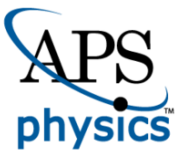

Universality and Specificity in Protein Fluctuation Dynamics

J. Copperman, M. Dinpajooh, E. R. Beyerle, and M. G. Guenza

Phys. Rev. Lett. **119**, 158101 — Published 11 October 2017

DOI: 10.1103/PhysRevLett.119.158101

Universality and specificity in protein fluctuation dynamics

J. Copperman[a]

Department of Physics, University of Oregon, Eugene, Oregon 97403

M. Dinpajooh, E. R. Beyerle, M.G. Guenza[b]

Department of Chemistry and Institute of Theoretical Science,

University of Oregon, Eugene, Oregon 97403

(Dated: September 13, 2017)

Abstract

We investigate the universal scaling of protein fluctuation dynamics with a site-specific diffusive model of protein motion, which predicts an initial subdiffusive regime in the configurational relaxation. The long-time dynamics of proteins is controlled by an activated regime. We argue that the hierarchical free energy barriers set the timescales of biological processes and establish an upper limit to the size of single protein domains. We find it compelling that the scaling behavior for the protein dynamics is in close agreement with the Kardar-Parisi-Zhang scaling exponents.

[a] Co-corresponding author. Electronic mail: copperma@uwm.edu

[b] Co-corresponding author. Electronic mail: mguenza@uoregon.edu

Proteins are molecular machines whose structure and dynamics have been evolutionarily designed to perform functional roles. While random sequences of amino-acids exhibit disorder and frustration, native sequences possess funnel-like free energy landscapes, and at physiological temperature can reversibly fold to unique global configurations.[1, 2] In their folded state, proteins possess specific dynamical pathways that allow the motion required for biological function. In the picosecond regime, proteins fluctuate around a single structural minimum characterized by the topology of protein connectivity.[3] In the nanosecond to microsecond regime, fluctuations lead to transition between metastable states which are of nearly equal free energy. The slowest modes are biologically relevant for protein-protein recognition and enzymatic activity.[4, 5] The longest timescale is set by the kinetics of protein folding and unfolding, which is in the millisecond regime and longer.

Protein foldability and functionality occur in a narrow range of temperature, salt concentration, pressure, and volume: this implies a self-similarity in their dynamics. Simultaneously, proteins have exquisitely precise biological activity, which requires a high level of specificity.[6] It is of paramount relevance to establish a precise relation between dynamical pathways, which are protein specific, and the dynamical observables shared across proteins as a class.

Dynamical normal-mode pathways are well-described by the Langevin Equation for Protein Dynamics (LE4PD), which is the starting point of this study.[7–9] The LE4PD represents proteins as chains of coarse-grained sites interacting through an effective potential-of-mean-force and the hydrodynamic interaction. In its normal mode representation, the free energy landscape is decomposed into a set of linearly-independent, mode-specific, rough free-energy surfaces, which display the relevant free energy barriers and kinetic pathways.[10]

In this Letter we investigate the emergence of general scaling exponents in the specific dynamical behavior by studying fourteen equilibrium molecular dynamics (MD) simulations covering twelve various proteins. Input to the LE4PD ranges from 50 ns to 1.23 ms equilibrium MD trajectories. The starting structures were taken from NMR or X-ray structures.[11–19] More details about the proteins and MD simulation protocols can be found in the supplementary material (SM).[17]

In a set of recent papers the LE4PD was shown to provide protein-specific and site-specific *quantitative* predictions for the pico- to nano-second dynamics measured by nuclear magnetic resonance (NMR) spin-lattice relaxation time, T_1, spin-spin relaxation time, T_2, and nuclear

Overhauser effect (NOE), which qualifies the LE4PD as a reliable representation of protein dynamics close to the folded state.[7–9] Theoretical predictions correlate with allosteric, catalytic, and binding activity reaching the microsecond regime.[10] Here we address the shared dynamical properties across proteins as a class, spanning the short-time diffusive regime out to the longest timescales of folding and unfolding.[20]

