Appendix 3B

Summary of Professional Accomplishments

Beata Greb-Markiewicz, BEng, PhD
ORCID: 0000-0002-2831-439X

Department of Biochemistry, Molecular Biology and Biotechnology
Faculty of Chemistry
Wroclaw University of Science and Technology

Wroclaw 2023

TABLE OF CONTENTS

1. Name

Beata Greb-Markiewicz

2. Diplomas, degrees conferred in specific areas of science

Diploma of postgraduate study: 'Management of research project and commercialization of research results', Faculty of Computer Science and Management, Wroclaw University of Science and Technology, 2011.

PhD degree, Institute of Organic Chemistry, Biochemistry and Biotechnology, Faculty of Chemistry, Wroclaw University of Technology, June 1999.
PhD dissertation title: Searching for factors engaged in the infection and intoxication process of insects by entomopathogenic fungi belonging to *Zoophthora* and *Paecilomyces* genera.
Thesis supervisor: Prof. Andrzej Zabża.

MSc degree in Biotechnology, Faculty of Fundamental Problems of Technology, Wroclaw University of Technology, June 1993.
MSc thesis title: Culture conditions development and characterization of neutral lipids and sugars from *Neopandora s*p. belonging to the *Zoophthora (Entomophthoraceae)* genus.
Thesis supervisor: Prof. Andrzej Zabża.

3. Information on employment in research institutes

2019-present Department of Biochemistry, Molecular Biology and Biotechnology, Faculty of Chemistry, Wroclaw University of Science and Technology, assistant professor.
2016 (1.03-31.07) Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, guest scientist.
2004-2019 Department of Biochemistry, Faculty of Chemistry, Wroclaw University of Science and Technology, assistant professor.
2002-2004 Department of General Zoology and Endocrinology, Ulm University, Germany, assistant professor (*Wissenschaftliche Angestellte* position).
1999-2004 Laboratory of Ecological Chemistry, Faculty of Chemistry, Wroclaw University of Technology, assistant professor.
1993-1999 Laboratory of Ecological Chemistry, Institute of Organic Chemistry, Biochemistry and Biotechnology, Faculty of Chemistry, Wroclaw University of Technology, research assistant.

1995/1996 maternity leave

4. Description of the achievements, set out in art. 219 para 1 point 2 of the Act

In the section **4.1**, the first scientific achievement under the title „Characteristics of molecular determinants influencing localization of bHLH transcription factors in the cell" is presented. The scientific achievement has been documented in a series of papers including 4 original papers and 1 invited review paper published in the years 2011-2019 in peer-reviewed scientific journals with IF. In all papers I was the originator of the idea as well as the first and the corresponding author. All mentioned papers were published after obtaining my PhD. In the further part of the Professional Accomplishments papers presented as the first scientific achievement are referenced as **P1-P5**.

In the section **4.2**, the second scientific achievement under the title "Molecular characterization of intrinsically disordered regions and their importance for the functioning of proteins from the family of transcription factors bHLH-PAS" is presented. The scientific achievement has been documented in a series of works including 3 original papers and 1 invited review paper published in the years 2016-2021 in peer-reviewed scientific journals with IF. In all papers I was the originator of the idea as well as the last and the corresponding author. All mentioned papers were published after obtaining my PhD. In the further part of the Professional Accomplishments papers presented as the second scientific achievement are referenced as **P6-P9**.

4.1 The first scientific achievement submitted for evaluation

4.1.1 Title of the scientific achievement

Characterization of molecular determinants influencing localization of bHLH transcription factors in the cell.

4.1.2 Publications documenting the presented scientific achievement

P1. **Greb-Markiewicz B**., Orłowski M., Dobrucki J. & Ożyhar A. (2011) Sequences that direct subcellular traffic of the Drosophila methoprene-tolerant protein (MET) are located predominantly in the PAS domains. *Mol Cell Endocrinol 345*: 16–26.

P2. Greb-Markiewicz B., Sadowska D., Surgut N., Godlewski J., Zarębski M. & Ożyhar A. (2015) Mapping of the sequences directing localization of the drosophila germ cell-expressed protein (GCE). *PLoS One 10*: e0133307.

P3. **Greb-Markiewicz B.**, Zarębski M. & Ożyhar A. (2018) Multiple sequences orchestrate subcellular trafficking of neuronal PAS domain–containing protein 4 (NPAS4). *J Biol Chem 293*: 11255–11270.

P4. Greb-Markiewicz B., Kazana W., Zarębski M. & Ożyhar A. (2019) The subcellular localization of bHLH transcription factor TCF4 is mediated by multiple nuclear localization and nuclear export signals. *Sci Rep 9*: 15629.

P5. Greb-Markiewicz B. & Kolonko M. (2019) Subcellular Localization Signals of bHLH-PAS Proteins: Their Significance, Current State of Knowledge and Future Perspectives. *Int J Mol Sci 20*: 4746.

A description of my individual contribution to papers presented as the scientific achievement is included in **Appendix 4B**.

Statements of the co-authors specifying the individual contribution to papers presented as scientific achievement are included in **Appendix 5**.

Copies of papers presented as the scientific achievement are included in **Appendixes 6.1-6.5**

4.1.3 Scientometric parameters of the presented scientific achievement

Scientometric parameters of the publications documenting the first scientific achievement are presented in the Table 1.

Table 1. Parameters on 15.08.2023

Publication	Impact factor IF*	Points MEiN**	All citations Web of Science	Citations excluding selfcitations Web of Science	All citations Scopus	Citations excluding selfcitations Scopus
P1	4,192	30_{2011}	18	14	20	15
P2	3,057	40_{2015}	11	6	10	5
P3	4,106	35_{2018}	5	2	5	2
P4	3,998	140_{2019}	5	5	5	5
P5	4,556	140_{2019}	7	6	6	5
Total	**19,909**	**385**	**46**	**33**	**46**	**32**

*IF according to Journal Citation Report from the year of publication.
** Points were calculated according to the list of scientific journals of the Ministry of Education and Science (MEiN, previously MNiSW) from the year of publication

4.1.4 Discussion of the scientific achievement

a) Introduction

The bHLH family

The basic Helix-Loop-Helix (bHLH) proteins are the superfamily of transcription factors (TFs) that are present in all eukaryotes and regulate the expression of genes involved in organism development and cell differentiation. In the vertebrate organisms, bHLH TFs are key determinants of neural cell fate specification and differentiation (Dennis et al., 2019). In plants, in addition to plant growth and metabolism, bHLH TFs are also engaged in the regulation of plant response to stress (X. Sun et al., 2018). The highly conserved bHLH domain recognizes the core E-box DNA sequence (5'-CANNTG-3'). In addition to the bHLH domain, these proteins may also contain other domains like Leucine Zipper (LZ), Orange (O), and Per-Arnt-Sim (PAS), which provide specificity of transcriptional regulation and additional functionality (Murre, 2019). Currently, a few independent classification systems exist for the bHLH family. Evolutionary classification based on phylogenetic studies (Ledent et al., 2002), classification based on the complete amino acid sequence analyses, and presented below, the natural method of classification proposed by Murre, which divides the bHLH proteins into seven classes based on the presence of additional domains, expression patterns and performed transcriptional function (Murre, 2019).

Class I (known as E proteins) and **class II** of the bHLH TFs do not possess domains additional to the bHLH. Contrary to class I, which is expressed in many tissues, the expression of proteins from class II is tissue specific. Members of class II are dimerization partners for the class I TFs.

Class III comprises bHLH-LZ TFs, for example, sterol regulatory element-binding proteins (SREBPs) critical for the regulation of cholesterol and fatty acid homeostasis in animals. Important members of this class are also proteins belonging to the MYC subfamily, which regulate oncogenic transformation, apoptosis, and cellular differentiation.

Class IV contains MAD and MAX (only bHLH domain) which can dimerize with MYC and regulate its activity. Also, MAD/MAX can create homo- and heterodimers with each other and influence transcription in differentiated ways, though these TFs do not possess a transcription activation domain (TAD).

Class V contains transcriptional inhibitors ID1-3 which cannot bind DNA (HLH domain, lack of basic region) and act by the sequestration of other bHLH TFs, and ID4 functioning as an inhibitor of ID1-3.

Class VI comprises bHLH-O TFs which perform regulatory functions in various developmental processes including cell differentiation and maintaining pluripotency. For this

reason, they are often linked to cancer development. Examples of bHLH-O TFs are HES (hairy/enhancer of split) and HES-related proteins which regulate the highly conserved Notch signalling pathway and were suggested as participating in cancer development (H. Sun et al., 2007; Weber et al., 2014).

Class VII comprises bHLH-PAS TFs. These proteins in addition to the bHLH domain contain Per-Arnt-Sim (PAS) domains, which serve as sensors of signals and interaction hubs. The bHLH-PAS TFs are signal transduction proteins responsible for sensing environmental signals like the presence of xenobiotics (Aryl hydrocarbon receptor, AHR), hypoxia (hypoxia-inducible factor, HIF), or regulation of circadian rhythms (CLOCK, CYCLE). The members of bHLH-PAS subclass II TFs - Aryl hydrocarbon receptor translocator 1-4 (ARNT1-4) proteins are general dimerization partners of the subclass I members (Dougherty & Pollenz, 2010).

PAS motif was identified for the first time by sequence homology of approximately 50 amino acids in the *Drosophila* proteins period (Per), single-minded (Sim), and the vertebrate aryl hydrocarbon receptor nuclear translocator (Arnt). Additional conserved residues located C-terminally to that region were identified later as PAS-domain associated (PAC) motifs. PAS domains are conserved across different kingdoms - the first structure of PAS was presented for the photoactive yellow protein (PYP) from bacteria *Halorhodospira halophila.* The structure showed that the PAS and PAC motifs adopt a single globular fold comprising approximately 100 residues and they are currently referred to as the PAS domain or PAS_fold. Importantly, the PAS domain was shown to be responsible for the specificity of the bHLH-PAS TFs functioning (Kolonko & Greb-Markiewicz, 2019; Möglich et al., 2009; Ponting & Aravind, 1997).

Subcellular localization of proteins

The ability to localize and translocate proteins to specific cellular compartments is fundamental to the organization and functioning of all living cells. For transcription factors responsible for the regulation of gene expression in response to signal reception, translocation from the cytoplasm to the nucleus is an important event, enabling the transcription factors to recruit co-activators and bind to DNA. Also, translocation from the nucleus to the cytoplasm when transmission of the signal is no longer needed is often used as a way of TFs deactivation. Importantly, some transcription factors were shown to play additional non-genomic functions in the cytoplasm. A member of the bHLH-PAS TFs family, AHR as a part of the cytoplasmic complex was thought to be an inactive form from which AHR is released after ligand binding,

which results in AHR translocation to the nucleus and dimerization with ARNT, which in turn results in xenobiotic response elements (XRE) binding and transcription initiation. Currently, it is known that in addition to its function as TF, AHR also regulates protein ubiquitination and phosphorylation. For this reason, the control of AhR localization is a regulatory mechanism modulating both transcriptional activity and ubiquitin ligase function. Currently it is known that both genomic and non-genomic signaling pathways contribute to the AHR-dependent modulation of macrophage activation (Großkopf et al., 2021). The other representative of the bHLH-PAS family, BMAL1 (ARNT3), was shown to pair with CLOCK in the cytosol to translocate the active, heterodimeric complex to the nucleus. Also in this protein case, an additional non-genomic function as a translation factor enhancing protein synthesis during the metabolically active night-time period was revealed (Michael et al., 2015).

Nuclear pore complexes (NPCs) form channels across the nuclear envelope by fusing the inner and outer nuclear membranes. NPCs allow the passive diffusion of ions and small proteins (<40 kDa) through the nuclear envelope. In contrast, the passage of larger molecules is restricted to those containing appropriate targeting signals recognized by the nuclear transport receptors known as importins and exportins belonging to the karyopherin's family. Importin-α and importin-β bind to the nuclear localization signals (NLSs) and facilitate their passage through the central channel of the NPC. At the nuclear site of the NPC, the binding of RAN·GTP to the import complex releases the cargo into the nucleoplasm. Similarly, exportin 1 (XPO1; also known as Chromosome region maintenance 1 protein (CRM1) functions in the nuclear export pathway, during which cargos containing nuclear export signals (NESs) bind to XPO1 in a RAN·GTP-dependent manner, travel to the cytoplasm and are released on GTP hydrolysis (Beck & Hurt, 2017).

NLSs/NESs represent short linear motifs (SliMs, usually 3-10 aa long sequences) located within intrinsically disordered regions (IDRs) of protein and responsible for low-affinity docking onto globular domains during protein-protein interactions (PPIs) creating signaling network of the cell (Wigington et al., 2020). For a long time, scientists believed that spontaneous folding into a well-defined and stable tertiary structure is required for protein action. However, it is currently known that most of eukaryotic proteins contain intrinsically disordered regions (IDRs) that do not have a stable tertiary structure in physiological conditions, but at the same time still perform important biological functions. The lack of a defined structure and flexibility of IDRs is important for playing the role of hubs for PPIs (Wright & Dyson, 2014). Dynamic regulation of SLiMS activity can be easily done by post-translational

modifications (PTMs) like phosphorylation, that influence the region conformation. Masking/unmasking NLS/NES motifs, which impact protein cellular localization also depends on PPI and PTMs, the factors sensitive to various stimuli (J. Lu et al., 2021) (Wubben et al., 2020).

The best-studied nuclear localization signals (NLSs) recognized by importins are the classical NLSs consisting of monopartite (discovered in the simian virus 40 (SV40) large T-antigen, PKKKRKV) and bipartite (discovered in nucleoplasmin, KRPAATKKAGQAKKKK) motifs rich in basic amino acid residues. An example of nonclassical NLS is PY-NLS class comprising P and Y residues.

In contrast to NLSs, NES sequences are very diverse. Usually are 8-16 residues long and contain 4–5 hydrophobic residues (often L/V/I/F/M) that bind hydrophobic pockets in a hydrophobic groove of XPO1 (Fung et al., 2017). Though many NLS/NES predictors exist, non-classical signals have remained difficult to predict. In addition, not all predicted *in silico* signals are active in the cell. For this reason, experimental studies are indispensable for the final verification of the active NLS/NES motifs in protein.

Subcellular localization and shuttling of proteins between nuclear and cytoplasmic compartments are consequences of a dynamic balance between operational strengths of nuclear localization signals (NLSs) and nuclear export signals (NESs) (J. Lu et al., 2021). Dysregulation of protein localization can affect signaling pathways responsible for tumor growth, inflammatory response, cell cycle, and apoptosis. Restoring normal protein localization resulted in the proper functioning of tumor suppressors and oncoproteins (L. Wang et al., 2015). For this reason, nucleo-cytoplasmic transport in cancer cells was proposed as a therapeutic target. As an example, synthetic peptide with EPS8 NLS sequence was shown to present anticancer activity against acute myeloid leukemia (Y. Chen et al., 2018). Also, many viruses are dependent on entering the host nucleus to replicate their genome, or to suppress the host antiviral response. For this reason, inhibitors of nuclear transport leading to improper localization of host/viral proteins important for the viral life cycle, also emerged as promising antiviral therapeutics (Wubben et al., 2020). An example is Ivermectin shown to be an effective inhibitor of mediated specifically by α/β importins protein import to the nucleus. This ability was used for effective inhibition of HIV-1 and dengue virus replication. Recently, Ivermectin was proposed as an agent useful against a wider spectrum of viruses including Zika and SARS-CoV2 (Patil et al., 2022). In addition to importins, exportins have also been considered as therapeutic targets. The first discovered small-molecule competitive inhibitor of nuclear export is Leptomycin B (LMB) that binds to XPO1 residues responsible for interaction with NES

sequences. Unfortunately, though being effective, LMB is highly toxic. For this reason, LMB analogues and other inhibitors presenting the high potency of LMB, and much lower toxicity in vivo were later searched (Gravina et al., 2014). Currently, LMB is often used as a tool for the verification of NES activity in a studied protein (Cheng et al., 2019).

b) Main objectives of the presented scientific achievement

My interest in the bHLH family has developed during my stay in Ulm University, after obtaining PhD, when I decided to rearrange my scientific career into biochemistry and molecular biology. That time, I switched my interests from low molecular compounds and chemical characterization of entomopathogenic fungi/medicinal plants and had to learn new methodologies. The first bHLH protein I had known was *Drosophila melanogaster* Methoprene tolerant protein belonging to the bHLH-PAS subfamily. Later, its paralog Germ cell-expressed was discovered. I have chosen MET and GCE, recognized as JH receptors, as my first subjects. I aimed to verify my hypothesis that different NLS/NES patterns can be linked to MET and GCE functional differentiation. Performed extensive literature studies led me to the interest in mammalian family members linked to human diseases. I was especially interested in proteins engaged in neuronal system development. As mentioned, the precise regulation of protein shuttling in the cell is important for the proper functioning of signal transduction. Any dysregulation of the signaling pathway can result in serious disease. However, most of the papers documented single NLS/NES without systematic studies and results were not updated. For this reason, I decided to identify the NLSs/NESs of selected representatives of the bHLH TFs family. I have chosen NPAS4 as a neuroprotective protein with no known mechanism of functioning. My goal was to see if the complex pattern of NLSs/NESs shown for MET and GCE is more general for the bHLH-PAS subfamily. If so, it could shed light on the propensity of NPAS4 to fulfil various functions documented in recent papers. Next, I selected TCF4 representing the E-class of bHLH TFs to enable the comparison between bHLH-PAS and the other class member of bHLH family NLSs/NESs pattern. Importantly, TCF4 was linked to the mental disorder Pitt-Hopkins Syndrome (PTHS) and I wanted to test my hypothesis concerning NLS/NES mutations in bHLH domain area with PTHS development. As the summary of my studies, I discussed in a review the current state of the knowledge with existing gaps in the context of the importance and the need of more detailed future studies for other members of bHLH-PAS TFs using as example results of my previous studies. I asked a question about the

general mechanism of regulation of bHLH-PAS proteins localization and how precise and systematic were studies reported in the area.

c) Description of papers presented as the scientific achievement

P1. Greb-Markiewicz B., Orłowski M., Dobrucki J. & Ożyhar A. (2011) Sequences that direct subcellular traffic of the Drosophila methoprene-tolerant protein (MET) are located predominantly in the PAS domains. *Mol Cell Endocrinol 345*: 16–26.

P2. Greb-Markiewicz B., Sadowska D., Surgut N., Godlewski J., Zarębski M. & Ożyhar A. (2015) Mapping of the sequences directing localization of the drosophila germ cell-expressed protein (GCE). *PLoS One 10*: e0133307.

Drosophila melanogaster has been established as a valuable model organism to study a diverse range of multicellular organism pathways including how genes direct the development of an embryo from a single cell to a mature organism (Jennings, 2011). Because of the high complexity of vertebrate hormone signaling pathways, *D. melanogaster* which possessed only two known physiologically active hormones, the steroid hormone 20-hydroxyecdysone (20E) and the sesquiterpenoid juvenile hormone (JH) was a good candidate for performing preliminary studies of hormone signaling pathways. At that time, contrary to the well-understood mechanism of functioning of the ecdysone receptor, the JH receptor was unknown. 20E and JH were shown to regulate insect development, reproduction, and most of the important biological processes (Dubrovsky, 2005). In contrast to the extensively studied molecular mechanisms of 20E action (Beckstead et al., 2007), mechanisms of JH signaling were poorly understood, and many basic questions remained unanswered during the time of P1 and P2 preparation. Heterodimeric components of the 20E receptor: ecdysteroid receptor (EcR) and ultraspiracle (USP) were known while the receptor of JH was still searched (Henrich et al., 2003; Riddiford, 1996). In addition, it has been shown that the regulation of the response to JH can be modified by 20E, showing that the interplay between JH and 20E was more complex than it was assumed (Beckstead et al., 2007).

At the time of P1 preparation, one of the main candidates for the function of the JH receptor was Methoprene-tolerant protein (MET) (Miura et al., 2005). MET has been discovered in *D. melanogaster* by screening for mutants resistant to the JH analogue, Methoprene (Wilson & Fabian, 1986) and classified as a member of the family of bHLH-PAS TFs by a comparison of the MET sequence to the sequences of proteins deposited in databases (Ashok et al., 1998; Shemshedini & Wilson, 1990). Interestingly, the viability of the Drosophila Met-null mutant undermined the credibility of the hypothesis that MET functions as a JH receptor (Wilson & Fabian, 1986). However, the subsequent discovery of the MET paralog: the

germ cell-expressed protein (GCE) presenting partial functional redundancy explained this discrepancy (Abdou et al., 2011; Baumann, Barry, et al., 2010). Currently, MET and GCE have the status of the Drosophila JH receptors (Jindra et al., 2015a). Interestingly, in contrast to *Drosophila*, most insect species possess only one paralog, which was classified by phylogenetic analysis as more like GCE. The course of evolution of both genes across the *Drosophila* genus suggests Met as a product of the duplication of the Gce-like ancestor gene during early dipteran evolution (Baumann, Fujiwara, et al., 2010).

Importantly, Dubrovsky et al. showed that the activation of E75A, which is JH target gene, is dependent on the presence of GCE- but not MET-dependent, indicating for the first-time existing differences between MET and GCE and providing proof that GCE is not merely a MET substitute (Dubrovsky et al., 2011). A hypothesis about intricate MET/GCE-specific functions was supported by the study showing that the expression of MET in *D. melanogaster* is higher than the expression of GCE (Abdou et al., 2011). However, GCE binds JH with a higher affinity than MET(Charles et al., 2011). Also, overexpression of MET results in the lethal effect in contrast to the overexpression of GCE not causing *Drosophila* mortality (Baumann, Barry, et al., 2010).

The starting point of my studies presented in P1 and P2 was the observation that scientists studying *Drosophila* MET and GCE neglected previously shown functional differences between these proteins. Often, results of a study performed experimentally only for MET were discussed also in the context of GCE, using just MET/GCE entry (Jindra et al., 2015b). The published data on the role of JH reception, including review reports focused primarily on the function of MET and left the function of GCE largely unknown (Bernardo & Dubrovsky, 2012; Jindra et al., 2013; Riddiford, 2012). For this reason, I aimed to indicate that MET and GCE should be studied independently as separate proteins, and results obtained for one of them should not be extrapolated to another one, otherwise, such simplifications could result in erroneous conclusions. Protein function can be influenced by its subcellular localization which assures controlled access to molecular interaction partners. Thus, knowledge of protein localization regulation is fundamental for characterizing the cellular function of newly discovered proteins. For this reason, I decided to perform a detailed characterization and comparison of MET and GCE in the context of sequence motifs and other factors responsible for the regulation of their localization.

Both MET and GCE were classified into the family of bHLH-PAS TFs whose activity could be regulated by the shuttling between cytoplasm and nucleus in response to signals.

Translocation of TF to the nucleus enables binding of DNA and recruiting co-activators which enable fulfilling protein function. However, the knowledge about the subcellular trafficking of MET at the time of P1 preparation was very limited and not conclusive. The first study on the localization of the MET protein was done in 1990 by analyzing JH binding in larval fat bodies and showed MET localization in the cytosol of the fat body (Shemshedini & Wilson, 1990). Later, Shemshedini and Wilson used the photo-affinity labelling technique, which resulted in the localization of MET both in cytoplasmic and nuclear fractions of *D. melanogaster* cells (Shemshedini & Wilson, 1993). In contrast, immunolocalization studies of MET in a variety of *D. melanogaster* tissues showed this protein as exclusively nuclear (Pursley et al., 2000) which was confirmed in Drosophila Schneider cells using GFP (green fluorescent protein) labelled MET (Miura et al., 2005).

As previously used methods, which were performed in insects or insect cells, did not result in consistent information, I decided to use mammalian cells, which were devoid of juvenile hormone, 20-hydroxyecdysone, and some insect cell-specific endogenous proteins, which could influence previous results. First, I prepared cDNA vectors, which allowed me to express full-length wild-type MET tagged N- /C-terminally to yellow fluorescence proteins (YFP) and cyan fluorescence protein (CFP) in COS-7, Hela, and CHO-1 cells. It is worth noting that fluorescent proteins were previously shown as not influencing protein localization and function (Chalfie et al., 1994; Nardozzi et al., 2010). The distribution of expressed YFP–MET, MET-YFP, CFP-MET and MET-CFP was analyzed using confocal microscopy 20- 24 and 48 h after transfection. I used YFP as a control. In addition, the expression of MET tagged with the YFP/CFP in COS-7 cells was also analyzed by Western blot to prove expression and test the stability of the protein. I observed that the fusion of the YFP/CFP to the C-terminus of MET resulted in the absence of band corresponding to the respective fusion protein, probably due to the process of degradation. In contrast, the fusion proteins consisting of MET and CFP/YFP attached to the N-terminus of MET appeared to be more stable. Since the MET derivative containing YFP appeared as a single band corresponding to the full-length YFP–MET, the N-terminal tagged YFP derivatives were further used in all the experiments as presenting the highest stability. Similarly, the stability of the MET derivatives also in CHO-K1 and HeLa cells was checked. As the results were essentially the same, in further experiments only COS-7 cells were used. Initial microscopy observations and western blot analysis of cells were performed 48 h after transfection, but a high number of apoptotic cells and high degradation of proteins were observed. For this reason, I decided to perform further experiments 24 h after transfection.

Expression of YFP-MET resulted in the presence of protein exclusively or predominantly in the nucleus for more than 90% of the analyzed cells. However, in some cells YFP–MET localized predominantly in the cytoplasm, sometimes in the form of clusters (Fig. 1B). Observed discrepancies supported my theory that MET contains NLS/NLSs active in all investigated cell lines and responsible for its translocation to the nucleus. The appearance of cells where localization was predominantly cytoplasmic and the lack of cells with a homogeneous distribution (Fig. 1) suggested additionally the presence of an NES/NESs in MET.

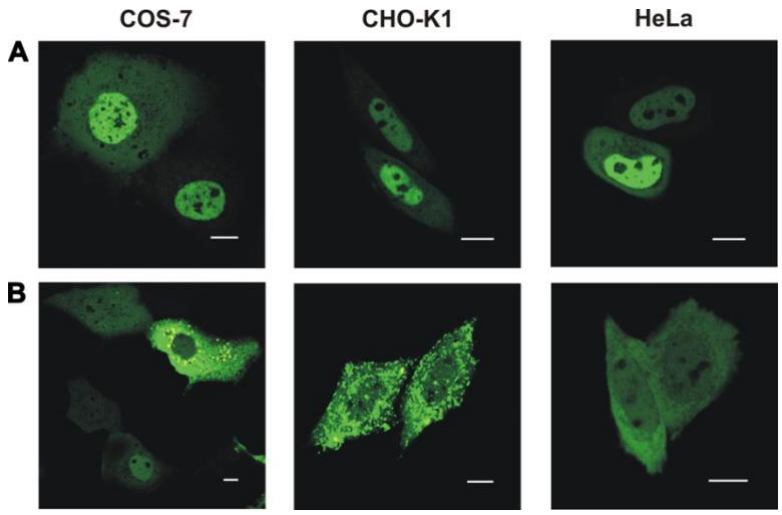


Figure 1. Analysis of the subcellular distribution of full-length MET tagged with YFP in COS-7, CHO-K1 and HeLa cells. (**A**) Typical confocal images of cells expressing fusion proteins 20 h after transfection. (**B**) Confocal images of cells presenting an atypical cytoplasmic localization of expressed fusion proteins. Bar, 10 μm [P1].

First, I performed putative NLS/NES motifs predictions using available programs. Considering that some of the predicted signals could be not active and opposite, some of the existing non-classical signals could be difficult or even not possible for detection by *in silico* analysis, experimental screening/verification of localization signals was indispensable. I decided to divide MET for structural/functional fragments and to express these fragments in COS-7 cells as YFP-fused proteins. Regions of truncations were planned out according to the putative secondary structure motifs of MET in such a way that the fragments encompassed whole domains of the MET protein (Ashok et al., 1998). In addition, some amino acid residues were included based on the PROSITE/SMART databases and PONDR secondary structure predictor to assure that no structural motif was broken. Consequently, I divided the MET sequence (Fig. 2A a) into six fragments: (1) the N-terminal fragment without known function

(MET/1-33, Fig. 2A c), (2) the bHLH domain (MET/34-97, Fig. 2A d) responsible for DNA binding and dimerization, (3) fragment referred to as PAS-A domain (MET/98–190, Fig. 2A e) containing additional amino acid residues from the linker between bHLH and PAS-A known in bHLH-PAS proteins as an additional dimerization surface responsible for partner binding specificity (4) area without defined structure and function located between PAS-A and PAS-B domains (MET/191-402, Fig. 2A f), (5) PAS-B (MET/403-508, Fig. 2A g) shown previously for other bHLH-PAS family representatives as responsible for ligand binding, (6) the C-terminal part of MET (MET/509–716, Fig. 2A h) without defined structure nor a known function in time of P1 preparation. To ensure that no structural motif was broken, which could influence observed localization, I also prepared and expressed mutants comprising two neighboring domains/fragments fused to YFP: YFP-MET/1-97, YFP-MET/34-190, YFP-MET/98-402, YFP-MET/191-508 and YFP-MET/403-716 (Fig. 2A i-m). In contrast to the full-length protein, MET fragments tagged with YFP revealed clear patterns of localization in the COS-7 cells (Fig. 2B c-j). The expression of these fragments in CHO-K1 and HeLa cells showed the same results. Especially interesting was nuclear localization of the PAS-A domain and cytoplasmic localization of the PAS-B domain and C-terminus of MET. Collectively, these results further confirmed that MET could contain both NLS and NES signals, which might regulate the distribution of MET in cells.

To prove the presence of motifs responsible for subcellular shuttling of MET, I prepared a set of YFP-fused point mutants with substitution of amino acid residues predicted as important for NLS/NES activity and expressed them in COS-7 cells. Change of localization was expected as experimental proof of the presence of active NLS/NES. The presence of NLS in the PAS-A domain was verified by YFP-MET34-190/R98A/K102A mutant. Interestingly, cytoplasmic expression of this mutant suggested the presence of NES in this area. For this reason, I prepared an additional, YFP-MET34-190/R98A/K102A/L130A/L133A mutant resulting in ubiquitous expression as expected. Very intriguing was the discovery that NES detected in PAS-A encompassed the LXXLL motif. LXXLL motifs are known to mediate the protein-protein interaction of transcriptional cofactors with nuclear receptors (Plevin et al., 2005). Also, one of the AhR LXXLL motifs was shown to be important for the regulation of localization and transcriptional activation of this protein (Ikuta et al., 2002). Considering that JH and 20EOH pathways cross each other, I hypothesized that the MET LXXLL motif could be responsible for interaction with EcR/USP. However, Bitra and Palli published study suggesting that heterodimerization of MET with ECR and USP is not dependent on the LXXLL motif (Bitra &

Palli, 2009). To date, the molecular mechanism of interaction and pathway crossing is not fully understood.

