# Stabilization of a protein conferred by an increase in folded state entropy

Shlomi Dagan[a], Tzachi Hagai[b,1], Yulian Gavrilov[b], Ruti Kapon[a], Yaakov Levy[b,2], and Ziv Reich[a,2]

Departments of [a]Biological Chemistry and [b]Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel

Edited by Peter G. Wolynes, Rice University, Houston, TX, and approved May 20, 2013 (received for review February 4, 2013)

**Entropic stabilization of native protein structures typically relies on strategies that serve to decrease the entropy of the unfolded state. Here we report, using a combination of experimental and computational approaches, on enhanced thermodynamic stability conferred by an increase in the configurational entropy of the folded state. The enhanced stability is observed upon modifications of a loop region in the enzyme acylphosphatase and is achieved despite significant enthalpy losses. The modifications that lead to increased stability, as well as those that result in destabilization, however, strongly compromise enzymatic activity, rationalizing the preservation of the native loop structure even though it does not provide the protein with maximal stability or kinetic foldability.**

protein folding | loop closure entropy | molecular dynamics

Reducing the difference in entropy between the unfolded and folded states can increase the thermodynamic stability of a protein. This is commonly accomplished by strategies that act to restrict the conformational space available for the unfolded state (e.g., decreasing loop length, macromolecular crowding, and backbone cyclization) (1). In principle, changes that increase the entropy of the folded state can also lead to its stabilization provided that they exceed the loss of enthalpic contributions, if stabilizing interactions are perturbed by the modifications.

The current work originally aimed at studying the effects exerted by changes in the length of a loop region on protein stability and folding. As a model, we chose a loop in human muscle acylphosphatase (hmAcP), a small (~100 aa) enzyme that catalyzes the hydrolysis of the carboxyl–phosphate bond in various acylphosphate compounds and presents an open α/β-sandwich structure (ref. 2, and see, e.g., refs. 3–5). The folding stability and dynamics of hmAcP have been studied extensively and are well characterized (ref. 6–10 and references therein). Excluding a minor *cis–trans* prolyl isomerization phase, it folds in a two-step process, albeit very slowly (due to abundance of long-range interactions; ref. 9), through a relatively compact, native-like transition state. The loop we chose for the modifications (hereafter referred to as L4) connects between the second helix and the fourth β-strand of the protein (Fig. 1) and possesses multiple internal and external contacts (refs. 3 and 4 and our own contact analysis). The latter contacts are formed predominantly with residues located in the first loop of the protein (L1), which runs along L4 and is involved in the binding of the phosphate group of the substrate (4, 11–15).

Characterizing the properties of hmAcP mutants carrying deletions or insertions in L4 we found that the thermodynamic stability of mutants in which the loop was shortened is increased to an extent significantly larger than that predicted by polymer models for loop closure entropy. The increased stability is predominantly due to a decrease in the unfolding rate and is attained despite the fact that shortening of the loop is accompanied by considerable losses in enthalpy. Given the above, we postulated that the surplus stabilization arises from an increased conformational entropy of the folded state ensemble. We provide evidence that supports this hypothesis and show that the enhanced dynamics are accompanied by only relatively small rearrangements in secondary and tertiary structure. The thermodynamically stabilized mutants, as well as the destabilized variants carrying elongated loops, are all deficient in enzymatic activity, suggesting that the native structure of the loop has evolved primarily for function.

## Results and Discussion

The WT and mutant proteins used in the study were derived from the C21S variant of hmAcP, which was used to avoid complications associated with free cysteine residues (7). This mutant is referred to in the manuscript as WT hmAcP. The structure of hmAcP has not been determined. However, the horse ortholog, for which a solution structure is available (3), differs from hmAcP in only five amino acids, all of which are located outside L4. We used this structure to generate models of WT hmAcP, as well as of the various loop-length variants used in the experiments and in the molecular dynamics (MD) simulations (see also refs. 10, 16, and 17).

Table 1 lists the L4 deletion and insertion mutants that were constructed. The site chosen for the modifications encompasses the two adjacent serine residues at positions 72 and 73, which are located in a β-bulge at the tip of L4 that points toward the solvent. This location was considered favorable in terms of end effects and potential interactions between the inserts and the protein. The deletion series consists of three mutants in which two to six residues were removed in a pairwise fashion, starting with the two serine residues and progressing symmetrically outward. These mutants are designated Δ2, Δ4, and Δ6. The deletions made in the Δ4 and Δ6 mutants include Pro71, the substitution of which to alanine was previously shown to result in destabilization of the protein (9). However, as discussed below, all of the deletion mutants we constructed are more stable than the WT. The elongated loop mutants, designated ∇2–∇12, contain stretches of 2, 4, 6, 8, and 12 glycine residues (marked in boldface in Table 1), which were inserted between the two serine residues. Far-UV circular dichroic (Fig. S1*A* and Table S1) and fluorescence emission spectra (Fig. S1*C*) recorded for the mutants revealed only small variations from the WT protein, indicating that the modifications in L4 do not lead to major alterations in secondary or tertiary structure. All mutants have a relative contact order similar to that of the WT (about 20%) and kinetically folded and unfolded in a two-state manner (excluding the proline isomerization phase).

