Final report of Project No. K119359

(S100 PROTEINS AND CELL SIGNALING: STRUCTURE-FUNCTION STUDIES)

Our main research aim was to further explore the structure-function relationships and the interactome of the S100 protein family. Previously, we have characterized the S100A4-NM2A and the S100B-Rsk1 complexes and determined their high-resolution crystal structure (Kiss et al, PNAS 2012, Gogl et al, JBC 2016). In this project all our successful experiments originate from the above two work. Since basic research follows a meandering path, in the final report I will describe our results organized in chapters not following of the original research plan, but presenting them broadly in chronology. Some of the planned experiments were not successful, however others happened to be more fruitful as far as research progress is concerned. I will focus mainly on published result and only mention a few unpublished ones.

1. Regulation of non-muscle myosin filaments

The assembly of non-muscle myosin II paralogs (NM2A, 2B, 2C) into homo- and heterotypic bipolar filaments in living cells is primarily regulated by phosphorylation of the N-terminally bound regulatory light chain. In our studies we provided evidence that the equilibrium between these filaments and single NM2A and NM2B molecules can be controlled via calcium-dependent S100 protein interactions and phosphorylation at the C-terminal end of the myosin heavy chains. Furthermore, we showed that in addition to S100A4, other members of the S100 family can also mediate disassembly of homotypic NM2A filaments. Importantly, these proteins can selectively remove NM2A molecules from heterotypic filaments. We also found that tail phosphorylation (at Ser-1956 and Ser-1975) of NM2B by casein kinase 2, as well as phosphomimetic substitutions at sites targeted by PKC and TRPM7 down-regulates filament assembly in an additive fashion. Tail phosphorylation of NM2A had a comparatively minor effect on filament stability. S100 binding and tail phosphorylation, therefore preferentially disassemble NM2A and NM2B, respectively. These two distinct mechanisms are likely to contribute to the temporal and spatial sorting of the two NM2 paralogs within heterotypic filaments. The existence of multiple NM2A-depolymerizing S100 paralogs offers the potential for diverse regulatory inputs modulating NM2A filament disassembly in cells and provides functional redundancy under both physiological and pathological conditions. (**Ecsédi et al, JBC 2018**).

One of our goals was to obtain experimental evidence that S100-dependent NM2A regulation affects cell motility. To find an appropriate system, finally we succeeded with a multicellular contractility assay using Matrigel patterning. In collaboration with András Czirok's group, we were able to provide the first functional demonstration that overexpression of S100A4, which is frequently overexpressed in metastatic tumors, effectively reduces cell contractility by inhibiting NMIIA activity and inducing filament disassembly. In this paper we not only demonstrated that the Matrigel patterning assay is a highly sensitive tool to evaluate cell contractility within a soft ECM environment (using A431 epithelial carcinoma cell line), we also proposed a computational model to explore how cell-exerted contractile forces can tear up the cell-Matrigel composite material and gradually remodel it into a network structure. The assay was calibrated by inhibiting NMII activity directly with blebbistatin or indirectly by a Rho kinase inhibitor, besides the overexpressed S100A4 (**Méhes et al, PLoS Comp Biol 2019**).

2. Role of extracellular S100A4

We have finished our experiments regarding the effect of extracellular S100A4 on cell adhesion. The results are written up and ready to submit for publication. According to our results, S100A4 enters cells through a caveolin-mediated, dynamin-dependent endosomal pathway and accumulates in the

cytoplasm and nuclei. Extracellular S100A4 specifically reduces cell adhesion in a concentration-dependent manner, demonstrating that an intact binding interface and a specific receptor are necessary during the process. Our studies shed light on the internalization process and the extracellular role of this promising therapeutic target protein. The receptor to which S100A4 binds remains to be determined. (B Biri-Kovács, H Vadászi, B Kiss, G Gógl, E Lajkó, O Láng, B Szeder, L Buday, L Kőhidai and L. Nyitray: Internalization of extracellular S100A4 protein reduces cell adhesion of carcinoma cells, FEBS L, in preparation).

3. S100 interactions and structural studies of S100 complexes

We have characterized S100A4 interaction with ezrin. We have shown that S100A4 binds independently to the N-terminal and also to the C-terminal domain of ezrin (no ternary complex is formed). S100A4 co-localizes with ezrin in HEK-293T cells, however it very weakly binds to full-length ezrin in vitro indicating that the interaction of the two proteins requires other regulatory events such as protein phosphorylation and/or membrane binding. As both proteins play an important role in promoting metastasis, the characterization of their interaction could shed more light on the molecular events contributing to this pathological process. (**Biri-Kovács et al, PLoS One 2017**).

We have made considerable progress in structural characterization of S100 complexes using a diverse methodological arsenal, X-ray crystallography NMR-spectroscopy and structural modeling.