In the LE4PD each molecule is represented by a set of N coarse-grained sites, located in the α-carbons along the primary sequence. Given the bond l_i with $i = 1, ..., N-1$ the structural coupling between bonds is represented by the $\mathbf{U}$ matrix $(\mathbf{U}^{-1})_{ij} = \langle \vec{l}_i \cdot \vec{l}_j \rangle / (\langle |\vec{l}_i| \rangle \langle |\vec{l}_j| \rangle)$. Bond dynamics is coupled through the hydrodynamic interaction matrix, $\mathbf{H}$, which in its preaveraged form enters the LE4PD equation through the matrix product $\mathbf{L} = \mathbf{aHa}^T$, where $\mathbf{a}$ is the matrix of the backbone connectivity. The preaveraged hydrodynamic interaction is $H_{ij} = (\overline{\zeta}/\zeta_i)\delta_{ij} + (1-\delta_{ij})\overline{r}^w \langle r_{ij}^{-1} \rangle$, with $\overline{\zeta} = N^{-1} \sum_{i=1}^{N} \zeta_i$ the site-averaged friction coefficient, r_{ij} the distance between a pair of $\alpha-$carbons, and $\overline{r}^w$ the hydrodynamic radius corresponding to the fraction of a CG-site surface exposed to water. The stochastic equation of motion for bond $\vec{l}_i(t)$ is

$$\frac{\partial \vec{l}_i(t)}{\partial t} = -\sigma \sum_{j,k} L_{ij} U_{jk} \vec{l}_k(t) + \vec{v}_i(t) \ , \tag{1}$$

with $\sigma = 3k_{\mathrm{B}}T/(l^2\overline{\zeta})$, T the temperature, k_{B} the Boltzmann constant, and $\vec{v}_i(t)$ the random, delta-correlated bond velocity. This linear Langevin equation is used for constructing a convenient basis for projecting the non-linear dynamics from the atomistically-detailed simulations. Once transformed into the diffusive normal mode coordinates, mode-dependent energy barriers are used to rescale the dynamics using a Kramers type of approach.[8, 21]

Input parameters are the system conditions (temperature, solvent viscosity) and the equilibrium structural properties from the configurational ensemble spanned by the atomistic simulations. The diffusive normal mode representation is obtained from the diagonalization of the matrix product $\mathbf{LU}$, with $\mathbf{Q}$ the eigenvectors and λ_a the eigenvalues. In the diffusive modes Eq. 1 reduces to a set of $N-1$ uncoupled linear equations where the modes, $\vec{\xi}_a(t) = \sum_i Q_{ai}^{-1} \vec{l}_i(t)$, uniquely define the instantaneous conformation of the macromolecule. One may notice that any anisotropy inherent to the dynamics of $\vec{l}_i(t)$ in the MD trajectory is inherited by the vector $\vec{\xi}_a(t)$ (see SM). The mode basis spans the same space as the bond vector basis with near linearity: $\langle \vec{\xi}_a \cdot \vec{\xi}_b \rangle \cong \delta_{ab} l^2/\mu_a$ with $\mu_a = \sum_{i,j} Q_{ai}^{-1} U_{ij}^{-1} Q_{ja}$, which can then decompose the free energy landscape into a convenient set of mode-dependent energy

maps.[8] The length of the mode is defined as $L_a^2 = l^2/\mu_a$, while the associated diffusive timescale $\tau_a = (\sigma\lambda_a)^{-1}$. It is worth mentioning that our formalism is distinct from the mechanistic view of the dynamics in terms of the frequencies of normal modes of molecular vibrations often adopted in molecular biophysics, which lacks the dissipative dynamics and neglects hydrodynamics and internal energy barriers that are important at long time.[3, 22, 23] Internal dissipation due to fluctuations in the hydrophobic region is included in our formalism by accounting for an effective protein internal viscosity and considering the relative exposure of each amino acid to the hydrophobic region.[7]

The first three global modes of the LE4PD, in most cases, describe the protein rotation tensor. Internal modes, with index $p = 1, ..., N_p$ and $N_p = N-3$, are characterized by a small number of metastable minima whose depth, or the barriers between them, are largest for the low mode numbers corresponding to the most collective, large-amplitude fluctuations.[9] We evaluate the barrier height for mode p as the Median Absolute Deviation (MAD)[24] from the global minimum on the mode orientational free energy surface , where the MAD is an appropriate metric that is robust to outliers and unevenly sampled distributions, and captures the dispersion in a set of data (see SM[17]). Figure 1 shows that the free energy barriers scale with mode length as $E_p^\dagger = (\epsilon L_p)^\gamma$ with $\gamma = 0.93 \pm 0.20$ and $\epsilon = 6.5$ (kcal mol^{-1})$^{1/\gamma}$ nm^{-1}.

