Molecular model of human tropoelastin and implications of associated mutations

Anna Tarakanova[a], Giselle C. Yeo[b,c], Clair Baldock[d], Anthony S. Weiss[b,c,e], and Markus J. Buehler[a,1]

[a]Laboratory for Atomistic and Molecular Mechanics, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 01239; [b]School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia; [c]Charles Perkins Centre, The University of Sydney, Sydney, NSW 2006, Australia; [d]Wellcome Trust Centre for Cell-Matrix Research, Division of Cell Matrix Biology and Regenerative Medicine, School of Biological Sciences, Manchester Academic Health Science Centre, The University of Manchester, Manchester M13 9PT, United Kingdom; and [e]Bosch Institute, The University of Sydney, Sydney, NSW 2006, Australia

Edited by William A. Goddard III, California Institute of Technology, Pasadena, CA, and approved May 23, 2018 (received for review January 22, 2018)

Protein folding poses unique challenges for large, disordered proteins due to the low resolution of structural data accessible in experiment and on the basis of short time scales and limited sampling attainable in computation. Such molecules are uniquely suited to accelerated-sampling molecular dynamics algorithms due to a flat-energy landscape. We apply these methods to report here the folded structure in water from a fully extended chain of tropoelastin, a 698-amino acid molecular precursor to elastic fibers that confer elasticity and recoil to tissues, finding good agreement with experimental data. We then study a series of artificial and disease-related mutations, yielding molecular mechanisms to explain structural differences and variation in hierarchical assembly observed in experiment. The present model builds a framework for studying assembly and disease and yields critical insight into molecular mechanisms behind these processes. These results suggest that proteins with disordered regions are suitable candidates for characterization by this approach.

tropoelastin | elastic fiber | structural protein | disordered protein | molecular dynamics

Protein folding remains an outstanding challenge for experiment and simulation, in particular for large, highly dynamic, and largely disordered molecular systems. Proteins with a high degree of disorder represent >30% of all proteins in eukaryotic cells (1) and play important functions, from cell signaling and membrane transport to protein folding and macromolecular assembly processes, including misfolding and toxic aggregation in disease (2, 3). An absence of tools to predict fully atomistic structure for this class of proteins therefore presents a significant obstacle for probing their function (4).

Molecular dynamics (MD) simulations have supplemented experimental approaches such as X-ray crystallography. However, ab initio protein folding via MD has been achieved only for small molecules, where folding times may be on the order of microseconds, and is becoming more accessible to MD methods through an increase in computing capabilities via massive parallelism and graphics processing unit (GPU) computing (4). Large proteins beyond the length of short peptides remain elusive, because folding time scales are on the order of seconds, inaccessible to MD methods.

The main obstacle of traditional MD is ergodicity, stemming from the challenge of describing a system with a vast number of degrees of freedom that results in a complex free-energy landscape. MD methods, and even accelerated-sampling algorithms based on collective variable biasing, temperature acceleration, and tempering have not successfully reproduced ab initio protein folding for systems with hundreds of amino acid residues. We hypothesized that the dynamic nature and flat-energy landscape of largely disordered proteins would render them amenable to folding via accelerated-sampling MD (5). We show here that replica exchange MD (REMD) can be used to fold a 698-residue tropoelastin protein, validate the structure by comparison with experimental data, and use it to study protein functionality and disease etiology associated with mutations in the *ELN* gene.

Tropoelastin is the precursor molecule of elastin, which, together with microfibrils, forms elastic fibers, an essential component of the extracellular matrix. Elastic fibers provide elasticity and resilience to vertebrate tissues, in particular the skin, lungs, and connective and vascular tissue (6, 7). The tropoelastin protein is encoded by a single gene, *ELN*, with 34 exons, giving rise to a mature molecular structure with a molecular weight of ∼60 kDa (7). Within the tropoelastin sequence, hydrophobic domains rich in repetitive motifs of glycine, valine, and proline residues are arranged between hydrophilic, cross-linking domains rich in lysine and alanine or proline residues. Tropoelastin is the most elastic and distensible monomer protein known, extending to eight times its length (8). After unraveling fully, tropoelastin recoils back without hysteresis (8, 9). Its elasticity, although reduced upon cross-linking and assembly, is propagated within elastic fibers that can withstand a lifetime of extension and relaxation cycles (6). Beyond its remarkable mechanical properties, tropoelastin and elastin degradation products interact with cell-surface receptors and stimulate cell responses, including chemotaxis, adhesion, and proliferation (10). Elastin's biocompatibility, coupled with its responsiveness to external triggers (11), has inspired its use in biomedical applications, such as drug delivery (12–14) and tissue engineering (15). On the other hand, mutations in the elastin gene

Significance

Tropoelastin, the precursor molecule to elastic fibers, is a large, flexible elastic protein whose structure has been the subject of investigation and debate over several decades. Here, we present the fully atomistic structure of human tropoelastin, based on molecular dynamics simulations, and validate it with experiments. We explore the functional role of two key residues by inserting alanine substitutions and explain conformational changes and variations in hierarchical assembly. We also predict essential dynamics of the molecule by building elastin network models, to explain experimentally observed differences in assembly. Finally, we study the structural and dynamic molecular changes associated with the acquired cutis laxa disorder. The approach developed here is applicable for studying structure and function of other highly disordered proteins.

Author contributions: A.T., G.C.Y., C.B., A.S.W., and M.J.B. designed research; A.T., G.C.Y., C.B., A.S.W., and M.J.B. performed research; A.T. and A.S.W. contributed new reagents/analytic tools; A.T., G.C.Y., C.B., A.S.W., and M.J.B. analyzed data; and A.T., G.C.Y., C.B., A.S.W., and M.J.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: mbuehler@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1801205115/-/DCSupplemental.