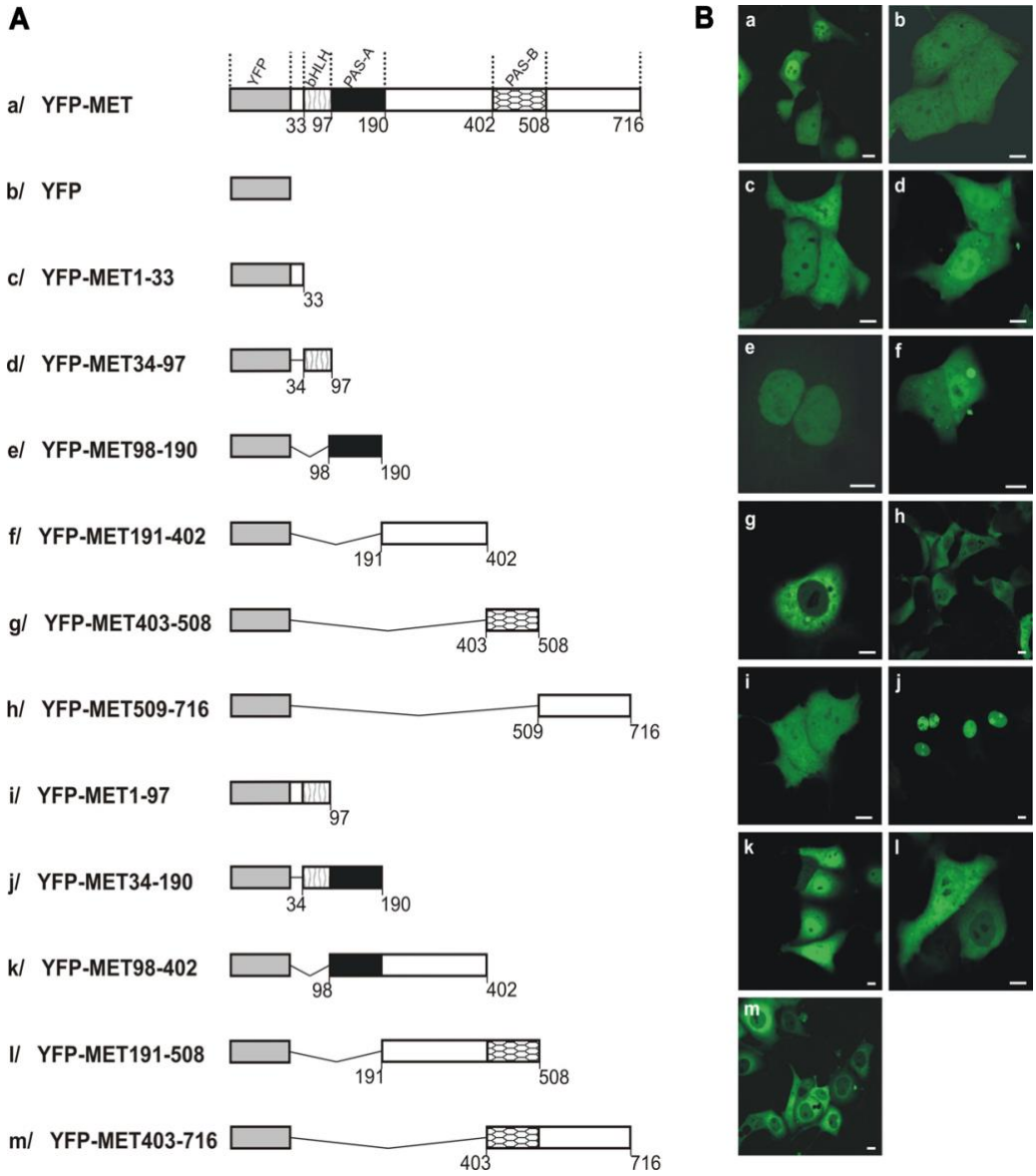


Figure 2. Analysis of the subcellular distribution of MET deletion mutants tagged with YFP in COS-7 cells. (**A**) Schematic representation of full-length MET and its derivatives tagged with YFP. Individual regions of deletion were defined according to Ashok et al. (1998) and PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). The length of each domain in the diagram is relative. (**B**) Confocal images of the subcellular distribution of the MET derivatives are depicted in (A). Bar, 10 μm [P1].

In contrast to the cytoplasmic location of the wild-type PAS-B domain, the expression of its point mutant: YFP–MET403–508/V446A/I450A resulted in a fluorescence observed unexpectedly only in the nucleus. The change of localization confirmed switching off the NES but also suggested the presence of an NLS in the tested area. For this reason, I prepared an additional mutant, YFP–MET403–508/V446A/I450A/G491A/K494A/H496A, whose expression displayed equal distribution in transfected cells and confirmed the hypothesis about

NLS activity. Thus, I demonstrated that the PAS-B domain of MET possesses active NES and NLS.

Finally, I decided to make attempts to detect NES in the C-terminal fragment of MET. While predictors were not able to detect NES in this area, I performed the sequence analysis of the presence of a Leucine-rich region, which resulted in the preparation of YFP–MET509–694 and YFP–MET695–716 deletion mutants. Homogeneous distribution of expressed MET695-716 throughout the cell excluded this area from further research and suggested that putative NES resides rather in the MET509-694 area. The alignment of known NES sequences with the sequence of MET C-terminus using CLUSTAL_X helped me to identify a fragment in the region of interest that presented some similarity to the retinoid X receptor NES. Unfortunately, the expression of YFP–MET508–716/R550A/V552G/ L555A mutant resulted in exclusively cytoplasmic fluorescence showing that NES was still active. Finally, successful detection of this NES was performed and explained in P2.

Considering that MET acts as TF and JH receptor I supposed that JH binding should initiate the signal transmission, and the relative strength of NLSs to NESs and regulation of their balance to ensure the proper translocation of a protein is fundamental for such transmission. For this reason, I decided to test how the addition of JH impacts the localization of full-length MET in comparison to the truncated MET comprising PAS-A and PAS-B domains (98-508 aa). Importantly, my experiments documented that NLS activity in the PAS-A linker (98-102 aa) was dominant and JH independent, while the activity of the second discovered NLS, located in the PAS-B a domain (482-498 aa), was dependent on the presence of JH, which is consistent with the expected function of PAS-B as a ligand-binding domain.

To date of P2 preparation, no localization studies were performed for GCE. The high degree of homology between the GCE and MET proteins (S. Wang et al., 2007) raised the possibility that the trafficking of GCE in the cell is mediated by some of the localization signals identified for MET [P1]. On the other hand, the sequence homology applied only to the bHLH (78%) and PAS domains (PAS-A 68%, PAS-B 86%). Using the methodology analogical to previous MET studies, I confirmed that GCE possesses active signals detected previously in MET and located in PAS-A (NES) and PAS-B domains (NLS and NES).

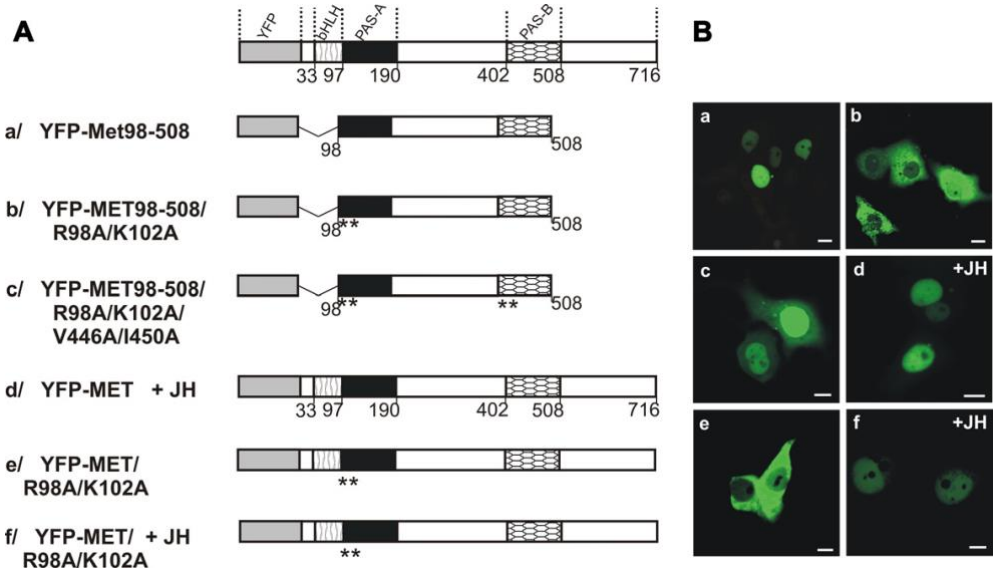


Figure 3. Subcellular localization of the MET NLS mutants. (**A**) The influence of JH on the subcellular distribution of the wild type and NLS point mutants of the bHLH-PAS-B area and full length MET in COS-7 cells. The length of each domain in protein schematic illustration is relative. Regions of MET are depicted using different patterns. Asterisks indicate mutated residues. (**B**) Representative images (single confocal plane) present the typical subcellular distribution of the MET derivatives. Bar, 10 μm [P1].

In contrast to the nuclear MET, full-length GCE was observed both in the nucleus and the cytoplasm of analyzed cells. This observation could be explained by the discovery that GCE did not possess dominant NLS independent of JH which in MET was in the linker region between bHLH and PAS-A domains (Fig. 4A, 5AB). It is worth noting that no homology between MET and GCE was observed for C-termini; however, GCE/MET sequence alignment followed by experimental verification resulted in discovering a short, conserved motif (Fig. 4B), documented as previously searched for MET active NES (Fig. 5AB). Interestingly, in addition to the NES, also NLS was detected and documented in the C-terminal part of GCE (Fig. 5B).

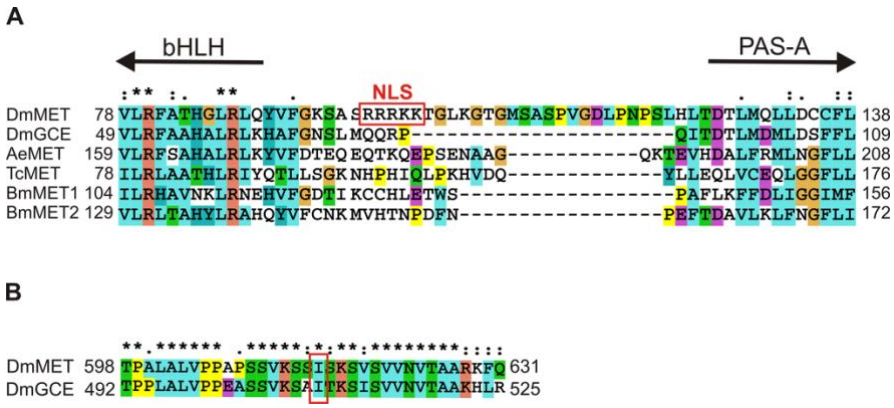
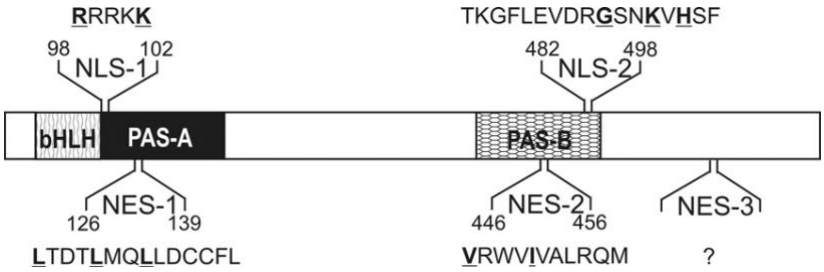


Figure 4. Alignment of the MET/GCE sequences. (**A**) The area that links bHLH and PAS-A domains from MET/GCE *Drosophila melanogaster*, MET *Aedes aegypti*, MET *Tribolium castaneum* and MET1/MET2 *Bombyx mori*. The red box indicates NLS unique for *Drosophila melanogaster*. (**B**) Drosophila MET598-631/GCE492-525 encompassing predicted NES. The red box indicates GCE residue I509 predicted as NES [P2].

Since GCE had been classified as TF, which should be transported to the nucleus after activation, I had asked the question about mediators of GCE shuttling. The ubiquitous distribution of full-length GCE both in the presence and absence of JH (Fig. 6 c) was changed in the case of GCE mutant (YFP-GCE/K582A/K585A) with switched off C-terminal NLS activity (NLS-2) and resulted in exclusively cytoplasmic expression independent on the presence/absence of JH (Fig. 6 d). Interestingly, expression of YFP-GCE69-390 comprising PAS-A and PAS-B domains resulted in cytoplasmic (without JH) or prevalent nuclear localization (with JH presence) (Fig. 6 a).

A



B

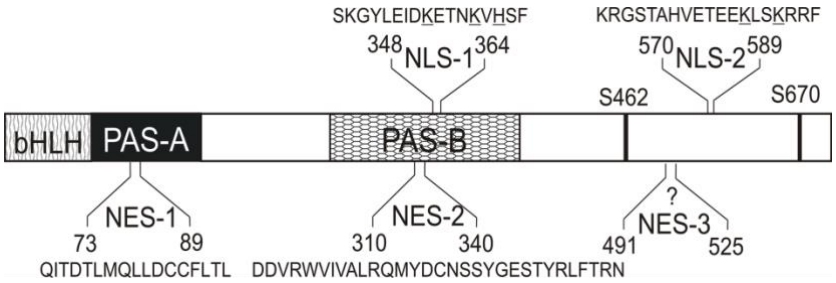


Figure 5. Schematic representation of NLSs and NESs residing within the MET (**A**) and GCE (**B**) proteins. Bold and underlined letters in sequences indicate residues that were mutated in studies [P1, P2].

This difference was no longer observed for YFP-GCE69-390/K356A/K360A/H362 mutant with switched off NLS in the PAS-B domain which was in the cytoplasm of analyzed cells independent of the JH (Fig. 6 b). These results show that NLS located in PAS-B (NLS-1) is JH dependent similarly to MET. This is consistent with the knowledge that PAS-B in the bHLH-PAS TFs family is responsible for ligand binding. Importantly, the activity of NLS-2 from the C-terminus seems to be necessary for the JH-dependent transport of GCE to the nucleus and contributes to the functioning of the JH-dependent NLS-1 in PAS-B. However, the mechanism of such regulation still awaits further studies and explanation.

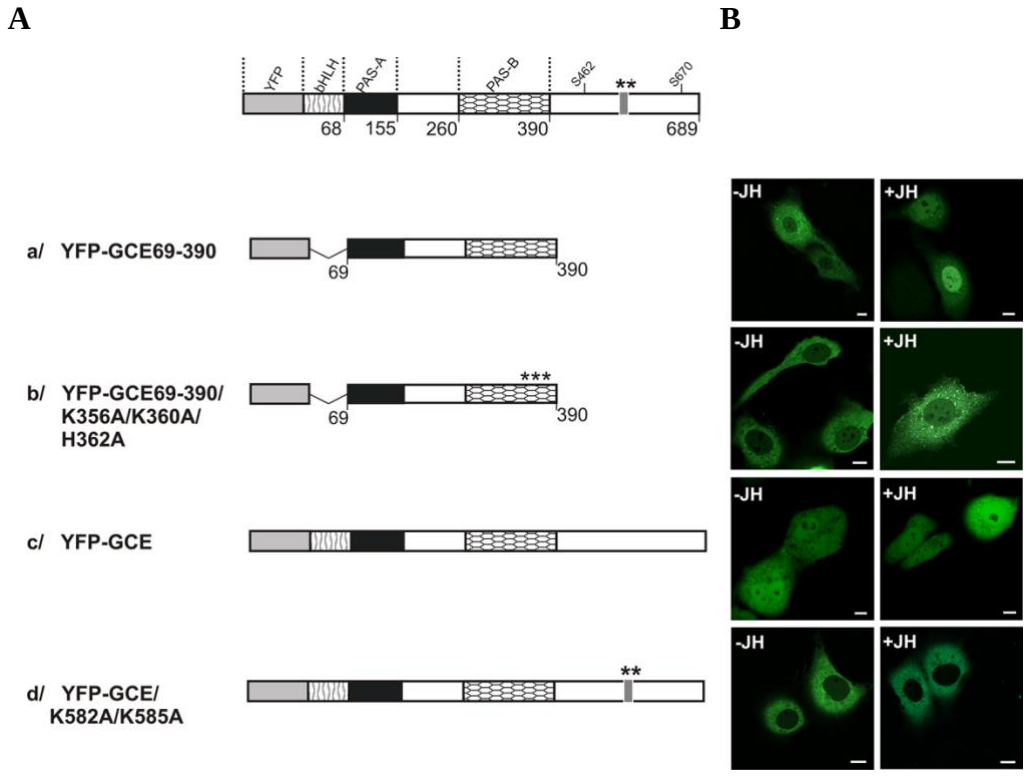


Figure 6. Subcellular localization of the GCE NLS mutants. (**A**) The influence of JH on the subcellular distribution of the wild type and NLS point mutants of the bHLH-PAS-B area and full-length GCE in COS-7 cells. The length of each domain in the protein schematic illustration is relative. Regions of GCE are depicted using different patterns. Asterisks indicate mutated residues. (**B**) Representative images (single confocal plane) present the typical subcellular distribution of the GCE derivatives. Bar, 10 μm [P2].

When searching for additional factors regulating the localization of GCE, I used bioinformatic tools for *in silico* GCE sequence analysis like ELM (Dinkel et al., 2014),Scansite 2.0 (Obenauer, 2003), Disphos 1.3 (Iakoucheva et al., 2004), and NetPhos 2.0 (Blom et al., 1999). Highly interesting was the prediction of S462 and S670 located in the C-terminus of GCE as targets of phosphorylation and a part of recognition motifs for the 14-3-3 proteins. It is worth noting that these regions were not conserved in MET (Fig. 7A). The ubiquitous family of 14-3-3 proteins is known to be involved in the regulation of signal transduction by modulation of subcellular localization, structure, stability, and molecular interactions of partner proteins (Obsilova et al., 2014). Results of performed predictions were supported by previous reports that 14-3-3 proteins could be involved in the regulation of signal transduction by changing the partner protein subcellular localization (Muslin & Xing, 2000). Shortly before P2 preparation, a study showing the dependence of the signal transduction specificity on the 14-3-3 interaction with intrinsically disordered regions (IDRs) of partner proteins was published (Bustos, 2012). For this reason, I also performed prediction of disorder in GCE, which identified the C-terminal part of GCE comprising S462 and S670 residues as IDR, which substantiated my hypothesis that S462 and S670 could serve as binding sites of 14-3-3. For experimental

verification, I prepared YFP fused GCE mutants with S462A and/or S670A substitutions. All expressed mutant proteins were observed in the cytoplasm without JH (Fig. 7C a-c). However, the addition of JH redistributed the mutants to the nucleus (Fig. 7C a-c), which indicated that there was dependence on JH. These results enabled me to suggest a possible role for 14-3-3 binding to GCE as modifying GCE localization.

It is worth noting that research was performed in mammalian cells lacking proteins responsible for the 20E hormonal pathway, while the ubiquitous 14-3-3 family is highly conserved between vertebrates and invertebrates. For this reason, I proposed in P2 that possible inhibition of 14-3-3 binding to GCE could be caused by the activity of other factors occurring in Drosophila, like de-phosphorylation by phosphatases or the masking of 14-3-3 binding sites by other partners of GCE protein. The lack of 14-3-3 interaction with the mutant GCE results in the unmasking of the C-terminal NLS and translocation to the nucleus. In turn, cytoplasmic localization of GCE mutants in the absence of JH is the result of the dominant activity of NES's located in several domains. The presence of JH activates JH-dependent NLS in PAS-B which acts synergistically with unmasked NLS located in the C-terminal fragment of protein and transports GCE mutant to the nucleus. 14-3-3 proteins typically bind to a partner as a dimer recognizing two binding sites, however interaction with a single site is also possible, although much weaker (Sluchanko et al., 2012; Sluchanko & Gusev, 2012). In the case of GCE, neither of the two predicted 14-3-3 binding sites alone was sufficient for interaction with these proteins, and both seem equally important for binding.

Summary of P1 and P2:

-P1 and P2 studies were the first detailed characterization of the subcellular traffic of *D. melanogaster* MET and GCE.

-In contrast to MET, which localized predominantly in the nucleus, wild type GCE was shown as ubiquitously localized throughout the cell independent on the presence of JH which supported my thesis about the different functioning of both proteins.

-I confirmed the homology between GCE and MET in their bHLH domain. Interestingly, I have not detected an NLS signal inside the bHLH domain, which was typical for other bHLH-PAS proteins. Also, it is worth noting that the number of basic amino acids in the bHLH domain of these proteins is lower in comparison to other bHLH proteins.

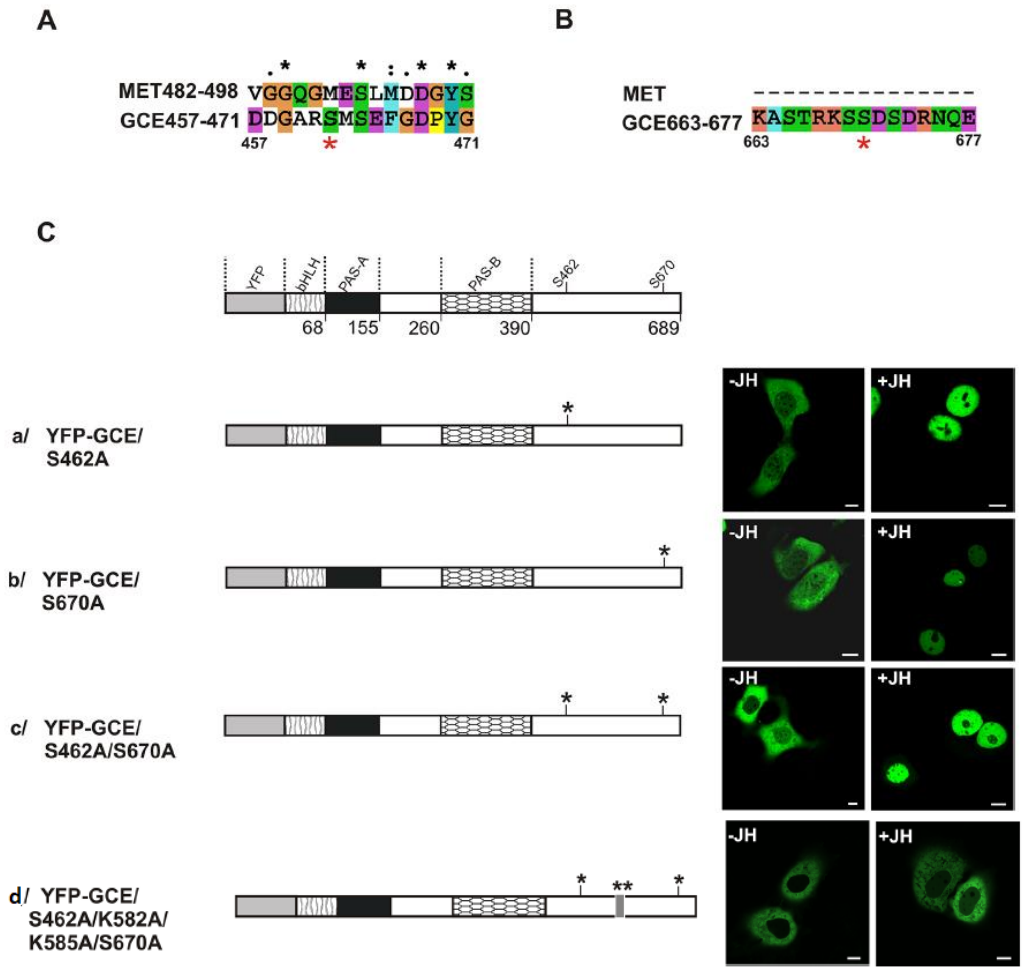


Figure 7. Subcellular distribution of the S462A and S670A mutants. (**A**) ClustalX alignment of the MET482-498/GCE457-471 encompassing GCE predicted 14-3-3 binding motif. (**B**) ClustalX alignment of the MET/GCE663-677 encompassing GCE predicted 14-3-3 binding motif. Asterisks indicate identical residues, and colons indicate similar residues. Residues numbered according to the sequence of GCE are given at the bottom of the alignment. Red asterisks indicate mutated residues. (**C**) The influence of JH on the subcellular distribution of the putative 14-3-3 binding sites and C-terminal NLS mutants of GCE. The length of each domain in the protein schematic illustration is relative. Regions of GCE are depicted using different patterns. Asterisks indicate mutated residues. Representative images (single confocal plane) present the typical subcellular distribution of the GCE derivatives. Bar, 10 μm [P2].

- detected an important difference in the linker between the bHLH and PAS-A domains. GCE does not possess the dominant and JH-independent NLS (RRRKK) that is present in MET (NLS-1, Fig. 5AB) which results in dominant nuclear localization of MET. Importantly, a comparative analysis of homologous sequences in other insect species revealed that this signal is a unique feature of *D. melanogaster* MET (Fig. 4A).

- I confirmed the homology between GCE and MET in PAS domains and verified the presence of active signals both in MET and GCE: NES-1 in PAS-A, NES-2 in PAS-B, and a JH-dependent NLS in PAS-B (GCE, NLS-1; MET NLS-2) (Fig. 5AB).

- I have shown that the C-terminal fragment of GCE is localized exclusively in the nucleus of the cell in contrast to the C-terminal fragment of MET, which was in the cytoplasm.

- I have identified the bipartite NLS 570-KRGSTAHVETEEKLSKRRF-589 (NLS-2; Fig 5B) in the C-terminal area of GCE which is absent in MET.

- I showed that the C-terminal NLS activity of GCE (NLS-2) is necessary for the JH-dependent transport of this protein to the nucleus and contributes to the functioning of the JH-dependent NLS in PAS-B.

- In contrast to MET, the C-terminus of GCE was predicted as interacting with the 14-3-3 proteins family. I showed that introduction of point mutations in the C-terminus of GCE (S462A and S670A) being predicted as targets of phosphorylation and parts of 14-3-3 protein-binding motifs changed the localization of GCE, which strongly suggests that 14-3-3 proteins could be one of modulating partners in addition to JH.

- MET and GCE were shown for the first time as shuttling proteins whose final intracellular localization can depend on a new and complex mode of multi-step regulation determined by the combined impact of nuclear localization signals, whose activity could be modified by other proteins/partners, JH, and post-translational factors.

- I have shown that regulation of MET and GCE localization through the existence of some similarities is differentiated. The final localization of GCE seems to be regulated in a much more complex manner than it is observed for MET. The proposed interaction of GCE with 14-3-3 proteins could be an important factor regulating the distribution and activity of this protein. The distinctive nature of the C-terminal parts of GCE and MET may be one of the main reasons for the functional divergence of these two proteins.

P3. Greb-Markiewicz B., Zarębski M. & Ożyhar A. (2018) Multiple sequences orchestrate subcellular trafficking of neuronal PAS domain–containing protein 4 (NPAS4). *J Biol Chem 293*: 11255–11270.

As the continuation of my studies, I was interested whether the complex pattern of localization signals shown for GCE and MET also occurs in mammalian representatives of bHLH-PAS family. Neuronal PAS domain-containing protein 4 (NPAS4) belonging to the first subclass of bHLH-PAS TFs was discovered in neurons in the neuronal layer of the mammalian hippocampus, however, later was also detected in non-neuronal tissues. NPAS4 was shown to be highly induced by activity and stress in pancreatic cells, to reduce insulin content, to impair

responsiveness to glucagon-like peptide 1 (GLP-1), and to protect pancreatic cells from ER stress (Sabatini et al., 2013), which resulted in the suggestion of NPAS4 as a therapeutic target in type 2 diabetes (Sabatini & Lynn, 2015) and treatment during pancreas transplantation (Speckmann et al., 2016). NPAS4 was also shown to promote neuroprotection in the injured brain following ischemic stroke and was proposed as a component of new therapies for stroke (Choy et al., 2015) which is the leading cause of long-term disability in developed countries and one of the top causes of mortality worldwide (Phipps & Cronin, 2020). NPAS4 has been proposed also as a therapeutic target for depression, neurodegenerative diseases associated with synaptic dysfunction (Zhang et al., 2014) and Alzheimer's disease (Fan et al., 2016). Among others, neurological disorders are major causes of death and long-term disability. Previously, many promising neuroprotective drugs have failed when translated to clinical trials. Performed evaluation of neuroprotection studies led to the hypothesis that one of the major reasons for failed neuroprotective strategies in cerebral ischemia/reperfusion injury was a single-target treatment without cerebral revascularization, which limits neuroprotectant delivery. For this reason, endovascular treatment was proposed as an important element of neuroprotective strategy (Xiong et al., 2018). Importantly, expression of NPAS4 was documented as a regulator of blood vessel formation and a factor necessary for endothelial cell function such as migration, sprouting, and branch formation in constructing new vasculature (Esser et al., 2017).

Despite presented important roles of NPAS4 and its potential use in therapy, the mechanism of action of this protein was not known. This led me to choose this protein as my research object to investigate molecular determinants responsible for regulation protein shuttling. The knowledge about the subcellular localization of NPAS4 during the time of P3 preparation was not conclusive. Previous studies performed in COS-7 and later in HEK293 cells 48 h after transfection reported strict nuclear localization of this protein (Moser et al., 2004; Shamloo et al., 2006). However, Sullivan et al. showed by immunofluorescent staining of fixed HEK293T cells 24 h after transfection that NPAS4 in addition to the nucleus was also present in the cytoplasm (Sullivan et al., 2016). Importantly, western blot analysis of samples obtained by subcellular fractionation of rat coronal brain tissue revealed NPAS4 expression in the nucleus and in the cytoplasm of cells. Ischemic insult resulted in partial translocation of NPAS4 from the cytoplasm to intracellular membranes. Moreover, in that study, NPAS4 was also detected in microsomal and synaptosomal fractions (Shamloo et al., 2006). These findings suggested that NPAS4 might contain motifs that direct its localization to different cellular compartments.

I decided to continue the methodology successfully used for GCE and MET. In addition to previously used COS-7 cells, I also used mouse albino neuroblastoma N2a cells which were shown as a suitable model to study neuronal differentiation and signaling pathways associated with this process (Namsi et al., 2018). Twenty-four hours after transfection of COS-7 and N2a cells, I analyzed the subcellular localization of YFP–tagged full-length NPAS4 and its derivatives by confocal and fluorescence microscopy. The expression of YFP-tagged NPAS4 and its derivatives in COS-7 cells was confirmed by Western blot analysis using an anti-GFP antibody. Because NPAS4 was linked to diabetes, I decided to perform comparative experiments in low and high glucose Dulbecco's modified Eagle's medium (DMEM). The low-glucose medium used in our experiments (4.5 mM) corresponded to the concentration of glucose in normal metabolism, whereas the high-glucose concentration (25 mM) corresponded to diabetes.

In experiments using the low-glucose medium for the expression of full-length NPAS4, a higher amount of fluorescent protein was observed in the cytoplasm (over 80% in COS-7 and 90% in N2a cells) than in the nucleus (less than 20% in COS-7 and 10% in N2a cells 24h after transfection. In the high-glucose medium, the proportion of cells with nuclear/cytosolic localization of NPAS4 was higher and achieved a ratio of 80/20 in COS-7 cells (Fig. 8 a) and 30/70 in N2a cells 24 after transfection. After 48 hours after transfection, this proportion reached 90/10 in COS-7 and 50/50 in N2a cells. Such differentiation of obtained results dependent on the used media and cell lines indicated that NPAS4 is a shuttling protein, and the concentration of glucose and specific cellular factors are engaged in the regulation of this process. Obtained results confirmed also that NPAS4 might possess motifs functioning as NLSs and NESs. To verify the hypothesis regarding the presence of an NES in NPAS4 I used LMB. The addition of LMB to COS-7 (Fig. 8 b) and N2a cells resulted in strictly nuclear localization of NPAS4, indicating the inhibition of protein transfer between the nucleus and the cytoplasm and switching off existing NES/ NESs which confirmed the hypothesis about the presence of active NLS. Experiment with methanol (LMB solvent) addition verified that the LMB solvent was not responsible for changes in protein localization after LMB addition [P3].