Thermodynamic parameters derived for the proteins from equilibrium urea denaturation and differential scanning calorimetry (DSC) are listed in Table S2 and drawn in Fig. 2. Going from the mutant carrying the longest loop, ∇12, one sees only a very

Author contributions: S.D., T.H., Y.G., Y.L., and Z.R. designed research; S.D., T.H., and Y.G. performed research; S.D., T.H., Y.G., R.K., Y.L., and Z.R. analyzed data; and R.K., Y.L., and Z.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Department of Microbiology and Immunology, University of California, San Francisco School of Medicine, San Francisco, CA 94158.

[2]To whom correspondence may be addressed. E-mail: koby.levy@weizmann.ac.il or ziv.reich@weizmann.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1302284110/-/DCSupplemental.

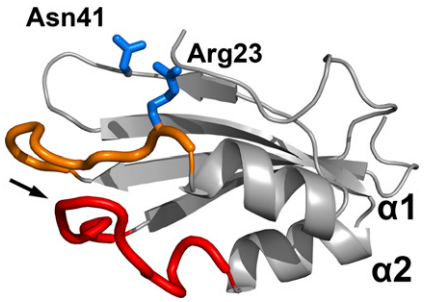


**Fig. 1.** Model of hmAcP, constructed based on the solution structure of horse muscle AcP (Protein Data Bank ID code 1APS). L4 (residues 64–77) and the catalytic loop (L1, residues 15–24) are shown in red and orange, respectively. The conserved Arg23 and Asn41 residues, which respectively function in the binding of the substrate's phosphate group and the catalytic water molecule, are shown in stick representation. The arrow marks the location of Ser72 and Ser73, which served as the site for modifications in L4.

small change in stability (0.5 kJ/mol) up to the ∇6 insertion mutant (Fig. 2*A*). As the length of the loop decreases further, the resistance of the proteins to chemical denaturation increases significantly more rapidly, leading eventually to an additional stabilization of ∼18 kJ/mol. The dependence on loop length of the thermal stability of the proteins exhibits a pattern that is very similar to the one derived from the chemical denaturation experiments (Fig. 2*B*), finally bringing the heat resistance of the mutant carrying the shortest loop (Δ6) to the range of thermophilic proteins.

Fig. 2*C* shows the dependence of the calorimetric unfolding enthalpy, $\Delta H_m$, on loop length. Moving from the WT to Δ6, $\Delta H_m$ decreases rapidly, eventually dropping by more than 50 kJ/mol. This decrease is most likely due to disruption of interactions involving residues in L4 upon the truncations, effected directly, by the removal of residues engaged in interactions, or indirectly, as a result of alterations in loop geometry. Elongation of L4 by inserts containing two, four, and six glycine residues results in $\Delta H_m$ values higher than those measured for the WT. The gain in enthalpy in these loop-insert mutants may reflect optimization of existing contacts or the formation of new ones, within or outside L4. Further elongation of the loop does not lead to additional significant changes. Overall, $\Delta H_m$ decreases by almost 80 kJ/mol between ∇6 and Δ6. This decrease, which concurs with decreasing loop length, opposes the measured change in conformational stability of the proteins, which increases as the loop becomes shorter. The gain in stability upon loop shortening must therefore be of an entropic origin.

Increasing loop length leads to deceleration of folding and acceleration of unfolding of the proteins (Table S3 and Fig. 3 *A* and *B*). In both cases, the plots of the natural logarithms of the rate constants vs. loop length are nonlinear and tend to saturate as the length of the loop increases. The magnitude of the changes, however, differs considerably between the folding and unfolding reactions. The effect of loop length on the folding rate of the proteins is quite small, with the difference between the fastest (Δ6) and slowest (∇12) variants being about threefold. Such a relatively small effect is in line with results obtained from phi-value analysis (9) and Monte Carlo and MD simulations (10, 16), which indicate that native contacts are mostly absent from L4 in the folding transition state ensemble. By contrast, the changes in the unfolding rates of the proteins extend over a range of ∼250-fold. These changes therefore dominate the changes in stability along the loop-length series.

**Table 1. Loop-length variants used in this study**

| Variant | Sequence of L4 |
|---|---|
| Δ6 | WLSKVG------DR |
| Δ4 | WLSKVGS----IDR |
| Δ2 | WLSKVGSP--RIDR |
| WT | WLSKVGSPSSRIDR |
| ∇2 | WLSKVGSPS**GG**SRIDR |
| ∇4 | WLSKVGSPS**GGGG**SRIDR |
| ∇6 | WLSKVGSPS**GGGGGG**SRIDR |
| ∇8 | WLSKVGSPS**GGGGGGGG**SRIDR |
| ∇12 | WLSKVGSPS**GGGGGGGGGGGG**SRIDR |

Boldface indicates glycine residues.

