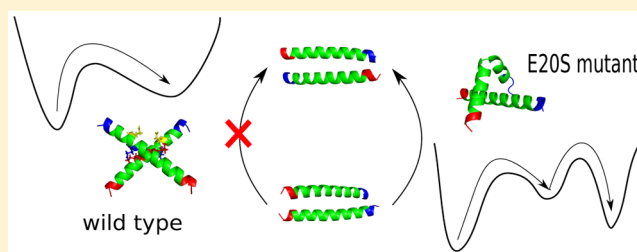Article

Transforming the Energy Landscape of a Coiled-Coil Peptide via Point Mutations

Konstantin Röder* and David J. Wales*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, U.K.

Ⓢ *Supporting Information*

ABSTRACT: We analyze the effect of point mutations on the energy landscape of a coiled-coil peptide, GCN4-pLI, where the native state is a parallel tetrameric configuration formed from two identical dimers. Experimentally, a single mutation, E20S, supports both antiparallel and parallel structures. Here, we analyze the potential energy landscapes of the dimeric units for the parent sequence and four mutants, namely E20S, E20A, E20P, and E20G. Despite sharing characteristic funnels containing the parallel and antiparallel structures, the point mutations change some parts of the landscape quite dramatically, and we predict new intermediate structures and characterize the associated heat capacities. For the mutants we predict that kinked intermediate structures facilitate the transition between parallel and antiparallel morphologies, in contrast to the parent sequence. Furthermore, we predict a change from a multifunnel energy landscape in the E20S mutant to a landscape dominated by an underlying single funnel in the parent sequence, with accompanying heat capacity signatures. Our results imply that changes in the landscape due to mutations might provide useful tools for functional protein design.

1. INTRODUCTION

In principle, the energy landscape of a protein contains all the information required to understand its structure, thermodynamics, and kinetics.[1−11] An exploration of the energy landscape therefore provides insight into experimentally observable properties.[12−14] The computational potential energy landscape framework employs discrete path sampling[15,16] to construct kinetic transition networks,[17,18] exploiting efficient geometry optimization procedures. The use of geometry optimization facilitates sampling that is largely independent of kinetic traps caused by high energy barriers, and hence permits us to explore a wide range of time scales. Previous studies have used this approach for various molecules, ranging from intrinsically disordered proteins[19] and protein folding[20,21] to the assembly of biomolecules.[22,23]

The changes a protein or peptide undergoes upon mutation will cause systematic changes in the underlying potential energy landscape. Here we analyze one particular example in detail. Some studies have hinted at the inherent stability of the topography of energy landscapes for naturally occurring proteins with respect to perturbations, such as point mutations.[24−28] One suggestion is that if a stable fold is still achieved after a mutation, this folded mutated structure will closely resemble the wild type native state.[24] However, we should also consider the possibilities presented by a multifunnel energy landscape, where there exist competing potential energy funnels, leading to alternative morphologies. A point mutation in such a system may not change the overall topography, but is likely to change the depth (enthalpy) and width (entropy) of funnels, as well as the relative energies, which in turn will

perturb the observed behavior. Understanding such changes is important, as many human diseases are affected or caused by mutations.[29−32]

The example we consider in the present work is a coiled-coil peptide. Coiled-coils are a common motif observed in biological systems,[33,34] and consequently they have been the target of many studies to gain insight into the structure and further the design of functional systems.[35−37] Particular attention has been paid to the assembly of coiled-coils in oligomers of different sizes, the electrostatic and nonpolar interactions involved in the formation of parallel and antiparallel states, and the influence of mutations on the stability.[38−42] These studies agree that there is competition between different oligomer sizes, as well as between parallel and antiparallel configurations, which is likely caused by a multifunnel energy landscape.[37] FRET experiments even allowed the observation of two such competing states for one example.[43]

The system investigated in this study is a peptide derived from the leucine zipper of the yeast transcription factor GCN4.[44] The peptide forms tetramers composed of two identical dimers of helices, where each monomer has the same sequence (MKQIED-KLEEILS-KLYHIEN-ELARIKK-LLG). Experiment[45] has shown that a single point mutation, E20S, in the helices leads to crystals supporting parallel and antiparallel structures, whereas the parent sequence exclusively crystallizes in a parallel arrangement. In addition to their

experimental work, Yadav et al.[45] used replica exchange molecular dynamics (MD) simulations to investigate crystallization. Their results suggest an inherent tendency to form parallel structures in the parent sequence, which they attributed to the formation of a strong salt bridge between Lys15 and Glu20. This interaction favors the parallel structure, whereas in the E20S mutant the absence of the salt bridge lowers the stabilization relative to the antiparallel alignment. The close match in energy, and the similar hydrophobic interaction surfaces, then permit both structures to be observed. This delicate balance indicates that the key interactions are unlikely to be optimized within a single conformation and strongly suggest the existence of multiple funnels on the energy landscape. It was also suggested that the existence of multiple topologies and the changes observed upon mutation are similar in other coiled-coil systems, and further evidence has been provided by more recent studies.[43]

Here we consider the global energy landscape of coiled-coils to analyze the competition between configurations, and the effect of mutations on its organization. We used discrete path sampling[15,16] for the parent sequence and the E20S mutant, along with three other mutations, namely E20A, E20P, and E20G. These variations were suggested by the E20S results as discussed below. In addition, we investigate intermediate structures and possible folding paths between configurations, providing new insight into the effects of mutations.

This investigation also provides further evidence that multifunnel energy landscapes give rise to more complex behavior, and that changes due to a point mutation can be elucidated using the computational potential energy landscape framework. Additionally, it is clear that multifunnel energy landscapes, which are characterized by high energy barriers and broken ergodicity, will pose a significant challenge to conventional simulation techniques.

2. METHODS

2.1. Starting Points and Force Field. Initial structures for the E20S mutant were taken from PDB entries 2CCE and 2CCF.[45] The lowest energy parallel and antiparallel structures from the E20S landscape were mutated to yield starting points for the parent sequence and the three additional mutants. We only use the resolved residues in the PDB entries, which means that each helix contains 30 amino acids, and all helices are identical in sequence. The strands were left uncapped and no restraint force was applied between the helices. All our calculations used the AMBER ff99SB force field,[46−49] properly symmetrized,[50,51] with an implicit generalized Born solvation model ($igb = 2$)[52,53] using infinite interaction cutoffs and the Debye−Hückel approximation for salt (0.1 M).[54] As the use of an implicit solvent model might affect the structural ability, we ran MD simulations in explicit solvent for key structures and probed changes in configuration as well as changes in relative energy (see Supporting Information for protocol and results). Unless otherwise stated, the results from these simulations confirm the picture derived from discrete path sampling.

