Macromolecular Bioscience
www.mbs-journal.de

Tropoelastin is a Flexible Molecule that Retains its Canonical Shape

*Anna Tarakanova, Giselle C. Yeo, Clair Baldock, Anthony S. Weiss, and Markus J. Buehler**

Tropoelastin is the dominant building block of elastic fibers, which form a major component of the extracellular matrix, providing structural support to tissues and imbuing them with elasticity and resilience. Recently, the atomistic structure of human tropoelastin is described, obtained through accelerated sampling via replica exchange molecular dynamics simulations. Here, principal component analysis is used to consider the ensemble of structures accessible to tropoelastin at body temperature (37 °C) at which tropoelastin naturally self-assembles into aggregated coacervates. These coacervates are relevant because they are an essential intermediate assembly stage, where tropoelastin molecules are then cross-linked at lysine residues and integrated into growing elastic fibers. It is found that the ensemble preserves the canonical tropoelastin structure with an extended molecular body flanked by two protruding legs, and identifies variations in specific domain positioning within this global shape. Furthermore, it is found that lysine residues show a large variation in their location on the tropoelastin molecule compared with other residues. It is hypothesized that this perturbation of the lysines increases their accessibility and enhances cross-linking. Finally, the principal component modes are extracted to describe the range of tropoelastin's conformational fluctuation to validate tropoelastin's scissor-twist motion that was predicted earlier.

1. Introduction

Tropoelastin is the molecular –precursor to elastin and is a dominant component of elastic fibers, key structural elements in elastic tissues including blood vessels, skin, and lungs.[1,2] The mature form of the secreted human tropoelastin monomer is 60 kDa[3] and is composed of an alternating pattern of repetitive, hydrophobic glycine-, valine-, and proline-rich domains spread out between lysine-containing cross-linking domains.[4] Traditionally described as a largely disordered molecule, recent studies suggest that tropoelastin is a flexible molecule with a distinct nanostructure[5–7] that determines its structural and cell-adhesive properties.

Tropoelastin's distinct nanostructure determines its propensity to assemble into higher-order structures (**Figure 1**) through elastogenesis.[8] Elastogenesis begins with self-assembly of the ≈15 nm monomers, a process called coacervation.[9] Self-assembly, naturally occurring at physiological pH and salt conditions at 37 °C, results in spherical structures from 200 nm to 6 µm in diameter.[9–11] Spherule formation through coacervation begins at the cell surface by surface-tethered tropoelastin molecules, until released for integration into the growing elastic fiber. Tropoelastin aggregates adhere to the cell surface through GAG- and integrin–mediated interactions at specific interaction sites.[12–18] Spherules are deposited on a microfibrillar scaffold where they are stabilized

Dr. A. Tarakanova, Prof. M. J. Buehler
Laboratory for Atomistic and Molecular Mechanics
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
02139 Cambridge, MA, USA
E-mail: mbuehler@MIT.EDU
Dr. G. C. Yeo, Prof. A. S. Weiss
School of Life and Environmental Sciences
The University of Sydney
2006 Sydney, NSW, Australia
Dr. G. C. Yeo, Prof. A. S. Weiss
Charles Perkins Centre
The University of Sydney
2006 Sydney, NSW, Australia
The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/mabi.201800250.
Prof. C. Baldock
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
M13 9PL Manchester, UK
Prof. A. S. Weiss
Bosch Institute
The University of Sydney
2006 Sydney, NSW, Australia