Previous studies conducted on flexible loop regions revealed that the energetic consequences of changing loop length are predominantly related to the entropic cost of ordering a loop upon folding (18–22). This effect can be approximated by simple polymer models of the form

$$\Delta\Delta G(n) = cRT\ln(n/n_{ref}), \quad [1]$$

where $n$ denotes the number of residues in the loop of a given variant and $n_{ref}$ is the number of residues in the loop of a reference mutant (usually, the one carrying the largest number of residues). The coefficient $c$ is a correction factor that is related to the persistence length and which depends on the nature of the polymer and on the length and composition (in our case, the amino acid content) of the loop (23, 24). To estimate the contribution of loop closure entropy to the changes in the stability of the hmAcP variants, we used Eq. 1 with two extreme values of $c$, one ($c$ = 1.5) corresponding to an ideal random-walk chain and the other ($c$ = 2.44) taking into account excluded volume effects for both the loop and the residues that flank it (18, 23). The results, depicted in Fig. 4, show that this term can account for only a fraction of the observed differences in thermodynamic stability within the ∇6–Δ6 range, with the relative contribution becoming increasingly smaller as the length of the loop decreases.

The analysis described above was complemented by native topology-based (Gō-model) simulations. In these studies, the loop-length series was generated by insertions of stretches of Cα beads mimicking oligoglycine chains into the loop possessed by the Δ6 mutant, which was used as a reference. To isolate topological effects, attractive noncovalent interactions involving residues within the modified loop region were omitted, so that the folded state of all variants is defined by the same set of native contacts. These settings resemble the polymer-based models in the sense that the effects produced by the changes in loop length are mostly confined to the unfolded and transition states. Consistent with the results derived from the kinetic analyses, increasing the length of L4 slowed folding and facilitated unfolding (Fig. S2). However, the magnitude of the effect on the unfolding reaction was not captured by the simulations, which revealed only a twofold increase in the unfolding rate upon elongation of the loop by 26 residues. In both cases, the model predicts a linear dependence of the rate constants on loop length, deviating from the experimentally observed behaviors.

What, then, is responsible for the surplus entropic stabilization of the proteins within the ∇6–Δ6 range when the length of L4 is decreased? The data presented so far strongly indicate that the effect is related to the properties of the folded state. One such property is configurational entropy: If shortening of L4 leads to an increase in the entropy of the folded state, it may stabilize the proteins by decreasing the entropic gain of unfolding. To test this possibility, we performed all-atom MD simulations on the WT protein and the Δ6 mutant. We first addressed the structural consequences of the deletion. The matrix shown in Fig. 5*A* represents the differences in interresidue distances between the two proteins $\left(\langle d_{ij}^{\Delta 6}\rangle - \langle d_{ij}^{WT}\rangle\right)$, where $\langle d_{ij}\rangle$ is the mean pairwise distance

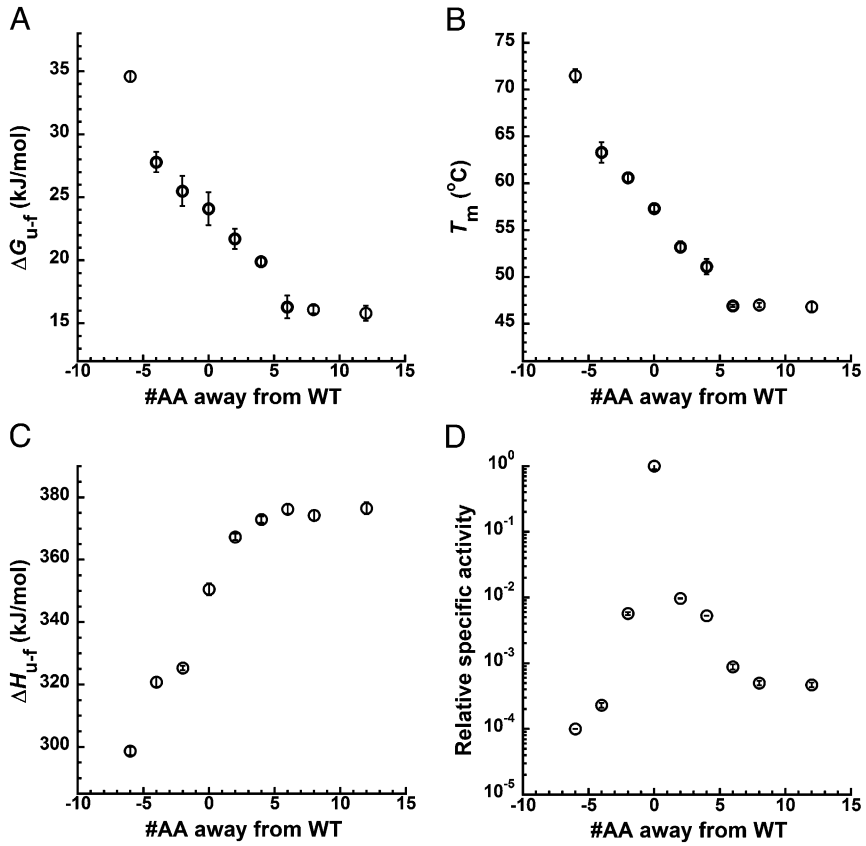


**Fig. 2.** Effects of L4 length on hmAcP stability and activity. The *x*-axes denote the difference in the number of residues in L4 between the loop-length mutants and the WT. (*A*) Unfolding free energy (25 °C). (*B*) Midpoint temperature of unfolding. (*C*) Calorimetric unfolding enthalpy. (*D*) Specific activity (25 °C). The values in *D* were normalized with respect to the activity of the WT protein. The data shown in this figure, as well as in Fig. 3, represent the means ± SD of at least three independent determinations.