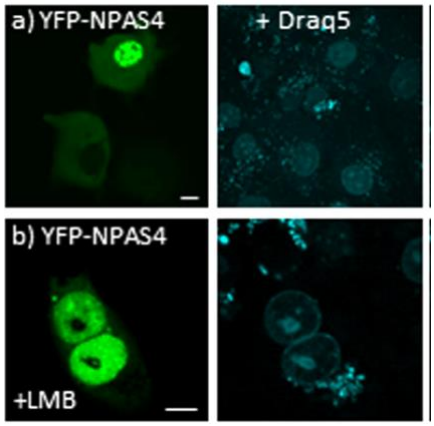


Figure 8. Subcellular distribution of the full-length NPAS4. Representative images of the subcellular distribution of YFP-NPAS4 in COS-7 cells in a high-glucose medium 24h after transfection showed under normal conditions (**a**), after LMB addition (**b**). Nuclei and nucleoli were stained by Draq5. Bar, 10 μm [P3].

To search for NLS/NLSs and NES/NESs in different parts of NPAS4, analogically to P1 and P2, I performed a set of experiments using a series of deletion mutants. To ensure the functionality of the structural motifs within the NPAS4 sequence, truncated regions were designed based on *in silico* analysis: subcellular localization motifs with all available in time of P3 preparation predictors for NLS (NucPred, PSORTII, cNLSMapper NLStradamus), nucleolar localization signal (NoLS) - NoD and NES (NetNes 1.1 server, ValidNES, NES Finder 2.0 and LocNES), the secondary structure prediction with Protein Structure Prediction Server (PSIPRED), the domain prediction with Simple modular architecture tool (SMART) and Database of protein families and domains PROSITE) [P3]. The *in silico* analysis suggested the presence of NLSs in the bHLH domain, the PAS-1 domain, and the C-terminal region of NPAS4. Accordingly, putative NESs were expected in the bHLH domain, the PAS-2 domain, and the C-terminal region of NPAS4. Interestingly, the predicted NLS and NES motifs were often located within a very close distance or partially overlapping each other. I divided NPAS4 in several parts: bHLH domain (1-59 aa), PAS-1 (60-144 aa), linker between PAS domains (145-194 aa), PAS-2 (195-276 aa and 277-340 aa) and C-terminal fragment (341-460 aa, 461-580 aa, 581-650 aa, 651-720 aa, 721-801 aa) as schematically presented in Fig. 9.

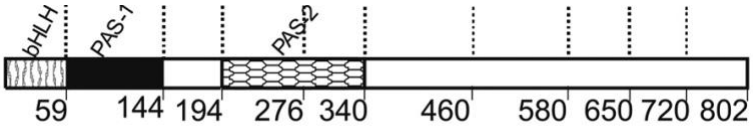


Figure 9. Schematic representation of NPAS4. Regions of NPAS4 are depicted using different patterns [P3].

Expression of the YFP-tagged bHLH domain resulted in nuclear and nucleolar fluorescence signals (Fig.10a), though NoLS prediction for NPAS4 was negative. Performed analysis suggested the presence of partially overlapping NLS and NES in this area. To verify the theory of overlapping signals, I prepared constructs with further truncations and point mutations (Fig. 10). Based on the results of prediction and experiments, I concluded that the

bHLH domain contains multipartite NLS motifs comprising 10–52 aa (10-KARRDQINAEIRNLKELLPLAEADKVRLSYLHIMSLACIYTRK-52) and NES comprising 26-45 aa (26-LLPLAEADKVRLSYLHIMSL-45). Detailed explanation and discussion of results are presented in P3.

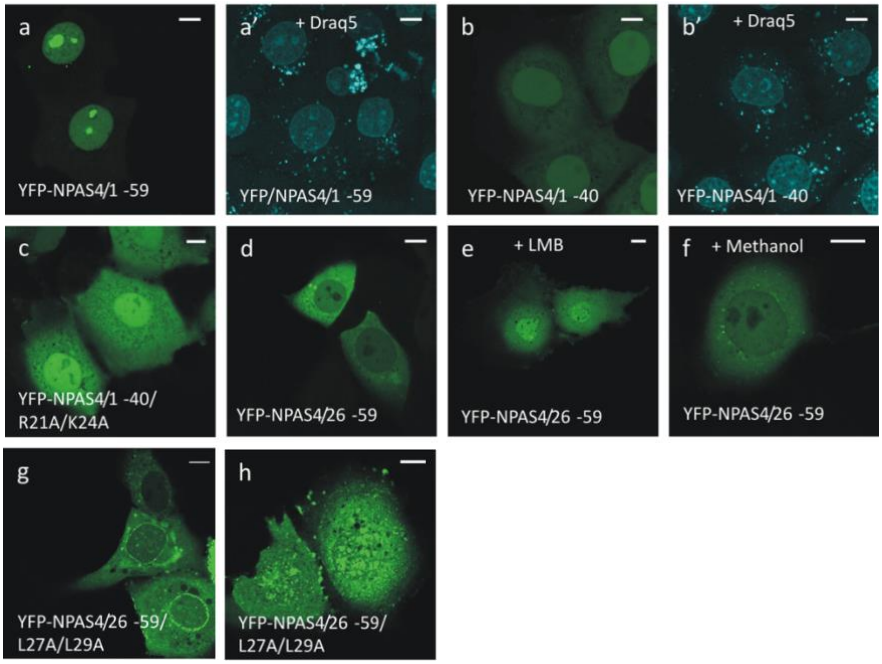


Figure 10. Subcellular distribution of the bHLH domain of NPAS4 and their derivatives in COS-7 cells in a low-glucose medium. Representative images of typical (presented by more than 95%) subcellular distributions of the YFP-tagged derivatives. Nuclei and nucleoli were stained with Draq5. Bar, 10μm. Shown are YFP-NPAS4/1–59 (a and a'), YFP-NPAS4/1–40 (b and b'), YFP-NPAS4/1–40/R21A/K24A (c), and YFP-NPAS4/26–59 (d) under normal conditions; YFP-NPAS4/26–59 after LMB addition (e); YFP-NPAS4/26–59 after methanol addition (f) YFP-NPAS4/26–59/L27A/L29A (g, h) [P3].

Then, I continued searching for NLS and NES in the PAS-1 and PAS-2 domains. Interestingly, I observed that the YFP-NPAS4/60–144 fragment containing the PAS-1 domain expressed in COS-7 cells in the low-glucose medium was localized predominantly to the cell nucleus (Fig. 11 a-a'), although no putative NLS was predicted. However, I could not detect any basic aa rich cluster and the expression of this fragment in all tested cells in high-glucose conditions resulted in ubiquitous distribution and the fragment was not further tested. In the next tested fragment, the linker located between the PAS-1 and PAS-2 domains presented a signal of fluorescence exclusively in the nucleus of all analyzed cells independent of glucose concentration (Fig. 11 b-b'), which confirmed the presence of predicted NLS activity. The cytoplasmic shift of the expressed point mutant enabled confirmation of 158-RRQSAGNKLVLIRGRFHAHPPGAYWAGNPVFTAF-191 sequence as an active NLS. The

distribution of the N-terminal PAS-2 fragment was cytoplasmic (Fig. 11 d) which corresponds to the presence of predicted NES in this area. The shift of fluorescence from exclusively cytoplasmic to distributed evenly in both compartments after LMB addition (Fig. 11 e-f) confirmed NES activity. I performed the final verification by substitution of hydrophobic residues: L230, I232, and L237 by A. Surprisingly, the resulting protein was located around the nuclear membrane in some of the analyzed cells (Fig. 11 g-h'). This result together with observations presented in Shammlo et al. (Shamloo et al., 2006) led me to the hypothesis that NPAS4 could contain transmembrane or interacting with membrane fragments. Performed in silico analysis seemed to support this thesis. I plan to continue the subject and verify this in the future (described in the section scientific plans).

The expressed C-terminal part of PAS-2 localized exclusively in the nucleus of all analyzed cells (Fig. 11 i), which was consistent with the cNLS Mapper prediction of the 285-RLQAKHGGWTWIYCMLYSDGPEGPITANNYPI-316 sequence as an NLS. The presence of NLS was confirmed by the expression of YFP-NPAS4/257–340/R285A/K289A mutant resulting in the fluorescence shift from exclusively nuclear to predominantly nuclear. Interestingly, a distribution pattern in the nucleoplasm was changed.

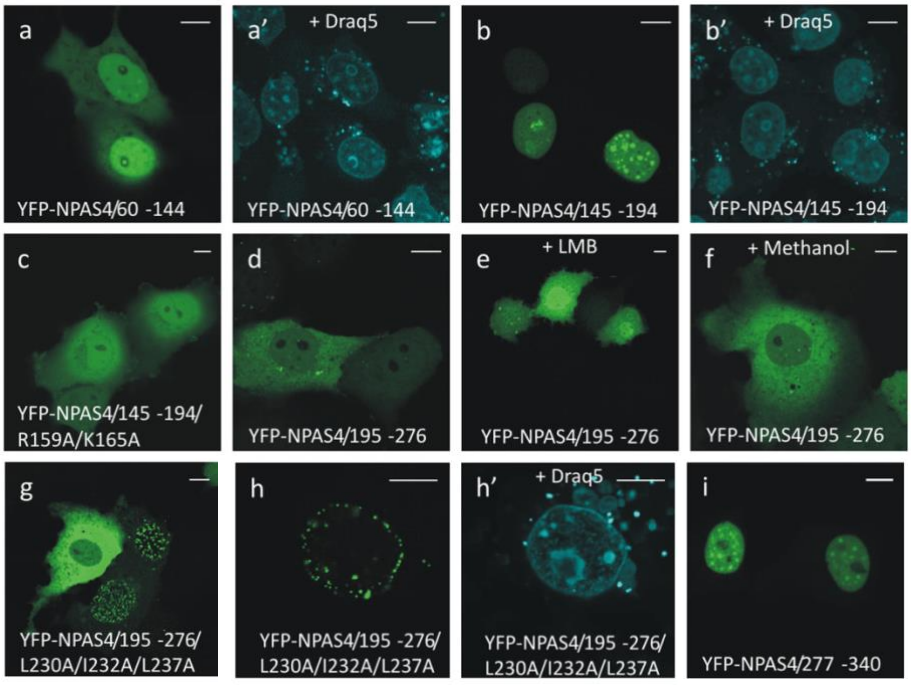


Figure 11. Subcellular distributions of the PAS domains of NPAS4 and their derivatives in COS-7 cells in a low-glucose medium. Representative images of typical (presented by more than 95% of cells unless stated otherwise) subcellular distributions of the YFP-tagged derivatives of the NPAS4 PAS domains. Nuclei and nucleoli were stained with Draq5. Bar, 10μm. Shown are YFP-NPAS4/60–144 (a and a'), YFP-NPAS4/145–194 (b and b'), YFP-NPAS4/145–194//R159A/K165A (c), and YFP-NPAS4/195–276 (d) under normal conditions; YFP-

NPAS4/195–276 after LMB addition (e); YFP-NPAS4/195–276 after methanol addition (f); prevailing patterns of YFP-NPAS4/195–276/L230A/L232A/L237A distributions (g, h, and h'), YFP-NPAS4/277–340 (i) [P3].

To test the presence of subcellular localization signals in the C-terminal region of NPAS4 encompassing residues 341- 802, YFP-tagged fragments: 341-460 aa, 461-580 aa, 581-650 aa, 651-802 aa, and 341-580 aa, 461-650 aa, 581-720 aa, 651-802 aa (Fig. 12) were expressed. YFP- NPAS4/341–460 was distributed ubiquitously throughout the whole cell as expected because no NLS/NES was predicted for this fragment (Fig. 12 a). Surprisingly, YFP-NPAS4/461–580 localized to the cytoplasm despite negative results of NES predictions for this fragment (Fig. 12 b). The addition of LMB shifted the fragment localization to the nucleus (Fig. 12 c), confirming the presence of an active NES interacting with exportin 1, while methanol addition (Fig. 12 d) verified that methanol had no localization impact. Interestingly, expression of the YFP-NPAS4/341–580 fragment, comprising both previously tested fragments, resulted in the evenly distributed fluorescence signal (Fig. 12 e). I hypothesized that some residues in the 341–460–aa fragment deactivate the putative NES in the 461–580–aa fragment by interacting with an unknown factor, thus masking the NES. As a typical NES sequence is rich in L residues, I performed additional analyses of the sequence of this region (461–580 aa). However, I did not find any obvious cluster of L residues. For this reason, I was not able to specify the putative NES sequence in this area.

The localization of YFP-NPAS4/581–650 (Fig. 12 f) was cytoplasmic in all analyzed cells, both in low and high-glucose medium. In addition to positive results of LMB experiments confirming the presence of motif interacting with exportin-1, I performed expression of the YFP-NPAS4/581–650/L591A/L594A/L598A mutant (Fig. 12 i). The shift of fluorescence from the cytoplasm to the nucleus proved the deactivation of the predicted NES 591-LAQLRGPLSV-600 sequence. Interestingly, nuclear localization of this mutant supported results of cNLS Mapper prediction suggesting 593-QLRGPLSVDVPLVPEGLLTPEASPVKQSFF-622 sequence as a putative NLS. This sequence partially overlaps with a region containing NES (591–600 aa). To test the influence of additional amino acid residues on NLS and NES activities, I generated YFP-NPAS4/461–650 and YFP-NPAS4/581–720 constructs containing both detected signals.

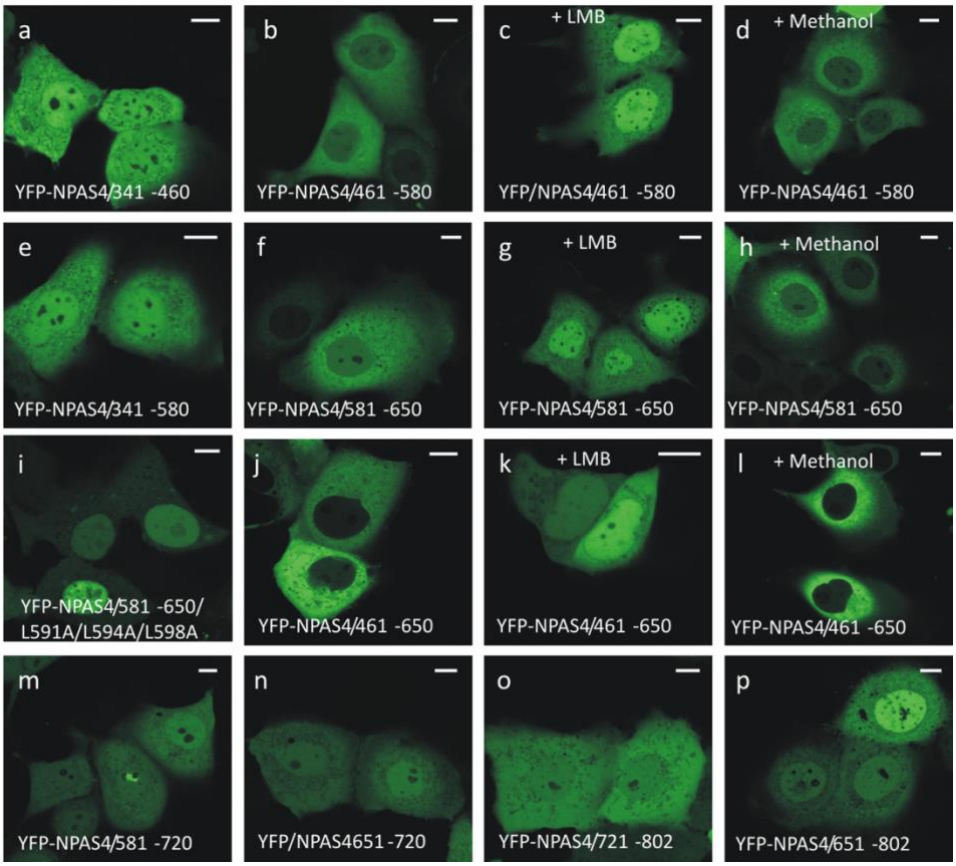


Figure 12. Subcellular distribution of the C-terminal part of NPAS4 deletion mutants in COS-7 cells in a low-glucose medium**.** Representative images of typical (presented by more than 95% of cells) subcellular distributions of the YFP-tagged derivatives of the NPAS4 C terminus are presented. Bar, 10 μm. Shown are YFP-NPAS4/341–460 (a) and YFP-NPAS4/461–580 (b) under normal conditions; YFP-NPAS4/461–580 after LMB addition (c); YFP-NPAS4/461–580 after methanol addition (d); YFP-NPAS4/341–580 (e) and YFP-NPAS4/581–650 under normal conditions (f); YFP-NPAS4/581–650 after LMB addition (g); YFP-NPAS4/581–650 after methanol addition (h); YFP-NPAS4/581–650/L591A/L594A/L598A (i) and YFP-NPAS4/461–650 under normal conditions (j); YFP-NPAS4/461–650 after LBM addition (k); YFP-NPAS4/461–650 after methanol addition (l); YFP-NPAS4/581–720 (m); YFP-NPAS4/651–720 (n); YFP-NPAS4/721–802 (o); and YFP-NPAS4/651–802 (p) [P3].

Expression of the YFP-NPAS4/461–650 resulted in exclusive cytoplasmic fluorescence in COS-7 cells both in low- (Fig 12 j) and high-glucose conditions and in N2a cells, thus predominantly presenting characteristics of the NES in the tested region. The addition of LMB resulted in the uniform localization of the mutant in both compartments of COS-7 cells in low-glucose medium (Fig. 12k) and in N2a cells, however, in high-glucose DMEM, I observed an evident shift in localization to the nucleus after LMB addition. I expected exclusive nuclear localization of YFP-NPAS4/461–650 in LMB presence because the L594–F622 sequence present in this fragment was predicted as NLS and was putatively responsible for the nuclear localization of the YFP-NPAS4/581–650/L591A/L594A/L598A mutant. Therefore, I had no clear explanation for the different localization behavior observed in COS-7 cells depending on

glucose concentration. One possibility is that the presence of additional amino acid residues (461–580 aa) could result in NLS activity inhibition in a low-glucose medium via an unknown mechanism. Expression of YFP-NPAS4/651–720, YFP-NPAS4/721–802 and YFP-NPAS4/651-802 fragments resulted in the fluorescence distributed evenly in the nucleus and the cytoplasm (Fig. 12 n, o, p) confirming the absence of active NLS/NES signal in this area.

Detection of NLSs/NESs within different NPAS4 domains led me to ask a question about their relative strength resulting in the hierarchy of NLSs and NESs in NPAS4. For this purpose, I prepared a wide spectrum of NPAS4 truncation mutants encompassing different sets of documented signals and performed their expression in COS-7 and N2a cells. Results are presented and discussed in detail in P3. Importantly, I hypothesized that the relative strength of signals may be dependent on the cell type and the glucose concentration.

To date of P3 preparation, ligand binding by the PAS domains of NPAS4 has not been tested. When searching for the explanation of the biological significance of the presence of NLSs and NESs in PAS domains, I performed predictions of NPAS4 ligand binding using the 3DligandSite server. The predictions resulted in the suggestion that the PAS-1 and PAS-2 domains bind heme and flavin mononucleotide (FMN), which is consistent with the results of Wu et al. (Wu et al., 2016), who found that mammalian bHLH-PAS transcription factors bind multiple ligands. My hypothesis of interaction with heme was substantiated by the fact that heme binding was documented for the PAS domains of another neuronal PAS domain–containing protein, NPAS2. Interestingly, the binding of heme by NPAS2 forms a gas-regulated sensor; however, the natural ligand of this protein is not known. In turn, FMN is present in light–oxygen–voltage (LOV) proteins, which belong to the PAS family, as a chromophore involved in blue-light absorption that is important, for example, in regulating circadian rhythms (Röllen et al., 2016). The predicted ligand binding by the PAS-2 domain occurs partially via amino acids located in regions of detected NLS and NES sequences which suggests ligand binding–dependent regulation of NPAS4 subcellular localization. Also, the presence of multiple opposing and partially overlapping localization signals suggests a complex mode of regulation of NPAS4 shuttling in response to different factors.

NLS and NES sequences are short linear motifs (SLIMs). The preferential location of these motifs in the intrinsically disordered regions (IDRs) of proteins has been found to enable flexible and easily accessible interactions with their binding partners (Dinkel & Sticht, 2007). To substantiate the results, I performed *in silico* analysis of NPAS4 showing that 67% of the NPAS4 sequence (mostly C-terminal part) was disordered. IDRs are known targets of intensive posttranslational modifications (PTM). Modifications such as phosphorylation, especially near

the NLS or the NES, have been shown to regulate the intracellular distribution of proteins by activating or deactivating the localization motifs (Jans & Hübner, 1996). I performed predictions for NPAS4 phosphorylation sites and found that many NPAS4 serine residues could be phosphorylated with a high probability. Some of the predicted serine residues were found in the NLSs or NESs, or near these motifs which suggested that phosphorylation is an influencing factor of NPAS4 localization. Western blot analysis showed the presence of dual bands for expressed YFP-NPAS4/277–340 (the area comprising NLS-3) and YFP-NPAS4/581–650 (the area comprising NES-4 and NLS-4) which supported the hypothesis regarding the role of PTMs.

The ratio of nuclear to cytoplasmic NPAS4 in the presented studies was dependent on the glucose concentration in the medium. The low-glucose medium used in our experiments (4.5 mM) corresponded to the concentration of glucose in normal metabolism, whereas the high-glucose concentration (25mM) corresponded to diabetes. These results are consistent with the proposed role of NPAS4 as a therapeutic target in type 2 diabetes protecting pancreatic cells from ER stress (Sabatini et al., 2013). Importantly, NPAS4 is one of the immediate early genes (IEG) that activate mechanisms related to the first defense against many cellular stresses (Fowler et al., 2011). IEGs are regulated by a specific stimulus without de novo protein synthesis (Greenberg et al., 1986). Shamloo et al. proposed that cytoplasmic NPAS4 in brain cells serves as a source for translocation to other cellular compartments after ischemia (Shamloo et al., 2006), a finding that could suggest and explain a role for NPAS4 shuttling in cells.

Summary of P3

- I showed that the bHLH domain of NPAS4, responsible for ARNT binding and interaction with DNA, contains NLS-1 (10-52 aa) and overlapping NES-1 (26-45 aa) (Fig. 13A). Interestingly, NLS-1 seems to be linked directly with NoLS. Based on the results, I hypothesized that the bHLH domain is an important player in the dynamic exchange of NPAS4 between subcellular compartments.
- I documented the presence of NLS-2 (158-191 aa), and NES-2 (227-242 aa) located near the PAS-2 domain, which is usually responsible for ligand binding and sensing environmental signals. In the C-terminal region of the PAS-2 domain, I hypothesized the presence of NLS-3 (285-316 aa) (Fig. 13A).
- I performed predictions of ligand binding by NPAS4 and proposed that the PAS-1 (PAS-A) and PAS-2 (PAS-B) domains bind to heme and flavin mononucleotide (FMN). The predicted ligand binding by the PAS-B domain occurs partially via amino acids located in regions of

NLS-2, NES-2, and NLS-3 suggesting that regulation of NPAS4 subcellular localization could be dependent on the ligand binding.

- NLS and NES sequences are often considered as short linear motifs (SLIMs). The preferential location of these motifs in the intrinsically disordered regions (IDRs) of proteins has been found to enable flexible and easily accessible interactions with their binding partners (31). I performed *in silico* analysis of NPAS4 showing for the first time that 67% of the NPAS4 sequence (mostly C-terminal part) is disordered.

- NPAS4 structure during P3 preparation was unknown, so I generated a 3D model of NPAS4 using Phyre2 (Kelley et al., 2015) to predict the surface exposure of defined NLSs and NESs and visualized by PyMol (Fig. 13B). NLSs and NESs from the bHLH and PAS domains were in or adjacent to short IDRs, which provided some degree of flexibility and accessibility, especially NLS-1 and NES-1 in the bHLH domain are easily accessible for interaction with the partners. In contrast, other signals are in areas that can exist in multiple conformations adopted by the disordered NPAS4 C-terminus which can influence the exposure and accessibility of signals located in this part of the protein. I believe that the location serves to enhance flexibility and precise mechanistic control of NPAS4 activity in the context of interactions with ligand and NPAS4 coactivators/corepressors.

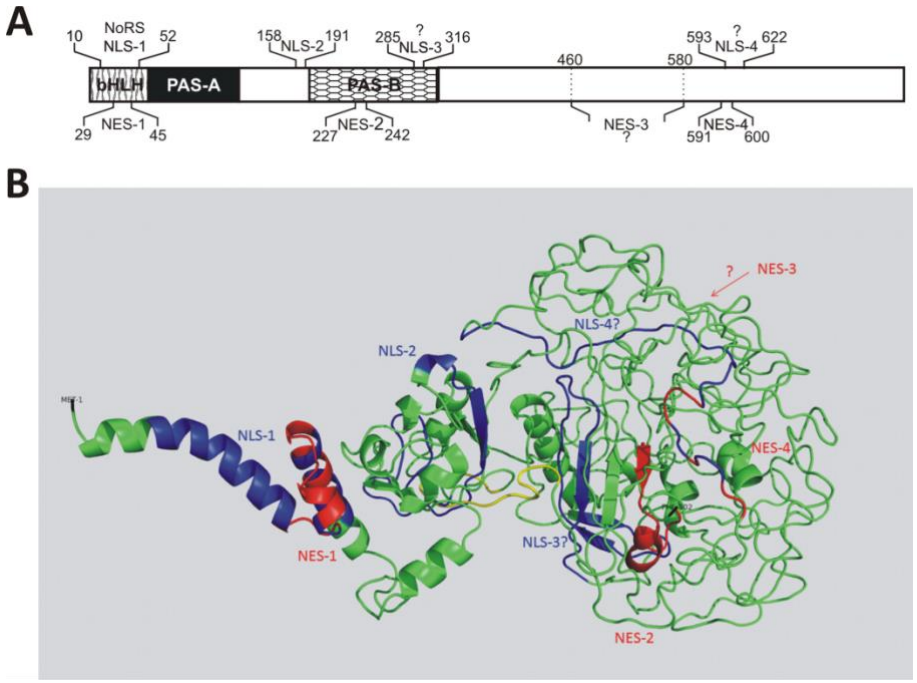


Figure 13. Summary of NLSs and NESs in NPAS4. (**A**) schematic representation of NLSs and NESs residing within the NPAS4 protein structure. (**B**) NLSs and NESs presented in a 3D model of NPAS4 generated by Phyre2 (Kelley et al., 2015) and visualized using PyMol (PyMOL Molecular Graphics System). Residues 1–340 encompassing the bHLH and PAS domains of NPAS4 were modelled with confidence higher than90%using the templates c4f31B, c4zprB, C4f31A, c4zp4D, c5sy5B and c5sy7B. The C-terminal region of NPAS4 lacking any

template and predicted as disordered was modelled ab initio (67% of the protein is predicted as disordered). NLS-1, NLS-2, and putative NLS-3 and NLS-4 are shown in blue. NES-1, NES-2, NES-3, and NES-4 are shown in red. The unique mammalian NPAS4 with a sequence rich in PG and PR repeats, where the Pro201 residue was predicted (probability 0.85, HydroxyPred) as the site of hydroxylation (197-PRPRPGPGPGPGPGP-211), is shown in yellow [P3].

- IDRs are known targets of intensive posttranslational modifications (PTMs). Modifications such as phosphorylation, especially near the NLS or the NES, have been shown to regulate the intracellular distribution of proteins by activating or deactivating the localization motifs. My predictions for NPAS4 phosphorylation sites showed that 62 serine residues of NPAS4 are phosphorylated with a high probability. Some of the predicted serine residues were found in the NLSs or NESs, and some were found to be near these motifs which suggested that phosphorylation could be an influencing factor in NPAS4 localization. The occurrence of PTMs of NPAS4 was supported by the presence of dual bands in western blot analysis.
- The importance of NPAS4 nuclear localization as TF is obvious, however the role of NPAS4 in many different cellular processes suggests an important nongenomic role for the cytoplasmic NPAS4. The proportion of nuclear to cytoplasmic NPAS4 in our studies was dependent on the glucose concentration in the medium. The low-glucose medium used in our experiments (4.5 mM) corresponded to the concentration of glucose in normal metabolism, whereas the high-glucose concentration (25mM) corresponded to diabetes. Obtained results indicate that glucose concentration could be a factor influencing NPAS4 localization by regulating NPAS4 activity as a transcription and cytoprotective factor.
- The linker between PAS-1 and PAS-2 domains presents the unique for mammalian NPAS4 sequence rich in PG and PR repeats, where the P201 residue is putative site of hydroxylation (197-PRPRPGPGPGPGPGP-211). Elucidating the role of this interesting fragment is one of my future goals presented in section 4.4.
- I showed that NPAS4 possesses multiple opposing and partially overlapping localization signals and is a shuttling protein with a very complicated and precise regulation of subcellular localization. The disordered character of the NPAS4 C terminus may serve as a platform for simultaneous interactions with many partners, which may be responsible for the very precise and multifactor-dependent regulation of NPAS4 trafficking and may be crucial for the ability of NPAS4 to act as a biological sensor and switch to modulate the crosstalk of different signaling pathways. For this reason, understanding of this mechanism can be an important step in establishing NPAS4 as a therapeutic target.

P4. Greb-Markiewicz B., Kazana W., Zarębski M. & Ożyhar A. (2019) The subcellular localization of bHLH transcription factor TCF4 is mediated by multiple nuclear localization and nuclear export signals. *Sci Rep 9*: 15629.

Since previously studied proteins represented class VII of the bHLH family, I found it important to take as an object of study also protein representing other class of the bHLH family. I have chosen the TCF4 protein as an example of class I (E proteins) that play a very important role in the development of the nervous system by regulation of neurogenesis and cell specialization. TCF4 binds the canonical enhancer-box (E-box) sequence 5′-CANNTG-3′ and influences chromatin remodeling and transcription through the recruitment of p300 with histone acetyltransferases (HATs) activity (Kennedy et al., 2016). TCF4 is known to be the object of alternative splicing of 5′ exons (Sepp et al., 2011) which generates many isoforms differing in the presence of activation and repressor domains (Sobrado et al., 2009). Decreased TCF4 expression has been linked to schizophrenia and bipolar disorder (characterised by alternating episodes of depression and mania) (Blake et al., 2010). In 2007 the TCF4 gene was linked to the presence of the mental disorder Pitt-Hopkins Syndrome (PHS). Most mutations that cause PHS are located within the C-terminally located bHLH domain of the protein. For this reason, it was suggested that interaction with DNA and other proteins is impaired (Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007). Additionally, common TCF4 gene variants are risk factors for non-psychotic disorders like Fuchs Corneal endothelial dystrophy (FECD) (Hill et al., 2014; Lau et al., 2014) and primary sclerosing cholangitis (PSC) (Ellinghaus et al., 2013).