DOI: 10.1002/mabi.201800250

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

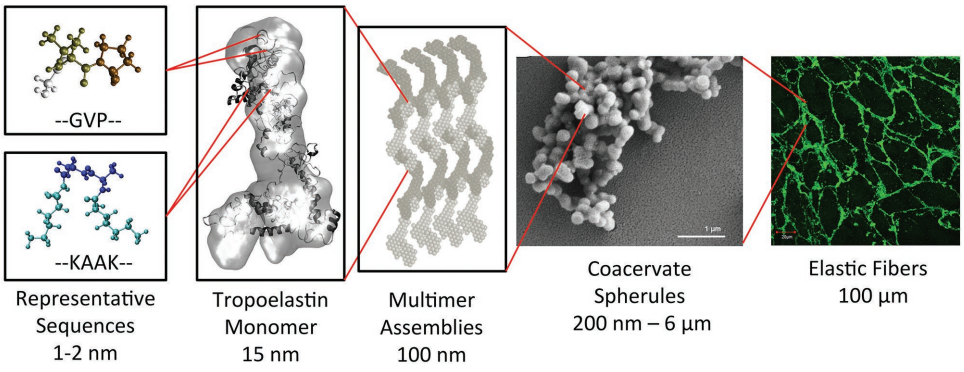

Figure 1. Overview of the hierarchical structure of elastic fibers. Tropoelastin is composed of alternating GVP-rich hydrophobic domains and hydrophilic K-containing cross-linking domains. The tropoelastin monomer represents the dominant building block for elastic fibers (cartoon representation of molecular structure is embedded in a SAXS-derived outline, based on ref. [7]). Reproduced with permission.[7] Copyright 2018, National Academy of Sciences. Tropoelastin interacts to form multimer assemblies (image based on SAXS-derived structures from ref. [5]). Reproduced with permission.[5] Copyright 2011, National Academy of Sciences. Tropoelastin coacervates into spherical aggregate structures. Reproduced with permission.[10] Copyright 2010, Elsevier. Spherules assemble onto a microfibrillar scaffold to form an elastic fiber.

by cross-links facilitated by lysyl oxidase[13,19–21] and eventually integrated into a mature elastic fiber. Elastic fibers confer the strength and elasticity needed for repetitive mechanical deformation of elastic tissues over a lifetime.

The structure of elastin and associated molecular mechanisms linked to elastin's mechanical properties and biological roles have been subject to decades of research. From the earliest random polymer network model of elastin,[22,23] investigation into elastin's structure has revealed preferential secondary structure formation and function.[24–32] A number of computational models have made strides in elucidating the molecular nature of elastin, although the focus of these works has been on shorter elastin-like peptides or assemblies thereof.[33–41] For molecules with a high degree of disorder, we and others have shown that enhanced sampling methods such as replica exchange molecular dynamics (REMD)[42] are effective in predicting molecular structure.[7,43–47] Following on from our recent work to develop a fully atomistic model of the tropoelastin monomer,[7] here we consider the structural ensemble accessible to the molecule at body temperature, that is, 37 °C. Characterization of the structural ensemble is key for capturing the specific flexibility of the structure, and provides a means to define bounds for the spread of conformations accessible to a dynamic molecule. Various methods of ensemble analysis exist for different available structural data sets, in particular in application to intrinsically disordered proteins (IDPs).[46,48–52]

A useful approach used here for capturing molecular fluctuation that describes multiple states accessible to the molecule is principal component analysis.[53] This method transforms the atomic coordinates from the Cartesian system to a new system of collective coordinates, identifying the dominant direction of structural changes within the set. Within the new coordinate system, the largest variance in the data set is along the first principal component (PC1) axis, and the second largest variance is along the second principal component (PC2), and so forth. Thus, a structural spread can be identified by considering the distribution of structures across dominant principal components.

Our simulations capture intrinsic conformational characteristics of tropoelastin, suggesting a bias toward tropoelastin's canonical nanostructure. We describe the conformational heterogeneity of tropoelastin within the confines of its broader nanostructure that determines the molecule's wide range of functions. Our results are discussed in the context of previous studies to rationalize functional implications of the conformational ensemble of the molecule.

2. Results and Discussion

We analyze the ensemble of 1000 tropoelastin structures at 37 °C derived from replica exchange molecular dynamics simulation as described previously.[7] The structures considered correspond to the last 2 ns of REMD simulation in explicit water. To characterize the structural variability of tropoelastin (also referred to the *motions* and *dynamics* of the molecule throughout the text, not to be confused with time-dependent dynamics not accessible to REMD), we perform principal component analysis on this set of structures. The variance for the first 20 principal component modes is shown in **Figure 2**a. We note that the first two principal component modes have variances of 23% and 19%, respectively, accounting together for 42% of the variation in the structural ensemble. The variance of higher modes drops considerably after the second mode, indicating a lower contribution to structural variation. As we will show in the forthcoming discussion, higher modes do not affect the relative flexibility of tropoelastin domains.

While the variance of the principal components allows us to identify dominant directional shifts in the distribution of all considered structures, we perform an independent clustering analysis to extract and analyze representative structures within the ensemble. We use k-means clustering, a method that employs a fixed cluster radius based on Cartesian coordinate root mean square deviation (RMSD) between structures, where RMSD is set to 5 Å. There are a total of 39 clusters identified through this approach. A histogram of clusters, ranked from most populated (Cluster 1) to least populated (Cluster 39) is shown in Figure 2b. The lowest energy structure is extracted from each of the 39 clusters. We next consider the distribution