2.2. Exploration of the Landscape. We explored the potential energy landscapes using discrete path sampling[15,16] to construct kinetic transition networks[17,18] consisting of local minima and the transition states that connect them. Transition states were located using the doubly nudged[55] elastic band[56,57] algorithm and hybrid-eigenvector-following.[58] Connected minima were obtained by characterizing approximate steepest-descent paths. Stationary points were optimized using the

L-BFGS minimizer with a RMS force convergence of 10^{-6} kcal mol^{-1}. After an initial discrete path between the end points was found, the landscape was further sampled to remove artificially high barriers and kinetic traps using the SHORTCUT[22,59] and UNTRAP[22] schemes within PATH-SAMPLE.[60] As the systems under consideration are relatively large, GPU acceleration[61,62] was used to speed up calculations via the GPU interface[63] for OPTIM[64] and AMBER12.

Usually all new stationary points are accepted and included in the database in the above procedure, which has been used in a number of previous studies. However, before an initial path between products and reactants has been found, a list of Euclidean distances between the minima needs to be maintained, which becomes expensive as the database grows. Instead, after each set of connection attempts between two minima, the best path between them was computed with a Dijkstra-based selector[65−67] using the transition states and minima local to the connection attempt. Only stationary points on this best path were then included in the database. This procedure corresponds to the new DUMPBESTPATH keyword in OPTIM.[64] Additionally, for larger systems it seems to be beneficial to bridge bigger gaps between structures before smaller ones. Only missing connections with a Euclidean distance between the locally permutational aligned[68] end points larger than a threshold value, in this case 20 Å, were considered in initial connection attempts. When no such gaps remained, the threshold was removed and sampling was continued as usual. This procedure corresponds to the new MINGAP keyword in PATHSAMPLE.[60] It leads to a sparser initial sampling, which improves the efficiency and reduces the computational effort, and produces a smaller number of stationary points in the initial database. However, this effect is mitigated by more extensive sampling after an initial path has been found.

2.3. Tetramer configurations. The dimers form four-helix bundles (tetramers) for both the E20S mutant and the parent sequence with their helix axis parallel to a C_2 axis at the tetramer center.[45] Exploration of the entire tetramer landscape would be quite expensive, and therefore a different approach was chosen to compare tetrameric structures. Using the dimers already characterized, two dimers were combined using PyMOL[69] and the resulting structures were locally minimized. Different dimer configurations from various parts of the landscape were used to increase the structural diversity. As discussed below, the structures obtained closely match the observed crystal structures, both in the parent sequence and the mutant, and appear to provide a concise explanation for the experimentally observed behavior.[45]

3. RESULTS AND DISCUSSION

3.1. Order Parameters. The potential energy landscape is visualized using disconnectivity graphs.[70,71] Two structural order parameters were defined to differentiate the different regions of the landscape. The first order parameter, q_1, reports on whether a structure has antiparallel or parallel components. For each helix, a vector is defined along its length as follows: the two points defining the vector are the geometrical centers of the α and the carbonyl carbons of residues 6 to 9 and 24 to 27, respectively, as shown in Figure 1. The angle, ϕ, between the two vectors then defines q_1 as
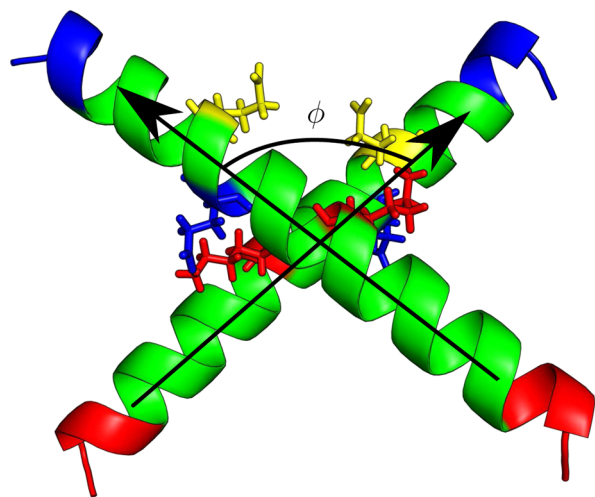
$$q_1 = \frac{1}{2}(1 - \cos \phi) \tag{1}$$

Figure 1. Vectors defining q_1 are shown for both helices connecting residues 6 to 9 to residues 24 to 27. The figure was created using PyMOL.[69]

The second order parameter, q_2, is designed to identify a subclass of dimeric structures, characterized by a kink in one of the helices, which wraps partially around the second helix. An example is shown in Figure 2a. This kinking partially unwinds the helix, and hence extends the backbone, as shown in Figure 2b. Using the difference in distance, d, between the α carbons in residues 17 and 21, q_2 is defined as

$$q_2 = \frac{d - d_{\min}}{d_{\max} - d_{\min}} \tag{2}$$

where $d_{\max}$ and $d_{\min}$ denote the maximum and minimum difference, respectively, which scales the order parameter to the range [0,1].

3.2. Energy Landscape of the E20S Mutant. The disconnectivity graph for the E20S sequence is shown in Figure 3. The coloring schemes are derived from the order parameters q_1 and q_2. This potential energy landscape exhibits three distinct regions. The parallel and antiparallel configurations are similar in energy, with the parallel structures slightly lower. Furthermore, the intermediate structures, which are mainly kinked, mediate the transition between the two states with low associated energy barriers. The kinking decreases the solvent exposed area of the protein and allows for more contacts

between the strands, but incurs an energy penalty for helix breaking. For conformational switches the kinetics play an important role. While there needs to be an energy balance between antiparallel and parallel configurations to allow for the population of both states, small enough energy barriers are needed to prevent kinetic trapping in one configuration. This topography means that the two structures ought to be observed in equilibrium, with both configurations significantly populated.

3.3. Energy Landscape for Dimers of the Parent Sequence. Disconnectivity graphs for the parent sequence are illustrated in Figure 4. The main difference for the E20S mutant is the occurrence of kinked structures in the landscape, which do not exist for the parent sequence. These structures provide a low-energy pathway between parallel and antiparallel configurations, which is not possible for the parent sequence. The likely reason for this difference is that the kink occurs near the point mutation. For the parent sequence this feature means that, besides the energy penalty for helix breaking, it is also necessary to expose more hydrophobic residues to the solvent. Mutating kinked structures found for the E20S mutant back to the parent sequence, and then optimizing the geometry, results in structures with potential energies up to 100 kcal mol^{-1} higher than for the low energy parallel dimers. Comparing the shortest pathway between both states on either landscape shows that they appear similar in terms of q_1 (see Supporting Information Figure S5). However, using q_2 to distinguish minima along the path shows that the kinked structures play a significant role in the high energy rearrangements (see Supporting Information Figure S6).

The presence of different salt bridges causes the appearance of small funnels in the antiparallel region. These salt bridges correspond to Lys15–Glu20, as observed in the parallel structures, and the formation of a new salt bridge Lys15–Glu22. If the Lys15–Glu20 salt bridge is retained, one helix must rotate, which explains the lack of hydrophobic core formation in the tetramer structures. However, the formation of the salt bridges does not stabilize the antiparallel structures significantly. Comparing the energies of intermediate structures, and the bottom of the parallel and antiparallel funnels, the barriers for the E20S mutant are more symmetrical (see Supporting Information page S3 and S4).

