Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

The present and the future of motif-mediated protein–protein interactions

Moon-Hyeong Seo[1,2] and Philip M Kim[1,3,4]

Protein–protein interactions (PPIs) are essential to governing virtually all cellular processes. Of particular importance are the versatile motif-mediated interactions (MMIs), which are thus far underrepresented in available interaction data. This is largely due to technical difficulties inherent in the properties of MMIs, but due to the increasing recognition of the vital roles of MMIs in biology, several systematic approaches have recently been developed to detect novel MMIs. Consequently, rapidly growing numbers of motifs are being identified and pursued further for therapeutic applications. In this review, we discuss the current understanding on the diverse functions and disease-relevance of MMIs, the key methodologies for detection of MMIs, and the potential of MMIs for drug development.

Addresses
[1] Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, ON M5S 3E1, Canada
[2] Natural Constituents Research Center, Korea Institute of Science and Technology (KIST), Gangneung 25451, Republic of Korea
[3] Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 3E1, Canada
[4] Department of Computer Science, University of Toronto, Toronto, ON M5S 3E1, Canada

Corresponding author:

Current Opinion in Structural Biology 2018, **50**:162–170

This review comes from a themed issue on **Sequences and topology**

Edited by **Joseph Marsh**

https://doi.org/10.1016/j.sbi.2018.04.005

Introduction

As proteins perform pivotal roles in cells, tightly and accurately controlled interactions among proteins orchestrate essential cellular processes, from differentiation to cell death. Thus not surprisingly, PPIs have caught particular attention for deeper understanding of cellular mechanisms as well as for the development of novel strategies to target various types of diseases. Valuable efforts with systematic high-throughput screening (HTS) methods, such as yeast two-hybrid (Y2H) [1••] and affinity purification–mass spectrometry (AP–MS) [2••,3], have led to interactome networks of several organisms with increasing coverage [4–9]. Concurrently, a growing number of PPIs are known to be mediated by short linear motifs (SLiMs), which may even comprise the majority of the overall PPIs [10]. Such short motifs typically are a few to 10 residues in length and are most often found in intrinsically disordered regions (IDR). Different types of linear motifs have been discovered and known to be relevant to key biological processes, which include localization, binding, modification, and cleavage [11]. Moreover, MMIs are regarded as potential drug targets, as their small interfaces render them more likely to be druggable with small molecules or peptides [12–15]. Conversely, it is much more difficult to imagine inhibition of typical interactions between globular domains with small molecules.

Nevertheless, because of a bias toward stable interactions in current experimental methodologies, MMIs have been underrepresented in PPI datasets; interactions between well-defined globular proteins account for the vast majority of instances. Considering the incompleteness of PPI datasets (estimated still less than 20% of all potential pairwise interactions in human proteome [16]) many instances of motif-domain pairs are yet to be discovered. As these short motifs have been shown to play vital roles particularly in dynamic cellular processes, for instance through allostery or posttranslational modulation [11], comprehensive studies of these types of interactions will have a large impact on our understanding of cellular networks. The number of putative binding motifs (less than 15 amino acids in length) in the human proteome was conservatively estimated to be more than 100,000 [10,11]. There are large databases of MMIs, such as the Eukaryote Linear Motifs (ELM, elm.eu.org) [17••], Linear Motif mediated Protein–protein interaction (LMPID, bicresources.jcbose.ac.in/ssaha4/lmpid) [18] or Minimotif Miner (MnM, minimotifminer.org) [19], among those, however, experimentally validated motif instances do not exceed more than 4000 as of September 2017 [17••], underlining the scale of the lack of our current understanding of this type of interactions.

This review discusses recent studies of motif-based PPIs and their biological significances. We then introduce experimental methods for systematic detection of domain-motif interactions, followed by a discussion of their therapeutic potential.

