THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

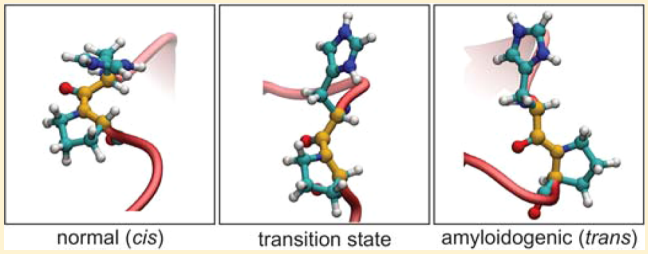
pubs.acs.org/JPCB

Energetics and Mechanism of the Normal-to-Amyloidogenic Isomerization of β2-Microglobulin: On-the-Fly String Method Calculations

Spencer T. Stober and Cameron F. Abrams*

Department of Chemical and Biological Engineering, Drexel University, Philadelphia, Pennsylvania 19104, United States

ABSTRACT: We use on-the-fly finite temperature string method in collective variables to study the transition from a normal to an amyloidogenic conformation of β2-microglobulin. We show that the protonation state of two histidine residues is of key importance and that under acidic (protonating) conditions, the transition to the amyloidgenic form is facilitated by both displacement of N-terminal residues to disrupt a hydrophobic pocket and by side-chain/side-chain electrostatic attraction, both of which facilitate a cis−trans prolyl isomerization. The free energy barriers for the normal-to-amyloidogenic isomerization are found to be 14.9 and 7.1 kcal/mol for the neutral and protonated cases, respectively, consistent with enhanced amyloidgenesis at low pH observed both in vitro and in hemodialysis-associated amyloidosis, and somewhat lower than experimentally determined barriers for bare prolyl cis−trans isomerization. We suggest specific mutagenesis experiments which could be used to further validate the mechanism observed.



INTRODUCTION

Protein misfolding and amyloidosis can disrupt normal cellular function and accompany many neurodegenerative diseases (e.g., Alzheimer's, Parikinson's, transmissible spongiform encephalopathy).[1,2] Understanding the mechanisms by which normal proteins begin to polymerize into amyloid is key to understanding these diseases. β2-Microglobulin (β2M) is an invariant part of the class-I major histocompatibility complex and also exists as a serum protein.[3] β2M tends to accumulate in the serum and synovial fluid of patients undergoing dialysis, where it can polymerize into amyloid fibrils, leading to a condition known as hemodialysis-associated amyloidosis (HAA).[4,5]

High concentration is not sufficient for β2M aggregation, because normal β2M does not form amyloid fibrils in vitro even at concentrations in excess of those found in synovial fluid of sufferers of HAA.[6] Truncation mutants lacking the six N-terminal residues (ΔN6-β2M), however, polymerize readily in vitro,[7] and wild-type (WT) β2M polymerizes at low pH[8] through an apparently short-lived (and structurally uncharacterized) amyloidogenic intermediate state. Recently, the NMR structure of β2M in both a normal and a ΔN6 mutant aggregation-prone state were determined (Figure 1).[6] Crucially, these structures reveal that an unusual *cis*-prolyl peptide bond (between His31 and Pro32) flips into trans as part of the transition to the amyloidogenic state. Exactly how this flip is triggered and/or catalyzed and the degree to which this depends upon interactions with N-terminal residues and/or pH is not known in detail. Experimental limitations also make it problematic to determine the energetics of this conformational transition, at least for WT β2M.

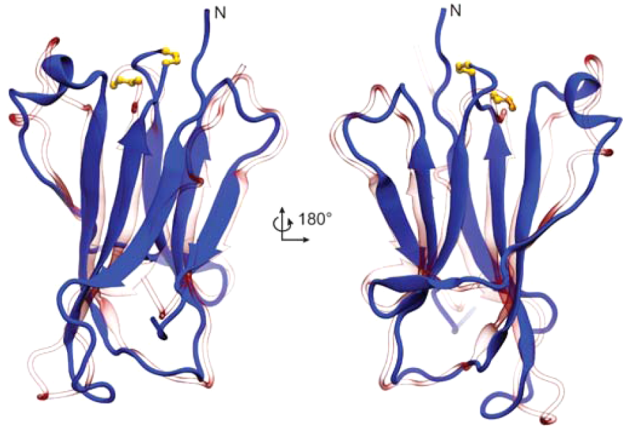


Figure 1. Two views of normal (blue) and amyloidogenic (red, transparent) β2M. The His31−Pro32 peptide bond torsion angle determinants (Cα−C−N−Cα) are shown in yellow.

The purpose of this article is to report on the use of advanced molecular simulation methods to study the details of this transition. We use an all-atom model of WT β2M in explicit-solvent molecular simulations to model this transition at both neutral and acidic pH using the finite temperature on-the fly string method (OFSM) in collective variables.[9−11] The principal curves between these two states that result from our calculations provide both new details about this transition and a

Received: May 17, 2012
Revised: July 6, 2012
Published: July 14, 2012

meaningful estimate of its energetics. In particular, we provide a mechanistic role of the protonation state of two histidines (His31 and His84) on determining the relative stabilities and barrier-crossing energies between normal and amyloidogenic forms of β2M.

METHODS

Finite Temperature On-the-Fly String Method: General Considerations. The string method is an algorithm for identifying discretized pathways of minimal energy through configuration space, or, in a collective variable description used here, pathways of minimal *free* energy through collective variable space.[10] However, the minimum free energy path is based on local features in the free energy landscape that may contain thermal-scale roughness or very narrow reaction tubes, which do not reflect the true transition pathway of the reaction mechanism. In the finite temperature version of string method, the addition of a thermostat term serves to average out local thermal-scale roughness in the free energy landscape and/or favor broader reaction tubes that although less favored energetically, may be more favored entropically.[11,14] This pathway is referred to as the *principal curve* and depends on global features in the free energy landscape.

The "string" of string method is a set of simultaneous, coupled (MD) systems (termed "images") equally spaced along a curve $z(s,t)$ in collective variable (CV) space; s is the image location along the string ($s \in [0,1]$) and t is simulation time. The basic idea of the string method is to use short MD simulations of each image restrained at its respective position on the string to compute forces used to update the image positions, thereby "evolving" the string as a finite-length object along the free energy landscape of CV space. This force measurement/position update is performed iteratively until the positions no longer change, at which point the string has relaxed to its principal curve. The string method has been applied to the analysis of myosin VI and the transmembrane domain of Gloeobacter violaceus channel (GLIC),[15,16] and recently to the activation of the insulin-receptor kinase.[17]

In OFSM, rather than an iterative update approach, image updates happen in lockstep with the MD updates of atomic positions.[9] That is, auxiliary variables z that realize image positions in CV space and atomic variables x of all images evolve under the coupled dynamics (see also the Addendum of Maragliano and Vanden-Eijnden[9]):

$$m_i\ddot{x}_i(s,t) = -\nabla_i U(\mathbf{x}(s,t)) - m_i\gamma\dot{x}_i(s,t) + \sqrt{2\gamma k_B T}\eta_x(t) - \kappa\sum_j [\theta_j[\mathbf{x}(s,t)] - z_j(s,t)]\frac{\partial\theta_j[\mathbf{x}(s,t)]}{\partial x_j} \quad (1)$$

$$\overline{\gamma}\ddot{z}_j(s,t) = \sum_k [\tilde{M}_{jk}(\mathbf{x}(s,t))\kappa[\theta_k(\mathbf{x}(s,t)) - z_k(s,t)]] + \sqrt{2\overline{\gamma}k_B\overline{T}}\eta_z(t) + \lambda(s,t)\frac{\partial z_j}{\partial s} \quad (2)$$

Equation 1 is Langevin dynamics of the atomic variables subject to interatomic potential U, friction γ, and white noise η_x. The final term in eq 1 represents harmonic forces that restrain instaneously CV's $\theta(x)$ to the auxiliary variables z with a spring constanct κ. Equation 2 represents Langevin dynamics evolution of z's. Here, $\tilde{M}_{jk}(x(s,t))$ is a metric tensor mapping distances on the manifold of atomic coordinates to the manifold of CV space, and $\lambda(s,t)(\partial z_j/\partial s)$ represents a "reparametrization" force tangent to the string that is sufficient to maintain equidistant images along the string. The second term in eq 2 represents a thermostat at temperature $\overline{T}$ that in our implementation is set to $T = \overline{T} = 310$ K (n.b., setting $\overline{T} = 0$ gives steepest descent dynamics for z's and represents zero temperature string method whereas setting $\overline{T} > 0$ represents finite temperature string method[11]). The dynamics of eqs 1 and 2 converges the composite set of variables z that define the string to the principal curve. Straightforward thermodynamic integration using forces measured on a static string is then used to compute the potential of mean force along the principal curve.

Specific Protocols. To begin, we generated three fully solvated systems, which were then equilibrated under constant temperature conditions for 50 ns with standard MD: (i) nonprotonated normal β2M (PDB code 2XKS); (ii) protonated normal β2M (PDB code 2XKS modified by protonation of HIS residues); (iii) protonated amyloidogenic ΔN6-β2-microglobulin (ΔN6-β2M) (PDB code 2XKU). A requirement of OFSM is that all of the images be topologically equivalent (i.e., they contain the same number of atoms and the same bonding configuration). However, the 2XKS and 2XKU crystal structures are not topologically identical because 2XKU has a truncated N-terminus. To generate topologically identical images for OFSM, targeted molecukar dynamics (TMD) was used to generate an amyloidogenic conformation of the full-length β2M by driving the equilibrated protonated and nonprotonated 2XKS structures into alignment with the equilibrated 2XKU structure. In this way, we create end points that are conformationally nearly identical to 2XKU, yet topologically equivalent to 2XKS (1.1 Å and 1.0 Å root-mean-square deviation (rmsd) for protonated and nonprotonated 2XKS, respectively, compared to 2XKU). Two initial strings of 27 images were then generated by selecting conformations from the TMD trajectories.

For each string, the 27 images were arranged into a $3 \times 3 \times 3$ cube to construct a single simulation system representing all string images. Each of these composite systems was approximately 450 000 atoms and 4×10^6 Å^3 in volume. Solvent water was explicitly included using the TIP3P model[19] and the system was solvated to physiological ion concentration using VMD.[20] We used the CHARMM 22 force field with CMAP (cross term map) corrections for proteins.[21,22] An 8.5 Å switching distance and 9.0 Å cutoff distance were used for nonbonded interactions. Long-range electrostatics were handled using particle-mesh Ewald summation with a 1 Å grid spacing. Equilibration of the composite system was performed using 50 000 steps of minimization with fixed protein atoms followed by 1000 steps of minimization with no fixed atoms. The system was then slowly heated using MD with a 1 fs time step under NPT conditions to equilibrate cell size prior to OFSM simulation.

Our set of CVs is 21-dimensional and is composed of (a) 20 distinct state-determining inter-α-carbon (Cα) distances and (b) the dihedral angle of the prolyl peptide bond between His31 and Pro32. The residue pairs defining the twenty inter-Cα distances were selected on the basis of a statistical comparison of each pair of structures (normal and amyloidogenic), following the approach we recently developed for the acetylcholine binding protein.[18] These are (22,32); (11,32); (63,86); (86,92); (8,16); (32,88); (18,70); (11,63); (21,63); (27,86); (33,61); (31,39); (56,60); (18,42); (10,17); (18,94);

dx.doi.org/10.1021/jp304805v | J. Phys. Chem. B 2012, 116, 9371−9375

(32,78); (17,86); (31,48); and (8,86). These are illustrated in Figure 2.

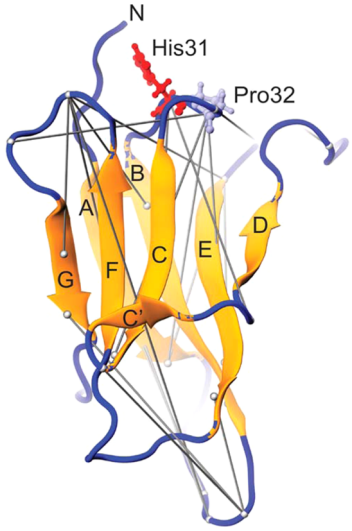


Figure 2. Amyloidogenic form of the β2M structure with labeled β strands. Pro32 and His31 are rendered in ball and stick representation. Inter-Cα distance CVs are shown as gray bonds between Cα. Here the pairs of the inter-Cα distance CVs are composed of residues 22,32; 11,32; 63,86; 86,92; 8,16; 32,88; 18,70; 11,63; 21,63; 27,86; 33,61; 31,39; 56,60; 18,42; 10,17; 18,94; 32,78; 17,86; 31,48; and 8,86.

We implemented OFSM using the NAMD 2.7[23] MD engine patched with a modified PLUMED[24] plug-in. The string was initially equilibrated with fixed ends (by restraining the CVs of the first and last image at their initial values) to prevent potentially unstable end states drifting away from their TMD-generated shape (see Table 1 for details). After 10.25 ns of OFSM simulation, the ends of the string were released followed by 3 ns of additional equilibration to ensure that steady state had been achieved.

To compute the potential of mean force along each converged string, update of the z variables is disabled (by setting $1/\bar{\gamma} = 0$) and a time average (denoted by $\langle ... \rangle$) of the restraining force is calculated for each CV in every image along the string:

Table 1. OFSM[a]

time (ns)	κ dihedral (kcal/mol/rad^2)	κ distance (kcal/mol/Å^2)	$10^{-6}\bar{\gamma}$ (fs^{-1})
0–500	10	10	0.5
0.5–2	10	10	1.0
2–3.5	15	15	2.0
3.5–3.75	100	100	10
3.75–4.75	1000	1000	20
4.75–6.25	1000	1000	40
6.25–8.25 (NPT)	1000	1000	∞
8.25–10.25	1000	1000	40
10.25–13.25[b]	1000	1000	40
500 ps force accumulation	1000	1000	∞

[a] $\bar{\gamma} = \infty$ implies that updates to z variables have been disabled for thermodynamic integration. *NVT* ensemble used for all segments except where noted to re-equilibrate cell size. [b] At this point, restraints on the CVs for the first and last image are released.

$$\frac{dF(z_\alpha(s))}{dz_\alpha} = \kappa\langle\theta_\alpha(x(s,t)) - z_\alpha(s,t)\rangle \quad (3)$$

These are then used in thermodynamic integration to compute the PMF:

$$F(z(s)) - F(z(0)) = \int_0^s \sum_{\alpha=1}^{m} \frac{dF(z_\alpha(s))}{dz_\alpha}\frac{dz_\alpha(\alpha)}{d\alpha}\,ds \quad (4)$$

RESULTS AND DISCUSSION

The major results of our OFSM calculations are the identification of princpal reaction pathways through a multi-dimensional CV-space, and subsequent (TI) calculations provide the energetics along those pathways. The variables that constitute our chosen CV space are illustrated in Figure 2; they include a set of state-discriminating inter-Cα distances and the His31–Pro32 peptide torsion angle. In the following, we first discuss the energetics of the normal-to-amyloidogenic isomerization of β2M, followed by a detailed discription of the mechanisms observed along the princpal reaction pathways.

Energetics of β2M Isomerization. Figure 3 shows the free energy as a function of distance along the princpal pathways for

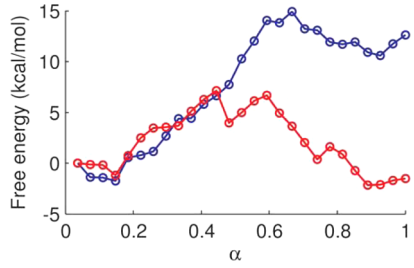


Figure 3. Free energy profiles from string-method simulations of the normal-to-amyloidgenic conformation of nonprotonated (blue) and protonated (red) β2M.

protonated (red) and nonprotonated (blue) β2M. Note that we are not reporting free energy difference *between* the protonated and nonprotonated systems, but instead free energy differences along the normal-to-amyloidgenic transition for *each* system independently. First, these data show that the amyloidogenic conformation is less stable by 12.6 kcal/mol than is the normal conformation when unprotonated, whereas the amyloidogenic conformation is more stable by 2.2 kcal/mol when His31 and His84 are protonated. Given these free energy differences, a simple thermodynamic equilibrium argument would therefore predict amyloidogenic β2M at concentrations in the subppb range under nonprotonating conditions, and at a few percent under protonating conditions. Therefore, it is not surprising that experimental evidence demonstrates nonprotonated wild-type β2M does not form amyloid fibrils in vitro.[6,7,12]

Second, the free energy barrier for transition to the amyloidogenic conformation for the unprotonated β2M is 14.9 kcal/mol, whereas in the protonated case, the barrier is reduced to to 7.1 kcal/mol. The experimentally determined free energy barrier for cis–trans isomerization of X-Pro peptide bonds in small proline-containing molecules ranges between about 19 and 21 kcal/mol.[13] This suggests an autocatalytic process in β2M that lowers the free energy barrier for cis–trans isomerization, especially in the case of low pH. Indeed, lowering pH to values below the pK_a of histidine results in relatively rapid amyloid growth of β2M in vitro.[6,7,12]

dx.doi.org/10.1021/jp304805v | J. Phys. Chem. B 2012, 116, 9371–9375

Mechanism of β2M Isomerization. Because the energetics of our pathways seem consistent with experiments with respect to the effects of histidine protonation, we turn now to our observations on the detailed mechanism of the transition. Selected intermediate conformations of the normal-to-amyloidogenic isomerization for protonated and nonprotonated β2M from our OFSM calculations appear in Figure 4. First, we note

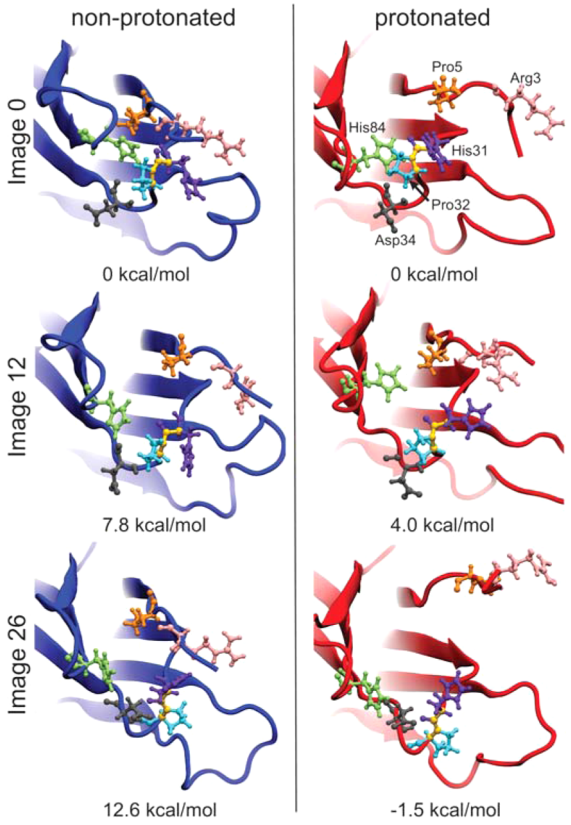


Figure 4. Normal to amyloidogenic reaction for nonprotonated (left) and protonated (right) β2M, illustrated using snapshots of string images. Important residues are shown in ball-and-stick rendering: Pro32, cyan; His84, green; His31, purple; Asp34, black; Arg3, pink; Pro5, orange (N.b., Arg3 and Pro5 are located on the N-terminus).

that models of the nonprotonated and protonated normal state differ significantly, primarily in the placement of the N-terminus. In the nonprotonated normal state, His84 and His31 are attracted to Pro5 and Arg3, respectively. These interactions anchor the position of the N-terminus proximal to the hydrophobic pocket formed by the F-G loop, the B-C loop, and the N-terminus itself. Upon histidine protonation in the normal state, however, equilibration results in the N-terminus breaking away from the pocket. This favorably exposes the protonated His31 and His84 side-chains.

The His31−Pro32 peptide bond appears to be stabilized in cis by interaction of Pro32 with His84. In both the protonated and nonprotonated case, rotation of this bond to trans requires breaking this interaction. We observe the loss of this interaction when the His31−Pro32 peptide bond rotates past about 100° (0 is cis and 180° is trans), causing the Pro32 side chain to point away from His84. Figure 5A shows the average potential energy of interaction between Pro32 and His84 as a function of distance along the principal pathway α from normal (α = 0) to amyloidogenic (α = 1). The transition requires this energy be sacrificed, and the interaction is slightly stronger when His84 is

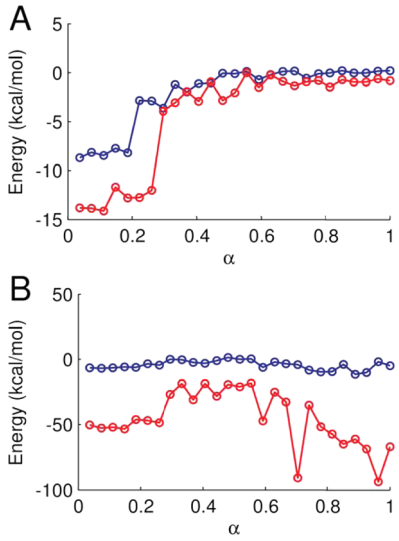


Figure 5. Average potential energy of interaction (A) between Pro32 and His84, and (B) between Asp34 and His84 as functions of position along princpal pathway α for nonprotonated (blue) and protonated (red) β2M.

protonated. As the His31−Pro32 peptide bond rotates further toward trans, protonated His84 begins to interact strongly with the anionic Asp34 side chain (Figure 5B). This interaction appears to promote the rotation of His31−Pro32 peptide bond into trans because it pulls the B−C loop inward, which inhibits rotation of the bond back into cis. In the case of nonprotonated His84, there is no interaction with Asp34 and passage into the trans state is not facilitated.

This finding provides a mechanistic explanation of why acidic environments (that protonate histidine), or truncation of the N-terminus, potentiate fibril formation: both circumstances disrupt the hydrophobic pocket that stabilizes the *cis* proline. In addition, experimental results have shown that increasing acidity further accelerates fibril formation.[12] This result is also consistent with our findings because the trans-stabilizing interaction of protonated His84 and Asp34 does not occur for the nonprotonated case. This mechanism would be increasingly favored as the environment becomes more acidic, leading to a larger population of amyloidogenic β2M molecules in solution.

Future investigation using mutagenesis could be used to further support our predictions. It has already been shown that mutation of Pro5 to Gly stabilizes the amyloidogenic state,[12] presumably by facilitating dislocation of the N-terminus that either precedes or stabilizes protonated histidines. As an additional example, mutation of Asp34 to a neutral or positively charged residue could help verify that preventing interaction with protonated His84 destabilizes the amyloidogenic state compared to WT-β2M. Alternatively, mutation of Pro5 and/or Arg3 to a negatively charged residue would enhance interaction with protonated His84 and His31. This mutation may stabilize the position of the N-terminus in the nonamyloidogenic state for the case of protonated histidine and slow or prevent formation of the amyloidogenic state. Future work should also address the nature of how β2M evidently catalyzes its own isomerization into the growth of fibrils.

dx.doi.org/10.1021/jp304805v | J. Phys. Chem. B 2012, 116, 9371−9375

CONCLUSION

We have used all-atom OFSM calculations to simulate the transition from a normal to an amyloidogenic form of β2M under both neutral and acidic conditions. Our main conclusions follow. (i) Protonation of histidine residues in β2M stabilizes a conformation consistent with the experimentally determined amyloidogenic state of the ΔN6 mutant. (ii) This state is formed in a two-step process: first the N-terminus breaks away disrupting the hydrophobic pocket, and then rotation of the His31-Pro32 peptide bond to a trans conformation is facilitated by interaction of the protonated His84 side chain with Asp34.

More generally, this contribution illustrates a convenient and efficient method to implement OFSM calculations in all-atom, explicit-solvent protein simulations for the study of the energetics of conformational transitions. This bodes well for the method's eventual use in more complicated settings involving protein−protein interactions and allosterics.

AUTHOR INFORMATION

Corresponding Author

*E-mail: cfabrams@drexel.edu. Phone: 1 (215) 895-2231. Fax: 1 (215) 895-5837.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Computational resources were provided by the DRACO GPU cluster in the Department of Physics at Drexel University (National Science Foundation AST-0959884). This work was funded in part by the Pennsylvania Department of Health, Commonwealth Universal Research Enhancement C.URE Award to Drexel University, and the National Institutes of Health (NIGMS) under grant 1R01GM100472. We are grateful to L. Maragliano for a critical assessment of the manuscript.

REFERENCES

(1) Gidalevitz, T.; Ben-Zvi, A.; Ho, K. H.; Brignull, H. R.; Morimoto, R. I. *Science* **2006**, *311*, 1471−4.

(2) Rubin, E., Gorstein, F., Schwarting, R., Strayer, D., Eds. *Rubin's Pathology: Clinicopathologic Foundations of Medicine*, 4th ed.; Lippincott Williams & Wilkins: Boston, 2004.

(3) Gussow, D.; Rein, R.; Ginjaar, I.; Hochstenbach, F.; Seemann, G.; Kottman, A.; Ploegh, H. *J. Immunol.* **1987**, *139*, 3132−3138.

(4) Gorevic, P. D.; Munoz, P. C.; Casey, T. T.; DiRaimondo, C. R.; Stone, W. J.; Prelli, F. C.; Rodrigues, M. M.; Poulik, M. D.; Frangione, B. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 7908−12.

(5) Padrini, R.; Canova, C.; Conz, P.; Mancini, E.; Rizzioli, E.; Santoro, A. *Kidney Int.* **2005**, *68*, 2331−7.

(6) Eichner, T.; Kalverda, A. P.; Thompson, G. S.; Homans, S. W.; Radford, S. E. *Mol. Cell* **2011**, *41*, 161−72.

(7) Esposito, G.; Ricagno, S.; Corazza, A.; Rennella, E.; Gümral, D.; Mimmi, M.; Betto, E.; Pucillo, C.; Fogolari, F.; Viglino, P.; et al. *J. Mol. Biol.* **2008**, *378*, 887−97.

(8) Trinh, C. H.; Smith, D. P.; Kalverda, A. P.; Phillips, S. E. V.; Radford, S. E. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9771−6.

(9) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2007**, *446*, 182−190.

(10) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 24106.

(11) E, W.; Ren, W.; Vanden-Eijnden, E. *J. Phys. Chem. B* **2005**, *109*, 6688−93.

(12) Eichner, T.; Radford, S. E. *J. Mol. Biol.* **2009**, *386*, 1312−1326.

(13) Cheng, H. N.; Bovey, F. A.; Laboratories, B.; Murray, H. *Biopolymers* **1977**, *16*, 1465−1472.

(14) Vanden-Eijnden, E.; Venturoli, M. *J. Chem. Phys.* **2009**, *130*, 194103.

(15) Ovchinnikov, V.; Karplus, M.; Vanden-Eijnden, E. *J. Chem. Phys.* **2011**, *134*, 085103.

(16) Zhu, F.; Hummer, G. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 19814.

(17) Vashisth, H.; Maragliano, L.; Abrams, C. F. *Biophys. J.* **2012**, *102*, 1979−1987.

(18) Stober, S.; Abrams, C. F. *Protein Sci.* **2012**, *21*, 307−317.

(19) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.

(20) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33−38.

(21) MacKerell, A., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R., Jr.; Evanseck, J.; Field, M.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(22) Mackerell, A. D.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400−15.

(23) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−802.

(24) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.