# MS8 CIB TIP-UPJS know-how increasing - written report based on WP1 deliverables

Name of project: Fostering high scientific quality in protein research in Eastern Slovakia (**CasProt**)

No. 952333

# MS8 CIB TIP-UPJS know-how increasing - written report based on WP1 deliverables

Name of project: Fostering high scientific quality in protein research in Eastern Slovakia (**CasProt**) No. 952333

| | |
|---|---|
| Deliverable No: | Milestone 8 |
| Work package No. and title: | WP1 – Research activities and know-how transfer |
| Nature: | Report |
| Dissemination level: | Public |
| Lead Beneficiary: | UPJS |
| Author(s): | doc. RNDr. Erik Sedlák, DrSc. |
| Date of publication: | 30 September 2023 |

WP1 leader: doc. RNDr. Erik Sedlák, DrSc.

Project coordinator: doc. RNDr. Erik Sedlák, DrSc.



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement N° 952333

## EXECUTIVE SUMMARY

This report addresses the know-how increase in four specific topics of directed evolution of GPCRs: (i) Selection and ad hoc design of suitable GPCR candidates, (ii) Expression and purification of GPCRs (e.g., -opioid receptor, NTR), (iii) Rational identification hot spots in a selected GPCRs for evolution, and (iv) Application of protein evolution methods, e.g., yeast display, ribosome display to enhance the solubility of proteins. The report further addresses three specific topics of the development of single-molecules force assays: (i) Technical assistance and know-how support in the development of specific biological assays for membrane proteins, preparing materials and modifications for single-molecule experiments, (ii) Hands-on design and case-study of instrument development for single-molecule force/fluorescence applications, and (iii) Advanced instrument and software development for force and fluorescence experiments with selected GPCRs.

### Selection and ad hoc design of suitable GPCR candidates

Based on our numerous experimental analyses of evolution of two types of GPCRs – human kappa opioid receptor (hKOR) and neurotensin binding receptor 1 (NTR1), and analyses of published data by other scientific groups, we concluded that “an evolvable” GPCR system consisted of a pair of the protein and corresponding ligand should have the following minimal properties:

(i) The crystal structure of the GPCR is known,

(ii) The GPCR protein is monomeric in its physiological state,

(iii) The GPCR protein ligand does not carry significant electrostatic charge.

These minimal properties are critical for:

(i) a proper design of a library of the GPCR as the known protein crystal structure enables to determine suitable amino acids positions for a replacement. The hydrophobic amino acids buried in the phospholipid bilayer need to be replaced by water soluble amino acids, which will lead to an increased aqueous solubility of the evolved GPCR,

(ii) a proper selection process because a formation of dimeric states would critically affected and even prevented the evolution of such GPCR,

(iii) a reduction of a nonspecific interaction of the members of protein library with the surface on plates or magnetic beads in the process of evolution.

One more critical criterion for successful evolution is a proper design of the protein library. This significantly depends on a number of positions of which the hydrophobic amino acids need to

be replaced by the water-soluble amino acids. Based on our experiences, this modification has to cover a majority of the protein surface exposed to the hydrophobic environment of the lipid bilayer.

**Expression and purification of GPCRs**

As it has been shown by the group of Prof. Plückthun, certain member of GPCR protein family can be expressed in the bacterial hosts such as *Escherichia coli (E.coli*). This approach has been successfully demonstrated also in our laboratories by using the *E.coli* strain BL21 (DE3) in purification of hKOR and NTR1. For purification of these GPCRs we have: (i) created a conjugate of the GPCR with maltose-binding protein (MBP) and (ii) developed a matrix conjugated with DARPIN off7, which has been developed at UZH in Prof. Plückthun laboratory for specific and high-affinity binding of MBP. Created affinity column has been successfully used for the purification of the GPCR proteins in one-step purification. While we showed feasibility of this approach, we are in the process of an establishment of the laboratory in which we will use the expression system of baculovirus-insect cells for expression of the GPCR proteins. The necessary knowledge for an establishment of this system in CIB laboratories had been obtained during the study visits at UZH in the frame of the CasProt project. The baculovirus-insect cells expression system provides larger versatility of expressing not only of different GPCR proteins but also other eukaryotic proteins.

**Rational identification hot spots in selected GPCRs for evolution**

One of the ambitious goal of the project has been a determination of basic rules for a rational design/identification of hot spots in selected GPCRs. While we believe that this goal is achievable, in the course of the project we were unable to reach it. We believe that such goal is possible to attain by an applying of artificial intelligence (AI). However, for learning of AI we would need to have many more data obtained from an extensive analysis the GPCR evolution experiments. For that, we would need larger capacity including more researchers and time to obtain sufficient amount of primary sequence with corresponding conformational analysis of high-throughput purified GPCRs.