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

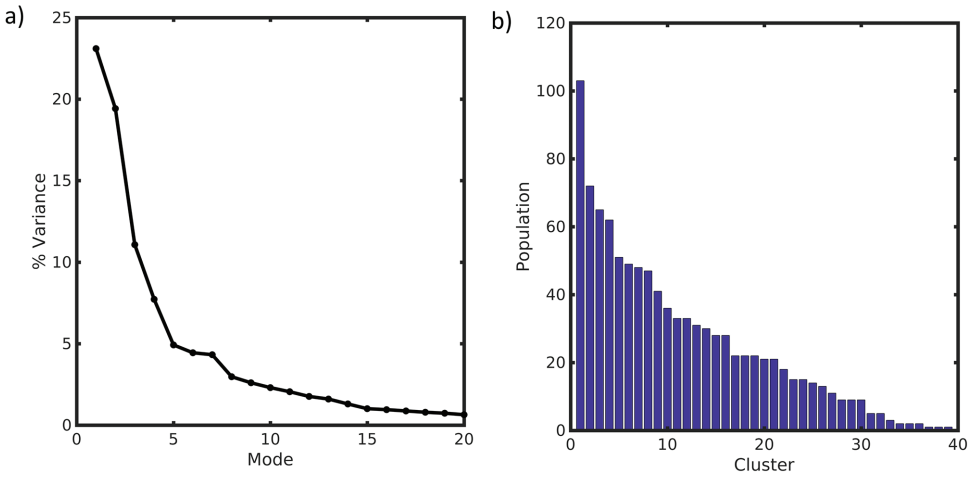

Figure 2. a) Percentage variance for the first 20 modes from principal component analysis. Variance corresponds to the fractional contribution of each mode to structural variation within the ensemble. b) Population of structures in 39 clusters through k-means clustering, with an RMSD of 5 Å, where clusters are ranked from most (1) to least (39) populous.

of the structures within the subspace spanned by the first two principal components (**Figure 3**a) as a representative spread of the observed structures of tropoelastin at 37 °C. All 1000 structures are projected onto PC1-PC2 space. Thirteen representative structures from k-means clustering analysis are labeled on the principal component map (Figure 3a) according to their cluster affiliation. The specific location of structures 1–10, 17, 20, and 24 is shown within the subspace. Representative structures affiliated with the most populous clusters (1–10) tend to reside in the densest regions of the map. Additional structures 17, 20, and 24 are chosen as a means to consider less frequently sampled, yet still accessible states of tropoelastin. The representative structure from the most populous cluster (structure 1) is considered as a reference. Six structures (2, 5, 8, 17, 20, 24) are considered in more detail to analyze the range of accessible conformations available to the molecule (Figure 3b–g). These six structures are chosen based on their location on the principal component map, as representative of greatest structural variability found in tropoelastin at 37 °C. They are superimposed on structure 1 by considering the minimal RMSD based on all atoms.

We find that despite an intrinsic flexibility displayed by the molecule, the range of motion accessible by tropoelastin preserves its canonical structure,[5,6,54] with an extended molecular body and two protruding legs (Figure 3b–g). The first four domains of the molecule (D2–5) are characteristically positioned at the head of the structure, in the N-terminal region. The location is consistent with earlier studies based on small angle X-ray and neutron scattering experiments[5] that estimate the locations of these domains by comparing full-length and truncated tropoelastin constructs. This region exhibits broad flexibility as observed by its twisted position in structures 2, 5, and 24 (Figure 3b,c,g). A more localized twist of domains 2 and 3 (D2 and D3) is seen in structure 8 (Figure 3d). The repositioning of these domains is consistent with the natural tendency toward a twisting motion in the upper region of the molecule described in earlier elastic network models and normal mode analysis of tropoelastin.[6,54]

The preserved position of the first four domains is supported by adjacent domain 6, which contains aspartate in position 72, the only negatively charged residue in the N-terminal region of the molecule, that may function to stabilize the upper region of the molecule.[54,55] We have shown previously that domain 6 partially forms a stable α-helix that is locked in place by a salt bridge between aspartate 72 and three neighboring lysine residues.[54] In fact, the location of domain 6 (D6) shows regularity across the tropoelastin ensemble. We note that in structure 17 (Figure 3e), domains 6 and 7 (D6 and D7) are shifted, yet the location of D2–5 remains consistent with structure 1. A minor shift in domain 6 is seen also in structure 20 (Figure 3f), again without significant change to the location of the first domains of the molecule.