Published online June 26, 2018.

are associated with debilitating diseases such as cutis laxa and supravalvular aortic stenosis (16).

The structure of tropoelastin has been the subject of decades of research, constrained by the difficulty of isolating the monomer due to extensive cross-links (7). As a result, many studies focused instead on recombinant tropoelastin fragments, soluble elastin, and elastin-like peptides (17–23). The expression of recombinant human full-length tropoelastin allowed probing of its complete molecular structure (8, 24). Recent small-angle X-ray scattering studies determined the global molecular shape of tropoelastin to be an asymmetric molecule, with an extended, coil-like, N-terminal region, supported by a hinge that props a bridge linking the N-terminal region to the cell-interactive C-terminal foot region (8). Experiments on supramolecular tropoelastin-based structures and dynamic molecular models revealed a requirement for a subtle balance between structure and flexibility within the molecule (25).

In this article we present the fully atomistic molecular structure of human tropoelastin, based on computationally intensive MD simulations. We validated the structure against available experimental molecular data. We then considered two alanine-substitution mutations, R515A and D72A, within key regions of the molecule. Arginine (R) 515 is a highly conserved residue that stabilizes the tropoelastin bridge (26). Aspartate (D) 72 is the sole negatively charged residue in the first half of tropoelastin and functions to stabilize the N-terminal region (27). Mutations at either site are associated with global conformational changes and impaired functional assembly. We considered the molecular geometry, secondary structure, and the location and exposure of hydrophobic domains to explain observed differences in structure and hierarchical assembly. We created an elastic network model based on the fully atomistic structure of the molecule to predict the essential dynamics and explain differences in assembly. We then introduced a mutation associated with acquired cutis laxa to study the associated structural and dynamic changes.

Results

Fully Atomistic Model of Human Tropoelastin. The fully atomistic model of tropoelastin was determined through a series of replica exchange MD simulation stages (28) in implicit and explicit solvent (Fig. 1*A*). The resulting ensemble of structures was clustered by structural similarity based on rms deviation. A representative structure of tropoelastin was identified by extracting the lowest energy structure from the most populated cluster. This final structure, representative of the most frequently occurring state within the ensemble, was compared with predictions from small-angle X-ray scattering (SAXS) based on work by Baldock et al. (8) (Fig. 1 *B* and *C*). Both models displayed common features, characterized by a long molecular torso, with two diverging branches toward the molecule's bottom (the foot region) linked by a bridge region, and with the N and C termini located in their expected positions. Molecular dimensions correlated closely between the atomistic model and the SAXS-derived structure (Fig. 1 *D* and *E*). No significant differences were seen in the narrowest and widest horizontal lengths of the molecule, with a preserved general envelope of the molecule. We found 6–9% α helix, 5% β sheet and strand, 2% 3_{10} and π helix, and 84–87% coil and turn structure in simulation using the structure identification (STRIDE) algorithm. Helices were counted for a minimum length of 7–10 consecutive residues in helical formation, with a helix longevity >1 ns, based on the observation that helical propensity is closely dependent on sequence length (29). Calculations were performed from a 50-ns production run. Our findings are in close agreement with experimental data that found 35–50% β sheet (including β turns), a small (10%) contribution from α helices, and 40–55% unassigned to regular helices or β structures (6). A detailed domain-by-domain comparison with experimental data is included in *SI Appendix*.

Tropoelastin's conserved global shape can be coarsely divided into three main regions of interest: the elastic region extending down from the N terminus, the foot region containing the cell-interactive C terminus, and the bridge linking them, which has been shown to play a key role in modulating fiber assembly (26). To demonstrate how the fully atomistic model may be used to assess the functionality of different regions within the molecule, and to study disease etiology, we selected three residues of interest: arginine 515 (R515), aspartate 72 (D72), and glycine 685 (G685). R515 is a highly conserved residue in mammalian tropoelastin that stabilizes the bridge region (26). D72 is the sole negatively charged residue in the first half of tropoelastin and functions to secure the N-terminal region (27). Finally, a mutation of G685 to aspartate (G685D) is associated with acquired cutis laxa, a disease characterized by pendulous inelastic skin and inflammation in the vasculature (30), as well as early-onset chronic obstructive pulmonary disease (31).

