FEBS Letters xxx (2015) xxx–xxx

journal homepage: www.FEBSLetters.org

Review

Order and disorder in intermediate filament proteins

Micha Kornreich [a], Ram Avinery [a], Eti Malka-Gibor [b], Adi Laser-Azogui [a], Roy Beck [a,*]

[a] School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel
[b] Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel

ARTICLE INFO

Article history:
Received 27 May 2015
Revised 21 July 2015
Accepted 22 July 2015
Available online xxxx

Edited by Wilhelm Just

Keywords:
Intermediate filament
Intrinsically disordered protein
Cytoskeleton
Polymer brush
Self-assembly

ABSTRACT

Intermediate filaments (IFs), important components of the cytoskeleton, provide a versatile, tunable network of self-assembled proteins. IF proteins contain three distinct domains: an α-helical structured rod domain, flanked by intrinsically disordered head and tail domains. Recent studies demonstrated the functional importance of the disordered domains, which differ in length and amino-acid sequence among the 70 different human IF genes. Here, we investigate the biophysical properties of the disordered domains, and review recent findings on the interactions between them. Our analysis highlights key components governing IF functional roles in the cytoskeleton, where the intrinsically disordered domains dictate protein–protein interactions, supramolecular assembly, and macro-scale order.

1. Introduction

The cytoskeleton of animal eukaryotes is composed of three principal components. Two, microfilaments (composed of actin subunits) and microtubules (composed of tubulin subunits), are well-characterized. The third cytoskeletal component, intermediate filaments (IFs), is found in cells of multicellular animal species. IF diameter is approximately 10 nm, intermediate in size between actin microfilaments and microtubules. Originally thought to be mechanical scaffolds that maintain the structural and mechanical integrity of cells and tissues, IFs were later found to participate in many important physiological functions, such as distribution of organelles, signal transduction, cell polarity and gene regulation [1–4]. Changes in IF expression have been correlated with various diseases, including cancer progression and acquisition of a metastatic phenotype [5–11].

Human IF proteins are encoded by 70 genes [12]. IFs are differentially expressed during embryonic development, and display cell- and tissue specificity upon maturation, indicating their distinct functions (Fig. 1A) [13,14]. Unlike actin and tubulin isoforms, IF proteins are highly divergent in sequence, and vary greatly in molecular weight.

Below, we review IF classification, as evidenced by their differential expression and self-assembly properties. We will focus on the structural–functional significance of the IF intrinsically disordered regions (IDRs). To assess the physical properties of the IDRs, we will briefly introduce the theoretical realization of IFs as interacting bottlebrushes. We utilize sequence analysis to characterize the IDRs and find unifying properties of the tail and head domains. We review recent in vitro experiments which demonstrate the universal physical properties and their role as mediators for IF inter- and intra-filament interactions. We will conclude with future directions and open questions in the field.

1.1. Intermediate filament classification

IFs are classified into six types, based on their sequence similarities (Fig. 1A) [14–16]. The largest group consists of the type I (acidic) and type II (basic) keratins, which include 54 subunits of obligate heteropolymers that are typically expressed in epithelial cells [17,18]. Keratin filaments are composed of specific combinations of type I and type II keratins in a 1:1 ratio, and are both tissue- and cell-specific [16,19]. Keratins can be further subdivided into two clusters: seventeen proteins are found in "hard" epithelial tissues such as nails and hair, while the remaining 37 proteins are more often found in epithelial tissues.

Type III IFs include the proteins vimentin, desmin, glial fibrillary acidic protein (GFAP), and peripherin. Type III proteins can form homopolymeric as well as heteropolymeric IFs, in combination with other type III or type IV proteins. The most widely distributed IF protein is vimentin, which is typically expressed in leukocytes,

* Corresponding author.
E-mail address: roy@post.tau.ac.il (R. Beck).

