

ESTUDOS DE INTERAÇÃO ENTRE A CO-CHAPERONA RPAP3 DE SORGO E DEMAIS PROTEÍNAS DO COMPLEXO R2TP

Palavras-Chave: CHAPERONAS, Hsp90, BIOINFORMÁTICA

Autores:

Gustavo Henrique Martins – Instituto de Química, UNICAMP

Prof. Dr. Carlos Henrique Inácio Ramos (orientador*) – Instituto de Química, UNICAMP

*cramos@unicamp.br

INTRODUÇÃO:

Descoberto inicialmente no organismo *Saccharomyces cerevisiae* e constituído pelas proteínas Tah1 (RPAP3); Pih1 e as helicases Rvb1/Rvb2, o complexo multiproteico R2TP desempenha diversas funções moleculares ao se associar com a chaperona Hsp90. Os papéis relacionados a esse sistema estão compreendidos desde o remodelamento da cromatina, passando por degradação de mRNA até a montagem e estabilização da RNA polimerase 2, entre outros (Kakihara e Houry, 2012).

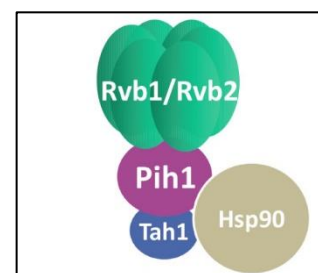


Figura 1. Representação esquemática do complexo R2TP. (Retirado de Kakihara e Houry, 2012).

Com o avanço dos estudos sobre esse complexo em diferentes organismos, a proteína Tah1 recebe destaque devido a sua capacidade de interagir com outras subunidades do complexo R2TP conectando-o à chaperona molecular. Em humanos, por exemplo, a proteína RPAP3 (homóloga de Tah1) interage com Pih1, Rvb2 e, através de seu domínio conservado intitulado como TPR, se liga ao motivo MEEVD presente na Hsp90 (Henri *et al.*, 2018; Maurizy *et al.*, 2018).

A caracterização dessas interações pode fornecer informações quanto às funções atribuídas à família das chaperonas moleculares e sua relação com a homeostase proteica, sobretudo nas ações da Hsp90. Diante disso e, pelo fato desse sistema não estar muito bem esclarecido em plantas, este trabalho se soma a resultados experimentais previamente obtidos acerca da proteína de interesse em quotas PIBIC anteriores e traz como proposta o estudo inicial das interações proteína-proteína no complexo R2TP a partir da caracterização da RPAP3 de *Sorghum bicolor* (SbRPAP3) olhando para seus domínios preditos por ferramentas de bioinformática.

METODOLOGIA:

A sequência de aminoácidos da RPAP3 de *Sorghum bicolor* foi um importante parâmetro para obter diversas predições estruturais/funcionais. Desse modo, nas plataformas computacionais mencionadas a seguir, a sequência proteica foi o principal *input* para obter os resultados.

Para avaliar os tipos de estrutura secundária presentes na proteína de interesse, utilizou-se a ferramenta PSIPRED (Jones, 1999) e, na tentativa de detectar alguns domínios importantes e suas posições ao longo da cadeia polipeptídica, as plataformas Prosite (Sigrist *et al.*, 2012) e InterPro (Hunter *et al.*, 2009) foram empregadas.

Com a finalidade de obter o grau de similaridade entre a sequência da SbrPAP3 com outras proteínas homólogas, foram realizadas buscas e alinhamentos múltiplos utilizando as ferramentas BLAST (Altschul *et al.*, 1990) e Clustal Omega (Sievers *et al.*, 2011). A comparação inicial do alinhamento obtido com a estrutura tridimensional de alta resolução da proteína homóloga em humanos foi realizada com a plataforma ESPript 3.0 (Robert e Gouet, 2014).

As estruturas tridimensionais dos domínios previamente preditos para a proteína de interesse foram construídas através de modelagem nos servidores online I-TASSER (Yang e Zhang, 2015; Zhang *et al.*, 2017); Swiss Model (Waterhouse *et al.*, 2018) e, posteriormente, foram submetidos à métricas de validação na plataforma SAVES (Laskowski *et al.*, 1993; Colovos *et al.*, 1993). Para avaliar as modelos resultantes em nível de comparação, foi realizado o alinhamento estrutural com modelos experimentais depositados no banco de dados *Protein Data Bank* (PDB) e, também, do servidor AlphaFold (Jumper *et al.*, 2021).

RESULTADOS E DISCUSSÃO:

1. Predição de estrutura secundária e detecção de domínios

A predição de estrutura secundária obtida a partir da sequência de aminoácidos da SbrPAP3 sugere que esta é composta majoritariamente por hélice-alfa (figura 2A). Esse dado se encontra em conformidade com os resultados obtidos para a proteína de interesse em experimentos prévios de espectropolarimetria de dicroísmo circular, os quais indicaram que SbrPAP3 se constitui com 53% desse tipo de estrutura.