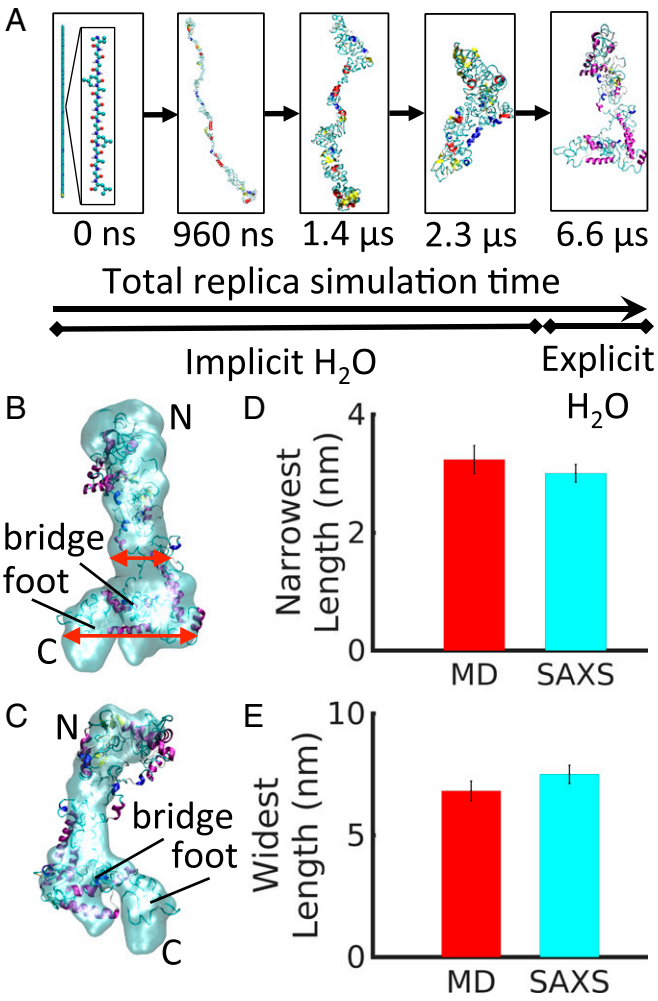

Fig. 1. (*A*) Protein folding from a straight extended chain, through 2.3 μs of total replica time, in implicit solvent, followed by 4.3 μs in explicit solvent. (*B*) Front view and (*C*) back view of the representative structure of the tropoelastin molecule. The atomistic model is shown in cartoon representation, with the α helix in purple, the 3_{10} helix in blue, the π helix in red, the extended β structure in yellow, the β bridge in tan, the turn in cyan, and the coil in white. The SAXS model is shown as a light-blue envelope. N and C indicate N and C termini, respectively. Arrows show the narrowest and widest lengths. (*D*) The narrowest length and (*E*) the widest horizontal length of tropoelastin in the MD model and the averaged SAXS model. Error bars are based on the average of the last 50 ns of the simulation. A 5% error bar is included for comparison in the SAXS model.

R515A Mutation. To evaluate the roles of R515 and D72, we artificially mutated these residues into alanine (R515A and D72A, respectively) to mute their effects. Alanine substitutions

were inserted into the WT molecule and the structures were equilibrated to convergence. The R515A mutant retained the overall shape of the WT molecule, with the exception of a noticeable shift in the foot, which displayed reduced extension (Fig. 2 *A* and *B*). The shift of the foot in the mutant may be explained by a change in secondary structure that propagates into domains 26–36 below the junction where residue 515 is located. An increased α helix and β bridge content and decreased coil content between WT and R515A were observed (Fig. 2*C*). A transformation to a more-ordered secondary structure suggests that local secondary structure changes in the bridge region may influence the flexibility of the foot, resulting in a more compact shape. The analogous dislocation of the mutant foot is visible in the SAXS structures, which show confinement of this region to the body of the R515A molecule (Fig. 2 *D* and *E*). Experimentally, the exposure of the R515A C terminus, detected by an antibody specific for this region, was lower compared with WT (26), consistent with the structural change observed in the molecular models (Fig. 2*F*).

The R515A mutation is also associated with a lower propensity for tropoelastin self-assembly (26). While both WT and R515A displayed temperature-dependent coacervation, WT fully coacervated at 35 °C, compared with 40 °C for R515A (Fig. 2*G*). To understand these effects at the molecular scale, we measured the solvent accessible surface area (SASA) of hydrophobic domains in WT and R515A (Fig. 2*H*), since tropoelastin coacervation is primarily driven by hydrophobic interactions. SASA is an indication of exposure and interactive potential with the surrounding environment. R515A exhibited significantly less exposed hydrophobic SASA than WT (Fig. 2*H*) as a consequence of structural shifts in the molecule arising from the residue substitution, which may help explain the difference in coacervation patterns observed in experiment (Fig. 2*I*).

Collective molecular motion may further direct self-assembly. To examine this, we built elastic network models based on fully atomistic structures, with α-carbons interconnected by uniform springs, and calculated normal modes of motion. Low-frequency modes represent the most concerted motions, where large structural subunits move with respect to other subunits. These modes are most energetically favorable on the multidimensional energy landscape and correspond to intrinsically accessible motions within the molecule (32).

The current model supported our earlier description of WT dynamics (25), represented by a characteristic twist in the N-terminal region and a scissors-like motion about the bridge, between the legs of the molecule (Fig. 2*J* and Movie S1). The consistency between the current model and our earlier dynamic model based on SAXS structures is instructive, as it shows the strong dependency of normal modes on the global molecular shape. Furthermore, the model is consistent with the observation that dynamics driving coacervation may simultaneously preserve intrinsic disorder, as has been suggested by recent NMR measurements (33). The R515A mutant exhibited altered collective dynamics (Fig. 2*K* and Movie S2), displaying a bend across the axis perpendicular to the body of the molecule. The WT is more dynamic on the whole, showing higher rms fluctuation per residue (Fig. 2*L*). We propose that the elevated level of mobility in the WT predisposes it more strongly to intermolecular assembly.