3.4. Tetramer Formation from Dimers. The assembly of tetramers from dimers in solution is based on the formation of a hydrophobic core, which excludes solvent molecules and entropically favors the assembly of tetramers, along with strong
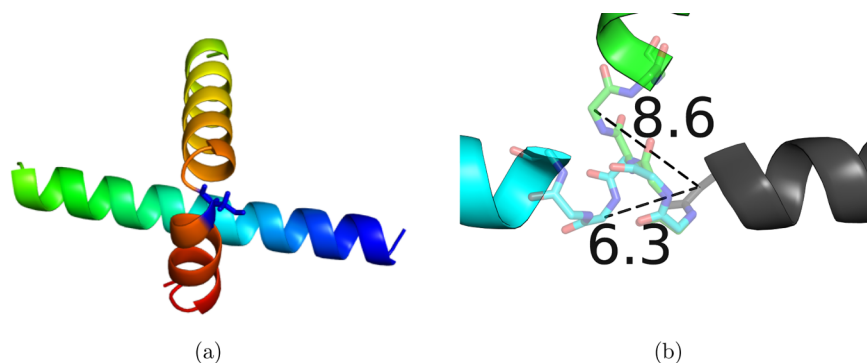

(a)

(b)

Figure 2. A number of kinked structures, similar to the one shown as structure a, are found in the intermediate region between the antiparallel and parallel regions on the landscape of the E20S mutant. The position of the mutation is indicated by the blue colored residue. (b) The distances defining q_2 for the kinked structure (green) and the straight helix (cyan) in Å. The gray parts of the helices are superimposed to allow a comparison of the distances. The figures were created with PyMOL.[69]
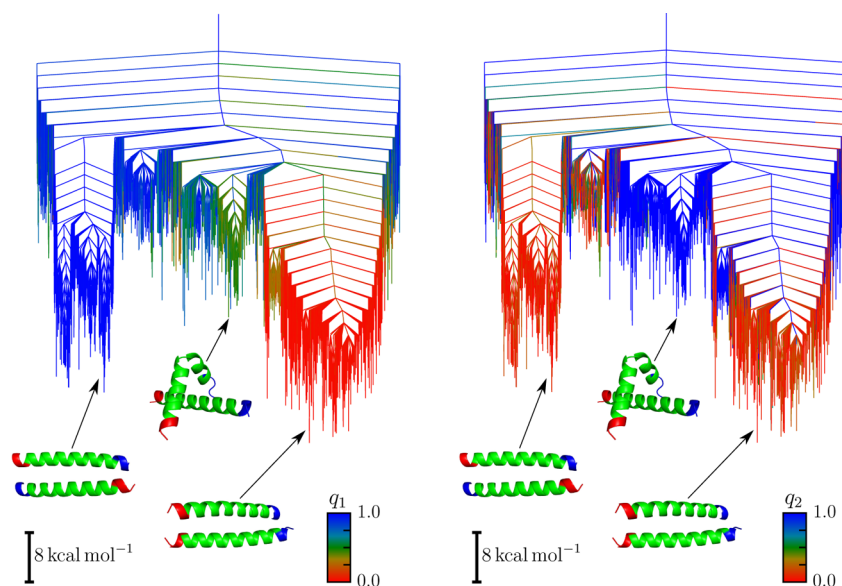
Figure 3. Disconnectivity graphs for the E20S mutant. (Left) order parameter q_1: separation into two funnels for the parallel (red) and antiparallel (blue) configurations is observed with some intermediate structures (green). (Right) order parameter q_2: the kinked structures observed in the intermediate region (blue), but not in the parallel and antiparallel funnel (red), are clearly visible.
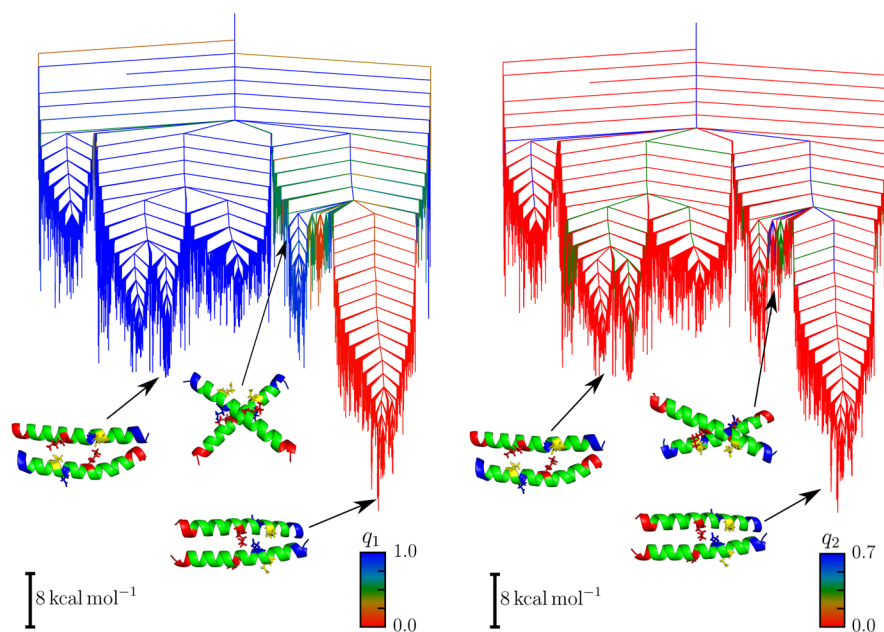

Figure 4. Left: The energy landscape of the parent sequence exhibits many antiparallel features (blue) for the dimer. The lowest energy funnel still contains the parallel configurations. Right: Using order parameter q_2, we see that no kinked structures exist in the energy landscape of the parent sequence.

interhelical side chain interactions that favor assembly from an energetic perspective.

3.4.1. Parent Sequence Tetramers. Considering the low energy parallel structures, two dimers can be combined to form a tetramer with four identical faces based on the salt bridge between Lys15 and Glu20. This regular pattern is shown in Figure 5.

An analysis of all the polar interactions[69] between the helices shows a much larger number of contacts than for the antiparallel tetramers, with the Glu20-Lys15 interaction as the central unit. The faces of the tetramer are essentially identical and exhibit a large number of stabilizing interactions, as shown in Figure 6.