Através de preditores diferentes, foi possível localizar dois perfis de repetições característicos dos motivos TPR entre os resíduos 39 - 72 e 72 - 139, formando um domínio TPR para a proteína de estudo entre os resíduos T39 e N139 e, além disso, um desses preditores identificou, também, um possível domínio RPAP3 entre os resíduos K264 e S352 (figura 2B).

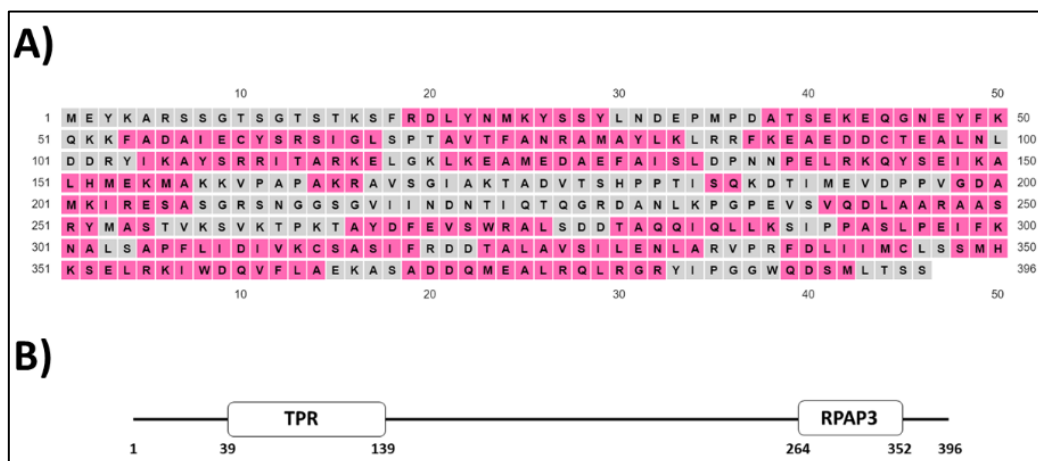


Figura 2. Sequência proteica. **A) Sequência de aminoácidos da proteína RPAP3 de *Sorghum bicolor* (acesso NCBI XP_002440525).** As regiões previstas com estrutura secundária do tipo hélice- α se encontram destacadas em rosa e, em cinza, se encontram as estruturas do tipo *coil*. Ao todo, esta sequência se apresenta com 396 aminoácidos, resultando numa massa molecular teórica de 44 kDa. **B) Representação esquemática dos domínios proteicos.** Domínio TPR da SbRPAP3 previsto do resíduo 39 ao 139 e um possível domínio RPAP3 identificado entre os resíduos 264 e 352.

2. Predição tridimensional dos domínios TPR e RPAP3

Os modelos tridimensionais dos domínios da SbRPAP3 construídos através do preditor que utiliza a metodologia por *threading* (I-TASSER) se assemelham muito às estruturas resolvidas experimentalmente da proteína homóloga em humanos.

O domínio TPR apresenta um perfil estrutural característico contendo um padrão de repetição em hélice-volta-hélice. Os resíduos da SbRPAP3 envolvidos na ligação ao motivo MEEVD presente no C-terminal da Hsp90 são altamente conservados (K42, N46, N76, K83, K106 e R110) (Henri *et al.*, 2018) e, tanto para o modelo construído da proteína de sorgo quanto para estrutura tridimensional experimental da homóloga em humanos, têm-se a cadeia lateral desses aminoácidos apontando para o mesmo lado (figura 3A).

Quanto ao domínio RPAP3, a predição estrutural da região entre os resíduos 264 – 352 se aproxima muito àquela da porção C-terminal da proteína RPAP3 humana, a qual possui um domínio intitulado RBD (do inglês, *Rvb2-binding domain*). O RBD contém os resíduos responsáveis (R623, M626, F630 e S632) pela interação com o domínio de atividade ATPásica de uma das helicases que constitui o complexo R2TP (Maurizy *et al.*, 2018). O alinhamento entre as sequências da proteína aqui estudada e a homóloga humana mostra que tais resíduos são parcialmente conservados entre as respectivas espécies, entretanto, ao sobrepor suas estruturas, as cadeias laterais permanecem voltadas ao mesmo plano espacial (figura 3B).