The central region of the molecule through domain 19, characterized by a conserved arrangement of alternating hydrophobic and hydrophilic domains, is contained within the elongated torso of the molecule. The largest structural oscillation in this region is apparent in structures 5 and 8 (Figure 3c,d), where domains 10–19 and 15–19, respectively, are twisted leftward. This shift is again consistent with the predicted twist in the torso of the molecule.[6,54] More localized shifts are also characteristic of this region. Domain 19 displays flexibility away from structure 1 (in structures 2, 17, 24, Figure 3b,e,g), as do domains 15 and 17 (in structure 20, Figure 3f). The flexibility of the large central region of the molecule suggests a twofold functionality. First, the central spring-like coil region has been previously proposed to account for tropoelastin's superb elasticity, allowing the molecule to extend to eight times its resting length.[5] A wide structural flexibility available to these domains in the relaxed state is consistent with an entropic model of tropoelastin's elasticity, where extension restricts the number of conformations available to the molecule which then drives its return to the resting state.[56–58] Second, the variability in domain positioning may promote cellular interaction. Earlier work suggests that the central region of the molecule interacts with cell surface receptors including glycosaminoglycans (GAGs) and integrins, promoting cell binding and cell spreading.[17,18] In particular, domains 17 and 18 contain a glycosaminoglycan-binding

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

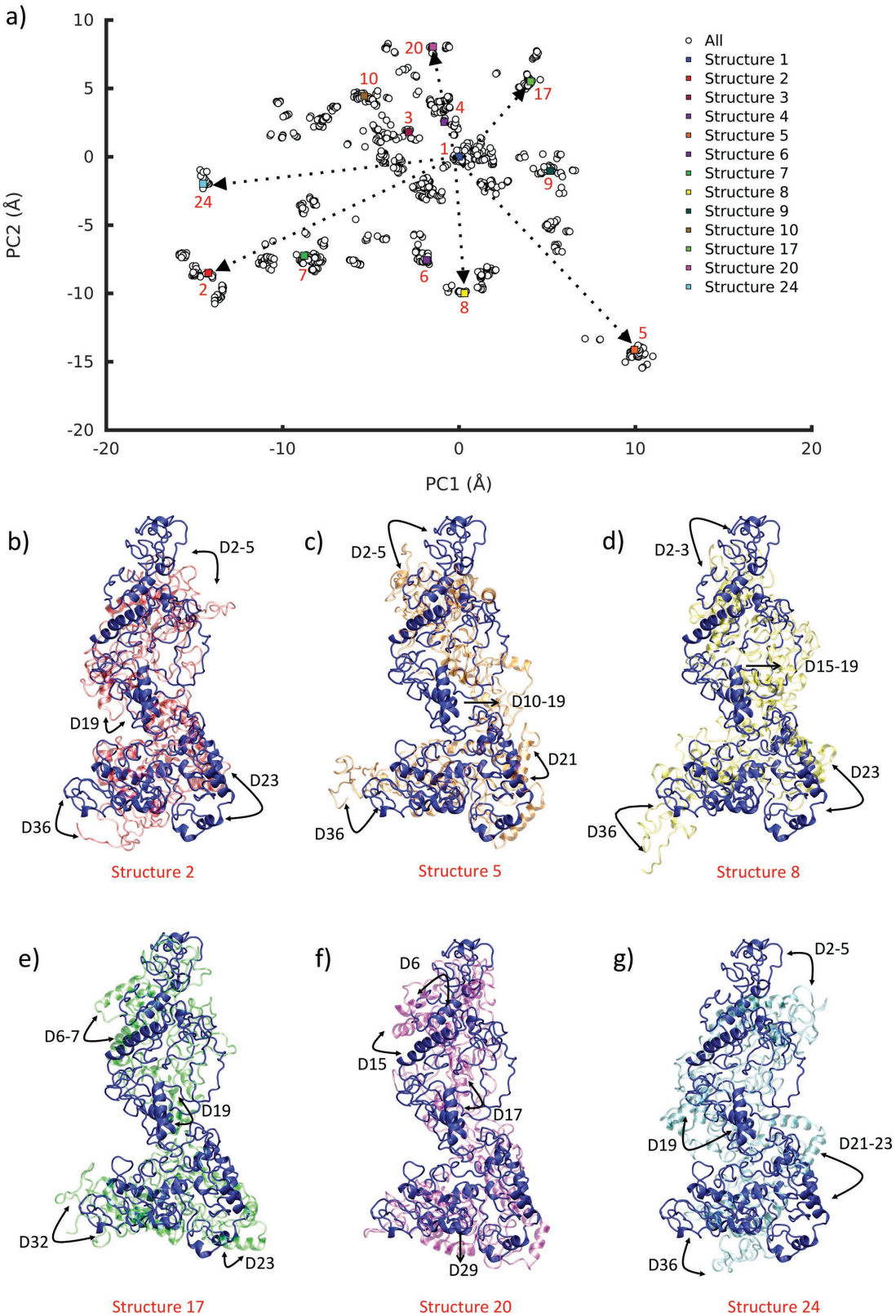

Figure 3. a) Projection of each of the 1000 structures of tropoelastin at 37 °C onto PC1 and PC2 (white circles). Structures 1–10, 17, 20, 24 represent the lowest energy structures from corresponding clusters (colored circles with red numerical labels). Structure 1 is the reference structure, and structures 2, 5, 8, 17, 20, and 24 are considered for characterizing the spread of structural variation on PC1-PC2 space, indicated by arrows. Structure 1, in blue, overlaid with b) structure 2, c) 5, d) 8, e) 17, f) 20, and g) 24. Specific domain shifts are indicated by arrows in (b–g).