The direct proportionality between the free energy barrier and mode length appears to extend from the 1 Å scale out to the nanometer scale, which characterizes the overall size of the protein. On the local length scale, where the specific chemical nature of the protein primary sequence is most important, the free energy barriers are still protein specific.

The observed scaling law is consistent with the hierarchical nature of the protein free energy landscape. On the local scale (large p) the bonds fluctuate independently, while large-amplitude correlated fluctuations occur when the bonds, which are dynamically correlated on the lengthscale of the mode, transition collectively (small p).[25] The model connects the complex hierarchical nature of the free energy landscape of a protein in solution to the structure of a glassy fluid.[26, 27]

In the *short-time regime*, the protein fluctuates around the minima of the free energy landscape: the dynamics are diffusion controlled and well represented by the LE4PD with free energy barriers not included. The LE4PD is an extension of the traditional Rouse-Zimm approach to the dynamics of macromolecules in solution,[28, 29] while it also includes local

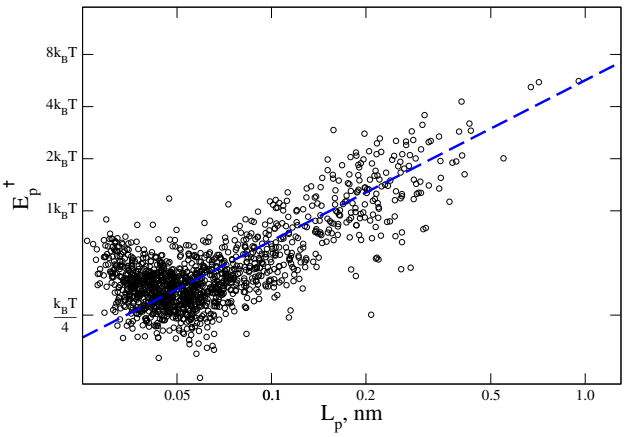

FIG. 1: Mode-dependent free energy barrier, $E_p^\dagger = (\epsilon L_p)^\gamma$, mode lengthscale L_p for all proteins with $\gamma = 0.93 \pm 0.20$ and $\epsilon = 6.5$ (kcal mol^{-1})$^{1/\gamma}$ nm^{-1} (blue dashed line).

semiflexibility and non-linear connectivity, anisotropic rotational dynamics, and the effect of the hydrophobic core on the hydrodynamic interaction: these play a role in the dynamics at short time.[9, 10]

The mode lengthscale scales with internal mode number p as $L_p \propto p^{-\beta}$ with $\beta = 0.41 \pm 0.06$, for all the proteins in this study (see left panel of Figure 2). The scaling indicates a greater stability than in the case of the completely flexible unfolded chain where $\beta = 1$.[28]

In the diffusive regime the average mean-squared displacement (MSD) for one amino acid inside a protein can be written in the center-of-mass frame as

$$\frac{1}{N}\sum_{i=1}^{N}\left\langle\left(\vec{R}_i(t)-\vec{R}_i(0)\right)^2\right\rangle = \frac{2}{N^2}\sum_{i=1}^{N}\sum_{j=i+1}^{N}\sum_{k,l=i}^{j-1}\sum_{a=1}^{N-1} Q_{ka}Q_{la}\left(\langle\vec{\xi}_a^2\rangle - \langle\vec{\xi}_a(t)\cdot\vec{\xi}_a(0)\rangle\right) \ , \quad (2)$$

where R_i is the position of the α-carbons in the coarse-grained model. At zero time difference, the bead position lengthscale for internal mode p is
$R_p^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=i+1}^{N}\sum_{k,l=i}^{j-1} Q_{kp}Q_{lp}L_p^2 \approx C_s^2 L_p^3$. The diffusive mode timescale is predicted by the LE4PD to scale with mode length as $\tau_{0,p} = C_\tau L_p^\alpha$, with $\alpha = 2.00 \pm 0.41$, where the index 0 indicates that free energy barriers are not included in the short-time LE4PD. Approximating R_p and τ_p with their scaling forms, and the discrete sum as an integral

$$MSD \approx 2C_s^2 L_{\max}^3 \int_{p=1}^{N_p} dp\ p^{-3\beta}\left(1-\exp\left[-\frac{tp^{\beta\alpha}}{C_\tau L_{\max}^\alpha}\right]\right) \ , \quad (3)$$

where $L_{\max}$ is the largest internal mode length.

