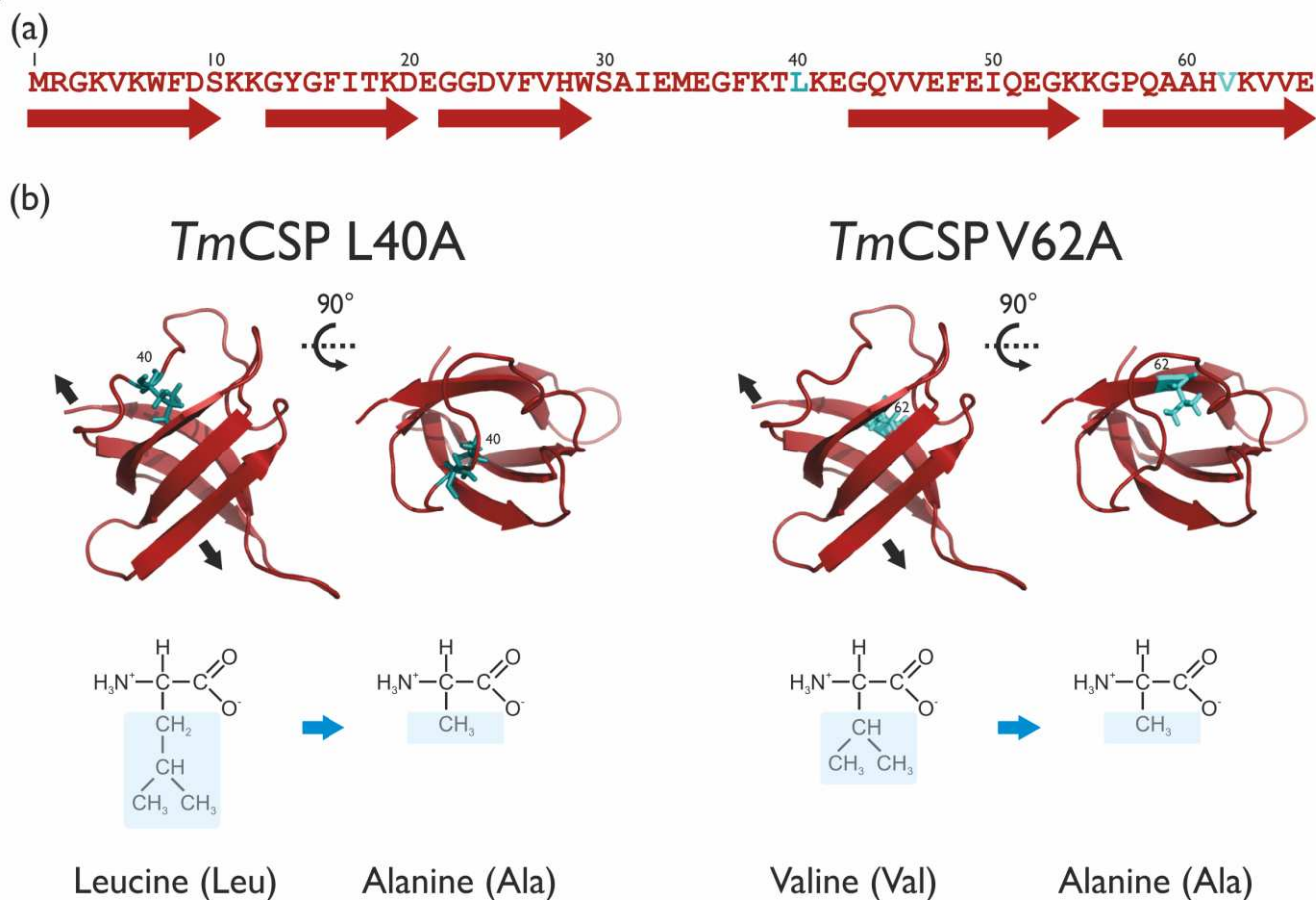
Figure 1. (a) Primary amino acid sequence of TmCSP highlighting the positions of the two mutations, L40A (dark turquoise) and V62A (light turquoise). (b) The two variant of the cold shock protein B from the hyperthermophilic bacterium Thermotoga maritima (TmCSP, PDB code 1G6P) and the positions of the mutations L40A (dark turquoise) and V62A (light turquoise). A side and top view are shown. The closed structure of the CSP β-barrel is formed when the two β-sheets twist around one another with contacts made between β-strand 1 and β-strand 4 through hydrogen bonding. For information, the direction of the applied force in single molecule force spectroscopy experiments is shown as black arrows (c) In the mutant L40A a leucine in position 40 is mutated to an alanine and in V62A a valine in position 60 is mutated to an alanine.

Molecular dynamics simulations measure hydrophobic core packing and loop flexibility. Native state equilibrium simulations show that all three proteins remain stable during the lifetime of the simulations. The β-strand structure remains largely unaffected upon mutation while V62A shows loss in the small helix that follows β-strand 3 in TmCSP (Fig. S1 and S2 and Table S1). The residues that make up the hydrophobic core of the protein were monitored throughout the simulation by analyzing the solvent exposure of the side-chain atoms. The results of this analysis are shown in Fig. 2(a). For clarity, the residues are separated into three groups: the 'hard-core' of the protein, those residues with very limited solvent accessibility (<1%); the 'soft-core', those residues with a 1–10% accessible side-chain; and more exposed residues, those having a >10% accessible side-chain. The number of hard and soft-core residues for each protein is: TmCSP – 8 hard-core, 3 soft-core, TmCSP L40A – 6 hard-core, 4 soft-core and TmCSP V62A – 5 hard-core, 5 soft-core. The change in the number of hard- and soft-core residues in each mutant relative to the wild-type protein resulted directly from the mutation performed (see Supporting Information for full details). Overall, the simulations indicate that reducing the size of the hydrophobic residue (V62A or L40A) destabilizes its position within the core making it more accessible to solvent during the simulations com-

pared to the wild-type protein. An analysis was made of numbers of hydrogen bonds (HBs) between the different regions of the protein (i.e. the 5 β-strands and the long loop that incorporates the small helix, see Fig. 2b). For the three proteins, the numbers of HBs between β-strands were largely the same within error. However, L40A showed noticeably fewer HBs on average between β-strands 1–2 than in TmCSP. TmCSP also showed more HBs on average (3 HBs) between the long loop and β5 than in either of the two mutants (2 HBs). The pattern of salt bridges within the proteins was very similar, albeit the partially occupied salt bridge E33-K63 in TmCSP, also between the long loop and β5, was disrupted in both mutants (Fig. S3).

