Journal of Molecular Graphics and Modelling 29 (2010) 2–12

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Temperature-induced unfolding of epidermal growth factor (EGF): Insight from molecular dynamics simulation

Chunli Yan, Varun Pattani, James W. Tunnell**, Pengyu Ren*

Department of Biomedical Engineering, University of Texas, Austin, TX 78712, USA

ARTICLE INFO

Article history:
Received 18 January 2010
Received in revised form 15 March 2010
Accepted 29 March 2010
Available online 4 April 2010

Keywords:
Human epidermal growth factor (EGF)
Thermal denaturation
Replica exchange molecular dynamics simulations
Free-energy landscape
Molten globule
Transition state
Unfolding kinetics

ABSTRACT

Thermal disruption of protein structure and function is a potentially powerful therapeutic vehicle. With the emerging nanoparticle-targeting and femtosecond laser technology, it is possible to deliver heating locally to specific molecules. It is therefore important to understand how fast a protein can unfold or lose its function at high temperatures, such as near the water boiling point. In this study, the thermal damage of EGF was investigated by combining the replica exchange (136 replicas) and conventional molecular dynamics simulations. The REMD simulation was employed to rigorously explore the free-energy landscape of EGF unfolding. Interestingly, besides the native and unfolded states, we also observed a distinct molten globule (MG) state that retained substantial amount of native contacts. Based on the understanding that which the unfolding of EGF is a three-state process, we have examined the unfolding kinetics of EGF (N → MG and MG → D) with multiple 20-ns conventional MD simulations. The Arrhenius prefactors and activation energy barriers determined from the simulation are within the range of previously studied proteins. In contrast to the thermal damage of cells and tissues which take place on the time scale of seconds to hours at relatively low temperatures, the denaturation of proteins occur in nanoseconds when the temperature of heat bath approaches the boiling point.

1. Introduction

The central theme of structural biology is that the protein function depends on the structure. Thermal denaturation is a process in which a protein unfolds from its native structure upon heating. The energy threshold needs to be surpassed for the protein to reach the point where either the secondary and/or tertiary structure of the protein, will break apart, thus arriving at an unfolded state [1]. Protein denaturation has been found to be the primary cause of cell death [2], and this protein denaturation can be induced by a number of external stresses including the application of strong acids or bases, organic solvents, and heat. The latter mechanism provides the basis for a range of medical treatments that make up thermal therapy, whereby tissue is heated to high temperatures (above 50 °C), resulting in cell death. Traditionally, heat is generated via conduction, radiofrequency energy, or laser sources, and has found numerous applications in including the removal of birthmarks, tissue cautery during surgery, and treatment of cancer. Many of these techniques induce widespread tissue heating to large tissue volumes (e.g. whole tumor); however, recent advances in nanotechnology allow for confined heating of tissues on the nanometer scale. Photothermal therapy, a specific thermal therapy procedure, utilizes a foreign probe such as nanoparticles that has the property of changing light into heat therefore localizing the heating in the nanometer range [3]. This technique has become important in research for such fields as cancer therapy [4,5]. With this emerging technique, thermal damage at the molecular level becomes possible. It is therefore important to understand the kinetics of thermal denaturation at high temperature and short timescale to apply the local-heating to thermal therapy.

A number of different experimental approaches have been used to study the structure and energetics of unfolding, folding and transition states [6–10]. Molecular dynamics simulations have been widely used to examine details of the folding process of proteins that are otherwise not readily available experimentally [7,8,11–16]. Because protein folding and unfolding reactions typically occur on a time scale out of reach by the current computational power except for very small systems, significantly elevated temperatures have been used to induce unfolding. The results from such temperature-induced unfolding simulations are generally in reasonable agreement with experiments [11,12,17]. Replica exchange molecular dynamics (REMD) method [18] has been used to provides enhanced sampling of protein folding [15,17,19–22]. The free-energy landscapes of protein folding in water are believed to be at least partially rugged. At room temperature, protein systems

* Corresponding author. Tel.: +1 512 232 1832; fax: +1 512 471 0616.
** Co-corresponding author.
E-mail addresses: jtunnell@mail.utexas.edu (J.W. Tunnell), pren@mail.utexas.edu (P. Ren).

1093-3263/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2010.03.011

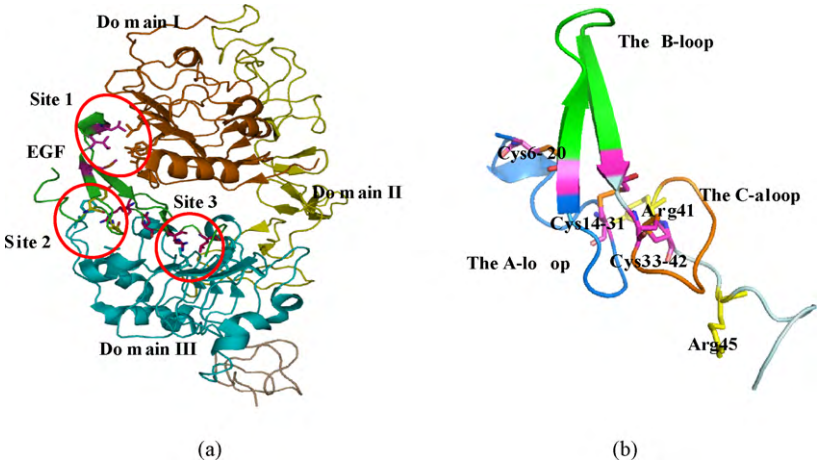

Fig. 1. EGF molecule used for the simulation. (a) Human EGF–EGFR complex (PDB entry: 1IVO). The EGF chains are colored in cyan. Three binding sites in the interface are outlined. (b) EGF. The three loops (A loop: blue, B loop: green, and C loop: brown) are indicated in colors. Arg41 of EGF interacts with the site 2 and the C-terminal region around Arg45 interacts with the site 3 are shown in sticks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

