


5

DESVENDANDO A ANCORAGEM PROTEÍNA-PROTEÍNA

Autores 5.1

Autores: Luana Luiza Bastos 

Revisão: Isaac Farias Cansação  , Ana Carolina Silva Bulla 

Resumo 5.1

O conhecimento das interações proteína-proteína é de fundamental importância para o entendimento dos mecanismos moleculares, dos processos bioquímicos e das vias celulares. No entanto, a resolução experimental de complexos proteína-proteína é desafiadora, uma alternativa para as limitações, no número de complexos experimentais resolvidos e para o conhecimento tridimensional (3D) das interfaces são as abordagens computacionais, como as análises de ancoragem molecular. O presente artigo, realiza uma introdução à ancoragem proteína-proteína, às metodologias utilizadas, bem como suas aplicações.

As proteínas são macromoléculas biológicas, presentes em todos os seres vivos. Participam de inúmeros processos celulares e possuem diversas funções, que vão desde o transporte de pequenas moléculas até o envolvimento no processo de sinalização celular [1]. Neste contexto, o entendimento das interações proteína-proteína (IPPs) é de extrema importância, uma vez que as proteínas exercem diversas funções biológicas a partir das suas interações.

O conhecimento das IPPs possibilita o melhor entendimento dos processos bioquímicos e celulares, abre possibilidade para o desenvolvimento de fármacos e o melhor entendimento de processos patológicos. Como exemplo, podemos citar os estudos realizados a respeito do SARS-CoV-2, o vírus causador da COVID-19.

A COVID-19 é uma doença infecciosa, que foi responsável por cerca de 664.516 óbitos e mais de 30.617.786 (dados obtidos em abril de 2022) casos só no Brasil, durante a pandemia mundial que se prolonga desde o início de 2022 [2]. A partir de diversos estudos[3, 4], descobriu-se que a entrada do vírus na célula é feita principalmente por meio da ligação da proteína spike de SARS-CoV-2 com a ECA2 (enzima conversora da angiotensina 2), uma proteína transmembrana, expressa na superfície de diversas células do corpo, como o epitélio do sistema respiratório. Essa interação desencadeia a mudança conformacional pré/pós-fusão, responsável pela entrada do vírus na célula hospedeira (Figura 5.1). Com o entendimento mais detalhado das interações da proteína spike e seu domínio RGD

(região da spike que interage com ECA2) com a ECA2, foi possível, por exemplo, a elaboração de vacinas que possuem anticorpos que interagem diretamente nesta região.

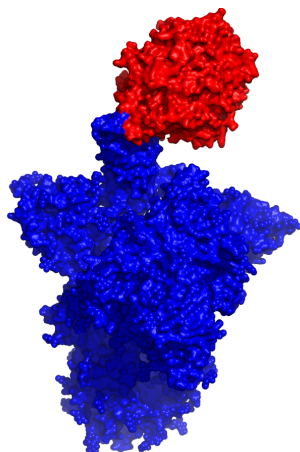


Figura 5.1: Modelo da estrutura tridimensional da ECA2 (em vermelho) ligada ao RBD da proteína spike de SARS-CoV-2 (em azul). Fonte: o próprio autor.

Contudo, o conhecimento das estruturas tridimensionais dos complexos proteína-proteína, ainda é limitado. Uma vez que experimentos de bancada para obtenção de modelos tridimensionais, como a cristalografia de raio X, conta com inúmeros desafios. Principalmente relacionados ao tamanho das estruturas, uma vez que um complexo de duas proteínas ligadas tende a ser bem maior do que uma proteína ligada a uma pequena molécula, por exemplo [4]. Apesar da inovação obtida com o uso da crio-microscopia eletrônica, que possibilita a resolução de complexos proteicos maiores, o número de estruturas depositadas em bancos de dados como o PDB (*Protein Data Bank*) [5] e EMDB (*Electron Microscopy Data Bank*) [6] ainda é bem pequeno se comparado ao número de estruturas únicas.

Como alternativa a essa quantidade ainda reduzida de interfaces IPPs experimentalmente resolvidas, temos as metodologias computacionais, como a ancoragem (ou *docking*) proteína-proteína. Os primeiros ensaios computacionais de ancoragem de duas proteínas datam de 1978 e foram realizados por S.J. Wodak e J. Janin [7], utilizando modelos de baixa resolução (em uma escala de resíduos de aminoácidos). Desde então, inúmeras ferramentas e metodologias

têm sido desenvolvidas, a fim de melhorar a qualidade e precisão dos experimentos realizados.