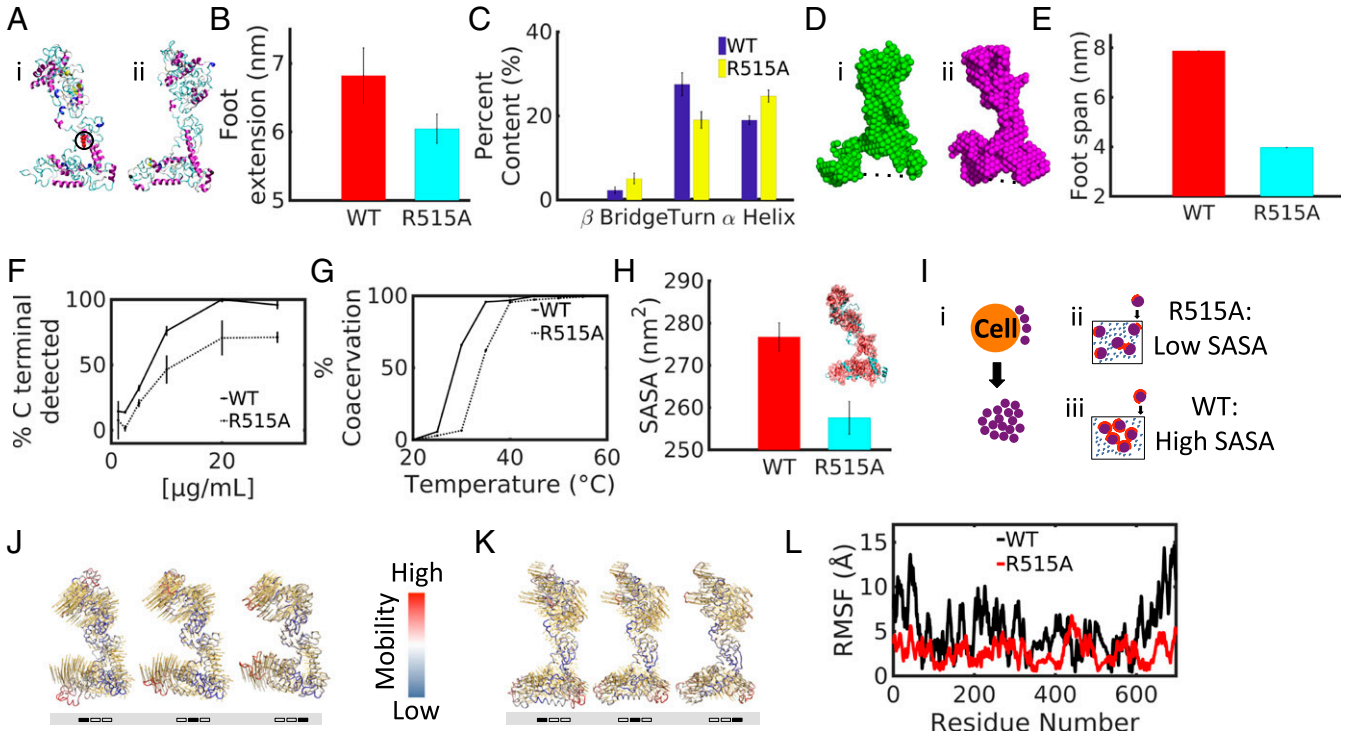

Fig. 2. (*A*, *i*) WT with R515 in red, circled; (*ii*) R515A mutant. (*B*) Extension of foot region for WT and R515A. (*C*) Secondary structure composition of WT and R515A in domains 26–36 shows structural changes in β bridge, turn, and α helix content downstream of the mutation. (*D*, *i*) WT and (*ii*) R515A representative structures from SAXS based on ref. 26. (*E*) Foot span indicated by dotted lines in the SAXS structures in *D*. (*F*) Percentage of C termini detected in WT and R515A for different concentrations. (*G*) WT and R515A coacervation at different temperatures. (*H*) SASA of hydrophobic domains in WT and R515A. (*Inset*) WT with hydrophobic SASA shown as a red envelope around the cyan molecular structure in cartoon representation. (*I*, *i*) Model of tropoelastin molecules in purple secreted from the cell. (*ii*) Model of R515A self-assembly into sparser clusters, given a lower percentage of hydrophobic SASA. Molecules are indicated by purple circles, where red represents hydrophobic SASA and dark blue represents water molecules surrounding protein. (*iii*) Model of WT self-assembly into denser clusters, with a higher percentage of hydrophobic SASA. (*J*) WT and (*K*) R515A dynamic snapshots based on the linear combination of the first six lowest-frequency modes. Eigenvectors per residue are shown as yellow arrows. (*L*) RMSF of residues in WT and R515A. The average was taken over last 50 ns for *B*, *C*, and *H*. Data were modified from ref. 26 for *F* and *G*. RMSF, root-mean-square fluctuation.

