Florian DE POL*, Elise PIERSON, Johan WOUTERS
Laboratoire de Chimie Biologique Structurale (CBS), Namur Research Institute for Life Sciences (NARILIS), University of Namur (UNamur), 5000 Namur, Belgium - florian.depol@vuw.ac.nz

Protein oligomerisation: why and new avenues for drug discovery

1. Introduction

Protein self-assembly is one of Nature's most elegant ways to provide cellular multi-tasking. Protein oligomerisation is indeed ubiquitous in all three domains of life and is crucial for a wide spectrum of biological functions [1-5]. In this article, we aim to briefly outline why protein oligomerisation occurs by using several illustrative examples from reviewed literature. We will then discuss how the in-depth study of quaternary structures creates exciting avenues for drug discovery and the current challenges of this emerging strategy.

2. Why protein oligomerisation occurs: examples of biochemical advantages for the cell

For the cell, protein self-assembly offers a plethora of evolutionary advantages that has seen monomeric proteins evolve to more sophisticated structures. Evolution has naturally driven the formation of oligomeric proteins for two main reasons. The first biological profit of protein oligomerisation is the ability to form large and stable macromolecules without extending the amount of genetic material to transcript [6]. Indeed, a large protein complex composed of short polypeptidic sequences is more likely to be error-free transcripted than a longer sequence of comparable size. Additionally, oligomeric proteins have an enhanced overall thermal stability thanks to the enthalpic gain of forming stabilising protein-protein interactions (PPIs) between the subunits and the augmentation of entropy following the shielding of hydrophobic residues from the solvent [7]. In the complex, the solvent accessible surface of the monomeric subunits is indeed reduced, which could make them less prone to denaturation. In turn, this may allow the oligomers to carry out their function for a longer cellular timeframe, which reduces the need to translate DNA to replenish the cellular environment with fresh proteins. The second main biological profit of protein oligomerisation is the extended ability for the cell to adjust protein function and activity. It is now well established that the correct assembly of a precise number of subunits is required for the oligomeric protein to carry its physiological function. Disrupting protein interfaces through either small molecules binding, temperature, pH or ionic strength variations is a natural way of regulating oligomeric protein function. This is the interesting case of *moonlighting proteins*, defined as proteins able to carry more than one physiologically relevant function

https://doi.org/10.52809/cn2022/YHHW7080

with a single polypeptidic chain. In that sense, protein oligomerisation provides an efficient way for the cell to multitask. For example, the enzyme ubiquitin carboxy-terminal hydrolase L1 (UCHL-1), an enzyme studied in our laboratory, shows interesting moonlighting properties: when in a monomeric form, it is a hydrolase that catalyses the deubiquitination of proteins. Under its dimeric form however, it has been shown to catalyse the reverse reaction, which makes it a ligase [8]. Additionally, allosteric regulation can be delivered through the formation of oligomers. A striking example is the human ribonucleotide reductase (hRNR), which is naturally found as heterodimeric. An allosteric regulatory pocket is located at the dimeric interface and the enzymatic activity is regulated according to the molecule that binds to it: the binding of adenosine triphosphate (ATP) to the allosteric site triggers the reduction of cytidine diphosphate (CDP) while the binding of thymidine triphosphate (TTP) to the allosteric site stimulates the reduction of guanosine diphosphate (GDP) [9].