O experimento computacional consiste em gerar um grande número de poses para encontrar a orientação preferencial de duas proteínas, com base no menor índice de energia livre. Para entender melhor como esse procedimento funciona, vamos relembrar o conceito de energia livre de Gibbs (ΔG). Basicamente, a energia livre é definida como a diferença entre a entropia e a entalpia multiplicada pela temperatura. Sendo assim, o ΔG corresponde a energia que “sobra” após uma reação química. Logo, uma reação só é espontânea se a energia livre for menor do que zero. Na ancoragem, o algoritmo vai buscar na interface de ambas as proteínas o local de ligação com o menor índice de energia livre. Quanto menor o ΔG , teoricamente, mais estável e firmemente ligadas estarão as duas proteínas [8].

A ancoragem proteína-proteína pode ser realizada, tanto quanto não se tem nenhuma informação do local de interação, ou seja não se tem informação prévia de onde na interface de ambas as proteínas ocorre a ligação. Quanto quando se tem informações a respeito do local de ligação das proteínas analisadas, realizando a ancoragem em um sítio previamente conhecido.

Quando as informações sobre a interface de interação não estão disponíveis, realizamos a ancoragem livre ou “cega”. O algoritmo, nesse caso, irá considerar toda estrutura tridimensional de ambas as proteínas em busca da combinação de interfaces com menor energia livre. Para entender melhor como esse procedimento é feito, imagine os movimentos de rotação e translação feitos pela Terra. O nosso sol será a proteína receptora (geralmente a maior proteína) e a nossa Terra será a proteína ligante. O algoritmo mantém a proteína maior fixa, e “gira” a proteína menor por toda a interface do receptor, como ocorre nos movimentos de translação da terra em volta do sol, gerando poses com base nas coordenadas atômicas x, y, z. Enquanto isso, as coordenadas espaciais da proteína ligante também sofrem alterações, como se ela “girasse” em torno de si mesma, imitando o movimento de rotação da terra. O algoritmo então computa milhares de poses

geradas com as diferentes combinações de interface receptor-ligante, e retorna as poses com menor energia livre [9, 10].

Na figura 5.2, temos um exemplo de resultados obtidos de uma ancoragem livre. Nesse caso, a proteína receptora, representada em vermelho, é a TNF- α , uma citocina pró inflamatória. E, como proteína ligante, foi utilizado o receptor nativo da citocina, a TNF-R2, onde as poses estão representadas em azul claro, azul escuro, lilás e marrom.

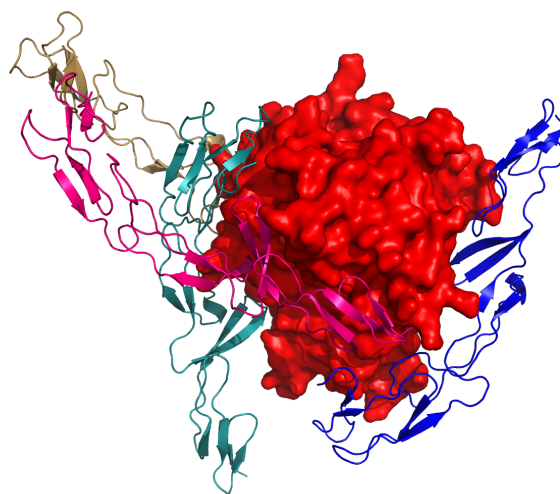


Figura 5.2: Resultado de ancoragem global utilizando as proteínas TNF- α (em vermelho) e TNF-R2 (em azul e verde) Fonte: o próprio autor.

Outra possibilidade é a realização de ancoragem orientada ou local, realizada quando se tem informações a respeito do local de interação de uma ou de ambas as proteínas. Nesse caso, o algoritmo irá gerar um grande número de poses, com base na região informada, buscando a menor energia livre local. Em relação à ancoragem orientada, a ancoragem livre tende a ser computacionalmente mais cara, uma vez que busca o menor ΔG em toda o espaço conformacional 3D de ambas as proteínas [11].

Um ponto a ser considerado na ancoragem proteína-proteína é que proteínas não são objetos estáticos, elas interconvergem entre confôrmeros de energia variada.. Isso dificulta o processo de ancoragem, uma vez que

considerar todas as possíveis alterações conformacionais é extremamente custoso computacionalmente, exigindo máquinas e algoritmos robustos [10].

