Dynamic charge interactions create surprising rigidity in the ER/K α-helical protein motif

Sivaraj Sivaramakrishnan*, Benjamin J. Spink*, Adelene Y. L. Sim†, Sebastian Doniach†‡, and James A. Spudich*§

Departments of *Biochemistry, †Applied Physics, and ‡Physics, Stanford University, Stanford, CA 94305

Contributed by James A. Spudich, June 30, 2008 (sent for review June 9, 2008)

Protein α-helices are ubiquitous secondary structural elements, seldom considered to be stable without tertiary contacts. However, amino acid sequences in proteins that are based on alternating repeats of four glutamic acid (E) residues and four positively charged residues, a combination of arginine (R) and lysine (K), have been shown to form stable α-helices in a few proteins, in the absence of tertiary interactions. Here, we find that this ER/K motif is more prevalent than previously reported, being represented in proteins of diverse function from archaea to humans. By using molecular dynamics (MD) simulations, we characterize a dynamic pattern of side-chain interactions that extends along the backbone of ER/K α-helices. A simplified model predicts that side-chain interactions alone contribute substantial bending rigidity (0.5 pN/nm) to ER/K α-helices. Results of small-angle x-ray scattering (SAXS) and single-molecule optical-trap analyses are consistent with the high bending rigidity predicted by our model. Thus, the ER/K α-helix is an isolated secondary structural element that can efficiently span long distances in proteins, making it a promising tool in designing synthetic proteins. We propose that the significant rigidity of the ER/K α-helix can help regulate protein function, as a force transducer between protein subdomains.

MD simulations | protein structure | single-molecule analysis | small-angle x-ray scattering

The α-helix has been found to be a ubiquitous secondary structural element in proteins since its presence was inferred by Pauling *et al.* (1). However, in naturally occurring proteins, α-helices are rarely thought of as independent protein structural elements, stable in solution in the absence of tertiary interactions (2). In folded proteins, α-helices are considered to be stabilized by exclusion of water from the hydrophobic protein core and/or by specific electrostatic interactions with neighboring secondary structures (3). Destabilization of α-helices in the absence of tertiary interactions is, thus, thought to be a consequence of competition for hydrogen bonding between the amide hydrogen in the α-helical backbone with either water oxygen in aqueous solutions or carbonyl oxygen in the backbone (4, 5).

However, there are a few known examples of α-helices that are stable in solution in the absence of tertiary interactions, including a fragment of the central region of smooth muscle caldesmon ($\approx$160 residues) (6), a segment of the tail domain of myosin X ($\approx$36 residues) (7), and the medial tail (MT) domain of myosin VI ($\approx$70 residues) (8). A common feature of these α-helices is an alternating pattern of approximately four negative (E) and approximately four positive (R/K) amino acids (6–8). Functionally, these α-helices have been suggested to bridge two different protein domains, as in caldesmon and myosin VI, and potentially extend the lever arm of the mechanoenzymes myosin VI and myosin X (6–8). How this pattern of charged residues confers stability on such α-helices, has not been clear.

Studies of the stability of isolated α-helices have almost exclusively used short ($<$20 residues) synthetic alanine-based α-helices (9–11). Polyalanine peptides form stable α-helices in solution, with α-helical content augmented by the presence of organic solvents (12). Alanine α-helices are relatively flexible structures that appear as fluctuating semibroken rods (11). Short alanine-based peptides containing an occasional single type of charged amino acid (E, R, or K) have enhanced α-helical content compared with α-helices consisting of alanine only (13, 14). The enhanced α-helical content of alanine peptides with occasional E, R, or K has been attributed to their ability to sequester water from the carbonyl oxygen and amide hydrogen groups in the α-helical backbone (4, 9). Experiments examining sparse E–K interactions in alanine α-helices suggest that E–K charge interactions do not augment α-helical content relative to alanine α-helices with only one of E, R, or K (10, 13, 15).

In contrast, the α-helical content of peptides containing only E and K (no alanine) depends on the specific distribution of E and K within the sequence (16). Specifically, the α-helical content of $(E_4K_4)_n$ is substantially higher than that of $(E_2K_2)_n$ for the same number of amino acid residues (16). This trend is reflected in the sequences of the few known naturally occurring stable α-helices of caldesmon, myosin VI, and myosin X, which follow the $(E_4K_4)_n$ pattern more closely than the $(E_2K_2)_n$ pattern (6–8). In these three proteins, it is seen that R can be substituted for K, giving alternating repeats of approximately four E residues and approximately four positively charged residues, a combination of R and K (which we denote R/K). Also, the E and R/K residues are sparsely interspersed with hydrophobic residues (A, L, I, V, and M), along with glutamine (Q) and aspartic acid (D). We refer to this general $(E_4(R/K)_4)_n$ pattern as the ER/K motif.

The dynamics of E and R/K side-chain interactions and their potential contributions to structural properties of ER/K α-helices are unknown. In this study, we use replica-exchange MD (REMD) simulations to examine the equilibrium conformations of ER/K α-helices. We use information derived from the MD simulations to construct a simplified theoretical model for the bending rigidity of ER/K α-helices. The stability and rigidity of ER/K α-helices are independently examined by a combination of SAXS and single-molecule optical-trap analyses of a sample ER/K α-helix, namely the MT of myosin VI.

Results and Discussion

ER/K α-Helix Is Recurrent in Nature. The number of both known and hypothetical or predicted protein sequences in the database with the ER/K motif is revealed by a BLAST search in the PubMed database of nonredundant protein sequences including all species. To ensure a nonbiased evaluation of search hits, $(E_4K_4)_2$ and $(E_4R_4)_2$ were used as search templates, and the list of sequences that scored better than any species of myosin X or myosin VI was examined to assess the relative number of proteins that contain the ER/K motif. Because tail fragments in myosin VI and myosin X are known to be isolated α-helices in solution, we propose that our search criteria provide a conser-

Author contributions: S.S. and J.A.S. designed research; S.S., B.J.S., and A.Y.L.S. performed research; S.D. contributed new reagents/analytic tools; S.S., A.Y.L.S., and J.A.S. analyzed data; and S.S., B.J.S., A.Y.L.S., S.D., and J.A.S. wrote the paper.

The authors declare no conflict of interest.

§To whom correspondence should be addressed. E-mail: jspudich@stanford.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806256105/DCSupplemental.

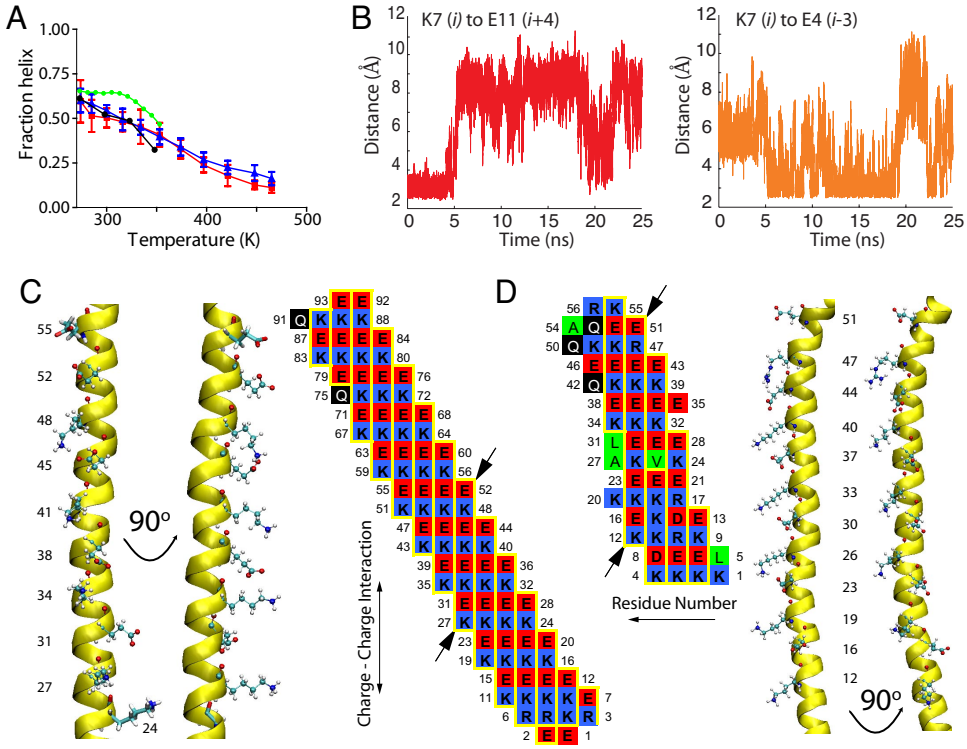

Fig. 1. Charge interactions along the ER/K α-helix backbone. (*A*) Fraction helix as a function of temperature from experiment (CD melt) and computation (REMD). Blue, REMD data starting with α-helical conformation; red, REMD data starting with random initial conformation; black, CD-melt data from Lyu *et al.* (16); green, CD melt data for the MT ER/K α-helix in myosin VI. CD data were converted to fraction helix as per Rohl and Baldwin (28). (*B*) With reference to an $(E_4K_4)_2$ helix (E1 to K16), sample variation of distance (Å) between N_ζ atom of K7 (*i*) and center of mass of O_ε atoms of E4 ($i-3$) (*Right*) and E11 ($i+4$) (*Left*) with time (ns) during MD simulations (*Left*). (*C*) (*Left*) Snapshot of two orthogonal views of a section of the ER/K α-helix (residues 24–55) in mannosyltransferase from the MD simulation, showing side-chain interactions and residue numbers. For reference, the backbone carbonyl oxygens corresponding to the illustrated side chains are also shown. (*Right*) The charge interaction map (see *SI Methods*) shown to the right illustrates the pattern of E↔R/K interactions along the vertical axis (see yellow boxed groups of residues in columns). Arrows show boundaries of the segment of the α-helix displayed in *Left*. (*D*) Snapshot of two orthogonal views of a section of the ER/K helix (residues 12–51) in upf2 regulator of nonsense transcripts homolog (*Right*), with its charge interaction map (*Left*).

vative estimate of similar α-helices in other protein structures. We find the ER/K motif is present in at least 123 known distinct proteins in 137 organisms ranging from archaea to humans, with at least 20 distinct proteins in mammals [e.g., supporting information (SI) Table S1]. As mentioned above, the ER/K motifs in the three proteins that are known to form stable α-helices in solution are sparsely interspersed with residues A, L, I, V, M, Q, and D. The relative content of residues E, R, and K in these helices is 70% in caldesmon, 78% in myosin X, and 79% in myosin VI. When we analyzed the results of the BLAST search for continuous ER/K regions matching the search criteria, with the boundaries of this region determined by the presence of amino acids not listed above, we found that the mean content of E, R, and K in helices obtained from the search is 80%. By using this approach, the ER/K motif is >30 residues in at least 40 distinct proteins with the longest continuous ER/K region being 346 residues for a Kelch motif family protein in *Trichonomis vaginalis G3*.

MD Simulations Reveal Dynamic Charge Interactions in ER/K α-Helices. To understand the effects of E and R/K side-chain interactions on the stability and function of ER/K α-helices, we performed REMD simulations on the sample sequence $(E_4K_4)_2$ in explicit solvent. We selected $(E_4K_4)_2$ as a suitable candidate because it allows the side chains to sample multiple interactions (*i* to $i \pm n$ where $n = 1$–7) and seek out the most energetically favorable ones. Also, $(E_4K_4)_2$ is known have high α-helical content in solution with known temperature, pH, and salt-concentration dependence, as measured by circular dichroism (CD) (16). These experimental measurements provide a quantitative benchmark of our REMD simulations. Last, the $(E_4K_4)_2$ sequence is well represented in the protein mannosyltransferase (*Saccharomyces cerevisiae*), which contains four continuous perfect repeats of $(E_4K_4)_2$. In the absence of known crystal structures of ER/K α-helices, we performed REMD simulations starting with the $(E_4K_4)_2$ peptides in both random (no secondary structure) and fully α-helical conformations (see *Methods*). Convergence of simulation was assessed by insignificant differences in fraction helix (see *SI Methods*) at all simulated temperatures (Fig. 1*A* and Fig. S1). The convergence of two different initial conditions to a unique solution is essential to obtain equilibrium conformations for both the peptide backbone and side-chain interactions, especially because the objective of these MD simulations is to study side-chain interactions.