3. Targeting quaternary structures for drug discovery: extended druggability and selective toxicity

As we have exemplified in the previous section, the function, activity, and regulation of an oligomeric protein is inherently related to the correct assembly of a precise number of subunits to form a fully functional complex. Since 30-50% of known proteins oligomerise [6], the targeting of quaternary structures was recently regarded with great interest for therapeutic purposes. Indeed, one of the most important criteria for the validation of a drug target is its druggability (i.e. its ability to bind small molecules with high affinity and specificity). It is not uncommon for potential therapeutic targets to be abandoned due to the non-druggability of their active site. Thus, considering oligomeric interfaces as potential allosteric regulatory sites widens the scope of possibilities in terms of druggability. Take Rad52, a cyclic heptameric protein involved in DNA repair, as an example. This protein is considered as a therapeutic target for certain breast cancers chemotherapies. However, its DNA binding site has been considered undruggable. Later, it was shown that 6-hydroxy-DL-dopa could bind at the protein-protein interfaces of the Rad52 heptamer. The binding of this small molecule to the heptameric structures induces the dissociation of the heptameric ring, subsequently resulting in a loss of the DNA repair function. From this perspective, Rad52 could be validated as a therapeutic target [10]. Targeting quaternary structures also creates exciting avenues to achieve high levels of selective toxicity towards human pathogens. Indeed, it is common for the active site of the same protein family to be highly conserved due to the similarity in the chemical reaction they catalyse. Targeting the active site therefore leads to undesirable effects because the inhibitor also affects the other proteins of the same family. In contrast, it has been shown that oligomeric interfaces are often less conserved than the active site [11]. In that sense, protein-protein interactions constitute an extended pool of molecular targets to be exploited to enhance selectivity for a single member of a given protein family. For example, protein kinases are a family of proteins that share a highly conserved ATP binding site. Therefore, developing a selective inhibitor for a particular protein kinase is a major challenge. In contrast, targeting their oligomeric interfaces offers interesting prospects in terms of selectivity. A particularly relevant example is B-Raf, a dimeric protein kinase whose disfunction in tumor cells can lead to abnormal hyperactivated Mitogen Activated Protein Kinase (MAPK) signalling. Knowing the nature of the protein-protein interactions at the dimeric interface of B-Raf, Gunderwala A.Y. *et al* [12] designed a small peptide, namely "*Braftide*", capable of selectively lodging itself at the protein-protein interface of the B-Raf dimer without affecting the normal activity of other kinases of the same family. Disruption of B-Raf into its constituting monomers inhibits the phosphorylation of MAPK and additionally, triggers the proteasome mediated degradation of B-Raf – Figure 1.

