Extreme Mechanics Letters 20 (2018) 112–124

Contents lists available at ScienceDirect

Extreme Mechanics Letters

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

Multiscale modeling of keratin, collagen, elastin and related human diseases: Perspectives from atomistic to coarse-grained molecular dynamics simulations

Jingjie Yeo [a,b], GangSeob Jung [a], Anna Tarakanova [a], Francisco J. Martín-Martínez [a], Zhao Qin [a], Yuan Cheng [b], Yong-Wei Zhang [b], Markus J. Buehler [a,*]

[a] Laboratory for Atomistic and Molecular Mechanics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[b] Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), Singapore 138632, Singapore

ARTICLE INFO

Article history:
Received 13 October 2017
Received in revised form 11 January 2018
Accepted 22 January 2018

Keywords:
Keratin
Collagen
Elastin
Molecular dynamics
Human disease phenotypes
Mechanics
Mechanobiology
Mutation
Defect
Failure
Elasticity

ABSTRACT

Scleroproteins are an important category of proteins within the human body that adopt filamentous, elongated conformations in contrast with typical globular proteins. These include keratin, collagen, and elastin, which often serve a common mechanical function in structural support of cells and tissues. Genetic mutations alter these proteins, disrupting their functions and causing diseases. Computational characterization of these mutations has proven to be extremely valuable in identifying the intricate structure–function relationships of scleroproteins from the molecular scale up, especially if combined with multiscale experimental analysis and the synthesis of model proteins to test specific structure–function relationships. In this work, we review numerous critical diseases that are related to keratin, collagen, and elastin, and through several case studies, we propose ways of extensively utilizing multiscale modeling, from atomistic to coarse-grained molecular dynamics simulations, to uncover the molecular origins for some of these diseases and to aid in the development of novel cures and therapies. As case studies, we examine the effects of the genetic disease Epidermolytic Hyperkeratosis (EHK) on the structure and aggregation of keratins 1 and 10; we propose models to understand the diseases of Osteogenesis Imperfecta (OI) and Alport syndrome (AS) that affect the mechanical and aggregation properties of collagen; and we develop atomistic molecular dynamics and elastic network models of elastin to determine the role of mutations in diseases such as Cutis Laxa and Supravalvular Aortic Stenosis on elastin's structure and molecular conformational motions and implications for assembly.

1. Introduction

Keratin, collagen, and elastin belong to an important category of proteins within the human body, known as fibrous proteins or scleroproteins, which adopt filamentous, elongated conformations in contrast with typical globular proteins [1,2]. They typically consist of tandemly repeating units of amino acid sequences, such as the heptad repeats in keratins [3], the G–X–Y motifs in collagen [4], and the V–P–G–V–G motifs in elastin [5]. These fibers provide structural support to cells and tissues due to their unique mechanical properties conferred by the repeating motifs and the resultant secondary structures, as well as their tendency to aggregate and form functional filaments [2]. Examples of these filaments include the helical coiled-coil bundles of keratins, the triple-helical fibers of collagen I, and the elastin microfibrils formed through coacervation and crosslinking. Given the remarkable importance of these proteins as structural elements, it is of no surprise that many disease phenotypes arise as a result of mutations in these proteins that lead to misaggregation that diminishes or completely disrupts their ability to form fibrillar structures, thereby dramatically reducing their mechanical functions. Due to the molecular nature of these disruptions, experimental characterization of diseases related to fibrous proteins has been significantly assisted by the proliferation of computational studies that are able to probe the intricate structure–function relationships of these scleroproteins on the molecular scale. Here, we review the molecular origins of numerous critical diseases that are related to keratin, collagen, and elastin, through several case studies. We propose ways of extensively utilizing multiscale modeling, from atomistic to coarse-grained molecular dynamics simulations, to uncover the underling molecular mechanisms for some of these diseases, helping to pave the way towards novel cures and therapies.

* Corresponding author.
E-mail address: mbuehler@MIT.EDU (M.J. Buehler).

https://doi.org/10.1016/j.eml.2018.01.009
2352-4316/

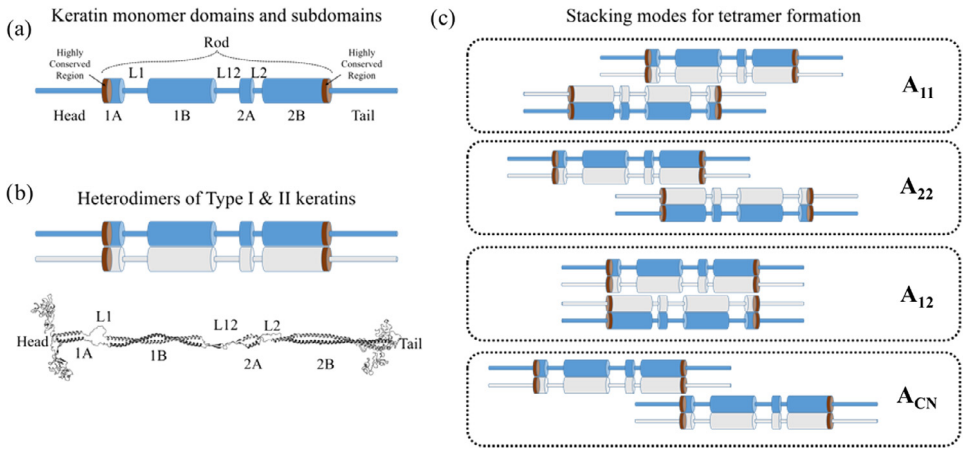

Fig. 1. (a) Schematics of a keratin monomer secondary structure at the molecular scale. (b) Parallel coiled-coil heterodimers of keratin Types I (blue) and II (gray) stylized from the proposed tertiary structure of heterodimeric coiled-coil keratins (bottom). (c) Different stacking modes of keratin tetramers during the formation of keratin filaments. Thin cylinders denote random coils, thick cylinders denote α-helices, and brown regions denote the highly-conserved regions of keratins. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