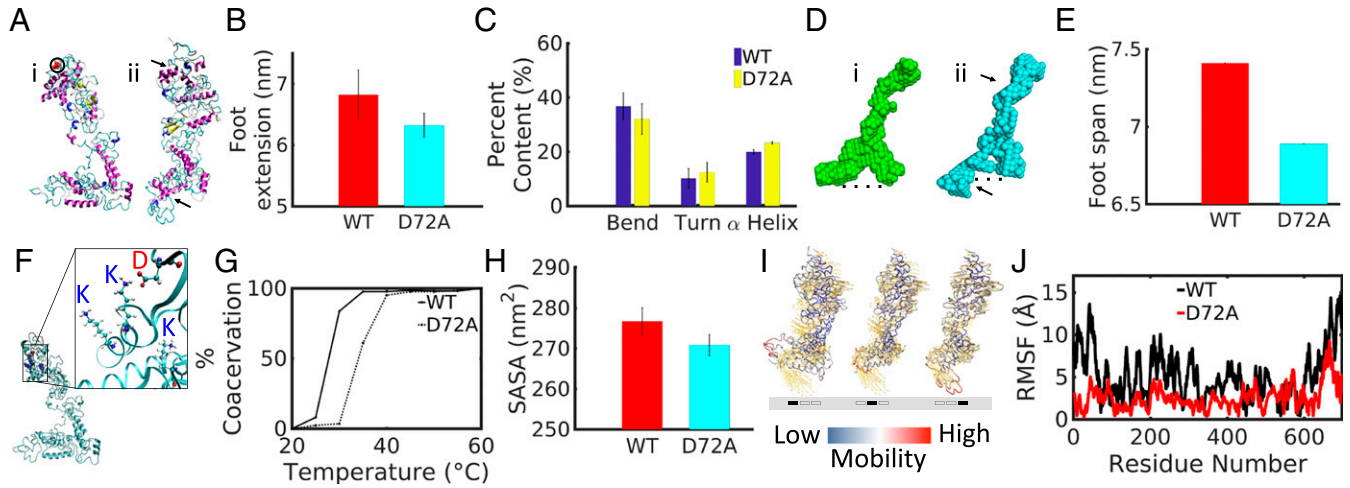

Fig. 3. (*A*, *i*) WT with D72 in red, circled; (*ii*) D72A mutant. Arrows indicate domain shifts. (*B*) Extension of foot region for WT and D72A. (*C*) Secondary structure composition of WT and D72A in domain 6 shows structural changes in bend, turn, and α helix content within the domain of the mutation. (*D*) WT (*i*) and D72A (*ii*) representative structures from SAXS based on ref. 27. (*E*) Foot span of SAXS structures in *D*, as indicated by dotted lines. (*F*) WT with lysines 78, 81, and 114 in blue and aspartate 72 in red. (*Inset*) Close-up of lysine (K) and aspartate (D) residues that may be involved in salt bridges to stabilize the N-terminal region. (*G*) WT and D72A coacervation at different temperatures. (*H*) SASA of hydrophobic domains in WT and D72A. (*I*) D72A dynamic snapshots based on the linear combination of the first six lowest-frequency modes. Eigenvectors per residue are shown as yellow arrows. (*J*) RMSF of residues in WT and D72A. The average was taken over 50 ns for *B*, *C*, and *H*. Data modified from ref. 27 for *G*. RMSF, root-mean-square fluctuation.

D72A Mutation. We performed a similar analysis on the D72A mutant. D72A displayed a downward shift in the C-terminal region, leading to a restricted foot. This conformational change is driven by a cascading shift in local secondary structure starting from domain 6, and a twist in the upper N-terminal region (Fig. 3 *A*–*C*). Analogous shifts were observed in SAXS structures (Fig. 3 *D* and *E*). We predict the role of D72 in stabilizing the N-terminal region of the molecule, as it is capable of forming salt bridges with adjacent lysine (K) residues. In particular, D72 can interact with K78 and K81 in domain 6 and, less likely, with K114 in domain 8 (Fig. 3*F*). A salt bridge develops between D72 and K78 or K81 21% and 30% of the time, respectively, and between D72A and K114 2% of the time. Overall, D72 is involved in a salt bridge >50% of the time, which would contribute significantly to the stability of the N-terminal region. The absence of a prominent salt bridge in D72A leaves the region to rotate freely into the position at which it appears in the mutant. In D72A, the polyalanine region in domain 6 increases from four to six consecutive alanines, resulting in an increase in α helix and hydrogen-bonded turn content and a reduction in bend structure in domain 6. Increased helical content may lock into place the more flexible N-terminal region, which is no longer held in place through a salt bridge. The importance of D72 in stabilizing the molecule is further emphasized in the observation that the conformational changes associated with D72A propagate down the body of the molecule, inducing a prominent foot displacement. The position of domains 15 and 16, spanned by a long helical region below domain 6, seems to be determined by steric exclusion. Based on this observation, a shift in domain 6 may trigger global shape changes by providing space for other regions to occupy. While WT coacervated fully at 35 °C, D72A coacervated fully only at 40 °C (Fig. 3*G*). Experimental findings are supported by SASA measurements of hydrophobic residues in the WT and D72A models (Fig. 3*H*). WT exhibited more exposed hydrophobic SASA than D72A, which may contribute to the difference in coacervation patterns. D72A also produced altered dynamics, displaying a pronounced bending motion (Fig. 3*I* and Movie S3) and reduced molecular fluctuation (Fig. 3*J*), consistent with experimental observations of reduced coacervation and impaired elastic fiber assembly compared with WT (27).

G685D Acquired Cutis Laxa Mutation. We next considered a clinically relevant mutation associated with acquired cutis laxa and early-onset chronic obstructive pulmonary disease (30, 31). We studied a tropoelastin isoform with a glycine-to-aspartate substitution in position 685 (G685D), derived from a mutation in the *ELN* gene.

We found that the general molecular shape is preserved for the mutant protein (Fig. 4*A*), although G685D has a bulkier, displaced N-terminal region and an inwardly shifted, shortened foot region (Fig. 4 *A* and *B*). G685D further stabilizes the salt bridge between D72 and K78, which is present almost 100% of the time. The shape change associated with the mutation introduces two additional salt bridges, between glutamate 345 (E345) and K356 (98% of the time) and between D685 and K647 (56% of the time) (Fig. 4*C*). The salt bridges act to elongate and tighten the molecule's structure and are accompanied by a decrease in bend structure and a corresponding increase in α helix content in domains 27–36 (Fig. 4*D*), effectively drawing the foot in. The change in structure reduces the exposed hydrophobic surface area as in R515A and D72A (Fig. 4*E*). We propose that the altered molecular shape and reduced hydrophobic SASA would result in changes in self-assembly of mutant tropoelastin, as illustrated in the schematic in Fig. 4*F*. The dynamics of the mutant molecule are also transformed, displaying a more pronounced bending motion in contrast to the scissors-twist movement of WT (Fig. 4 *G* and *H* and Movie S4). We find that G685D has comparable residue-by-residue fluctuation as WT, excluding domains 33 and 36 (the last 29 residues), where fluctuation in G685D is reduced (Fig. 4*H*).