Properties of MMIs

SLiMs and their interactions generally have distinct characteristics compared to interactions between globular

domains. First, SLiMs are evolutionary dynamic due to their short length and the relative simplicity of sequence pattern, with only few positions being under strong selection pressure. Consequently, SLiMs can easily appear or disappear by a simple mutation independently in multiple proteins [20], which gives rise to the higher complexity of the PPI networks as well as the functional diversity of proteins. As biological complexity is presumably related to the complexity of interactions among proteins [21], it is worth noting the contribution of SLiMs. Second, the vast majority of linear motifs are known to reside within IDR. IDRs are estimated to account for 30% of the human proteome. Therefore, IDRs that are mostly without function annotation constitute a significant fraction of the unexplored interaction surfaces of the human proteome [22,23••]. Third, because of their short length, SLiMs have smaller interfaces (~500 Å^2) with binding domains [12,13] usually resulting in low affinity (often in the single-digit or double-digit micromolar range), and, in turn, in transient and reversible interactions that can respond to the quickly changing cellular environment. Another factor limiting affinity is the entropic penalty resulting from their unstructured nature in the unbound state. Rapid and spontaneous association or dissociation is directed to initiation or termination of signals. Likewise, fine-tuned affinity is a key for proper function of signaling proteins, which could be achieved by varying stoichiometry of motifs (avidity) or substitution of hot spot residues; too high or low affinity disturbs proper functions [24,25]. Finally, affinity or specificity of MMIs can vary by several factors. Other than a few conserved key residues that directly interact with the binding domains, flanking regions can enhance or interrupt the binding. Consequently, many instances of promiscuous binding have been reported [11,26], in which many binding domains can bind similar but diverse sequences. In addition, posttranslational modification (PTM) of a motif is another critical determinant for its binding to complementary domains. Several SLiMs are activated or inactivated by PTM of a residue [27]. Furthermore, motif sequence is not the only determinant for interactions. Spatiotemporal coincidence must be acquired for both a motif and a binding domain to have *in vivo* functional relevance [11,22]. These multiple factors governing the specific interactions complicate the identification of authentic novel MMIs.

Functions of MMIs

Their transient nature endows MMIs with crucial roles in signal transduction and acute regulation. Motifs do not just function as binding interfaces but also mediate many regulatory processes of each protein, such as subcellular localization, enzyme recruitment and PTM. In Figure 1, we briefly summarize functional classifications of ligand motifs with representative cases for each group based on both the ELM [11,17••] and Minimotif miner [19] classes.

As binding modules, SLiMs facilitate complex formation among distant proteins that contain either ligand motifs or motif-binding domains. Despite the low affinity of motifs, the cooperative use of multiple SLiMs provides biologically relevant interactions and can form stable complexes with high-avidity; such cooperativity can also determine the specificity of interactions. Various stoichiometries of interactions are governed by SLiMs from binary interactions (e.g., SH2 or SH3 domains) (Figure 1a) to multiprotein scaffolds including Epsin-1 [28] and AKAP [29] that assemble functionally connected proteins at once with multiple motifs.

Furthermore, SLiM-meditated interactions determine correct subcellular localization of proteins, which is a critical factor for proteins involved in signaling or functions occurring in a specific subcellular compartment. Dysregulated localization of proteins by mutations in trafficking motifs or pathogenic uses often results in several diseases [11,30]. KDEL motifs [31] and several nuclear export or localization signals [32] (Figure 1b) are among the most actively studied and utilized examples, which determine transport of proteins between the endoplasimic reticulum and the Golgi apparatus or across the nuclear membrane, respectively. Besides the trafficking motifs, tethering proteins with anchoring motifs to specific cellular compartments or molecules, such as membrane, microtubules or scaffold proteins, enables distinctive spatial roles for the proteins that harbor them. For instance, microtubule plus-end-tracking proteins (+TIPs) associate with growing microtubule plus ends, which is mediated by the interaction of their SxIP motifs with end-binding homology (EBH) domain of end-binding (EB) proteins. As such, +TIPs control microtubule dynamics as well as attribute to linkages between microtubule ends and other cellular structures during cell division, migration, and morphogenesis [33] (Figure 1c).

Enzymatic regulations of proteins are controlled by SLiMs, which include enzyme recruitment and protein degradation. Docking motifs that are distinct from the target modification sites of, for example, kinases and phosphatases provide specificity or increase affinity of the enzymes [34]. Since catalytic sites of enzymes are generally promiscuous to modification target sites in several target proteins, docking motifs often determine specific modification of target sites. MAPK [35] (Figure 1d) and CDK [36] or PP2A [37], Calcineurin [38] are among the well-studied cases for kinases and phosphatases, respectively. When it comes to an ubiquitin ligase, for example, APC/C [39] or MDM2 [40] (Figure 1e), SLiMs determine protein stability and they thereby can control cell cycles or apoptosis [41]. E3 ubiquitin ligases recognize degron motifs of target proteins for proteasomal degradation by polyubiquitylation, whereas deubiquitylating enzymes bind to other docking sites to antagonize the activity of ubiquitin ligases.