The thermal melt curves obtained from REMD (Fig. 1*A*) are very similar to experimental data obtained for $(E_4K_4)_2$ by using CD (16). CD data for the MT segment of myosin VI are shown in Fig. S2 with the fraction helix displayed in Fig. 1*A*. The fraction helix for the myosin VI MT with ≈70 residues is higher than that for $(E_4K_4)_2$ with 16 residues, as expected due to the increased length of the peptide chain (11).

We used REMD simulations to equilibrium sample a large range of protein conformations by using finite computing power (17). However, due to periodic exchange between different replicas, REMD does not provide trajectories of side-chain interactions in real time. To obtain this information, we randomly selected $(E_4K_4)_2$ replicas that were completely folded from random initial configuration and equilibrated them in

BIOPHYSICS

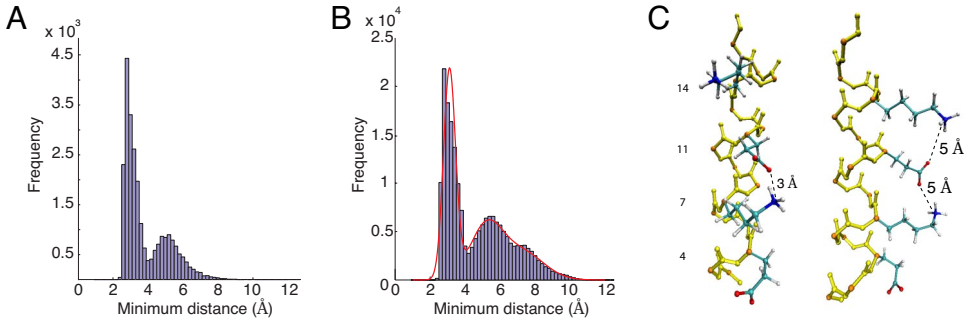

Fig. 2. Side-chain interaction distances in the $(E_4K_4)_2$ α-helix. (*A*) Histogram of minimum distance, at every time step of the MD simulation, between K7-N_ζ and the centers of mass of O_ε in neighboring E residues. For the entire simulation time, K7-N_ζ is closest to one of E3, E4, E10, and E11. (*B*) Histogram of minimum distances, at every time step of the MD simulation, between N_ζ of each K residue and centers of mass of O_ε in all neighboring E residues and vice versa. The red line shows a maximum likelihood fit to the sum of three Gaussians, which yields the relative time spent under the 3, 5, and 8-Å peaks along with their standard deviations. (*C*) Snapshot of the $(E_4K_4)_2$ α-helix showing sample side-chain conformations during 3-Å (*Left*) and 5-Å (*Right*) separations between $N_\zeta$and O_ε.

explicit solvent for 25 ns at 274 K. These simulations were continued for an additional 25 ns with trajectories of all protein atoms monitored over this time period. The N_ζ atom of residue K7 (i) oscillates between the center of mass of the two O_ε atoms of E4 ($i - 3$) and E11 ($i + 4$) (Fig. 1*B*). The center of mass of the two O_ε atoms of E10 (i), alternately, oscillates between the N_ζ atom of residue K6 ($i - 4$) and K13 ($i + 3$) (data not shown). This preference for ($i - 3$)/($i + 4$) interactions at K(i) residues and ($i - 4$)/($i + 3$) interactions at E (i) residues is seen for all of the center eight residues (K5-E12) of the $(E_4K_4)_2$ helix. The preference for these interactions is reflected in the histograms of the frequency of all distances for $i - 3/4$ and $i + 3/4$ interactions, at each of the eight residues K5–E12. These histograms reveal clear 3-Å peaks for the specific interactions listed above (Fig. S3). Residues in the first turn of the helix (E1–E4) do not have $i - 3$ and $i - 4$ interactions, whereas residues in the last turn of the helix (K13–K16) do not have $i + 3$ and $i + 4$ interactions.

Closer examination of the instantaneous structures of the folded $(E_4K_4)_2$ α-helices show that the particular pattern of interactions for E and K residues arises from a combination of the following: permissible dihedral angles between the peptide backbone and the C_α–C_β side chain of the E and K residues; and the right-handed structure of the α-helix causing residues to rise left to right from the N to the C terminus of the peptide backbone of the α-helix (Fig. 1 *C* and *D*). For example, K27, E31, K34, E38, K41, E45, K48, and E52 form a network of interactions that spans nearly 4 nm along one face of the α-helix (Fig. 1*C*). This alternating pattern of + 4 then + 3 closely approximates 3.6 residues per turn of an α-helix. Therefore, the ER/K motif allows for continuous series of interactions along the same face of the α-helical backbone. We use this alternating pattern of + 4 then + 3 to draw a map of preferred interactions along the α-helical backbone, both for the extended $(E_4K_4)_n$ helix and for ER/K α-helices in general (Fig. 1*C* and Fig. S4). A striking feature in such maps is several continuous series of interactions in all ER/K α-helices (Fig. 1 *C* and *D*; boxes in yellow). In the $(E_4K_4)_n$ α-helix, eight residues form a chain of oscillating interactions, whereas in upf-2 regulator of nonsense transcripts (Table S1) we see a maximum of 12 residues (Fig. 1*D*, arrows).

The oscillation of E and K side chains between their $i - 3/4$ and $i + 3/4$ neighbors implies that at every time step of the MD simulation, the minimum distance between each N_ζ of K and the centers of mass of O_ε of all its E neighbors (and vice versa) is the dominant contributor to the interaction between E and K side chains. Fig. 2*A* shows the histogram of these minimum distances collected >25 ns for residue K7, and Fig. 2*B* shows the combined histogram of these minimum distances for residues K7 through K12. The preference for ($i - 3$)/($i + 4$) interactions at K(i) residues and ($i - 4$)/($i + 3$) interactions at E(i) residues, described above, holds true in this combined histogram, in that for 90% of the simulation time the closest pairing residue for K or E follows this preferred pattern. The histogram can be best characterized by a maximum-likelihood fit to the sum of three Gaussian distributions (*SI Methods*) with centers located at 3, 5, and 8 Å. The relative times spent in the three modes of interaction, in turn, are given by the areas under the individual Gaussian distributions to be 45%, 37%, and 18%, respectively. Detailed examination of interatomic distances reveals that the interaction centered at 3 Å corresponds to hydrogen bonding between the H_ζ of K and O_ε of E (Fig. 2*C* and Fig. S5). For the interaction centered at 5 Å, the E and K side chains are located approximately equidistant from their $i - 3/4$ and $i + 3/4$ neighbors (Fig. 2*C*). For these conformations, we find that the O_ε atom of E is located 3–6 Å from H atoms in the K side chain, suggesting a combination of direct interaction through Coulomb forces between opposite charges and indirect interactions through solvent-separated salt-bridge contacts (Fig. S5). The third peak centered at 8 Å has contributions from the following: the continuous series of alternating E and K interactions located seven residues apart implies that residues K8 and E9 do not have $i - 3$ and $i + 3$ neighbors, respectively (Fig. S3); and due to oscillations of side chains between their $i - 3/4$ and $i + 3/4$ pairs, each side chain spends a small fraction of its time (<10%) without an interacting partner.

Side-Chain Interactions Significantly Influence The Stability and Rigidity of ER/K α-Helices. The stability of isolated α-helices in the presence of polar solvents such as water is limited by disruption of hydrogen bonds in the α-helical backbone (4, 5). Previous comparisons of MD simulations of alanine-based α-helices with SAXS measurements suggest that these helices undergo continuous transient breaks in the α-helical backbone that reduce the end-to-end distance of the α-helix (11). Our MD simulations show that there are multiple E and R/K side-chains interactions for each turn of the ER/K α-helix, which prevent transient breaks in backbone hydrogen bonds from unraveling the α-helix. Additionally, the ER/K α-helix partially excludes water molecules from the backbone (Fig. 3*A*), approximately to the same extent as the alanine-based F_s peptide (simulation data not shown). This partial water exclusion is potentially a combination of space occupied by the long E and R/K side chains and the hydrophobic methylene groups in them. Therefore, outer charges in the E and R/K side chains that interact favorably, along with inner hydrophobic groups that exclude water molecules, offer a unique combination to stabilize the backbone conformation of ER/K α-helices in the aqueous cellular environment.

There are two potential sources of bending rigidity of isolated

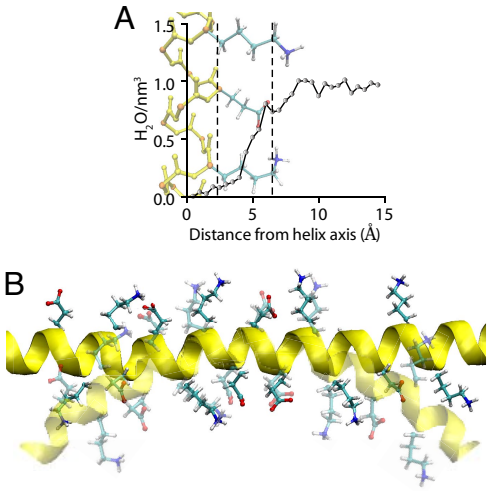

Fig. 3. Lateral bending and stability of ER/K α-helix. (*A*) Water density as a function of distance from the α-helix axis for the $(E_4K_4)_2$ α-helix. The vertical dotted line at 2.2 Å shows the average location of C_α atoms, whereas the vertical dotted line at 6 Å shows the average location of N_ζ/O_ε atoms of the K/E side chains. Note the exclusion of water from the α-helical backbone resulting from the presence of side chains. (*B*) Exaggerated schematic representation showing bending of the ER/K α-helix in mannosyltransferase. The α-helix backbone is shown in ribbon representation (yellow), along with selected E and K side chains. To illustrate the effect of bending on side-chain interaction distances, the bent α-helix is shown as a transparent structure. The α-helix bending approximately the neutral axis results in an increase in E↔K distances on the convex side (top face), with a corresponding decrease in E↔K distances on the concave side (bottom face).

ER/K α-helices, the α-helical backbone and side-chain interactions. Previous computational analyses of alanine α-helices that ignore transient breaks in the α-helical backbone report a bending stiffness of ≈1 pN/nm (scaled for a 10-nm long α-helix) (18, 19). However, in the presence of polar solvent, with partial water exclusion, the contribution of the α-helical backbone to the bending stiffness is unclear. With regard to side-chain contribution to bending stiffness, the backbone carbonyl oxygen, amide hydrogen, and C_α atoms are of 2.2 Å from the helix axis, whereas N_ζ and O_ε of the K and E side chains, respectively, are located on average at 6 Å from the axis of the ER/K α-helix. Thus, bending forces on the ER/K α-helix will first be "sensed" and, therefore, resisted by side-chain interactions before they propagate to the backbone (Fig. 3*B*). In this regard, the E and R/K side-chain interactions resemble a continuous narrow tube of radius 6 Å that reinforces the bending rigidity of ER/K α-helices. To examine the contribution of E and R/K side-chain interactions to the bending rigidity, we use a simplified model for the bending of the ER/K α-helix as a cantilever beam with side-chain interactions representing the elastic elements that resist bending (Fig. S6). Our model is valid only for external mechanical forces applied over time scales that are significantly greater (10 μs) than the dynamic time scale for side-chain interactions shown here (100 ns). Fig. 3*B* shows an exaggerated schematic of the bending of the α-helix with the resulting deformations along the α-helical backbone.