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

site responsible for cell adhesion. This sequence works synergistically with integrins, where GAG-binding allows the engagement of the α_v integrin family to induce cell spreading.[18] Domains that contribute to integrin binding reside in regions adjacent to the GAG-binding sequence spanning domains 17 and 18.[18] We propose that the flexibility of the central region may therefore facilitate cell–receptor interactions by exposing relevant domains.

Continuing down the length of the molecule, a variability in the location of domains 21 and/or 23 is observed in five of the six structures considered (Figure 3b–e,g). Due to alternative splicing of tropoelastin mRNA, domain 22 is excluded from human tropoelastin so that two hydrophilic domains 21 and 23 are adjoined, forming a hinge region,[59] previously identified computationally and empirically through NMR and small-angle X-ray scattering.[60–62] Elastic network models of domains 21–23 suggest a flexible peptide capable of existing in closed and open conformations.[6] The high mobility of this region is supported by the findings in the present analysis and suggests that domains 21 and 23 contribute to tropoelastin's elasticity, as flanking domains to the central elastic region of the molecule. The scissors-like bending of domains 21–23 is consistent with the dynamic models,[6,7] suggesting an additional role in multimeric assembly via domain repositioning and conformational space sampling to facilitate cross-linking.

The C-terminus is highly conserved across mammalian species, terminating with a charged RKRK sequence.[3] It is key for tropoelastin incorporation into elastic fibers through elastogenesis, as well as GAG- and integrin-mediated cell adhesion.[12–16] In tropoelastin's structural ensemble, the C-terminal domain 36 shows significant flexibility in structures 2, 5, 8, and 24 (Figure 3b–d,g). Adjacent domains, in particular domains 32 and 29 in structures 17 and 20, respectively (Figure 3e,f) can also undergo displacement. The flexibility of the C-terminus has been predicted previously through elastic network models and normal mode analysis.[6,54] This flexibility suggests functionality by widening the conformational space available to this region for improved interactions with cell-surface receptors and elastic fiber-associated proteins.

The displacement profile of tropoelastin domains confirms the observations noted by analyzing specific representative structures. **Figure 4**a shows the displacement profile as a weighted contribution of movements along the top two PCA modes (black curve) and the top six PCA modes (red curve). The top two PCA modes account for 42% of structural variation, while the top six PCA modes account for 70% of the structural variation. The similar trend for two and six modes supports the dominance of the first two modes, validating our approach to consider the structural projection onto PC1-PC2 space. Largest displacements are found in domains 2–5 (residues 1–51), 10–19