between the Cα atoms of residues *i* and *j*) derived from analysis of 1,500 snapshots observed during 10 20-ns runs. The major differences relate to residues located within L4 and L1, and the helix adjacent to L1, α1. They amount to a small shift in the position of α1 (which also becomes slightly kinked) in the mutant, which moves α1 away from the β-sheet as well as from the second helix (α2). The shift is induced by a conformational change in L1, which results in a movement of its C terminal toward L4. This repositioning is enabled in Δ6 because the β-bulge in L4 that lies below L1 is removed by the deletion. Overall, however, the structure of the two proteins is similar, consistent with the far-UV CD and fluorescence emission data. This is further supported by near-UV CD analysis, which is extremely sensitive to changes in the microenvironment around aromatic residues. As can be seen (Fig. S1*B*), the major difference between the spectra recorded for the WT and Δ6 is a lower amplitude of the signal up to ∼275 nm in the spectrum of the latter, but the positions of the maxima and minima are conserved. The decrease in the amplitude at short wavelengths may reflect perturbations of interactions between L4 and Tyr25 in α1 (our own contact analysis) and/or change in the distance ($d_{wt}$ = 2.75 Å) between

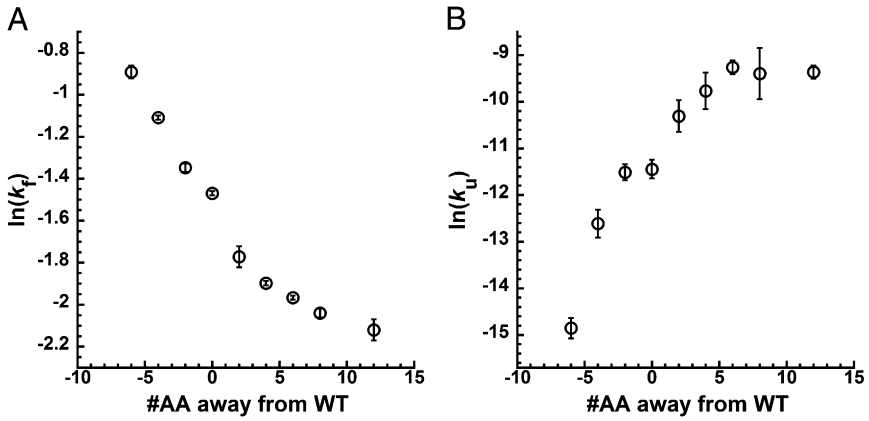


**Fig. 3.** Dependence on L4 length of folding (*A*) and unfolding (*B*) rates, measured at 25 °C.

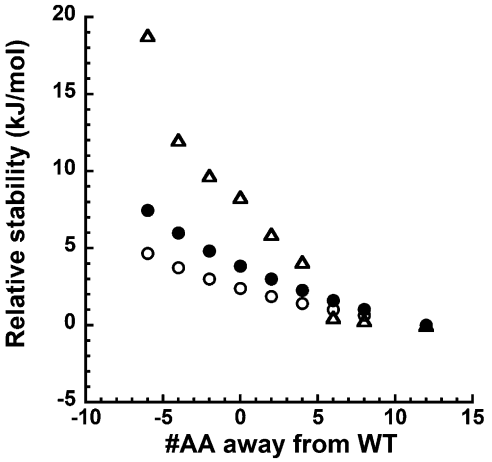


**Fig. 4.** Changes in stability predicted by the loop entropy model, using $c$ values of 1.5 (○) and 2.44 (●). The experimentally determined differences in unfolding free energy are marked by (△). Values were normalized with respect to the stability of the ∇12 mutant.

Tyr25 and Trp64 in L4. As discussed below, it may also reflect the altered dynamics of the folded state of the mutant. The differences between the spectra recorded for the other deletion mutants and the WT are even smaller (Fig. S1*B*).

Next, we examined the impact of the six-residue truncation in L4 on the conformational flexibility of the protein. Fig. 5*B* shows the differences in the SDs of the interresidue distances between Δ6 and the WT. The matrix reveals that the variance of the distances is generally higher in the mutant. The most prominent increases are apparent in regions corresponding to distances between residues located in L1 (C terminal) and α1 and residues in α2, the first (β1), second (β2), third (β3), and fifth (β5) strands of the sheet, and the loop that connects the fourth strand (β4) to β5. These increases most likely result from the removal of the β-bulge in L4 and the breakage of hydrogen bonds it makes with L1 by the deletion, lessening restrictions on the motion of L1 and the adjacent helix. Other, more localized, increases are observed for residues located in α2 and the β-turn connecting strands 2 and 3. An opposite effect is exhibited mainly by residues located at the edges of L4, particularly at the N terminal, for which the SDs are mostly lower in the mutant. Notably, albeit being overall more flexible, Δ6 has, on average, a smaller solvent-accessible surface area (Fig. S3). Results obtained from experiments using the fluorescence probe 8-anilino-1-naphthalene sulfonate (ANS) are in line with this notion, showing that binding of the dye to the loop-length variants decreases with decreasing loop length (Fig. S1*D*). These results are consistent with the decrease observed between Δ6 and ∇6 in the $m$-values (Tables S2 and S3), which correlate to the difference in the accessible surface area between the folded and unfolded states (refs. 25–29 and references therein).