http://dx.doi.org/10.1016/j.febslet.2015.07.024
0014-5793/

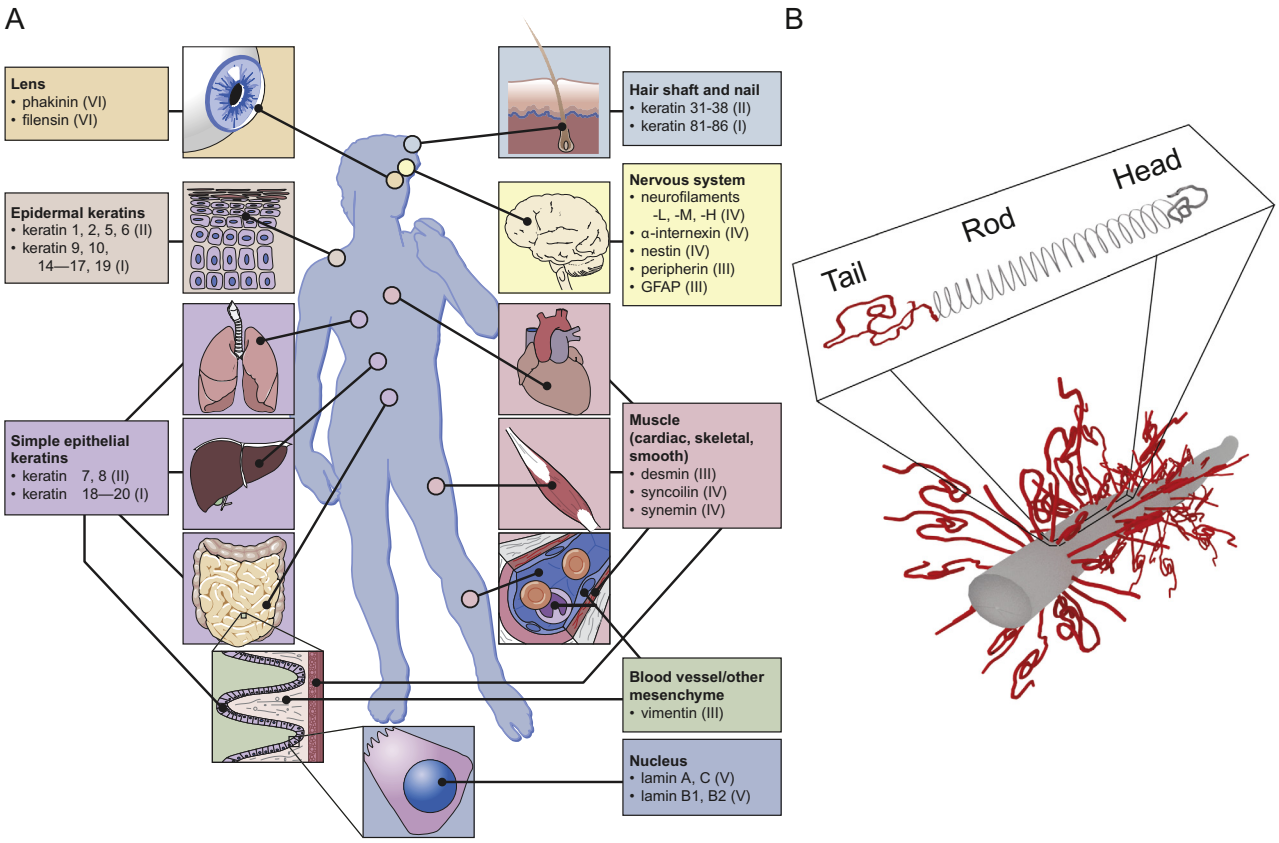

Fig. 1. (A) Distribution of IF proteins in the human body. IF proteins include six types; most of them are expressed in the cytoplasm and are tissue-specific. Lamins (type V) are the only IF found in the nucleus of most mammalian cells. The complexity and cell-specific expression of IFs are demonstrated in the intestine, which expresses keratins (epithelial layer), vimentin (vasculature), desmin (smooth muscle layer) and neuronal IFs (nervous system of the gastrointestinal tract, not shown). For other tissues and organs, representative major IFs are listed. Not all tissues are displayed, and the complex IF tissue expression profile is not fully demonstrated. The figure is adapted with permission from Toivola et al. [4]. (B) Schematic bottlebrush architecture of an IF assembled from subunit proteins composed of intrinsically disordered head and tail domains flanking an α-helical rod domain.

blood vessel endothelial cells, some epithelial cells, and mesenchymal cells such as fibroblasts [20,21]. Other type III proteins have a more limited distribution: desmin filaments, for example, are found in muscle cells, and are responsible for stabilizing sarcomeres in contracting muscle. GFAP forms filaments in glial cells that surround neurons, and in astrocytes. Peripherin is found in neurons of the peripheral nervous system, as well as in neurons of the central nervous system that project toward peripheral structures, such as spinal motor neurons [22].

The neurofilament (NF) triplet subunits NF-L, NF-M and NF-H, together with α-internexin (α-Inx), nestin, synemin and syncoilin, comprise the fourth type of IF proteins. The NF triplet and α-Inx are expressed in most neurons, although in different stoichiometries and expression levels during development [23–25]. Nestin is expressed primarily in dividing precursor cells of developing and regenerating tissues [26]. The two synemin isoforms are found in muscle cells, where they co-assemble with desmin into heteropolymeric IFs [27]. Syncoilin is mostly expressed in skeletal and cardiac muscle [28,29], where it links the extracellular matrix and the cytoskeleton.

Type V proteins, known as lamins, are found exclusively in cell nuclei, where they form a network that supports the nuclear membrane [14]. The sixth type bears little resemblance to other IF proteins. Members of this class are two lens-specific proteins that form beaded filaments: beaded filament structural protein 1 (Bfsp1; also known as filensin), and Bfsp2, also known as phakinin and CP49 [30].

1.2. From molecular structure to self-assembled supramolecular bottlebrushes

All IF proteins share a common tripartite structure consisting of a central α-helical rod domain, and disordered N-terminal head and C-terminal tail domains. The size and sequence of the rod domain is conserved among all IFs except for lamins, in which it is slightly longer [31]. Within the rod domain, amino acids are organized in heptad repeats, with every first and fourth amino acid being hydrophobic. This order drives the association of like molecules into coiled coil dimers [14].

Structural differences in the organization of the rod domain give rise to three distinct assembly groups: keratins (type I and type II IF proteins) constitute assembly group 1, forming obligate heterodimers composed of one acidic (type I) and one basic (type II) keratin. Type III and type IV IF proteins constitute assembly group 2. Here, some members are able to form homodimers, while others require specific partners for dimerization. Lamins (type V) constitute assembly group 3. Members from different assembly groups cannot co-assemble into IFs; rather, they completely segregate into distinct networks, even within the same cell [1].

While the conserved and ordered rod domains interact with each other to form the core of the filament, the head and tail domains are responsible for filament assembly and network organization. Importantly, both head and tail domains of most IF proteins are unstructured [32], except for lamin A/C tails which form an immunoglobulin-like fold [33]. The head and tail domains are

Please cite this article in press as: Kornreich, M., et al. Order and disorder in intermediate filament proteins. FEBS Lett. (2015), http://dx.doi.org/10.1016/j.febslet.2015.07.024