can be trapped in the local energy minima during conventional MD simulations. In REMD, a set of simulations are performed independently at different target temperatures, and exchanges are attempted according to the Metropolis criterion, therefore permitting random walks in the temperature space and escape from local energy traps. REMD has been successfully applied to the folding studies of β-hairpin [23], three-strand β-sheet [24], helical peptides [25], and small proteins [26,27].

A typical small protein can fold and unfold reversibly *in vitro*, undergoing a first-order (i.e., two-state) transition between its denatured state D and its "native" (i.e., folded) state N, depending on temperature, pH, denaturant concentration, etc. [28,29]. Under certain conditions, proteins can also exhibit a collapsed state with partial order known as the "molten globule" (MG) intermediate that possesses native secondary and tertiary characteristics but lacks well-packed side chains [30–32]. There is evidence for some proteins, e.g. apomyoglobin, that the molten globule present under equilibrium conditions resembles kinetic intermediates formed during protein folding [33,34]. Moreover, there have been suggestions that the molten globule and other non-native states of some proteins are functionally important and that the partially denatured species play a role in the transition to amyloid Fibrils [35,36]. It is therefore of great interest to understand the structural and thermodynamic properties of these non-native species which are not easily accessible by experiments as their NMR spectra have broad lines.

It has been shown that many epithelial cancer cells, which account for over 85% of all cancers, overexpress human epidermal growth factor receptor (EGFR), a clinically relevant biomarker [37]. This has lead to the development of several drugs that bind to EGFR and inhibit its activation. Recent examples include cetuximab for the treatment of colorectal cancer and herceptin for the treatment of breast cancer. As a result, EGF is commonly used as a mechanism to actively target these carcinoma cells, and numerous imaging and therapeutic agents are being developed to bind to the EGFR on the cell surface. EGF is a small mitogenic peptide of 53 amino acid residues that was first characterized and sequenced in 1972 [38]. EGF stimulates the growth of epidermal and epithelial cells by binding to the EGF receptor. The EGF-like domain has been found in a large number of functional unrelated proteins. Since the discovery of EGF, more than 300 EGF-like sequences have been identified, mostly as domains of larger proteins, ranging from urokinase, E-selectins, lipoprotein receptor, type-R transforming growth factor (TGF-R), heregulin, and tissue plasminogen activator [39]. These proteins have been associated with a diverse range of functions, including blood coagulation, fibrinolysis, neural development, and cell adhesion [39]. The corresponding receptor (EGFR) is a trans-membrane protein comprising 1186 residues. EGF induces dimerization of EGFR by binding to the extracellular region of the receptor, which leads to a 2:2 EGF–EGFR (Fig. 1a). A recent experimental structure has suggested a receptor-mediated mechanism for the receptor dimerization [40]. In this mechanism, an EGF molecule binds to the extracellular domains (L1 and L2) of a receptor molecule, which induces conformational changes of EGFR so that its dimerization surface in the S1 domain is exposed. After dimerization, the cytoplasmic tyrosine kinase domains in the two EGFR molecules are close enough for autophosphorylation, which will activate the intrinsic tyrosine kinase activity and trigger numerous downstream signaling events to regulate cell proliferation and differentiation [40].

EGF is stabilized by three disulfide bonds with the pairing pattern of (1–3, 2–4, and 5–6) (Cys6–Cys20, Cys14–Cys31, and Cys33–Cys42). The presently known experimental structures of human EGF are very similar [40–42]. Also, the EGFs from different species such as murine have essentially the same fold as human EGF [43–45]. The primary structure of EGF comprises three distinct loops (A, B, and C loops), which are divided by three disulfide bridges (Cys6–Cys20, Cys14–Cys31, and Cys33–Cys42), as shown in Fig. 1b. The *N*-terminal A-loop (residues 6–19) that is fastened by the disulfide bond Cys6–Cys20 contains an α-helical fragment. The so-called B-loop (residues 20–31) comprises a two-stranded anti parallel β-sheet, and the C-loop (residues 33–42) is constrained by the third disulfide bond Cys33–Cys42. The three loops in EGF interact extensively with three specific sites in EGFR (Fig. 1a). The B-loop (residues 20–31) of EGF interacts with the site 1 in domain I (Fig. 1a), the region containing the A-loop (residues 6–19); Arg41 of EGF interacts with the site 2 in domain III (Fig. 1a); the C-terminal region around Arg45 interacts with the site 3 in domain III. Fully reduced and denatured EGF is able to refold via disulfide oxidation to form the native conformation spontaneously and quantitatively. However, the folding mechanism of EGF displays unique properties that are not shared by other small disulfide proteins [46].

Although the EGF has been intensively investigated experimentally and theoretically [1,46,47], EGF folding/unfolding mechanism is not yet fully understood. Indeed, some experiments suggest a two-state behavior in chemical denaturation [48], while others point to the presence of intermediates [49]. In this work, we investigate the folding and unfolding of protein EGF using the explicit