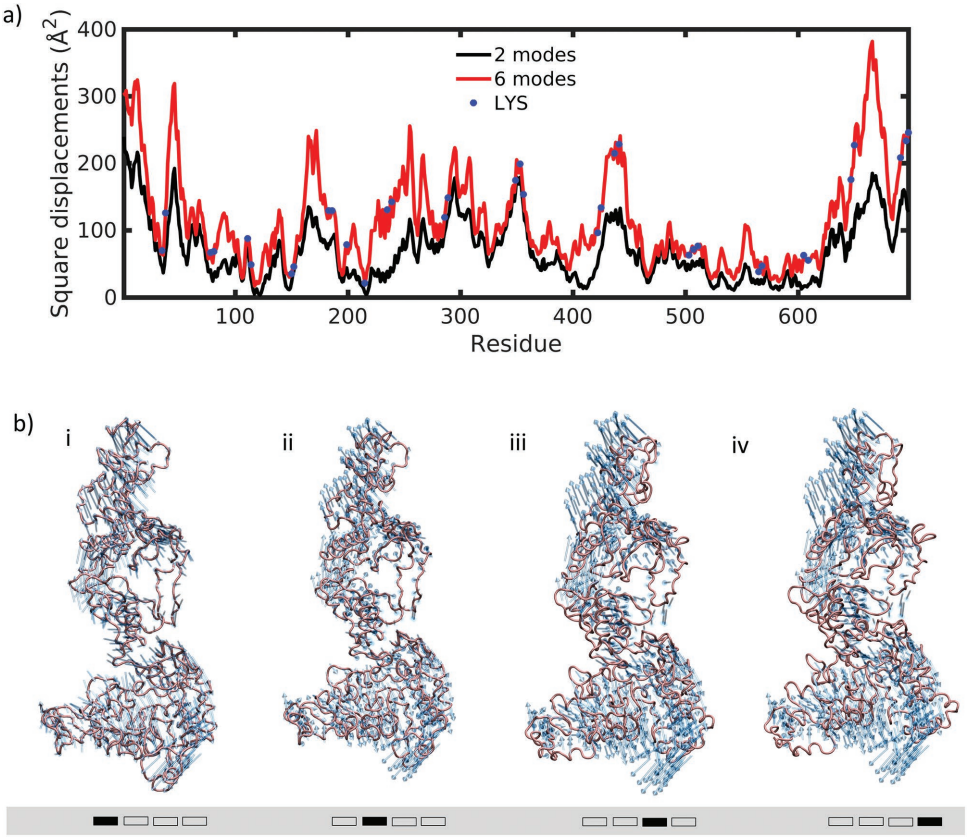

Figure 4. a) Comparison of the sum of square displacements from principal component modes 1 and 2 (black curve) and modes 1–6 (red curve) for 1000 structures. Location of lysine residues is indicated in blue. b) Snapshots (i–iv) represent the motion associated with the linear combination of the top six PCA modes, scaled by their variance. Integrated mode vectors are shown in blue, backbone of structure 1 is shown in red. The gray bar represents four time snapshots of molecular motion.

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

(residues 133–357), 21–23 (residues 413–445), and domain 36 (residues 685–697) (Figure 4a). Lysine residues, marked as blue dots on Figure 4a, are scattered throughout the tropoelastin structure and involved in inter- and intramolecular cross-links. They are almost uniformly found in highly mobile domains on the molecule. In fact, 83% of the 35 lysine residues found in tropoelastin reside at the high peaks of relative displacement among their nearest neighbor residues. The large displacements of the lysine residues suggest a mechanism for high conformational space sampling toward efficient cross-linking during elastogenesis.

Finally, large domain dynamics of the molecule are considered in Figure 4b. Snapshots (i–iv) in Figure 4b show the motion associated with the linear combination of the top six PCA modes, scaled by their variance. The molecule tends toward a twist in the upper region (N-terminal) and displays a scissors-like bend in the two protruding legs. We propose that this motion can be maintained as a result of the relatively constrained domains 6–9 (residues 52–132) and domains 24–30 (residues 446–638) (Figure 4a). The PCA-derived fluctuations correspond closely to previously predicted motions based on elastic network models and normal mode analysis,[6,7] suggesting a dynamic model to drive multimolecular interactions.

3. Conclusions

The distribution of structures accessible to human tropoelastin is characterized via principal component analysis. The distribution preserves the canonical tropoelastin structure, while oscillations in structure give rise to predicted molecular motions driving a twist at the N-terminal region and a scissors-like bending in the more C-terminal regions of the molecule. Enhanced fluctuation of specific domains, and more specifically lysine residues within the molecule, is discussed to explain regional functionality and implications for elastogenesis. The current model supports a paradigm shift for considering tropoelastin's structure. Despite a high degree of disorder, tropoelastin maintains a uniform nanostructure, and local dynamics are tied to specific regional functions of the molecule.

4. Experimental Section

The REMD-based methods for the development of the structural ensemble of tropoelastin were previously discussed.[7] The ensemble at 37 °C was characterized using the MMTSB toolset[63] and ProDy PCA analysis tools.[64]

Acknowledgements

Support from NIHU01 HS 4976 and ONRN00014-16-1-2333 is acknowledged. This work utilized the Extreme Science and Engineering Discovery Environment (XSEDE),[65] which is supported by National Science Foundation grant number 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. C.B. acknowledges grant support from the BBSRC (BB/N015398/1 and BB/R008221/1).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

disordered protein, elastic fiber, principal component analysis, structural ensemble, structural protein, tropoelastin

Received: July 3, 2018
Revised: September 3, 2018
Published online:

[1] B. Vrhovski, A. S. Weiss, *Eur. J. Biochem.* **1998**, *258*, 1.
[2] F. W. Keeley, C. M. Bellingham, K. A. Woodhouse, *Philos. Trans. Royal Soc. B* **2002**, *357*, 185.
[3] S. G. Wise, A. S. Weiss, *Int. J. Biochem. Cell Biol.* **2009**, *41*, 494.
[4] Z. Indik, H. Yeh, N. Ornsteingoldstein, U. Kucich, W. Abrams, J. C. Rosenbloom, J. Rosenbloom, *Am. J. Med. Genet.* **1989**, *34*, 81.
[5] C. Baldock, A. F. Oberhauser, L. Ma, D. Lammie, V. Siegler, S. M. Mithieux, Y. Tu, J. Y. Chow, F. Suleman, M. Malfois, S. Rogers, L. Guo, T. C. Irving, T. J. Wess, A. S. Weiss, *Proc. Natl. Acad. Sci.* **2011**, *108*, 4322.
[6] G. C. Yeo, A. Tarakanova, C. Baldock, S. G. Wise, M. J. Buehler, A. S. Weiss, *Sci. Adv.* **2016**, *2*, e1501145.
[7] A. Tarakanova, G. C. Yeo, C. Baldock, A. S. Weiss, M. J. Buehler, *Proc. Natl. Acad. Sci.* **2018**, *115*, 7338.
[8] J. E. Wagenseil, R. P. Mecham, *Birth Defects Res. C Embryo Today* **2007**, *81*, 229.
[9] A. W. Clarke, E. C. Arnspang, S. M. Mithieux, E. Korkmaz, F. Braet, A. S. Weiss, *Biochemistry* **2006**, *45*, 9989.
[10] Y. Tu, S. G. Wise, A. S. Weiss, *Micron* **2010**, *41*, 268.
[11] G. C. Yeo, F. W. Keeley, A. S. Weiss, *Adv. Colloid Interface Sci.* **2011**, *167*, 94.
[12] P. L. Brown, L. Mecham, C. Tisdale, R. P. Mecham, *Biochem. Biophys. Res. Commun.* **1992**, *186*, 549.
[13] S. G. Wise, S. M. Mithieux, M. J. Raftery, A. S. Weiss, *J. Struct. Biol.* **2005**, *149*, 273.
[14] P. B. Augsburger, T. Broekelmann, J. Rosenbloom, R. P. Mecham, *Biochem. J.* **1996**, *318*, 149.
[15] T. J. Broekelmann, B. A. Kozel, H. Ishibashi, C. C. Werneck, F. W. Keeley, L. J. Zhang, R. P. Mecham, *J. Biol. Chem.* **2005**, *280*, 40939.
[16] D. V. Bax, U. R. Rodgers, M. M. M. Bilek, A. S. Weiss, *J. Biol. Chem.* **2009**, *284*, 28616.
[17] P. Lee, D. V. Bax, M. M. Bilek, A. S. Weiss, *J. Biol. Chem.* **2014**, *289*, 1467.
[18] P. Lee, G. C. Yeo, A. S. Weiss, *FEBS J.* **2017**, *284*, 2216.
[19] F. Sato, H. Wachi, M. Ishida, R. Nonaka, S. Onoue, Z. Urban, B. C. Starcher, Y. Seyama, *J. Mol. Biol.* **2007**, *369*, 841.
[20] P. Brown-Augsburger, C. Tisdale, T. Broekelmann, C. Sloan, R. P. Mecham, *J. Biol. Chem.* **1995**, *270*, 17778.
[21] S. M. Mithieux, S. G. Wise, M. J. Raftery, B. Starcher, A. S. Weiss, *J. Struct. Biol.* **2005**, *149*, 282.
[22] C. A. J. Hoeve, P. J. Flory, *J. Am. Chem. Soc.* **1958**, *80*, 6523.
[23] C. A. J. Hoeve, P. J. Flory, *Biopolymers* **1974**, *13*, 677.
[24] T. Weis-Fogh, S. O. Anderson, *Nature* **1970**, *227*, 718.
[25] W. R. Gray, L. B. Sandberg, J. A. Foster, *Nature* **1973**, *246*, 461.
[26] J. M. Gosline, *Biopolymers* **1978**, *17*, 677.
[27] J. M. Gosline, *Adv. Exp. Med. Biol.* **1977**, *79*, 621.
[28] C. M. Venkatachalam, D. W. Urry, *Macromolecules* **1981**, *14*, 1225.