We have published our work on interaction of S100 proteins with annexin A2 and its regulation by phosphorylation. We have reported the crystal structure of S100A4-(full-length)-ANXA2 complex (the third asymmetric complex of the S100 family, determined by our research group), an ANXA2 variant phosphorylated at Tyr24, and a phosphomimicking mutant (S26E). Our results shed light on how the alterations in structure and dynamics of the N-terminal domain modulate ANXA2 membrane aggregating function. (**Ecsédi et al, Structure 2017**).

Using the crystallization strategy applied for the S100A4-annexin-A2 complex, we were able to produce suitable crystals of the long-sought complex between S100A4 and the TAD domain of p53. As expected, it is an asymmetric and „fuzzy" complex where the N-terminal region of TAD forms an alpha-helix, while the C-terminal region has a considerably high B-factor. A particularly important methodology development of the whole project is the introduction of annexin A2 as a new crystallization chaperon for structural studies of protein complexes. We compared the crystallization chaperon characteristics of ANXA2 with the most often used MBP and found that the ANXA2 system is highly predictable and has even better crystal forming capability than MBP. We expect that this novel chaperon becomes a useful tool in the protein crystallization community (**Ecsédi et al, Structure 2020**).

In structural biology, solving a high resolution structure of a protein complex with either X-ray crystallography or NMR spectroscopy has drawbacks, therefore getting a particular structure obtained by both methods could strongly validate the conclusions. This is why we initiated in parallel of our crystallization approach, an NMR spectroscopy approach to validate the fuzzy complex (meaning that part of the complex has high internal mobility thus it is invisible in the crystal form) of S100A4 with the intrinsically disordered TAD domain of p53. We were able to get independent and additional structural insight of this fuzzy complex by a hybrid NMR/MD (molecular dynamic simulation) approach. Our results represent an example of how to combine NMR and MD in order to obtain realistic information about complex systems that represent challenges for crystallization; and also when even in solution studies obtaining crucial information is not too straightforward. The atomic level structural characterization of the p53TAD–S100A4 complex by two distinct approaches can contribute to the identification of p53 inhibitor molecules based on S100A4 binding. Our findings

emphasize that, despite the difficulties, structural studies of fuzzy complexes by such hybrid methods can be crucial to understanding the complex functions of hub proteins similar to p53 **(Dudás et al, Chembiochem 2020**). (Note that our results were selected as an illustration on the front cover of the issue of the publishing journal.)

We also succeeded to solve the structure of two S100A6 complexes, with p53TAD and FOP (unpublished). The S100A6-p53TAD complex is particularly interesting (resolution 2.1 Å), since it is a higher affinity complex compared to S100A4-p53TAD that can be explained by the fact that the p53 sequence makes specific interactions with the dimer interface of S100A6 not only with the "canonical" ligand binding sites, and therefore the complex is not fuzzy. The importance of the S100A6-FOP complex is that the FOP (a centrosome protein) has no known structure with any S100 proteins (in the current state the resolution of our structure is only 3.4 Å, which needs to be improved).

In the prestigious Springer Methods in Molecular Biology series we have contributed to the volume titled: "Calcium-Binding Proteins of the EF-hand Superfamily. From Basics to Medical Application". In our chapter we have presented detailed protocols and practical notes how to produce all known human S100 proteins in heterologous expression systems, how to purify them. Moreover, we have provided a detailed description of three in vitro methods to determine the affinity, stoichiometry, and kinetics of S100 protein-protein interactions. (**Kiss et al, Methods Mol Biol 2019**).

We summarized our knowledge of the structural feature of S100 complexes in a review paper, emphasizing their often "fuzzy" and asymmetric character and the methodological arsenal to experimentally study such complexes, which is a difficult task. Depending on the number of partners binding to a dimeric S100, symmetric or asymmetric interactions can occur. We found that the stoichiometry is an indicator of fuzziness and the asymmetric complexes are prone to a high degree of fuzziness (see our published structures of ANXA2–S100A4, NM2A–S100A4, RSK1–S100B and more to be published, such as S100A6-FOP and some others, for which data collection at SLS will be conducted later during this year). We noted that the binding modes of S100 proteins are reminiscent of the interaction properties of 14-3-3 proteins, which can also form homo- or heterodimers and have a specific binding pocket in each monomer that can accommodate various sequences. As for the methods to study these complexes, disappearance of electron density in X-ray structures could indicate fuzzy regions, and, alternatively, preferred conformations of a heterogeneous ensemble can also be trapped in crystal structures; NMR spectorscopy is more ideal for detecting dynamic flexible region, especially, if complemented with molecular dynamic simulations (as exemplified in the case of the S100A4-p53TAD complex). Use of multiple techniques (besides the above, CD, ITC, SAXS) is important because in numerous cases (like in the S100 family), fuzziness indicates the existence of multiple binding partners, and by analyzing each interaction thoroughly, partner-specific surfaces could be recognized. The detailed knowledge of these surfaces could aid the development of inhibitor variants, which could be used to selectively limit the binding of the targeted protein, leaving other functions intact (**Ecsédi et al, Front Mol Biosci 2021)**.