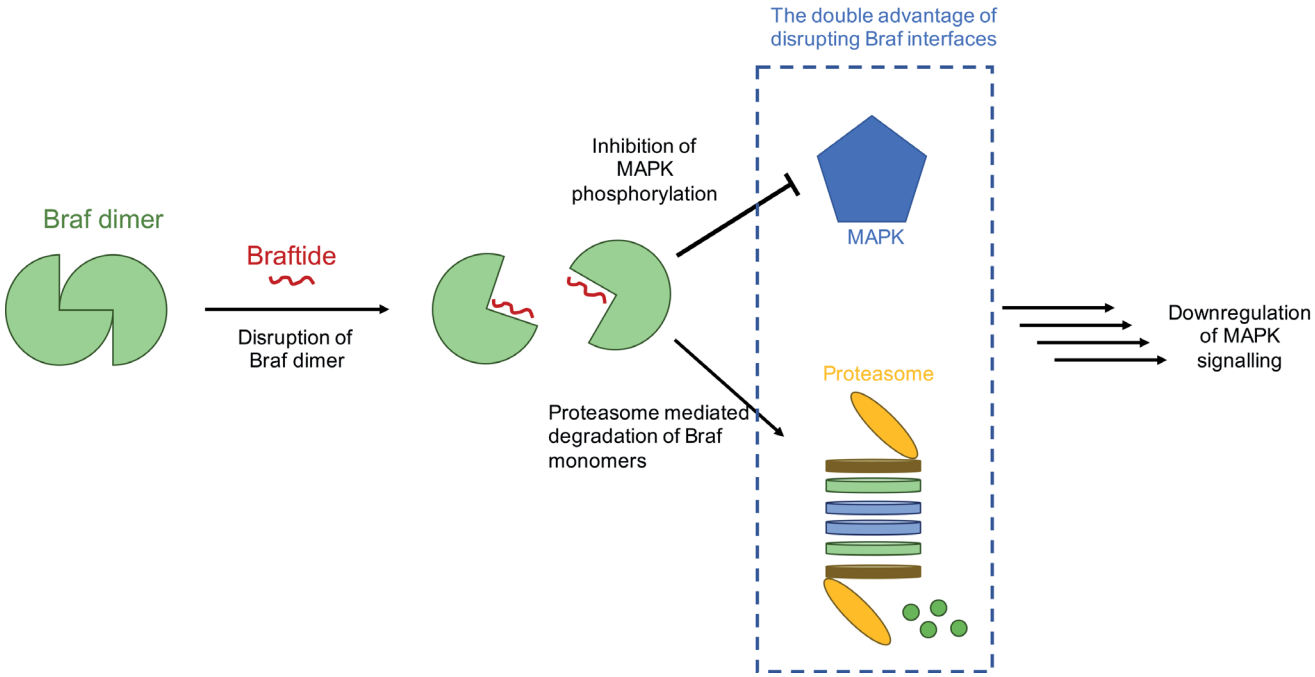

Figure 1 – Mechanism of action of Braftide. By mimicking the dimeric interactions in B-Raf, the small peptide called Braftide can disrupt the dimer of B-Raf, which leads to the inhibition of MAPK phosphorylation and the degradation of B-Raf monomer. The combined effects lead to the downregulation of the MAPK signaling pathways. Adapted from [12]

Additionally, targeting protein quaternary structures circumvents a frequent problem when it comes to targeting a protein from a pathogenic microorganism: if the therapeutic target has a human homologue, it is crucial for the drug to interact selectively with the pathogen's protein. In the context of research against Chagas disease, whose pathogen is the microorganism Trypanozoma cruzi (TP), the triose phosphate isomerase has been identified as an interesting therapeutic target because it is involved in the fifth step of glycolysis [13]. It is therefore essential for the survival of TP. However, this protein has a human equivalent with a highly conserved active site. Unsurprisingly, targeting the active site of this enzyme with an inhibitor therefore proved to be not selective towards the human homologue. This represents a major problem as it could lead to the co-lateral inhibition of the human counterpart and thus induce unwanted side-effects during the treatment. However, Tellez-Valencià *et al*. found that benzothiazole derivatives could bind with high selectivity to a non-catalytic pocket located at the dimeric interface of the TP triose-phosphate isomerase, which resulted in its full inhibition. By X-ray diffraction (XRD), it was shown that these derivatives could interact with residues at the dimeric interface that are not redundant in the human counterpart [14].

These examples illuminate the growing interest to use protein-protein interactions and interfaces as a means to synthetically control protein functions for therapeutic purposes. Especially, a major milestone was reached in 2016 with the approval by the FDA of Venetoclax, the first completely synthetic drug targeting PPIs designed with a fragment-based approach [15]. Venetoclax is used in the treatment of chronic lymphocytic leukemia : in tumor cells, pro-apoptotic proteins like BAK or BAX are bound to anti-apoptotic proteins like BCL-2 [16]. Venetoclax works by selectively disrupting these interactions and thereby releasing the pro-apoptotic BAK/BAX proteins, which further promotes cell death. Fig. 2. By being approved by the FDA and showing outstanding results in the treatment of patients, Venetoclax serves as a first proof of concept that therapeutics can indeed be designed to target