Using the radius of gyration (R_{gyr}) as a measure of protein packing (Table S1), the simulations indicate that while TmCSP and TmCSP L40A have similar values, TmCSP V62A is less tightly packed. Looking at contact order values (Table S1), TmCSP and TmCSP L40A again have similar values, while TmCSP V62A shows a reduction in contact order. To compare flexibility within the proteins we measured the root mean square fluctuations (RMSF) of Cα atoms from each residue in the three cold-shock proteins (Fig. 2c). Here, a large RMSF

5

value indicates a more flexible structure. All three proteins show regions of the sequence with comparatively high RMSF, which relate to the loop regions between successive β-strands. Both mutants show greater flexibility, compared to TmCSP, with the most enhanced flexibility seen in the long loop region (and also in the short loop between strands 4 and 5) in TmCSP L40A. V62A shows slightly higher RMSF values across the N-terminal half of the protein. In line with the hydrogen bond/salt bridge data this feature suggests that in TmCSP the long loop is better constrained than in either mutant.
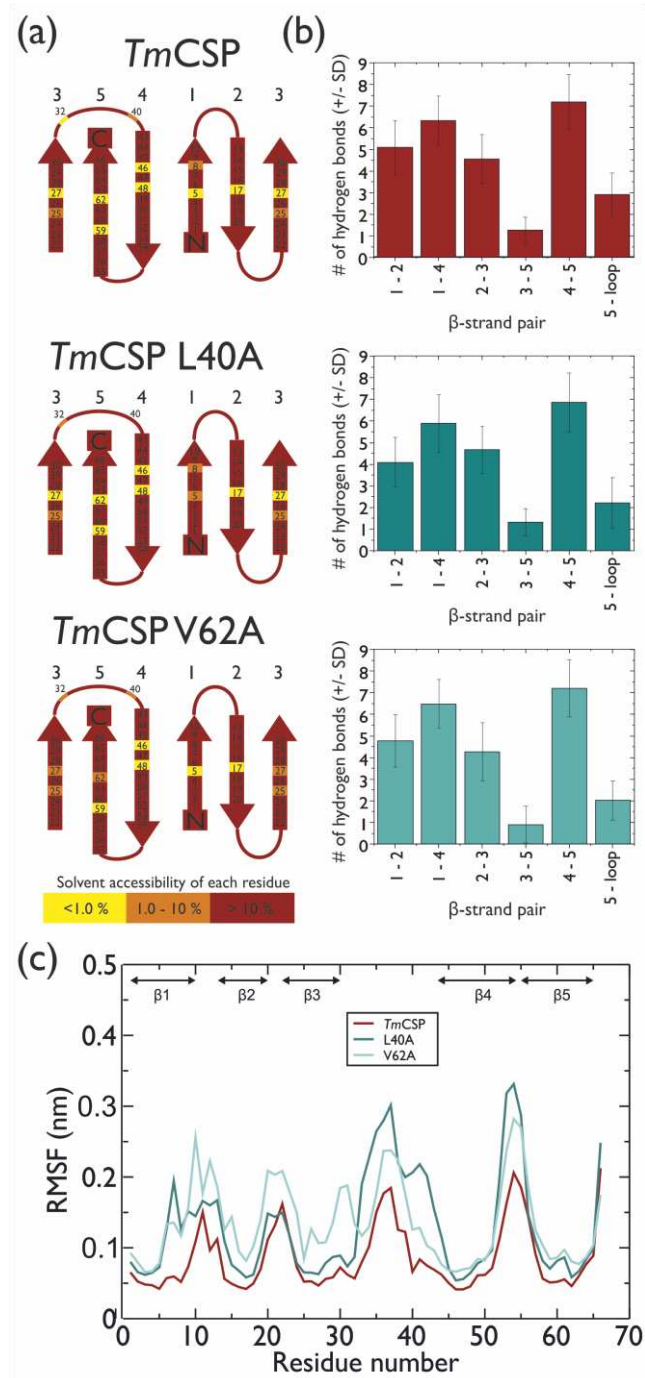

Figure 2. Results from MD simulations. (a) Topology diagram of the three cold shock protein domains TmCSP (top) TmCSP L40A (middle) and TmCSP V62A (bottom) highlighting the β-strands (labelled 1-5) and the solvent accessibility of each residue with a probability of <1% (yellow) 1–10% (orange), and >10% (red) averaged over the simulations. Note, β-strand 3 is repeated to show the connectivity in the protein (b) The mean number of hydrogen bonds ± the standard deviation (SD, shown as error bars) during the simulation between selected regions in TmCSP (red), TmCSP V62A (light turquoise) and TmCSP L40A (dark turquoise). (c) Comparison of the root mean square fluctuations (RMSF) of Cα atoms from each residue in simulations of TmCSP and the two hydrophobic core mutants (L40A and V62A). Positions of the β-strands are indicated.

We determined the contacts made by residues 40 and 62 with other residues during the simulations of all three proteins (Fig. S4 and S5). In TmCSP, the side chain of residue L40, which is in the long loop, acts to link together the edge of the long loop with strand 1. This pattern is also seen in the mutant TmCSP V62A. In the TmCSP L40A mutant, however, the contacts made by A40 are reduced leaving only the (sequence-wise) neighboring residue T39 as a significant contact. In TmCSP, the side chain of residue V62, which is on β-strand 5, provides links to strand 4 and the loop, and also weakly to strand 3. In the TmCSP L40A mutant, these contacts remain largely unchanged. Surprisingly, in the TmCSP V62A mutant the much smaller alanine side chain is still able to maintain a very similar range of contacts. This feature suggests that the structure adapts to maintain contacts with residues across the sequence despite the loss in atoms. On visual inspection of the simulation trajectory, and in accord with the helicity results, the contacts (particularly between A62 and I32) are maintained by loss in the helicity of residues 29–31.