4. S100 interactome: specificity mapping

We have established the first specificity map of the S100ome (20 isoforms). Using a high-throughput fluorescence polarization assay (and validation by ITC measurements), extended functional redundancy was revealed in the S100 family. Based on measuring >250 individual interactions, the S100ome can be grouped into two distinct classes: promiscuous and orphan. In the first group, members bound to several ligands (>4-5) with comparable high affinity, while in the second one, the paralogs bound only one partner weakly, or no ligand was identified. Our results demonstrate that FP

assays are highly suitable for quantitative interaction profiling of selected protein families. Moreover, we provide evidence that PPI-based phenotypic characterization (hierarchical clustering using UPGMA) can complement or even exceed the information obtained from the sequence-based phylogenetic analysis of the S100ome, an evolutionary young protein family (**Simon et al, FEBS J 2020**).

Although numerous S100 binding partners are known, they are rather restricted to a small subset of the protein family, therefore we applied an alternative approach to map the binding profile of the S100 family. We used a quantitative holdup assay (similar in design as we used first for a different purpose in our lab; see in Gogl et al FEBSJ, 2018) to measure affinity profiles of S100 proteins against a library of chemically synthetized foldamers (so called local surface mimetics). The profiles allowed us to quantitatively map the binding preference of each member towards the foldamer library. Since the library was designed to systematically contain most binary natural amino acid side chain combinations, the data also provide insight into the quantified promiscuity (for the first time in the literature) of each S100 protein towards all potential naturally-occurring S100 partners in the human proteome. Such information will be precious for future drug design of modulators of S100 pathological activities. **(Simon et al, Sci Rep, 2022**).

Protein surface mimetics can be used as chemical tools that target protein–protein interactions (PPIs) that are undruggable with small molecules. Tamás Martinek is an expert in this field, and our new results, described in the previous paragraph was a collaboration with his group. In a more general project, we used the same local surface mimetic library and a bottom-up pulldown methodology to probe several proteins whether these foldamers can distinguish targets by functions. Target proteins that display different levels of interface similarity and promiscuity in their own PPI network were selected (calmodulin, two S100 proteins, a lectin and DNA helicase). After mapping LSM binding fingerprints, we found that the proposed probes displayed biomimetic features on the protein test set, suggesting the suitability of short foldameric sequences to probe protein surfaces and serve as surface mimetic building blocks. (**Tököli et al, Chem Sci, 2020**)

5. PDZ domain interactions

After characterizing the S100B-RSK1 complex, our attention from the S100B binding linear motif shifted to the very C-terminal end of RSK1, comprising an uncharacterized PDZ binding motif. Note here that our attempts to determine the high-resolution structure of the full RSK1 with bound S100B by cryo-EM (Gergő Gogl's experiments in T. Bakes's lab at UCSD supported from an EMBO Fellowship) failed. However, we were able to describe a rather novel phosphoswitch-based mechanism dynamically regulating RSK complexes. Specifically, we showed that autophosphorylation of the disordered C-terminal tail of RSK1 promotes the formation of an intramolecular charge clamp, which indirectly hinders ERK binding. This feedback mechanism, which was structurally characterized and validated in living cells, forms the structural basis for the rapid dissociation of ERK2-RSK1 and RSK1-MAGI-1 PDZ domain complexes under EGF stimulation. We propose that interaction of phosphorylated residues with positively-charged residues in disordered regions (called charge clamp) is likely to be a common mechanism of phospho-regulation. (**Gógl et al FEBS J, 2017**).

Together with our French collaborators (R. Vincentelli and G. Travé), using their HTP PDZ interaction assay we identified a bunch of previously suggested as well as novel RSK1 PDZ domain interactions. Some of these were selected for validation in cell-based assays using a novel a binary fragment complementation assay (NanoBIT) amenable for detecting transient PPIs like the ones involving PDZ domains (Bilics et al, abstract). One of the most interesting partner is ARHGEF12/LARGE, a Rho GTPase playing a fundamental role in numerous cellular processes. We have we found that

phosphorylation dramatically altered the interaction of RSK with the PDZome (266 human PDZ domains), by strengthening or weakening numerous interactions to various degree. The RSK-PDZome interactome analyzed in this study revealed how linear motif-based phospho-switches convey stimulus-dependent changes in the context of related network components. (**Gógl et al. J Mol Biol, 2019**).