To estimate the stiffness of individual side-chain interactions, we interpret the two Gaussian distributions centered at 3 and 5 Å as representing two favored potential wells. The width of these three Gaussian distributions can be used to estimate the harmonic spring constants of 3,200 and 440 pN/nm, respectively, for these potential wells. Whereas the hydrogen bonding interaction at 3 Å is approximately 10-fold stronger than the 5-Å one, the softer 5-Å interaction has an approximately 3-fold greater reach due to larger standard deviation (0.3 vs. 0.9 Å). For side-chain-interaction distances corresponding to the 5-Å potential well, each E (i) interacts simultaneously with K ($i - 4$) and K ($i + 3$). However, for the 3-Å potential well, each E (i) interacts with only one of K ($i - 4$) and K ($i + 3$). Thus, the 5-Å potential well, in addition to its longer reach, provides a continuous sheet of interactions without breaks along the α-helical backbone. Therefore, we model the side-chain interactions as a continuous narrow tube with stiffness corresponding to the 5-Å potential well. We thus estimate a stiffness of 0.5 pN/nm entirely from side-chain interactions (Fig. S6 and *SI Methods*). Our model suggests that the ER/K α-helix can act as a relatively stiff structural element in mechanoenzymes such as myosin VI and X, based on side-chain interactions alone.

Single-Molecule Analyses and SAXS Are Consistent with High Rigidity of ER/K α-Helices. We studied the contribution of the MT of myosin VI (MT ER/K α-helix) to its effective lever arm stroke under load applied by a dual beam optical trap (see *Methods*) (20). The myosin VI lever arm strokes through ≈180° parallel to the F-actin filament (21), resulting in a lateral bending force transmitted through the MT ER/K α-helix (Fig. 4*A*). A measure of the stroke size of two myosin VI constructs, one that extends to the beginning and another to the end of the MT ER/K α-helix, enabled us to measure the deflection of the MT ER/K α-helix under lateral bending load.

To characterize the stroke of myosin VI, it is essential to understand the structural elements involved in the myosin VI lever arm. We have previously described the tail domain of myosin VI as being composed of three segments, a globular proximal tail (PT), a central region containing the MT ER/K α-helix, and a terminal globular segment containing the distal tail (DT) and cargo binding domain (8). We had previously shown by using SAXS analysis that the PT alone is an ≈3-nm globular structure, whereas the SAXS envelope for the MT-DT (120 residues) is suggestive of a single α-helix (MT) coupled with a folded domain (DT).

Here, we use SAXS analysis to get insight into the contribution of the PT to the myosin VI lever arm, along with a SAXS envelope for the MT ER/K α-helix alone. To understand the contribution of the PT to the myosin VI lever arm, we obtained a SAXS envelope structure for a segment of the myosin VI molecule starting with the second IQ motif (with bound calmodulin) to the end of the PT (Fig. 4*B Upper*). Our structure for this complex comes from a fusion of the IQ-calmodulin structure-based on previously modeled myosin VI structures (22) and Rosetta prediction of the PT (see *Methods*). The IQ and PT were fused at Pro-835 and the backbone dihedral angles at the IQ–PT junction were adjusted manually to fit the structure optimally into the SAXS envelope. The modeled structure was used to predict a 15-nm stroke size of the myosin VI lever arm to the beginning of the MT ER/K α-helix (Fig. 4*A*). This prediction is verified by a 15 ± 3-nm stroke size measured in an optical trap (Fig. 4*C Upper*).

The SAXS envelope of the MT ER/K α-helix reveals an ≈10-nm long envelope that matches the predicted length of a single isolated α-helix containing ≈70 residues comprising the MT (Fig. 4*B Lower*). Interestingly, the SAXS envelope has two distinct bends whose locations along the α-helix correspond to two highly conserved hydrophobic patches along the length of the MT (Fig. 4*B Lower*). These SAXS results would be most consistent with these hydrophobic patches being relatively rigid kinks rather than flexible nodes. To understand the role of hydrophobic patches on stability of the α-helix, we performed MD simulations of different segments of the MT folded into complete α-helices and then simulated at 400 K (Fig. S7). We found that all segments of the MT remain complete α-helices, whereas the peptide $(E_2K_2)_4$ unfolds to a random coil (Fig. S7), suggesting that these hydrophobic patches in the MT do not unravel the ER/K α-helix. However, the simulations do show that regions that contain these hydrophobic residues undergo tran-

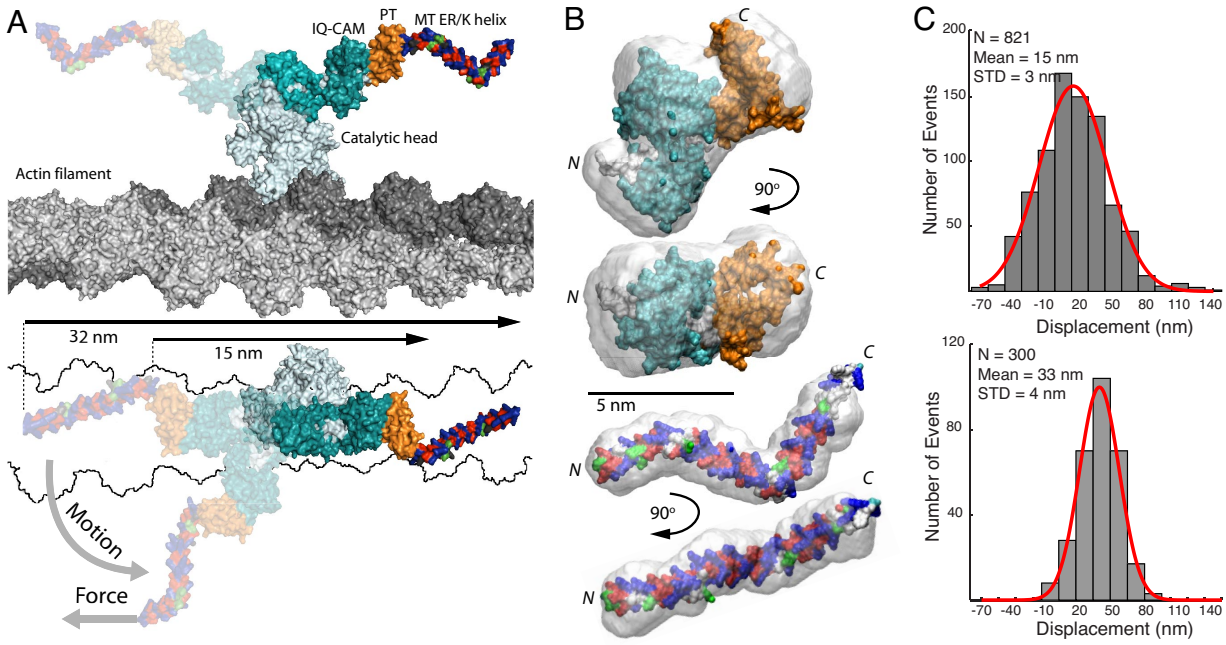

Fig. 4. Experimental verification of stability and bending rigidity of the MT ER/K α-helix in myosin VI. (*A*) Schematic of the lateral bending of the MT ER/K α-helix during the 180° swing of the myosin VI lever arm. (*Upper*) The poststroke state is indicated with the prestroke state shown as a transparent structure (side view along the actin filament). (*Lower*) The poststroke state is indicated with the prestroke state and lever arm in transition shown as transparent structures (top view along the actin filament). Distances shown in the schematic representation are derived from the known myosin VI crystal structure with modeled IQ-CAM (22) and predicted structures of the PT and MT. Although the angle between the PT and MT is unknown, the orientation shown in the schematic representation is consistent with the SAXS structures of PT-MT and PT-MT-DT published previously (8), and aligns the MT parallel to the F-actin filament to allow dimerization of the cargo binding domains, resulting in ≈36-nm processive stepping of myosin VI. (*B*) Two orthogonal views of the predicted structure of a fragment of myosin VI including the IQ motif and PT with bound calmodulin docked into its SAXS envelope. The PT is predicted to contribute only ≈1.5 nm to the lever arm of myosin VI (*Upper*). Orthogonal views of the predicted structure of the MT ER/K α-helix in myosin VI docked into its SAXS envelope. The end–end distance of the SAXS envelope (10 nm) is consistent with the ER/K motif forming an isolated, stable α-helix in solution. N and C termini are marked (*Lower*). (*C*) Histogram of bead displacement in response to binding of myosin VI to an actin filament, followed by the myosin VI power stroke. Mean displacement is a measure of the stroke size of the myosin VI lever arm without (*Upper*) and with (*Lower*) the MT ER/K α-helix.

sient unfolding and refolding more frequently than for completely charged ER/K repeats. The functional significance of this observation remains to be elucidated.

The maximal load applied in the optical trap is the stiffness of the optical trap multiplied by the stroke size (displacement of the bead). Our theoretical model predicts that side-chain interactions contribute significant stiffness to ER/K α-helices. We performed optical trapping measurements at a trap stiffness of ≈0.005 pN/nm, which should result in small deflections (<1 nm) of ER/K α-helices with a stiffness of ≈0.5 pN/nm estimated from our model (Fig. S6). We found that the stroke size of a myosin VI construct that extends to the end of the MT ER/K α-helix is 33 ± 4 nm, which is the same as the predicted stroke size of 33 nm measured from the modeled structure of the same myosin VI construct. It should be emphasized that the angle between the PT and MT ER/K α-helix is unknown. However, the orientation shown in the schematic illustration is consistent with the SAXS structures of PT-MT and PT-MT-DT published previously (8) and aligns the MT ER/K α-helix parallel to the F-actin filament to allow dimerization of the cargo binding domains and the ≈36-nm processive stepping of myosin VI (23).

In summary, we show here that the ER/K α-helix forms a relatively rigid and stable isolated secondary structural element in proteins. The ER/K α-helix is much more prevalent in protein sequences than previously reported. We find a unique oscillating interaction of charged side chains distributed along the α-helical backbone. This dynamic interaction of side chains can tolerate significant change in interaction distances that result from lateral bending of the α-helix, compared with locked interactions resulting from salt-bridge formation. We propose a model of the ER/K α-helix as a cantilever beam, which despite its many simplifications is sufficient to account for the high stiffness of the ER/K α-helix observed in both SAXS and single-molecule optical-trapping experiments. The ER/K α-helix in myosin VI, myosin X, and caldesmon acts as a bridge between protein subdomains. Specifically, in myosin VI a load applied at the end of the cargo binding domain must be transmitted through the ER/K α-helix to the head, where it is known to alter the kinetics of nucleotide binding and release (24). The rigidity of the MT ER/K α-helix is likely to impact the force transduced to the myosin VI catalytic head and, therefore, mechanosignaling of myosin VI. Last but not least, the ability of the ER/K α-helix to efficiently span large distances as an isolated protein structural element makes it a promising tool for use in designing synthetic proteins.

Methods

MD Simulations. REMD of the E_4K_4 peptide (CH_3CO-EEEEKKKKEEEEKKKK-$NHCH_3$) were performed with GROMACS (25, 26) by using the AMBER2003 force field (27). Two simulations, one starting with the peptide in complete α-helix form and the other starting with the peptide as a random coil, were performed for 80 and 42 ns, respectively, for each replica. Trajectories of side-chain interactions were determined from 50 ns simulations of individual replicas at 274 K. Details of simulation parameters and data analysis are described in *SI Methods*.

Biophysical Characterization of Segments of Myosin VI. IQ-PT (814–918 *Homo sapiens*) bound to human calmodulin in an equimolar ratio along with myosin VI MT (916–981 *H. sapiens*) were purified by using standard protein purification techniques (*SI Methods*). SAXS measurements, data analysis, and reconstructions were carried out as described in Spink *et al.* (8) (see *SI Methods* and

Fig. S8 for specific details). CD, multiangle laser light scattering (MALLS), and dynamic light scattering (DLS) measurements were carried out as described in Spink *et al.* (8) (see *SI Methods* for details).