**Application of protein evolution methods to enhance the solubility of proteins**

The obtained amino acid sequences of the sequenced hKOR mutants were also analysed using available web servers designed for the prediction of protein properties. The CamSol method (https://www-cohsoftware.ch.cam.ac.uk) was used to predict the impact of individual mutations on protein solubility and the tendency to aggregate. In its algorithm, this method takes into account, in addition to the hydrophobicity of individual amino acids, their total charge, as well as

their tendency to form the secondary structure of the protein and other external and internal factors affecting the overall solubility. The result of this method is an overall protein solubility profile that allows the identification of critical aggregation positions within the protein sequence.

Another tool used in bioinformatics analysis was ProtParam (https://web.expasy.org/protparam/). It is a tool that, based on the inserted protein sequence, enables the calculation of various physical and chemical parameters of the given protein, such as molecular weight, theoretical pI or extinction coefficient. In our case, we focused on the parameters related to the solubility and stability of proteins, namely the overall GRAVY hydropathy index and the instability index. GRAVY's total hydropathy index is used to express the hydrophobicity of the protein, while for the calculation it uses the sum of the hydropathy values of all amino acids, which is divided by the length of the sequence. In the case of hydrophilic proteins, the value of the GRAVY index acquires negative values, on the contrary, in the case of hydrophobic proteins, these values are positive. The instability index provides an estimate of the stability of a protein in a test tube. It is assumed that a protein that acquires index values less than 40 is considered stable, index values above 40 indicate protein instability. We compared these parameters, which were determined for all our mutants, with the same parameters determined for selected membrane and water-soluble proteins. From the correlation between GRAVY and the instability index, we concluded an increase in solubility of certain selected GPCRs in comparison to other hydrophobic proteins. We could also identify several mutants whose instability index values indicate increased stability, or only moderate instability compared to these values for other membrane proteins.

Above described bioinformatics tools have been very helpful in identification a progress in the GPCR evolution towards water-soluble proteins and also parts of GPCRs, which were responsible for formation of oligomeric structures of evolved proteins.

Ribosome display proved as a strong directed protein evolution method in such ambitious goal. Its particular advantage is a capacity to use large libraries but on the other side, its disadvantage is low efficiency in a determination of evolved proteins with increased solubility. Consequently, we are in the process of an establishing the yeast display method in our laboratories, in close collaboration with UZH, with an intention to apply this method in the evolution of the water-soluble GPCR. Thanks to the project of the structural funds, we purchased FACS sorter, which inevitable for the yeast display. The yeast surface display method has two critical advantages in comparison with the ribosome display: (i) the fact that the yeast surface will reach only correctly folded GPCRs as a result of yeast chaperone system and (ii) high-throughput analysis of yeasts expressing GPCRs.

The such ambitious project, which involves multiple amino acids replacements, is in the presence unsolvable without utilization of the directed evolution methods such as ribosome and yeast displays.

**Technical assistance and know-how support in the development of specific biological assays for membrane proteins, preparing materials and modifications for single-molecule experiments**

Our preliminary work with nanodiscs has produced some intriguing results, but these findings have not yet been published due to the fact that the work is still in its infancy. Due to the preliminary nature of these findings, further extensive experimental work is required for validation. It is essential to determine whether or not our observations coincide with these preliminary results. The cleaning of the nanodiscs is going to be an essential part of these additional investigations. It is necessary for us to ascertain whether or not they contain a single embedded protein. For this confirmation, we need to cross-reference our data using a different method that is not dependent on us. This will ensure that our findings are accurate. The improvement of DNA-protein tether capture is another area that calls for attention. Increasing the productivity of our experiments using single-molecule tweezers requires that we streamline this process as much as possible.

In order for these experiments to be successful, there are a number of variables that need to be optimized. Due to the conditions of the solvent, careful calibration is required. In addition to this, careful consideration should be given to the existence of detergents as well as the optimum levels of protein concentration. It is especially important to do this in order to avoid the possibility of multiple tethers forming between the microspheres and our protein construct, which can cause the results to be inaccurate.

Another factor of the utmost importance is making sure that the constructions we are working with are stable. This stability needs to be cross-verified using conventional biochemical or biophysical experiments in order to guarantee that the constructs will maintain their integrity throughout the process of the experimentation.

In addition, despite the fact that we understand the significance of quality control in relation to our methodology, it is essential to point out that the control checks that are already in place might not be directly applicable. The solubility of proteins was a primary consideration in the development of many of these methods. Due to the complexities of membrane proteins, it is

possible that these methods cannot be directly transferred; therefore, we will need to develop new protocols that are specifically adapted to meet our requirements.