TCF4 expression was reported in different organs: brain, heart, kidney, lungs, muscle, spleen, and testis, and in lesser amounts in the liver, prostate, and ovaries (Muir et al., 2006). A particularly high-level expression of TCF4 is observed in the embryonic central nervous system, mesoderm, and adult brain (Wirgenes et al., 2012). TCF4 regulates the neurogenesis and differentiation of cells by forming heterodimers with proneuronal activators belonging to the bHLH family, like achaete-scute homolog 1 (ASCL1), protein atonal homolog 1 (ATOH1) and neurogenic differentiation factor 1 (NEUROD1). Conversely, the interaction of TCF4 with the DNA-binding protein inhibitor ID-2 (ID2) results in the formation of an inactive heterodimer, thus blocking TCF4 binding with activators (M. P. Forrest et al., 2013). To date, not much is known about the role of TCF4 in the central nervous system. However, in situ tests have shown that TCF4 is an important factor in the regulation of glial cell differentiation, especially the maturation of oligodendrocyte progenitors, and it also plays an important role in the regulation of the nuclei development of the pons involved in motor activity (Flora et al., 2007; Fu et al., 2009; Sweatt, 2013). In addition, TCF4 is involved in the differentiation of

plasmacytoid dendritic cells forming part of the adaptive immune system. TCF4 is also crucial for the development of lymphoid progenitor cells, giving rise to T- and B-lineage cells (Blake et al., 2010). Additionally, TCF4 regulates the development of Sertoli cells, myogenesis, melanogenesis and epithelial-mesenchymal transition (Sepp et al., 2011).

The subcellular distribution of TCF4 was proposed previously as one of the mechanisms regulating the function of this protein (Onions et al., 2000). The first report concerning subcellular localization of TCF4 showed this transcription factor to be only in the nucleus (Corneliussen et al., 1991) while related proteins from the class I bHLH family, like E2a and HEB, were detected both in nuclear and cytoplasmic fractions of human embryonic stem cells (Yoon et al., 2011). Detailed studies examining the expression of isoforms of TCF4 that differ in their N-terminal and internal sequences revealed that isoforms vary in terms of their subcellular localization. The longest isoform B presented a strictly nuclear localization, while others were present in both the nuclei and cytoplasm of HEK 293 cells and human brain tissue (Sepp et al., 2011). Later, Brandl et al. investigated TCF4 distribution in human colorectal carcinomas cells. They showed that TCF4 localization in the cytoplasm was positive in 76% of cases, while localization in the nucleus was positive in 51% of cases. Importantly, the cytoplasmic localization of TCF4 was correlated with better survival (Brandl et al., 2015). Cytoplasmic localization of TCF4 was also observed by d'Rosario et al. (D'Rozario et al., 2016). In their research, TCF4 was only detected in the nuclei of some neurons, while in the cytoplasm (soma and dendrites) it was seen in most post-mitotic neurons. Sepp et al. (Sepp et al., 2017) did not agree with these results. They presented results of the overexpression of different isoforms of TCF4 in the primary neurons of rats, which were fully consistent with previously published results (Sepp et al., 2011). The authors also declared that the regulation of TCF4 by neuronal activity cannot be attributed to its signal-dependent nuclear import. Finally, in 2018, Jung et al. showed that TCF4 expression in the lateral (LA) and basolateral (BLA) nuclei in the amygdala of young adult mice was nuclear, while in the central amygdala nucleus (CEA) the TCF4 was in fibers and not in the nucleus. The authors suggested the possibility that TCF4 activity could be modulated by nuclear-cytoplasmic translocation (Jung et al., 2018).

Due to the discrepancies presented above, I decided to systematically analyze the presence of putative NLSs and/or NESs and their role in the subcellular trafficking of TCF4-B^{+} isoform. The distribution of TCF4 fused with YFP in more than 95% of COS-7 cells was strictly nuclear (Fig. 14A a, a') however, there were some cells (less than 5%) presenting fluorescence

not only in the nucleus, but also in the cytoplasm in the puncta that were non-uniformly distributed in the cytoplasm of cells (Fig. 14B). In contrast, the expression of YFP-TCF4 in N2a cells was exclusively nuclear (Fig. 14A b, b'). Differences in localization of TCF4, depending on the cell type and literature data, led me to the hypothesis about the putative presence of the composite system of NLSs and NESs in TCF4. This hypothesis was supported by a previous report indicating that related proteins, E2a and HEB, were observed in the nucleus and cytoplasm of human embryonic stem cells (hESCs).

To search for NLS/NLSs and NES/NESs in TCF4, I performed a set of experiments using a series of deletion mutants like previously. To design truncations, I performed *in silico* analysis: subcellular localization motifs with all available in time of P4 preparation predictors for NLS (NucPred, PSORTII, cNLSMapper NLStradamus), nucleolar localization signal (NoLS) - NoD and NES (NetNes 1.1 server, ValidNES, NES Finder 2.0 and LocNES), the secondary structure prediction with Protein Structure Prediction Server (PSIPRED), the domain prediction with Simple modular architecture tool (SMART) and Database of protein families and domains (PROSITE) [P4].

A **B**

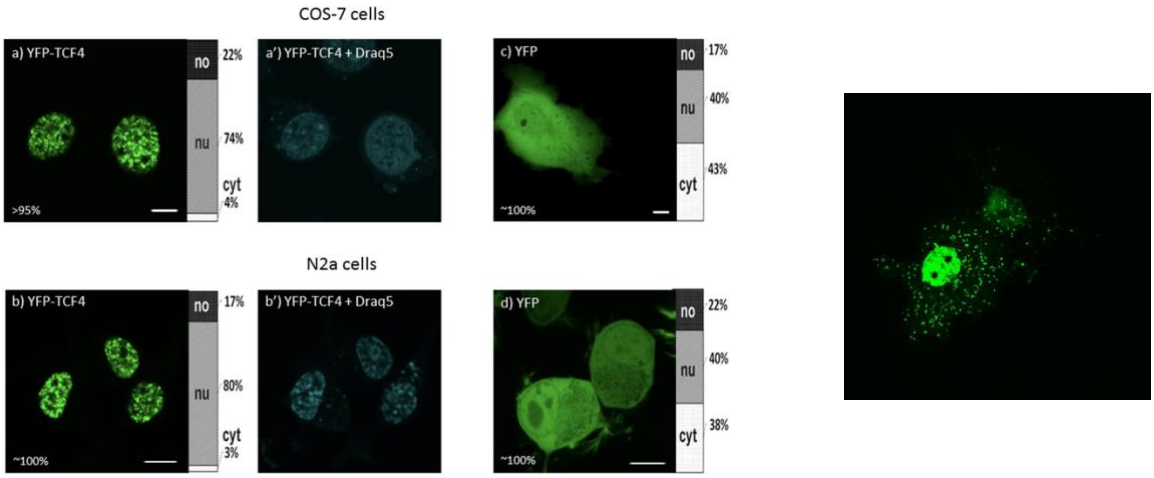


Figure 14. Subcellular distribution of the full-length TCF4 isoform B. Subcellular localization of YFP-tagged TCF4 was analyzed by confocal microscopy 20-24 h after transfecting COS-7 and N2a cells. Draq5 was added to the cells for DNA visualization. Ratios between mean fluorescence intensity in cytoplasmic, nuclear, and nucleolar compartments are presented as an accumulated bar graph (no- nucleolus; nu- nucleus, cyt- cytoplasm). (**A**) Representative images (single confocal plane) of the subcellular distribution of YFP-TCF4 (a-b') and YFP (c-d) in COS-7 and N2a cells. Bar, 10 μm. (**B)** Some of the COS-7 cells expressing YFP-TCF4-B presented additionally to nuclear, also specific fluorescence patterns in the cytoplasm [P4].

Navarrete et al. suggested that the N-terminal region of TCF4 could contain an active NES, however experimental verification of prediction was not performed (Navarrete et al., 2013). I have used NES predictors available during our study and obtained analogical positive

results for 10–19 amino acid residues. To test experimentally this prediction, deletion fragments comprising 1–30, 31–146, and 1–146 aa fragments fused C-terminally to YFP. The expression of YFP-TCF4/1–30, YFP-TCF4/31–146 and YFP-TCF4/1–146 in the COS-7 and N2a cells resulted in the ubiquitous distribution of fluorescence throughout the whole cell showing that the predicted motif was not an active NES. The lack of activity can be explained by results of NetSurfP *in silico* analysis, showing that leucine residues important for classical NES are predicted to be not exposed to the surface of the protein.

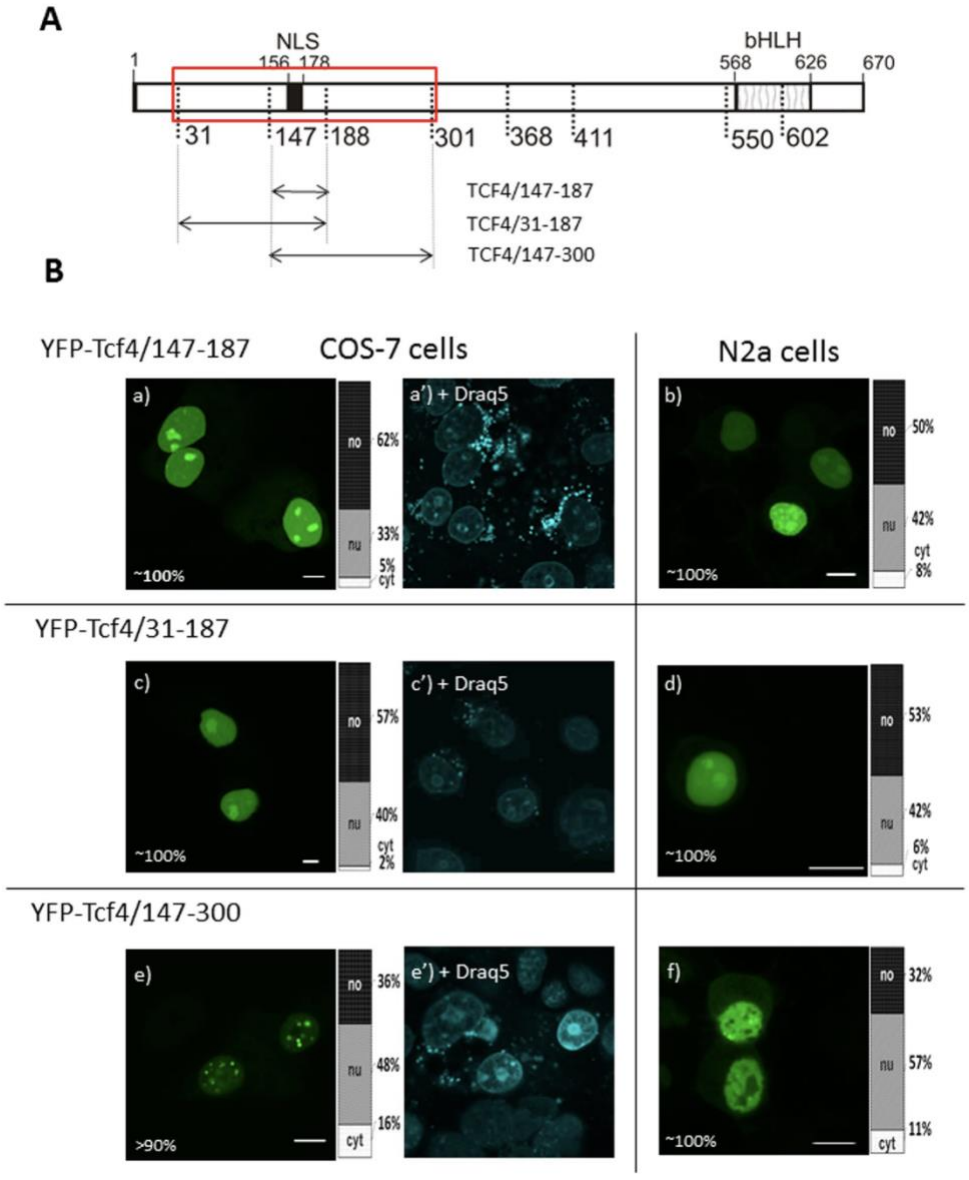


Figure 15. Subcellular distribution of deletion mutants (147-300aa) of TCF4. (**A**) Schematic representation of TCF4 protein. Region of the studied area of TCF4 is shown by the red rectangle. Expressed deletion mutants of TCF4 are depicted as arrows. The length of each domain in the diagram is relative (**B**) Subcellular distribution of deletion mutants of TCF4. Subcellular localizations of the expressed proteins were analyzed by confocal microscopy 20-24 h after transfecting COS-7 and N2a cells. Representative images (single confocal plane) of subcellular distribution of the derivatives of the TCF4/31-300 area. Bar, 10 μm. Draq5 was added to the cells for DNA visualization. Ratios between mean fluorescence intensity in cytoplasmic, nuclear, and nucleolar

compartments are presented as an accumulated bar graph (no- nucleolus; nu- nucleus, cyt- cytoplasm). (a, a', b) YFP-TCF4/147-187, (c,c',d) YFP-TCF4/31-187, (e-f) YFP-TCF4/147-300 [P4].

Importantly, Sepp with co-workers identified bipartite NLS in the region comprising aa residues 156–178 in the TCF4-B isoform (Sepp et al., 2011). This NLS was shown to be conserved among E-proteins. We prepared YFP-tagged truncation mutants of TCF4 containing this area (Fig. 15A). The expression of YFP-TCF4/147–187 resulted in a strictly nuclear fluorescence in the COS-7 cells confirming NLS activity (Fig. 15B a). Importantly, we observed nucleolar localization (Fig. 15B a'). Moreover, in the case of YFP-TCF4/147–187 expression in the N2a cells, we observed fluorescence in both the nuclei and nucleoli (Fig. 15B b). I performed FRAP analysis (not published) showing rapid recovery of expressed YFP-147-187 fluorescence in the nucleolus. Interestingly, the expression of YFP-TCF4/31–187 resulted in the same results for both the COS-7 (Fig. 15B c, c') and N2a (Fig. 15B d) cells, while YFP-TCF4/147–300 presented only nuclear and not nucleolar localization. Instead, some punctate pattern appeared in the nuclei of the COS-7 cells (Fig. 15B e, e'), and to a lesser extent in the N2a cells (Fig. 15B f). This suggests that the NoLS activity is regulated by an unknown mechanism (potentially posttranslational modification or interaction with a masking NoLS activity partner) that could use some of the amino acid residues present in the 188–300 area of the TCF4. This result is the start point of further studies I plan to perform.

To test the presence of NLS/NES in the central part of TCF4 (area comprising aa residues 188-549), deletion mutants: YFP-TCF4/188–300, YFP-TCF4/301–367, YFP-TCF4/368-410, YFP-TCF4/411-549 and YFP-TCF4/301-410, YFP-TCF4/368-549 were used. Their ubiquitous distribution in the COS-7 and N2a cells after expression, and lack of positive prediction proved that no active localization signal exists in this area.

Previous expression of bHLH domain located in the C-terminus of TCF4 resulted in detection of fluorescence both in the nucleus and cytoplasm of the HEK293 cells, while the expression of the separated N- and C-terminal fragments of this domain resulted in strictly cytoplasmic localization of both truncations (Sepp et al., 2011). As previously published data did not define the molecular features that affect subcellular distribution of the bHLH domain, we asked the question about the presence of putative NESs and NLSs in fragments containing bHLH using different predictors. According to CDD (conserved domain database), a resource for the annotation of functional units in proteins, homology area of TCF4 bHLH was shown for aa residues 553-634 in contrast to predicted by Smart (568-626 aa) and Prosite (567-620 aa). Due to this, I decided to add additional aa residues to both the N- and C-terminus of the

predicted bHLH to be sure that any structural motif would be not destroyed. I have chosen to refer to the bHLH domain in the presented paper as area 550-670 aa, additionally divided into the N-terminal (550-601 aa) and C-terminal (602-670 aa) parts. Results of the subcellular localization signals predictions suggested the presence of two NESs (one in the N-terminal and one in the C-terminal part of the bHLH) and NLS (569-603 aa) overlapping predicted NESs.

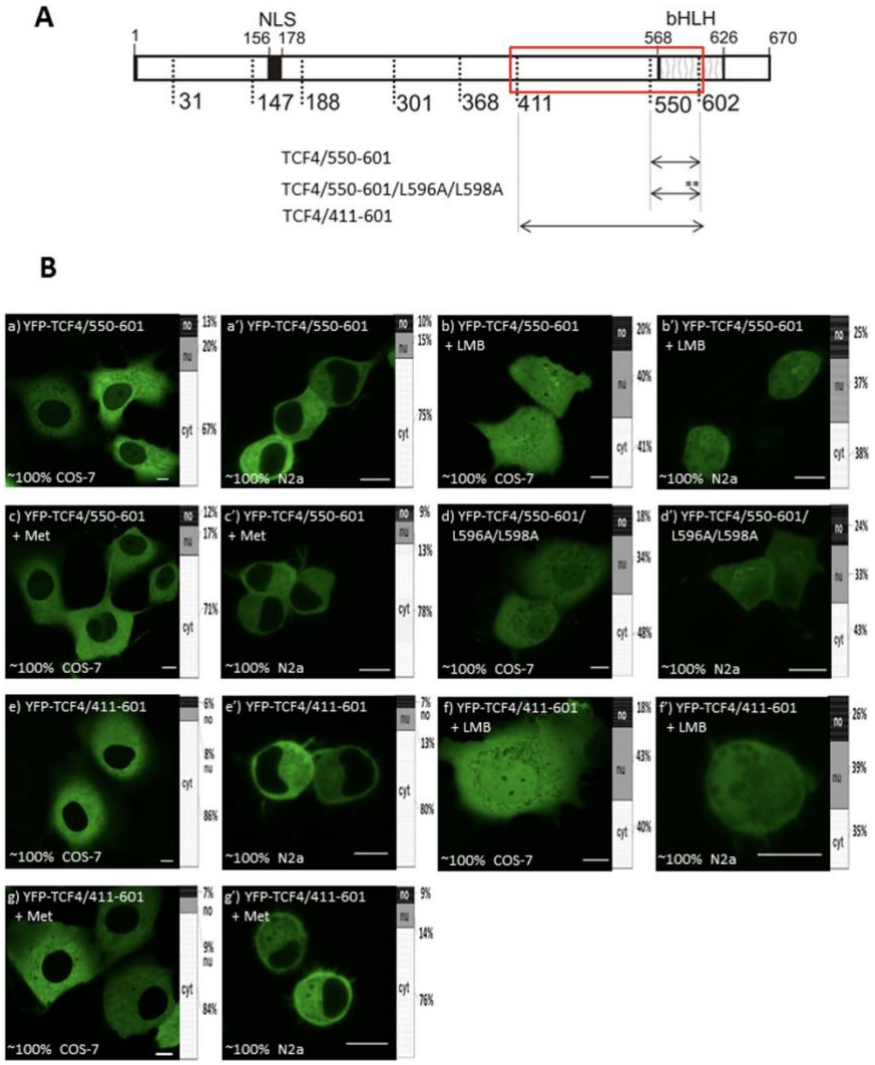


Figure 16. Subcellular distribution of deletion mutants (411-601aa) of TCF4 (A) Schematic representation of TCF4 protein. Expressed deletion mutants of TCF4 are depicted as arrows. Point mutations are depicted as stars. (B) Representative images (single confocal plane for confocal microscopy) of subcellular distribution of the derivatives of TCF4/411-601 area in the absence or presence of additional factors such as LMB or methanol in COS-7 and N2a cells. Bar, 10 μm. Ratios between mean fluorescence intensity in cytoplasmic, nuclear, and nucleolar compartments are presented as an accumulated bar graph (no- nucleolus; nu- nucleus, cyt- cytoplasm). (a,a') YFP-TCF4/550-601, (b,b') YFP-TCF4/550-601 after LMB addition, (c,c') YFP-TCF4/550-601 after methanol addition, (d,d') YFP-TCF4/550-601/L596A/L598A, (e,e') YFP-TCF4/411-601, (f,f ') YFP-TCF4/411-601 after LMB addition, (g,g') YFP-TCF4/411-601 after methanol addition [P4].

To study experimentally the activity of the predicted signals, I prepared the YFP-TCF4/550-601 construct (Fig. 16A) comprising the N-terminal fragment of the bHLH. The

cytoplasmic localization of this fragment both in COS-7 and N2a cells (Fig. 16B a, a') supported the presence of predicted leucine-rich NES. Verification with LMB (Fig. 16B b, b') and methanol (Fig. 16B c, c') substantiated this hypothesis. Based on results of performed prediction, I proposed 589-598 aa as a motif responsible for NES activity in this fragment. Expressed mutant with L/A substitution (YFP-TCF4/550-601/L596A/L598A) was distributed in both the nucleus and cytoplasm of the cell. A slightly higher fluorescence signal in the cytoplasm in the COS-7 and N2a cells (Fig. 16B d, d'), suggested some residual activity of NES. The N-terminal extension of NES containing the 550-601 fragment by neighboring 411-549 aa (YFP-TCF4/411-601, Fig.16A) presented cytoplasmic localization in both the used cell lines (Fig. 16B e, e'), demonstrating an active NES. As expected, the addition of LMB shifted localization to being equally distributed in the nucleus and cytoplasm in the COS-7 and the N2a cells (Fig. 16B f, f'), while Methanol had no influence (Fig. 16B g, g').

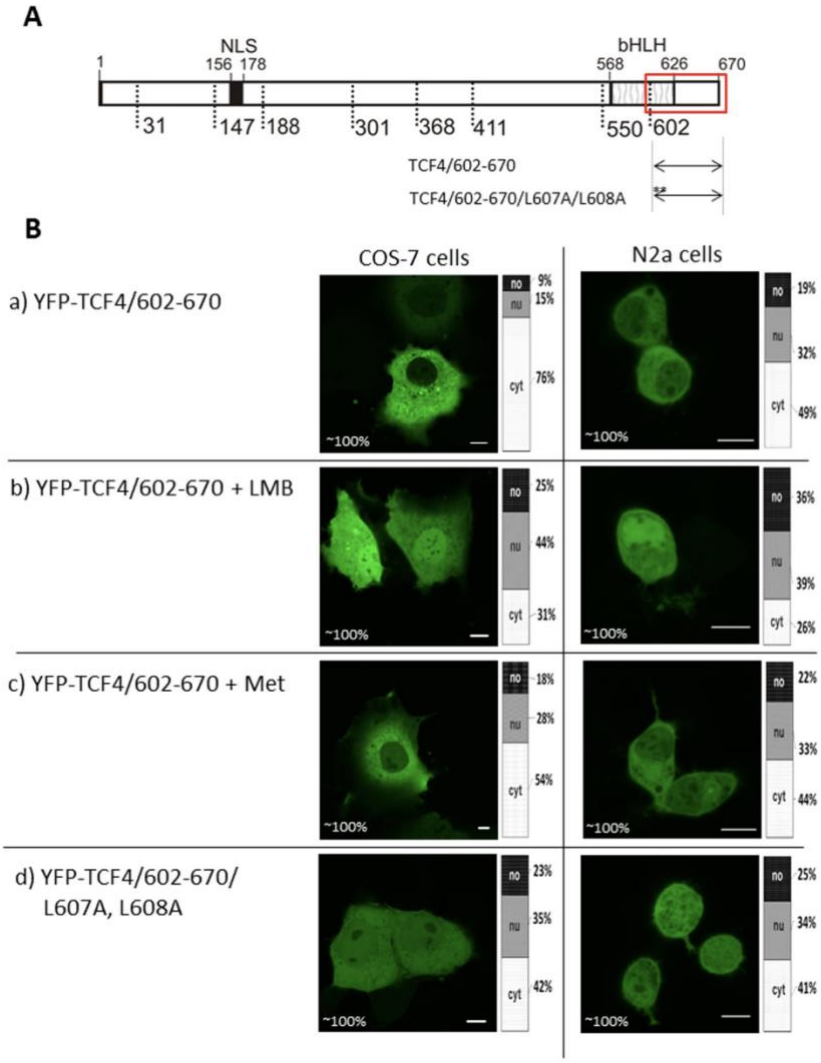


Figure 17. Subcellular distribution of the deletion mutant (602-670aa) of TCF4. (**A**) Schematic representation of TCF4 protein. Expressed deletion mutants of TCF4 are depicted as arrows. Point mutations are depicted as stars. (**B**) Representative images (single confocal plane for confocal microscopy) of subcellular distribution of the

derivatives of TCF4/602-670 area in the absence or presence of additional factors such as LMB or methanol in COS-7 and N2a cells. Bar, 10 μm. Ratios between mean fluorescence intensity in cytoplasmic, nuclear, and nucleolar compartments are presented as an accumulated bar graph (no- nucleolus; nu- nucleus, cyt- cytoplasm). (a) YFP-TCF4/602-670, (b) YFPTCF4/602-670/L607A/L608A (c) YFP-TCF4/602-670 after methanol addition, (d) YFP-TCF4/602-670 after LMB addition [P4].

Then, I verified the activity of the predicted NES in the C-terminal part of the bHLH domain, by expression of prepared YFP-TCF4/602-670 construct (Fig. 17A). The expression of this construct resulted in fully cytoplasmic fluorescence in the COS-7 cells and cytoplasm prevailing in the N2a cells (Fig. 17B a). Experiments with LMB (Fig. 17B b) and Methanol (Fig. 17B c) addition supported the hypothesis about the second NES. After detailed NES prediction analysis, I have chosen the sequence comprising 604-618 aa as a hypothetical NES. To investigate this assumption, a mutant where L607 and L608 were substituted by A (YFP-TCF4/602-670/L607A/L608A) was obtained. Localization of this mutant changed from being exclusively cytoplasmic, observed for the unmodified fragment, to being distributed equally in both compartments of the cell for both COS-7 and N2a cells (Fig. 17B d).

Surprisingly, the results of YFP-TCF4/550-670 expression comprising bHLH domain were not consistent, as the COS-7 cells presented a highly variable pattern of localization from exclusively cytoplasmic, by distributed ubiquitously, to mainly nuclear. Knowing that the N-terminal part of the bHLH (550-601 aa) presented strictly cytoplasmic localization, I deduced that basic amino acid residues K602 (predicted as part of NLS) and K606 (not predicted as part of the NLS, however situated near the predicted NLS) in the C-terminal part of the bHLH are indispensable for this NLS activity. The expression of the point mutant with K606A substitution (YFP-TCF4/55-670/K606A) in the COS-7 cells and N2a cells proved the importance of this amino acid residue for the detected NLS. Detailed explanation of studies is presented in P4.

Importantly, it has been previously shown that point mutations of basic amino acid residues in the bHLH domain of TCF4 were often associated with the Pitt-Hopkins Syndrome. The first identified PTHS-associated heterozygous point mutations were R576W and R576Q (Sepp et al., 2011). Finally, PTHS-association was reported for G358V, D535G, R569W, R572G, R574P, R574H, R576G, R576W, R576Q, R578P, R578H, R580W, R580Q, R582P, R587P, A610V and A614V TCF4 human mutants (Giurgea et al., 2008; Sepp et al., 2012; Whalen et al., 2012; Zweier et al., 2008). All the detected basic residue point mutants are localised in the NLS-2, while A610V and A614V are seen in the NES-2 sequence, thus suggesting the importance of localization signals in PTHS. Localization studies of PTHS-connected point mutants performed previously, did not give consistent results, which was probably due to the different conditions of the performed experiments. Sepp et al. expressed

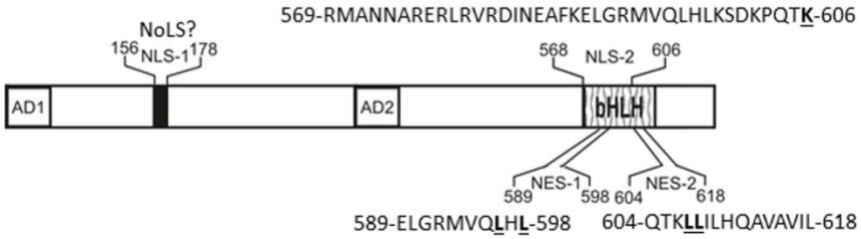
wild type and R576Q, R578H, R580W, R582P and A614 PTHS-associated mutants of TCF4-B^- proteins (isoform lacking RSRS) in HEK 293 cells, getting an analogical nuclear localization for both the wild type (wt) and all the mutants (Sepp et al., 2012). The expression of the bHLH tagged with GFP resulted in ubiquitous localization for wild type and all the tested mutants. On the other hand, Forrest et al. performed the expression of G358V, D535G, R578P, R580W and A614V mutants of the TCF4-B^+ isoform tagged with GFP in the COS-7 cells. Interestingly, while fluorescence for all the expressed proteins was visible only in the nucleus, the mutants R578P, R580W, and to a lesser extent A614V formed a small, spherical punctate, whose localization differed markedly from the distribution of the wild-type TCF4, G538V and D535G mutants. Similar results were obtained in SH- SY5Y neuroblastoma cells (M. Forrest et al., 2012). I hypothesize that point mutations in the bHLH domain of TCF4, known for PTHS patients, influence the pattern of distribution in the nucleus and function of TCF4 probably by changes in DNA binding specificity. However, these single substitutions are not able to inhibit NLS-2 activity. In contrast, our experiments have led to the identification of amino acid residue K606 being indispensable for this NLS activity. Interestingly, this residue has not to date been connected to the PTHS disorder. Importantly, human aa residues located in the DNA binding region: R565, R576, E576, R578, R580, R587, K603 and K607 refer to murine aa residues R564, R575, E576, R577, R579, R586, K602 and K606 (Fig. 18B), which were pointed out by InterPro as being responsible for DNA binding. Localization signals can be structurally exposed, accessible to transport machinery or masked by interacting partners. To visualize the arrangement of the detected NLSs and NESs in the TCF4, we generated a 3D model of the TCF4 using Phyre2 (Fig. 19). We also performed *in silico* analysis of TCF4 structure disorder, and 89% of the TCF4 sequence was predicted to be intrinsically disordered.

IDRs are known to be the targets of intensive posttranslational modifications. As I mentioned previously, modifications such as phosphorylation, especially close to the NLS or NES, have been shown to regulate the intracellular distribution of proteins by activating or deactivating localization motifs (Jans & Hübner, 1996). Interestingly, phosphorylation of S448, located very close to the bHLH domain, by protein kinase A (PKA) was documented by Sepp and co-workers as being necessary for TCF4 transcriptional activity in cultured neurons and in the developing brain in vivo (Sepp et al., 2017). I performed predictions of TCF4 phosphorylation sites and found many putative sites of phosphorylation, especially in the N-terminal part of protein with a probability higher than 0.5, located in the NLS-1 area (156-178

aa). Also, some of the S and T residues of TCF4, predicted as putative phosphorylation sites with a probability higher or equal to 0.8, are in the bHLH domain or near this domain.

As previously mentioned, the ubiquitous family of 14-3-3 proteins is involved in the regulation of signal transduction and binding of 14-3-3 can modulate subcellular localization, structure, stability, and molecular interactions of partner proteins (Obsilova et al., 2014). Performed in P4 prediction of 14-3-3 binding sites in TCf4 pointed motif 542-SITRSRSSNND-552 as putative interaction site. This is substantiated by prediction with a high probability phosphorylation o S548 being a part of the motif. In this aspect, I found very interesting the observation by Liu et al. in U1240MG cells, that TCF4-B^{-} isoform encoding full-length protein without motif 545-RSRS-548, was able to significantly reduce the activity of the promoter containing E-box, while the splice variant TCF4-B^{+} with RSRS sequence did not affect this promoter activity (Liu et al., 1998). Taking all this into account, I hypothesize the importance of 14-3-3 interaction for the regulation of TCF4 localization, and in consequence, also this protein activity. On the other hand, I cannot exclude that a lack of phosphorylated S548 is the reason for the unbalanced and highly differentiated subcellular localization of the expressed TCF4/550-670.