Single-Molecule Optical-Trap Assays. Optical trapping was performed by using myosin VI molecules extending to the beginning (1–918 *Sus scrofa*) and end (1–981 *S. scrofa*) of the MT (see *SI Methods* for details). Optical trapping was performed as described in Bryant *et al.* (21).

ACKNOWLEDGMENTS. We thank M. A. Hartman (Stanford University) for technical help with protein purification and manuscript review; S. Seifert and D. S. Pavlichin for help with SAXS data collection; J. Lipfert for help with SAXS reconstructions and manuscript review; J. Liao, A. Dunn, Z. Bryant, V. S. Pande, M. Levitt, and R. L. Baldwin for critical discussions and manuscript review; and R. Sowdhamini for discussions and use of the National Center for Biological Sciences computing cluster in Bangalore, India, for preliminary simulations. Use of the Advanced Photon Source was supported by the Office of Science and Office of Basic Energy Sciences U.S. Department of Energy under Contract DE-AC02-06CH11357. The Bio-X^2 computing cluster is supported by National Science Foundation Award CNS-0619926. This work was supported by an American Cancer Society postdoctoral fellowship (to S.S.), the Agency for Science, Technology, and Research, Singapore (A.Y.L.S.), and National Institutes of Health Grants T32 GM008294 (to B.J.S.), PO1 GM066275 (to S.D.), and GM33289 (to J.A.S.).

1. Pauling L, Corey RB, Branson HR (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* 37:205–211.
2. Dill K, Ozkan S, Shell M, Weikl T (2008) The Protein Folding Problem. *Annu Rev Biophys* 37:289–316.
3. Daggett V, Fersht A (2003) The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol* 4:497–502.
4. Vila JA, Ripoll DR, Scheraga HA (2000) Physical reasons for the unusual alpha-helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides. *Proc Natl Acad Sci USA* 97:13075–13079.
5. Cammers-Goodwin A, *et al.* (1996) Mechanism of Stabilization of Helical Conformations of Polypeptides by Water Containing Trifluoroethanol. *J Am Chem Soc* 118:3082–3090.
6. Wang CL, *et al.* (1991) A long helix from the central region of smooth muscle caldesmon. *J Biol Chem* 266:13958–13963.
7. Knight PJ, *et al.* (2005) The predicted coiled-coil domain of myosin 10 forms a novel elongated domain that lengthens the head. *J Biol Chem* 280:34702–34708.
8. Spink B, Sivaramakrishnan S, Lipfert J, Doniach S, Spudich J (2008) Long single alpha helical domains bridge the gap between structure and function of myosin VI. *Nat Struct Biol* 15:591–597.
9. Garcia AE, Sanbonmatsu KY (2002) Alpha-helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Natl Acad Sci USA* 99:2782–2787.
10. Ghosh T, Garde S, Garcia AE (2003) Role of backbone hydration and salt-bridge formation in stability of alpha-helix in solution. *Biophys J* 85:3187–3193.
11. Zagrovic B, Jayachandran G, Millett IS, Doniach S, Pande VS (2005) How large is an alpha-helix? Studies of the radii of gyration of helical peptides by small-angle X-ray scattering and molecular dynamics. *J Mol Biol* 353:232–241.
12. Scholtz JM, Baldwin RL (1992) The mechanism of alpha-helix formation by peptides. *Annu Rev Biophys Biomol Struct* 21:95–118.
13. Marqusee S, Baldwin RL (1987) Helix stabilization by Glu-.Lys+ salt bridges in short peptides of de novo design. *Proc Natl Acad Sci USA* 84:8898–8902.
14. Marqusee S, Robbins VH, Baldwin RL (1989) Unusually stable helix formation in short alanine-based peptides. *Proc Natl Acad Sci USA* 86:5286–5290.
15. Huyghues-Despointes BM, Scholtz JM, Baldwin RL (1993) Helical peptides with three pairs of Asp-Arg and Glu-Arg residues in different orientations and spacings. *Protein Sci* 2:80–85.
16. Lyu PC, Gans PJ, Kallenbach NR (1992) Energetic contribution of solvent-exposed ion pairs to alpha-helix structure. *J Mol Biol* 223:343–350.
17. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151.
18. Adamovic I, Mijailovich SM, Karplus M (2008) The elastic properties of the structurally characterized myosin II S2 subdomain: A molecular dynamics and normal mode analysis. *Biophys J* 94:3779–3789.
19. Choe S, Sun SX (2005) The elasticity of alpha-helices. *J Chem Phys* 122:244912.
20. Spudich JA, Rice S, Rock R, Purcell T, Warrick H (2008) *Optical Traps to Study Properties of Molecular Motors* (Cold Spring Harbor Laboratory Press, New York).
21. Bryant Z, Altman D, Spudich JA (2007) The power stroke of myosin VI and the basis of reverse directionality. *Proc Natl Acad Sci USA* 104:772–777.
22. Menetrey J, Llinas P, Mukherjea M, Sweeney HL, Houdusse A (2007) The structural basis for the large powerstroke of myosin VI. *Cell* 131:300–308.
23. Rock RS, *et al.* (2001) Myosin VI is a processive motor with a large step size. *Proc Natl Acad Sci USA* 98:13655–13659.
24. Altman D, Sweeney HL, Spudich JA (2004) The mechanism of myosin VI translocation and its load-induced anchoring. *Cell* 116:737–749.
25. Berendsen HJC, Vanderspoel D, Vandrunen R (1995) Gromacs - a message-passing parallel molecular-dynamics implementation. *Comput Phys Commun* 91:43–56.
26. Lindahl E, Hess B, Vanderspoel D (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model* 7:306–317.
27. Wang JM, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074.
28. Rohl CA, Baldwin RL (1997) Comparison of NH exchange and circular dichroism as techniques for measuring the parameters of the helix-coil transition in peptides. *Biochemistry* 36:8435–8442.

Supporting Information

Sivaramakrishnan *et al.* 10.1073/pnas.0806256105

SI Methods

Molecular Dynamics Simulations. ***Simulation setup.*** Two PDB structures of the E_4K_4 peptide (CH_3CO-EEEEKKKKEEEEKKKK-$NHCH_3$), one starting as a complete α-helix and the other as a random coil, were generated with the molecular modeling program SYBYL (Tripos). SYBYL output to AMBER PDB was used to facilitate use with the AMBER2003 force field. The PDB files were converted to GROMACS (1, 2) coordinate and topology files by using the GROMACS utility pdb2gmx (version 3.3.1). Hydrogen atoms in the PDB structure were ignored, new protons were added back by pdb2gmx. The AMBER2003 force field (3) ported for use with GROMACS was used. The structure was solvated in a cuboid box of dimensions 4.231 nm $\times$ 5.463 nm $\times$ 7.224 nm with 5547 TIP3P water molecules. The genion utility in GROMACS was used to add Na^+ and Cl^- atoms to a final ionic strength of 150 mM at neutral pH. Glutamic acid (E) and lysine (K) side chains were charged corresponding to neutral pH of the solution. For replica exchange, 48 replicas were created with exponentially distributed temperatures from 274 K to 465 K with final exchange rates for both configurations ranging from 12–24%. Solvated replicas were subject to a preliminary energy minimization step by using the method of steepest descent with 1.4-nm cutoff for neighborlists, and 9-Å cutoff for Coulombic and van der Waals interactions, until a maximum value of less that 10,000 kJ mol^{-1} nm^{-1} was achieved. Long-range electrostatics are treated throughout with the particle-mesh Ewald (PME) algorithm (4). The system was equilibrated at 274 K through 10 ps of molecular dynamics, by using 2-fs time steps with protein-atom coordinates frozen and water-bond lengths constrained. Initial velocities were assigned randomly from a Maxwell-Boltzmann distribution. For both equilibration and the complete MD simulation, a grid-based neighborsearch to 9 Å was conducted every 5 steps. Linear and angular motion of the protein and solvent groups about the center-of-mass were removed at every step. A cutoff of 9 Å was used for both Coulombic and van der Waal's interactions. The LINCS algorithm with default GROMACS parameters was used to constrain all bond lengths. Periodic boundary conditions were used. REMD simulations for both random and α-helical conformations were continued until the fraction helix at 274 K did not show appreciable change in fraction helix with further simulation time (Fig. S1). Fraction helix was computed from the backbone (ϕ,ψ) by using the Lifson-Roig model, where $\phi = -60 \pm 30$ and $\psi = -47 \pm 30°$ for 'n' consecutive angles results in a helical segment of length 'n-2'. The $(E_4K_4)_2$ peptide has an acetyl group at the N terminus and a methyl-amide group at the C terminus and, therefore, can have a maximum helical length of 14. For the random initial configuration each replica was simulated for 80 ns for a total REMD simulation time of 3.8 μs. For the α-helical conformation, each replica was simulated for 42 ns for a total REMD simulation time of 2 μs. As shown in Fig. S1 and Fig. 1*A*, both simulations converge to the same fraction helix at all temperatures simulated. To study the dynamics of side-chain interactions in real time, three folded replicas were randomly selected from the random initial conformations at 274 K and simulated for 50 ns each. With increasing simulation temperature, fraying of the helix ends becomes more prominent and could disrupt the dynamic pattern of side-chain interactions seen at lower temperatures. Because naturally occurring ER/K α-helices are not seen at the ends of proteins but bridge gaps between protein subdomains, we propose that fraying of ends is an artifact of simulating short peptides due to computational limitations. Fraction helix at 274 K is only 6% higher than fraction helix at the temperature used for the SAXS and single-molecule analyses (295 K). Data from the last 25 ns of simulation were used to analyze distances between side-chain atoms. Distance histograms obtained for each of the three simulations were found to be virtually identical, indicating that we sampled sufficient times to capture different modes of interaction.

Simulation hardware. All simulations were performed on the Bio-X^2 cluster at Stanford University.

Data Analysis. Analysis of side-chain interactions was performed on explicit atom trajectories output from GROMACS by using custom Matlab (The Mathworks) computer programs. Maximum likelihood estimator (MLE) fit to the sum of three Gaussians was also performed by using custom Matlab code. The goodness of fit was assessed by generating multiple random datasets based on MLE parameters and estimating their likelihood relative to the experimental dataset. We find that the likelihood estimates generated from MLE parameters are equally distributed about the experimental likelihood, verifying goodness of fit. Water density distribution (Fig. 3*A*) around the helix backbone was calculated by using data from the last 1 ns of the simulation. TIP3P water coordinates were output explicitly in GROMACS trajectory files. The axis of the α-helix was defined by a least-squares-fit to backbone C_α coordinates. Coordinate transformation was performed to align the axis of the α-helix with the z axis. Distances of water oxygen atoms, between the minimum and maximum z-coordinate of the peptide, were computed by using a custom Matlab program. Water atoms were binned into 0.5-Å wide annuli around the helix axis and normalized to the volume of the annuli to get the water density distribution. Distances of C_α, N_ζ, and O_ε from the axis of the α-helix (Fig. 3*A*) were calculated as the average for 25 ns of simulation.