To quantify the differences apparent in Fig. 5*B*, we followed ref. 30 (see also ref. 31) and used the covariance matrices of the atomic fluctuations observed in the MD trajectories to estimate the configurational entropy of the folded state ensemble of the two proteins. The analysis yielded values of 1.61 and 1.72 kJ/mol·K for WT and Δ6, respectively. At 25 °C, the consequence of this difference in internal entropy is stabilization of the folded state of Δ6 over that of the WT by 32.8 kJ/mol—almost an order of magnitude larger than the stabilization conferred by the corresponding difference in loop closure entropy (3.63 kJ/mol, with $c = 2.44$). Using Kirchhoff's relation, the folding enthalpies of the WT and Δ6 at 25 °C are estimated to be *ca.* −151 and −127 kJ/mol. Given this and the differences in configurational and loop closure entropies, the Δ6 mutant is expected to be stabilized (at 25 °C) by ∼12 kJ/mol compared with the WT, not too far from the experimentally determined difference of 10.5 kJ/mol (Table S2). Further analysis reveals that the secondary structure that contributes the most to the increase in the configurational entropy of Δ6 is α2 (Table S4). The first helix, β3, and β4 each contributes less than 50% compared with α2; contributions made by β1, β2, and by the catalytic loop (L1) are yet smaller, in the latter case likely reflecting the high flexibility of this loop in the native protein in the absence of ligand (4). The terminal strand of the sheet (β5) is found to be more restricted in the mutant than in the WT. Separating the contributions of backbone and side-chain atoms to the overall increase in folded state entropy of the mutant, we found the relative weight of backbone disordering to be similar to or larger than side-chain mobility for α1, α2, β2, and L1. This is reversed in the first, third, and fourth strands of the sheet, where side-chain mobilization makes a comparatively larger contribution.

The catalytic pocket of hmAcP is located in a cleft between L1 and the β-turn connecting β2 and β3, which lies above L1 (Fig. 1). L1 adopts a cradle-like conformation, characteristic of phosphate-binding pockets, and functions in the binding of the phosphate moiety of the substrate and the stabilization of the transition intermediate. The active site is completed by a conserved arginine residue (Arg23), located at the interface between L1 and α1, which serves as the primary binder of the substrate phosphate group, and by an asparagine residue (Asn41), located in the β-turn, which binds the catalytic water molecule (4, 12–15). As shown in Fig. 2*D*, alterations in L4 length, caused either by deletion or by insertion of residues, lead to a marked decrease in enzymatic activity. As was observed for the thermodynamic and kinetic parameters, the effect is sharper and more extensive in the deletion mutants.

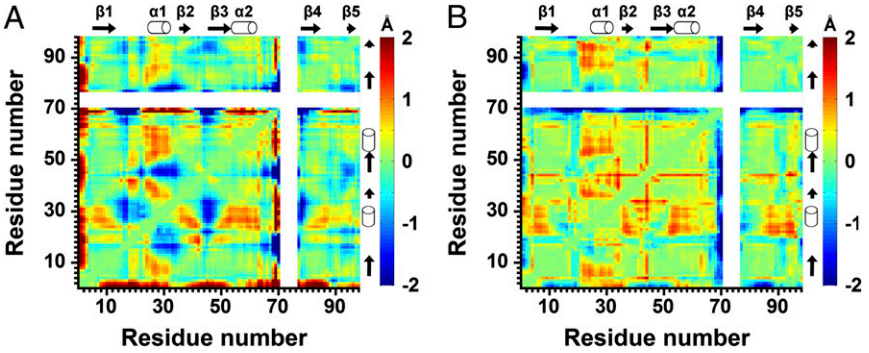


**Fig. 5.** Differences in mean interresidue distances $\left(\langle d_{ij}^{\Delta 6}\rangle - \langle d_{ij}^{WT}\rangle\right)$ (*A*) and their SDs $\left(\mathrm{SD}d_{ij}^{\Delta 6} - \mathrm{SD}d_{ij}^{WT}\right)$ (*B*) between Δ6 and WT hmAcP, as derived from analysis of snapshots from all-atom MD simulations. The white areas correspond to residues in L4 that are absent in the mutant. Secondary structures (defined by DSSP): α1 (residues 25–33), α2 (residues 55–63), β1 (residues 6–14), β2 (residues 36–40), β3 (residues 46–54), β4 (residues 78–85), and β5 (residues 94–97).

Uniquely here, however, the largest changes occur between the WT and the first deletion (Δ2) and insertion (∇2) mutants, suggesting an important role for L4 in substrate binding and/or hydrolysis, probably by affecting L1 conformation and dynamics, as indicated by the all-atom MD simulations.

