Some specific interactions are now at the center of drug discovery research

Protein–Protein Interactions

Lessons from the Past and Current Directions

Richard A. Stein, M.D., Ph.D.

Protein–protein interactions, the basis of cellular structure and function, are critical for virtually all pathways underlying development, health, and disease. The past few decades have witnessed major advances in the ability to interrogate and exploit interactions between proteins and, in parallel, the field has developed better capabilities to target them therapeutically. As a result, certain interactions that were initially viewed as intractable have emerged more recently at the forefront of drug discovery efforts. "For a long time, many people believed that membrane-spanning segments are inert anchors that simply insert proteins into membranes or hold proteins there," says Daniel DiMaio, MD, PhD, Waldemar Von Zedtwitz Professor of Genetics, professor of molecular biophysics and biochemistry, professor of therapeutic radiology at Yale University, and deputy director of the Yale Cancer Center.

It is estimated that up to one third of eukaryotic proteins, including many drug targets, are embedded into either the plasma membrane or

intracellular membranes. Many transmembrane protein domains interact with each other in a highly specific way to ensure proper protein folding or to mediate protein oligomerization or complex formation.

"The idea that transmembrane domains can make highly specific interactions is something that much of the scientific community does not yet appreciate," says DiMaio. While the ability of transmembrane proteins to dimerize has been studied, and certain critical motifs that establish these interactions have been identified, less work has focused on the ability of different transmembrane domains to interact with one another. "Our work has confirmed that these can be very highly specific interactions," continues DiMaio.

In a recent study published in *eLife*, DiMaio and colleagues analyzed 26-amino-acid artificial transmembrane proteins known as traptamers, which contain one transmembrane helix and consist exclusively of leucine and isoleucine residues, two structurally very similar hydrophobic amino acids. When expressed in cells, some traptamers activate the erythropoietin receptor, and others activate the platelet-derived growth factor receptor.

Getty Images / Hero Images

The researchers found that changing a single methyl group at specific positions in the traptamer changed its target specificity. "This finding was surprising and made us wonder what a conservative mutation is, because one cannot be more conservative than a leucine to isoleucine change, yet this very small chemical change had a very profound effect on activity," explains DiMaio.

The focus on these transmembrane protein domains in DiMaio's group began with their finding that bovine papillomavirus E5, a short transmembrane protein, turns on the

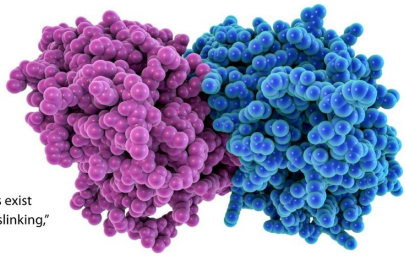
PDGF receptor by its use of a transmembrane interaction. "That started the idea to create artificial transmembrane proteins, and we wonder if there are similar proteins in cells that people never looked for," says DiMaio. Some of these proteins are small and could easily be missed or ignored during genomic or biochemical searches, but may have biological value. "There may be a whole universe of proteins out there that has not been studied very much, and this conceivably could represent a whole new class of therapeutics," concludes DiMaio.

"We got engaged in the idea of chemical crosslinking to look at protein–protein interactions early on, but we identified some major stumbling blocks," says James E. Bruce, PhD, professor of genome sciences at the University of Washington. Crosslinking has been successfully used for decades, and made it relatively easy to identify binding partners from small protein databases or for few interacting proteins. Comparatively, using conventional crosslinking for large-scale applications has been challenging. "With conventional crosslinking the identification of each crosslinked peptide has been a real bottleneck," says Bruce, "and the task rapidly can become intractable for most approaches."

Identifying Crosslinked Peptides

In a recent study, Bruce and colleagues developed a protein interaction reporter–based crosslinking approach combined with mass spectrometry, and used it to identify crosslinked peptides generated from the mitochondrial interactome. This strategy provided unprecedented insights into the structure of many mitochondrial proteins, including the ones involved in oxidative phosphorylation. Some of the peptide pairs that were identified supported the existence of the respirasome. "The underpinning of this whole approach is to learn about what proteins are interacting and what conformational features of proteins exist in the system when we do the crosslinking," explains Bruce.