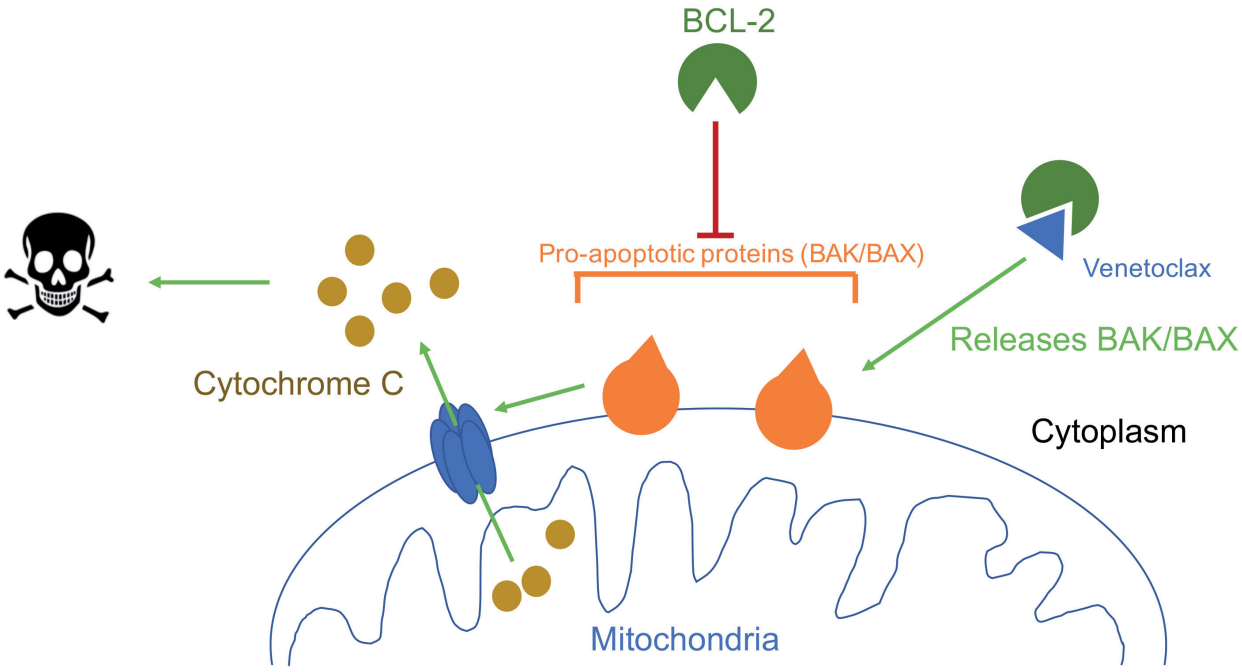

Figure 2 – Mechanism of action of Venetoclax. In tumor cells, the anti-apoptotic protein BCL-2 antagonises the pro-apoptotic proteins BAK and BAX to block the apoptotic cascade. Binding of Venetoclax to BCL-2 releases BAK and BAX, which in turn trigger the apoptotic cascade leading to the release of cytochrome C in the cytoplasm and, hence, cellular death. Adapted from [17].

protein-protein interfaces with a similar level of efficiency and safety as those classically designed to target active sites.

4. Discussion

Protein oligomerisation is crucial for a large range of cellular processes. This phenomenon has been retained during evolution as it provides several advantages for the cell. In the case of oligomeric proteins, the assembly of a defined number of subunits in a precise architecture is critical to enable activity, and a switch or disruption of the oligomeric architecture may disable functional properties. For this reason, the last few years have seen numerous works showing that the targeting of quaternary structures for drug purposes is a viable strategy that provides an ingenious alternative to the classic active site inhibition approach - especially when the latter cannot be properly drugged. This revolutionary paradigm currently stands at the edge of innovation and bears golden opportunities for drug discovery, including extended druggability and selective toxicity towards pathogens [16]. However, this nascent strategy is still facing several challenges. Arguably, a major obstacle is related to the lack of reliability regarding the *in silico* modelling of quaternary structures. Indeed, the currently available computational tools are unable to support experimental data with accurate and reliable quaternary structures predictions. Recent developments of Alpha Fold Multimer - although yielding ground-breaking results against benchmark protein-protein docking algorithms - are still in their optimization phase and struggle with modelling 50% of its homomeric targets. Particularly, sequence-based predictions of oligomeric targets are particularly unreliable when no homologue structures exist under the same oligomeric state. Additionally, although progress is being made in that direction, targeting protein oligomerisation still lacks a thorough and well-established workflow for the drug design of small molecules or peptides targeting quaternary structures.