All these polar interactions are on the outside of the tetramer; that is, they are solvent exposed. The hydrophobic core is formed excluding all solvent, while interactions of the polar groups with solvent are still possible. Furthermore, the observed interactions closely resemble those seen in the crystal structure (PDB entry 1GCL),[44] with an RMSD of 1.241 Å, and the main deviation is seen in the terminal region, which is the most flexible. The formation of these highly symmetrical structures is also an example of the principle of maximum symmetry.[72] Higher symmetry in biomolecules may be associated with deeper funnels on the landscape,[73] potentially leading to a stronger preference for such structures. As the dimers exhibit a high degree of symmetry, and only one
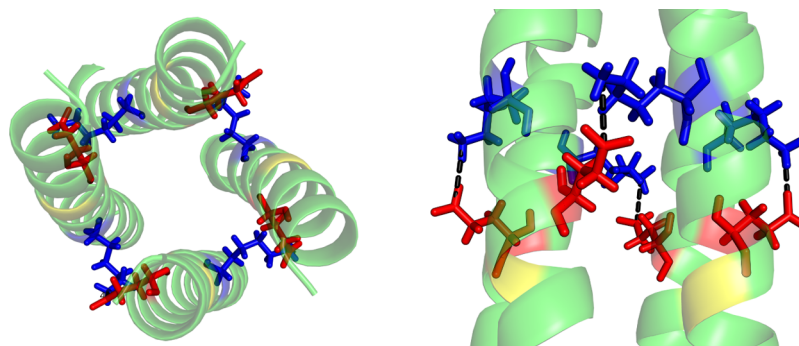
Figure 5. Pattern of salt bridges based on Glu20 (red) and Lys15 (blue) is shown from different angles. The arrangement has relatively high symmetry, with essentially identical tetramer faces. The yellow-colored residues in the helix are the Glu22 residues. Figure created with PyMOL.[69]
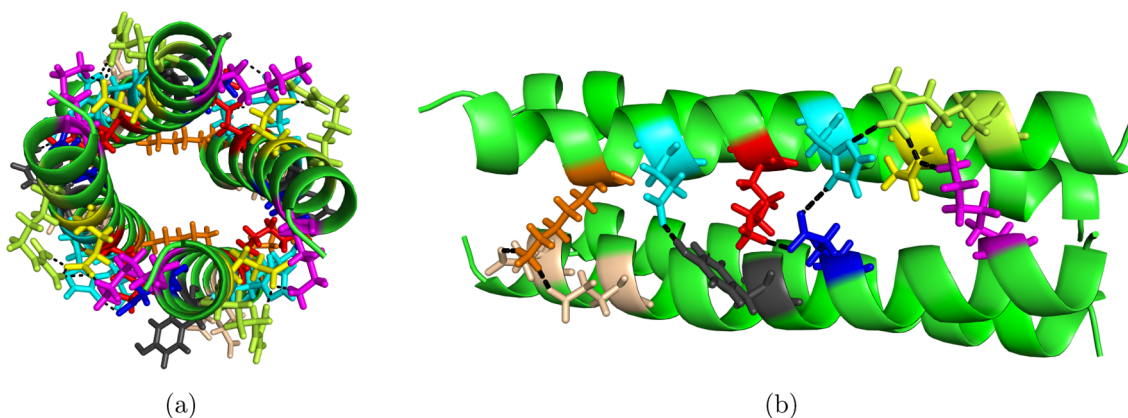

(a) (b)

Figure 6. Different views of the parallel tetramer with all the polar interactions indicated (black, dashed lines) reveal identical faces and a fully formed hydrophobic core. For clarity only residues in one face are shown in structure b. Figure created with PyMOL.[69]
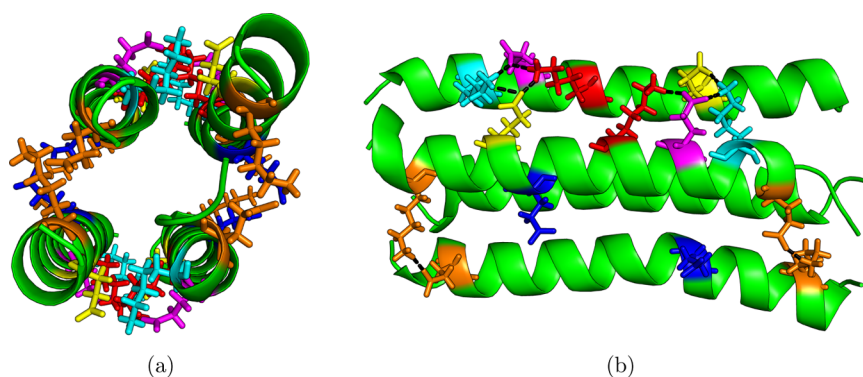

(a) (b)

Figure 7. Different views of the antiparallel tetramers of the first category with all polar interactions shown (black dashed lines). There are two distinct faces of the tetramer with distinctly different interactions. For clarity only residues in two faces are shown in structure b. Figure created with PyMOL.[69]

possible alignment, there are no competing structures for the parallel arrangement, and a deep funnel in the landscape is the likely result.

The antiparallel structures are more complicated and more heterogeneous, due to the different dimer configurations observed. This variety is based on symmetry breaking, which allows for competing antiparallel conformations in the dimers, and consequently in the tetramers. In general, the morphologies fall into two categories. Structures in the first category have a hydrophobic core similar to the parallel configurations, but there is a change in the polar interactions observed. An example is shown in Figure 7.

Despite the formation of the hydrophobic core, the number of favorable interactions is much lower than for the parallel configuration. In particular, there are fewer additional interactions upon formation of the tetramer than in the corresponding parallel configuration. This lack of native contacts leads to an instability for the tetramers, which in turn results in a higher predicted occupancy for the corresponding dimer configuration. Hence such structures are inherently less stable than the parallel configurations in solution. This type of antiparallel structure is formed for the dimers with salt bridges between Lys15 and Glu22. The second type is formed by the antiparallel structures close to the parallel structures on the energy landscape, in terms of rearrangements
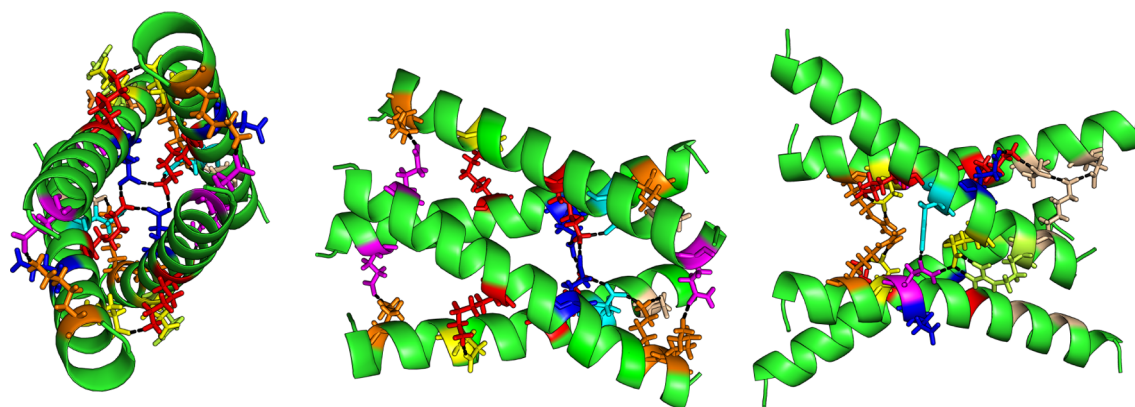
Figure 8. Different views of the antiparallel tetramers of the second category illustrate polar interactions within the core of the assembled structures (black dashed lines). There is no hydrophobic core, but instead a variety of bonding patterns are observed. Figure created with PyMOL.[69]

via single transition states. These configurations exhibit very different behavior due to the geometrical restrictions caused by the Lys15−Glu20 salt bridge. Some representative structures are shown in Figure 8.