Discussion

Fully Atomistic Structure Prediction of the Tropoelastin Molecule. The fully atomistic molecular structure of tropoelastin was successfully predicted using a series of REMD simulations, and validated against experimental data. Long-time REMD simulations were essential to effectively sample a structure of the size of the tropoelastin molecule. The need to conduct ergodic sampling of the energy landscape stemmed from the observation that the landscape for a system of the size and complexity of tropoelastin would present local energy minima and encounter energy barriers

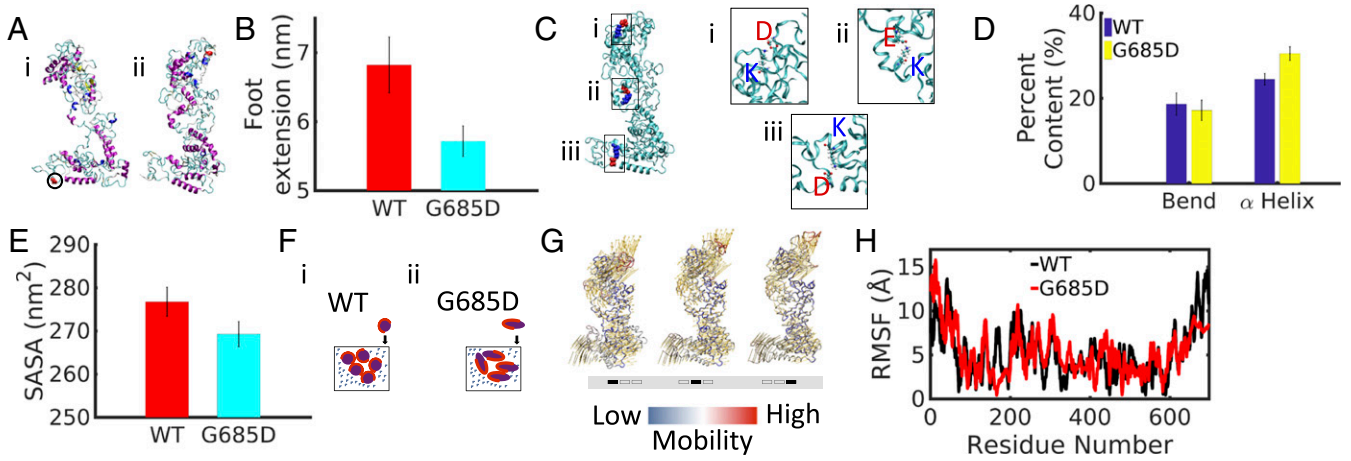

Fig. 4. (*A*, *i*) WT with G685 in red, circled; (*ii*) G685D mutant. (*B*) Extension of foot region for WT and G685D. (*C*) G685D with lysines 78, 356, and 647 in blue and aspartate 72, glutamate 345, and aspartate 685 in red. (*Inset*) Close-up of lysine (K), aspartate (D), and glutamate (E) residues that may be involved in salt bridges. (*D*) Secondary structure composition of WT and G685D in domains 27–36 shows structural changes in α helix and bend content. (*E*) SASA of hydrophobic domains in WT and G685D. (*F*, *i*) Model of WT self-assembly; (*ii*) model of G685D self-assembly (protein in purple, hydrophobic SASA in red, water molecules in blue). (*G*) G685D dynamic snapshots based on the linear combination of the first six lowest-frequency modes. Eigenvectors per residue are shown as yellow arrows. (*H*) RMSF of residues in WT and G685D. The average was taken over 50 ns for *B*, *D*, and *E*. RMSF, root-mean-square fluctuation.

that would be difficult to cross at ambient temperatures within a reasonable simulation time.

REMD resolved this problem by running a series of discrete, independent replicas in varied temperature ensembles, where coordinates of the replica structures were periodically exchanged between ensembles. High-temperature replicas provided a means to effectively cross energy barriers, while ambient-temperature replicas presented an ensemble from which sampling is desired. In each temperature ensemble, there was a distribution of structures with associated potential energies. Probability distributions of potential energy were normally distributed. Adjacent temperature ensembles were chosen so that there was sufficient overlap in the probability distributions, implying that a particular structure could exist in both. Metropolis-style Monte-Carlo moves were constructed whereby coordinates of adjacent replicas could exchange based on the probability of being observed in both ensembles. Improved ergodicity was achieved by ensuring the following key considerations: (*i*) sufficient overlap, so that exchanges produce valid ensembles; (*ii*) adequately high temperatures to establish exhaustive ensemble space sampling; and (*iii*) high probability of successful exchange attempts. In the first simulation stage, the initial input structure was an extended linear chain built based on the amino acid sequence of the tropoelastin molecule. The protein folded progressively from a linearly extended initial conformation to a compact final structure.

The representative structure from the initial phase was used as an input for the second stage of REMD simulation in explicit solvent. The role of water has been widely discussed for its contributions to the unique mechanical properties and inverse temperature transition capabilities of the elastin protein (19, 34–36). Therefore, using an explicit solvent model was essential for local structure refinement of the tropoelastin molecular model.

The similarity in the predicted fully atomistic model and the SAXS-derived model is striking, identifying common features, including a protruding foot region connected to a bridge (Fig. 1 *B* and *C*), corresponding molecular dimensions (Fig. 1 *D* and *E*), and matching secondary structure distributions. We note that any local inconsistency observed in the comparison between the SAXS envelope and the fully atomistic structure in Fig. 1 *B* and *C* should be expected, as the molecular structure embodies local and global dynamic properties responsible for tropoelastin's function and assembly properties (25). Additional representative structures are shown in *SI Appendix*, Fig. S15, to emphasize this point.

We observed that domain 2 and domain 36, the first and last domains in the mature form of human tropoelastin, maintained their expected locations at the head of the N-terminal region and at the tip of the C-terminal foot region, respectively (*SI Appendix*, Fig. S1). This positioning was consistent with earlier studies that estimated these locations through structural comparisons of full-length and truncated tropoelastin constructs using small-angle X-ray and neutron scattering experiments (8). Domain 36, corresponding to the last 14 residues in the sequence, appeared to be the most mobile of all tropoelastin domains based on rms fluctuation (Figs. 2*L*, 3*J*, and 4*H*), consistent with its role in interacting with cells (37, 38) and other elastic fiber proteins (39).