Theoretical Model for Bending of The ER/K α-Helix. To understand the contribution of side-chain interactions to bending stiffness of the ER/K α-helix, we constructed a simplified theoretical model wherein each side-chain interaction is considered to be a harmonic spring at distance d_o = 6 Å from the axis of the α-helix (Fig. S6*A*). The side-chains interact through a combination of hydrogen bonding (3 Å) and intermediate (5 Å) modes (Fig. 2*C*). However, these two modes are not independent of each other, because they derive from the same distribution. Bending of the α-helix changes the side-chain interaction distances (Fig. 3*B*). The hydrogen-bonding interactions centered at 3 Å can accommodate only very small extensions (0.3 Å). But, the presence of a second interaction centered at 5 Å with a larger reach of 0.9 Å, that follows the hydrogen-bonding interaction, prevents the α-helix from falling apart in the presence of lateral bending forces, such as those that occur during the swing of the myosin VI or myosin X lever arm. Also, the harmonic spring constant for the 3-Å interaction is 10-fold larger than that for the 5-Å interaction. For side-chain interaction distances corresponding to the 5-Å potential well each E (i) interacts simultaneously with K ($i - 4$) and K ($i + 3$). However, for the 3-Å potential well each E (i) interacts with only one of K ($i - 4$) and K ($i + 3$). Thus, the 5-Å potential well, in addition to its longer reach, provides a continuous sheet of interactions without breaks along the α-helical backbone. Therefore, we model the side-chain interactions as a continuous narrow tube with stiffness corresponding

to the 5-Å potential well. For mechanoenzymes such as myosin VI or X, the ER/K α-helix will experience a lateral bending force at one end (for example the cargo binding domain of myosin VI), which is then transmitted to the other end held fixed by stiff protein structural elements (for example the proximal tail of myosin VI). Therefore, we model the ER/K α-helix as a cantilever beam (Fig. S6*A*) (5). The force and deflection of a cantilever beam are related by Eq. **3** (Fig. S6). The ER/K α-helix has continuous series of interacting residues along multiple faces of the α-helical backbone (Fig. 1 *C* and *D*). Therefore, each lateral cross section of the ER/K α-helix is stabilized by either three or four side-chain pairs. A pitch of approximately $N = 3.6$ residues for the α-helix results in the location of the side-chain pairs, in successive cross sections from its N to C terminus, staggered relative to each other (Fig. S6*B*). Therefore, we model the side-chain pairs at each cross section as N springs of stiffness $k_{5Å} = 440$ pN/nm (corresponding to the 5-Å peak), uniformly distributed at $d_o = 6$ Å from the α-helix axis. The stiffness of these springs provides for the Young's modulus (E) and their distribution provides for the second moment of area (I) (Fig. S6, Eqs. **1** and **2**). Cantilever beam theory (Fig. S6, Eq. **3**) was then used to estimate a force per unit deflection (stiffness) of 0.46 pN/nm for the ER/K α-helix, limited to small deflections (≈1 nm for a 10-nm α-helix).

Simplifications and Uncertainties. The theoretical model assumes that the α-helix can be treated as a cantilever beam, with the side-chain interactions represented as harmonic springs that resist the deflection of the α-helix. To construct this model, we interpret the two prominent Gaussian peaks seen in the histogram of minimum distances as harmonic potential wells, in which side-chain interactions prefer to reside 82% of the time in equilibrium MD simulations. This interpretation is based on the assumption that the attractive forces between oppositely-charged atoms in the side chains would resist change in distance between them during bending of the α-helix. The 3-Å potential well represents a hydrogen-bonding interaction between E and R/K side chains, but the source of the 5-Å potential well is not completely clear. Importantly, we note that there is no evidence that the 5-Å potential well will resist bending. However, this is a reasonable assumption given that for conformations corresponding to the 5-Å potential well, several charged H atoms in the R/K side chain are in close proximity to the electronegative O_ε atom of E contributing to attractive forces between side chains. Our model imposes the simplification of treating the bending of the ER/K α-helix as a cantilever beam based on neutral axis bending theory (5). Despite these caveats, our model successfully accounts for the observed stiffness of the ER/K α-helix based on physically reasonable side-chain interactions.

Charge Interaction Map of ER/K α-Helices. ER/K sequences are arranged such that R/K at i are aligned vertically with E at $i - 3$ or $i + 4$. This implies also that E at i is aligned vertically with R/K at $i - 4$ or $i + 3$. To construct this map (see Fig. 1*C* and Fig. S4) the sequence is written from right to left and bottom to top starting with the first residue of the ER/K motif at the bottom right hand corner. For most ER/K sequences this alignment is easily done by switching lines when the charge changes sign (E↔R/K). On changing lines, the subsequent R/K or E residue is placed above its interacting partner. In regions where multiple interactions are possible, the map is aligned to maximize the number of displayed interactions. Vertical lines through this map show charged residues that can interact with each other along the α-helix backbone.

SAXS Experiments. ***Protein expression and purification.*** For IQ-PT-CAM the IQ-PT human myosin VI cDNA (814–918) was cloned into a pGEX 6P1 vector (GE Healthcare, Piscataway NJ) with human calmodulin (CAM) cloned into a pET28a vector (EMD Chemicals, San Diego, CA). Myosin VI medial tail (MT) [human myosin VI cDNA 916–981] was cloned into a modified pET28a vector (EMD Chemicals, San Diego, CA). This specific MT construct was designed with additional sequences to enable dye/ligand conjugation at each end. The N terminus has the additional sequence EEEEKKC providing a highly reactive cysteine. The C terminus has the additional sequence KK-*DS*LEFIASKLA-KWK with the 11 central residues for labeling with serine phosphoryl transferase (sfp) at the first serine (*S*) (6). The modified vector contains a His_6 tag, maltose binding protein and a TEV protease cleavage site 5′ to the cloning site. IQ-PT and CAM vectors were coexpressed in *Escherichia coli* Rosetta (DE3) cells (EMD Chemicals, San Diego CA). pET28a provides kanamycin resistance, whereas pGEX 6P1 provides ampicillin resistance. Hence, cells were grown in triple selection media (including chloramphenicol for Rosetta cells). The MT vector was expressed independently in *E. coli* Rosetta (DE3) cells. Cells were lysed by sonication in the presence of 1–1.5 mg/ml lysozyme. Lysates were clarified by centrifugation at 100,000 × *g* for 30 min. IQ-PT-CAM lysates were bound to a GSH column (GE Healthcare, Piscataway NJ), followed by elution with reduced glutathione. Eluted proteins were dialyzed in the presence of Precision protease to cleave the GST from the IQ-PT-CAM domain. Dialysis products were further purified by GSH column chromatography to remove GST. After the GSH column chromatography, the relative amounts of IQ-PT and CAM were assessed by Coomassie stained SDS/PAGE gels. Purified human CAM was added back to obtain approximate equimolar amounts of IQ-PT and CAM, followed by dialysis of the sample into scattering buffer containing 20 mM phosphate, 100 mM NaCl, 1 mM DTT, 1 mM EGTA, and 1 mM EDTA. Samples were further purified on a superdex 200 gel fitration column (GE Healthcare) to obtain a single species of the protein complex containing equimolar amounts of IQ-PT and CAM. The presence of a single species of protein complex was verified independently by using analytical size exclusion column (Superdex 75), multiangle laser light scattering (MALLS), and dynamic light scattering (DLS) measurements, from which we estimated sample dispersity at <1%. MT lysates were bound to Ni-NTA resin (Qiagen). The eluted proteins were dialysed in the presence of TEV protease (1–100 by weight) overnight to cleave the His_6 and MBP portions from the MT domains. MT domains were further purified by running the cleavage reaction over a Ni-NTA column to remove His_6 containing fragments, followed by a superdex 200 column (GE Healthcare) by using either Circular Dichroism (CD) buffer (10 mM phosphate, pH 7.4/25 mM NaCl) or scattering buffer (see above).

SAXS measurements. SAXS measurements were carried out at the XOR/BESSRC undulator beam line 12-ID of the Advanced Photon Source (Argonne, IL) employing a sample-detector distance of 2 m and CCD detector read out (MAR USA). The data were collected by using a custom-made sample cell (7) at an x-ray energy of 12 keV. Details of the beam line were as described previously (7–9)

Samples were clarified by centrifugation at 10,000 × *g* for 10 min before measurement and five 0.5 s exposures were obtained. Data were image corrected, normalized by incident flux and circularly averaged. The five profiles for each condition were averaged to improve signal quality. Buffer profiles were collected by using identical procedures and subtracted for background correction. The data showed no signs of radiation damage, based on comparison of consecutive scattering profiles from the same sample (data not shown).

SAXS data analysis. Scattering intensities as a function of the momentum transfer were obtained at different protein concentrations. The SAXS profiles for the IQ-PT-CAM complex were superimposable after scaling by forward scattering intensity,

suggesting that there was no detectable aggregation or interparticle interference effects (Fig. S8*E*). The MT construct has a highly reactive cysteine, attached for use in future experiments, which could react to form a dimer under very high concentrations; hence results by using lower protein concentration were used (Fig. S8*A*). Radii of gyration (R_g) were determined from Guinier analysis of the low angle scattering data (10) (Fig. S8 *B* and *F*). R_g obtained from Guinier analysis agree within experimental errors with the values from the real space distribution function $P(r)$ computed by using the regularized transform method implemented in the program GNOM (11). The R_g for the IQ-PT-CAM complex and the MT are ≈2.3 and 3.2 nm, respectively. Our modeled structures (Fig. 4*B*) have corresponding R_g values of 2.2 and 3.1 nm, respectively.

SAXS Structure Reconstructions. The programs DAMMIN (12) and GASBOR (13) were used to construct 3-D bead models that fit the scattering data (Fig. S8 *C*, *D*, *G*, and *H*). Both programs employ a simulated annealing procedure and a compactness criterion. Ten independent DAMMIN and GASBOR runs were performed for each scattering profile, by using default parameters, the "slow" mode for DAMMIN, no symmetry assumptions (P1 symmetry), and the full recorded scattering profiles. The models resulting from independent runs were superimposed and compared by using the program SUPCOMB (14), based on the normalized spatial discrepancy (NSD) criterion. Models with NSD values <1 were considered similar. For all data presented in the main text, the 10 independent repeat runs yielded models with pairwise NSD values <1, indicating that the algorithms converged reproducibly to similar structures. The 10 independent structures for each scattering profile were subsequently averaged, and "filtered" consensus models were computed by using the program DAMAVER with default settings (15). Consensus models constructed with DAMMIN and GASBOR gave similar results. For visualization, the reconstructed bead models were converted to electron density maps with the program Situs (16).

CD. CD Spectra were acquired by using an Aviv 62DS instrument (Aviv Biomedical) with a 1-mm path length cell in CD buffer. Spectra were taken at 0°C with data collected every 1 nm with a 10 second averaging time, and are the average of 2 repeat scans. Protein concentrations ranged from 1–17 μM. Melt data were collected every 1°C with a 30 s averaging time and a 2 min equilibration. In all cases the reverse melt showed at least 90% reversibility. CD data were converted from mean residue ellipticity (MRE) to fraction helix by using equations outlined in Rohl and Baldwin (17, 18).

MALLS. In solution molecular mass of the IQ-PT-CAM complex was determined by using a size exclusion chromatography system coupled to a multiple angle light scattering detector. Protein concentrations were determined with an Optilab rEX refractive index detector, and scattering was detected with a Dawn 18 angle MALLS light scattering instrument (Wyatt Technology Corporation). MALLS molecular mass of the IQ-PT-CAM complex was found to be 27.8 ± 0.3 kDa at a maximum concentration of 100 μM with <5% polydispersity. The expected molecular mass of the IQ-PT-CAM complex is 29 kDa.

DLS. DLS measurements were made by using a DynaPro instrument (Protein Solutions) running Dynamics version 6 software. IQ-PT-CAM samples at 1–10 mg ml^{-1} in scattering buffer were assayed at 25°C with an acquisition interval of 10 s. Samples were spun at 15,000 × *g* for 10 min immediately before analysis. Results were derived by using a regularization fit and with PBS buffer settings. DLS measured a hydrodynamic radius (R_h) of 2.3 ± 0.2 nm for the IQ-PT-CAM complex.

Single-Molecule Optical-Trap Assays. ***Protein expression and purification.*** Two different porcine myosin VI sequences, beginMT [1–918] and endMT [1–981], followed by $(GSG)_4$, eYFP, and an N-terminal FLAG tag (DYKDDDDK), were cloned into pBiex-1 (Novagen) and expressed by transfection of plasmid DNA into Sf-9 cells (19).