Our ANXA2 crystallization chaperon approach turned to be very suitable to solve the structure of PDZ domains and their complexes (more than a dozen of new atomic resolution structures in the PDB). Together with Gilles Travé and Gergő Gógl, we measured the binding affinities of native, phosphorylated, and phosphomimetic variants of both PDZ binding motifs toward the 266 human PDZ domains. We co-crystallized all the motif variants with a selected PDZ domain to characterize the structural consequence of the different modifications. Finally, we elucidated the structural basis of PBM capture by 14-3-3 proteins (**Gogl et al Str, 2020).** Moreover, we introduced a simple and cost-effective strategy for the crystallization of PDZ domains and their complexes. A human ANXA2 fusion tag was used to determine the structures of nine PDZ domains, including five domains that had not yet been solved before **(Cousido-Siah et al, Acta Cryst D, 2022)**.

We have also published a methodology article in the prestigious Springer Methods in Molecular Biology series titled: "PDZ Mediated Interactions: Methods and Protocols". In our chapter we presented we present in detail highly effective biophysical (fluorescence polarization, isothermal calorimetry) and cellular (protein-fragment complementation) methods to study the effect of phosphorylation on RSK1-PDZ interactions that can be also applied to investigate phosphor-regulation of other PPIs in signaling pathways (**Simon et al, Methods Mol Biol, 2022**)

6. Novel NMR methods to study IDPs and cis/trans isomers in proteins

During the studying of S100A4 interactions with NM2A and p53, we noticed that both ligands contain X-Pro dipeptidyl residues, which could be, in theory both in trans (energetically preferred) and cis (minor component) conformation. The conformational heterogeneity of IDPs is often associated with cis/trans isomerization of these proline residues. Together with Andrea Bodor, who is an NMR-expert, we introduced sensitive Hα-detected NMR methods that are able to observe the low concentration minor species, and help in determining the isomer form. Using the proline rich p53TAD peptide, we specify several minor forms, and both cis and trans conformers were observed. These minor forms do not show distinct secondary propensities, however their concentration varies between 4-15%. Using literature information, we established a database that reveals which type of amino acids in the proline surrounding have the higher influence for the formation of the cis conformation. Our results indicate that aromatic residues at (i+2) and/or (i-1) positions induce a higher frequency of the cis form. By applying casein kinase 2 enzyme, we prove a stepwise, slow phosphorylation of T55, S46, S6 sites in p53 TAD1-60 and found that phosphorylation effects the conformer distribution of neighboring prolines. These changes could have functional consequences in the protein-protein interactions of p53 (**Sebák et al, Ang Chem, 2022**)

The methodological background of the above NMR experiments were established in a separate paper (collaboration of Andrea Bodor with Burkard Luy in Karsluhe). In that work the importance (and technical details) of pure shift HalphaCalpha correlations were described using, among other IDPs, the TAD peptide of p53 (**Bodor et al, Anal Chem 2020**).

7. Other published results

We were involved in structural studies to characterize, in collaboration with Laszlo Buday's group, the effect of phosphorylation on linear motive-dependent PPIs of SH3 domains. We have determined

the first crystal structures of tyrosine-phosphorylated human SH3 domains derived from the Abelson-family kinases ABL1 and ABL2 at 1.6 and 1.4 Å resolutions, respectively. The structures revealed that phosphorylation of tyrosine residues induces only minor structural perturbations. Instead, the phosphate groups sterically blocked the ligand-binding grooves, thereby strongly inhibiting the interaction with proline-rich peptide ligands. Our conclusion is that tyrosine phosphorylation might be a mechanism involved in the regulation of the human SH3 interactome. (**Mérő et al JBC, 2019, Mérő et al Int J Mol Sci, 2021**).

Previously, we have extensively studied the interactome of the LC8 dynein light chain, a highly conservative small hub protein, which interacts with a large number of proteins via linear motifs (see e.g. our review article, Rapali et al, FEBS J, 2011). Our more recent contribution was, together with Zsuzsa Dosztányi's bioinformatics group, to predict additional binding partners of LC8 binding (**Erdős et al, PLoS Comut Biol, 2018**). In another collaboration with Judit Ovádi's group, we identified two novel interacting partners of human LC8, HDAC6 and TPPP/p25, which are regulatory proteins of microtubules (MT) dynamics displaying different structural and functional characteristics. Interestingly there is no consensus LC8 binding sequence in TPPP/p25. Instead, we hypothesize that the complex formation is a consequence of their multivalent interactions with the unstructured TPPP/p25 similarly that has been reported by the Ovádi group in the case of its interaction with α-synuclein. The multiple interactions of the MT-related proteins, TPPP/p25, HDAC6 and LC8, may display a role in the fine-tuning of the dynein derived trafficking process partly by the modulation of the acetylation level of the MT network (**Oláh et al BBA Mol Cell Res, 2019**).