The antiparallel tetramers in this case have no hydrophobic core and instead a variety of polar interactions can exist within the assembled structure. These changes result in destabilization of the structure, unfavorable interactions with the solvent, and little entropic driving force for assembly. Hence tetrameric, antiparallel configurations are unstable compared to the parallel conformation, and antiparallel parent sequence structures do not exist. Analogous destabilization due to hydrophilic groups has previously been observed and modeled in other systems.[74,75]

3.4.2. E20S Tetramers. The tetramers formed with the E20S mutation are very similar to the crystal structures previously observed.[45] The pattern of contacts between the helices is symmetrical for the parallel structures, and there are two different faces for the antiparallel configuration. The absence of the Lys15−Glu20 salt bridge destabilizes the parallel configuration and lowers the relative energy of the two observed conformations. The combination of a parallel or antiparallel dimer with a kinked structure is also stable in implicit and explicit solvent, allowing for kinked intermediates providing transition pathways for tetramers similar to the dimer intermediates.

3.5. Heat Capacity Curves and Landscape Entropies. The databases of minima and transition states constructed for the E20S mutant and the parent sequence provide further insight into some features of the underlying landscape, as well as a connection to observable properties.

The first property considered here is the landscape entropy[76−78] of the two systems. The landscape entropy is the heat capacity due to the density of local minima, excluding vibrational degrees of freedom. For the parent sequence a broad spectrum with two peaks is observed, whereas the E20S mutant only exhibits one sharp peak (Figure 9). This clear feature is associated with two distinct, observable states, which corresponds to well separated funnels on the energy landscape. This view is supported by the distribution of potential energy for the local minima (see Supporting Information, Figures S7 and S8), which shows a bimodal distribution for the E20S mutant, clearly distinguishing the parallel and antiparallel structures, with the kinked structures populating the intermediate energies. In contrast, the parent sequence shows,
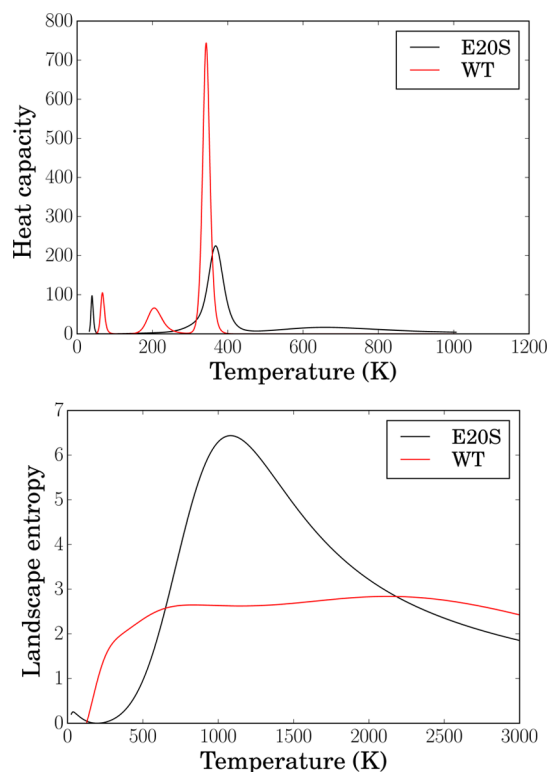

Figure 9. Plots of the landscape entropy (top) defined by the density of local minima, and the corresponding heat capacity curves (bottom) for the parent sequence and the E20S mutant, reveal clear differences in the underlying energy landscape, which should be observable in experiment. The temperature scale reflects the energy density of local minima. The temperature range of 3000 K is roughly 10 kcal mol^{-1}, which corresponds to the energy gap between the parallel and antiparallel funnels in the mutant landscape.

not only a broader peak in terms of the landscape entropy, but also a unimodal distribution of minima. This difference provides evidence that the E20S mutant possesses a multifunnel energy landscape, whereas the parent sequence, although it exhibits some distinct funnels, is closer to an energy landscape with one underlying funnel. This result shows how point mutations may induce a transition from multifunnel landscapes to landscapes dominated by an underlying single funnel, and vice versa, which might be exploited in engineering
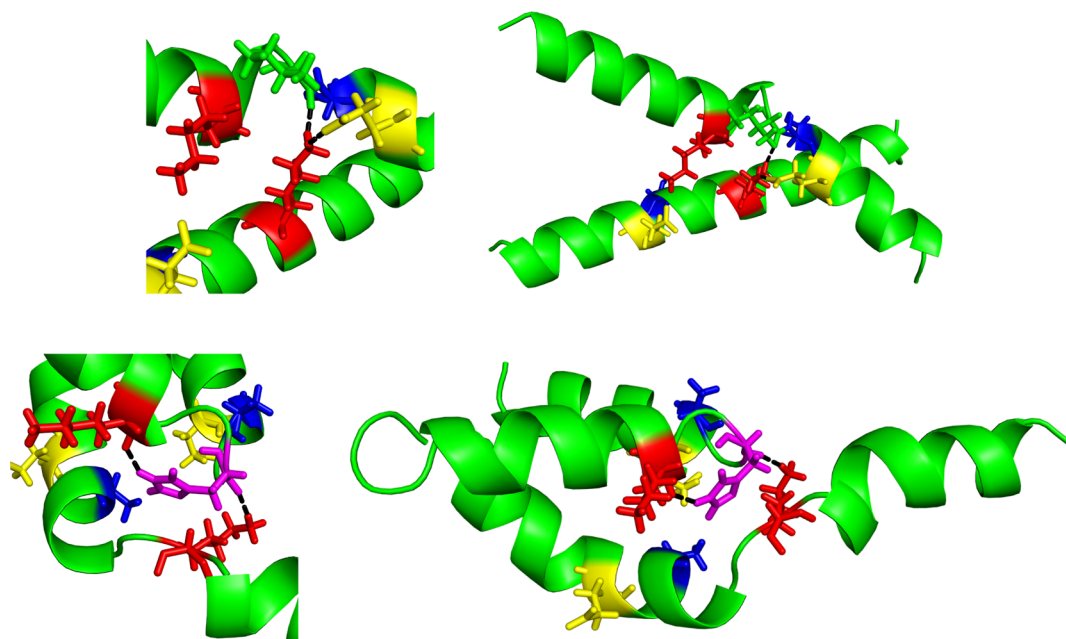
Figure 10. Comparison of kinked intermediates for the E20S mutant (top) and the E20A mutant (bottom). Figure created with PyMOL.[69]

applications, but at the same time could prove to be important in disease-related mutations.

A similar pattern is observed for the normal-mode frequency distributions (see Supporting Information Figures S9 and S10). The E20S mutant exhibits a bimodal frequency distribution, corresponding to the two funnels observed. This pattern reinforces the suggestion proposed above that a transition is observed from a multifunnel landscape to a landscape, dominated by an underlying single funnel.