## Hands-on design and case-study of instrument development for single-molecule force/fluorescence applications

Provision of technical expertise and knowledge transfer in formulating specific biological assays for proteins/membrane proteins by establishing experiments using lipid nanodiscs, and in preparing and modifying materials for single-molecule experiments had been undertaken. We found useful to separate this research task into two interrelated parts:

1. successful development of biological assays and modifications for laser optical tweezers using multi-domain protein Hsp70
2. modification and assay development using lipid nanodiscs as a prerequisite for successful optical manipulation of membrane proteins.

**Significance**. Structural changes in proteins are essential for the regulation of their function. Hsp70 is a two-domain protein that binds ATP nucleotide in its nucleotide-binding domain (NBD). Protein substrates are bound in the second, so-called substrate-binding domain (SBD). Hsp70 hydrolyses ATP to ADP; by doing this, Hsp70 undergoes significant changes in the SBD. Using a single Hsp70 molecule, we showed how the mechanics of Hsp70 is changed during ATP hydrolysis by using laser optical tweezers. We find that DnaK is mechanically stable and rigid in an ADP-bound state. However, in the ATP-state, one domain is mechanically destabilized as the result of interdomain docking, followed by the partial unfolding of the SBD. By observing the folding state of the SBD, we observe the enzyme's continuous ATP/ADP cycling in real-time with a single molecule. The SBD lid closure is strictly coupled to the chemical steps of the ATP hydrolysis cycle even in the presence of peptide substrate.

Large proteins are made of multiple domains that are organized like beads on a string. In some cases, domains can interact with each other in a ligand-dependent manner, which results in the regulation of the function of the distal domain. The result of such long-range domain-domain interaction can be, for example, a change in the affinity of a ligand, binding site accessibility, and activation/ suppression of a protein activity. These examples are well described by the concept of allostery. The existence of allosteric control of functional properties of proteins is long well-known from a large number of biochemical and biophysical studies. For some intensively studied proteins, molecular mechanisms and pathways were elucidated using mutagenesis. However, revealing fundamental mechanical principles and energetics of long-range allosteric ligand-induced coupling

between subdomain interfaces is challenging. The scale of conformational changes in proteins varies significantly; conformational changes can be local at residue level (e.g., side-chain rotations), native-state rearrangements (e.g., monomer/dimer equilibria and opening/closing of binding sites), and extensive conformational transitions (e.g., unfolding of polypeptide chain). Under favorable circumstances, single-molecule force spectroscopy can provide an efficient tool to observe mechanical and structural changes in proteins at the sub-nanometer scale. The development of highly sensitive differential detection dual-beam laser tweezers enabled observing folding/unfolding transitions of many proteins and observing subnanometer-sized opening and closing of the active site of adenylate kinase.

The primary benefit of the force technique is the possibility of controlling the state of the protein structure and determining its propensity to undergo structural re-arrangements. The immediate outcomes of single-molecule assays using optical tweezers measurements are two parameters describing the observed conformational changes: force and distance. Under equilibrium conditions, force and distance fluctuations completely specify the underlying energy landscape of the observed conformational changes, their minima, barriers, and transition path times. In our case studies (in details described in D1.4 report), we employed laser optical tweezers to observe and monitor ATP-induced allosteric changes in the domain conformations of a well-studied, two-domain protein, DnaK.

**Advanced instrument and software development for force and fluorescence experiments with selected GPCRs**

During the most recent phase of our investigation, we have only recently begun to carry out experiments that contribute to our mechanistic comprehension of how nanodiscs interact with proteins. It is essential to note that despite the fact that our preliminary findings have been promising, these discoveries have not yet been conclusively validated or published. This is primarily due to the fact that this research is still in its early stages.

Purifying the nanodiscs to check that each one only has a single protein molecule embedded in it will be an essential part of the upcoming research that we are conducting. It is of the utmost importance that the constructs we use be checked for both their stability and their integrity. An essential realization is that a lot of our methods, which are based on the solubility of proteins, might not directly apply to the complexities of membrane proteins. This is an important realization. This is a ground-breaking insight, which hints at the requirement for new protocols that are individualized. Our investigation has also delved into the inner workings of enzymes such as DnaK,

illuminating the mechanisms by which ATP can induce mechanical and structural changes in single molecules. This was particularly useful when studying the ATP/ADP cycle in the presence of a peptide substrate. The technical challenges involved in capturing these intricate changes have been significant; however, we have been able to overcome them by making several advancements in our methodology, such as optimizing DnaK variants for single-molecule force experiments.

Despite the fact that we have made significant headway in comprehending the workings of a two-domain DnaK during an ATP/ADP cycle, it has not yet been possible to apply our findings in a practical manner to GPCR, also known as G Protein-Coupled Receptors. This is an astonishing fact. GPCRs are an important category of molecular targets, and the work that we have done so far has the potential to pave the way for further research involving GPCRs. The progress we have made in understanding nanodiscs and the way in which they interact with proteins has the potential to serve as the stepping stone that will be required to make significant advances in GPCR research in the near future.