Optical trapping. Optical trapping was performed by using the custom dual-beam optical trap (20–22). All trapping was done without feedback. Dumbbells of actin were formed by using a streptavidin-biotin link between streptavidin coated beads and 100% biotinylated F-actin. Phalloidin was added to a final concentration of 10 μM to stabilize actin filaments. The actin dumbbell was pulled taut with trap stiffness of ≈0.005 pN/nm. Motors were attached to nitrocellulose coverslips with monoclonal anti-GFP antibody (Chemicon MAB3580). Motor dilutions were chosen such that <10% of tested platforms showed myosin VI binding to actin, to minimize binding of more than one myosin head to the F-actin filament. Binding events were determined from data traces, by eye, by using both decrease in bead position variance and decrease in bead-to-bead correlation.

1. Berendsen HJC, Vanderspoel D, Vandrunen R (1995) Gromacs - a message-passing parallel molecular-dynamics implementation. *Comput Phys Commun* 91:43–56.
2. Lindahl E, Hess B, Vanderspoel D (2001) GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J Mol Model* 7:306–317.
3. Wang JM, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 21:1049–1074.
4. Robertson A, Luttmann E, Pande VS (2008) Effects of long-range electrostatic forces on simulated protein folding kinetics. *J Comput Chem* 29:694–700.
5. Beer F, Johnston E, Jr (1981) *Mechanics of Materials* (McGraw Hill, New York).
6. Yin J, Lin AJ, Golan DE, Walsh CT (2006) Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nat Protoc* 1:280–285.
7. Lipfert J, Millett IS, Seifert S, Doniach S (2006) A Sample Holder for Small-Angle X-ray Scattering Static and Flow Cell Measurements *Rev Sci Inst* 77:046108.
8. Beno MA, *et al.* (2001) Basic Energy Sciences Synchrotron Radiation Center Undulator Sector at the Advanced Photon Source *Nucl Instr Meth Phys Res A* 467–468(1):690–693.
9. Seifert S, Winans RE, Tiede DM, Thiyagarajan P (2000) Design and performance of a ASAXS instrument at the Advanced Photon Source. *J Appl Crystallogr* 33:782–784.
10. Guinier A (1939) La diffraction des rayons X aux tres petits angles: Application'a l'etude de phenomenes ultramicroscopiques. *Ann Phys* 12:161–237.
11. Svergun D (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 25:495–503.
12. Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* 76:2879–2886.
13. Svergun DI, Petoukhov MV, Koch MH (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* 80:2946–2953.
14. Kozin MB, Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* 34:33–41.
15. Volkov VV, Svergun DI (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr* 36:860–864.
16. Wriggers W, Chacon P (2001) Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering. *J Appl Crystallogr* 34:773–776.
17. Luo P, Baldwin RL (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water. *Biochemistry* 36:8413–8421.
18. Rohl CA, Baldwin RL (1997) Comparison of NH exchange and circular dichroism as techniques for measuring the parameters of the helix-coil transition in peptides. *Biochemistry* 36:8435–8442.
19. Bryant Z, Altman D, Spudich JA (2007) The power stroke of myosin VI and the basis of reverse directionality. *Proc Natl Acad Sci USA* 104:772–777.
20. Altman D, Sweeney HL, Spudich JA (2004) The mechanism of myosin VI translocation and its load-induced anchoring. *Cell* 116:737–749.
21. Rice SE, Purcell TJ, Spudich JA (2003) Building and using optical traps to study properties of molecular motors. *Methods Enzymol* 361:112–133.
22. Spudich JA, Rice S, Rock R, Purcell T, Warrick H (2008) *Optical Traps to Study Properties of Molecular Motors* (Cold Spring Harbor Laboratory, New York).

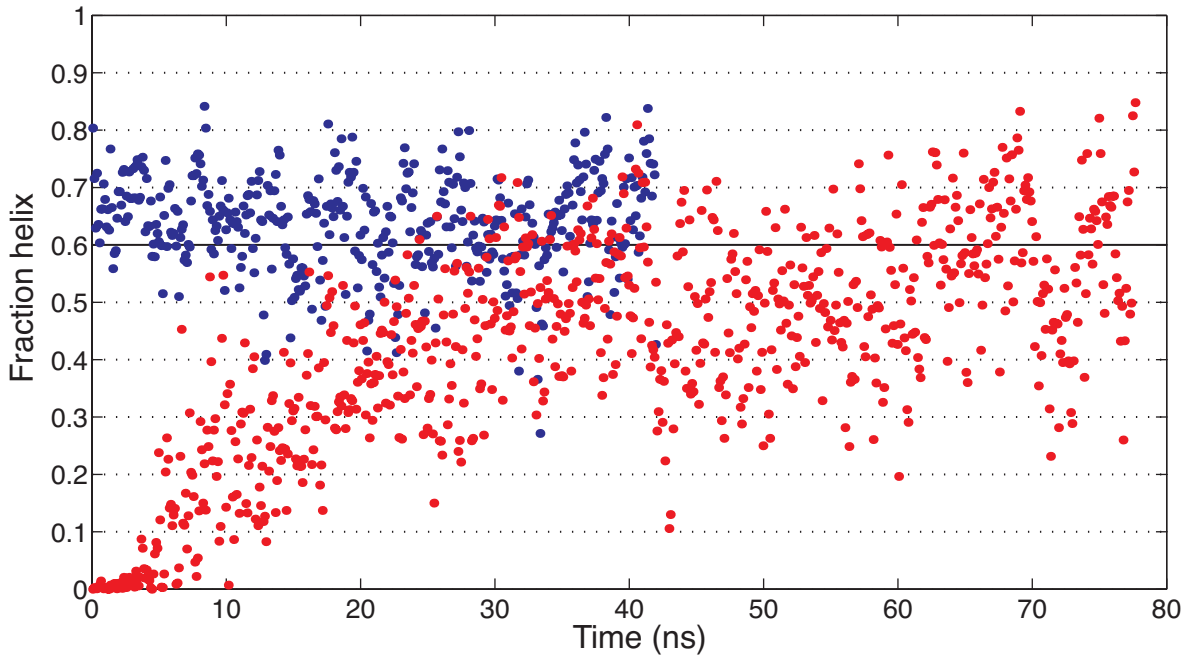

Fig. S1. Convergence of fraction helix in REMD simulation from α-helical and random initial conformations. Change in fraction helix at 274 K was computed by using the Lifson-Roig model as a function of time during REMD simulation. Data are block averaged over 0.2-ps intervals. Blue symbols, starting with fully α-helical conformation; red symbols, starting with random initial conformation. The convergence of simulations is seen by near zero slope toward the end of each simulation. Additionally, both simulations converge to the same final value.

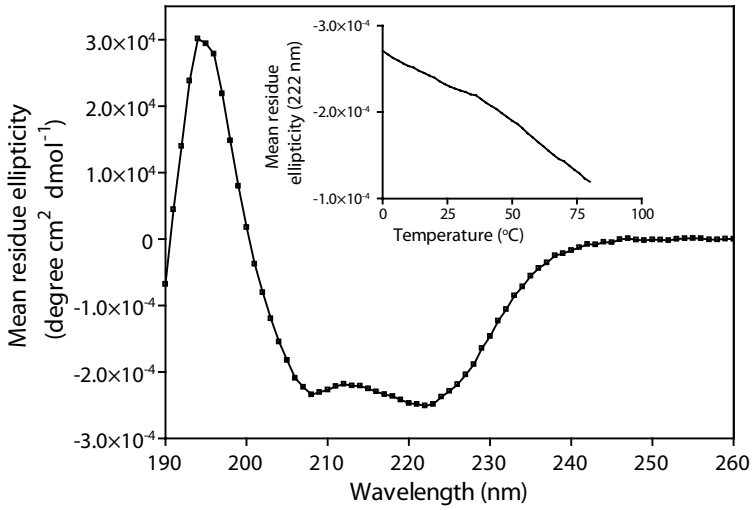

Fig. S2. The ER/K motif in the medial tail of myosin VI is α-helical in solution. MRE measured by using circular dichroism at 273 K (0°C) for the medial tail of myosin VI shows characteristic minima at 208 and 222 nm. The inset shows the MRE at 222 nm as a function of temperature (thermal melt).

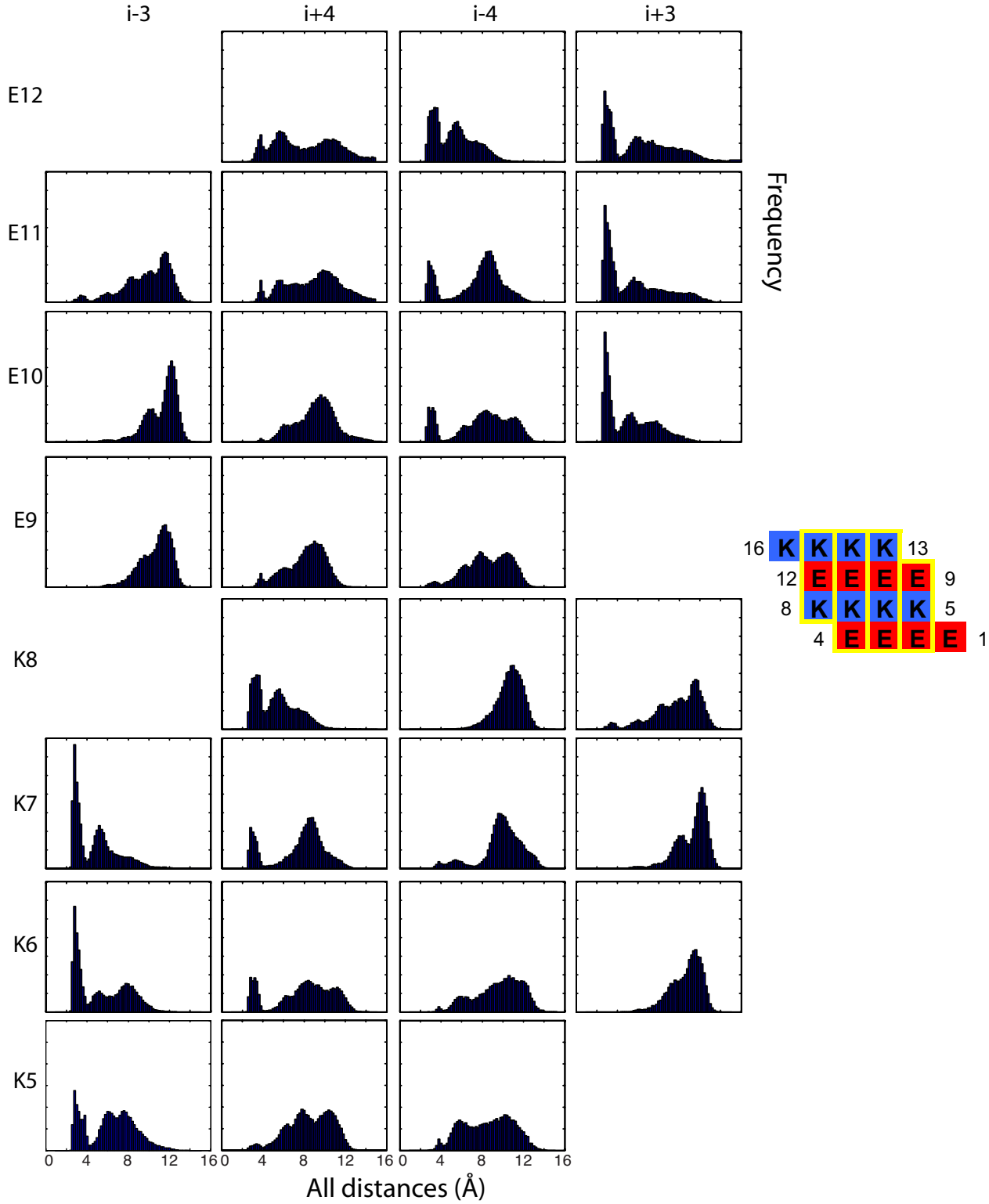

Fig. S3. K residues at *i* preferentially interact with E at $i-3$ and $i+4$ positions, whereas E residues at *i* preferentially interact with K at $i-4$ and $i+3$ positions. Histograms of distance distributions between N_ζ/O_ε atom of residue *i* at positions K5-E12 with O_ε/N_ζ atom at $i-4$, $i-3$, $i+3$, and $i+4$ positions. The *x*-axes are distance in Å. Data were acquired over 25 ns of simulation time. Columns left to right are $i-3$, $i+4$, $i-4$, and $i+3$ distances, respectively, whereas rows correspond to residues K5-E12 bottom to top. Note the clear 3-Å peaks with at $i-3$ and $i+4$ for K residues, whereas E residues have 3-Å peaks with $i-4$ and $i+3$ residues. Importantly, residue K8 (*i*) does not interact with E4 [$i-4$], and residue E9 (*i*) does not interact with K13 ($i+4$), as evidenced by the absence of 3-Å peaks in the corresponding distributions. Also, K8 (*i*) does not have a partner at $i-3$ (K5), whereas E9 (*i*) does not have a partner at $i+3$ (E12).