A



B

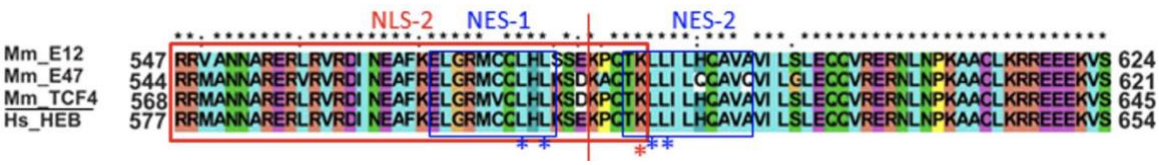


Figure 18. Summary of NLSs and NESs in TCF4. (**A**) Schematic representation of NLSs and NESs residing within the TCF4 protein structure. (**B**) ClustalX sequence alignment of TCF4/568-645 region encompassing NLS-2, NES-1 and NES-2 in the bHLH domain with corresponding sequences of E12 (547-624 aa), E47 (544-621 aa) and HEB (577-654 aa) [P4].

Summary of P4

- I showed that the previously documented NLS, located in the N-terminal part of the TCF4 protein (156-178 aa), also presents the activity of NoLS. Interestingly, additional sequences may inhibit this motif activity.

- Additionally, I identified the activity of two NESs (NES-1 and NES-2) and an overlapping NLS (NLS-2) in the bHLH domain of TCF4 (Fig. 18A). I performed alignment of TCF4-B+ isoforms of mouse and human TCF4, which present a high sequence identity (Fig. 18B), pointing out the conservation of the mosaic pattern of overlapping localization signals with opposing activities between these two organisms. The presence of the NLS and NES in proximity, or overlapping in the bHLH domain, was reported previously for the bHLH-PAS transcription factor aryl hydrocarbon receptor (AhR) (Kawajiri & Fujii-Kuriyama, 2017) and NPAS4 [P3] while two NES motifs were shown in the bHLH of Sima (Drosophila homolog of HIF-1α) (Romero et al., 2008). This suggests a more general function of the bHLH domain in the regulation of bHLH TFs shuttling.

- I performed alignment of the TCF4/568-645 region, encompassing NLS-2, NES-1 and NES-2 in the bHLH domain with corresponding sequences of E12 (aa 547-624), E47 (aa 544-621) and HEB (aa577-654). Sequence homology in these areas is very high (Fig. 18B), and we therefore propose that not only NLS-2, confirmed by Lingbeck et al. (Lingbeck et al., 2003), but also NES-1 and NES-2 activities might be present in all mammalian E proteins. I therefore believe that these signals are important elements that regulate dynamic exchange between the subcellular compartments of not only TCF4, but generally bHLH class I transcription factors.

- Point mutations of basic amino acid residues in the bHLH domain of TCF4 were often associated with the Pitt-Hopkins Syndrome. All the detected basic residue point mutants are localized in the NLS-2, or in the NES-2 sequence, thus suggesting the importance of localization signals in PTHS. I hypothesize that point mutations in the bHLH domain of TCF4, known for PTHS patients, influence the pattern of distribution in the nucleus and function of TCF4 probably by changes in DNA and/or partner proteins binding specificity. However, presented single substitutions are not able to inhibit NLS-2 activity. In contrast, our experiments have led to the identification of amino acid residue K606 being indispensable for this NLS activity. Interestingly, this residue has not to date been connected to the PTHS disorder. My hypothesis requiring verification is that the reason could be the lethality of these mutations.

- I hypothesize that a system that regulates nucleocytoplasmic shuttling of these TF families is very complex and relies on many mutually dependent factors, both exportin-1 dependent and

independent, like signal masking/unmasking by posttranslational modification, and interaction with partner proteins such as 14-3-3 or calmodulin. This enables both precise and flexible signal transduction. Such complex multi-element regulation might also be the reason for the differentiated response to the LMB presence in the COS-7 cells during our experiments: high sensitivity in the case of separated NES-1 and NES-2 in the bHLH of TCF4, and low sensitivity in the case of the whole bHLH domain.

- To visualize the arrangement of the detected NLSs and NESs in the TCF4, I generated a 3D model of the TCF4 using Phyre2 and PyMol (Fig. 19). *In silico* analysis of TCF4 resulted in prediction 89% of the TCF4 sequence as intrinsically disordered. The area with a tendency to possess a more rigid structure and the highest surface accessibility was the bHLH domain. More precise analysis of the selected C-terminal part of the protein (550-670aa) comprising bHLH domain showed that short IDRs could exist in the external parts of this domain. Importantly, the preferential location of NLSs and NESs in IDRs of proteins was found to enable flexible and easily accessible interactions with their binding partners (Dinkel & Sticht, 2007).

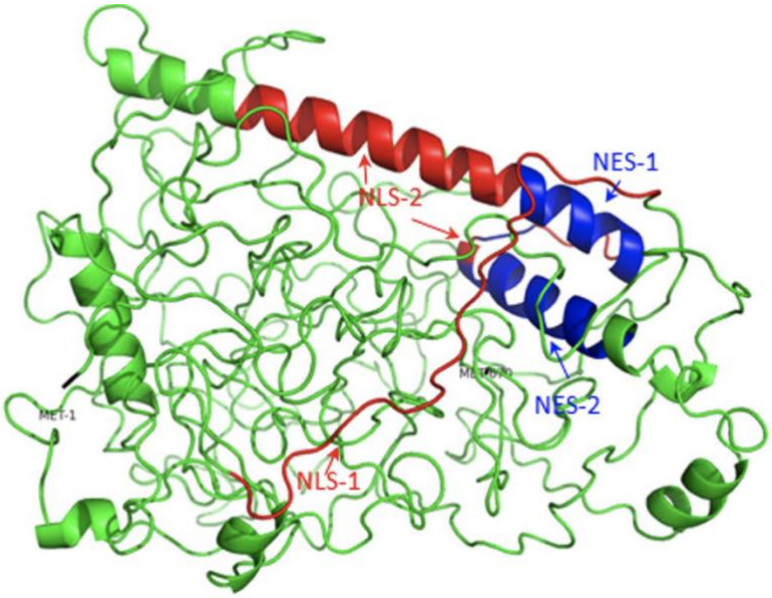


Figure 19. Model of TCF4 generated by Phyre2 and visualized by PyMol. 113 residues (17% of protein) encompassing mainly the bHLH domain (aa 567-620) of TCF4 were modelled with a confidence higher than 90% using the templates: c2mh0A, c5c3lA, c2ql2A and c2ypbB. The remaining 601 residues (89%) of TCF4 lacking any template and predicted as disordered were modelled by ab initio. NLS-1 and NLS-2 are shown in blue. NES-1 and NES-2 are shown in red [P4].

- Location of full-length TCF4 in some of tested COS-7 cells in punctate pattern in the cytoplasm (Fig. 14B) led me to the hypothesis not presented in P3 about TCF4 propensity to LLPS. This is substantiated by published previously observations that in PTHS-linked mutants' localization of PTHS-linked mutants differed from the distribution of the wild-type TCF4 and

fluorescence was visible in the nucleus, in the form of small, spherical punctate. The further studies of structure and LLPS propensity of TCF4 studies become one of my further scientific goals.

- The complex interplay of regulatory mechanisms of the intracellular localization of TCF4 and other I class bHLH transcription factors awaits further studies in vivo.

P5. Greb-Markiewicz B. & Kolonko M. (2019) Subcellular Localization Signals of bHLH-PAS Proteins: Their Significance, Current State of Knowledge and Future Perspectives. Int J Mol Sci 20: 4746.

In previous papers I have presented a detailed experimental study of motifs directing distribution of selected representatives of bHLH-PAS TFs family. First, I selected *D. melanogaster* paralogs MET [P1] and GCE [P2] as JH receptors responsible for insect growth and presenting an interesting and to date not solved question about partial redundance of MET and GCE function. Then, I selected mammalian representatives known as cytoprotective protein NPAS4 with not known understood way of functioning [P3]. In all cases I have determined a complexed pattern of overlapping signals with opposing activity located in bHLH and PAC domains as well as in C-termini. My results were the first, to suggest the presence of multiple localization signals regulating bHLH-PAS TFs shuttling and their complex localization pattern. The results of studies published in my previous papers [P1-P3] and in-depth analysis of the literature on the subject led me to the hypothesis that the knowledge concerning regulation of bHLH-PAS family activity is largely incomplete. To perform their function as sensors, indispensable is receiving the signal in cytoplasm and its transduction, by translocation, to the nucleus. It leads to the activation of gene transcription. Importantly however, some of bHLH-PAS TFs were documented to perform additional nongenomic functions (for example AHR) dependent on the location in the cytoplasm. For this reason, in P5, I decided to present the current state of knowledge with existing gaps, and to draw attention to the importance of more detailed characterization of bHLH-PAS molecular determinants that influence these proteins' shuttling. I asked a question about the general mechanism of bHLH-PAS proteins localization regulation and how precise and systematic were studies reported in the area.

To test the hypothesis, I performed bioinformatic analysis allowing me to predict putative NLSs and NESs. I used cNLS Mapper, NLStradamus, PSORTII, NucPred and SeqNLS servers for NLS predictions, while NetNES, NESFinder0.2 and LocNES for NES predictions.

Additionally, I used The Eukaryotic linear motif resource ELM for both putative NLS and NES searching [P5].

Simultaneously, for all selected proteins I was searching for papers documenting the existence of NLSs/NESs. I have studied human representatives of bHLH/PAS class I: AHR (Aryl hydrocarbon Receptor), HIF-1 (Hypoxia Inducible Factor), HIF-2 and HIF-3, SIM1 (Single-minded homolog) and SIM2, NPAS1-4 (Neuronal PAS domain-containing protein), CLOCK, and representatives of class II: ARNT (Aryl Hydrocarbon Nuclear Translocator), ARNT2, ARNTL1 (Aryl hydrocarbon receptor nuclear translocator-like protein 1, BMAL1, ARNT3), ARNTL2 (BMAL2, ARNT4). Additionally, I analyzed Drosophila class I proteins: MET, GCE, SIMA (Similar) and class II representative: TANGO, being a homolog of mammalian ARNT and CYCLE (homolog of mammalian BMAL1). Then, I visualized results in tables presenting predicted and documented signals. The distribution of NLSs/NESs in the best characterized proteins (MET, GCE, NPAS4 as results of P1, P2 and P4) and AhR (Ikuta et al., 2004b, 2004a; Ikuta, T, Namiki T, 2009; Kawajiri & Ikuta, 2004) was presented graphically (Fig. 20) to show complexity of regulation of bHLH-PAS TFs regulation. I showed that the activity of the most predicted signals was not verified experimentally to date of P5 preparation.

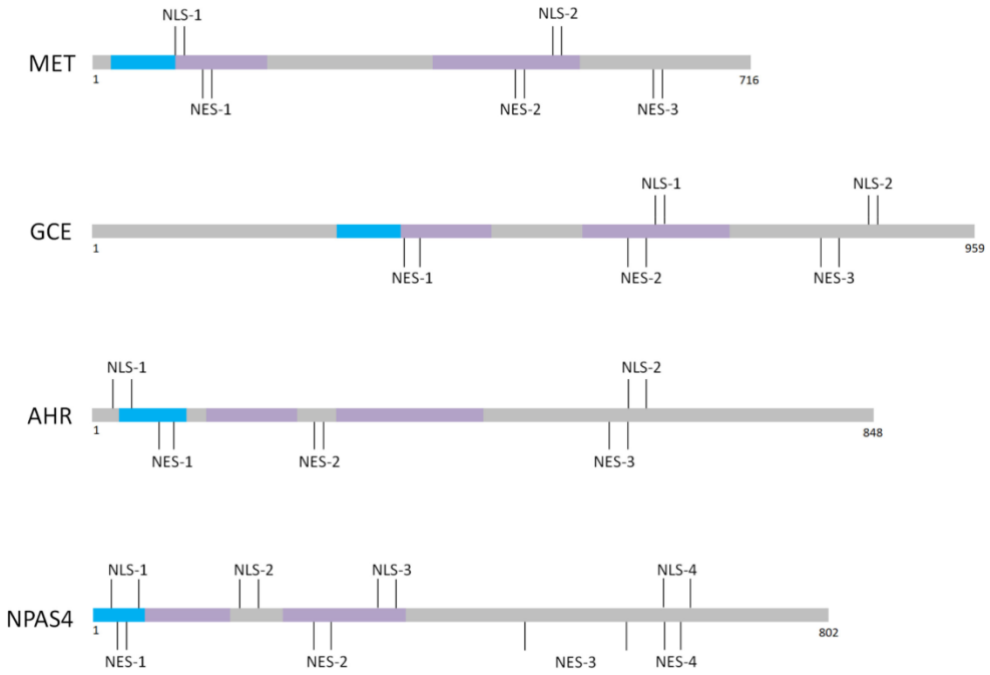


Figure 20. Schematic representation of documented NLSs/NESs distribution in selected bHLH-PAS proteins: MET, GCE, AHR, NPAS4, bHLH domains are shown in blue, PAS domains are shown in violet [P5].

The results of predictions, presented in P5, compared to published studies confirmed that a big set of data concerning the presence of NLSs/NESs regulating translocation of bHLH-PAS TFs as response to stimulus is missing and only its completion will allow for a better understanding of the bHLH TFs functioning. Main goal of P5 was to rekindle interest in the subject, as many of the proteins were studied not systematically leading to single NLS and/or NES sequences often accidentally found over a dozen years ago, that was not analyzed later and verified since then.

d) The importance of the presented scientific achievement

Currently it is known that the nonhomogeneous distribution of cellular components is an important feature of all cells enabling cellular signaling and their inappropriate localization can lead to the development of serious human diseases (Florindo et al., 2016). 1999 year's Nobel Prize in Physiology or Medicine has been awarded to Günter Blobel (the Rockefeller University) for 'the discovery that proteins have intrinsic signals that govern their transport and localisation in the cell' leading to the comments stating that the prize is 'a landmark in its recognition of the fundamental importance of basic research in molecular cell biology for modern medicine and the fight against disease' ("A Nobel Prize for Cell Biology," 1999). That time scientists realized how important for the health of living organisms can be the proper localization of protein related to its activity and possibility to perform expected functions. Improper localization of proteins can result in diseases that range from metabolic disorders to cancer. Later, the possibility to use the knowledge about nucleocytoplasmic transport to deliver drugs was proposed (Davis et al., 2007). However, understanding the ways of complicated and balanced regulation is indispensable to obtain predictable successful results. Analyzing literature of the subject, I realized that years 1999-2006 exploded with papers documenting usually single sequences of NLS or NES sequences in studied protein, however no detailed and systematic research of the whole proteins to recognize exact patterns of signals distribution was performed generally and specifically in the case of bHLH-PAS family. Most papers from the recent decade focus on the multi-protein complex interactions, leaving individual proteins at the stage of characterization carried out earlier in the context of localization signals. Being aware of this, I decided to conduct a thorough characterization of proteins from the bHLH family, with particular emphasis on the bHLH-PAS class, which contains proteins of particularly high, documented importance for the functioning and health of eukaryotic organisms, including humans, and indicated the need to verify the results for members of this

family. For this reason, I believe that presented papers constitute a significant contribution to understanding the molecular basis of the functioning of representants of important bHLH family, which is confirmed by including in the FlyBase for Met (P1) and GCE (P2), and the Uniprot database for NPAS4 (P3) and TCF4 (P4).

4.2 The second scientific achievement submitted for evaluation

a) Title of the scientific achievement

Molecular characterization of intrinsically disordered regions and their importance for the functioning of proteins from the family of transcription factors bHLH-PAS.

b) Publications documenting the presented scientific achievement

P6. Kolonko, M., Ożga, K., Hołubowicz, R., Taube, M., Kozak, M., Ożyhar, A., & **Greb-Markiewicz, B**. (2016). Intrinsic disorder of the C-terminal domain of drosophila methoprene-tolerant protein. *PLoS ONE, 11*(9), e0162950.

P7. Kolonko, M., Bystranowska, D., Taube, M., Kozak, M., Bostock, M., Popowicz, G., Ożyhar, A., & **Greb-Markiewicz, B**. (2020). The intrinsically disordered region of GCE protein adopts a more fixed structure by interacting with the LBD of the nuclear receptor FTZ-F1. *Cell Communication and Signaling, 18*(1), 180.

P8. Kolonko, M., & **Greb-Markiewicz, B**. (2019). bHLH–PAS Proteins: Their Structure and Intrinsic Disorder. *International Journal of Molecular Sciences*, *20*(15), 3653.

P9. Kolonko-Adamska, M., Uversky, V. N., & **Greb-Markiewicz, B**. (2021). The Participation of the Intrinsically Disordered Regions of the bHLH-PAS Transcription Factors in Disease Development. *International Journal of Molecular Sciences*, *22*(6), 2868.

Description of my individual contribution to each of the above-mentioned publications is included in the **Appendix 4B.**

Copies of papers documenting the scientific achievement are included in **Appendixes 7.1-7.4**

c) Scientometric parameters of the presented scientific achievement

Scientometric parameters of the publications documenting the first scientific achievement are presented in the Table 2.

Table. 2 Parameters on 15.08.2023

Publication	Impact factor IF*	Points MEiN**	All citations Web of Science	Citations excluding selfcitations Web of Science	All citations Scopus	Citations excluding selfcitations Scopus
P6	2,8	35	6	4	6	4
P7	5,716	140	5	5	6	5
P8	4,556	140	15	13	16	14
P9	6,208	140	1	1	2	2
Total	**19,28**	**445**	**27**	**23**	**30**	**25**

*IF according to Journal Citation Report from the year of publication.
** Points were calculated according to the list of scientific journals of the Ministry of Education and Science (MEiN, previously MNiSW) from the year of publication.

d) Discussion of the scientific achievement

P6. Kolonko, M., Ożga, K., Hołubowicz, R., Taube, M., Kozak, M., Ożyhar, A., & Greb-Markiewicz, B. (2016). Intrinsic disorder of the C-terminal domain of drosophila methoprene-tolerant protein. *PLoS ONE, 11*(9), e0162950.

P7. Kolonko, M., Bystranowska, D., Taube, M., Kozak, M., Bostock, M., Popowicz, G., Ożyhar, A., & Greb-Markiewicz, B. (2020). The intrinsically disordered region of GCE protein adopts a more fixed structure by interacting with the LBD of the nuclear receptor FTZ-F1. *Cell Communication and Signaling, 18*(1), 180.

After completing the internship, I started to work at the Department of Biochemistry headed by Prof. A. Ożyhar, who allowed me to choose the objective of my future research. At that time, contrary to the well-understood mode of functioning of the ecdysone receptor, the JH receptor was unknown. While there were earlier suggestions for the USP as the JH receptor, these have not been supported experimentally. It was also not known how the signal transduction pathways of 20E and JH, on which proper insect development depends, are interconnected. During my internship in Ulm, I had the opportunity to learn about publications of Prof. T. Wilson, who discovered *D. melanogaster* MET and proposed it as the JH receptor (Wilson & Fabian, 1986). The subsequent discovery of the MET paralog - GCE, which was shown to be partially redundant to MET, and non-existing in most insect species (Baumann, Fujiwara, et al., 2010), was highly enigmatic. I was interested in studying these proteins and finding the reason for their functional differentiation. The characterization of molecular determinants responsible for subcellular trafficking of MET and GCE was described previously

as part of scientific achievements (P1-P2) in section 4c. I have found that both proteins differed in NLSs and NESs pattern, especially in the C-terminal fragments (MET 509-716 aa and GCE 391-689 aa), predicted as IDRs. In addition, C-termini of MET and GCE presented the highest sequence differentiation. As continuation of the presented as scientific achievements research, I decided to perform the structural characterization of these proteins. Importantly, at that point of time no C-terminal region of the bHLH-PAS proteins was structurally studied. Though it was known that C-termini often contains transactivation domains (TADs) responsible for PPIs and regulation of protein activity. For this purpose, in the years 2015-2020 I continued the subject, taking care of Mrs. Marta Kolonko-Adamska as an assistant promoter. The main supervisor of the PhD thesis was Prof. A. Ożyhar, and the topic of the work was 'Molecular analysis of C-terminal fragments of receptors: Methoprene tolerant and Germ cell-expressed from Drosophila melanogaster'. In the years 2018-2022 I was also the contractor of the "Preludium" grant obtained by M. Kolonko-Adamska.

In collaboration with Prof. M. Kozak (Department of Macromolecular Physics, Faculty of Physics, Adam Mickiewicz University, Poznan and National Synchrotron Radiation Centre SOLARIS, Jagiellonian University) and Prof. G. Popowicz (Biomolecular NMR and Center for Integrated Protein Science Munich at Department Chemistry, Technical University of Munich, and Institute of Structural Biology, Helmholtz Zentrum München), we showed experimentally by CD, SAXS analysis and NMR studies that according to *in silico* analysis C-terminal fragments of MET (METC) and GCE (GCEC) are intrinsically disordered regions (IDRs). Intrinsic disorder is characterized by the highly extended, ellipsoidal conformation and the lack of stable tertiary structure of protein/region. We showed that GCEC, with unstructured-like (U-like) IDR characteristics, shows a slightly higher degree of asymmetry than the premolten globule-like (PMG-like) METC, characterized by the presence of some stable, defined dimensional structure. Simultaneously, GCEC, contains more molecular recognition elements (MoREs), and in consequence can probably interact with more physiological partners supporting my hypothesis presented in P2 about a more complex manner of GCEC function regulation. Importantly, the conformational changes within the bHLH and PAS domains, induced by ligand binding or interaction, may trigger structural changes in GCEC, but to a lesser extent than in the case of METC.

The partial redundancy of MET and GCE functioning was supported by Dubrovski et al. study, showing that nuclear receptor Fushi Tarazu factor-1 (FTZ-F1) interacts both with MET and GCE, but GCE interaction is JH-independent (Dubrovsky et al., 2011). Ligand

binding domain (LBD) FTZ-F1 encompassing AF2 motif was shown to interact with novel type of NR box (LIXXLL motif) located in C-termini of MET and GCE. We aimed to test if intrinsically disordered GCEC alone can create an interaction with the LBD of the FTZ-F1. Results of verification was positive. Importantly, multidimensional NMR analysis using peptide GCE^{PEP} encompassing the novel NR-box (LIQNL) led us to define specific residues directly responsible for the interactions of GCEC and LBD Ftz-F1. To confirm the results and to determine the effect of GCEC and LBD FTZ-F1 interactions on the subcellular localization of these proteins, we performed pull-down assays in COS-7 cells. We observed, that GCEC can induce LBD FTZ- F1 transition towards the nucleus and modulate its activity independent on the JH presence. We hypothesize that GCEC is slightly separated from the bHLH and PAS domains and can act more independently in comparison to METC. The established interactions can directly affect the subcellular localization, stability and structure of the full-length proteins and thus modulate their function and activity and assure the functional differentiation of MET and GCE. The described analysis contributes to a better understanding of the molecular basis of the functions of the C-terminal fragments of the bHLH-PAS family since GCE and MET are the first described hormone receptors in this transcription factor family. Our study might be helpful in explaining the conformation and mode of action and the regulation of IDRs. Results of these research were published [P6] [P7]. I was the last and corresponding author.

P8. Kolonko, M., & Greb-Markiewicz, B. (2019). bHLH–PAS Proteins: Their Structure and Intrinsic Disorder. *International Journal of Molecular Sciences*, *20*(15), 3653.

I performed comprehensive literature searching of structural characterization of bHLH-PAS TFs and deduced that our studies were unique. The available structure characterization of bHLH–PAS proteins is limited to the relatively well-conserved domains bHLH, PAS-A, and PAS-B. C-termini, comprising an extensive part of proteins and fulfiling important roles, have not yet been structurally characterized. Due to the highly variable amino acid sequence and the lack of predefined domains, C-termini are believed to be responsible for the specific modulation of the functioning of bHLH–PAS proteins and the recognition of partner proteins necessary for their unique action. Continuing collaboration with M. Kolonko-Adamska, we discussed the known functions of the presented C-termini proteins according to their disorder character. We pointed that the structural properties of subsequent IDRs predicted in the sequences of bHLH–PAS transcription factors (mainly C-termini) need to be resolved for a full understanding of the way of bHLH–PAS family transcription factors function. We presented our observations and conclusions in paper [P8]. As the author of the concept of work, and the person performing

bioinformatic analysis of protein structures along with taking an active part in writing the first version of the manuscript and its correction, I am the last and corresponding author.

P9. Kolonko-Adamska, M., Uversky, V. N., & Greb-Markiewicz, B. (2021). The Participation of the Intrinsically Disordered Regions of the bHLH-PAS Transcription Factors in Disease Development. *International Journal of Molecular Sciences, 22*(6), 2868.

The aim of the P9 work was to lay the foundations for future experimental studies devoted to the analysis of the effects of mutations affecting the functionality of bHLH-PAS proteins. As IDRs are suggested as the most important regulatory regions for proteins, I was interested in finding out if there is a pattern of the distribution of disease associated missense mutations among ordered and disordered regions in bHLH-PAS protein family members. Are the missense mutations observed more frequently in IDRs prone to PTMs, LLPS or aggregation? To address this problem, I collaborated with Dr. M. Kolonko Adamska and Prof. V. Uversky (Department of Molecular Medicine, USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa,). We decided to analyze the known aa missense mutations listed in the HuVarBase database and to compare their localizations with the localizations of documented PTMs (PhosphoSitePlus database) and predicted MoRFs (Anchor server), simultaneously with the *in silico* analysis of protein's LLPS (catGranule and PScore servers) and amylogenic propensity (Waltz predictor). Based on the results, we assume that most of the disease-associated missense mutations are localized in IDRs of analyzed and selected bHLH-PAS family representatives [P9]. I believe that showing such a relationship is important for directing a new research pathway using low-molecular LLPS regulators as therapeutics in cases of diseases currently difficult to treat.

4.3 Discussion of other scientific and research achievements

4.3.1 Before obtaining PhD degree

After obtaining an MSc title under the supervision of Prof. A. Zabża, I was offered the opportunity to continue my scientific research concerning entomopathogenic fungi as a research assistant. At that time, the first of my goals was to elaborate conditions of an efficient culture of selected strains of entomopathogenic fungi, with particular emphasis on the Entomophthorales family, which are the obligatory pathogens that cause difficulties in their cultivation in the laboratory using artificial media. Unlike species such as *Beauveria* or *Metarhizium*, the Entomophthorales family are highly specialized, sometimes even up to one species of insect. For this reason, this family could be used as a safe and effective way to combat selected species of harmful insects. However, for this purpose, it is necessary not only to

develop their effective cultivation but also to understand the mode of action of these fungi and to know the metabolites they produce. For this reason, after choosing conditions and media for the culture of selected fungi, my next step was searching for active substances, presenting biological activity towards insects (for example causing death), secreted by these fungi during growth. For this purpose, properly prepared samples of post-culture medium and mycelium extracts were introduced by injection into the larvae of *Galleria mellonella*, as a model insect commonly used for this purpose, and in intervals of 24, 48, 72, 96 hours, and one week after injection, I controlled the number of individuals dead or phenotypically different from the control. Later, I made attempts to isolate by extraction of the compounds from the post-culture medium or mycelium from the day when the highest activity against the larvae was observed, and purification using HPLC system equipped in normal-phase or RP18 columns. In collaboration with Prof. Zbigniew Szewczuk (Wroclaw University), I conducted MS of selected samples. Unfortunately, due to the very small amounts of compounds secreted and hardware limitations at that time it was only possible to confirm the presence of a cyclic dipeptide, without knowing its exact chemical structure. Interestingly, the secretion of the peptide by fungi took place only during small-scale culturing. In addition, working for several years, I observed a recurring period of effective fungus growth on selected media (between May and December each year) and a practically complete lack of growth (between January and early May) regardless of the medium used. Considering at least 3 weeks of fungal growth in a single culture such behavior created a significant problem.

In the case of the potential application of fungi as insecticides, consideration should also be given to the possibility of interaction between fungi and chemicals used for crop protection. On the one hand, fungi could lead to the decomposition or transformation of compounds into effective/harmful forms, however also the negative impact of chemicals on the activity and viability of fungi cannot be excluded. There are reports of a noticeable increase in the number of pests in the following years after the use of chemical insecticides. On one side, this may be due to the problem of resistance that occurs in insects to a chemical agent, on the other side, however, this may be due to the loss of natural enemies of insects (like fungi) in chemical agent action. For this reason, an additional direction of my research was to determine the activity of oxidoreductases from selected entomopathogenic fungi. For this purpose, I performed biotransformation processes with particular emphasis on the campholenic ketones. As substrates were used: racemic methyl α-campholenyl ketone (**I**), (+)(-)-campholene nitryl(**7**) and both enantiomers of 4-(2',2',3'-trimethyl-cyclopent-3'-en-l'-yl)-buty-2-enoate (+)-(4 R)-

(**11**) and (-)-(4 S)-(**12**) (Fig. 1 and 2). The transformation was carried out for 10 days, and then the products were extracted with chloroform. The extracts were dried over anhydrous MgSO, and, after solvent evaporation, subjected to flash chromatography followed by HPLC for compound separation. Finally, CD spectra, GC /MS, and NMR were used for the identification of purified products. Results confirmed earlier observations that fungi of the *Zoophthora* genus present strong oxidative properties. The main products of biotransformations were optically active α,β-unsaturated ketones **4** and **10**, and a racemic mixture of epoxides **2**,**2'**, **3**,**3'** from (+-)-l, and **8**, **8**, **9**, **9'** from (+-)-**7**) (Fig. 21 and 22). Smaller amounts of two alcohols, keto-alcohol **6** and epoxy-alcohol **5** were also isolated. The structure of compound 6 was proved by its MS spectrum. The formation of epoxy-alcohol 5 was especially interesting because it was a single example of the reduction of the carbonyl group by fungi of this genus. The structure of compound **5** was confirmed by NMR spectra. Interestingly, both enantiomers of unsaturated esters **11** and **12** were not transformed. After 10 days of biotransformation almost theoretical amounts of unreacted substrates were recovered.

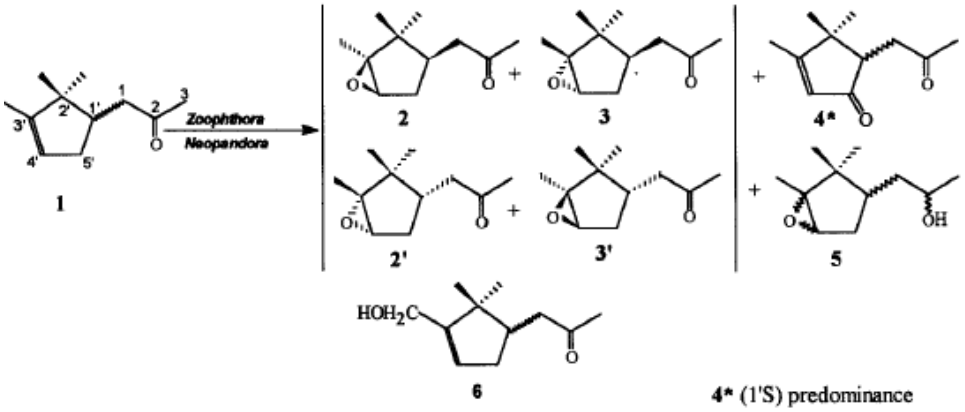


Figure 21. Transformation of methyl-a-campholenyl ketone (**1**) by *Zoophthora (Neopandora) phyllobii* [A**1**].