The changes in the landscape also affect the heat capacities. As seen in Figure 9, the peaks shift in temperature. The shifts for the heat capacity peaks, as well as the energy scale for the landscape entropy, can be understood in terms of the entropy contribution of the vibrational normal modes. For high energy minima the modes are softer, leading to the observed shifts for the high temperature peaks in C_v.

To obtain an experimentally relevant temperature scale for the entropy we can convolute the landscape entropy with the contribution from local vibrational densities of states.

Analyzing the heat capacity peaks further, we can attempt to identify the local minima contributing to the phase transition associated. Using the changes in occupation probabilities with temperature, as detailed elsewhere,[79] it is possible to compare the three transitions observed for the two sequences. All three peaks for the parent sequence are associated with changes within the parallel funnel. Our sample of local minima does not produce a melting peak for the parent sequence. In contrast, for the E20S mutant a melting peak can be identified at around 700 K. While the lowest peak at around 35 K corresponds to the transition from the global minimum to other low lying parallel structures, the peak at about 340 K is associated with a transition from parallel structures to kinked intermediates (appropriately colored disconnectivity graphs are provided in the Supporting Information, Figure S11).

3.6. Conformational Switches. Switching between the parallel and antiparallel conformations seems to depend on the two properties described above, that is, interhelical interactions and low-energy intermediates. Thermodynamically, the states

need to be close in energy for the equilibrium occupation probabilities to be comparable. The energy difference is largely determined by the interaction between the two coils, and for the parent sequence, by one strong interaction, namely the Glu20-Lys15 salt bridge, which facilitates alignment with a range of favorable contacts, in agreement with the work of Yadav et al.[45] The absence of Glu20 leads to a destabilization of the parallel configuration due to the missing salt bridge, and stabilization of the antiparallel configuration is observed because destabilizing arrangements formed by salt bridges are not present. Experimental observation of alternative conformations also depends on the kinetics, and only if the barriers involved are sufficiently low will transitions occur on relevant time scales.

The changes caused by mutations may have observable effects beyond the parallel to antiparallel interconversion. The kinking observed for the E20S mutant coincides with instabilities in the helices themselves. Hence any changes may destabilize the helices in such a way that no stable dimer is observed. To investigate this relationship further, the effect of a second alternative mutation was probed with the aim of lowering the energy barrier for kinked structures. Suitable effects were achieved by introducing proline (E20P), which favors distortions in helices,[80−82] glycine (E20G), and alanine (E20A), which will not be able to form strong interhelical contacts.

3.6.1. The Energy Landscapes of the E20G, E20A, and E20P Mutants. All three energy landscapes still exhibit a clear separation of configurations into antiparallel and parallel funnels, as seen for the E20S mutant and the parent sequence (see Supporting Information, Figure S2−S4). Despite this preservation of the key features, clear differences emerge as well. First, most kinked structures exhibit even higher degrees of deformation, including kinking introduced in both helices. As expected, the energy barriers for kinking are lowest for the proline mutation, but they are also lowered for the E20G and E20A mutants (see Supporting Information, p S3). The lower barrier is caused by the lower propensity for helical structures

for proline and the lack of interhelical salt bridges that prevent kinking. The lack of interhelical interactions is also reflected in the fact that the absolute energy difference between parallel and antiparallel structures is reduced (see Supporting Information, p S3).

3.6.2. Kinking Mechanism of E20S. The mutations considered above have additional effects. The landscapes for the E20G and E20A mutants both show low-lying funnels competing with the helical dimers, which exhibit a high degree of deformation and a loss of helicity. Despite being able to introduce more kinking, and hence lower energy intermediate structures, the loss in structural definition of the parallel and antiparallel states, which corresponds to broader, shallower funnels and the appearance of competing funnels in the landscape, inhibits the switching behavior. The key factor here is that the loss of interhelical interactions required to produce lower energy intermediates must be conserved to some degree for structural stability. This conservation occurs for the E20S mutant, but not for the E20G or E20P mutants, and only to a small degree for the E20A mutant. This effect can also be seen in the intermediate structures (see Figure 10).

The E20S mutant supports one specific kinked structure, with the serine mutation exhibiting sufficiently strong interactions between strands to conserve the helical dimers, but still allows for a kinked intermediate that is stabilized by interactions of the backbone and Glu22 with Lys15. In contrast, the E20A mutant exhibits a stronger degree of kinking, leading to a larger number of feasible structures and to stabilization via different residues, for example, His18, which in turn causes changes in the backbone arrangements for both chains. A necessary condition for the switchability of coiled-coil structures therefore seems to be mutations that weaken the bond strength of the interhelical interactions. This change lowers the energies of intermediate structures and the barriers that separate them, while maintaining the stability of the end points.

4. CONCLUSION

Our results for the GCN4-pLI peptide illustrate how analysis of energy landscapes can be used to explore the effect of point mutations. Within this framework it is possible to explain experimental observations, find a structural basis for emergent properties, and predict new phenomena that correspond to a multifunnel potential energy landscape.

For this coiled-coil system it was shown that the energy landscapes of the parent sequence and E20S mutant dimers exhibit similar topographies, with separation into two main regions, corresponding to parallel and antiparallel arrangements. The main difference between the two systems is the appearance of kinked structures in the intermediate region around the mutated residue, leading to a different interconversion pathway. Furthermore, the energy difference between the antiparallel and parallel configurations is lowered. This result agrees with the experimental observation that the parent sequence preference for the parallel arrangement is based on interhelical interactions, especially the Lys15−Glu20 salt bridge. A salt bridge lowers the energy of the parallel configurations, and destabilizes the antiparallel configurations. This salt bridge is not present in the E20S mutant, leading to the possibility of antiparallel structures for the dimer. We predict that the assembly of tetramers from these dimers can easily occur for the mutant protein. For the parent sequence the

antiparallel tetramers are unstable, and only the parallel configurations are expected.

Probing the importance of the intermediate, kinked structures and the interhelical interactions further, our study of the E20G, E20A, and E20P mutants reveals that the observations for the parent sequence and its E20S mutant can be generalized to other mutations. The energy landscapes are relatively robust toward mutations, all exhibiting two well-separated regions associated with antiparallel and parallel structures. However, the mutations still lead to significant changes in the distribution of energy barriers, the number of structures with particular structural properties observed, and their relative energies. The removal of the salt bridge can lead to a switchable system (E20S), but also to unstable helices (E20A, E20G, and E20P), indicating a delicate balance in the interactions.

Our results also imply that a change in the multifunnel character can shift the energy and normal-mode frequency distributions to reflect a single underlying funnel. These changes will have significant effects on the thermodynamic and kinetic properties. Multifunnel protein energy landscapes hinder thermodynamic equilibrium, so that nonequilibrium configurations determined by kinetics may be observed, with interesting consequences for stability with respect to point mutations.

A better understanding of the factors involved in the stability and pathways between parallel and antiparallel structures may be useful in extending the understanding of protein−protein interactions, in providing insight into the role of mutations in self-assembly and aggregation, and in the future design of proteins based on the coiled-coil motif.[83−86]

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b00024.