Thermodynamic stability curves of TmCSP, TmCSP L40A and TmCSP V62A. To determine the impact of side-chain reduction on the thermodynamic stability of TmCSP we completed thermal unfolding titrations in the presence of 0–3 M guanidine hydrochloride. We obtained the temperature dependence of the thermodynamic stability (ΔG_{D-N}) for TmCSP (red), TmCSP L40A (dark turquoise) and TmCSP V62A (light turquoise) (Fig. 3). Inspection of the results shows that the side-chain reduction in the variants TmCSP L40A and TmCSP V62A both result in a lower melting temperature, T_M, and a lower overall thermodynamic stability (ΔG_{D-N}) over all temperatures. The T_M decreased from 81.9 °C for TmCSP to 60.9 °C for TmCSP V62A and to 53.8 °C for TmCSP L40A (Figure 3, circles). While both of the CSP variants are maximally stable just below room temperature, similar to the hyperthermophilic TmCSP, their maximal stabilities are considerably smaller than that of TmCSP. At 23 °C, the changes in ΔG_{D-N} on mutation are 16 and 11 kJ mol^{-1} for TmCSP L40A and TmCSP V62A, respectively (Table S2).
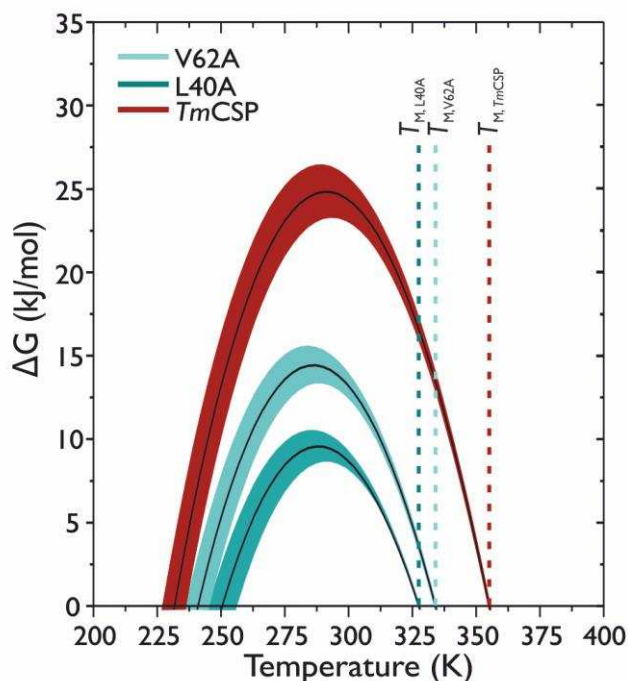
Figure 3. Protein stability curves for TmCSP (top) TmCSP L40A (middle) and TmCSP V62A, showing the change in Gibbs free energy for protein unfolding (ΔG_{D-N}) as a function of temperature (T). ΔG_{D-N} (T) was calculated as described (Materials and Methods). The errors in ΔG_{D-N} were calculated using the experimental errors in the enthalpy, difference in

specific heat capacity and melting temperature for each protein. The dashed vertical lines highlight the melting temperature, T_m, of each protein

The measured enthalpy, ΔH, of the proteins decreased from 270.62 ± 3.7 kJ mol^{-1} in TmCSP to 196.7 ± 4.0 kJ mol^{-1} in TmCSP V62A and even lower to 168.3 ± 5.8 kJ mol^{-1} for TmCSP L40A. This reduction in ΔH is likely to be a measure of the impact of side-chain reduction on the native-state interactions. While TmCSP has a section of the loop region with reduced solvent accessibility that is part buried within the hydrophobic core, this contribution to the hydrophobic interaction is no longer present in the variant TmCSP L40A. This is consistent with the hydrogen bonding patterns and positional fluctuations observed in the simulations (Fig 2). In addition, the contribution to the hydrophobic interaction is also reduced for V62A (i.e. residue 62 is shifted from being in the 'hard core' 0–1% solvent accessibility to being in the 'soft core' 1–10% solvent accessibility, Fig. 2a). The reduction of hydrophobic interactions in the native state of TmCSP L40A and TmCSP V62A is concomitant with a significant reduction in the thermodynamic stability (Fig. 3). This points to a relation between hydrophobic interactions and thermodynamic stability in this hyperthermophilic protein, in agreement with previous studies. The measured change in heat capacity at constant pressure, ΔC_p, is similar for all three proteins: 3.86 ± 0.10 kJ mol^{-1} K^{-1} in TmCSP, 3.89 ± 0.17 kJ mol^{-1} K^{-1} in TmCSP V62A, and 3.81 ± 0.26 kJ mol^{-1} K^{-1} in TmCSP L40A (Fig. S6). These experiments all three proteins are stable and folded at room temperature. Next we measured their mechanical properties by SMFS.
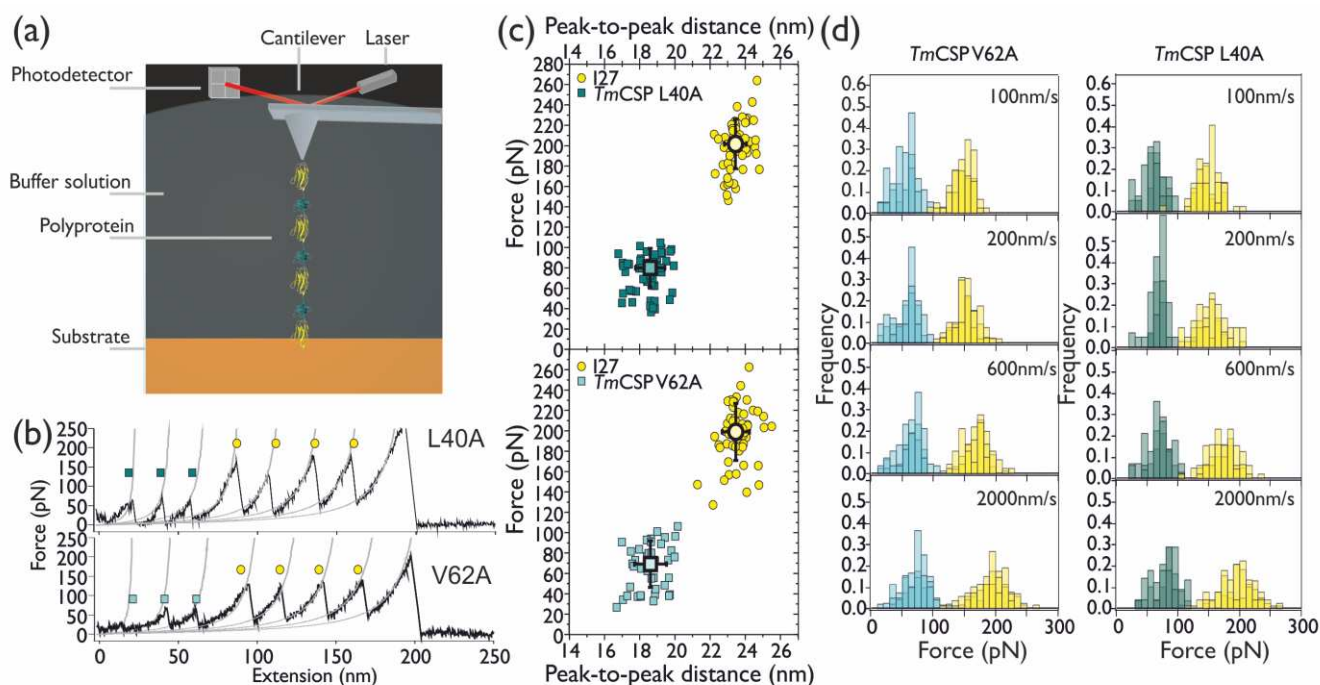

Figure 4. SMFS mechanically unfolds a mutated cold shock protein . (a) The schematic shows an AFM cantilever picking up a polyprotein immobilized on a gold substrate in solution. ((I27-TmCSP L40A)$_3$-I27) contains four I27 domains (yellow) and three TmCSP L40A domains (dark turquoise). (b) Example force-extension data shows the mechanical unfolding of the polyprotein at a constant velocity of 200 nm s^{-1}. In both examples, three CSP proteins unfold first (turquoise squares) followed by the subsequent unfolding of four I27 proteins (yellow circles) for the TmCSP L40A (upper trace) and TmCSP V62A (lower trace). (c) The CSP- and I27-specific unfolding forces and inter-peak distances are displayed as scatter plots for the ((I27-TmCSP L40A)$_3$-I27) polyprotein (upper scatter plot) and ((I27-TmCSP V62A)$_3$-I27) polyprotein (lower scatter plot) both at a constant velocity of 2000 nm s^{-1}.

50 TmCSP L40A unfolding events (dark turquoise squares) and 54 I27 unfolding events (yellow circles), and 39 TmCSP V62A unfolding events (light turquoise squares) and 51 I27 unfolding events (yellow circles). (d) The measured unfolding forces are displayed as histograms at pulling speeds of 100, 200, 600 and 2000 nm s^{-1} for the ((I27-TmCSP V62A)$_3$-I27) polyprotein (left) and ((I27-TmCSP L40A)$_3$-I27) polyprotein (right). At each pulling velocity three different histograms are shown, obtained from the triplicate experiments. The histograms of unfolding forces for I27 and the TmCSP variants are distinct, with I27 always displaying greater unfolding forces than the TmCSP variants. Gaussian fits to histograms for each data set are used to obtain a measure of the unfolding forces.

Effect of side-chain reduction on the mechanical stability of TmCSP. We used SMFS experiments to unfold the chimeric polyproteins (I27-TmCSP L40A)$_3$-I27 and (I27-TmCSP V62A)$_3$-I27 to measure their mechanical unfolding forces (Fig. 4(a)). Stretching the chimeric polyproteins resulted in force-extension (FX) data, allowing measurement of the mechanical unfolding forces needed to unfold each protein. Example FX traces for (I27-TmCSP L40A)$_3$-I27 and (I27-TmCSP V62A)$_3$-I27 are shown in Fig. 4(b), where each individual unfolding peak corresponds to the mechanical unfolding of individual CSP variants or I27 domains. The previously studied I27 protein acts as an internal mechanical marker in identifying the single domains being unfolded in FX traces.[57] The FX data contain two distinct sets of peaks, which differ in both their unfolding forces (F_U) and the distances between them (x_{p2p}) (Fig. 4(b) and (c)). For the variant TmCSP L40A the force-distance scattergram (Fig. 4(c)) shows two distributions centered around distances of 18.6 nm and 23.5 nm and forces around 80.0 pN and 201.9 pN. For the variant TmCSP V62A the distance-frequency histogram (Fig. 4(d)) shows two distributions centered around 18.6 nm and 22.9 nm and the force-frequency histogram shows two distributions around 70.0 pN and 199.4 pN. Both values of x_{p2p} for both chimeric polyproteins are close to the previously published x_{p2p} values for TmCSP and I27 (19.0 and 23.7, respectively[28]). The FX data were fitted with the WLC model (dashed lines in Fig. 4(b)) to obtain the increase in contour length of the polyprotein with each unfolding event, ΔL_C. At 2000 nm s^{-1}, an unfolding peak with a ΔL_C ~ 28.0 nm and a F_U of ~ 195 pN (±3 pN, standard deviation between the median values of the triplicate datasets) is the mechanical fingerprint for the I27 proteins in the chimeric polyprotein. In the FX data, initial unfolding peaks are observed with ΔL_C and F_U values of ~23.5 nm and 83 pN (±3 pN) for TmCSP L40A and ~23.5 nm and 73 pN (±4 pN) for TmCSP V62A, respectively. These correspond to the unfolding of the smaller (and weaker) TmCSP L40A and TmCSP V62A. While the ΔL_C is the same as that measured for TmCSP, the F_U is 6% lower for TmCSP L40A and 17% lower TmCSP V62A. This reduction in the average unfolding force (Fig. 4(d)), despite a similar ΔL_C, suggests that the reduction in side-chain length has an impact on the mechanical stability. Inspection of the unfolding force distributions for TmCSP L40A and TmCSP V62A (Fig. 4(d)) and comparing with those of TmCSP[28] show that while F_U decreases for the variants, the width of the distributions is relatively unchanged. The width of the F_U distributions is related to the distance to the unfolding transition state Δx_U in the Zhurkov-Bell model.[45] This suggests that the position of the mechanical unfolding transitions state for the variants is not shifted along the unfolding reaction coordinate with respect to TmCSP.

In addition to SMFS experiments, we completed simulations of constant velocity protein unfolding using MD. A total of 50 simulations were performed for each protein: TmCSP, TmCSP L40A and TmCSP V62A. In the force extension trajectories the initial rupture force was measured at the first peak in the

pathway that was accompanied by a sudden increase in length of the protein and a loss in the number of hydrogen bonds between a pair of β-strands (Fig. S8). The mean forces measured, including standard deviation (SD), were 103 ± 23 pN for TmCSP, 99 ± 21 pN for TmCSP L40A and 91 ± 19 pN for TmCSP V62A which is 3.9% and 11.7% lower than for TmCsp, respectively. The end-to-end distances at the peak position were the same for the three proteins: 1.54 ± 0.14 nm (TmCSP), 1.53 ± 0.13 nm (TmCSP L40A) and 1.55 ± 0.18 nm (TmCSP V62A). The MD simulations therefore show the same mechanical hierarchy as that measured in the experiments, with the TmCSP displaying the highest rupture force, followed by TmCSP L40A and then TmCSP V62A. The simulations also show that the localized region of the protein that resists unfolding (the mechanical clamp) in the two variants is in the same position as that in TmCSP[11b]
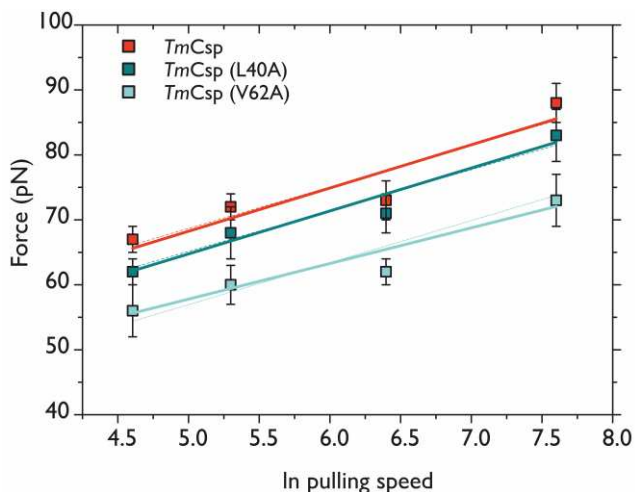

Figure 5. Impact of hydrophobic side-chain reduction on the free energy landscape of TmCSP. The mechanical unfolding force is shown for the cold shock proteins for the two TmCSP variants, (I27-TmCSP L40A)$_3$-I27) (dark turquoise squares) and ((I27-TmCSP V62A)$_3$-I27) (light turquoise squares) as a function of the logarithm of the pulling velocity and compared with TmCSP in the chimeric polyprotein (I27-TmCSP)$_3$-I27) (red squares). For each pulling speed, the data points show the median value of the unfolding force for the CSPs from three different experiments. The error bars for each data point show the standard deviation between the three different experiments at each pulling velocity. The line of best fit to the data (solid line) and the Monte Carlo fits to the experimental data (dashed line) are also shown.

Pulling speed dependence of unfolding force for TmCSP L40A and TmCSP V62A. To uncover details about the unfolding energy landscape of the two CSP variants we completed experiments at pulling speeds of 100, 200, 600 and 2000 nm s^{-1} to obtain the pulling speed dependence of F_U (Table S3 and S4). We measured the F_U for each unfolding peak in the FX data, and made separate histograms for each triplicate ex-

periment at each pulling velocity (Figure 4(d)). The histograms were fitted with Gaussian distributions and the median values of F_U for I27 and both TmCSP L40A and TmCSP V62A from each of the three replicate experiments were found. The natural logarithm of the pulling speed against the mean F_U (calculated from the medians of the triplicate experiments) for the two populations observed in the F_U histograms was plotted, allowing examination of the pulling speed dependence of F_U for each protein.Fig. 5 shows F_U for TmCSP, TmCSP L40A and TmCSP V62A, and Fig. S9 shows F_U for I27 from data on unfolding (I27-TmCSP)$_3$-I27, (I27-TmCSP L40A)$_3$-I27 and (I27-TmCSP V62A)$_3$-I27. These data were compared with those obtained previously for TmCSP under the same experimental conditions.[28] The values of F_U for I27 are in good agreement with previous studies of this I27 variant with similar domain numbers and scaffold design.[28, 30, 37, 50] It is clear from Fig. 5 that the pulling speed dependence of the two variants displays a clear shift with respect to that of TmCSP. At all pulling speeds, the side-chain reduction in each variant results in a decrease in the mechanical stability of the TmCSP, with the variant TmCSP V62A displaying a greater reduction in F_U at all pulling speeds. While the F_U is reduced for each variant, we measure a similar slope as compared with TmCSP. These experimental results suggest that the reduced mechanical stability of TmCSP L40A and TmCSP V62A is due to a decrease of the activation energy barrier height (ΔG^*), while the distance to the unfolding transition state (Δx_U) appears unchanged.

Protein	Mutation location	$\Delta\Delta G_{D-N}$, kJ/mol[‡]	Δx_U, nm	k_U, s^{-1} (±SD)	Monte Carlo[¶]	
					$\Delta\Delta G^*$, kJ/mol	ϕ_u
TmCSP			0.70	0.00095 (±0.00018)		
L40A	Loop	16	0.70	0.00160 (±0.00015)	1.29	0.08
V62A	β-strand 5	11	0.70	0.00550 (±0.00046)	4.35	0.39

Table 1. Summary of mechanical free energy parameters for TmCSP, TmCSP L40A, TmCSP V62A and I27, where $\Delta\Delta G_{D-N}$ = ΔG_{D-N}(TmCSP) − ΔG_{D-N}(mutant), $\Delta\Delta G_U^*$ = ΔG_U^*(TmCSP) − ΔG_U^*(mutant) and ϕ_u = $\Delta\Delta G_U^*/\Delta\Delta G_{D-N}$. [¶] The Monte Carlo fit uses a fixed value for x_U of 0.70 nm for the CSPs and fixed value for x_U of 0.32 nm and k_U of 0.00150 s^{-1} for I27.[11b] [‡]ΔG_{D-N} was obtained from chemical denaturation experiments at 23°C (see Fig. S7 and Table S2).

From the information in Fig. 5 we can access the unfolding rate, k_U and the distance from the native state to the unfolding transition state, Δx_U using a Monte Carlo simulation procedure described previously.[57] Given that the slope of the pulling speed dependence is similar for all three proteins (Fig. 5) we assume the distance to the unfolding transition state (Δx_U) is unchanged. This allows us to extract information about k_U and calculate mechanical ϕ_u-values, as described below. Δx_U for TmCSP was recently determined to be 0.70 nm,[11b] therefore Δx_U was fixed to this value for all three proteins. The values of Δx_U and k_U are shown in Table 1. In the analysis of the best fit parameters, the slopes of the lines of best fit to the experimental data were also assumed to be fixed (6.3 ± 0.7 pN for Csp).[38] These results show that the two variants TmCSP L40A and TmCSP V62A have a higher k_U than TmCSP. Constant velocity protein unfolding simulations using MD show the same mechanical hierarchy as that measured in the experiments, with the TmCSP displaying the highest rupture force, followed by TmCSP L40A and then TmCSP V62A. The simulations show that the mechanical clamp in the two variants is in the same position as that in TmCSP. In all three cases rupture of strand pairs β1–β4 or β4–β5 or the near simultaneous rupture of both pairs occurs in synchrony with the initial peak in force and subsequent extension of the protein.

Calculation of mechanical ϕ values. Using a method applied previously,[38] we calculated the mechanical ϕ value for the two variants[58, 59] to probe their role in the unfolding transition state of TmCSP. The mechanical ϕ value for unfolding (ϕ_u) is defined as the ratio of the loss of stability of the transition state and native state on mutation, using the native state as a reference. The ϕ_u value is given by the ratio of the change in the height of the mechanical activation energy barrier $\Delta\Delta G_U^*$ to the change in the thermodynamic stability of the protein $\Delta\Delta G_{D-N}$, ϕ_u = $\Delta\Delta G_U^*/\Delta\Delta G_{D-N}$.[38] The value of $\Delta\Delta G_U^*$ (Table 1) is measured from the mechanical unfolding kinetics, where $\Delta\Delta G_U^*$ = -RTln(k_U(TmCSP)/k_U(mutant)). The value of $\Delta\Delta G_{D-N}$ (Table 1) is measured from the difference in the thermodynamic stability at room temperature (Fig. 3). This method provides a measure of structure disruption in the transition state on a scale from 0 to 1, where 0 indicates that the mutated site remains fully structured and 1 indicates the site is completely disrupted, relative to the native state. Using the measured mechanical unfolding rates k_U (Table 1) and the thermodynamic stability ΔG_{D-N} (Fig. 3) we calculated the mechanical ϕ_u values for the two TmCSP variants. For TmCSP L40A where a mutation has been made to a residue in the loop region of the protein, the ϕ_u value is close to zero (0.08). This suggests that there is little disruption of the native interactions of L40 in the transition state. For TmCSP V62A, where a mutation was made in the fifth β-strand, a larger mechanical ϕ_u value of 0.39 was measured. An intermediate ϕ_u value such as this can indicate either that the residue forms a fraction of its native contacts in the transition state, or that there are alternative transition states in which the native contacts are either formed or unformed. The latter of these two scenarios is very much in accord with our simulation results.
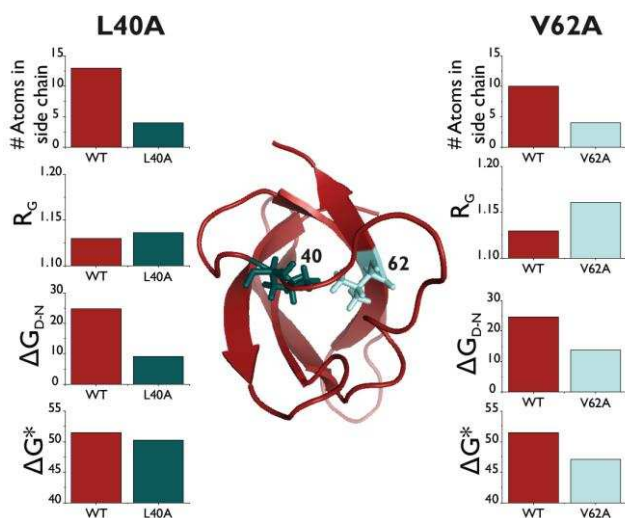
Figure 6. Differential effects of hydrophobic side-chain reduction on the flexibility and stability of TmCSP. Mutations in positions 40 and 62 of TmCSP result in distinct effects on the stability of the protein. The bar charts show the number of atoms in the hydrophobic core residue which is mutated, the radius of gyration R_G, the thermodynamic stability ΔG_{D-N} and mechanical stability ΔG_U^* for TmCSP and the two mutants L40A (left) and V62A (right).

CONCLUSIONS

The TmCSP protein forms a β-barrel structure comprising five β-strands and a loop region (Fig. 6). To determine the impact of side-chain reduction of buried hydrophobic residues on the thermodynamic stability of TmCSP we completed thermal unfolding titrations in the presence of a chemical denaturant. The results showed that the side-chain reduction in the variants TmCSP L40A and TmCSP V62A both result in a lower melting temperature, T_M, and a lower overall thermodynamic stability (ΔG_{D-N}) over all temperatures (Fig. 6). While both of the CSP variants are maximally stable just below room temperature, like the hyperthermophilic TmCSP, their maximal stabilities are considerably smaller than that of TmCSP. At 23 °C, the changes in ΔG_{D-N} on mutation are 16 and 11 kJ mol[-1] for TmCSP L40A and TmCSP V62A, respectively (Table 1). A search of single point mutations (T = 20–30 °C, pH 7–8) in proteins on the Protherm mutant database[60] indicates that L to A mutations destabilize proteins by 10.2 ± 5.5 kJ mol[-1] whilst V to A mutations destabilize proteins by 8.9 ± 5.8 kJ mol[-1]. The loss of a greater number of atoms in the L to A mutants causes, on average, a larger destabilization of the protein to denaturation. Our measured values agree well with these values from the database, although it is noteworthy that the L40A change is at the high end of the range. (We also estimated the stability change on mutation for the specific changes to TmCSP using I-Mutant[61] which resulted in similar values of 10.3 and 8.4 kJ mol[-1] for TmCSP L40A and TmCSP V62A, respectively.)

The force-bearing β-strands, 1–4 and 4–5, are antiparallel and constitute a shear topology. Interestingly, this is a typical feature of mechanically stable proteins.[62] These force-bearing strands are connected by hydrogen bonds; our simulations show approximately six hydrogen bonds on average between β-strands 1–4 and seven between β-strands 4–5 (Fig.

2b), and they may represent the 'mechanical clamp' of the TmCSP protein. This structural element provides mechanical stability and is the rate-limiting step for the unfolding of the protein. While the mechanical clamp motif of the protein is important for mechanical stability, other non-covalent interactions play a role in conferring stability. For example, previous studies have shown that protein mechanical stability depends on the interactions which occur between the surfaces which are sheared apart upon forced unfolding.[37, 38] Here, we have shown that by reducing the hydrophobic core and increasing the loop flexibility in TmCSP we can significantly modulate protein stability without changing the mechanical unfolding pathway (Fig. 6). We show that reduced loop flexibility confers

thermodynamic stability in TmCSP and may play a role in minimizing potential initiation sites for thermal denaturation.[32] We show that the hydrophobic core of TmCSP is important for conferring mechanical stability, particularly when it involves residues central to the mechanical clamp region (Fig. 6). The mutant TmCSP V62A involves a reduction in the side chain of a residue in β-strand 5, which is part of the structural motif of the mechanical clamp of the TmCSP protein. Our SMFS results show that a reduction in the hydrophobic core packing upon mutation (Fig. 2) results in a decrease in the mechanical stability of the TmCSP V62A (Fig. 5), providing evidence that interactions mediated by this hydrophobic residue at the interface of two shearing motifs (β-strands 4-5) plays an important role in determining the mechanical stability of TmCSP. Single molecule force spectroscopy measures the mechanical stability (ΔG^*) and shows a greater reduction for TmCSP V62A ($\Delta\Delta G^*$ = 9.2%) than TmCSP L40A ($\Delta\Delta G^*$ = 2.6%). We use mechanical ϕ_u value analysis to reveal details of the mechanical unfolding pathway of the protein, revealing the important role of regions in the vicinity of the mechanical clamp motif.

The ability to make conservative mutants which do not alter the structure of the protein or the mechanical unfolding pathway provides a unique platform with which to quantitatively measure the impact of specific interactions on protein mechanical and thermodynamic stability. Such insight provides opportunities to rationally design proteins with specific and optimized stabilities for exploitation in bionanotechnology applications.

AUTHOR INFORMATION
Corresponding Author
L.Dougan@leeds.ac.uk

1. Horikoshi, K.; Grant, W. D., Extremophiles:Microbial life in extreme environments. Wiley-Liss: 1998.

2. Rothschild, L. J.; Mancinelli, R. L., Life in extreme environments. Nature **2001,** 409, (6823), 1092-1101.

3. Scandurra, R.; Consalvi, V.; Chiaraluce, R.; Politi, L.; Engel, P. C., Protein thermostability in extremophiles. Biochimie **1998,** 80, (11), 933-941.

4. Siddiqui, K. S.; Thomas, T., Protein Adaptation in Extremophiles. Nova Biomecial Books: New York, 2008.

5. Razvi, A.; Scholtz, J. M., Lessons in stability from thermophilic proteins. Protein Science **2006,** 15, (7), 1569-1578.

6. Hiller, R.; Zhou, Z. H.; Adams, M. W. W.; Englander, S. W., Stability and dynamics in a hyperthermophilic protein with melting temperature close to 200 degrees C. Proceedings of the National Academy of Sciences of the United States of America **1997,** 94, (21), 11329-11332.

7. Vieille, C.; Zeikus, G. J., Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol R **2001,** 65, (1), 1-+.

8. Sterner, R.; Liebl, W., Thermophilic adaptation of proteins. Crit Rev Biochem Mol **2001,** 36, (1), 39-106.

9. Russell, R. J. M.; Taylor, G. L., Engineering Thermostability - Lessons from Thermophilic Proteins. Curr Opin Biotech **1995,** 6, (4), 370-374.

10. Jaenicke, R.; Bohm, G., The stability of proteins in extreme environments. Curr Opin Struc Biol **1998,** 8, (6), 738-748.

11. Hart, K. M.; Harms, M. J.; Schmidt, B. H.; Elya, C.; Thornton, J. W.; Marqusee, S., Thermodynamic System Drift in Protein Evolution. Plos Biol **2014,** 12, (11).

12. Zhou, H. X.; Dong, F., Electrostatic contributions to the stability of a thermophilic cold shock protein. Biophysical Journal **2003,** 84, (4), 2216-2222.

13. Wells, S. A.; Crennell, S. J.; Danson, M. J., Structures of mesophilic and extremophilic citrate synthases reveal rigidity and flexibility for function. Proteins-Structure Function and Bioinformatics **2014,** 82, (10), 2657-2670.

14. Sawle, L.; Ghosh, K., How Do Thermophilic Proteins and Proteomes Withstand High Temperature? Biophysical Journal **2011,** 101, (1), 217-227.

15. Marcos, E.; Mestres, P.; Crehuet, R., Crowding Induces Differences in the Diffusion of Thermophilic and Mesophilic Proteins: A New Look at Neutron Scattering Results. Biophysical Journal **2011,** 101, (11), 2782-2789.

16. Mamonova, T. B.; Glyakina, A. V.; Galzitskaya, O. V.; Kurnikova, M. G., Stability and rigidity/flexibility—Two sides of the same coin? Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **2013,** 1834, (5), 854-866.

17. Lazaridis, T.; Lee, I.; Karplus, M., Dynamics and unfolding pathways of a hyperthermophilic and a mesophilic rubredoxin. Protein Science **1997,** 6, (12), 2589-2605.

18. Dominy, B. N.; Perl, D.; Schmid, F. X.; Brooks, C. L., The effects of ionic strength on protein stability: The cold shock protein family. Journal of Molecular Biology **2002,** 319, (2), 541-554.

19. Chan, C.-H.; Wilbanks, C. C.; Makhatadze, G. I.; Wong, K.-B., Electrostatic Contribution of Surface Charge Residues to the Stability of a Thermophilic Protein: Benchmarking Experimental and Predicted pKa Values. PLoS ONE **2012,** 7, (1), e30296.

20. Robinson-Rechavi, M.; Alibés, A.; Godzik, A., Contribution of Electrostatic Interactions, Compactness and Quaternary Structure to Protein Thermostability: Lessons from Structural Genomics of Thermotoga maritima. Journal of Molecular Biology **2006,** 356, (2), 547-557.

21. Gromiha, M. M.; Pathak, M. C.; Saraboji, K.; Ortlund, E. A.; Gaucher, E. A., Hydrophobic environment is a key factor for the stability of thermophilic proteins. Proteins: Structure, Function, and Bioinformatics **2013,** 81, (4), 715-721.

22. Gera, N.; Hussain, M.; Wright, R. C.; Rao, B. M., Highly Stable Binding Proteins Derived from the Hyperthermophilic Sso7d Scaffold. Journal of Molecular Biology **2011,** 409, (4), 601-616.

23. Priyakumar, U. D., Role of Hydrophobic Core on the Thermal Stability of Proteins - Molecular Dynamics Simulations on a Single Point Mutant of Sso7d. J Biomol Struct Dyn **2012,** 29, (5), 961-971.

24. Priyakumar, U. D.; Ramakrishna, S.; Nagarjuna, K. R.; Reddy, S. K., Structural and Energetic Determinants of Thermal Stability and Hierarchical Unfolding Pathways of Hyperthermophilic Proteins, Sac7d and Sso7d. J Phys Chem B **2010,** 114, (4), 1707-1718.

25. Xu, X. J.; Su, J. G.; Chen, W. Z.; Wang, C. X., Thermal Stability and Unfolding Pathways of Sso7d and its Mutant F31A: Insight from Molecular Dynamics Simulation. J Biomol Struct Dyn **2011,** 28, (5), 717-727.

26. Wintrode, P. L.; Zhang, D. Q.; Vaidehi, N.; Arnold, F. H.; Goddard, W. A., Protein dynamics in a family of laboratory evolved thermophilic enzymes. Journal of Molecular Biology **2003,** 327, (3), 745-757.

27. Sadeghi, M.; Naderi-Manesh, H.; Zarrabi, M.; Ranjbar, B., Effective factors in thermostability of thermophilic proteins. Biophysical Chemistry **2006,** 119, (3), 256-270.

28. Hoffmann, T.; Tych, K. M.; Brockwell, D. J.; Dougan, L., Single-Molecule Force Spectroscopy Identifies a Small Cold Shock Protein as Being Mechanically Robust. J Phys Chem B **2013,** 117, (6), 1819-1826.

29. Tych, K. M.; Batchelor, M.; Hoffmann, T.; Wilson, M.; Paci, E.; Brockwell, D. J.; Dougan, L., Tuning protein mechanics through an ionic cluster graft from an extremophilic protein. Soft matter **2016.**

30. Tych, K. M.; Hoffmann, T.; Brockwell, D. J.; Dougan, L., Single molecule force spectroscopy reveals the temperature-dependent robustness and malleability of a hyperthermophilic protein Soft Matter **2013,** 9, (37), 9016-9025.

31. Takahashi, Y. T.; Sasaki, H.; Takayama, S. I. J.; Mikami, S. I.; Kawano, S.; Mita, H.; Sambongi, Y.; Yamamoto, Y., Further enhancement of the thermostability of Hydrogenobacter thermophilus cytochrome c(552). Biochemistry-Us **2006,** 45, (36), 11005-11011.