Nonetheless, we can be enthusiastic about the exciting perspectives that this approach brings to drug discovery. Further research in that area will shake the status quo and break new grounds. We believe that advances in structural biology regarding the formation of quaternary structures and the development of new drugs targeting protein-protein interfaces will co-fertilise and accelerate discoveries in both areas, for the better.

Part of our work at the *Laboratoire de Chimie Biologique structurale* (CBS) focuses on proteins implied in various diseases, including cancer [18-20], neurodegenerative diseases and infectious diseases [21-24]. Our efforts concentrate on understanding the structural properties of our target enzymes and how their regulated catalysis works. Our objective is to synthetically control protein activity for therapeutic purposes. Recently, an essential metabolic enzyme from *Mycobacterium tuberculosis* gained our attention and was thoroughly characterized. Using experimental techniques (*Multi Angle Light Scattering, Small Angle X-Ray Scattering, Size Exclusion Chromatography, Native-PAGE and Site-Directed Mutagenesis*) together with bioinformatic tools (*protein-protein docking, homology modelling, in silico mutagenesis*) we deciphered a unique and unexpected oligomeric equilibrium that could be the natural mechanism for the allosteric regulation of the enzyme. Our findings paved the way to a novel route for the synthetic inhibition of the target enzyme and created inspiring alternatives to avoid the undesirable co-inhibition of the human homolog. For this work, we were awarded the *Prix de la Société Royale de Chimie* in 2022.

References

[1] Sara G., Princeton Insights, 2021, online : https://insights.princeton.edu/, visited 23/03/2023

[2] Robinson P. K., Essays Biochem., 59, 2015, 1-41

[3] Keller B., Plant. Physiol., 101, 1993, 1127-1130

[4] Oxender D.L., Biomembranes, 1974, 25-79

[5] Schenk P.W., et al, Biochim. Biophys. Acta, Mol. Cell Res., 1449(1), 1999, 1-24

[6] Mayssam H., et al, Bioorganic Med. Chem. Lett., 5, 2005, 5013–5020

[7] Honig B., et al, Adv. Protein Chem. Struct. Biol., 46, 1995, 27-58

[8] Zhang Z., et al, Bioinformatics, 28(5), 2012, 664-671

[9] Nozomi A., et al., Biochemistry, 55(2), 2016, 373–381

[10] Chandramouly G., et al, Chem. Biol., 22, 2015, 1-14

[11] Sridhar R., et al, Pharm. Res., 17(11), 2000, 1345-1353

[12] Gunderwala A.Y., et al, ACS Chem. Biol.,14, 2019, 1471-1480

[13] Aguilera E., et al, ChemMedChem, 11, 2016, 1328 – 1338

[14] Téllez-Valencia A., et al, J. Mol. Biol., 341(5), 2004, 1355-1365

[15] Green D.R., Cell, 165, 2016

[16] Scheffold A., et al., Small Molecules in Hematology, 3rd Edition, Springer, 2021, 215-242

[17] Thabault L., et al., Drug Disc. Today, 26(5), 2021, 1148-1163

[18] Mirgaux M., et al., Acta Cryst. D., 76, 2020, 1211-1221

[19] Mirgaux M., et al., Int. J. Tryptophan Res., 14, 2021

[20] Thabault L., et al., Eur. J. Med. Chem., 230(3), 114102

[21] Haufroid M., et al., Eur. J. Med. Chem., 245(8), 114935

[22] Haufroid M., et al., Pharmaceuticals, 12(66), 2019

[23] Pierson E., et al., Biophys. Biochem. Res. Commun., 530(4), 2020, 739-744

[24] Pierson E., et al., Molecules, 25(2), 2020, 415