The short-time expansion of Eq. 3 scales as $MSD \propto t^\nu$ with $\nu = (3\beta-1)/(\beta\alpha)$. The LE4PD gives the exponents $\beta = 0.41$ and $\alpha = 2.00$, which leads to a strongly subdiffusive

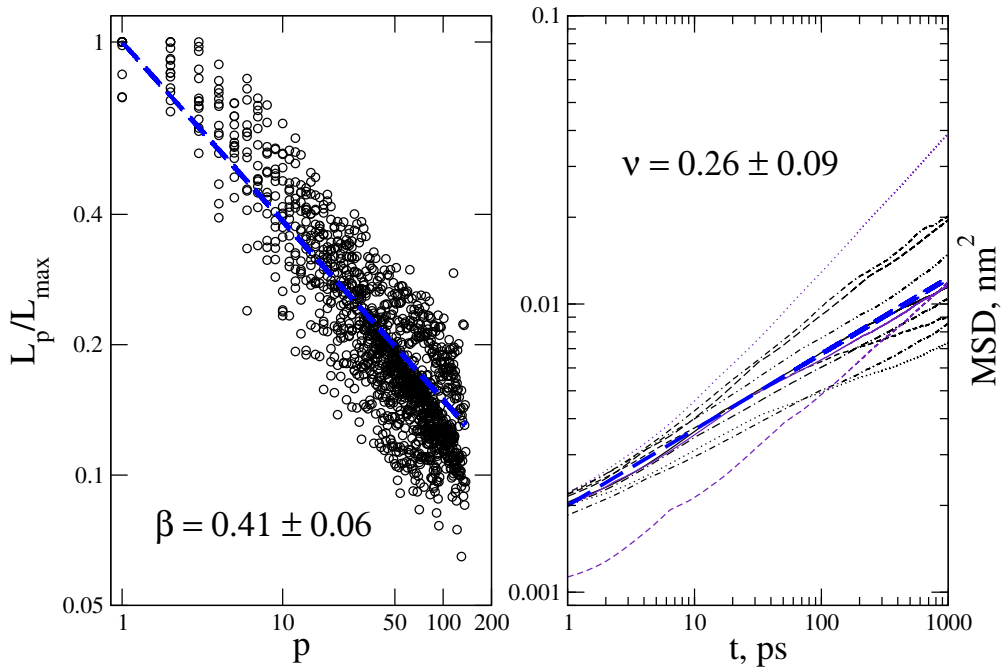

FIG. 2: Left: Mode length L_p and mode index for all proteins and the scaling $L_p = L_{\max} p^{-\beta}$ with $\beta = 0.41$ (blue dashed line). Right: Predicted subdiffusive mean-squared displacement $\langle R^2(t) \rangle \propto t^{\nu}$ with $\nu = 0.26$ (blue dashed line), the average MSD of all proteins (solid-line), and the MSD of each protein individually.

short-time regime characterized by the exponent $\nu = 0.28$. This exponent agrees well with the exponent $\nu = 0.26 \pm 0.09$ measured directly from MD simulations, see Figure 2, and with the exponent $\nu \sim 0.3$ that models thiyl radical recombination experiments, where the exponent is shown to be an inherent property of the polypeptide backbone, independent of the primary sequence of the protein.[30]

In the *long-time regime* the LE4PD accounts for the effect of the local free energy barriers on the internal dynamics, as the friction becomes mode dependent, including thermal activation over the mode-dependent free energy barrier[21] $\overline{\zeta} \to \overline{\zeta}\ exp[\langle E_p^{\dagger} \rangle / (k_{\mathrm{B}} T)]$, leading to the slowing of the mode timescale as

$$\tau_p = C_\tau L_p^{\alpha} \exp\left[\frac{(\epsilon L_p)^{\gamma}}{k_{\mathrm{B}} T}\right] . \tag{4}$$

This renormalization provides an average correction which approximately accounts for the local barrier crossing, in agreement with free energy landscape theories.[1, 26] Figure 3 compares the LE4PD mode-dependent relaxation times with and without inclusion of the free energy barriers.

At long times the dynamics are dominated by the largest internal mode lengthscale,

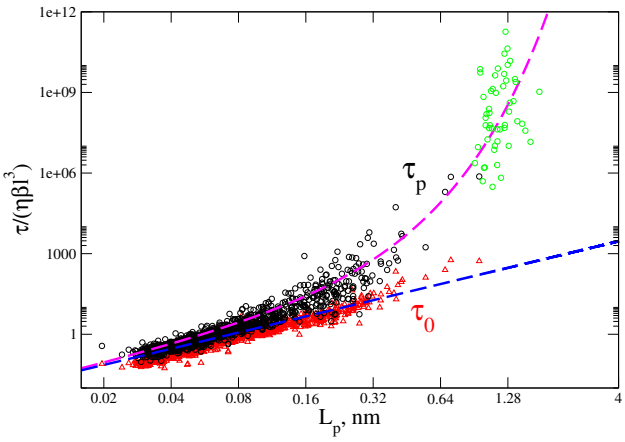

FIG. 3: Reduced mode timescales, $\tau/(\eta\beta l^3)$, for all proteins as a function of mode length scale, L_p, where η, β, and l are solvent viscosity, inverse temperature, and average effective bond length, respectively. The red triangles show the reduced mode timescales before free energy barrier correction and the scaling $\tau_{0,p} = C_\tau L_p^\alpha$ with $\alpha = 2.00 \pm 0.41$ (blue line). Free-energy barrier corrected, reduced mode timescales (black circles) and the scaling $\tau_p = C_\tau L_p^\alpha \exp[\frac{(\epsilon L_p)^\gamma}{k_B T}]$ (magenta line). Reduced timescales of folding for 52 proteins, plotted against the largest mode lengthscale, approximately the protein radius of gyration (green circles).[6]

$L_{\max}$. For $(\epsilon L_p)^\gamma < k_B T$, the relatively small barriers only slightly slow the mode relaxation; therefore, the diffusive and barrier rescaled timescales roughly coincide. As the fluctuations grow in size the free energy barrier correction causes the mode relaxation time to rapidly propagate out to folding timescales at the nanometer lengthscale. This lengthscale happens to be the typical size of single-domain proteins and the analysis of a dataset of 2-state folding times for 52 proteins,[6] shows clustering around the line representing Eq. 4. LE4PD relaxation times from the millisecond unfolding/folding trajectories of ubiquitin by Piana et al., also agrees with Eq. 4. This equation effectively extends to the length and timescales of protein folding, suggesting that the hierarchical roughness of the free energy landscape is the leading contribution to the dynamical slowing down of fluctuations and folding/unfolding.

When including the free energy barriers Eq. 3 becomes

$$MSD = 2C_s^2 L_{\max}^3 \int_{p=1}^{N_p} dp\; p^{-3\beta} \left(1 - \exp\left[-\frac{t p^{\beta\alpha}}{C_\tau L_{\max}^\alpha} \exp\left(-\frac{(\epsilon L_{\max} p^{-\beta})^\gamma}{k_B T}\right)\right]\right) . \quad (5)$$

Figure 4 compares the mean-squared displacement from the simulation with the theoretical LE4PD predictions. It can be observed that the initial subdiffusive power-law scaling at short times undergoes a long crossover to a barrier-dominated, slow growth. At 300 K,

below the folding temperature, the protein is stable and the MSD levels off at around a microsecond. The scaling form from the integral in Eq. 5 quantitatively reproduces the shape of mean squared displacement with the largest internal mode length of $L_{\max} = 0.6$ nm (left panel of Figure 4). Above the folding temperature, the MSD is of the order of the size of the protein, and levels off around the folding relaxation time of $\sim .1\ ms$, described by Eq. 5 with $L_{\max} = 1.2$ nm (right panel of Figure 4).

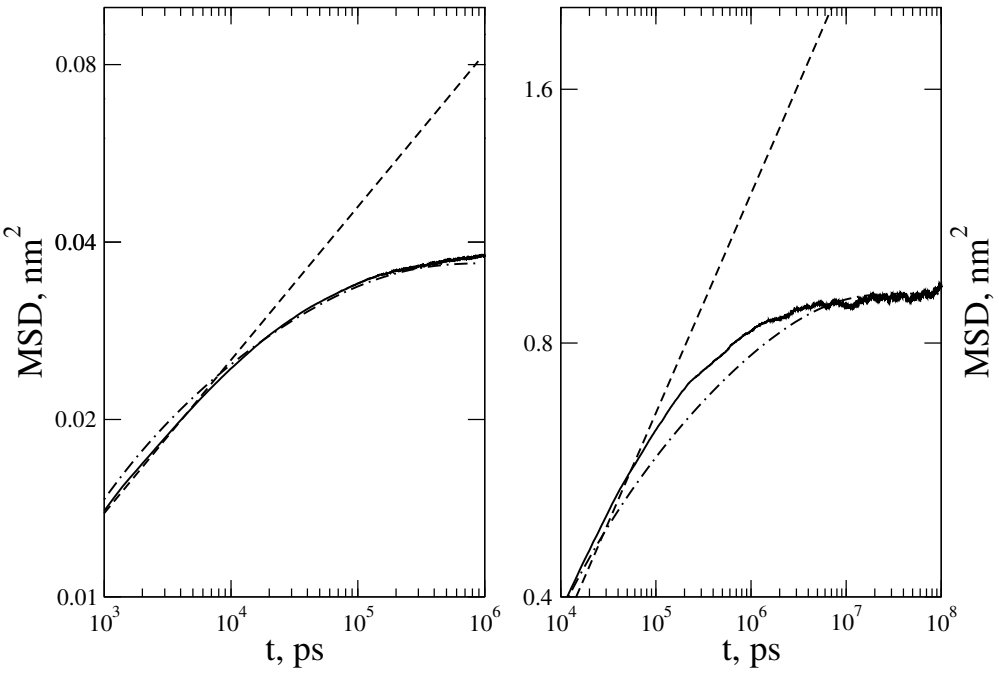

FIG. 4: Mean squared displacement of the ubiquitin protein at 300 K (left panel) and 390 K (right panel) from the millisecond scale simulations of Piana et al. (solid line).[20] The short-time subdiffusive $MSD \propto t^{\nu}$ with $\nu = 0.26$ (dashed line), and the long-time crossover to the activated regime from Eq. 5 (dot-dashed line).

The scaling of free energy barriers with fluctuation length sets an upper limit on the domain size in proteins. The size distribution of protein domains found in biology is peaked at around 100 amino acids and near zero by 300 amino acids.[31, 32] This corresponds to proteins of maximum size $R_g \sim 2.0$ nm, with relaxation time of ~ 1 minute. This result indicates that protein domains larger than the typical size found in nature would have relaxation times exceeding biologically relevant timescales.

A simple representation that captures the proper scaling exponents predicted by our approach is the theory of motion on a random energy landscape. A general elastic manifold

embedded in a field of random pinning potentials results qualitatively in hierarchical free energy landscapes of the type observed here for proteins.[33] Following this analogy, the origin of the roughness in the free energy landscape of the proteins can be disorder, which is generated by the constantly forming and breaking of the hydrogen-bonding network between protein and solvent, on a timescale of picoseconds.

Protein function operates over a narrow temperature window bounded by an upper melting temperature and a lower glass transition temperature.[26] In this temperature range the hydrogen bonding network is poised at critical stability. The entropy-enthalpy compensation of individual hydrogen bonds results in the metastability of multiple protein configurations.[34–36] We conjecture that the critical nature of the hydrogen bond network serves as random energy perturbations to the protein, and that the energy-scale and length-scale involved in making and breaking of hydrogen bonds sets ϵ the energy per unit length.

We make the assumption that the motion of an average protein site, in the field of its complex, rapidly-varying, many-body environment, can be described by harmonic fluctuations subjected to a time-dependent energetic disorder. By making use of this reasonable assumption, the LE4PD may be mapped to the Directed Polymer in Random Media (DPRM) (see the SM),[17] which is the simplest model to describe finite-temperature motion on a d-dimensional random energy landscape.[37] Specific considerations of the folded state of proteins as fractal objects has led to the conjecture that folded proteins are poised on the edge of metastability where the space-filling dimension is $d_f > 2$ and the spectral dimension is $d_s < 2$,[38] while analysis of the vibrational spectrum of globular proteins predicts a general $d \sim 2$ dimensional folded state.[39] Thus protein dynamics maps onto the DPRM theory in roughly (2+1) dimension, with the extra dimension given by time.

Through a straightforward mathematical transformation the (2+1) DPRM model maps into the celebrated (d=2) Kardar-Parisi-Zhang (KPZ) equation describing the surface height of a solid growing by random deposition.[40] The effective free energy of the DPRM is directly proportional to the surface height of a growing solid in the KPZ model.[41, 42] Noting that the energetic disorder explains the universal features of the protein dynamics in this work, we conjecture that the origin of the general scaling of the free energy barriers observed for all proteins is the rough surface-height distribution of the KPZ universality class (see SM).[17] In fact, we find it compelling that the scaling exponents predicted by our formalism such as the barrier exponent, $0.38 = \gamma\beta$ ($E_p^\dagger \propto p^{-\gamma\beta}$) and the early-time exponent, $\nu = 0.26$ are in

close agreement with the roughening exponent $\chi = 0.39 \pm 0.01$ of the $d = 2$ KPZ-equation and the early-time growth exponent $L(t) \propto t^{\nu}$ of the (2+1) DPRM $\nu = 0.240 \pm 0.001$.[43, 44]

While biology demands specificity in protein motions, the narrow physical conditions in which proteins are biologically active suggests self-similarity and criticality in their dynamical behavior. Concealed by the evident complexity of protein dynamics should be simple universal laws. In this Letter we show that underlying the specific dynamical pathways of proteins is in fact a universal, hierarchical scaling that suggests an origin in the Directed Polymer in Random Media dynamical model, and the Kardar-Parisi-Zhang universality class. This interesting analogy may emerge once the dynamics are represented in the diffusive normal modes of the LE4PD theory because it decomposes the configurational free energy into linearly-independent modes, from which the universal scaling is readily identified. This hierarchical scaling in the free energy sets the overall time- and length-scales of biological processes, which involve the rearrangement of protein domains.

This work was supported by the National Science Foundation Grant CHE-1362500. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575. In particular, we acknowledge the D. E. Shaw group for providing us with the Anton trajectories.

[1] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins Struct. Funct. Bioinf. **21**, 167 (1995).

[2] K. A. Dill and H. S. Chan, Nature Struct. Biol. **4**, 10 (1997).

[3] I. Bahar, A. R. Atilgan, and B. Erman, Fold Des. **2**, 173 (1997).

[4] K.-I. Okazaki and S. Takada, Proc. Natl. Acad. Sci. U.S.A. **105**, 11182 (2008).

[5] O. F. Lange, N.-A. Lakomek, C. Fares, G. F. Schroder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, and B. L. de Groot, Science **320**, 1471 (2008).

[6] D. De Sancho, U. Doshi, and V. Munoz, J. Am. Chem. Soc. **131**, 2074 (2009).

[7] E. Caballero-Manrique, J. K. Bray, W. a. Deutschman, F. W. Dahlquist, and M. G. Guenza, Biophys. J. **93**, 4128 (2007).

[8] J. Copperman and M. G. Guenza, J. Phys. Chem. B **119**, 9195 (2014).

[9] J. Copperman and M. G. Guenza, J. Chem. Phys. **143**, 243131 (2015).

[10] J. Copperman and M. G. Guenza, J. Chem. Phys. **145**, 015101 (2016).

[11] W. Vijay-Kumar, S., Bugg, Ch., Wilkinson, K., Vierstra, R., Hatfield, P., and Cook, J. Biol. Chem. **262**, 6396 (1987).

[12] R. Ishima, D. A. Torchia, S. M. Lynch, A. M. Gronenborn, and J. M. Louis, J. Biol. Chem. **278**, 43311 (2003).

[13] G. I. Lee, Z. Ding, J. C. Walker, and S. R. Van Doren, Proc. Natl. Acad. Sci. U.S.A. **100**, 11261 (2003).

[14] G. Gao, V. Semenchenko, S. Arumugam, and S. R. Van Doren, J. Mol. Biol. **301**, 537 (2000).

[15] J. Lu, D. P. Cistola, and E. Li, J. Mol. Biol. **330**, 799 (2003).

[16] C. Williams, D. Rezgui, S. N. Prince, O. J. Zaccheo, E. J. Foulstone, B. E. Forbes, R. S. Norton, J. Crosby, A. B. Hassan, and M. P. Crump, Structure **15**, 1065 (2007).

[17] See Supplemental Material at [url], which includes Refs. [45-48] .

[18] D. E. Shaw, P. Maragakis, K. Lindorff-larsen, S. Piana, Y. Shan, and W. Wriggers, Science **330**, 341 (2010).

[19] K. Lindorff-Larsen, P. Maragakis, S. Piana, and D. E. Shaw, J. Phys. Chem. B **120**, 8313 (2016).

[20] S. Piana, K. Lindorff-Larsen, and D. E. Shaw, Proc. Natl. Acad. Sci. U.S.A. **110**, 5915 (2013).

[21] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, New York, 2001).

[22] I. Bahar, A. Atilgan, M. Demirel, and B. Erman, Phys. Rev. Lett. **80**, 2733 (1998).

[23] H. Na, G. Song, and D. Ben-Avraham, Phys. Biol. **13**, 016008 (2016).

[24] D. Ruppert, *Statistics and Data Analysis for Finacial Engineering* (Springer, New York, 2010).

[25] M. B. Jackson, J. Chem. Phys. **99**, 7253 (1993).

[26] Z. Luthey-schulten and P. G. Wolynes, Annu. Rev. Phys. Chem. **48**, 545 (1997).

[27] D. J. Wales, J. Chem. Phys. **142**, 130901 (2015).

[28] M. Doi and S. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, New York, 1986).

[29] A. Perico and M. Guenza, J. Chem. Phys. **83**, 3103 (1985).

[30] L. Milanesi, J. P. Waltho, C. A. Hunter, D. J. Shaw, G. S. Beddard, G. D. Reid, S. Dev, and M. Volk, Proc. Natl. Acad. Sci. U.S.A. **109**, 19563 (2012).

[31] D. Xu and R. Nussinov, Fold Des. **3**, 11 (1998).

[32] A. Heger and L. Holm, J. Mol. Biol. **328**, 749 (2003).

[33] L. Balents, J.-P. Bouchaud, and M. Mézard, J. Phys. I France **6**, 1007 (1996).

[34] M. Tarek and D. J. Tobias, Phys. Rev. Lett. **88**, 138101 (2002).

[35] S. K. Pal, J. Peon, and A. H. Zewail, Proc. Natl. Acad. Sci. U.S.A. **99**, 1763 (2002).

[36] Y. Levy and J. N. Onuchic, Annu. Rev. Biophys. Biomol. Struct. **35**, 389 (2006).

[37] M. Kardar and Y. C. Zhang, Phys. Rev. Lett. **58**, 2087 (1987).

[38] S. Reuveni, R. Granek, and J. Klafter, Phys. Rev. Lett. **100**, 208101 (2008).

[39] D. Ben-Avraham, Phys. Rev. B **47**, 14559 (1993).

[40] M. Kardar, G. Parisi, and Y.-C. Zhang, Phys. Rev. Lett. **56**, 889 (1986).

[41] T. Halpin-Healy and Y. C. Zhang, Phys. Rep. **254**, 215 (1995).

[42] S. Roux, A. Hansen, and E. L. Hinrichsen, J. Phys. A: Math. Gen. **24**, 295 (1991).

[43] B. M. Forrest and L. H. Tang, Phys. Rev. Lett. **64**, 1405 (1990).

[44] L. H. Tang, B. M. Forrest, and D. E. Wolf, Phys. Rev. A **45**, 7162 (1992).

[45] A. Amadei, M. A. Ceruso, and A. Di Nola, Proteins: Struct., Funct., Genet. **36**, 419 (1999).

[46] I. Daidone and A. Amadei, WIREs Comput. Mol. Sci. **2**, 762 (2012).

[47] A. Atilgan, S. Durell, R. Jernigan, M. Demirel, O. Keskin, and I. Bahar, Biophysical Journal **80**, 505 (2001).

[48] E. Lifshitz and L. Pitaevskii, *Statistical Physics, Part 2: Volume 9* (Butterworth-Heinemann, Oxford, 1980).