Proteins are subjected to various energetic and topological frustrations (32–34), which are often related to function (refs. 35, 36 and see, e.g., refs. 37 and 38). Buried polar or charged residues, as well as water molecules and prosthetic groups, all of which are generally energetically unfavorable, often serve in ligand binding or catalysis. Similarly, the possession of long, flexible loops increases the entropic penalty for folding, but such loops are frequently involved in binding or required to enable localized or long-range conformational motions during catalysis or movement. Although such frustrating motifs can often be engineered out to increase stability or to facilitate folding, this almost always comes at the expense of function, reflecting the highly specific nature of the requirements it imposes. This is also the case for L4, where the native structure of the loop does not confer hmAcP with maximal conformational stability or kinetic foldability, but is required for optimal activity. Clearly, the energetic penalties associated with frustrating motifs cannot be too large, and one way to achieve this is by placing the motives in regions that do not have a critical impact on the integrity of the folded state structure, such as solvent-protruding loops like L4. A somewhat similar situation exists in dry globules, where the relatively moderate (and, likely, nonuniform) loss of packing interactions in the semifolded, expanded structures is partially compensated by an increase in backbone and side-chain entropy (ref. 39 and references therein). Finally, many proteins exhibit large-amplitude conformational fluctuations, which are conventionally assigned functional significance. In light of the results presented in this work, we propose that these fluctuations could also act to increase the stability of the proteins, implying that, despite the above, thermodynamic stability and function may not always be in conflict.

## Materials and Methods

**Cloning and Purification.** A synthetic gene encoding for the C21S variant of human muscle acylphosphatase was purchased from GeneArt and was cloned and altered by standard procedures to produce the truncation and insertion loop variants. The identity of the genes encoding for the WT and mutant proteins was confirmed by DNA sequencing. Protein expression was carried out in *Escherichia coli* BL21 cells transformed with pET45, following isopropyl β-D-1-thiogalactopyranoside (IPTG) induction ($OD_{600}$ ~0.6, 0.1 mM). Bacteria were lysed by sonication in a mixture containing 50 mM Tris·HCl (pH 8), 500 mM NaCl, 1 mM PMSF, protease inhibitor mixture (Calbiochem), DNase (1 mg/mL), and lyzosyme (40 U/mL). Following Ni capture (HiTrap chelating HP; GE Healthcare) and size-exclusion chromatography (HiLoad 16/60 Superdex 200; GE Healthcare), the hexahistidine tag at the N terminus of the proteins was removed by tobacco etch virus (TEV) cleavage. The samples were then subjected to another Ni-affinity purification step to remove the His-tag-TEV fragment, the (His-tagged) protease, and uncleaved protein. The purified (~95%) proteins [verified by proteolytic digestion and mass spectrometry (Reflex III MALDI-TOF; Bruker)] were flash-frozen in a buffer containing 50 mM Tris·HCl (pH 8.0), 0.1 M NaCl, and 10% glycerol and stored at −80 °C. Immediately before use, they were dialyzed against 50 mM sodium acetate buffer (pH 5.5), which was used in all measurements performed in this study.

**CD and Fluorescence Spectroscopy.** CD spectra were recorded on a Chirascan spectropolarimeter (Applied Photophysics) calibrated with a solution of ammonium *d*-10-camphorsulfonate. Far-UV CD spectra were acquired using 1-mm path-length cuvettes, a step size of 0.5 nm, a bandwidth of 1 nm, and a time constant of 1 s. Protein concentration was 0.4 mg/mL. With the exception of the thermal denaturation experiments, the temperature was 25 °C. Near-UV CD spectra were recorded on 0.4 mg/mL protein solutions at 25 °C with 10-mm path-length cuvettes, a step size of 0.5 nm, a bandwidth of 1 nm, and a time constant of 1 s. Fluorescence emission spectra (0.4 mg/mL protein, 25 °C; $\lambda_{ex}$ = 280 nm) were recorded with a FluoroLog 3 (HORIBA Scientific) spectrofluorometer in 1-cm optical path-length cuvettes. In the experiments conducted in the presence of ANS (100 μM; protein concentration, 0.4 mg/mL) the excitation wavelength was 375 nm. Before recording, the proteins were let to interact with the dye for 1 h. Following acquisition, the spectra (CD and fluorescence) were corrected for buffer (and dye, in the experiments involving ANS) contributions, averaged, and smoothed using sliding windows of 1.5 nm (far- and near-UV CD) or 3 nm (fluorescence).

**Stability.** Equilibrium urea denaturation curves were obtained by monitoring the mean residue ellipticity at 222 nm (time constant, 60 s) of equilibrated protein samples (0.4 mg/mL) in 1-mm path-length cuvettes at 25 °C. The data were analyzed according to ref. 40. Differential scanning calorimetric measurements were conducted on degassed samples (protein concentration, 1 mg/mL) with a VP differential scanning microcalorimeter (GE Healthcare). Heating and cooling cycles were carried out over a temperature range of 20–90 °C at a scanning rate of 1 °C/min. The reversibility of the transition (determined by comparison of the endotherms exhibited by consecutive scans) was between >95% and ~80%, depending on the protein variant. Thermodynamic parameters ($T_m$, $\Delta H_m$, and $\Delta C_p$), were obtained from fitting of the molar heat capacity vs. temperature plots, assuming a temperature-independent $\Delta C_p$ (41); the β ($\Delta H_{vH}/\Delta H_m$) values were between 1.01 and 1.08. Denaturation/renaturation curves, obtained by following the change in the 222-nm ellipticity as a function of the temperature, yielded $T_m$ values that were essentially identical to those determined by the scanning calorimetric measurements. The temperature was increased (or decreased) over the range of 20–90 °C at a rate of 1 °C/min. Once the desired temperature was reached, the samples were let to equilibrate for 60 s and the ellipticity was recorded over an additional period of 60 s.