Mutation Studies. The model presents a lens through which to study local structural impact and the influence on molecular aggregation resulting from mutations in the protein structure. We considered three mutations in this study: R515A and D72A, corresponding to artificial mutations of arginine and aspartate to alanine in positions 515 and 72, respectively, and G685D, a glycine to aspartate substitution, associated with a clinically relevant mutation found in patients with acquired cutis laxa and early-onset chronic obstructive pulmonary disease (30, 31).

R515A and D72A mutations resulted in a local structural perturbation that drove changes in dynamics and coacervation behavior (Figs. 2 and 3). These studies not only expose the function of important molecular regions but also provide a means of validating the model and a framework for studying other mutations. In both the atomistic models and the SAXS-derived models of R515A, the mutation yielded a compaction of the foot region, which can be explained on the basis of the absence of a long, charged side chain of arginine, and a transformation to a more compact helical structure in the lower region of the mutant molecule. This change was further supported by an observed reduction in the detection of the C terminus in antibody experiments. The D72A mutation resulted in the perturbation of stabilizing salt bridges and a cascading shift in secondary structure from the top down, driving a twist at the N-terminal end of the molecule, and a downward shift toward the molecular body at the C-terminal end.

The structural changes observed in response to both mutations affected the coacervation, or self-assembly propensity, of the mutants, in part driven by a reduction in solvent-exposed

hydrophobic surface area. Dynamic models predicted a change in the quality and magnitude of concerted molecular oscillation contributing to differences in coacervation patterns.

A similar analysis was performed for the mutation G685D (Fig. 4). The shortened foot identified for the mutant and reduced fluctuation in the C terminus of the mutant obscure the exposure of the cell-interactive C-terminal region, explaining the observed decrease in cell adhesion in the mutant (30). The introduction of additional salt bridges throughout the molecule contributes to changes in the mechanical properties of tropoelastin. We propose that salt bridges can contribute to heightened molecular stability and increased molecular stiffness, corresponding to observations of coarser skin in patients with this mutation (30). It is noted that the molecule's inherent flexibility results in significant changes throughout the entire molecular structure in response to a single residue substitution, emphasizing the interplay of order and disorder within the molecule.

We propose that the shifted foot position, altered dynamics within the foot region, an extended molecular shape supported by salt bridges, and a reduction in exposed hydrophobic surface area all play a role in hindering fiber assembly in acquired cutis laxa and chronic obstructive pulmonary disease. This is supported by experimental observation of a subtle defect in the deposition of mutated tropoelastin into elastic fibers (30). We propose that the structural characterization of this mutation is a first step to developing targeted therapies for this and related diseases.

To conclude, we have shown how accelerated-sampling MD methods are used for ab initio folding of the 698-residue tropoelastin protein, a major flexible, elastic protein. We validated the model against experimental data, used it as a template to assess the function of local structural elements via artificial mutations, and presented a structural study of a disease-associated mutation. We identified structural shifts and variations in the molecule's dynamics in the presence of mutations, which may explain differences seen in the molecular coacervation and assembly patterns of mutants, and present mechanisms to elucidate the molecular basis of disease not previously possible from experiment alone. This work provides a framework to examine the functional roles of particular tropoelastin domains, forms a foundation for understanding the assembly of tropoelastin into elastic fibers, and can be extended to study other disease mutations. More generally, our approach would be suitable for probing the structure and function of other highly disordered proteins.

Materials and Methods

The tropoelastin structure is computed from the primary protein structure of mature, WT human tropoelastin corresponding to residues 27–724 of GenBank entry AAC98394. The WT and mutant structures are simulated through a series of replica exchange MD simulations. Structure dynamics are determined through elastic network models and normal mode analysis. Further details can be found in *SI Appendix*.

ACKNOWLEDGMENTS. Support from NIH Grant U01 HS 4976 and Office of Naval Research Grant N00014-16-1-2333 is acknowledged. This work utilized the Extreme Science and Engineering Discovery Environment (XSEDE) (40), which is supported by National Science Foundation Grant ACI-1053575. XSEDE resources Stampede 2 and Ranch at the Texas Advanced Computing Center and Comet at the San Diego Supercomputing Center through allocations TG-MSS090007 and TG-MCB180008 were used. A.S.W. acknowledges support from the Australian Research Council. The Wellcome Trust Centre for Cell-Matrix Research, University of Manchester, is supported by core funding from the Wellcome Trust (Grant 088785/Z/09/Z). C.B. is supported by Biotechnology and Biological Sciences Research Council funding (Reference: BB/N015398/1).