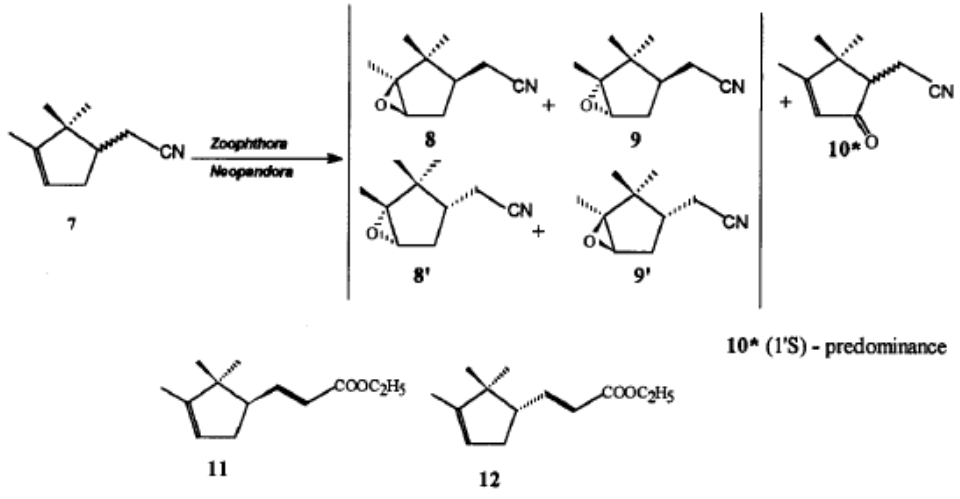


Figure 22. Transformation of a-campholenyl nitrile (**7**) by *Zoophthoru (Neopandora) phyllobii* [**A1**].

[A1] Zabża, A., **Greb-Markiewicz, B.,** Sodano, G., & De Riccardis, F. (1995). Microbial transformation of isoprenoid systems by means of fungi of Zoophthora genus — Microbial transformation of compounds with α-campholenic system by means of fungusZoophthora (Neopandora) sp. *Journal of Basic Microbiology*, *35*(6), 433–439. (IF=3.611, MNiSW$_{1995}$ 0.441, no citations)

I am also the author of a research project results presented in the Ecological Chemistry Laboratory's SPR reports in 1994, 1996, 1998, 1999.

4.3.2 Presentation of scientific achievements after obtaining the doctoral degree (excluding the achievements presented in the sections 4.1-4.2)

a) Years 1999-2002

After obtaining my PhD degree, I was employed at the Laboratory of Ecological Chemistry at the Faculty of Chemistry at the Wroclaw University of Technology as an assistant professor. At that time, the head of the laboratory, Prof. dr hab. A. Zabża, had changed the field of interest to medicinal plants. Therefore, in addition to an entomopathogenic fungi project [A**2**], I had to study the literature and learn new techniques related to the isolation and identification of metabolites from plants. The main subject of my research was *Euonymus europaea*. I was able to isolate and identify the unknown for this plant cardenolide. I presented the preliminary results of analyses during the 2nd Symposium of the Center for Biomonitoring, Biotechnology and Ecosystem Protection of Lower Silesia (II Sympozjum Centrum Biomonitoringu Biotechnologii i Ochrony Ekosystemów Dolnego Śląska) [A4]. I also collaborated with Prof. W. Peczyńska-Czoch (Wroclaw University of Technology) in the project testing conditions of laccases immobilization. Extracellular laccases produced by wood-rotting fungi: *Cerrena unicolor*, *Heterobasidion annosum* and *Trametes versicolor*, were immobilized differemnt carriers: DEAE-Granocel 500, CM-Granocel 500, and acrylic carriers. We showed that only DEAE-Granocel 500 carrier activated by divinyl sulphone, was a suitable matrix for the preserving enzyme activity. Among different tested laccases, only laccase produced by *C. unicolor* showed the best binding to the carrier and a satisfactory enzymatic activity. I performed chromatographic separation of samples obtained by Prof. W. Peczyńska-Czoch team using the HPLC system. The results of this collaboration were presented in the paper [**A5**].

At that time I was co-author of a research project results presented in the Ecological Chemistry Laboratory's SPR report in 2000.

I was a supervisor of two master's students and was taking care of master's students of Prof. A. Zabża. I also supervised a PhD student of Prof. A. Zabża, who performed studies concerning the dependence of *Sedum acre* metabolites content on the place of the plant growth. Preliminary results were presented during the 2nd Symposium of the Centre for Biomonitoring,

Biotechnology and Ecosystem Protection of Lower Silesia [A**6**]. These studies, however, were not finished due to foreign internships of both laboratory members and the closing down the Laboratory because of Prof. A. Zabża's retirement.

[A2] Zabża A., **Greb-Markiewicz B**. (2000). Entomopatogeniczne grzyby z rodzaju *Zoophthora* i *Paecilomyces*. *Biotechnologia 3*, 47-57.

[A3] Greb-Markiewicz B. (2001). Grzyby entomopatogeniczne jako bioinsektycydy. I Sympozjum Centrum Biomonitoringu, Biotechnologii i Ochrony Ekosystemów Dolnego Śląska, Wroclaw, 4-5 April 2001. Oficyna Wydawnicza PWr., 77-81.

[A4] Greb-Markiewicz B., Cierpicki T., Kasper P., Zabża A. (2002). Badania frakcji kardenolidowej z rośliny *Euonymus europeae*. II Sympozjum Centrum Biomonitoringu, Biotechnologii i Ochrony Ekosystemów Dolnego Śląska, Wroclaw, 22 May 2002 Oficyna Wydawnicza PWr, 111-117.

[A5] Al-Adhami J.-H., Bryjak J., **Greb-Markiewicz B**., W. Peczyńska Czoch (2002). Immobilization of wood-rotting fungi laccases on modified cellulose and acrylic carriers. *Process Biochemistry 37*(12), 1387-1394. (IF=1.143, no MNiSW points, citations74/74 Scopus, 63/63 Web of Science)

[A6] Kasper P., **Greb-Markiewicz B**., Zabża A. (2001). Badania porównawcze składu metabolitów izolowanych z roślin leczniczych Dolnego Śląska. I Sympozjum Centrum Biomonitoringu, Biotechnologii i Ochrony Ekosystemów Dolnego Śląska, Wroclaw, 4-5 April 2001. Oficyna Wydawnicza PWr., 111-114.

b) Years 2002-2004 scientific internship abroad

After obtaining a PhD in chemical sciences, I was looking for an internship enabling me to perform research that would allow me to better understand the functioning of insects that were targets of previously studied entomopathogenic fungi simultaneously with learning methodology used in biochemistry and molecular biology. For my internship, I have chosen the Institute of Zoology and Endocrinology of Ulm University in Germany, headed by Prof. K. D. Spindler. I was accepted after giving a lecture (INCHECO 2nd Conference, 2002, Calambrone, Pisa, Italy) and a successful interview for the '*Wissenschaftliche angestellte*' position. There, under the supervision of Prof. M. Spindler-Barth I started working on EcR and USP, proteins which heterodimer create *D. melanogaster* 20E receptor. It allowed me to switch to the use of protein biochemical characterization and cloning methods, which was an important step in my research career. Insects are good model organisms to study the effects of hormones on body development because, unlike mammals, they only use 2 hormones: 20-hydroxyecdysone (20E) and juvenile hormone (JH). Their relative amounts determine whether the insect remains at a given stage of development or undergoes the transition to the next stage. During the two-year foreign internship, when I got position of "Wissenschaftliche Angestellte" at the University of Ulm, Germany, I performed the expression in the yeast *Saccharomyces cerevisiae* system and

the purification of the ligand binding domains (LBDs) of the nuclear receptors EcR and USP, as well as examining the effect of selected residues mutation and the presence of a hormone on the oligomerization of these domains using gel filtration, SDS-PAGE, and western blot. The obtained results were published in the work in which I am the first author [A7]. I conducted all experiments myself, apart from studying the interaction of a protein with a hormone analogue. Also, I prepared all figures and participated in the writing and editing of the manuscript. In addition, I was involved in writing a review article presenting the current state of knowledge regarding the ecdysteroid receptor complex at the time of paper preparation [A8]. The additional methodology I got known included: nuclear extracts preparation, *in vitro* transcription assays, insect cells (Sf9) culture, ultracentrifugation, work with radioactivity.

During my internship I also had teaching and organizational responsibilities. It should be emphasized that at that time, on my initiative, the above Institute of Zoology and Endocrinology established a research collaboration with the Department of Biochemistry at the Wroclaw University of Technology, headed by Prof. dr hab. A. Ożyhar. This resulted in several joint publications of both units.

[A7] Greb-Markiewicz B., Fauth T. & Spindler-Barth M. (2005) Ligand binding is without effect on complex formation of the ligand binding domain of the ecdysone receptor (EcR). *Arch Insect Biochem Physiol 59*: 1–11. (IF=1.827, $MNiSW_{2010}$ 32, citations 2/2 Scopus, 2/2 Web of Science)

[A8] Spindler K.-D., **Greb-Markiewicz B.**, Polifke T., Spindler-Barth M. (2004) The functional ecdysteroid receptor complex. *Current Topics in Steroid Research 4*: 123-129. (no IF, LM)

c) Years 2004-to date

I. Charakterystyka molekularna regionów inherentnie nieuporządkowanych i ich znaczenie dla funkcjonowania białek z rodziny czynników transkrypcyjnych bHLH

As a continuation of previous studies, in cooperation with Dr. M. Kolonko-Adamska, we performed interaction analysis to studying more detailed the functional differences between Met and Gce by performing interaction analysis. For this, we performed the interaction analysis of MetC, GceC and peptides corresponding to the novel NR-boxes localized in MetC and GceC (Met^{PEP} and Gce^{PEP} respectively) with Ftz-F1 LBD (based on the Dubrovski et al study (17). We have also tested MetC and GceC interactions with 14-3-3 (the hypothesis presented in P2). It is worth to note, that the overexpression in bacterial cells and purification of the full-length JH receptors (similarly to other bHLH-PAS TFs) were not efficient, so we decided to use C-terminal fragments, considered as responsible for interactions with both Ftz-F1 LBD and 14-3-

3. During the nuclear magnetic resonance (NMR) studies, two sets of experiments were performed: MetC and GceC 15 N-HSQC spectra were obtained before and after incubation with partners, and Ftz-F1 LBD 15 N-HSQC spectra were measured before and after incubation with MetC/GceC or peptides Met^{PEP} and Gce^{PEP}. The dissociation constant (Kd) of Met and Ftz-F1 LBD was determined by both NMR and Isothermal titration calorimetry (ITC). Finally, a pull-down assay combined with fluorescence imaging was used to analyze the full-length Met and Gce interactions and to determine fluctuations in their subcellular localization forced by these interactions. Obtained results were consistent and clearly indicated a distinct interaction pattern of both Drosophila JH receptors. We were able to show, that in contrast to Met, Gce interacts with Ftz-F1 in JH independent manner because its C-terminus (GceC) comprising NR-box is easily accessible for partner protein. In the case of Met, binding of the hormone enables presentation of the C-terminally located NR-box and makes possible the interaction with Ftz-F1. As hypothesized previously, we showed that only Gce interacts with 14-3-3 protein, which affects its subcellular localization and activity. All this supports the hypothesis concerning functional differentiation of Met and Gce as Drosophila JH receptors. What more, we have observed different patterns of structural changes when GceC was incubated with Ftz-F1 LBD or 14-3-3. We assume that determined specific interactions can force MetC and GceC to adopt a more fixed specific structure that depends just on the interaction partner and determines specific receptor activity. Presented research is the first trial to explain why Drosophila Gce and Met are not fully redundant and how their roles could differ. Currently the manuscript presenting these studies is under preparation [A9].

[A9] Kolonko-Adamska M., Zawadzka-Kazimierczuk A., Bartosińska Marzec P., Koźmiński W., Popowicz G., Krężel A., Ożyhar A., **Greb-Markiewicz B**. (2023). Interaction patterns of Drosophila JH receptors Methoprene tolerant and Germ-cell expressed suggest significant differences in their functioning. *Frontiers in Molecular Biosciences*, 10:1215550. doi: 10.3389/fmolb.2023.1215550 (IF=5.0, $MEiN_{2023}$ 140, no citations, paper published 15th August 2023)

I would like to emphasize that we are the only group in Poland involved in research of GCE and MET proteins, which was reflected in the invitation in 2022 to write an article on the hormonal regulation of insects for the special issue of the journal *KOSMOS „*Insects in scientific research and practice" (pol. *Owady w badaniach naukowych i praktyce*). I invited Dr. M. Orłowski and Dr. M. Kolonko-Adamska to cooperate, which resulted in the review paper [A10].

[A10] Kolonko-Adamska M., Orłowski M., **Greb-Markiewicz B.** (2022). Rola kluczowych hormonów i ich receptorów w rozwoju owadów na przykładzie *Drosophila melanogaster* i *Aedes aegypti*. *Kosmos, 72* (3), 225-242. (no IF; $MEiN_{2021}$ list 20; no citation)

My interest in the link between bHLH transcrtion factors' structure and function, in accordance with recent discovery that IDRs could be engaged in liquid-liquid phase separation (LLPS) of proteins led me to the hypothesis, that this process can be especially important for proteins involved in transcription regulation. I invited Dr. A. Tarczewska to collaborate as an expert in the phase separation field (Wroclaw University of Science and Technology). I performed *in silico* analysis of selected proteins from the bHLH family simultaneously with comprehensive literature studies looking for experimental verification of my thesis. In the published work, we drew attention to the existence of large gaps in the knowledge of the time and discussed the directions of research that should be performed [A11].

[A11] Tarczewska, A., & **Greb-Markiewicz, B.** (2019). The Significance of the Intrinsically Disordered Regions for the Functions of the bHLH Transcription Factors. *International Journal of Molecular Sciences*, *20*(21), 5306. (IF=4,556; $MEiN_{2019}$ 140, citations 23/22 Scopus, 22/21 Web of Science)

II. The ability of viral proteins to undergo liquid-liquid phase separation

When the global Covid pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) began, scientists around the world were seeking the best possible solutions to stop it. As the university was closed, I was wondering how I can help to gain knowledge of the workings of Sars-Cov2 despite not having the opportunity to work in a laboratory with viruses. I performed a detailed search of literature on the subject and joined a Crowdfight group gathering volunteer scientists from around the world. In this way, I could send a request and obtained the cDNAs of some of the virus proteins. The SARS-CoV-2 genome encodes nearly 30 proteins, which are responsible for host cell penetration, viral genome replication, viral gene transcription, and many other functions (I. Chen et al., 2020). Important structural proteins necessary for assembling a complete viral particle and genomic RNA packing are the spike (S) protein, the membrane (M) protein, the envelope (E) protein and the nucleocapsid (N) protein. Using my knowledge concerning IDPs, I conducted *in silico* analysis of the N protein, which indicated the propensity of this protein for phase separation. I proposed that LLPS could be involved in the viral infection progress. Unfortunately, due to the time lapse between the hypothesis and the receipt of the cDNA, another laboratory (cDNA donor) published a paper that proved the ability of the N protein to LLPS before we could.

Nonetheless, our attempts at expression and purification of N protein undertaken in collaboration with Dr. A. Tarczewska and Dr. M. Kolonko-Adamska resulted in significant new observations. Considering the basic function of the protein, which is the binding of genomic RNA, we investigated whether and to what extent typical N protein isolation procedures enable a fully defined, homogeneous preparation to be obtained without contamination of nucleic acids from host cells. We performed a reconstruction of the procedures described in previously published papers (S. Lu et al., 2020; Perdikari et al., 2020; Ye et al., 2020; Zeng et al., 2020; Zinzula et al., 2020), which resulted in protein samples that were highly contaminated with nucleic acids from the host cells used for overexpression. We believed that using contaminated samples for the study could have a significant impact on the results. Therefore, we developed a novel protocol for purifying the N protein and showed that the obtained preparation is free of nucleic acid contamination. Next, we performed comparative structural analysis of the N protein samples obtained with a previously published protocol and our new protocol. We demonstrated that the N protein completely separated from nucleic acids can only assemble into well-defined dimers and not previously reported high oligomers (tetramers or octamers). Also, previously, the N protein was reported to form liquid condensates in the presence of different types of RNA (S. Lu et al., 2020). We did not observe such an effect under the conditions used in our study. However, we did observe the formation of liquid condensates formed by the N sample contaminated with nucleic acids at low ionic strength. It was shown that the N protein can bind a fragment of viral genomic RNA encoding the N protein (Carlson et al., 2020; Iserman, 2020).Therefore, it is possible that the N protein expressed in bacterial cells might be complexed with specific mRNA encoding N, rather than random prokaryotic sequences. In addition, we made important observations that N protein can form spontaneous more ordered condensates like gel-like or solid. Notably, SARS-CoV and SARS-CoV-2 are the only known coronaviruses with N proteins that do not possess cysteine residues, which are believed to be responsible for enhancing the stabilization of other coronavirus virions by forming disulfide bonds (Surjit & Lal, 2008). Therefore, we hypothesized that the more organized structures observed in our analysis might facilitate the compact character of RNP and stabilize the SARS-CoV and SARS-CoV-2 protein structures. Interestingly, in previous studies Naskalska *et al.* suggested a new way of viral entry into the host cell by interaction of M protein with heparan sulphate proteoglycans (Naskalska et al., 2019). However, the studies were performed in the presence of N protein. Considering that the N protein tightly interacts with M (He et al., 2004; S. Lu et al., 2021), we suggested that additional experiments should be performed to elucidate the role of the N protein in the proposed alternative pathway of cell penetration. As the

recruitment of host proteins into condensates might be an important factor in altering host cell metabolism, establishing a method to enable the selective dissolution of SARS-CoV-2 liquid condensates or the prevention of their formation may facilitate efforts to identify a means of combating viral infection. Results of this research were published [A12].

[A12] Tarczewska A., Kolonko-Adamska M., Zarębski M., Dobrucki J., Ożyhar A. & **Greb-Markiewicz B.** (2021) The method utilized to purify the SARS-CoV-2 N protein can affect its molecular properties. *Int J Biol Macromol* 188: 391–403. (IF=8,025, $MEiN_{2021}$ 100, citations 5/5 Scopus, 4/4, Web of Science).

When I continued my interest in the possibility to fight viral infections by preventing/influencing phase separation, I found by international Crowdfight group partners presenting different areas of scientific interests, presenting different countries. Also, I invited my younger colleague Dr. M. Kolonko-Adamska to the project. In collaboration we prepared a detailed review that asked the question whether the current state of the art in LLPS involvement in the development of viral diseases could allow us to use or discover new drugs preventing or treating infections. The work was published in a prestigious journal Trend in Microbiology. It should be mentioned that the drawing was selected on the cover of December Issue [A13].

[A13] Li H., Ernst C, Kolonko-Adamska M., **Greb-Markiewicz B**., Man J., Parissi V. & Ng B.W. (2022) Phase separation in viral infections. *Trends Microbiol. 30*(12):1217-1231. (IF=18.230, $MEiN_{2021}$ list 200, citations 11/11 Scopus, 9/9 Web of Science)

III. Entomopathogenic fungi as agents improving the growth of plants

Though switching of main scientific interest after obtaining PhD, still I believed that more detailed studies of entomopathogenic fungi resulting in better understanding mechanisms of their functioning would be beneficial for sustainable cultivation. As a supervisor of the BioTop student group, I introduced the students of the circle into the area of entomopathogenic fungi. I began cooperation with Prof. Cezary Tkaczuk (University of Natural Sciences and Humanities in Siedlce), obtaining strains of fungi. Also, I presented to the group of students how to work with these fungi. Recently, the ability of entomopathogenic fungi to act as endophytes and to improve nutrient uptake and plant growth was presented. The application of inorganic fertilizers in the last century highly increased agricultural production, however, it has also led to deterioration of the environment. For this reason, development of new alternatives in the form of naturally obtained, environmentally friendly and cost-effective biofertilizers such as bacteria or fungi with the capacity to add nutrients to the soil, fixing nitrogen and ensuring

solubilization of minerals to be used by the plant is advantageous. Considering the antagonistic activity of entomopathogenic fungi against insects and plant diseases in addition to their beneficial effect on nutrient uptake, it would be highly beneficial to develop a method for their use in sustainable agricultural production. In cooperation with students the effect of B. bassiana water extracts on the growth of wheat seedlings were tested. Selected dilutions of the crude extract obtained under different pH and temperature conditions were used to establish the optimal method of extraction. Plant length, total fresh weight, and chlorophyll composition, in addition to micro- and macro-elements and toxic metals, were evaluated. Additionally, the antibacterial activity of extracts was tested to exclude their negative impact on the beneficial soil microorganisms. Results were published in paper I am the last and corresponding author [A14]. It is important to mention that our study was the first trial to evaluate the suitability of entomopathogenic fungi water extracts as factors improving the growth of agriculturally important plants.

[A14] Kramski, D. J., Nowinski, D., Kowalczuk, K., Kruszyński, P., Radzimska, J., & **Greb-Markiewicz, B.** (2023). Beauveria bassiana Water Extracts' Effect on the Growth of Wheat. *Plants*, *12*(2), 326. (IF 04.658, MEiN$_{2021}$ list 70, no citation)

4.4 Research plans

While carrying out research related to the project presented as my main scientific achievements, I developed my knowledge about the bHLH family of transcription factors. Importantly, many of the results obtained were the starting points for further questions. I am generally interested in characterizing bHLH family members linked to human disease and would like to characterize this family more detailed. I am particularly interested in determining the basis of the role of TCF4 mutations in PTHS developments and NPAS4 as multifunctional protein and potential diagnostic/therapeutic target.

TCF4 protein

As mentioned in P4, TCF4 was shown as a protein associated with neuronal system development and linked to Pitt-Hopkins Syndrome (PTHS). I showed that associated with PTHS point mutations of basic amino acid residues in the bHLH domain of TCF4 are in area of the NLS-2 and/or the NES-2 sequence, suggesting the putative involvement of localization signals mutations in PTHS. The system that regulates nucleocytoplasmic shuttling of TCF4 seems to rely on many mutually dependent factors. In P4 I proposed that localization signal masking/unmasking is dependent on intrinsic disorder, posttranslational modification, and

interaction with partner proteins to find an answer how TF4 functioning is regulated, I decided to continue TCF4 molecular characterization in cooperation with Msc Nicola Sozańska (PhD student, supervisor Prof. dr hab. A. Ożyhar) and Aneta Tarczewska, PhD (assistant supervisor). We will especially test TCF4 disorder and propensity to LLPS for wild type TCF4 different isoforms and mutants linked to PTHS. In addition, I plan to verify the hypothesized that 14-3-3 could enable precise but also flexible signal transduction by TCF4.

NPAS4 protein

Despite the documented role of NPAS4 as the protective factor, the mechanism of its functioning is not understood. To date, no detailed structural or biochemical studies of NPAS4 were performed, though new manuscripts supporting the importance of this protein continuously appear. I believe that very important for therapeutic use of NPAS4 is the understanding of its cytoprotecting mechanisms. My main goal is to provide the missing knowledge and determine biochemical properties of NPAS4. To realize it I am trying to get funding. Also, I am establishing cooperation with students realizing diploma thesis and my younger college Marta Kolonko-Adamska. I want to verify experimentally thesis presented below.

Thesis 1. NPAS4 as IDP with LLPS propensity

Previously performed bioinformatic analysis led me to hypothesis that multifunctionality of NPAS4 depends on the intrinsically disordered character of its structure and the propensity to LLPS (Fig. 23). I believe that the conformational plasticity of IDRs, its extreme sensitivity to changes in the environment, its ability to interact with multiple partners, and consequently to fold in different ways could explain the multifunctionality of NPAS4 described in the literature. Recently, the ability to form liquid-liquid phase separation (LLPS) was proposed as important for neuronal development and synaptic plasticity. Mutations in areas responsible for LLPS were shown to lead to pathological aggregation and disease, like autism or cancer, also linked to NPAS4. Importantly, to date disordered character and LLPS propensity were not studied for NPAS4 nor other members of the bHLH-PAS family. The LLPS propensity of NPAS4 could have important implications as phase-separated condensates were proposed very recently as novel drug targets in a wide spectrum of diseases. We believe, that intrinsically disordered structure and LLPS propensity of the C-terminal part of NPAS4 is important for its function. Also, the presence of predicted as disordered with high propensity to LLPS, PR/PG rich sequence located between PAS domains, seems to be particularly interesting. Glycine and Proline with Arginine were proposed to serve as anti-aggregation agents, while arginine-rich

peptides were proposed as potent neuroprotective agents, with demonstrated efficacy in vitro and in animal models of stroke. I hypothesize, that NPAS4 functioning by organization of phase separated droplets could be regulated by RNA- binding. This theory is supported by results of our initial experiments, in which initialization of LLPS by synthesized PR/PG rich peptide occurred after RNA addition (Fig. 24). It is worth to note, that PR/PG rich sequence of NPAS4 differs significantly from any studied and described previously and can be found only in mammalian homologues of NPAS4.

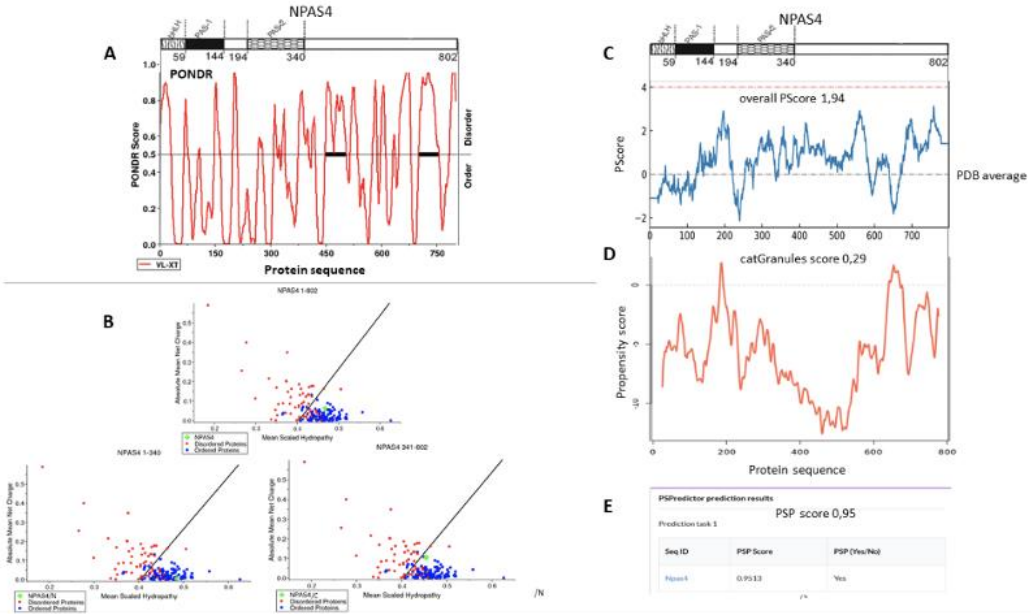


Figure 23. NPAS4 *in silico* analysis **(A)** The top panel represents the domain structure of NPAS4. The bottom panel presents a prediction of disorder regions based on the amino acid sequence of the NPAS4 protein. Calculations were performed using PONDR-VLXT, a score above 0.5 indicates a high probability of disorder. **(B)** Charge-hydropathy plot. The Uversky plot compares the absolute, mean net charge and the mean hydropathy of disordered (red circles) and ordered proteins (blue circles). The boundary between ordered and disordered proteins is shown. The green diamond corresponds to NPAS4, NPAS4/1-340 and NPAS4/341-802 relatively. **(C)** Prediction of NPAS4 propensity of LLPS formation performed with PScore. **(D)** Prediction of NPAS4 propensity of LLPS formation performed with catGRANULE. **(E)** Prediction of propensity of LLPS formation performed with PSPredictor.

Thesis 2. PAS domains of NPAS4 bind heme

NPAS4 structure, similarly to other bHLH-PAS proteins, comprises PAS domains responsible for ligand binding or signal sensing, however the way of signal receiving is not known. I performed in silico analysis using 3DLigandSite Web server, which resulted in indication of heme and FMN as putative ligands, which bind to both PAS domains of NPAS4. As the structure of NPAS4 is unknown, I used Phyre2 to generate a molecular model of NPAS4 to predict the surface exposure of sequences interesting for us. 3D structure prediction with heme/FMN putative locations was visualized with The PyMOL Molecular Graphics System (PyMOL) (Fig. 25). This hypothesis is supported by studies documenting that member of

bHLH-PAS family taking part in circadian rhythm regulation: NPAS2 and CLOCK, bind heme as a prosthetic group enabling these proteins to sense oxydo-redox potential. Interestingly, NPAS4 participation in circadian rhythm regulation was also proposed.

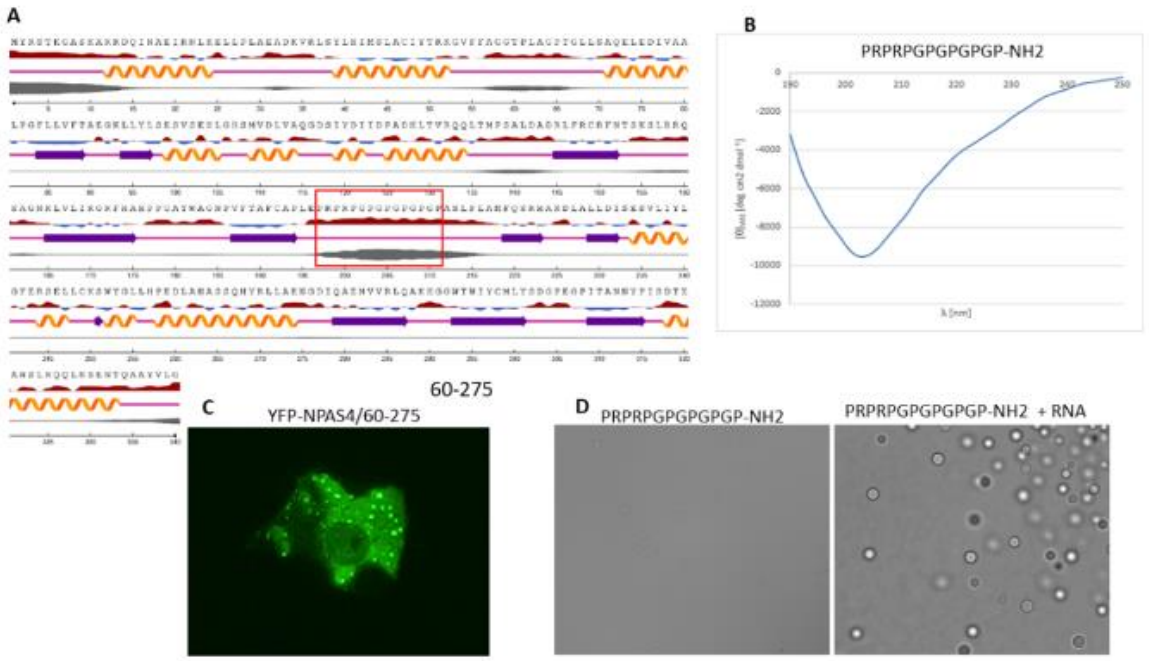


Figure 24. Wstępna analiza sekwencji bogatej w powtórzenia PR/PG. **(A)** NetSurfP-2.0 prediction of NPAS4/1-340 secondary structure. Red box indicates PR/PG rich sequence. **(B)** The far-UV CD spectrum of synthesized PR/PG rich peptide indicating disordered character of peptide. **(C)** Representative image of YFP-NPAS4/1-340 distribution in COS-7 cells. **(D)** The representative DIC images of condensates formed by PR/PG rich peptide in the absence (left) and presence (right) of RNA.