> Overview of potential energy differences between the funnels for all mutants, pathways for the E20S mutant and the parent sequence from parallel to antiparallel configurations, transition state energies for the highest transition state on the fastest path, RMSD for explicit solvent MD simulations and protocol for MD simulations, energy and normal-mode frequency distributions of minima and transition states, disconnectivity graphs for E20P, E20A, and E20G mutants (PDF)

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: kr366@cam.ac.uk.
*E-mail: dw34@cam.ac.uk.

■ REFERENCES

(1) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein folding funnels: a kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 8721−8725.

(2) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195.

(3) Wolynes, P.; Onuchic, J.; Thirumalai, D. Navigating the folding routes. *Science* **1995**, *267*, 1619−1620.

(4) Onuchic, J. N.; Socci, N. D.; Luthey-Schulten, Z.; Wolynes, P. G. Protein folding funnels: the nature of the transition state ensemble. *Folding Des.* **1996**, *1*, 441−450.

(5) Karplus, M. The Levinthal paradox: yesterday and today. *Folding Des.* **1997**, *2*, S69−S75.

(6) Dobson, C. M.; Karplus, M. The fundamentals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.* **1999**, *9*, 92−101.

(7) Wales, D. J. *Energy Landscapes*; Cambridge University Press, 2003.

(8) Wolynes, P. G. Energy landscapes and solved protein-folding problems. *Philos. Trans. R. Soc., A* **2005**, *363*, 453−467.

(9) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289−316.

(10) Truong, H. H.; Kim, B. L.; Schafer, N. P.; Wolynes, P. G. Funneling and frustration in the energy landscapes of some designed and simplified proteins. *J. Chem. Phys.* **2013**, *139*, 121908.

(11) Englander, S. W.; Mayne, L. The nature of protein folding pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 15873−15880.

(12) Dinner, A. R.; Šali, A.; Smith, L. J.; Dobson, C. M.; Karplus, M. Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.* **2000**, *25*, 331−339.

(13) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* **2002**, *420*, 102−106.

(14) Schuler, B.; Lipman, E. A.; Eaton, W. A. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **2002**, *419*, 743−747.

(15) Wales, D. J. Discrete path sampling. *Mol. Phys.* **2002**, *100*, 3285−3305.

(16) Wales, D. J. Some further applications of discrete path sampling to cluster isomerization. *Mol. Phys.* **2004**, *102*, 891−908.

(17) Noé, F.; Fischer, S. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr. Opin. Struct. Biol.* **2008**, *18*, 154−162.

(18) Wales, D. J. Energy landscapes: some new horizons. *Curr. Opin. Struct. Biol.* **2010**, *20*, 3−10.

(19) Chebaro, Y.; Ballard, A. J.; Chakraborty, D.; Wales, D. J. Intrinsically Disordered Energy Landscapes. *Sci. Rep.* **2015**, *5*, 10386.

(20) Prentiss, M. C.; Wales, D. J.; Wolynes, P. G. The Energy Landscape, Folding Pathways and the Kinetics of a Knotted Protein. *PLoS Comput. Biol.* **2010**, *6*, e1000835.

(21) Shang, C.; Whittleston, C. S.; Sutherland-Cash, K. H.; Wales, D. J. Analysis of the Contrasting Pathogenicities Induced by the D222G Mutation in 1918 and 2009 Pandemic Influenza A Viruses. *J. Chem. Theory Comput.* **2015**, *11*, 2307−2314.

(22) Strodel, B.; Whittleston, C. S.; Wales, D. J. Thermodynamics and Kinetics of Aggregation for the GNNQQNY Peptide. *J. Am. Chem. Soc.* **2007**, *129*, 16005−16014.

(23) Strodel, B.; Wales, D. J. Implicit Solvent Models and the Energy Landscape for Aggregation of the Amyloidogenic KFFE Peptide. *J. Chem. Theory Comput.* **2008**, *4*, 657−672.

(24) Nelson, E. D.; Onuchic, J. N. Proposed mechanism for stability of proteins to evolutionary mutations. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 10682−10686.

(25) Reddy, B. V.; Datta, S.; Tiwari, S. Use of propensities of amino acids to the local structural environments to understand effect of substitution mutations on protein stability. *Protein Eng., Des. Sel.* **1998**, *11*, 1137−45.

(26) Taverna, D. M.; Goldstein, R. A. Why are proteins so robust to site mutations? *J. Mol. Biol.* **2002**, *315*, 479−484.

(27) Taverna, D. M.; Goldstein, R. A. Why are proteins marginally stable? *Proteins: Struct., Funct., Genet.* **2002**, *46*, 105−109.

(28) Sapra, K. T.; Balasubramanian, G. P.; Labudde, D.; Bowie, J. U.; Muller, D. J. Point Mutations in Membrane Proteins Reshape Energy Landscape and Populate Different Unfolding Pathways. *J. Mol. Biol.* **2008**, *376*, 1076−1090.

(29) Berger, E. A.; Murphy, P. M.; Farber, J. M. Chemokine Receptors as HIV-1 Coreceptors: Roles in Viral Entry, Tropism, and Disease. *Annu. Rev. Immunol.* **1999**, *17*, 657−700.

(30) Selkoe, D. J. Alzheimer's Disease: Genes, Proteins, and Therapy. *Physiol. Rev.* **2001**, *81*, 469−480.

(31) Vogelstein, B.; Kinzler, K. W. Cancer genes and the pathways they control. *Nat. Med.* **2004**, *10*, 789−799.

(32) Clevers, H. Wnt/β-Catenin Signaling in Development and Disease. *Cell* **2006**, *127*, 469−480.

(33) Crick, F. H. C. The packing of α-helices: simple coiled-coils. *Acta Crystallogr.* **1953**, *6*, 689−697.

(34) Liu, J.; Rost, B. Comparing function and structure between entire proteomes. *Protein Sci.* **2001**, *10*, 1970−1979.

(35) Mason, J. M.; Arndt, K. M. Coiled coil domains: stability, specificity, and biological implications. *ChemBioChem* **2004**, *5*, 170−176.

(36) Lupas, A. N.; Gruber, M. The structure of alpha-helical coiled coils. *Adv. Protein Chem.* **2005**, *70*, 37−78.

(37) Grigoryan, G.; Keating, A. Structural specificity in coiled-coil interactions. *Curr. Opin. Struct. Biol.* **2008**, *18*, 477−483.

(38) Deng, Y.; Liu, J.; Zheng, Q.; Li, Q.; Kallenbach, N. R.; Lu, M. A Heterospecific Leucine Zipper Tetramer. *Chem. Biol.* **2008**, *15*, 908−919.

(39) Su, L.; Cukier, R. I. Hamiltonian Replica Exchange Method Studies of a Leucine Zipper Dimer. *J. Phys. Chem. B* **2009**, *113*, 9595−9605.

(40) Dunin-Horkawicz, S.; Lupas, A. N. Measuring the conformational space of square four-helical bundles with the program samCC. *J. Struct. Biol.* **2010**, *170*, 226−235.

(41) Ramos, J.; Lazaridis, T. Computational analysis of residue contributions to coiled-coil topology. *Protein Sci.* **2011**, *20*, 1845−1855.