1. Dunker AK, Obradovic Z (2001) The protein trinity–linking function and disorder. *Nat Biotechnol* 19:805–806.
2. Babu MM, Kriwacki RW, Pappu RV (2012) Structural biology: Versatility from protein disorder. *Science* 337:1460–1461.
3. Edwards YJK, Lobley AE, Pentony MM, Jones DT (2009) Insights into the regulation of intrinsically disordered proteins in the human proteome by analyzing sequence and gene expression data. *Genome Biol* 10:R50.
4. Freddolino PL, Harrison CB, Liu Y, Schulten K (2010) Challenges in protein folding simulations: Timescale, representation, and analysis. *Nat Phys* 6:751–758.
5. Chebaro Y, Ballard AJ, Chakraborty D, Wales DJ (2015) Intrinsically disordered energy landscapes. *Sci Rep* 5:10386.
6. Muiznieks LD, Weiss AS, Keeley FW (2010) Structural disorder and dynamics of elastin. *Biochem Cell Biol* 88:239–250.
7. Wise SG, Weiss AS (2009) Tropoelastin. *Int J Biochem Cell Biol* 41:494–497.
8. Baldock C, et al. (2011) Shape of tropoelastin, the highly extensible protein that controls human tissue elasticity. *Proc Natl Acad Sci USA* 108:4322–4327.
9. Mithieux SM, Wise SG, Weiss AS (2013) Tropoelastin: A multifaceted naturally smart material. *Adv Drug Deliv Rev* 65:421–428.
10. Rodgers UR, Weiss AS (2005) Cellular interactions with elastin. *Pathol Biol (Paris)* 53: 390–398.
11. Tarakanova A, Huang W, Weiss AS, Kaplan DL, Buehler MJ (2017) Computational smart polymer design based on elastin protein mutability. *Biomaterials* 127:49–60.
12. MacEwan SR, Chilkoti A (2014) Applications of elastin-like polypeptides in drug delivery. *J Control Release* 190:314–330.
13. McDaniel JR, Callahan DJ, Chilkoti A (2010) Drug delivery to solid tumors by elastin-like polypeptides. *Adv Drug Deliv Rev* 62:1456–1467.
14. Herrero-Vanrell R, et al. (2005) Self-assembled particles of an elastin-like polymer as vehicles for controlled drug release. *J Control Release* 102:113–122.
15. Nettles DL, Chilkoti A, Setton LA (2010) Applications of elastin-like polypeptides in tissue engineering. *Adv Drug Deliv Rev* 62:1479–1485.
16. Baldwin AK, Simpson A, Steer R, Cain SA, Kielty CM (2013) Elastic fibres in health and disease. *Expert Rev Mol Med* 15:e8.
17. Daamen WF, Hafmans T, Veerkamp JH, Van Kuppevelt TH (2001) Comparison of five procedures for the purification of insoluble elastin. *Biomaterials* 22:1997–2005.
18. Li B, Alonso DOV, Bennion BJ, Daggett V (2001) Hydrophobic hydration is an important source of elasticity in elastin-based biopolymers. *J Am Chem Soc* 123:11991–11998.
19. Li B, Alonso DOV, Daggett V (2001) The molecular basis for the inverse temperature transition of elastin. *J Mol Biol* 305:581–592.
20. Bochicchio B, Aït-Ali A, Tamburro AM, Alix AJP (2004) Spectroscopic evidence revealing polyproline II structure in hydrophobic, putatively elastomeric sequences encoded by specific exons of human tropoelastin. *Biopolymers* 73:484–493.
21. Tamburro AM, Bochicchio B, Pepe A (2005) The dissection of human tropoelastin: From the molecular structure to the self-assembly to the elasticity mechanism. *Pathol Biol (Paris)* 53:383–389.
22. Tamburro AM, Bochicchio B, Pepe A (2003) Dissection of human tropoelastin: Exon-by-exon chemical synthesis and related conformational studies. *Biochemistry* 42: 13347–13362.
23. Roberts S, Dzuricky M, Chilkoti A (2015) Elastin-like polypeptides as models of intrinsically disordered proteins. *FEBS Lett* 589:2477–2486.
24. Martin SL, Vrhovski B, Weiss AS (1995) Total synthesis and expression in Escherichia coli of a gene encoding human tropoelastin. *Gene* 154:159–166.
25. Yeo GC, et al. (2016) Subtle balance of tropoelastin molecular shape and flexibility regulates dynamics and hierarchical assembly. *Sci Adv* 2:e1501145.
26. Yeo GC, et al. (2012) Tropoelastin bridge region positions the cell-interactive C terminus and contributes to elastic fiber assembly. *Proc Natl Acad Sci USA* 109:2878–2883.
27. Yeo GC, Baldock C, Wise SG, Weiss AS (2014) A negatively charged residue stabilizes the tropoelastin N-terminal region for elastic fiber assembly. *J Biol Chem* 289: 34815–34826.
28. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151.
29. Chin DH, Woody RW, Rohl CA, Baldwin RL (2002) Circular dichroism spectra of short, fixed-nucleus alanine helices. *Proc Natl Acad Sci USA* 99:15416–15421.
30. Hu Q, Reymond JL, Pinel N, Zabot MT, Urban Z (2006) Inflammatory destruction of elastic fibers in acquired cutis laxa is associated with missense alleles in the elastin and fibulin-5 genes. *J Invest Dermatol* 126:283–290.
31. Kelleher CM, et al. (2005) A functional mutation in the terminal exon of elastin in severe, early-onset chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 33:355–362.
32. Bahar I, Lezon TR, Bakan A, Shrivastava IH (2010) Normal mode analysis of biomolecular structures: Functional mechanisms of membrane proteins. *Chem Rev* 110:1463–1497.
33. Reichheld SE, Muiznieks LD, Keeley FW, Sharpe S (2017) Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proc Natl Acad Sci USA* 114:E4408–E4415.
34. Dandurand J, Samouillan V, Lacabanne C, Pepe A, Bochicchio B (2015) Water structure and elastin-like peptide aggregation: A differential calorimetric approach. *J Therm Anal Calorim* 120:419–426.
35. Samouillan V, André C, Dandurand J, Lacabanne C (2004) Effect of water on the molecular mobility of elastin. *Biomacromolecules* 5:958–964.
36. Sun C, Mitchell O, Huang J, Boutis GS (2011) NMR studies of localized water and protein backbone dynamics in mechanically strained elastin. *J Phys Chem B* 115:13935–13942.
37. Bax DV, Rodgers UR, Bilek MMM, Weiss AS (2009) Cell adhesion to tropoelastin is mediated via the C-terminal GRKRK motif and integrin alphaVbeta3. *J Biol Chem* 284: 28616–28623.
38. Broekelmann TJ, et al. (2005) Tropoelastin interacts with cell-surface glycosaminoglycans via its COOH-terminal domain. *J Biol Chem* 280:40939–40947.
39. Nonaka R, Sato F, Wachi H (2014) Domain 36 of tropoelastin in elastic fiber formation. *Biol Pharm Bull* 37:698–702.
40. Towns J, et al. (2014) XSEDE: Accelerating scientific discovery. *Comput Sci Eng* 16: 62–74.