**Kinetic Analyses.** The folding and unfolding kinetics were followed at 25 °C by monitoring the intrinsic emission fluorescence changes of the proteins ($\lambda_{ex}$ = 280 nm; $\lambda_{em}$ > 315 nm), using Applied Photophysics SW-17 stopped-flow and SLM Aminco MC200 spectrofluorometers, respectively. The rate constants (at a given concentration of urea) were obtained by fitting the traces to single-exponential curves. The rate constants in the absence of denaturant and the kinetic *m*-values were derived from the plots of the natural logarithms of the rate constants vs. denaturant concentration.

**Activity Assays.** The (specific) activity of the proteins (50 μM) was determined at saturating substrate concentration (20 mM) by monitoring the hydrolysis of benzoyl phosphate (synthesized following ref. 42), as described (43).

**Molecular Dynamics Simulations.** Off-lattice Gō-model simulations were performed on a series of hmAcP loop-length variants with the Δ6-loop-truncated mutant serving as a template. The series was generated by insertion of stretches containing up to 26 Cα beads into the loop. Noncovalent attractive interactions made by residues within the modified region in the loop were omitted. Otherwise, the simulations were carried out as described (44, 45). Potential of mean force curves and stability analyses were carried out using the Weighted Histogram Analysis Method (46). Folding and unfolding rates were determined using the mean passage time method (45).

All-atom MD simulations were performed using GROMACS (v4.5.4; ref. 47). The proteins were solvated in a periodic box with pre-equilibrated TIP3P water molecules in the presence of sodium and chloride ions. Side-chain potentials were used as defined in the AMBER99SB-ILDN force field; covalent bonds were constrained using the Linear Constraint Solver algorithm. Equations of motion were integrated using a leapfrog algorithm with a time step of 2 fs. Before the runs, the proteins were subjected to short energy minimization, using steepest descent algorithm. The solvated systems were then equilibrated by gradually increasing the temperature to 300 K. Equilibration proceeded in two phases, with the first one conducted under an NVT ensemble and the second under an NPT ensemble, both for 100 ps. All simulations were carried out at constant pressure (1 atm) and temperature (300 K); the latter was controlled using a modified Berendsen coupling procedure. Periodic boundary conditions were used in all three dimensions. WT hmAcP and the Δ6 mutant were subjected to 10 runs of 20 ns. Analysis was performed on 1,500 snapshots (150 from each run) obtained during the last 15 ns of the trajectories. Solvent-accessible surface areas were calculated using a GROMACS module based on the double cube lattice method.

**ACKNOWLEDGMENTS.** We thank Drs. Shira Albeck, Orly Dym, Yoav Peleg, and Tamar Unger (Israel Structural Proteomic Center, Weizmann Institute of Science) for their help in producing the proteins used in this study; Drs. Vlad Brumfeld and Yosef Scholnik for assistance with the CD analyses; and Drs. Amnon Horovitz and Gideon Schreiber for critical discussions and helpful suggestions. This work was supported by the Carolito Stiftung (Z.R.), the Kimmelman Center for Biomolecular Structure and Assembly (Y.L. and Z.R.), and the Israel Science Foundation (Y.L.). Y.L. holds the Lillian and George Lyttle Career Development Chair.

1. Zhou HX (2004) Loops, linkages, rings, catenanes, cages, and crowders: Entropy-based strategies for stabilizing proteins. *Acc Chem Res* 37(2):123–130.
2. Stefani M, Taddei N, Ramponi G (1997) Insights into acylphosphatase structure and catalytic mechanism. *Cell Mol Life Sci* 53(2):141–151.
3. Pastore A, Saudek V, Ramponi G, Williams RJP (1992) Three-dimensional structure of acylphosphatase. Refinement and structure analysis. *J Mol Biol* 224(2):427–440.
4. Thunnissen MMGM, Taddei N, Liguri G, Ramponi G, Nordlund P (1997) Crystal structure of common type acylphosphatase from bovine testis. *Structure* 5(1):69–79.
5. Corazza A, et al. (2006) Structure, conformational stability, and enzymatic properties of acylphosphatase from the hyperthermophile Sulfolobus solfataricus. *Proteins* 62(1): 64–79.
6. Taddei N, et al. (1994) Equilibrium unfolding studies of horse muscle acylphosphatase. *Eur J Biochem* 225(3):811–817.
7. van Nuland NAJ, et al. (1998) Slow folding of muscle acylphosphatase in the absence of intermediates. *J Mol Biol* 283(4):883–891.
8. Chiti F, et al. (1998) Structural characterization of the transition state for folding of muscle acylphosphatase. *J Mol Biol* 283(4):893–903.
9. Chiti F, et al. (1999) Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding. *Nat Struct Biol* 6(11):1005–1009.
10. Vendruscolo M, Paci E, Dobson CM, Karplus M (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature* 409(6820):641–645.
11. Satchell DPN, Spencer N, White GF (1972) Kinetic studies with muscle acylphosphatase. *Biochim Biophys Acta* 268(1):233–248.
12. Taddei N, et al. (1994) Arginine-23 is involved in the catalytic site of muscle acylphosphatase. *Biochim Biophys Acta* 1208(1):75–80.
13. Taddei N, et al. (1996) Looking for residues involved in the muscle acylphosphatase catalytic mechanism and structural stabilization: Role of Asn41, Thr42, and Thr46. *Biochemistry* 35(22):7077–7083.
14. Taddei N, et al. (1997) Structural and kinetic investigations on the 15-21 and 42-45 loops of muscle acylphosphatase: Evidence for their involvement in enzyme catalysis and conformational stabilization. *Biochemistry* 36(23):7217–7224.
15. Cheung YY, et al. (2005) Crystal structure of a hyperthermophilic archaeal acylphosphatase from Pyrococcus horikoshii—structural insights into enzymatic catalysis, thermostability, and dimerization. *Biochemistry* 44(12):4601–4611.
16. Paci E, Vendruscolo M, Dobson CM, Karplus M (2002) Determination of a transition state at atomic resolution from protein engineering data. *J Mol Biol* 324(1):151–163.
17. Arad-Haase G, et al. (2010) Mechanical unfolding of acylphosphatase studied by single-molecule force spectroscopy and MD simulations. *Biophys J* 99(1):238–247.
18. Nagi AD, Regan L (1997) An inverse correlation between loop length and stability in a four-helix-bundle protein. *Fold Des* 2(1):67–75.
19. Viguera AR, Serrano L (1997) Loop length, intramolecular diffusion and protein folding. *Nat Struct Biol* 4(11):939–946.
20. Ladurner AG, Fersht AR (1997) Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *J Mol Biol* 273(1):330–337.
21. Scalley-Kim M, Minard P, Baker D (2003) Low free energy cost of very long loop insertions in proteins. *Protein Sci* 12(2):197–206.
22. Wang L, Rivera EV, Benavides-Garcia MG, Nall BT (2005) Loop entropy and cytochrome c stability. *J Mol Biol* 353(3):719–729.
23. Chan HS, Dill KA (1989) Intrachain loops in polymers: Effects of excluded volume. *J Chem Phys* 90(1):492–509.
24. Chan HS, Dill KA (1990) The effects of internal constraints on the configurations of chain molecules. *J Chem Phys* 92(5):3118–3135.
25. Tanford C (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv Protein Chem* 24:1–95.
26. Schellman JA (1978) Solvent denaturation. *Biopolymers* 17(5):1305–1322.
27. Myers JK, Pace CN, Scholtz JM (1995) Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci* 4(10): 2138–2148.
28. Pace CN (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131:266–280.
29. Otzen DE, Oliveberg M (2004) Correspondence between anomalous m- and DeltaCp-values in protein folding. *Protein Sci* 13(12):3253–3263.
30. Schlitter J (1993) Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem Phys Lett* 215(6):617–621.
31. Andricioaei I, Karplus M (2001) On the calculation of entropy from covariance matrices of the atomic fluctuations. *J Chem Phys* 115(14):6289–6292.
32. Onuchic JN, Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* 14(1): 70–75.
33. Dill KA (1999) Polymer principles and protein folding. *Protein Sci* 8(6):1166–1180.
34. Thirumalai D, Woodson S (1996) Kinetics of folding of proteins and RNA. *Acc Chem Res* 29:433–439.
35. Kapon R, Nevo R, Reich Z (2008) Protein energy landscape roughness. *Biochem Soc Trans* 36(Pt 6):1404–1408.
36. Gruebele M (2005) Downhill protein folding: Evolution meets physics. *C R Biol* 328(8): 701–712.
37. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG (2011) On the role of frustration in the energy landscapes of allosteric proteins. *Proc Natl Acad Sci USA* 108(9):3499–3503.
38. Li W, Wolynes PG, Takada S (2011) Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins. *Proc Natl Acad Sci USA* 108(9):3504–3509.
39. Thirumalai D, Liu Z, O'Brien EP, Reddy G (2013) Protein folding: From theory to practice. *Curr Opin Struct Biol* 23(1):22–29.
40. 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(21):8063–8068.
41. Chiti F, et al. (1998) Conformational stability of muscle acylphosphatase: The role of temperature, denaturant concentration, and pH. *Biochemistry* 37(5):1447–1455.
42. Camici G, Manao G, Cappugi G, Ramponi G (1976) A new synthesis of benzoyl phosphate: A substrate for acyl phosphatase assay. *Experientia* 32(4):535–536.
43. Ramponi G, Treves C, Guerritore A (1966) Continuous optical assay of acylphosphatase with benzoylphosphate as substrate. *Experientia* 22(11):705–706.
44. Clementi C, Nymeyer H, Onuchic JN (2000) Topological and energetic factors: What determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J Mol Biol* 298(5):937–953.
45. Hagai T, Levy Y (2008) Folding of elongated proteins: Conventional or anomalous? *J Am Chem Soc* 130(43):14253–14262.
46. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13(8):1011–1021.
47. Van Der Spoel D, et al. (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26(16):1701–1718.