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

[29] A. Pepe, D. Guerra, B. Bochicchio, D. Quaglino, D. Gheduzzi, I. P. Ronchetti, A. M. Tamburro, *Matrix Biol.* **2005**, *24*, 96.
[30] A. M. Tamburro, B. Bochicchio, A. Pepe, *Pathol. Biol.* **2005**, *53*, 383.
[31] A. M. Tamburro, B. Bochicchio, A. Pepe, *Biochemistry* **2003**, *42*, 13347.
[32] A. M. Tamburro, V. Guantieri, L. Pandolfo, A. Scopa, *Biopolymers* **1990**, *29*, 855.
[33] B. Li, D. O. V. Alonso, B. J. Bennion, V. Daggett, *J. Am. Chem. Soc.* **2001**, *123*, 11991.
[34] B. Li, D. O. V. Alonso, V. Daggett, *J. Mol. Biol.* **2001**, *305*, 581.
[35] N. K. Li, F. G. Quiroz, C. K. Hall, A. Chilkoti, Y. G. Yingling, *Biomacromolecules* **2014**, *15*, 3522.
[36] S. Rauscher, R. Pomes, *eLife* **2017**, *6*, e26526.
[37] R. Rousseau, E. Schreiner, A. Kohlmeyer, D. Marx, *Biophys. J.* **2004**, *86*, 1393.
[38] D. K. Chang, D. W. Urry, *Chem. Phys. Lett.* **1988**, *147*, 395.
[39] R. Glaves, M. Baer, E. Schreiner, R. Stoll, D. Marx, *Chem. Phys. Chem.* **2008**, *9*, 2759.
[40] A. Krukau, I. Brovchenko, A. Geiger, *Biomacromolecules* **2007**, *8*, 2196.
[41] M. Baer, E. Schreiner, A. Kohlmeyer, R. Rousseau, D. Marx, *J. Phys. Chem. B* **2006**, *110*, 3576.
[42] Y. Sugita, Y. Okamoto, *Chem. Phys. Lett.* **1999**, *314*, 141.
[43] N. Blinov, M. Khorvash, D. S. Wishart, N. R. Cashman, A. Kovalenko, *ACS Omega* **2017**, *2*, 7621.
[44] T. Terakawa, S. Takada, *Biophys. J.* **2011**, *101*, 1450.
[45] S. Patel, E. Vierling, F. Tama, *Biophys. J.* **2014**, *106*, 2644.
[46] Y. Chebaro, A. J. Ballard, D. Chakraborty, D. J. Wales, *Sci. Rep.* **2015**, *5*, 10386.
[47] G. H. Zerze, C. M. Miller, D. Granata, J. Mittal, *J. Chem. Theory Comput.* **2015**, *11*, 2776.
[48] T. Kosciolek, D. W. A. Buchan, D. T. Jones, *Sci. Rep.* **2017**, *7*, 6999.
[49] J. Kragelj, M. Blackledge, M. R. Jensen, *Adv. Exp. Med. Biol.* **2015**, *870*, 123.
[50] H. Gong, S. Zhang, J. Wang, H. Gong, J. Zeng, *J. Comput. Biol.* **2016**, *23*, 300.
[51] C. K. Fisher, A. Huang, C. M. Stultz, *J. Am. Chem. Soc.* **2010**, *132*, 14919.
[52] C. K. Fisher, C. M. Stultz, *Curr. Opin. Struct. Biol.* **2011**, *21*, 426.
[53] C. C. David, D. J. Jacobs, *Meth. Mol. Biol.* **2014**, *1084*, 193.
[54] A. Tarakanova, W. Huang, A. S. Weiss, D. L. Kaplan, M. J. Buehler, *Biomaterials* **2017**, *127*, 49.
[55] G. C. Yeo, C. Baldock, S. G. Wise, A. S. Weiss, *J. Biol. Chem.* **2014**, *289*, 34815.
[56] C. A. Hoeve, P. J. Flory, *Biopolymers* **1974**, *13*, 677.
[57] S. Rauscher, R. Pomes, *Biophys. J.* **2018**, *114*, 369a.
[58] B. Bochicchio, A. Pepe, A. M. Tamburro, *Chirality* **2008**, *20*, 985.
[59] M. Miao, J. T. Cirulis, S. Lee, F. W. Keeley, *Biochemistry* **2005**, *44*, 14367.
[60] J. Djajamuliadi, T. F. Kagawa, K. Ohgo, K. K. Kumashiro, *Matrix Biol.* **2009**, *28*, 92.
[61] A. M. Tamburro, A. Pepe, B. Bochicchio, *Biochemistry* **2006**, *45*, 9518.
[62] L. B. Dyksterhuis, C. Baldock, D. Lammie, T. J. Wess, A. S. Weiss, *Matrix Biol.* **2007**, *26*, 125.
[63] M. Feig, J. Karanicolas, C. L. Brooks 3rd, *J. Mol. Graph. Model.* **2004**, *22*, 377.
[64] A. Bakan, L. M. Meireles, I. Bahar, *Bioinformatics* **2011**, *27*, 1575.
[65] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, G. D. Peterson, R. Roskies, J. R. Scott, N. Wilkens-Diehr, *Comput. Sci. Eng.* **2014**, *16*, 62.

© 2018 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim