Chemical Physics Letters 475 (2009) 111–115

Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Molecular Dynamics insight into the role of tertiary (foldon) interactions on unfolding in Cytochrome c

M.Y. Tsai [a,b,*], A.N. Morozov [a,c], K.Y. Chu [d], S.H. Lin [a,e]

[a] National Chiao Tung University, 1001 Ta Hsuen Road, Hsinchu, Taiwan, ROC
[b] Department of Chemistry, National Taiwan University, Taipei 106, Taiwan, ROC
[c] Department of Chemistry and Biochemistry, Florida International University, 11200 S.W. 8th Street Miami, FL 33199, USA
[d] Department of Chemistry, National Taiwan Normal University, 88 Section 4, Tingchow Road, Taipei, Taiwan, ROC
[e] Institute of Atomic and Molecular Sciences, Academia Sinica, P.O. Box 23-166, Taipei, Taiwan, ROC

ARTICLE INFO

Article history:
Received 28 March 2009
In final form 12 May 2009
Available online 18 May 2009

ABSTRACT

Hydrogen exchange experiments reveal the existence of a foldon structure in Cytochrome c. We performed constant temperature Molecular Dynamics simulations of Cytochrome c to investigate the macroscopic behavior of the foldons. The results showed that the tertiary interactions in Cytochrome c play an important role in the stabilization of the N- and C-terminal helices. We also observed the macroscopic behavior of a black foldon in the process of thermal unfolding. Our results support that Cytochrome c folding occurs in accordance with the classical pathway concept of Levinthal.

1. Introduction

The foldon structure of Cytochrome c (Cyt c) was observed by Englander's group using hydrogen exchange (HX) experiments [1] and by our group using spectrometric experiments including UV absorption, circular dichroism (CD), fluorescence and small angle X-ray scattering (SAXS) [2]. These experiments showed that Cyt c consists of several highly cooperative groups, which includes structurally related amino acid residues (ex. α-helix, β-sheet, loop etc.). The residues of the *i*th group show approximately the same thermodynamic [1,3,4] and kinetic [5,6] behaviors, so one can describe them as a macro-unit that Englander called a foldon. In Cyt c, five foldons were identified by their stability. The blue foldon (N-ter helix, 1–16th res. and C-ter helix, 88–104th res.) is the most stable one, and it is followed by green (one helix, 61–70th res. and one loop, 17–37th res.), yellow (β-sheet, 38–40th and 58–60th res.), red (one loop, 71–87th res.) and black foldons (one infrared, or nested-yellow loop, 41–57th res.), respectively[1]. Here the black foldon is used for later graphical convenience (see Fig. 1). Such behavior of Cyt c allows a coarse-grained approach in theoretical modeling of Cyt c behavior [2,7]. Remarkably, these simple models provide a consistent explanation of the non two-state thermodynamics and multi-exponential kinetics of Cyt c [2,7] as a manifestation of the underlying foldon structure. Nevertheless, protein folding is governed by a large number of heterogeneous molecular interactions in a multidimensional, configurational space. Thus, there is a need to fill the gap between the micro and foldon levels of theoretical description of Cyt c. In this Letter, we present our early efforts using Molecular Dynamics (MD) simulations to detect foldon's cooperative behavior from microscopic molecular motions. Such a connection should prove useful for understanding the mechanism of Cyt c folding–unfolding in accordance with the classical pathway concept of Levinthal [8]. At this point it should be noted that the classical concept of Levinthal does not contradict the modern energy landscape framework. Rather the classical concept is the limiting case of the modern energy landscape theory, though each of them has diverse implications for protein-folding mechanisms.

The classical pathway concept suggests the existence of discrete intermediates; protein molecules eventually fold into the native state by visiting intermediates in the specific pathways. Although much effort has been made using energy landscape theory [9], which is based on the attractive and intuitive principle of minimal frustration that allows proteins to fold through a large number of intermediates (conformational ensemble) along a funnel-like pathway (non-specific pathway), experimental studies support the classical scenario for Cyt c [1] and some other proteins [10,11]. Kinetic spectroscopy measurements reveal the existence of intermediates upon folding of Cyt c [12–14]. It was not until the HX experiments performed by Englander's group, that detailed structural information about the intermediates was known [1,3–6,15–17]. Furthermore, their results also demonstrated that five foldons are sequentially put together to form the intermediates all the way to the native state, which is consistent with the classical Levinthal folding picture. This finding attracted biological physicists to design phenomenological models of Cyt c to show the classical scenario in protein folding [7,18]. While much research has been devoted to

* Corresponding author. Address: Department of Chemistry, National Taiwan University, Taipei 106, Taiwan, ROC.
E-mail address: victorleaf@gmail.com (M.Y. Tsai).

[1] For interpretation of color in Fig. 1, the reader is referred to the web version of this article.

0009-2614/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cplett.2009.05.027

Fig. 1. The structure of Cyt c (1HRC). Each foldon is colored and drawn in ribbon style. Heme group is presented in licorice style with a centered, red ball denoting the Fe(III) ion. The blue foldon's N-helix spans residues 1 to 16, and its C-helix spans residues 88 to 104. The green foldon's helix spans residues 61 to 70, and its loop spans residues 17 to 37. The yellow foldon spans residues 38 to 40 and 58 to 60. The red foldon spans residues 71 to 87. Finally, the black foldon spans residues 41 to 57. The transparent parts (only the yellow and black foldons are shown) show the macroscopic movements observed at 400 K (structure snapshot at time step 12 000 ps). (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

the folding mechanism of Cyt c, it is still a question of ongoing discussion.

The results presented here confirm that tertiary interactions of foldon units play a determining role in Cyt c folding. First of all, we found that foldons are unstable if removed from the context of the native Cyt c surrounding. Secondly, we observed the collective movement of the residues of the black foldon out of the Cyt c matrix in the process of temperature unfolding. Both of these results are in line with thermodynamically observed sequential folding and validate the use of coarse-grained models with foldons as interacting units. Since those models provided a good explanation of the thermodynamic and kinetic behavior of Cyt c [12–14], the results presented here should prove useful for understanding the basic mechanism of protein folding.

One is interested in answering the question of 'How can the Cyt c molecule quickly search all possible random trajectories along a unique pathway?' The plausible answer is the reduction of multi-dimensional space due to protein degrees of freedom into foldon subspaces; for Cyt c this was proposed for the first time by Bai et al. [1]. They hypothesized that each folding step of Cyt c requires exploring the reduced configurational space that results in sequential folding. This idea has been examined by Morozov et al. [7]. They found that taking into account five foldon units gives the folding time-scale of Cyt c around 1 s, which is a biologically sensible time-scale [7], provided that the time-scale of the microscopic backbone motions are around 50 ps [19,20]. Thus, sequential organization of foldon structures allows one to successfully resolve Levinthal's paradox. It should be noted, that Schlag et al. reported that the reduction of entropic barriers by solvent effect for microscopic backbone motions is needed as well to reach the 50 ps time scale, which is much shorter than that of the related motions in gas phase [21].

2. Methods

Classical MD simulations were carried out using the NAMD program, version 2.6 [22] with the CHARMM 27 force field [23]. A new charge set and CHARMM parameters for the Heme group in oxidized Cyt c was given by Luthey-Schulten's group [24]. These new charges and parameters are based on ab initio density functional theory (DFT) quantum calculation at the B3LYP/6-31G* level, which is supposed to provide more accurate simulations for Heme proteins. The initial structure of Cyt c was taken from the high-resolution X-ray structure (PDB code: 1HRC) [25]. The construction of the system, visualization and data analysis was conducted using VMD molecular visualization program [26]. Periodic boundary conditions together with an explicit water model (TIP3P) [27] was used. The system contained 14 859 atoms including 9257 water molecules, 1 Na^+ ion and 8 Cl^- ions. The purpose of adding the ions is to neutralize the side chain charges under physiological conditions (pH 7), which results in a concentration of 0.1 M (defined as $(\#Na^- + \#Cl^-)$/volume). An NpT ensemble was performed using the Langevin piston method [28] for keeping pressure at 1 atm with an oscillation period coefficient and a damping coefficient of 100 fs^{-1} and 50 fs^{-1}, respectively. Similarly, Langevin bath was taken for the constant temperature with a damping coefficient of 5 ps^{-1}. Moreover, the SHAKE algorithm [29] and a 2 fs integration time step were employed in all simulations. Finally, full electrostatics were calculated using the Particle-Mesh Ewald approach [30].

In order to demonstrate the dependence of tertiary interactions on the stability of foldons, two simulation sets were prepared. The first set was constructed to simulate the dynamics of the whole Cyt c molecule, which is aimed to provide the full tertiary interactions for each foldon. The initial structure of Cyt c (1HRC) was minimized for 2000 steps, followed by equilibration for 100 ps under NpT ensemble conditions. Three trajectories were obtained by

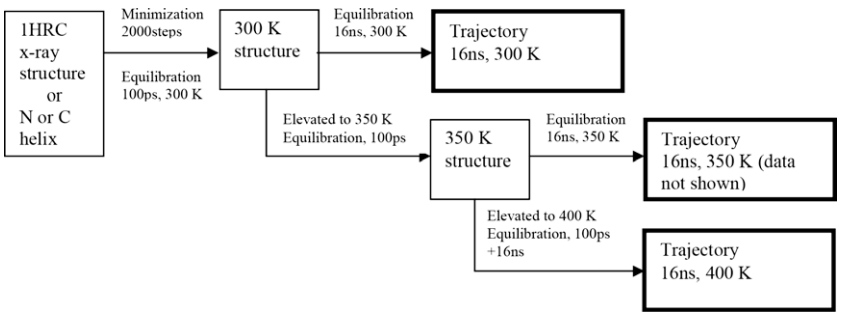

Fig. 2. The flow chart of the MD simulations for the whole Cyt c, single N-helix and C-helix.

performing MD simulations at temperatures of 300, 350 and 400 K. See Fig. 2 for a simulation flow chart. The other set was constructed to simulate the foldon stability without tertiary interactions. N-ter and C-ter helices were taken out of Cyt c. The MD simulation of each helix was prepared and performed with the same simulation condition mentioned above. See Fig. 2. Root mean square deviation (RMSD, relative to X-ray structure) and center-of-mass (COM) of the foldons (ex. α-helix, β-sheet, loop etc.) were calculated. The formulations of RMSD and COM are given by

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N_h}(\vec{r}_i - \vec{r}_{i,c})^2}{N_h}} \quad (1)$$

and

$$COM = \frac{\sum_{i=1}^{N} m_i \vec{r}_i}{\sum_{i=1}^{N} m_i}, \quad (2)$$

respectively.

N_h is the number of heavy atoms (C, N, O atom) in the protein backbone. N is the number of total atoms (including H atom). The $\vec{r}_i$ is the position vector of the ith atom during the simulation, and $\vec{r}_{i,c}$ denotes the position vector of the ith atom in the X-ray structure (1HRC). m_i is the mass of the ith atom.

3. Results and discussion

3.1. Stability of N and C helices due to tertiary interactions

The N and C helices are found to be the most stable foldons in Cyt c [16]. The stability of the secondary structure for the blue foldon with and without Cyt c surrounding has been examined. The RMSD value (relative to X-ray structure) for the blue foldon is given in Fig. 3. Overall, the fluctuation and absolute value of the RMSD for each helix at a higher temperature (400 K) showed a tendency to increase. It is obvious that this kind of structural fluctuation is due to temperature-induced local conformational fluctuations. At 300 K, the RMSD value of both the single N and C helices (RMSD 2–3), which neglects the Cyt c surrounding, deviated far from the RMSD value of the N and C helices in the native Cyt c surrounding (RMSD 0.5–1) over the simulation time studied. The structural difference between the helix (N or C) with and with-

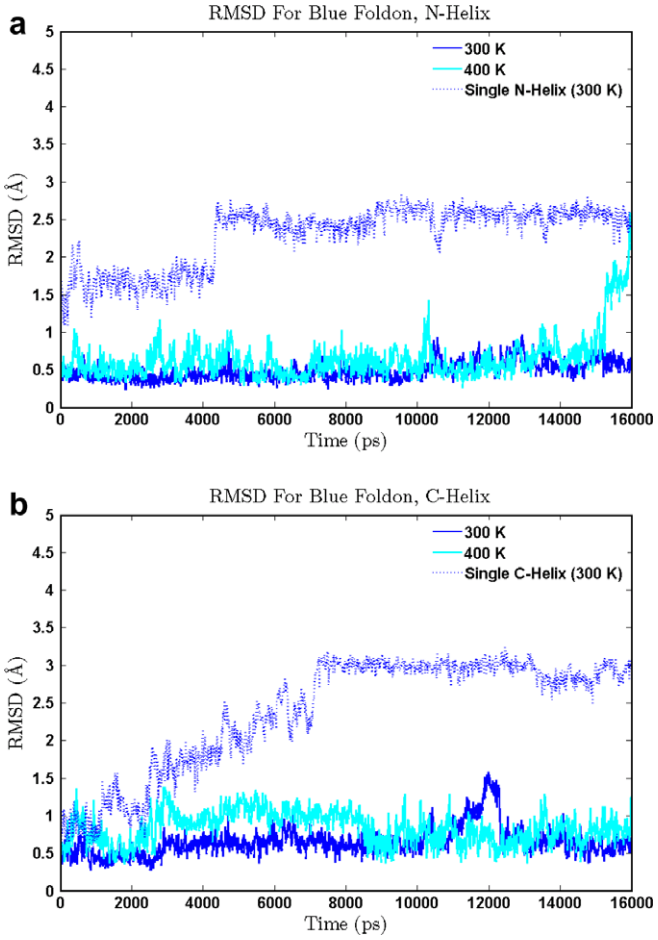

Fig. 3. RMSD for the blue foldon (relative to the X-ray structure, 1HRC). (a) To eliminate the frayed end effect, only residues 3 to 15 of the N-helix were included to compute the RMSD. (b) Likewise, for the C-helix, only residues 89 to 103 were included. The solid lines indicate that the helices are in the Cyt c native surrounding; the dark blue and cyan represent 300 K and 400 K, respectively; the dotted line denotes the single helices taken out of the Cyt c native surrounding. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

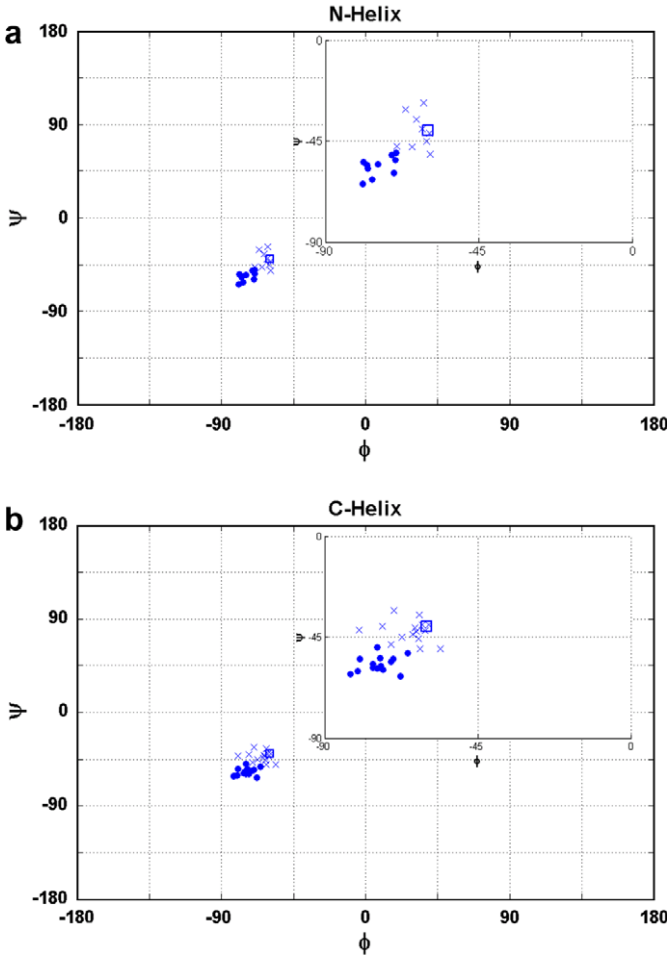

Fig. 4. Ramachandran plots generated from the N and C helices (not all residues are shown) (a) N-helix. Each solid circle denotes the average value of each (φ, ψ) pair for a single N-helix over the simulation time, 4000–16 000 ps, while each $\times$ symbol denotes the average value of each (φ, ψ) pair for the N-helix with Cyt c surrounding over the simulation time, 1–16 000 ps. (b) C-helix. The meaning of each solid circle and $\times$ symbol is the same as for (a) except it is for the C-helix, and the averages for the C-helix and the C-helix with Cyt c surrounding are taken over the simulation time 7000–16 000 ps and 1–16 000 ps, respectively. The squares shown in (a) and (b) denote the (φ, ψ) pair of the typical α-helix (-60, -40), and the inset plots in (a) and (b) are presented for a clearer demonstration.

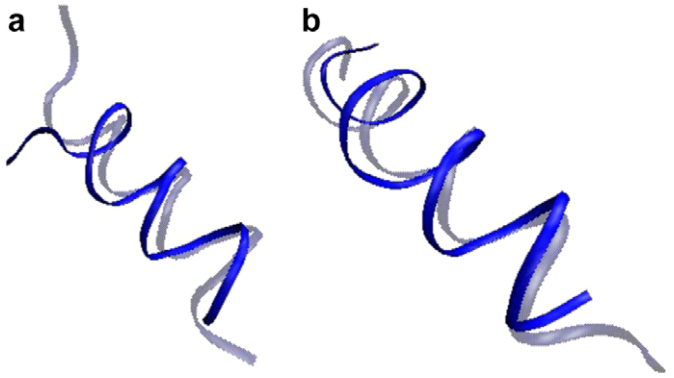

Fig. 5. Ribbon structure of a single N and C helix at 300 K. (a) The snapshot of a single N-helix. The dark blue ribbon describes the N-helix taken out of the Cyt c context at the simulation time step, 10 750 ps. (b) The snapshot of a single C-helix. The dark blue ribbon describes the C-helix taken out of the Cyt c native surrounding at the simulation time step, 13 740 ps. The transparent blue ribbons shown in (a) and (b) represent the N and C helices in the X-ray structure, respectively. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

out the native Cyt c surrounding can be further demonstrated by the Ramachandran plot given in Fig. 4. It was found that two groups (with and without Cyt c surrounding) are populated in different sub-regions in the helix domain. Interestingly, the single helix (N or C) without Cyt c surrounding populated in an area that is not typical for α-helices ($(\varphi, \psi) = (-60, -40)$), suggesting the existence of a local (meta-) stable state (whether it is 3_{10}-helix, π-helix or other species needs to be examined further). The corresponding structures in ribbon style are drawn in Fig. 5. We found that the helical (N or C) structures with and without the Cyt c surrounding do not overlap very well with each other; hence, it provides the evidence for a structural difference. These results can be explained by considering the importance of the native surroundings for the stability of the secondary structure. Moreover, our results are also consistent with those of Makhatadze et al. [31], who showed the stabilization of an α-helix in the process of ligand binding (viewed as a tertiary interaction). Although many studies have agreed on the stabilization of secondary structural elements by tertiary interactions [31,32], our results provide direct evidence from microscopic all-atom simulations that native tertiary interactions provide stability to secondary structures.

3.2. Macroscopic behavior of the black foldon and its role in the active site of Cyt c

Early on, the black foldon, originally named nested-yellow loop, was considered to be part of the yellow foldon [1]. It was then found that the nested-yellow loop reversibly folds and unfolds as a separate cooperative unit [33]. The black foldon has biologically functional activities and accounts for the first step (the least stable) in Cyt c unfolding. In this study, we observed the macroscopic behavior of this foldon. It was assumed that the macroscopic behavior of each foldon can be directly related to its center-of-mass (COM). The COM of each foldon was computed, and the macroscopic behavior of each foldon was measured at different temperatures by calculating the distance between each foldon's COM and the Heme's COM. The macroscopic behavior of each foldon is presented in Fig. 6. The results show that the distance of the blue, green, and red foldons were almost unchanged at 300 and 400 K during the simulation time; however, the distance of the yellow and black foldons started to increase at around 3000 ps, which resulted in the collective motions of the foldons. This large-scale collective motion of the black foldon was the first macroscopic behavior observed in our time-resolved MD simulations. It was then followed by the yellow foldon ~600 ps later (Fig. 1, transparent parts). This result is consistent with earlier findings suggesting that the black foldon is the last folding event in the sequential folding mechanism (or first folding event in unfolding). Meanwhile, we found that the yellow foldon formed before the black foldon (or unfolded after black foldon), which agrees with the sequential folding order proposed by Englander et al. [3]: blue → green → yellow → red → black. Although our results did not show the stepwise folding mechanism for all of the foldons, they do provide direct evidence of macroscopic foldon behavior from a microscopic approach.

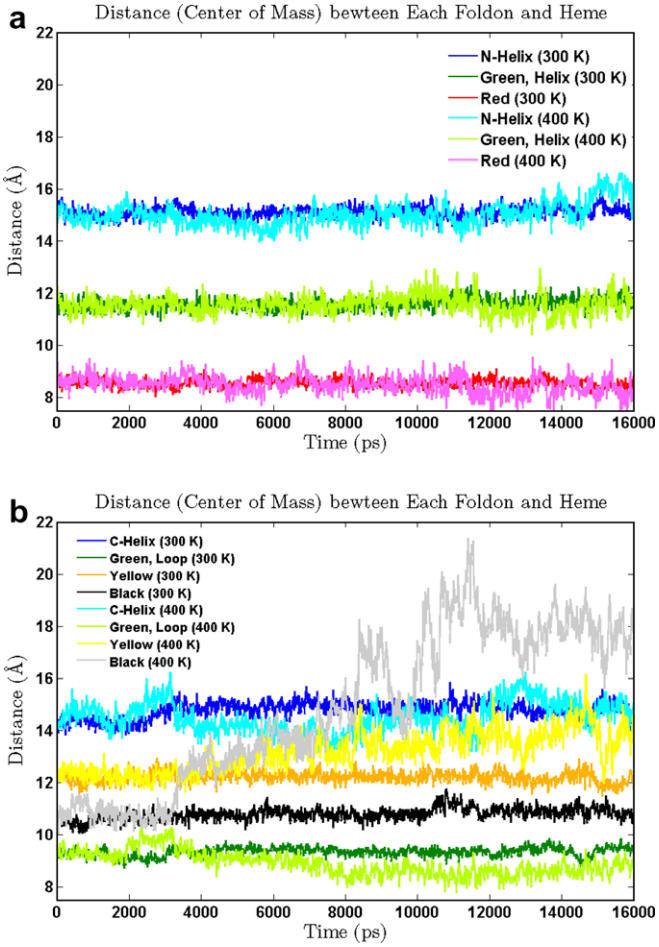

Fig. 6. The distance between the center-of-mass of each foldon and Heme. (a) Blue (N-helix), green (helix) and red foldons are colored by blue, dark green and red, respectively, at 300 K and cyan, light green and pink, respectively, at 400 K. (b) Blue (C-helix), green (loop), yellow and black foldons are blue, dark green, orange and black, respectively, at 300 K and cyan, light green, yellow and grey, respectively, at 400 K. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

In addition to determining the folding pathway, the black foldon of Cyt c probably serves as the gatekeeper to the active site of Cyt c. First, MD and NMR studies have been used to examine the flexible parts of Cyt c [34–36]. Recently, this flexibility was further confirmed by a SAXS experiment [2] (pH unfolding). Second, it was found that the de-protonation rate of buried Heme-related groups correlates well with the black foldon unfolding rate, suggesting that solvent species can access the Heme crevice [37]. In this study, our MD results showed the flexibility of the black foldon (see Fig. 6) and qualitatively demonstrated the opening of the black foldon, thereby exposing the Heme to the solvent (see Fig. 1). These results support the notion that the black foldon plays an important role in mediating the conformational gating reaction.

Acknowledgements

We are grateful to the Computer and Information Networking Center, National Taiwan University for their support of high-performance computing facilities. This research was supported by a Grant from the National Science Council (NSC), Taiwan.

References

[1] Y.W. Bai, T.R. Sosnick, L. Mayne, S.W. Englander, Science 269 (1995) 192.
[2] Y.J. Shiu et al., Biophys. J. 94 (2008) 4828.
[3] M.M.G. Krishna, H. Maity, J.N. Rumbley, Y. Lin, S.W. Englander, J. Mol. Biol. 359 (2006) 1410.
[4] L. Hoang, S. Bedard, M.M.G. Krishna, Y. Lin, S.W. Englander, Proc. Natl. Acad. Sci. USA 99 (2002) 12173.
[5] H. Roder, G.A. Elove, S.W. Englander, Nature 335 (1988) 700.
[6] M.M.G. Krishna, Y. Lin, L. Mayne, S.W. Englander, J. Mol. Biol. 334 (2003) 501.
[7] A.N. Morozov, Y.J. Shiu, C.T. Liang, M.Y. Tsai, S.H. Lin, J. Biol. Phys. 33 (2007) 255.
[8] C. Levinthal, J. Chim. Phys. Phys. Chim. Biol. 65 (1968) 44.
[9] J.D. Bryngelson, J.N. Onuchic, N.D. Socci, P.G. Wolynes, Proteins Struct. Funct. Genetics 21 (1995) 167.
[10] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Science 309 (2005) 2057.
[11] G. Baldini, F. Cannone, G. Chirico, M. Collini, B. Campanini, S. Bettati, A. Mozzarelli, Biophys. J. 92 (2007) 1724.
[12] T.Y. Tsong, Biochemistry 12 (1973) 2209.
[13] Y.P. Myer, J. Biol. Chem. 259 (1984) 6127.
[14] G.A. Elove, A.F. Chaffotte, H. Roder, M.E. Goldberg, Biochemistry 31 (1992) 6876.
[15] Y.J. Xu, L. Mayne, S.W. Englander, Nat. Struct. Biol. 5 (1998) 774.
[16] H. Maity, M. Maity, S.W. Englander, J. Mol. Biol. 343 (2004) 223.
[17] H. Maity, M. Maity, M.M.G. Krishna, L. Mayne, S.W. Englander, Proc. Natl. Acad. Sci. USA 102 (2005) 4741.
[18] P. Weinkam, C.H. Zong, P.G. Wolynes, Proc. Natl. Acad. Sci. USA 102 (2005) 12401.
[19] A.N. Morozov, S.H. Lin, J. Phys. Chem. B 110 (2006) 20555.
[20] W.S. Young, C.L. Brooks, J. Mol. Biol. 259 (1996) 560.
[21] S.Y. Sheu, H.L. Selzle, E.W. Schlag, D.Y. Yang, Chem. Phys. Lett. 462 (2008) 1.
[22] J.C. Phillips et al., J. Comput. Chem. 26 (2005) 1781.
[23] A.D. MacKerell et al., J. Phys. Chem. B 102 (1998) 3586.
[24] F. Autenrieth, E. Tajkhorshid, J. Baudry, Z. Luthey-Schulten, J. Comput. Chem. 25 (2004) 1613.
[25] G.W. Bushnell, G.V. Louie, G.D. Brayer, J. Mol. Biol. 214 (1990) 585.
[26] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. 14 (1996) 33.
[27] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926.
[28] S.E. Feller, Y.H. Zhang, R.W. Pastor, B.R. Brooks, J. Chem. Phys. 103 (1995) 4613.
[29] J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, J. Comput. Phys. 23 (1977) 327.
[30] T. Darden, D. York, L. Pedersen, J. Chem. Phys. 98 (1993) 10089.
[31] J.M. Richardson, M.M. Lopez, G.I. Makhatadze, Proc. Natl. Acad. Sci. USA 102 (2005) 1413.
[32] M. Sundaramoorthy, J. Terner, T.L. Poulos, Structure 3 (1995) 1367.
[33] M.M.G. Krishna, Y. Lin, J.N. Rumbley, S.W. Englander, J. Mol. Biol. 331 (2003) 29.
[34] A.E. Garcia, G. Hummer, Proteins Struct. Funct. Genetics 36 (1999) 175.
[35] J.S. Fetrow, S.M. Baxter, Biochemistry 38 (1999) 4480.
[36] S.M. Baxter, J.S. Fetrow, Biochemistry 38 (1999) 4493.
[37] L. Hoang, H. Maity, M.M.G. Krishna, Y. Lin, S.W. Englander, J. Mol. Biol. 331 (2003) 37.