Interestingly, heme is one of cofactors needed for induction of necrosis, while NPAS4 was shown to prevent necrosis during ischemia. Binding of heme predicted as putative ligand of NPAS4 PAS domains could result in decreasing free heme level in the cell. The activation of NPAS4 in the case of ischemia1 could supports this hypothesis and explain its mechanism of action. On the other hand, in the case of hypoxia sensor HIF-1α, Fe(II) atom connected to hypoxia-inducible factor prolyl hydroxylases is used instead of heme. To understand the basis how NPAS4 receives a low oxygen signal in the cell, I would like to verify experimentally the hypothesis about heme binding by the PAS domains. Importantly, heme-responsive sensors were shown to be involved in different important pathways. Interestingly, for human NPAS4 second isoform NPAS4 comprising aa 1-234 (only bHLH and PAS-1 domains; substitution V234 by G) was proposed (Fig.25 CD). However, to date there is no proof at the protein translation level and not known function. I believe that it would be important to shed light on the role of this putative isoform.

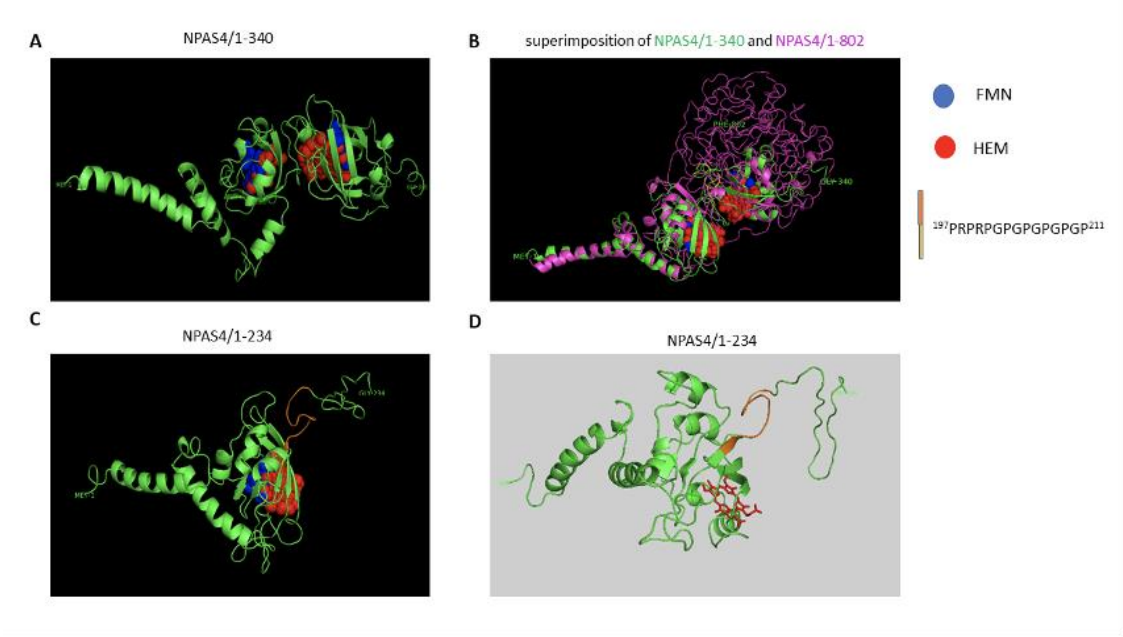


Figure 25. Schematic representation of the NPAS4 protein structure presented in a 3D model generated by Phyre2. **(A)** NPAS4/1–340 region encompassing the bHLH and PAS domains was modelled with confidence higher than 90% using the templates c4f31B, c4zprB, C4f31A, c4zp4D, c5sy5B and c5sy7B is presented in green. **(B)** Whole length NPAS4 (1-802aa) 3D model generated by Phyre (magenta) superimposed with NPAS4/1-340aa 3D model (green). The C-terminal region of NPAS4 lacking any template and predicted as disordered was modelled ab initio (67% of the protein is predicted as disordered). **(C)** NPAS4/1–234 region encompassing short isoform of NPAS4 (bHLH and PAS-1). The unique sequence rich in PG and PR repeats (197-PRPRPGPGPGPGPGP-211), is shown in yellow/orange. Blue and red balls represent potential localizations of predicted ligands, respectively FMN and heme. Ligand binding sites were predicted using the 3DLigandSite Web server (http://www.sbg.bio.ic.ac.uk/3dligandsite). **(D)** Npas4/1-234 isoform with one heme molecule position (selected from many possible). 3D structure prediction was visualized with PyMOL.

Thesis 3. NPAS4 interacts with calmodulin and 4-3-3 proteins

The testing of NPAS4 interaction with selected partner proteins and membranes is important for understanding the basics of NPAS4 functioning as a modulator of the cell death pathway. NPAS4 expression protects brain in the case of focal and generalized ischemic strokes by preventing necrosis and leading to cell apoptosis. Induction of necrosis depends on the presence of NO synthesized by Nitric oxide synthase (NOS), while NOS activation requires binding of Calmodulin (CaM)/Ca^{2+}. CaM family is a major class of calcium sensor proteins which play a crucial role in cellular signaling cascades through the regulation of numerous target proteins. Competitive binding of CaM/Ca^{2+} by highly expressed NPAS4 could explain its protective role in brain in the case of focal and generalized ischemic strokes. My in-silico analysis of NPAS4 sequence, resulted in prediction of multiple CaM binding sites.

As mentioned previously, the proteins from 14-3-3 family are highly conserved in all eukaryotic cells, and involved in the regulation of many different processes, like cell cycle, nervous system development and neurodegeneration. Importantly, dynamic 14-3-3/disordered

client protein interactions were shown to integrate survival and apoptotic pathways. Apoptosis regulation is of particular interest in the context of proposed NPAS4 protective effect by direction of cells after ischemia into apoptosis instead necrosis. I hypothesize that NPAS4 interaction with 14-3-3 modify its activity. Importantly, predicted calmodulin and 14-3-3 interaction sites overlap in the linker between PAS domains, suggesting mutually exclusive binding of these proteins.

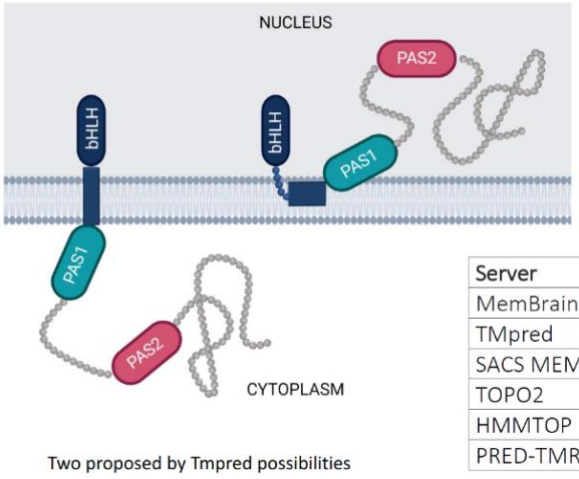
Thesis 4. Stress factors influencing NPAS4 behavior

I also plan to investigate the influence of the cellular stressors on the NPAS4 behavior in the cell. NPAS4 was shown to be one of the immediate early genes (IEGs) that activate mechanisms related to the first defense against many cellular stresses. Importantly, IEGs are regulated by a specific stimulus and can be activated without de novo protein synthesis. In P4 we showed for NPAS4 a complicated pattern partially overlapping molecular motifs responsible for the subcellular localization, which activation is dependent on glucose level and other putative factors such as ligand binding or interaction with partner proteins. Importantly, cytoplasmic localization was linked to nongenomic function of other bHLH-PAS TFs: AHR modulating immune response and BMAL1 functioning as a translation factor that enhances protein synthesis during the metabolically active nighttime period. This may confirm, that NPAS4 in addition to the genomic role as a TF also participates in the rapid non-genomic signal transduction, which could be beneficial for NPAS4 function as cell protector. Therefore, we will investigate the influence of the cellular stressor (e.g., Glucose level, low oxygen, DNA damage) on the NPAS4 behavior in mammalian cells. Glucose level impact may be significant in the context of NPAS4 connection to diabetes. During ischemia, the tissue is extremely deprived of oxygen and ATP. If so, the protein interaction studies also in hypoxia conditions are necessary.

Heme was shown to be an important cofactor for oxygen transfer, oxygen storage, however in addition to these prototypical heme proteins, there are emergent, critical roles of exchangeable/labile heme in signal transduction. Heme-responsive sensors were shown to be involved in DNA binding, transcriptional regulation, tRNA synthesis, microRNA splicing and protein synthesis. DNA repair after damage is important for organisms not only during stress conditions, but also for learning process. I believe in NPAS4 role in these processes. Based on presented results, we plan to perform experimental verification of the impact of mentioned

factors on NPAS4 propensity to localization, LLPS, heme binding and interaction with CaM/14-3-3 proteins.

The other aspect of NPAS4 behavior in stress, which I want to test is interaction with the membrane. Shamloo et al. (2006), confirmed translocation hypothesis by immunohistochemical studies demonstrating an accumulation of Npas4 signal around the cell membrane during ischemic reperfusion. The authors excluded the possibility of new protein synthesis. They suggested an active translocation of NPAS4 from the cytosol, as a simultaneous and proportional decrease in Npas4 signal in the cytosolic fraction was observed. To date no further studies of NPAS4 membrane localization were performed, performed prediction supports the hypothesis (Fig. 26). In cooperation with Prof. N. D'Amelio (Université de Picardie Jules Verne CNRS in Amiens), who has extensive experience in the characterization of the structure and dynamics of biomolecules involved in cellular processes. We will test interaction of selected NPAS4 fragments with biomimetic membranes by liquid state NMR of bicelles. Also, molecular dynamics simulations between NPAS4 fragments and nuclear phospholipids will be performed. Based on NPAS4 literature searching and my *in silico* analysis I hypothesize that the cytoprotective role of NPAS4 is dependent on the interaction with membrane. Cooperation with Prof. D'Amelio group that can create membrane models and to study experimentally the interaction of selected synthesized NPAS4 fragments with biomimetic membranes by NMR, is crucial for documentation and mapping sites responsible for NPAS4-membrane interaction which would be of high importance for understanding the way of NPAS4 functioning as cytoprotectant.



Server	Result [aa]
MemBrain 3.1	41-52
TMpred	45-68
SACS MEMSAT	53-71
TOPO2	53–71
HMMTOP	54-71
PRED-TMR	53-71

Figure 26. Most of used predictors suggested the location of transmembrane helix in the linker between bHLH AND PAS domain of NPAS4.

In cooperation with M. Zarębski, PhD from Prof. J. Dobrucki group (Jagiellonian University) we performed preliminary the light-induced DNA damage experiments in mammalian cells expressing YFP-tagged full-length NPAS4. After blue light exposition, depending on specific conditions of experiment, we could observe differentiated results: NPAS4 translocation from nucleus to cytoplasm (Fig. 27A), localized condensation of NPAS4 in target site of exposure (Fig. 27B) or nuclear stress bodies formation interestingly around nuclear membrane (Fig. 27C). Importantly, Prof. Dobrucki group has developed the methods allowing for DNA damage induction, and visualization.

I am convinced that performing of planned research may allow for a more complete understanding of NPAS4 functioning, which is crucial for the possibility of using this cytoprotective protein as a potential therapeutic factor in the previously presented diseases harmful for the human society.

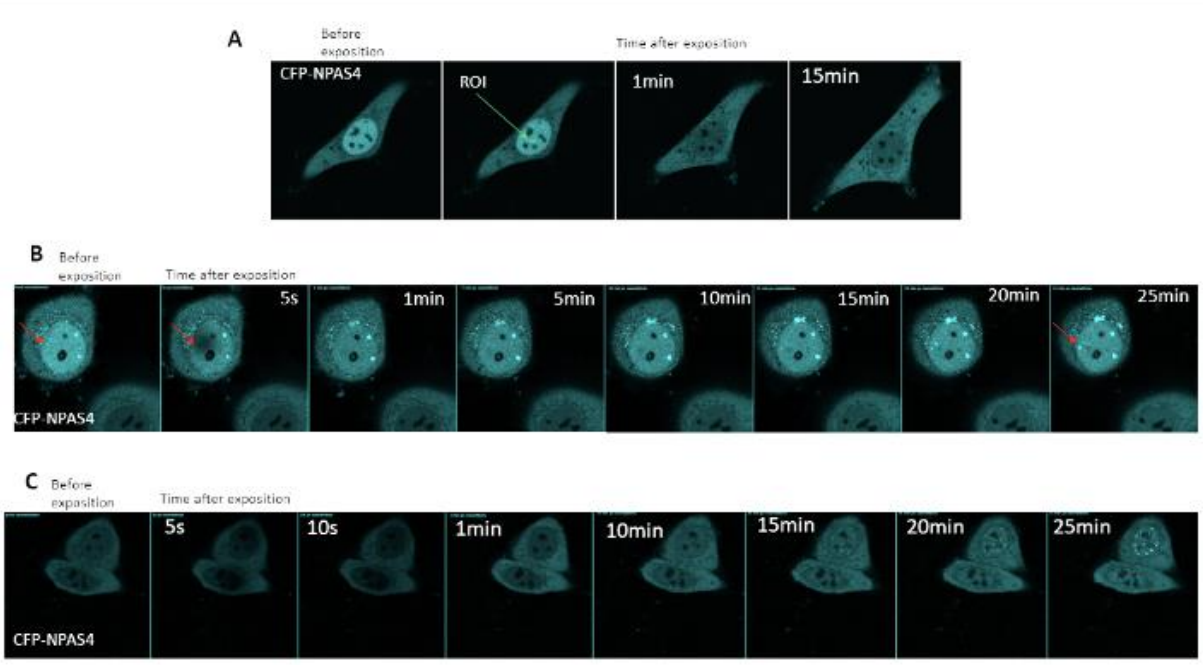


Figure 27. Preliminary studies of NPAS4 distribution in HeLa cells after DNA damage inflicted by blue light **(A)** NPAS4 translocation from nucleus to cytoplasm **(B)** localized condensation of NPAS4 in exposure target site **(C)** formation of nuclear stress bodies. Confocal microscopy was performed with a Leica SP5 microscope.

4.5 General scientometric parameters of scientific achievements

The total number of MEiN points is 1744 (including 204 to 2018 and 1570 from 2019), summary IF is 86,672 (status of August 15, 2023). Points are given according to the list of scientific journals of the Ministry of Education and Science (MEiN, previously the Ministry of

Science and Higher Education, MNiSW) from the year of publication. IF is given according to the Journal Citation Report for the year of publication.

According to the Web of Science, the number of all citations is 181, the number of citations without self-citations is 162 and the Hirsch index is 7 (status of August 15, 2023).
According to the Scopus, the number of all citations is 203, the number of citations without self-citations is 182 and the Hirsch index is 7 (status of August 15, 2023).
According to the Google Scholar the number of all citations is 297 (the number of citations without self-citations is not known) and the Hirsch index is 9 (status of August 15, 2023).

The number of points and the total IF and number of citations before obtaining the doctoral degree was 0.

The list of my total scientific achievements is presented in **Appendix 4B**.

4.6 References

A Nobel Prize for cell biology. (1999). In *Nature cell biology* (Vol. 1, Issue 7, pp. E169–E169). Nature Publishing Group.

Abdou, M. A., He, Q., Wen, D., Zyaan, O., Wang, J., Xu, J., Baumann, A. A., Joseph, J., Wilson, T. G., Li, S., & Wang, J. (2011). Drosophila Met and Gce are partially redundant in transducing juvenile hormone action. *Insect Biochemistry and Molecular Biology*, *41*(12), 938–945.

Amiel, J., Rio, M., Pontual, L. de, Redon, R., Malan, V., Boddaert, N., Plouin, P., Carter, N. P., Lyonnet, S., Munnich, A., & Colleaux, L. (2007). Mutations in TCF4, Encoding a Class I Basic Helix-Loop-Helix Transcription Factor, Are Responsible for Pitt-Hopkins Syndrome, a Severe Epileptic Encephalopathy Associated with Autonomic Dysfunction. *The American Journal of Human Genetics*, *80*(5), 988–993.

Ashok, M., Turner, C., & Wilson, T. G. (1998). Insect juvenile hormone resistance gene homology with the bHLH-PAS family of transcriptional regulators. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(6), 2761–2766.

Baumann, A., Barry, J., Wang, S., Fujiwara, Y., & Wilson, T. G. (2010). Paralogous genes involved in juvenile hormone action in Drosophila melanogaster. *Genetics*, *185*(4), 1327–1336.

Baumann, A., Fujiwara, Y., & Wilson, T. G. (2010). Evolutionary divergence of the paralogs Methoprene tolerant (Met) and germ cell expressed (gce) within the genus Drosophila. *Journal of Insect Physiology*, *56*(10), 1445–1455.

Beck, M., & Hurt, E. (2017). The nuclear pore complex: Understanding its function through structural insight. *Nature Reviews Molecular Cell Biology*, *18*(2), 73–89.

Beckstead, R. B., Lam, G., & Thummel, C. S. (2007). Specific transcriptional responses to juvenile hormone and ecdysone in Drosophila. *Insect Biochemistry and Molecular Biology*, *37*(6), 570–578.

Bernardo, T. J., & Dubrovsky, E. B. (2012). Molecular Mechanisms of Transcription Activation by Juvenile Hormone: A Critical Role for bHLH-PAS and Nuclear Receptor Proteins. *Insects*, *3*(4), 324–338.

Bitra, K., & Palli, S. R. (2009). Interaction of proteins involved in ecdysone and juvenile hormone signal transduction. *Archives of Insect Biochemistry and Physiology*, *70*(2), 90–105.

Blake, D. J., Forrest, M., Chapman, R. M., Tinsley, C. L., O'Donovan, M. C., & Owen, M. J. (2010). TCF4, schizophrenia, and Pitt-Hopkins Syndrome. *Schizophrenia Bulletin*, *36*(3), 443–447.

Blom, N., Gammeltoft, S., & Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology*, *294*(5), 1351–1362. https://doi.org/10.1006/jmbi.1999.3310

Brandl, L., Horst, D., de Toni, E., Kirchner, T., Herbst, A., & Kolligs, F. T. (2015). ITF-2B protein levels are correlated with favorable prognosis in patients with colorectal carcinomas. *American Journal of Cancer Research*, *5*(7), 2241–2248.

Brockschmidt, A., Todt, U., Ryu, S., Hoischen, A., Landwehr, C., Birnbaum, S., Frenck, W., Radlwimmer, B., Lichter, P., Engels, H., Driever, W., Kubisch, C., & Weber, R. G. (2007). Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Human Molecular Genetics*, *16*(12), 1488–1494.

Bustos, D. M. (2012). The role of protein disorder in the 14-3-3 interaction network. *Molecular BioSystems*, *8*(1), 178.

Carlson, C., Asfaha, J., Ghent, C., Howard, C., Hartooni, N., Safari, M., Frankel, A. D., & Morgan, D. (2020). Phosphoregulation of Phase Separation by the SARS-CoV-2 N Protein Suggests a Biophysical Basis for its Dual Functions. *Molecular Cell*, *80*, 1092–1103.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)*, *263*(5148), 802–805.

Charles, J.-P., Iwema, T., Epa, V. C., Takaki, K., Rynes, J., & Jindra, M. (2011). Ligand-binding properties of a juvenile hormone receptor, Methoprene-tolerant. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(52), 21128–21133.

Chen, I., Yuann, J. P., Chang, Y., Lin, S., Zhao, J., Perlman, S., Shen, Y., Huang, T., & Hou, M. (2020). *Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID- 19. The COVID-19 resource centre is hosted on Elsevier Connect, the company' s public news and information. January*.

Chen, Y., Xie, X., Wu, A., Wang, L., Hu, Y., Zhang, H., & Li, Y. (2018). A synthetic cell-penetrating peptide derived from nuclear localization signal of EPS8 exerts anticancer activity against acute myeloid leukemia. *Journal of Experimental and Clinical Cancer Research*, *37*(1), 1–15.

Cheng, J. H., Lai, G. H., Lien, Y. Y., Sun, F. C., Hsu, S. L., Chuang, P. C., & Lee, M. S. (2019). Identification of nuclear localization signal and nuclear export signal of VP1 from the chicken anemia virus and effects on VP2 shuttling in cells. *Virology Journal*, *16*(1), 1–9.

Choy, F. C., Klarić, T. S., Koblar, S. A., & Lewis, M. D. (2015). The Role of the Neuroprotective Factor Npas4 in Cerebral Ischemia. *International Journal of Molecular Sciences*, *16*(12), 29011–29028.

Corneliussen, B., Thornell, A., Hallberg, B., & Grundström, T. (1991). Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. *Journal of Virology*, *65*(11), 6084–6093.

Davis, J. R., Kakar, M., & Lim, C. S. (2007). Controlling protein compartmentalization to overcome disease. *Pharmaceutical Research*, *24*(1), 17–27. https://doi.org/10.1007/S11095-006-9133-Z/TABLES/1

Dennis, D. J., Han, S., & Schuurmans, C. (2019). bHLH transcription factors in neural development, disease, and reprogramming. *Brain Research*, *1705*, 48–65.

Dinkel, H., & Sticht, H. (2007). A computational strategy for the prediction of functional linear peptide motifs in proteins. *Bioinformatics*, *23*(24), 3297–3303.

Dinkel, H., Van Roey, K., Michael, S., Davey, N. E., Weatheritt, R. J., Born, D., Speck, T., Krüger, D., Grebnev, G., Kubań, M., Strumillo, M., Uyar, B., Budd, A., Altenberg, B., Seiler, M., Chemes, L. B., Glavina, J., Sánchez, I. E., Diella, F., & Gibson, T. J. (2014). The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Research*, *42*(D1), D259–D266.

Dougherty, E. J., & Pollenz, R. S. (2010). ARNT: A Key bHLH/PAS Regulatory Protein Across Multiple Pathways. In *Comprehensive Toxicology: Second Edition* (Vols. 2–14, Issue August, pp. 231–252).

Dubrovsky, E. B. (2005). Hormonal cross talk in insect development. *Trends in Endocrinology and Metabolism: TEM*, *16*(1), 6–11.

Dubrovsky, E. B., Dubrovskaya, V. A., Bernardo, T., Otte, V., DiFilippo, R., & Bryan, H. (2011). The Drosophila FTZ-F1 nuclear receptor mediates juvenile hormone activation of E75A gene

expression through an intracellular pathway. *The Journal of Biological Chemistry*, *286*(38), 33689–33700.

Ellinghaus, D., Folseraas, T., Holm, K., Ellinghaus, E., Melum, E., Balschun, T., Laerdahl, J. K., Shiryaev, A., Gotthardt, D. N., Weismüller, T. J., Schramm, C., Wittig, M., Bergquist, A., Björnsson, E., Marschall, H. U., Vatn, M., Teufel, A., Rust, C., Gieger, C., … Karlsen, T. H. (2013). Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology*, *58*(3), 1074–1083.

Esser, J. S., Charlet, A., Schmidt, M., Heck, S., Allen, A., Lother, A., Epting, D., Patterson, C., Bode, C., & Moser, M. (2017). The neuronal transcription factor NPAS4 is a strong inducer of sprouting angiogenesis and tip cell formation. *Cardiovascular Research*, *113*(2), 222–223.

Fan, W., Long, Y., Lai, Y., Wang, X., Chen, G., & Zhu, B. (2016). NPAS4 Facilitates the Autophagic Clearance of Endogenous Tau in Rat Cortical Neurons. *Journal of Molecular Neuroscience*, *58*(4), 401–410.

Flora, A., Garcia, J. J., Thaller, C., & Zoghbi, H. Y. (2007). The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(39), 15382–15387.

Florindo, C., Ferreira, B. I., Serrão, S. M., Tavares, Á. A., & Link, W. (2016). Subcellular Protein Localisation in Health and Disease. *ELS*, 1–7.

Forrest, M., Chapman, R. M., Doyle, A. M., Tinsley, C. L., Waite, A., & Blake, D. J. (2012). Functional analysis of TCF4 missense mutations that cause Pitt-Hopkins syndrome. *Human Mutation*, *33*(12), 1676–1686.

Forrest, M. P., Waite, A. J., Martin-Rendon, E., & Blake, D. J. (2013). Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PloS One*, *8*(8), e73169.

Fowler, T., Sen, R., & Roy, A. L. (2011). Regulation of primary response genes. *Molecular Cell*, *44*(3), 348–360.

Fu, H., Cai, J., Clevers, H., Fast, E., Gray, S., Greenberg, R., Jain, M. K., Ma, Q., Qiu, M., Rowitch, D. H., Taylor, C. M., & Stiles, C. D. (2009). A genome-wide screen for spatially restricted expression patterns identifies transcription factors that regulate glial development. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(36), 11399–11408.

Fung, H. Y. J., Fu, S.-C., & Chook, Y. M. (2017). Nuclear export receptor CRM1 recognizes diverse conformations in nuclear export signals. *ELife*, *6*.

Giurgea, I., Missirian, C., Cacciagli, P., Whalen, S., Fredriksen, T., Gaillon, T., Rankin, J., Mathieu-Dramard, M., Morin, G., Martin-Coignard, D., Dubourg, C., Chabrol, B., Arfi, J., Giuliano, F., Claude Lambert, J., Philip, N., Sarda, P., Villard, L., Goossens, M., & Moncla, A. (2008). TCF4 Deletions in Pitt-Hopkins Syndrome. *Human Mutation*, *29*(11), E242–E251.

Gravina, G., Senapedis, W., McCauley, D., Baloglu, E., Shacham, S., & Festuccia, C. (2014). Nucleo-cytoplasmic transport as a therapeutic target of cancer. *Journal of Hematology & Oncology*, *7*(1), 85.

Greenberg, M. E., Hermanowski, A. L., & Ziff, E. B. (1986). Effect of protein synthesis inhibitors on growth factor activation of c-fos, c-myc, and actin gene transcription. *Molecular and Cellular Biology*, *6*(4), 1050–1057.

Großkopf, H., Walter, K., Karkossa, I., von Bergen, M., & Schubert, K. (2021). Non-Genomic AhR-Signaling Modulates the Immune Response in Endotoxin-Activated Macrophages After Activation by the Environmental Stressor BaP. *Frontiers in Immunology*, *0*, 1008.

Hauser, J., Saarikettu, J., & Grundström, T. (2008). Calcium regulation of myogenesis by differential calmodulin inhibition of basic helix-loop-helix transcription factors. *Molecular Biology of the Cell*, *19*(6), 2509–2519.

He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., Cutts, T., Andonov, A., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., Baker, L., & Li, X. (2004). Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochemical and Biophysical Research Communications*, *316*(2), 476–483.

Henrich, V. C., Burns, E., Yelverton, D. P., Christensen, E., & Weinberger, C. (2003). Juvenile hormone potentiates ecdysone receptor-dependent transcription in a mammalian cell culture system. *Insect Biochemistry and Molecular Biology*, *33*(12), 1239–1247.

Hill, M. J., Forrest, M. P., Martin-Rendon, E., & Blake, D. J. (2014). Association of Transcription Factor 4 (TCF4) variants with schizophrenia and intellectual disability. *Current Behavioral Neuroscience Reports*, *1*(4), 206–214.

Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R., Sikes, J. G., Obradovic, Z., & Dunker, A. K. (2004). The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Research*, *32*(3), 1037–1049.

Ikuta, T., Kobayashi, Y., & Kawajiri, K. (2004a). Cell Density Regulates Intracellular Localization of Aryl Hydrocarbon Receptor. *Journal of Biological Chemistry*, *279*(18), 19209–19216.

Ikuta, T., Kobayashi, Y., & Kawajiri, K. (2004b). Phosphorylation of nuclear localization signal inhibits the ligand-dependent nuclear import of aryl hydrocarbon receptor. *Biochemical and Biophysical Research Communications*, *317*(2), 545–550.

Ikuta, T, Namiki T, F. (2009). AhR protein trafficking and function in the skin. *Biochemical Pharmacology*, *77*(4), 588–596.

Ikuta, T., Watanabe, J., & Kawajiri, K. (2002). Characterization of the LxxLL motif in the aryl hydrocarbon receptor: Effects on subcellular localization and transcriptional activity. *Journal of Biochemistry*, *131*(1), 79–85.

Iserman, C. (2020). Genomic RNA elements drive phase separation of the SARS-CoV-2 nucleocapsid. *Mol. Cell*, *80*(6), 1078-1091.e6.

Jans, D. A., & Hübner, S. (1996). Regulation of protein transport to the nucleus: central role of phosphorylation. *Physiological Reviews*, *76*(3), 651–685.

Jennings, B. H. (2011). Drosophila – a versatile model in biology & medicine. *Materials Today*, *14*(5), 190–195.

Jindra, M., Palli, S. R., & Riddiford, L. M. (2013). The juvenile hormone signaling pathway in insect development. *Annual Review of Entomology*, *58*, 181–204.

Jindra, M., Uhlirova, M., Charles, J.-P., Smykal, V., & Hill, R. J. (2015b). Genetic Evidence for Function of the bHLH-PAS Protein Gce/Met as a Juvenile Hormone Receptor. *PLOS Genetics*, *11*(7), e1005394.

Jung, M., Häberle, B. M., Tschaikowsky, T., Wittmann, M.-T., Balta, E.-A., Stadler, V.-C., Zweier, C., Dörfler, A., Gloeckner, C. J., & Lie, D. C. (2018). Analysis of the expression pattern of the schizophrenia-risk and intellectual disability gene TCF4 in the developing and adult brain suggests a role in development and plasticity of cortical and hippocampal neurons. *Molecular Autism*, *9*, 20.

Kawajiri, K., & Fujii-Kuriyama, Y. (2017). *The aryl hydrocarbon receptor: a multifunctional chemical sensor for host defense and homeostatic maintenance Introduction: A Short History of TCDD and AHR Research*. *66*(October 2016), 75–89.

Kawajiri, K., & Ikuta, T. (2004). Regulation of Nucleo-Cytoplasmic Transport of the Aryl Hydrocarbon Receptor. *Journal of Health Science*, *50*(3), 215–219.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858.

Kennedy, A. J., Rahn, E. J., Paulukaitis, B. S., Savell, K. E., Kordasiewicz, H. B., Wang, J., Lewis, J. W., Posey, J., Strange, S. K., Guzman-Karlsson, M. C., Phillips, S. E., Decker, K., Motley, S. T., Swayze, E. E., Ecker, D. J., Michael, T. P., Day, J. J., & Sweatt, J. D. (2016). Tcf4 Regulates Synaptic Plasticity, DNA Methylation, and Memory Function. *Cell Reports*, *16*(10), 2666–2685.