(42) Rämisch, S.; Lizatović, R.; André, I. Exploring alternate states and oligomerization preferences of coiled-coils by de novo structure modeling. *Proteins: Struct., Funct., Genet.* **2015**, *83*, 235−247.

(43) Gambin, Y.; Schug, A.; Lemke, E. A.; Lavinder, J. J.; Ferreon, A. C. M.; Magliery, T. J.; Onuchic, J. N.; Deniz, A. A. Direct single-molecule observation of a protein living in two opposed native structures. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 10153−10158.

(44) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **1993**, *262*, 1401−1407.

(45) Yadav, M. K.; Leman, L. J.; Price, D. J.; Brooks, C. L.; Stout, C. D.; Ghadiri, M. R. Coiled coils at the edge of configurational heterogeneity. Structural analyses of parallel and antiparallel homotetrameric coiled coils reveal configurational sensitivity to a single solvent-exposed amino acid substitution. *Biochemistry* **2006**, *45*, 4463−4473.

(46) Case, D.; Darden, T.; Cheatham, T., III; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A.; Kolossváry, I.; Wong, K.; Paesani, F.; Vanicek, J.; Wolf, R.; Liu, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D.; Mathews, D.; Seetin, M.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. *AMBER12*; University of California: San Francisco, CA, 2012.

(47) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **1986**, *7*, 230−252.

(48) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P.

AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dy- namics and free energy calculation to simulate the structural and energetic proper- ties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1−41.

(49) Hornak, V.; Abel, R.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.

(50) Małolepsza, E.; Strodel, B.; Khalili, M.; Trygubenko, S.; Fejer, S. N.; Wales, D. J. Symmetrization of the AMBER and CHARMM force fields. *J. Comput. Chem.* **2010**, *31*, 1402−1409.

(51) Małolepsza, E.; Strodel, B.; Khalili, M.; Trygubenko, S.; Fejer, S. N.; Wales, D. J. Erratum to: Symmetrization of the AMBER and CHARMM force fields. *J. Comput. Chem.* **2012**, *33*, 2209.

(52) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.

(53) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 383−394.

(54) Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theor. Chem. Acc.* **1999**, *101*, 426−434.

(55) Trygubenko, S. A.; Wales, D. J. A doubly nudged elastic band method for finding transition states. *J. Chem. Phys.* **2004**, *120*, 2082−2094.

(56) Henkelman, G.; Jónsson, H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.* **1999**, *111*, 7010−7022.

(57) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901−9904.

(58) Munro, L. J.; Wales, D. J. Defect migration in crystalline silicon. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 3969−3980.

(59) Carr, J. M.; Wales, D. J. Global optimization and folding pathways of selected alpha-helical proteins. *J. Chem. Phys.* **2005**, *123*, 234901.

(60) Wales, D. J. *PATHSAMPLE: A driver for OPTIM to create stationary point databases using discrete path sampling and perform kinetic analysis*; Cambridge, UK, 2016.

(61) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542−1555.

(62) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(63) Mantell, R. G.; Pitt, C. E.; Wales, D. J. GPU-Accelerated Exploration of Biomolecular Energy Landscapes. *J. Chem. Theory Comput.* **2016**, *12*, 6182−6191.

(64) Wales, D. J. *OPTIM−a program for optimizing geometries and calculating reaction pathways*; Cambridge, UK, 2016.

(65) Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Mater.* **1959**, *1*, 269−271.

(66) Evans, D. A.; Wales, D. J. Folding of the GB1 hairpin peptide from discrete path sampling. *J. Chem. Phys.* **2004**, *121*, 1080−1090.

(67) Carr, J. M.; Trygubenko, S. A.; Wales, D. J. Finding pathways between distant local minima. *J. Chem. Phys.* **2005**, *122*, 234903.

(68) Wales, D. J.; Carr, J. M. Quasi-Continuous Interpolation Scheme for Pathways between Distant Configurations. *J. Chem. Theory Comput.* **2012**, *8*, 5020−5034.

(69) *The PyMOL Molecular Graphics System*, version 1.6.x; Schrödinger, LLC, 2013.

(70) Becker, O. M.; Karplus, M. The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *J. Chem. Phys.* **1997**, *106*, 1495−1517.

(71) Wales, D. J.; Miller, M. A.; Walsh, T. R. Archetypal energy landscapes. *Nature* **1998**, *394*, 758−760.

(72) Wales, D. J. Symmetry, near-symmetry and energetics. *Chem. Phys. Lett.* **1998**, *285*, 330−336.

(73) Wolynes, P. G. Symmetry and the energy landscapes of biomolecules. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 14249−14255.

(74) Williams, M. A.; Thornton, J. M.; Goodfellow, J. M. Modelling protein unfolding: hen egg-white lysozyme. *Protein Eng., Des. Sel.* **1997**, *10*, 895−903.

(75) Vendruscolo, M.; Paci, E.; Karplus, M.; Dobson, C. M. Structures and relative free energies of partially folded states of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14817−14821.

(76) Sciortino, F.; Kob, W.; Tartaglia, P. Thermodynamics of supercooled liquids in the inherent-structure formalism: a case study. *J. Phys.: Condens. Matter* **2000**, *12*, 6525−6534.

(77) Bogdan, T. V.; Wales, D. J.; Calvo, F. Equilibrium thermodynamics from basin-sampling. *J. Chem. Phys.* **2006**, *124*, 044102.

(78) Wales, D. J. Surveying a complex potential energy landscape: Overcoming broken ergodicity using basin-sampling. *Chem. Phys. Lett.* **2013**, *584*, 1−9.

(79) Wales, D. J. Decoding Heat Capacity Features from the Energy Landscape. 2017, submitted.

(80) von Heijne, G. Proline kinks in transmembrane α-helices. *J. Mol. Biol.* **1991**, *218*, 499−503.

(81) Nilsson, I.; von Heijne, G. Breaking the camel's back: proline-induced turns in a model transmembrane helix. *J. Mol. Biol.* **1998**, *284*, 1185−1189.

(82) Cordes, F. S.; Bright, J. N.; Sansom, M. S. Proline-induced Distortions of Transmembrane Helices. *J. Mol. Biol.* **2002**, *323*, 951−960.

(83) Rose, A.; Meier, I. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. *Cell. Mol. Life Sci.* **2004**, *61*, 1996−2009, DOI: 10.1007/s00018-004-4039-6.

(84) Bromley, E. H. C.; Channon, K.; Moutevelis, E.; Woolfson, D. N. Peptide and Protein Building Blocks for Synthetic Biology: From Programming Biomolecules to Self-Organized Biomolecular Systems. *ACS Chem. Biol.* **2008**, *3*, 38−50.

(85) Kreuzer, S.; Elber, R. Coiled-Coil Response to Mechanical Force: Global Stability and Local Cracking. *Biophys. J.* **2013**, *105*, 951−961.

(86) Huang, P.-S.; Boyken, S. E.; Baker, D. The coming of age of de novo protein design. *Nature* **2016**, *537*, 320−327.