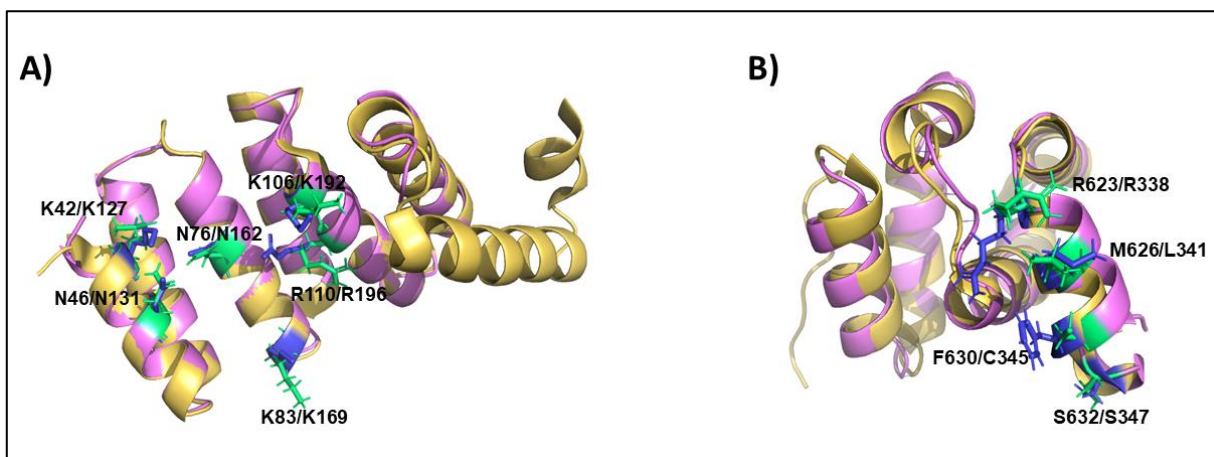


Figura 3. Alinhamento estrutural entre os modelos tridimensionais da proteína de interesse construídos pela plataforma I-TASSER (rosa) com estruturas de alta resolução obtidas experimentalmente dos domínios da RPAP3 humana (amarelo). A) Domínio TPR. Predição estrutural da região que engloba os resíduos 39 – 139 da proteína de estudo demonstrando um padrão de hélice-volta-hélice. Resíduos responsáveis pela ligação ao motivo MEEVD destacados em verde para a RPAP3 de sorgo e, em azul, para a homóloga em humanos (PDB 6GXZ); RMSD = 0,306 Å. **B) Domínio RPAP3.** Estrutura predita para a região entre os resíduos 264 – 352 da proteína de interesse. Resíduos responsáveis pela interação com a proteína Rvb2 destacados em verde para a RPAP3 de sorgo e, em azul, para a proteína homóloga em humanos (PDB 6EZ4). RMSD = 0,944 Å.

CONCLUSÕES:

A predição de estrutura secundária aponta SbRPAP3 como uma proteína composta majoritariamente por hélices- α , corroborando com o dado experimental que calcula 53% desse tipo de estrutura para a proteína em questão. Essa predição se relaciona com os resultados subsequentes de detecção e modelagem estrutural dos domínios contidos na SbRPAP3, os quais são muito importantes para a interação dessa proteína com a chaperona Hsp90 e a helicase Rvb2. As posições espaciais e a similaridade entre os resíduos responsáveis por essas interações proteína-proteína apontam a possibilidade desses eventos ocorrerem experimentalmente, os quais serão devidamente testados em futuros projetos de pesquisa.

BIBLIOGRAFIA

1. Kakahara, Y., Houry, W.A. (2012). The R2TP complex: discovery and functions. *Biochem. Biophys. Acta.*1823: 101-107.
2. Henri, J., Chagot, M-E., Bourguet, M., Abel, Y., Terral, G., Maurizy, C., Aigueperse, C., Georgescauld, F., Vandermoere, F., Saint-Fort, R., BehmAnsmant, I., Charpentier, B., Pradet-Balade, B., Verheggen, C., Bertrand, E., Meyer, F., Cianféroni, S., Manival, X., Quinternet, M. (2018). Deep Structural Analysis of RPAP3 and PIH1D1, Two Components of the Hsp90 Co-chaperone R2TP Complex. *Structure.* 26, 1196-1209.
3. Maurizy, C.; Quinternet, M.; Abel, Y. *et al.* (2018). The RPAP3-Cterminal domain identifies R2TP-like quaternary chaperones. *Nat Commun.* 9, 2093.
4. Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195-202.
5. Sigrist, C. J. A., de Castro, E., Cerutti, L., Cucho, B. A., Hulo, N., Bridge, A. J., Bougueleret, L., Xenarios, I. (2012). New and continuing developments at PROSITE. *Nucleic Acids Res.*, 41.

6. Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Mulder, N., Natale, D., Orengo, C., Quinn, A. F., Selengut, J. D., Sigrist, C. J. A., Thimma, M., Thomas, P. D., Valentin, F., Wilson, D., Wu, C. H., Yeats, C. (2009). InterPro: The integrative protein signature database. *Nucleic Acids Res.* (D1), D427-D432.
7. J Yang, Y Zhang. (2015). I-TASSER server: new development for protein structure and function predictions, *Nucleic Acids Research*, 43: W174-W181.
8. C Zhang, PL Freddolino, Y Zhang. (2017). COFACTOR: improved protein function prediction by combining structure, sequence and protein–protein interaction information. *Nucleic Acids Research*, 45: W291-W299.
9. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46(W1), W296-W303.
10. Laskowski R. A.; MacArthur M. W.; Moss D. S.; Thornton J. M. (1993). PROCHECK – a program to check the stereochemical quality of protein structures. *J. App. Cryst.* 26, 283-291.
11. Colovos C.; Yeates T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2(9):1511-9.
12. Jumper, J *et al.* (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*.
13. Altschul, S. F., Gish, W., Miller, W., Myres, E. W., Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
14. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol.* 7, 539.
15. Robert, X., Gouet, P. (2014). Deciphering key features in protein with the new ENDscript server. *Nucleic Acids Res.* (W1), W320-W324.