32. Kumar, S.; Tsai, C. J.; Nussinov, R., Factors enhancing protein thermostability. Protein Engineering **2000,** 13, (3), 179-191.

33. Thompson, M. J.; Eisenberg, D., Transproteomic evidence of a loop-deletion mechanism for enhancing protein thermostability. Journal of Molecular Biology **1999,** 290, (2), 595-604.

34. Brockwell, D. J.; Beddard, G. S.; Clarkson, J.; Zinober, R. C.; Blake, A. W.; Trinick, J.; Olmsted, P. D.; Smith, D. A.; Radford, S. E., The effect of core destabilization on the mechanical resistance of I27. Biophysical Journal **2002,** 83, (1), 458-472.

35. Fowler, S. B.; Best, R. B.; Herrera, J. L. T.; Rutherford, T. J.; Steward, A.; Paci, E.; Karplus, M.; Clarke, J., Mechanical unfolding of a titin Ig domain: Structure of unfolding intermediate revealed by combining AFM, molecular dynamics simulations, NMR and protein engineering. Journal of Molecular Biology **2002,** 322, (4), 841-849.

36. Ng, S. P.; Rounsevell, R. W. S.; Steward, A.; Geierhaas, C. D.; Williams, P. M.; Paci, E.; Clarke, J., Mechanical unfolding of TNfn3: The unfolding pathway of a fnIII domain probed by protein engineering, AFM and MD simulation. Journal of Molecular Biology **2005,** 350, (4), 776-789.

37. Sadler, D. P.; Petrik, E.; Taniguchi, Y.; Pullen, J. R.; Kawakami, M.; Radford, S. E.; Brockwell, D. J., Identification of a mechanical rheostat in the hydrophobic core of protein L. J Mol Biol **2009,** 393, (1), 237-248.

38. Bu, T. J.; Wang, H. C. E.; Li, H. B., Single Molecule Force Spectroscopy Reveals Critical Roles of Hydrophobic Core Packing in Determining the Mechanical Stability of Protein GB1. Langmuir **2012,** 28, (33), 12319-12325.

39. Hoffmann, T.; Tych, K. M.; Crosskey, T.; Schiffrin, B.; Brockwell, D. J.; Dougan, L., Rapid and Robust Polyprotein Production Facilitates Single-Molecule Mechanical Characterization of beta-Barrel Assembly Machinery Polypeptide Transport Associated Domains. Acs Nano **2015,** 9, (9), 8811-8821.

40. Moon, C. P.; Fleming, K. G., Using Tryptophan Fluorescence to Measure the Stability of Membrane Proteins Folded in Liposomes. Method Enzymol **2011,** 492, 189-211.

41. Pace, C. N.; Shaw, K. L., Linear extrapolation method of analyzing solvent denaturation curves. Proteins-Structure Function and Genetics **2000**, 1-7.

42. Becktel, W. J.; Schellman, J. A., Protein Stability Curves. Biopolymers **1987,** 26, (11), 1859-1877.

43. Rees, D. C.; Robertson, A. D., Some thermodynamic implications for the thermostability of proteins. Protein Science **2001,** 10, (6), 1187-1194.

44. Florin, E. L.; Rief, M.; Lehmann, H.; Ludwig, M.; Dornmair, C.; Moy, V. T.; Gaub, H. E., Sensing Specific Molecular-Interactions with the Atomic-Force Microscope. Biosens Bioelectron **1995,** 10, (9-10), 895-901.

45. Bell, G. I., Models for Specific Adhesion of Cells to Cells. Science **1978,** 200, (4342), 618-627.

46. Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., CHARMM: The Biomolecular Simulation Program. Journal of Computational Chemistry **2009,** 30, (10), 1545-1614.

47. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. J Mol Graph Model **1996,** 14, (1), 33-38.

48. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. Journal of Computational Chemistry **2005,** 26, (16), 1781-1802.

49. Haberthur, U.; Caflisch, A., FACTS: Fast analytical continuum treatment of solvation. Journal of Computational Chemistry **2008,** 29, (5), 701-715.

50. Wolny, M.; Batchelor, M.; Knight, P. J.; Paci, E.; Dougan, L.; Peckham, M., Stable Single alpha-Helices Are Constant Force Springs in Proteins. Journal of Biological Chemistry **2014,** 289, (40), 27825-27835.

51. Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A., Wordom: a program for efficient analysis of molecular dynamics simulations. Bioinformatics **2007,** 23, (19), 2625-2627.

52. Carter, P.; Andersen, C. A. F.; Rost, B., DSSPcont: continuous secondary structure assignments for proteins. Nucleic Acids Research **2003,** 31, (13), 3293-3295.

53. Dzubiella, J., Salt-specific stability and denaturation of a short salt-bridge-forming alpha-helix. J Am Chem Soc **2008,** 130, (42), 14000-14007.

54. Lee, B.; Richards, F. M., Interpretation of Protein Structures - Estimation of Static Accessibility. Journal of Molecular Biology **1971,** 55, (3), 379-&.

55. Hubbard, S. J.; Thornton, J. M. NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London. http://www.bioinf.manchester.ac.uk/naccess/

56. Ivankov, D. N.; Garbuzynskiy, S. O.; Alm, E.; Plaxco, K. W.; Baker, D.; Finkelstein, A. V., Contact order revisited: Influence of protein size on the folding rate. Protein Science **2003,** 12, (9), 2057-2062.

57. Hoffmann, T.; Dougan, L., Single molecule force spectroscopy using polyproteins. Chem Soc Rev **2012,** 41, (14), 4781-4796.

58. Best, R. B.; Fowler, S. B.; Herrera, J. L. T.; Steward, A.; Paci, E.; Clarke, J., Mechanical unfolding of a titin Ig domain: Structure of transition state revealed by combining atomic force microscopy, protein engineering and molecular dynamics simulations. Journal of Molecular Biology **2003,** 330, (4), 867-877.

59. Best, R. B.; Fowler, S. B.; Toca-Herrera, J. L.; Clarke, J., A simple method for probing the mechanical unfolding pathway of proteins in detail. Proceedings of the National Academy of Sciences of the United States of America **2002,** 99, (19), 12143-12148.

60. Bava, K. A.; Gromiha, M. M.; Uedaira, H.; Kitajima, K.; Sarai, A., ProTherm, version 4.0: thermodynamic database for proteins and mutants. Nucleic Acids Research **2004,** 32, D120-D121.

61. Capriotti, E.; Fariselli, P.; Casadio, R., I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Research **2005,** 33, W306-W310.

62. Hoffmann, T.; Tych, K. M.; Hughes, M.; Brockwell, D. J.; Dougan, L., Towards design principles for understanding protein mechanical stability. Physcial Chemistry Chemical Physics **2013,** 15, (38), 15767-15780.