1.1. Keratin-related diseases

The human skin, a complex multilayered structure encompassing almost the entirety of the human body, provides crucial biological functions, such as regulating the egress of body fluids, preventing the entry of foreign substances, and maintaining the body's temperature [6,7]. To perform those functions, cells known as keratinocytes undergo cellular differentiation and outward migration, leading to the formation of the final layer of the skin, known as the stratum corneum (SC) [8,9], which consists of a dead flat layer of corneocytes. In healthy human skin, the primary structural components in the SC corneocytes are keratins [9], which are long strands of proteins with a high molecular weight, accounting for the major portion of the total protein content in the epidermis [8]. Importantly, many inherited skin disorders have been found to explicitly link with mutations in keratin molecular structures [10], and have been listed in the Human Intermediate Filament database [9].

Keratins in the human skin are rich in glutamic acid, serine, leucine, and glycine [9,11]. The primary structure of keratins consists of chains of these amino acids with high sequence identity although the exact number and sequence of amino acids vary slightly. The amino acid sequences also display a periodicity in the arrangement of the residues, consisting of heptad repeats of $(a-b-c-d-e-f-g)_n$ such that residues a and d are typically hydrophobic and e and g are charged [3]. The precise amino acid sequence exerts immense influence on the functions, structures, and properties of keratins [12]. The secondary structure of keratins consists of multiple domains and subdomains, as depicted in Fig. 1a. Each keratin monomer consists of head, rod, and tail domains. The head and tail domains are largely non-helical and globular structures with some content of β-turns [13,14]. The rod domain is further divided into subdomains of right-handed α-helical segments (1A, 1B, 2A, and 2B) connected by strands of non-helical linkers (L1, L12, and L2).

The tertiary structures of keratin filaments consist of Type I and II keratins assembled as left-handed coiled-coil heterodimers [15] as shown in Fig. 1b, where the hydrophobic residues a and d from each keratin monomer are wrapped up within the coil, forming an amphipathic structure with hydrophilic residues on the surface. Homodimers of either type are highly unstable, and thus degrade swiftly, while the stability of heterodimers helps maintain a balanced ratio between Types I and II [10]. This is due to Type I keratins being acidic, while Type II keratins are neutral to basic. Finally, unit length filaments (ULFs) are formed through bundling the heterodimers through four different modes of antiparallel stacking [16,17] as illustrated in Fig. 1c. These stacking modes were identified after crosslinks were induced between human K1 and K10 epidermal keratin chains and the linked residues were isolated through trypsin digestion [16]. The precise reason that ULFs have these stacking modes are still unknown but it is speculated that these can lead to spatially efficient packing into a lattice structure [18] through electrostatic interactions between charged amino acids on their surfaces [11].

Epidermolysis Bullosa Simplex (EBS), Epidermolytic Palmoplantar Keratoderma (EPPK), and Epidermolytic Hyperkeratosis (EHK) are examples of inherited skin disorders that are explicitly linked to keratin mutations [15]. EBS and EHK display remarkable similarities in that the most frequent mutations occur in the highly-conserved regions in segments 1A and 2B (also known as mutation 'hot-spots') as highlighted in Fig. 1a. For EBS, mutations occur in both basal keratin 5 and 14 (K5 and K14), while EHK is a result of mutations in both suprabasal keratin 1 and 10 (K1 and K10), which are the predominant keratins found in human SC [11]. As there are 28 Type I and 26 Type II intermediate filament proteins, nomenclature of these keratins have been standardized according to the three categories of (1) epithelial keratins/genes, (2) hair keratins/genes, and (3) keratin pseudogenes, taking into account historical precedence [9,19]. Moreover, one of the most frequently observed mutations in EHK is the substitution of the positively-charged arginine residue in position 156 of K10 by six other residues [9]. Three of these are polar residues: cysteine, histidine, and serine; while the other three are hydrophobic residues: proline, leucine, and glycine. For ease of reference, the substitutions are simplified as three-letter protein codes, namely p.ARG156CYS, p.ARG156HIS, p.ARG156SER, p.ARG156PRO, p.ARG156LEU, and p.ARG156GLY respectively. Characteristics of EHK include skin redness coupled with broad thickening of the topmost layer of the skin and severe blistering [17]. These mutations also result in keratins failing to assemble into filament structures, leading to malformed clumps of keratins [12,16].

The exact molecular mechanisms behind the inability of the keratins to assemble after mutations have not been fully understood. Through experimental observations, Steinert et al. [16] proposed a surface lattice model of K1 and K10 packing and suggested that there likely exists a recurring pattern of ten to eleven