Outra questão que deve-se levar em consideração, é que algoritmos de ancoragem rígida possuem uma grande limitação, já que muitas informações estruturais não são consideradas. Como estratégia, geralmente, algoritmos que consideram a flexibilidade total ou parcial das proteínas são utilizados para a ancoragem local. Quando não se tem essa informação disponível, é realizado uma ancoragem inicial livre com uma ferramenta de ancoragem rígida, a melhor pose é selecionada e submetida a ancoragem local, em uma ferramenta que considera a flexibilidade das proteínas, assim otimizando o uso dos recursos computacionais [12].

Existe atualmente um grande número de métodos de ancoragem, bem como uma grande diversidade de ferramentas, com diferentes aplicações. Podemos citar no que se refere a metodologias de ancoragem, os métodos de correlação da transformada rápida de Fourier (FFT) que geralmente realizam uma busca sistemática global [13], os métodos geométricos de Wolfson e Nussinov [14], métodos de Monte Carlo[8] e o *docking* molecular de alta ambiguidade [15] (HADDOCK). O progresso é continuamente monitorado pela competição CAPRI (*Critical Assessment of Predicted Interactions*), uma competição que visa testar algoritmos computacionais em previsões cegas de estruturas 3D de complexos de proteínas determinadas experimentalmente [16].

Apesar da constante evolução das ferramentas e metodologias de ancoragem, a identificação de interfaces nativas com base em funções de pontuação de encaixe é limitada. A fim de contornar esse contornar esse problema, atualmente tem sido desenvolvido funções de pontuação mais robustas, que se baseiam não apenas na energia livre, mas também na complementaridade de forma, energia de solvatação, interface de contatos formada e área acessível ao solvente, o que tem melhorado os resultados obtidos [17]. E, apesar de suas limitações, o uso da metodologia computacional de ancoragem proteína-proteína, tem possibilitado o estudo de interfaces de interação, bem como o desenvolvimento de pesquisas a respeito de novos fármacos, processos biológicos e bioquímicos [18].

Bloco 5.1

Este artigo está disponível em <https://bioinfo.com.br/desvendando-a-ancoragem-proteina-proteina/>

5.1 Referências

- [1] Gray JJ. High-resolution protein–protein docking. *Curr Opin Struct Biol*, 1;16(2):183–93, 2006.
- [2] Painel Coronavírus. Coronavírus Brasil. Disponível em: <<https://covid.saude.gov.br/>>. Acesso em: abr 2022.
- [3] Hwang SS, Lim J, Yu Z, Kong P, Sefik E, Xu H, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483):1255–60, 2020.
- [4] Nesci S. SARS-CoV-2 first contact: Spike–ACE2 interactions in COVID-19. *Chem Biol Drug Des*, 98(2):207–11, 2021.
- [5] RCSB PDB: Homepage. Disponível em: <<https://www.rcsb.org/>>. Acesso em: maio 2022.
- [6] EMDDB < Home. Disponível em: <<https://www.ebi.ac.uk/emdb/>>. Acesso em: maio 2022.
- [7] Wodak SJ, Janin J. Computer analysis of protein-protein interaction. *J Mol Biol*, 124(2):323–42, 1978.
- [8] Gray J. J.; et al. Protein–Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. *J Mol Biol*, 1;331(1):281–99, 2013.
- [9] Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics*, 30(12):1771, 2014.
- [10] Yan Y, Zhang D, Zhou P, Li B, Huang SY. HDock: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res*, 45(W1):W365–73, 2017.
- [11] Yan Y, Tao H, He J, Huang SY. The HDock server for integrated protein–protein docking. *Nat Protoc* 2020, 15(5):1829–52, 2020.
- [12] Kozakov D; et al. The ClusPro web server for protein-protein docking. *Nat Protoc*, 12(2):255, 2017.

- [13] Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci U S A*, 89(6):2195, 1992.
- [13] R Nussinov, H J Wolfson. Efficient computational algorithms for docking and for generating and matching a library of functional epitopes I. Rigid and flexible hinge-bending docking algorithms. *Comb Chem High Throughput Screen*, 2(5):249-59, 1999.
- [14] Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc*, 125(7):1731-7, 2003.
- [15] Lensink MF, Wodak SJ. Docking and scoring protein interactions: CAPRI 2009, 78 (15):3073-84. 2010.
- [16] Moal IH, et al. IRaPPA: information retrieval based integration of biophysical models for protein assembly selection. *Bioinformatics*, 33(12):1806-13, 2017.
- [17] Kozakov D; et al. How good is automated protein docking? *Proteins*, 81(12):2159-66, 2013.