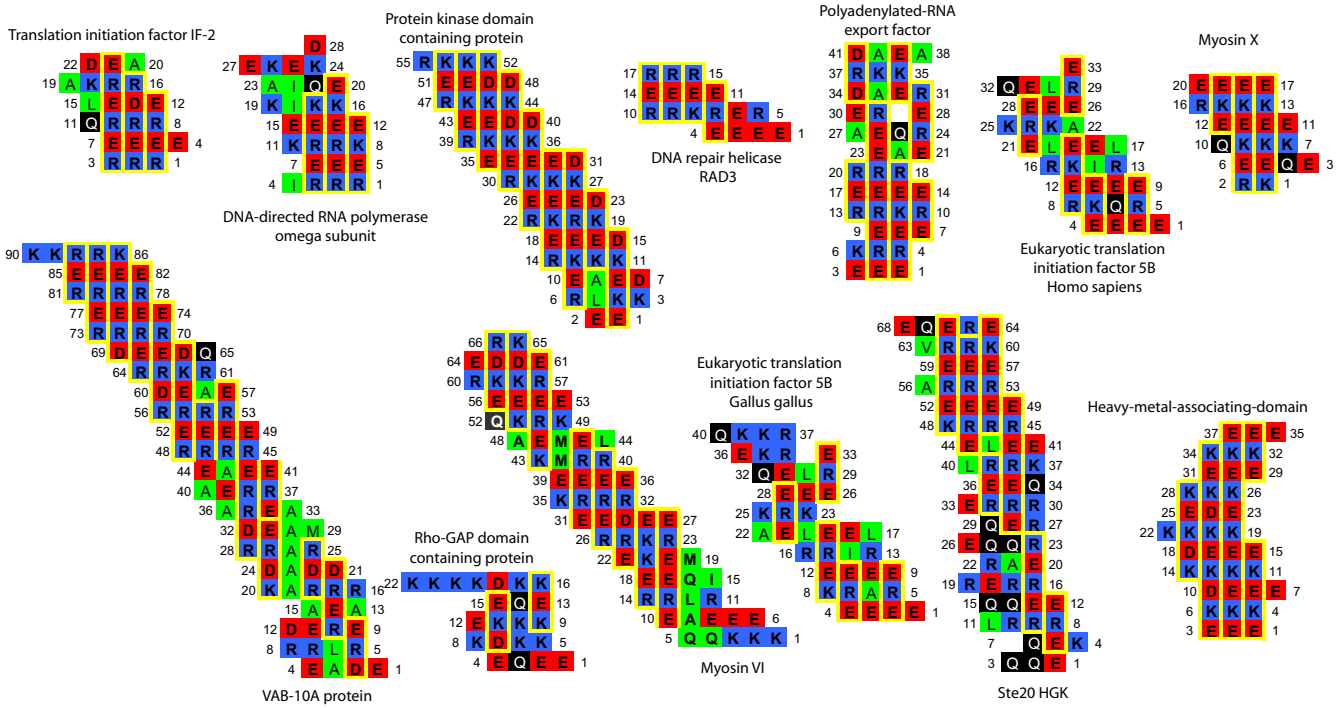

Fig. S4. Charge interaction maps for naturally occurring ER/K motifs. Charge interactions for proteins shown in Table S1 are displayed by using this interaction map (maps for mannosyltransferase and upf2 regulator of nonsense transcripts are shown in Fig. 1 *C* and *D*, respectively). Based on MD simulations of the $(E_4K_4)_2$ sequence residue E (*i*) interacts with R/K located at $i-4$ or $i+3$, and residue R/K (*i*) interacts with R/K located at $i-3$ or $i+4$. Sequences are shown with N termini at the bottom right-hand corner. Sequences are read right to left, bottom to top. Residue numbers relative to the first residue of the ER/K motif are listed beside the sequences. Successive rows of the map are aligned such that vertical lines represent the interacting residues. Most ER/K motifs have several columns with multiple residues that share interactions along the α-helical backbone (see yellow boxed groups of residues in columns).

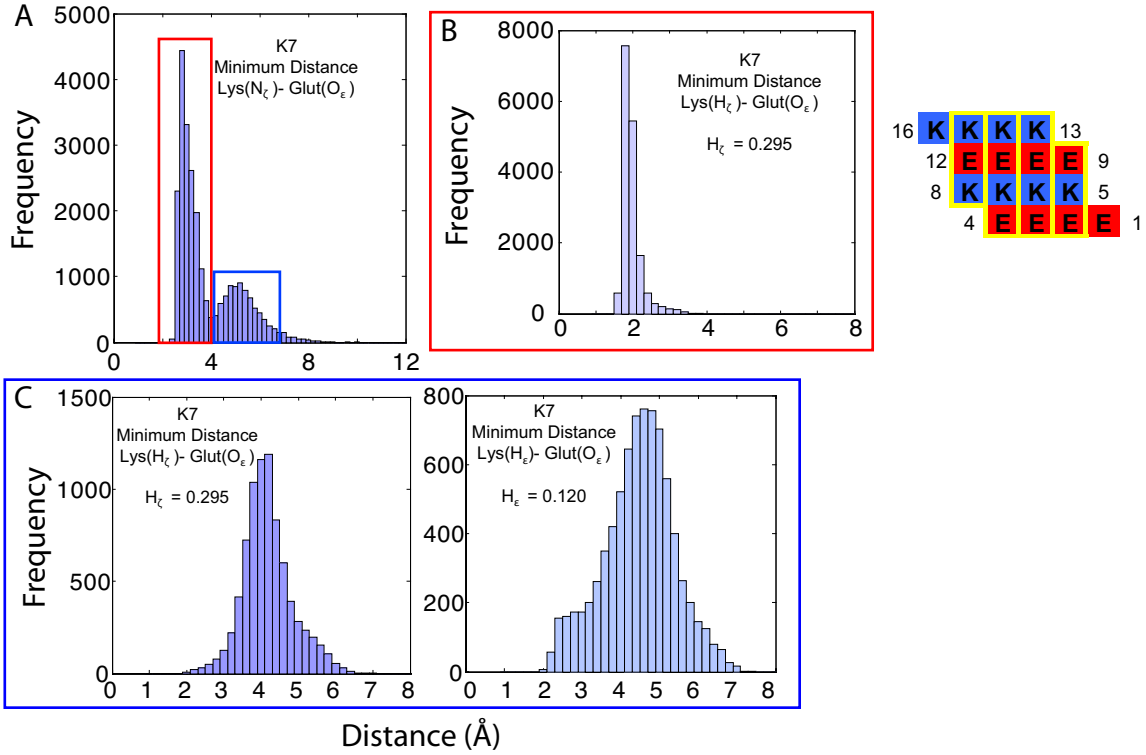

Fig. S5. The 5-Å preferred distance between N_ζ of K and the centers of mass of O_ε of E in the $(E_4K_4)_2$ α-helix is accompanied by multiple interactions between charged H-atoms in the lysine side chain and the O_ε atom of glutamic acid. (*A*) Histogram of minimum distances between K7-N_ζ (*i*) and the centers of mass of O_ε for E3 ($i-4$), E4 ($i-3$), E10 ($i+3$), and E11 ($i+4$). Red and blue boxes indicate the two peaks at 3 Å and 5 Å, respectively. (*B*) The smallest distance between three H_ζ atoms of K7 and two O_ε atoms of E3, E4, E10, and E11 for the portion of the histogram in (*A*) indicated by the red box. H_ζ and O_ε are at hydrogen-bonding distance for N_ζ–O_ε distances corresponding to the 3-Å peak in (*A*). (*C*) The smallest distance between three H_ζ atoms (*Left*) or two H_ε atoms (*Right*) of K7 and the two O_ε atoms of E3, E4, E10, and E11 for the portion of the histogram in (*A*) indicated by the blue box. For convenience, 4 Å is used as the demarcation between the red and blue boxes in (*A*). (*B*) and (*C*) list the partial positive charges of the H_ζ and $H_\varepsilon$atoms resulting from direct and indirect-bonding interactions with the electronegative N_ζ. Note that several H-atoms in the K7 side chain, each with significant partial positive charges are within close proximity to the highly electronegative O_ε atom.