This technique promises to unveil details of the interactions between drugs and proteins in complex systems. In a recent research effort that quantitatively examined cancer cells targeted with HSP-90 inhibitors, Bruce and colleagues identified conformational changes and protein–protein interaction, specifically in the HSP-90 network in a drug concentration–dependent and a drug mechanism of action–specific manner. "A new area that we are proposing, as to where this technology can go, is what we call 'systems structural biology,'



Getty Images / KATERYNA KON / SCIENCE PHOTO LIBRARY

"In cells, most proteins perform their functions by interacting with other molecules and we tried to take advantage of the protein–protein interaction networks to deduce the function of proteins for which we cannot identify functional homologies."
—Yang Zhang, PhD

and involves structural biology performed on intact systems," concludes Bruce.

Systems structural biology offers the opportunity to visualize protein complex structural features that can shed new light on disease pathologies, provide molecular-level insights on potential therapies, and greatly expand knowledge of many biological processes.

"We are trying to interpret what each protein does in the living cells through computational algorithms," says Yang Zhang, PhD, professor of computational medicine & bioinformatics and professor of biological chemistry at the University of Michigan.

Computer-based protein function annotation has been historically performed by relying on homology inference. This approach, established for years, involved an initial search for functional homology to a query protein, followed by transferring the functions of the known homologies (as template) to the unknown query protein. However, homology-based functional annotation becomes unreliable when the sequence identity between the query and template proteins is low, and typically lower than 30%.

"In cells, most proteins perform their functions by interacting with other molecules and we tried to take advantage of the protein–protein interaction networks to deduce the function of proteins for which we cannot identify functional homologies," says Zhang. Based on this idea, Zhang and colleagues developed a new method, MetaGO, for automated protein function annotations. "MetaGO predicts the gene ontology of a query protein using protein–protein interaction networks as an important component," says Zhang. As a hybrid method, MetaGO integrates three complementary pipelines—built on evolutionary biology, structural biology, and protein interaction networks—for multi-level function annotations.

"The idea of using protein–protein interaction networks to deduce new protein functions is not new," says Zhang. In previous approaches,

researchers tried to predict function from the direct interaction partner because it is believed that proteins interacting with each other should be involved in the same biological pathway and may share similar functions. Zhang and his colleagues went one step further and found that the homologies of the interaction partners can share similar function to the query protein. This helps MetaGO to get broader functional insights than that only from the direct interaction partners.

Low-resolution Protein Structure Prediction

Another unique aspect of Zhang's pipeline is the use of low-resolution protein structure prediction, a traditional focus of Zhang's laboratory, to help interpret protein functions. The interrogation of both global and local structural alignments built on the low-resolution predicted models makes MetaGO detect far more functional homologies than sequence-based homologous transferals. To validate MetaGO, Zhang and colleagues tested it on a large-scale set of 1,000 non-redundant proteins and demonstrated its superiority over peer approaches for functional prediction in the community-wide blind experiments, such as CAFA. "These results demonstrate the possibility to deduce new functional insights that go far beyond the traditional sequence homology-based predictions," says Zhang.

"Studying cryptic sides got a boost when molecular dynamics became computationally feasible, but when people identified large numbers of sites on protein surfaces the question became how useful those sites are," points out Sandor Vajda, PhD, professor of biomedical engineering and chemistry and director of the Biomolecular Engineering Research Center at Boston University. In a recent study that explored the relative contributions of conformational selection and induced fit, investigators in Vajda's group examined the structural origin of cryptic sites in proteins.

"Our study on cryptic sites changes the way druggability was seen longer ago," says Vajda. The study conducted in Vajda's group analyzed a representative group of 93 proteins in which the structure of the sites in unbound and ligand-bound structures differs substantially. "According to our paper, only a relatively limited number of those sites are actually relevant or druggable," says Vajda.

