Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Template-based structure modeling of protein–protein interactions

Andras Szilagyi[1] and Yang Zhang[2,3]

The structure of protein–protein complexes can be constructed by using the known structure of other protein complexes as a template. The complex structure templates are generally detected either by homology-based sequence alignments or, given the structure of monomer components, by structure-based comparisons. Critical improvements have been made in recent years by utilizing interface recognition and by recombining monomer and complex template libraries. Encouraging progress has also been witnessed in genome-wide applications of template-based modeling, with modeling accuracy comparable to high-throughput experimental data. Nevertheless, bottlenecks exist due to the incompleteness of the protein–protein complex structure library and the lack of methods for distant homologous template identification and full-length complex structure refinement.

Addresses

[1] Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Karolina út 29, Budapest 1113, Hungary

[2] Department of Computational Medicine & Bioinformatics, The University of Michigan, 100 Washtenaw Avenue, 2035B, Ann Arbor, MI 48109-2218, USA

[3] Department of Biological Chemistry, The University of Michigan, 100 Washtenaw Avenue, 2035B, Ann Arbor, MI 48109-2218, USA

Corresponding authors: Zhang, Yang (zhng@umich.edu)

Current Opinion in Structural Biology 2014, **24**:10–23

This review comes from a themed issue on **Folding and binding**

Edited by **James CA Bardwell** and **Gideon Schreiber**

For a complete overview see the Issue and the Editorial

Available online 11th December 2013

0959-440X/$ – see front matter, © 2013 Elsevier Ltd. All rights reserved.

http://dx.doi.org/10.1016/j.sbi.2013.11.005

Introduction

Proteins are important molecules involved in virtually all cellular functions, including structural support, signal transduction, bodily movement, and defense against pathogens. Most functions are mediated by interactions between proteins. To perform all their various biological functions, the protein–protein interactions must be extremely diverse in the three-dimensional structure: individual protein chains may form homomeric or hetero-oligomeric, obligate or non-obligate, and transient or permanent complexes. These interactions form an intricate and dynamic network, the interactome, in living cells. Due to the important role in cellular processes, vast efforts have been devoted to uncovering the interactome, primarily by high-throughput experimental techniques [1,2]. However, these methods can at best tell which proteins interact, but are unable to reveal the structural details of such interactions; the latter is essential to understanding the molecular basis of cellular functions and for designing new therapies to regulate these interactions. Therefore, a major long-term goal of modern structural biology is to create a detailed 'atlas' of protein–protein interactions [3], containing not only the full interactome but, more challengingly, the atomic-level 3D structures of all protein complexes.

The most accurate structures of protein complexes are provided by X-ray crystallography and NMR spectroscopy; however, these techniques are labor-intensive and time-consuming. There has been a large gap between the number of known interactions and the number of interactions with known structures. Despite significant efforts in traditional structural biology and the structural genomics projects that aim at high-throughput complex structure determination [4], the latest statistics show that only ∼6% of the known protein interactions in the human interactome have an associated experimental complex structure [5]. This number is quite low considering that we have a complete or partial experimental structure for ∼30% of human proteins. Moreover, while the estimated size of the human interactome ranges from ∼130 000 [6] to ∼650 000 [7], interactome databases currently contain only ∼41 000 binary interactions between human proteins, and many of them may be in error because of the inherent limitations of high-throughput experimental interaction discovery methods such as the yeast two-hybrid method [8]. Therefore, the development of efficient computational methods for discovering new interactions and in particular for large-scale, high-resolution structural modeling of protein–protein interactions is of paramount importance.

There are two distinct methods for the computational modeling of protein–protein complex structures (Figure 1). In protein–protein docking, complex models are constructed by assembling known structures of the interacting components, which are solved or predicted in the unbound form, through an exhaustive search and selection of various binding orientations (Figure 1a). The docking searches are often based on the shape and solvation matches of the surfaces of the component proteins, and work well for the protein complexes with an interface having obvious shape complementarity and with a large (>1400 Å^2) and predominantly hydrophobic

Figure 1

Two principal protocols for protein complex structure prediction. Red and blue represent sequences and structures of two individual chains. **(a)** Rigid-body protein–protein docking constructs protein complex structures by assembling known structures of monomer components which are usually solved (or modeled) in their unbound states. The final model is selected from those with the best shape complementarity, desolvation free energy and electrostatic matches between interfaces of the component structures [9–12]. **(b)** Template-based modeling (TBM) identifies complex structure templates by aligning the amino acid sequences of the target chains with the solved complex structures in the PDB library (shown on the left). The alignment can be generated based on sequence, sequence profile, or a combination of the sequence and structure feature information. The best template of the highest alignment score is selected; and the structure framework in the aligned regions is copied from the template protein which serves as a basis for constructing the structure model of the target [18•,21••,24,25]. Note that (b) only shows a typical protocol of homology-based template detection. There are variants of TBM which detect complex templates by query and template structure comparisons (see Figure 2) [19•,20••,22•,23,30].

interfacial area [9]. But one challenge for rigid-body protein docking is that the accuracy decreases rapidly when the protein chains undergo large conformational changes upon binding [10,11]. Additionally, docking can only be performed when monomer structures of the interacting components are provided; but the experimental structures are in fact unavailable for a major portion of protein domains (although structural models of the monomer proteins can be generated by computational structure prediction, the rigid-body docking accuracy is sensitive to the errors in the monomer models). The recent progresses in rigid-body protein docking are reviewed in [11,12].

The second method is template-based modeling (or TBM), which constructs protein complex structure of unknown targets by copying and refining the structural framework of other related protein–protein complexes whose structure has been experimentally solved (Figure 1b). The method of TBM has long been used to predict the tertiary structure of single-chain proteins, based on the principle that homologous proteins of similar sequences usually take the similar structure [13]; the method was later extended to model tertiary structure for distant homology proteins with the invention of the technique of threading [14], which aims to recognize the template structures without evolutionary relation to the target through incorporating structure information into sequence alignments. The general steps of TBM include finding one or more appropriate template(s); aligning the target sequence with the templates using sequence alignment, profile-based alignment, or threading; building an initial model for the target by copying the structural

Download English Version:

https://daneshyari.com/en/article/8320222

Download Persian Version:

https://daneshyari.com/article/8320222

Daneshyari.com