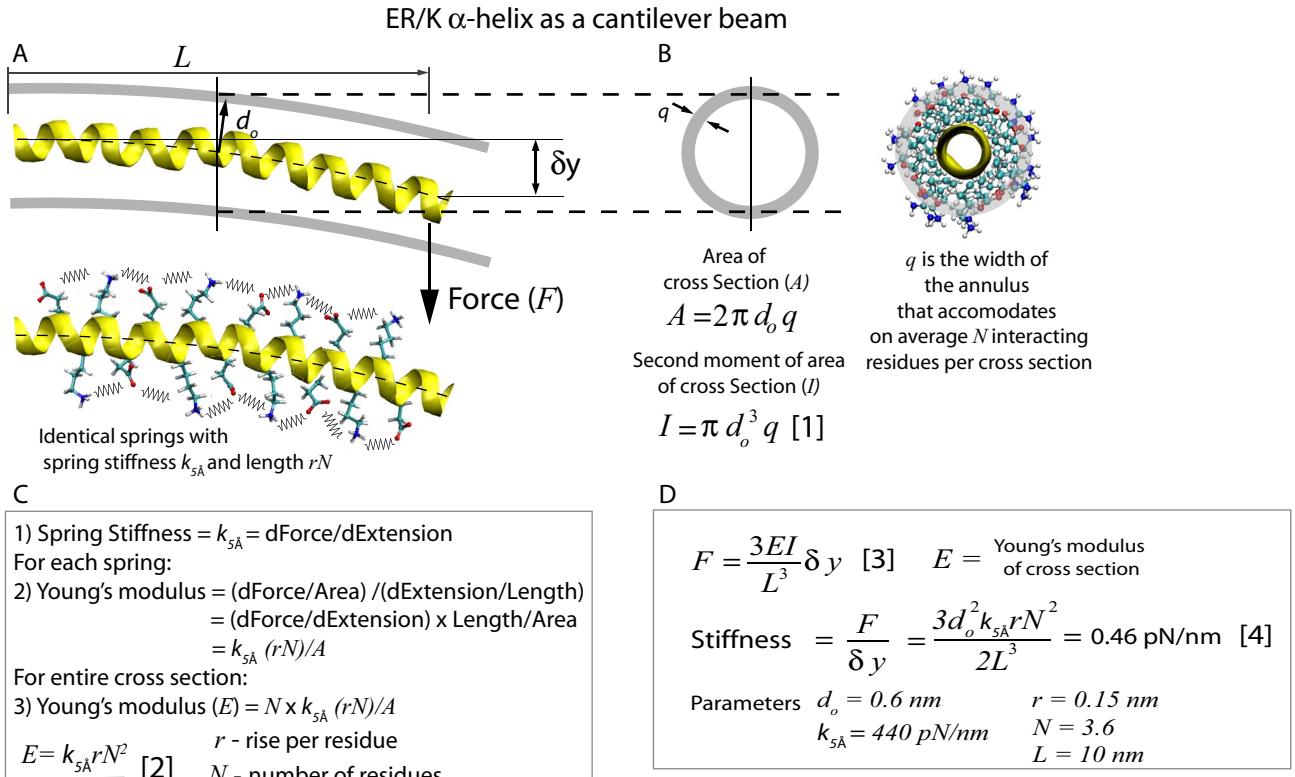

Fig. S6. Model of ER/K α-helix as a cantilever beam. (*A*) (*Upper*) The α-helical backbone is shown in ribbon representation (yellow) with the location of the interacting atoms of the E and R/K side chains shown as a gray tube surrounding the ER/K α-helix. *L* is the length of the α-helix, whereas *F* is the lateral bending force applied at one end, resulting in δ*y* deflection at that end. d_o is the average distance of side-chain N_ζ and O_ε atoms, of K and E respectively, from the axis of the α-helix. (*Lower*) Side-chain pairs are modeled as harmonic springs located at fixed distance d_o = 6 Å from the α-helix axis, each of which has a spring constant $k_{5Å}$ = 440 pN/nm, estimated from distribution of side-chain pair distances (Fig. 2*B*). Although the schematic representation only shows residues above and below the neutral axis for clarity, the model is based on side-chain pairs distributed continuously around the α-helix. (*B*) (*Left*) The cross section of the ER/K α-helix is approximated as an annulus of width *q* with area $A = 2\pi d_o q$ and second moment of area $I = \pi d_o^3 q$ (Eq. **1**). (*Right*) A view down the axis of a 29-aa long $(E_4K_4)_n$ α-helix (approximately eight turns) represents the staggered placement of residues. (*C*) The derivation of equivalent Young's modulus (*E*) (Eq. **2**) of our model as a function of harmonic spring stiffness of side-chain interactions is shown. (*D*) The force-deflection relationship is shown by using cantilever beam theory (Eq. **3**) (5), with *E* and *I* approximated for the ER/K α-helix as described in *A*–*C*. Substitution of *E* and *I* yields a bending stiffness of 0.46 pN/nm for small deflections (<1 nm) for a 10-nm long ER/K α-helix.

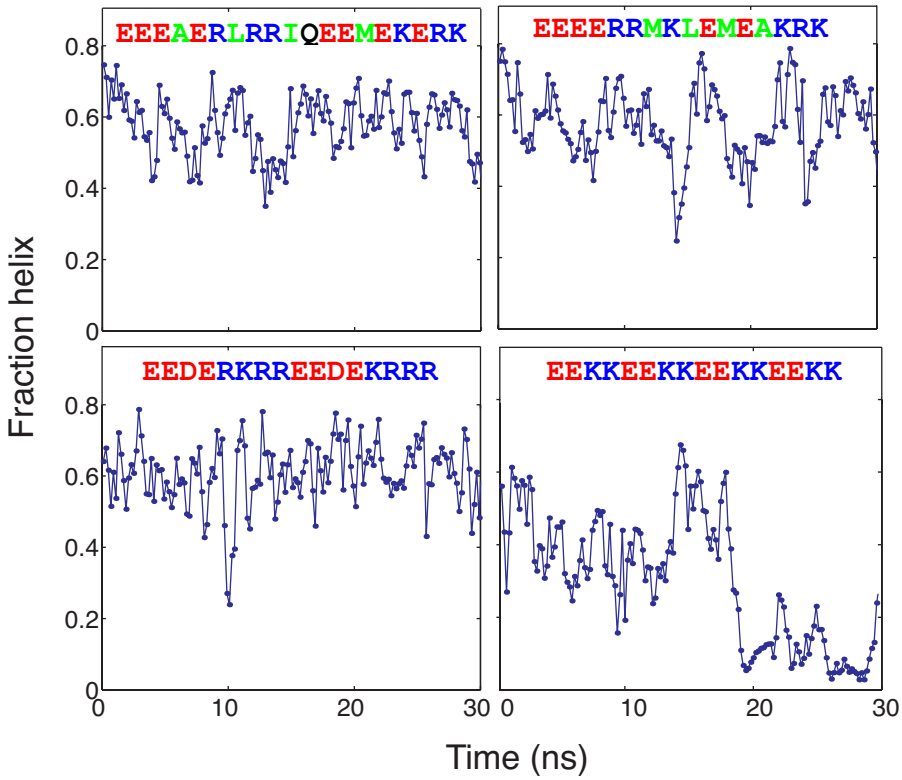

Fig. S7. Hydrophobic residues in the ER/K motif could reduce the local stability of the ER/K α-helix. Thermal melts are shown for different segments of the MT ER/K α-helix in myosin VI. Three different segments of the myosin VI medial tail domain containing different numbers of hydrophobic residues (A, L, I, and M) were folded into fully α-helical conformation and simulated in explicit solvent at 400 K to study their relative stabilities. Sequence $(E_2K_2)_4$ is simulated as a negative control as it is known to have low α-helical content. The $(E_2K_2)_4$ helix unravels completely within 20 ns, whereas the three segments of the myosin VI medial tail remain stable over the same time period. However, segments containing hydrophobic residues show transient reductions in fraction helix, which is a result of local unraveling and refolding of helix turns in the hydrophobic regions (data not shown).

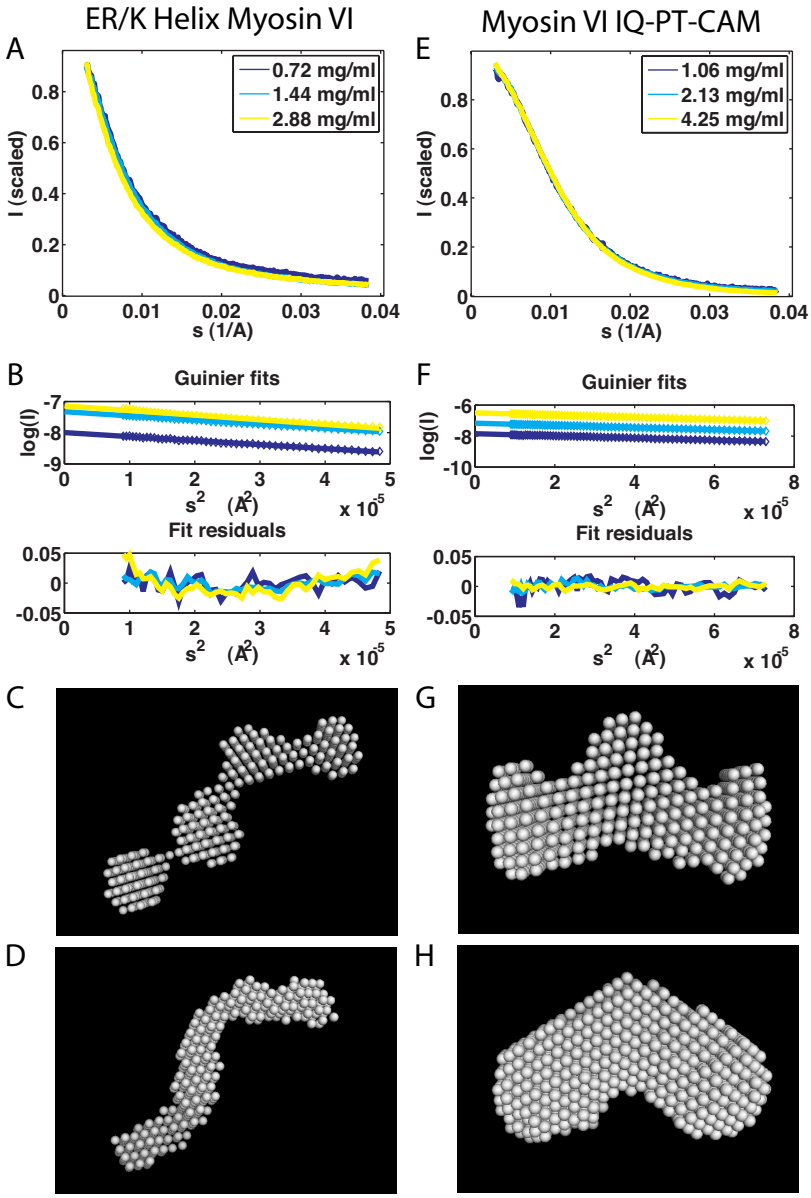

Fig. S8. SAXS results for myosin VI MT (*A–D*) and IQ-PT-CAM (*E–H*). (*A* and *E*) Scaled scattering profiles are shown of *I* (scattering intensity normalized by forward scattering intensity) versus *s* ($2 \sin\theta/\lambda$, where θ and λ are the scattering angle and x-ray wavelength, respectively) at indicated concentrations. (*B* and *F*) The linear Guinier fit to the scattering data are shown for the IQ-PT-CAM construct for different protein concentrations. The color scheme used is identical to (*A* and *E*). The fit residuals (*Lower*) do not show any significant deviation from zero, nor any consistent curvature. Fits to all three datasets gave consistent R_g values. SAXS bead reconstructions were obtained by using DAMMIN (*C* and *G*) and GASBOR (*D* and *H*) reconstruction algorithms.

Table S1. Examples of protein ER/K motifs in nature

Species	Protein Name	Sequence
Methylobacterium populi	Translation initiation factor IF-2	RRREEEERRRQEDELRRKAAED
Streptococcus pneumoniae	DNA-directed RNA polymerase omega subunit	RRRIEEEKRRKEEEEKKIKEQIAKEKED
Tetrahymena thermophila	Protein kinase domain-containing protein	EEKKLRDEAEKKKRDEEEKRKRDEEEKKKRDEEEEKKKRDDEEKKKRDDEEKKKR
Neurospora crassa	DNA repair helicase RAD3	EEEERERKRREEEEERRR
Chlamydomonas reinhardtii	Polyadenylated-RNA export factor	EEERRKEEERKRREEEEERRREAERQEAEREREADKKRAEAD
Saccharomyces cerevisiae	Mannosyltransferase	EERKRREKKKKEEEEKKKKEEEEKKKKEEEEKKKKEEEEKKKKEEEEKKKKEEEEKKKKEEEEKKKKEEEEKKKQEEEEKKKKEEEEKKKQEE
Dictyostelium discoideum	RhoGAP domain-containing protein	EEQEKKDKKKKEEQEKKDKKKK
Caenorhabditis elegans	VAB-10A protein	EDAERLRREREDAEARRRAKDDADRARRMAEDAERARREAEEAERRRREEEERRRREAEDRKRRqDEEDRRRREEEERRRREEEEKRRKK
Rattus norvegicus	UPF2 regulator of nonsense transcripts homolog	KKKKLEEDKRKKEDKERKKKEEEKVKAEEELKKKEEEEKKKQEEEERKKQEEQAKR
Mus musculus	Myosin VI	KKKQQEEEAERLRRIQEEMEKERKRREEDEERRRKEEEERRMKLEMEAKRKQEEEERKKREDDEKR
Gallus gallus	Eukaryotic translation initiation factor 5B	EEEERAKREEEERIRRLEELEAKRKEEERLEQERKERKKQ
Homo sapiens	Ste20 group protein kinase HGK	EQQKEQRRRLEEQQRREREARRQQEREQRRREQEEKRRLEELERRRKEEEERRRAEEEKRRVEREQE
Homo sapiens	Eukaryotic translation initiation factor 5B, isoform CRA_a	EEEERQKREEEERIKRLEELEAKRKEEERLEQE
Homo sapiens	Myosin X	KREQEEKKKQEEEEKKKREEEERERERERREAELRA
Arabidopsis Thaliana	Heavy-metal-associated domain-containing protein	EEEKKKEEEDKKKKEEEDKKKKEDEKKKEEEKKKEEE