Some of the advances that helped design therapeutics to target protein–protein interfaces include computation, molecular dynamics, and the progression of X-ray crystallography. These approaches made an increasing number of structures available under different conditions, including proteins co-crystalized with different ligands. "And very frequently even small molecules that are not targeted, such as crystallization oddities, can reveal the presence of a site," says Vajda. These molecules that are occasionally captured in X-ray crystal structures are frequently called crystallization artifacts. "Often these crystallization artifacts can tell something about the protein sites," adds Vajda.

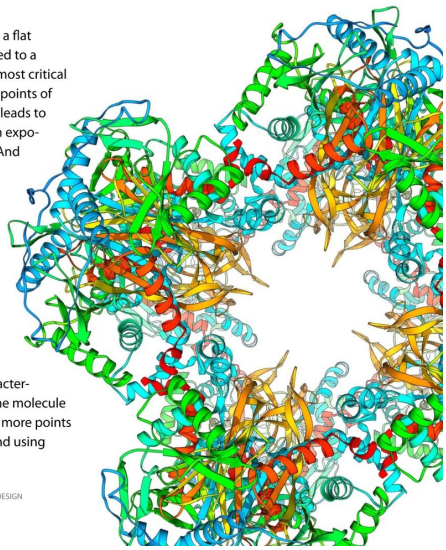
Macrocycles

When targeting protein–protein interactions, the main challenge is that typical interfaces do not have well-defined binding pockets, according to Dehua Pei, PhD, professor of chemistry and biochemistry at The Ohio State University. While small molecules are ideal for targeting deep binding pockets in proteins, they do not bind with sufficiently high affinity to flat interfaces, which are usually involved in protein–protein interactions. "This makes protein–protein interfaces undruggable by conventional methods," says Pei. This challenge can be solved by using other molecules, such as antibodies. "But the challenge is how get these large proteins into the cell," continues Pei.

A technique that provides a more general solution, and is used extensively in Pei's group, involves the use of macrocycles such as macrocyclic peptides, "which have a lot of advantages over other approaches and, in addition, they are smaller than proteins and easier to make," says Pei.

When small molecules interact with a flat protein–protein interface, as opposed to a protein binding pocket, one of the most critical aspects revolves around how many points of contact are being lost, because this leads to the loss of binding energy and to an exponential loss of the binding affinity. "And that is where small molecules get into trouble, because they don't have enough binding affinity," says Pei.

Macrocycles that are 3–5 times larger than conventional small-molecule drugs, due to their ability to make multiple points of contact with flat surfaces, can compete with proteins for target binding and yet retain several physicochemical characteristics of small molecules. "Because the molecule is larger, it has a larger interface and more points of contact, and that is the idea behind using macrocycles," says Pei.



Getty Images / LAGUNA DESIGN

Two features that make macrocyclic peptides desirable therapeutic options are the easiness of manufacture and their lower toxicity and immunogenicity as compared to other molecules. "The downside of macrocyclic peptides is that the cell membrane is impermeable to almost all of them, and that makes their target inaccessible," explains Pei. Therefore, introducing macrocycles into cells is a critical aspect of their therapeutic development. "We discovered a family of cyclic peptides that are very efficient in going into the cell and they can also take other cargo molecules with them," continues Pei.

Recently, Pei and colleagues synthesized cell-permeable bicyclic peptides in which the two rings contained different peptide sequences for target binding and cell-penetrating motifs. "We designed a combinatorial library in a way that guarantees cell entry and target engagement," says Pei. This strategy led to the identification of a potent and cell-permeable compound that inhibited NF-κB signaling and the proliferation of cisplatin-resistant ovarian cancer cells *in vitro*, emerging as a promising strategy to develop anti-inflammatory and anticancer drugs.

"One thing that could make a difference in the future are small molecules that can allosterically inhibit protein–protein interaction," notes Pei. Allosteric inhibitors bind sites that are distinct from the active site and lead to conformational changes that modulate protein function. "This will not be a general solution but certainly will provide some applications for certain protein–protein interaction targets," says Pei.

Another attractive set of therapeutic targets comprises molecules that can bind cryptic binding sites. These are binding sites that cannot be seen in the most stable conformation of a protein, but in alternative conformation in which the protein may exist for a fraction of the time. "That alternative conformation might have a binding pocket that one can inhibit," notes Pei. ■