Kolonko, M., & Greb-Markiewicz, B. (2019). bHLH–PAS Proteins: Their Structure and Intrinsic Disorder. *International Journal of Molecular Sciences*, *20*(15), 3653.

Lau, L. C. M., Ma, L., Young, A. L., Rong, S. S., Jhanji, V., Brelen, M. E., Pang, C. P., & Chen, L. J. (2014). Association of common variants in TCF4 and PTPRG with Fuchs' corneal dystrophy: a systematic review and meta-analysis. *PloS One*, *9*(10), e109142.

Ledent, V., Paquet, O., & Vervoort, M. (2002). Phylogenetic analysis of the human basic helix-loop-helix proteins. *Genome Biology*, *3*(6), 1–18.

Lingbeck, J. M., Trausch-Azar, J. S., Ciechanover, A., & Schwartz, A. L. (2003). Determinants of nuclear and cytoplasmic ubiquitin-mediated degradation of MyoD. *Journal of Biological Chemistry*, *278*(3), 1817–1823.

Liu, Y., Ray, S. K., Yang, X. Q., Luntz-Leybman, V., & Chiu, I. M. (1998). A splice variant of E2-2 basic helix-loop-helix protein represses the brain-specific fibroblast growth factor 1 promoter

through the binding to an imperfect E-box. *The Journal of Biological Chemistry*, *273*(30), 19269–19276.

Lu, J., Wu, T., Zhang, B., Liu, S., Song, W., Qiao, J., & Ruan, H. (2021). Types of nuclear localization signals and mechanisms of protein import into the nucleus. *Cell Communication and Signaling*, *19*(1), 1–10.

Lu, S., Ye, Q., Singh, D., Cao, Y., Diedrich, J. K., Yates, J. R., Villa, E., Cleveland, D. W., & Corbett, K. D. (2021). The SARS-CoV-2 nucleocapsid phosphoprotein forms mutually exclusive condensates with RNA and the membrane-associated M protein. *Nature Communications*, *12*(1).

Michael, A. K., Asimgil, H., & Partch, C. L. (2015). Cytosolic BMAL1 moonlights as a translation factor. *Trends in Biochemical Sciences*, *40*(9), 489–490.

Miura, K., Oda, M., Makita, S., & Chinzei, Y. (2005). Characterization of the Drosophila Methoprene - tolerant gene product. Juvenile hormone binding and ligand-dependent gene regulation. *The FEBS Journal*, *272*(5), 1169–1178.

Möglich, A., Ayers, R. A., & Moffat, K. (2009). Structure and Signaling Mechanism of Per-ARNT-Sim Domains. *Structure*, *17*(10), 1282–1294.

Moser, M., Knoth, R., Bode, C., & Patterson, C. (2004). LE-PAS, a novel Arnt-dependent HLH-PAS protein, is expressed in limbic tissues and transactivates the CNS midline enhancer element. *Molecular Brain Research*, *128*(2), 141–149.

Muir, T., Sadler-Riggleman, I., Stevens, J. D., & Skinner, M. K. (2006). Role of the basic helix-loop-helix protein ITF2 in the hormonal regulation of Sertoli cell differentiation. *Molecular Reproduction and Development*, *73*(4), 491–500.

Murre, C. (2019). Helix–loop–helix proteins and the advent of cellular diversity: 30 years of discovery. *Genes & Development*, *33*(1–2), 6–25. https://doi.org/10.1101/gad.320663.118

Muslin, a J., & Xing, H. (2000). 14-3-3 Proteins: Regulation of Subcellular Localization by Molecular Interference. *Cellular Signalling*, *12*(11–12), 703–709.

Namsi, A., Nury, T., Hamdouni, H., Yammine, A., Vejux, A., Vervandier-Fasseur, D., Latruffe, N., Masmoudi-Kouki, O., & Lizard, G. (2018). Induction of Neuronal Differentiation of Murine N2a Cells by Two Polyphenols Present in the Mediterranean Diet Mimicking Neurotrophins Activities: Resveratrol and Apigenin. *Diseases 2018, Vol. 6, Page 67*, *6*(3), 67.

Nardozzi, J. D., Lott, K., & Cingolani, G. (2010). *Phosphorylation meets nuclear import: a review*. 1–17.

Naskalska, A., Dabrowska, A., Szczepanski, A., Milewska, A., Jasik, K. P., & Pyrc, K. (2019). Membrane Protein of Human Coronavirus NL63 Is Responsible for Interaction with the Adhesion Receptor. *Journal of Virology*, *93*(19), 1–14.

Navarrete, K., Pedroso, I., De Jong, S., Stefansson, H., Steinberg, S., Stefansson, K., Ophoff, R. a., Schalkwyk, L. C., & Collier, D. a. (2013). TCF4 (e2-2; ITF2): A schizophrenia-associated gene with pleiotropic effects on human disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *162*(1), 1–16.

Obenauer, J. C. (2003). Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Research*, *31*(13), 3635–3641.

Obsilova, V., Kopecka, M., Kosek, D., Kacirova, M., Kylarova, S., Rezabkova, L., & Obsil, T. (2014). Mechanisms of the 14-3-3 protein function: regulation of protein function through conformational modulation. *Physiological Research / Academia Scientiarum Bohemoslovaca*, *63 Suppl 1*, S155-64.

Onions, J., Hermann, S., & Grundström, T. (2000). A Novel Type of Calmodulin Interaction in the Inhibition of Basic Helix−Loop−Helix Transcription Factors †. *Biochemistry*, *39*(15), 4366–4374.

Patil, V. M., Verma, S., & Masand, N. (2022). Prospective mode of action of Ivermectin: SARS-CoV-2. *European Journal of Medicinal Chemistry Reports*, *4*, 100018.

Perdikari, T. M., Murthy, A. C., Ryan, V. H., Watters, S., Naik, M. T., & Fawzi, N. L. (2020). SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *The EMBO Journal*, *39*(24), 1–15.

Phipps, M. S., & Cronin, C. A. (2020). Management of acute ischemic stroke. *BMJ*, *368*.

Plevin, M. J., Mills, M. M., & Ikura, M. (2005). The LXXLL motif: a multifunctional binding sequence in transcriptional regulation. *Trends in Biochemical Sciences*, *30*(2), 66–69.

Ponting, C. P., & Aravind, L. (1997). PAS: a multifunctional domain family comes to light. *Current Biology*, *7*(11), R674–R677.

Pursley, S., Ashok, M., & Wilson, T. G. (2000). Intracellular localization and tissue specificity of the Methoprene-tolerant (Met) gene product in Drosophila melanogaster. *Insect Biochemistry and Molecular Biology*, *30*(8–9), 839–845.

Quednow, B. B., Brzózka, M. M., & Rossner, M. J. (2014). Transcription factor 4 (TCF4) and schizophrenia: integrating the animal and the human perspective. *Cellular and Molecular Life Sciences*, *71*(15), 2815–2835.

Riddiford, L. M. (1996). Juvenile hormone: the status of its "status quo" action. *Archives of Insect Biochemistry and Physiology*, *32*(3–4), 271–286.

Riddiford, L. M. (2012). How does juvenile hormone control insect metamorphosis and reproduction? *General and Comparative Endocrinology*, *179*(3), 477–484.

Röllen, K., Granzin, J., Panwalkar, V., Arinkin, V., Rani, R., Hartmann, R., Krauss, U., Jaeger, K. E., Willbold, D., & Batra-Safferling, R. (2016). Signaling States of a Short Blue-Light Photoreceptor Protein PpSB1-LOV Revealed from Crystal Structures and Solution NMR Spectroscopy. *Journal of Molecular Biology*, *428*(19), 3721–3736.

Romero, N. M., Irisarri, M., Roth, P., Cauerhff, A., Samakovlis, C., & Wappner, P. (2008). Regulation of the Drosophila hypoxia-inducible factor alpha Sima by CRM1-dependent nuclear export. *Molecular and Cellular Biology*, *28*(10), 3410–3423.

Sabatini, P. V, Krentz, N. A. J., Zarrouki, B., Westwell-Roper, C. Y., Nian, C., Uy, R. A., Shapiro, A. M. J., Poitout, V., & Lynn, F. C. (2013). Npas4 is a novel activity-regulated cytoprotective factor in pancreatic β-cells. *Diabetes*, *62*(8), 2808–2820.

Sabatini, P. V., & Lynn, F. C. (2015). All-encomPASsing regulation of β-cells: PAS domain proteins in β-cell dysfunction and diabetes. *Trends in Endocrinology & Metabolism*, *26*(1), 49–57.

Sepp, M., Kannike, K., Eesmaa, A., Urb, M., & Timmusk, T. (2011). Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5' exon usage and splicing. *PloS One*, *6*(7), e22138.

Sepp, M., Pruunsild, P., & Timmusk, T. (2012). Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Human Molecular Genetics*, *21*(13), 2873–2888.

Sepp, M., Vihma, H., Nurm, K., Urb, M., Page, S. C., Roots, K., Hark, A., Maher, B. J., Pruunsild, P., & Timmusk, T. (2017). The Intellectual Disability and Schizophrenia Associated Transcription Factor TCF4 Is Regulated by Neuronal Activity and Protein Kinase A. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(43), 10516–10527.

Shamloo, M., Soriano, L., Von Schack, D., Rickhag, M., Chin, D. J., Gonzalez-Zulueta, M., Gido, G., Urfer, R., Wieloch, T., & Nikolich, K. (2006). Npas4, a novel helix-loop-helix PAS domain protein, is regulated in response to cerebral ischemia. *European Journal of Neuroscience*, *24*(10), 2705–2720.

Shemshedini, L., & Wilson, T. G. (1990). Resistance to juvenile hormone and an insect growth regulator in Drosophila is associated with an altered cytosolic juvenile hormone-binding protein. *Proceedings of the National Academy of Sciences*, *87*(6), 2072–2076.

Shemshedini, L., & Wilson, T. G. (1993). Juvenile hormone binding proteins in larval fat body nuclei of Drosophila melanogaster. *Journal of Insect Physiology*, *39*(7), 563–569.

Sluchanko, N. N., Artemova, N. V, Sudnitsyna, M. V, Safenkova, I. V, Antson, A. A., Levitsky, D. I., & Gusev, N. B. (2012). *Monomeric 14-3-3 ζ Has a Chaperone-Like Activity and Is Stabilized by Phosphorylated HspB6*.

Sluchanko, N. N., & Gusev, N. B. (2012). Oligomeric structure of 14-3-3 protein: what do we know about monomers? *FEBS Letters*, *586*(24), 4249–4256.

Sobrado, V. R., Moreno-Bueno, G., Cubillo, E., Holt, L. J., Nieto, M. A., Portillo, F., & Cano, A. (2009). The class I bHLH factors E2-2A and E2-2B regulate EMT. *Journal of Cell Science*, *122*(7), 1014–1024.

Speckmann, T., Sabatini, P. V, Nian, C., Smith, R. G., & Lynn, F. C. (2016). Npas4 Transcription Factor Expression Is Regulated by Calcium Signaling Pathways and Prevents Tacrolimus-induced Cytotoxicity in Pancreatic Beta Cells. *The Journal of Biological Chemistry*, *291*(6), 2682–2695.

Sullivan, A. E., Peet, D. J., & Whitelaw, M. L. (2016). MAGED1 is a novel regulator of a select subset of bHLH PAS transcription factors. *FEBS Journal*, *283*, 3488–3502.

Sun, H., Ghaffari, S., & Taneja, R. (2007). bHLH-Orange Transcription Factors in Development and Cancer. *Translational Oncogenomics*, *2*, 107–120.

Sun, X., Wang, Y., & Sui, N. (2018). Transcriptional regulation of bHLH during plant response to stress. *Biochemical and Biophysical Research Communications*, *503*(2), 397–401.

Surjit, M., & Lal, S. K. (2008). The SARS-CoV nucleocapsid protein: A protein with multifarious activities. *Infection, Genetics and Evolution*, *8*(4), 397–405.

Sweatt, J. D. (2013). Pitt-Hopkins Syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. *Experimental & Molecular Medicine*, *45*, e21.

Wang, L., Xing, J., Cheng, R., Shao, Y., Li, P., Zhu, S., & Zhang, S. (2015). Abnormal localization and tumor suppressor function of epithelial tissue-specific transcription factor ESE3 in esophageal squamous cell carcinoma. *PLoS ONE*, *10*(5), 1–14.

Wang, S., Baumann, A., & Wilson, T. G. (2007). Drosophila melanogaster Methoprene-tolerant (Met) gene homologs from three mosquito species: Members of PAS transcriptional factor family. *Journal of Insect Physiology*, *53*(3), 246–253.

Weber, D., Wiese, C., & Gessler, M. (2014). Hey bHLH transcription factors. In *Current Topics in Developmental Biology* (Vol. 110, pp. 285–315).

Whalen, S., Héron, D., Gaillon, T., Moldovan, O., Rossi, M., Devillard, F., Giuliano, F., Soares, G., Mathieu-Dramard, M., Afenjar, A., Charles, P., Mignot, C., Burglen, L., Van Maldergem, L., Piard, J., Aftimos, S., Mancini, G., Dias, P., Philip, N., … Giurgea, I. (2012). Novel comprehensive diagnostic strategy in Pitt-Hopkins syndrome: Clinical score and further delineation of the TCF4 mutational spectrum. *Human Mutation*, *33*(1), 64–72.

Wigington, C. P., Roy, J., Damle, N. P., Yadav, V. K., Blikstad, C., Resch, E., Wong, C. J., Mackay, D. R., Wang, J. T., Krystkowiak, I., Bradburn, D. A., Tsekitsidou, E., Hong, S. H., Kaderali, M. A., Xu, S. L., Stearns, T., Gingras, A. C., Ullman, K. S., Ivarsson, Y., … Cyert, M. S. (2020). Systematic Discovery of Short Linear Motifs Decodes Calcineurin Phosphatase Signaling. *Molecular Cell*, *79*(2), 342-358.e12.

Wilson, T. G., & Fabian, J. (1986). A Drosophila melanogaster mutant resistant to a chemical analog of juvenile hormone. *Developmental Biology*, *118*(1), 190–201.

Wirgenes, K. V, Sønderby, I. E., Haukvik, U. K., Mattingsdal, M., Tesli, M., Athanasiu, L., Sundet, K., Røssberg, J. I., Dale, A. M., Brown, A. A., Agartz, I., Melle, I., Djurovic, S., & Andreassen, O. A. (2012). TCF4 sequence variants and mRNA levels are associated with neurodevelopmental characteristics in psychotic disorders. *Translational Psychiatry*, *2*(5), e112.

Wright, P. E., & Dyson, H. J. (2014). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, *16*(1), 18–29.

Wu, D., Su, X., Potluri, N., Kim, Y., & Rastinejad, F. (2016). NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors. *ELife*, *5*, 1–15.

Wubben, J. M., Atkinson, S. C., & Borg, N. A. (2020). The Role of Protein Disorder in Nuclear Transport and in Its Subversion by Viruses. *Cells*, *9*(12), 1–24.

Xiong, X. Y., Liu, L., & Yang, Q. W. (2018). Refocusing neuroprotection in cerebral reperfusion era: New challenges and strategies. *Frontiers in Neurology*, *9*(APR), 249.

Ye, Q., West, A. M. V., Silletti, S., & Corbett, K. D. (2020). Architecture and self-assembly of the SARS-CoV-2 nucleocapsid protein. *Protein Science*, *29*(9), 1890–1901.

Yoon, S.-J., Wills, A. E., Chuong, E., Gupta, R., & Baker, J. C. (2011). HEB and E2A function as SMAD/FOXH1 cofactors. *Genes & Development*, *25*(15), 1654–1661.

Zeng, W., Liu, G., Ma, H., Zhao, D., Yang, Y., & Liu, M. (2020). Biochemical characterization of SARS-CoV-2 nucleocapsid protein. *Biochemical and Biophysical Research Communications*, *January*.

Zhang, Z., Fei, P., Mu, J., Li, W., & Song, J. (2014). Hippocampal expression of aryl hydrocarbon receptor nuclear translocator 2 and neuronal PAS domain protein 4 in a rat model of depression. *Neurological Sciences*, *35*(2), 277–282.

Zinzula, L., Basquin, J., Bohn, S., Beck, F., Klumpe, S., Bracher, A., Hartl, F. U., & Baumeister, W. (2020). High-resolution structure and biophysical characterization of the nucleocapsid

phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. *Biochemical and Biophysical Research Communications*, *January*.

Zweier, C., Peippo, M. M., Hoyer, J., Sousa, S., Bottani, A., Clayton-Smith, J., Reardon, W., Saraiva, J., Cabral, A., Göhring, I., Devriendt, K., de Ravel, T., Bijlsma, E. K., Hennekam, R. C. M., Orrico, A., Cohen, M., Dreweke, A., Reis, A., Nürnberg, P., & Rauch, A. (2007). Haploinsufficiency of TCF4 Causes Syndromal Mental Retardation with Intermittent Hyperventilation (Pitt-Hopkins Syndrome). *The American Journal of Human Genetics*, *80*(5), 994–1001.

Zweier, C., Sticht, H., Bijlsma, E. K., Clayton-Smith, J., Boonen, S. E., Fryer, A., Greally, M. T., Hoffmann, L., den Hollander, N. S., Jongmans, M., Kant, S. G., King, M. D., Lynch, S. A., McKee, S., Midro, A. T., Park, S.-M., Ricotti, V., Tarantino, E., Wessels, M., … Rauch, A. (2008). Further delineation of Pitt-Hopkins syndrome: phenotypic and genotypic description of 16 novel patients. *Journal of Medical Genetics*, *45*(11), 738–744.

5. Presentation of significant scientific activity carried out at more than one university, scientific institution, especially at foreign institutions

5.1 Long-term stays and short-term internships in research units abroad and other than Wroclaw University of Science and Technology

1-31.07.2000, Italy, Napoli, Istituto per la Chimica di Molecole di Interesse Biologico di Arco Felice (Pozzuoli); research training; Head: Prof. Guido Cimino. During this stay, I learned methods of purifying low-molecular secondary metabolites from marine organisms.

31.10-5.11.2001 Sweden, Alnarp, Institutionen för Växtvetenskap, Swedish University of Agricultural Sciences; short research training. During my stay, I learned how the university works and how to conduct research on the protection of potato crops against pests.

1.08.2002-31.07.2004 Germany, Ulm, Department of General Zoology and Endocrinology, Ulm University, assistant professor (Wissenschaftliche Angestellte); Head of Institute: Prof. Klaus Dieter Spindler. I described my activity during this stay in detail in the section 4.3.

01.07.2013-31.07.2013 Poland, Cracow, Department of Cell Biophysics, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, research training; Head of Department: Prof. Jerzy Dobrucki. During my stay, I learned the theoretical basics and the use of more advanced techniques of imaging mammalian cells using the Leica SP5 confocal microscope (e.g., sample preparation, FRAP method, images analysis using ImageJ. I also took part in imaging of fixed tissues/specimens prepared by other researchers).

01.03.2016-31.07.2016 Germany, Dresden, Max Planck Institute of Molecular Cell Biology and Genetics, Prof. Susanne Eaton group, Guest scientist. The first time I met Prof. Eaton at the Conference in Crete (2015) during which I presented orally the results of my research on the GCE protein. Then, we discussed the results and later Prof. Eaton invited me to do a research training. The aim of my stay was to learn the basics of working with the model organism

Drosophila melanogaster and to perform experiments to study the impact of silencing and overexpression of GCE and MET proteins during development of this model organism. I learned how to prepare wing discs of *Drosophila*, how to fix them and prepare for imaging with fluorescently labeled antibodies. As a result, I could observe that MET/GCE overexpression increased caspase expression leading to increase of cell apoptosis. During the research, I was able to observe differences between the effects of GCE and MET. In addition, I performed CRISPR/Cas9 genome editing of *Drosophila* to label endogenous GCE with GFP protein to be able to image a non-overexpressed protein. I learned much during these five months, still further research was necessary to publish the results, which was planned in the later term, but was not completed due to the tragic death of Prof. Eaton. Importantly, I subsequently taught students of the 'BioTop' student science group to breed wild type *D. melanogaster*. In the future, I plan to organize a GMO laboratory to be able to perform more advanced research not only on MET or GCE, but also on the insect homologue of the NPAS4 protein.

5.2 Scientific cooperation

For many years I have been actively cooperating with Prof. J. Dobrucki and Dr. M. Zarębski from the Department of Cell Biophysics (Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Cracow). This cooperation is documented by many papers published in the international journals with IF (P1-P4 and A12). Our current collaboration is focused on the experiments testing the contribution of NPAS4 in DNA repair, searching for NPAS4/TCF4 partners by PLA (ang. *proximity ligation assay*). A joint application for research funding is planned.

I also established cooperation with Dr. J. Godlewski during his stay in the laboratory of Prof. T.G. Wilson (Department of Entomology, Ohio State University). Our cooperation is documented by publishing the joint paper (P2).

I cooperated with Prof. V. Uverski (University of South Florida) to characterize unstructured regions of the bHLH-PAS proteins. The cooperation consisted in performing by Prof. Uversky as an expert some of more advanced *bioinformatic analysis of* selected proteins and was documented by the published article [P9].

Currently, an active cooperation is ongoing with Prof. Nicola d'Amelio (Université de Picardie Jules Verne). We will test interaction of selected NPAS4 fragments with biomimetic membranes by liquid state NMR of bicelles. Also, molecular dynamics simulations between

NPAS4 fragments and nuclear phospholipids will be performed. A joint application for research funding is planned.

As a part of cooperation with Prof. C. Tkaczuk (Institute of Agriculture and Horticulture, Faculty of Agrobioengineering and Animal Sciences, University of Natural Sciences and Humanities in Siedlce), I obtained strains of entomopathogenic fungi, that resulted in the publication [A14].

As a part of cooperation with Dr. Marta Polańska (Department of Animal Physiology, University of Warsaw) I obtained a model organism *Galleria melonella*.

International cooperation within the Crowdfight group, which brings together volunteer scientists from around the world, resulted in the acquisition of many cDNA templates of SARS-CoV-2 proteins (resulting in A14 paper). In addtion, cooperation resulted in the review paper published in high impact journal *Trends in Microbiology* (IF=18.230) described in the other section of the summary.

5.3 Active participation in the international and national scientific conferences

The list below includes only conferences where oral or poster presentations were given personally.
A complete list of all conferences and conference reports is presented in the **Appendix 4B**.

Oral presentations:

- 18th Congress of the Polish Biophysical Society (PTBF), 6-9 September 2022, Warsaw, Poland.
- invited lecture as the honour guest of the 4th National Scientific Conference „Primary and secondary metabolites of plants and fungi" 15 July 2021, virtual.
- 2nd International Insect Hormones, 12-18 July 2015, Chania, Crete, Greece.
- Ecdysone Workshop 2010, 19-23 July 2010, Ceske Budejovice, - Czech Republic.
- Ecdysone Workshop 2008, 20-24 July 2008, Ulm, Germany.
- 6th Parnas Conference 'Molecular mechanism of cellular signaling', 30 May-2 June 2007, Cracow, Poland.
- Cost Action 842: Biological Control of Pest Insects and Mites with Special Reference to Entomophthorales, Meeting of the Working groups 1 and 2, 18-20 April 2002, Heracleum, Crete, Greece.

- Lecture ‘Searching for agents involved in infection and intoxication process of insect’s organism by entomopathogenic fungi belonging to *Zoophthora* and *Paecilomyces* genera’. 9 April 2002, *Institut für Allgemeine und Spezielle Zoologie,* University of Giessen, Giessen, Germany.
- INCHECO 2nd Conference, 7-9 March 2002, Calambrone, Pisa, Italy.
I Polish Congress of Biotechnology, 20-25 September 1999, Wroclaw, Poland.
- Meeting of the working group Insect Pathogens and Entomoparasitic Nematodes, 10-15 August 1997, Kopenhagen, Denmark.

Poster presentations:

- EMBO Workshop - Recent advances in structural biology of membrane proteins, 29 November-1 December 2021, virtual.
- 14th EMBL Conference - Transcription and Chromatin, 27-29 August 2020, virtual.
- 43rd FEBS Congress ‘Biochemistry forever’, 7–12 July 2018, Prague, Czech Republic.
- 2nd Congress of Polish Biochemistry, Cell biology, Biotechnology and Bioinformatics “Bio2016 – Expanding beyond the limits”, 13-16 September 2016, Wroclaw, Poland.
- Seventh International Symposium on Molecular Insect Science, 13-16 July 2014, Amsterdam, the Netherlands.
- Polish-German Biochemical Societies Joint Meeting, Biochemistry for health and environment, 11-14 September 2012, Poznan, Poland.
- 33rd FEBS Congress ‘Biochemistry of cell regulation’, June 28- July 3, 2008, Athens, Greece.
- 32nd FEBS Congress ‘Molecular Machines and their Dynamics in Fundamental Cellular Functions’, 7-12 July 2007, Vienna, Austria.

5.4 International awards for scientific activity

- Participation after candidates’ selection in the EMBO Practical Course “Drosophila Genetics and Genomics” 11-15 January 2021, virtual.
- Positive selection and participation in the course ‘Bioinformatics of protein-protein interactions for wet lab scientist’, 3-7 April 2017, Pasteur Institute, Paris, France.
- FEBS bursary, 33rd FEBS Congress and 11th IUBMB Conference. ‘Biochemistry of cell regulation’, June 28-July 3, 2008. Athens, Greece.

- Swedish Institute 9-month scholarship in 2002; (finally not used because of obtaining a 2-year assistant professor position at Ulm University at the same time (2002-2004).

5.5 Research projects and grant efforts

April 2023 obtaining an internal faculty grant from the Dean of the Faculty of Chemistry of PWr for research on NPAS4 protein to prepare for NCN grant submission, 19900 PLN

June 2022 as the leader of the consortium, in cooperation with Agnieszka Bronowicka-Szydełko, PhD from the Medical University of Wroclaw submission of application to the Opus 23 call. 2022/45/B/NZ1/04262, the title “Study of the molecular basis of functioning and verification of the diagnostic significance of the cytoprotective protein NPAS4 belonging to the family of transcription factors bHLH-PAS”. After a positive formal assessment, the project was qualified for the first stage of substantive assessment. The proposal did not progress to the second stage of the substantive assessment, though the subject was considered as interesting. The justification was, among other things, too wide scope and difficulty of the project and too little experimental proof of hypotheses, which was associated with high risk, according to one of the evaluating experts. Importantly, the choice of NPAS4 as a research subject has been positively evaluated, therefore I plan to focus and refine the scope of the project and experimental work to apply for NCN funding again.

July 2021-December 2022 Leader of the SKN/SP/496912/2021 project entitled "Study of the interaction between entomopathogenic fungi and cultivated plants" carried out by the scientific club of biotechnology students "BioTop" at the Faculty of Chemistry of the WUST, as part of the Ministerial program "student clubs create Innovations". Funding requested 68000, funding received 18300 PLN, implementation period 12 months. The project was settled positively. As part of the project, the paper A14 was published.

June 2021 as a Principal Investigator - the Opus 21 call submission of application 2021/41/B/NZ1/0398. The project title 'Study of the molecular basis for the functioning of the cytoprotective factor NPAS4'. After a positive formal assessment, the project was qualified for the first stage of substantive assessment. The proposal did not progress to the second stage of the substantive assessment.

2018-2022 Participation as a contractor in the project PRELUDIUM 14, Principal Investigator Dr. Marta Kolonko-Adamska UMO-2017/27/N/NZ1/01783; Structural and functional analysis

of the C-terminal region of the Germ cell-expressed protein (GCE from Drosophila melanogaster as an inherently disordered region. Funding 140000 PLN
Project start date 2018-09-10; Project completion: 2022-09-09.

2017-2018 Leader of the project "Identification of active localization signals of NPAS4 protein and optimization of expression for NPAS4 purification. Registration No. 624/2017/KNOW
Source of funding Wrocław Biotechnology Centre by the KNOW program (National Scientific Leadership Centre) for years 2014-2018; Area 1: Scientific and didactic activity, subactivity 4: Support for scientific activities in new research projects as part of the scientific staff development, Funding 35000 PLN, Entity implementing Wroclaw University of Technology. Start date 2017-06-05; Completion date 2018-06-08

2015-2016 Leader of the project "Overexpression and purification of Drosophila melanogaster Methoprene tolerant (MET) protein. Registration NO. 4503/0001/14 KNOW
Source of funding Wroclaw Biotechnology Centre by the KNOW program (National Scientific Leadership Centre) for years 2014-2018. Area 1: Scientific and didactic activity, subactivity 4: Support for scientific activities in new research projects as part of the scientific staff development. Funding 19800 PLN, Entity implementing: Wroclaw University of Science and Technology. Start date 2015-06-15, Completion date 2016-06-15.

2015 as part of KNOW 2014-2018 I obtained financing for the costs of language correction and publication of the P2 work.

2014-2015 Biotechnologies and advanced health technologies" - EU funds. Operational Programme Innovative Economy Project No. POIG.01.01.02-003/08 Wroclaw Research Centre EIT+ "Diagnostics of bacterial diseases", participation in the project, leader Prof. A. Gamian.

2012 contractor of the statutory research of the Department of Biochemistry (Wroclaw University of Technology) "Structure-function relationship - N-terminal domains of nuclear receptors EcR and Usp - new targets of the specific gene expression regulators".

2007-2011 contractor of the statutory research of the Department of Biochemistry (Wroclaw University of Technology) "Disordered and globular proteins – differences in molecular determinants of function".

2005 contractor of the statutory research of the Department of Biochemistry (Wroclaw University of Technology) "Molecular mechanisms of the biological signal transduction".

1993-2002 participation as contractor in several research grants awarded by the Polish State Committee for Scientific Research (KBN) to Prof. dr hab. Andrzej Zabża resulted in several SPR reports. One of the State Committee for Scientific Research Grant (grant 6 6021 92 03) - resulted in the paper [A1].

6. Presentation of teaching and organizational achievements as well as achievements in the popularization of science

6.1 Preparation of new courses and teaching materials.

- Molecular biology - lecture prepared and conducted in English, 30h/semester from 2016.

- Biochemistry laboratory- laboratory prepared (translation of instructions) and conducted in English, 60h/semester, from 2015.

- 'Introduction to GMOs'- the lecture for II level students of the specialty Environmental Biotechnology conducted in Polish, 15h/semester, from 2019.

- Developing the lecture lecture „Application of chromatography in protein preparation"as a part of the course „Chromatographic methods in chemistry and biotechnology" for I level students of Faculty (6h/semester) from 2017.

- Developing and conducting a laboratory exercise entitled 'Analysis of insertion-deletion polymorphism of the angiotensin converting enzyme (ACE)' at the classes 'Genetic engineering in analysis and diagnostics' within the program 'Ordered directions'. The programme was financed from the state budget and structural funds of EU.

- Developing and conducting of a laboratory exercise entitled 'Isolation of genomic DNA from the model organism Drosophila melanogaster', as part of the classes 'Modern research methods and tools in biochemistry and molecular biology' conducted by the Department of Biochemistry in 2019 i 2022 under the Operational Programme Knowledge Education Development (POWER.03.01.00-00-K088/15). I prepared this exercise using the skills of working with *Drosophila melanogaster* acquired during an internship in Susanne Eaton's group in Max-Planck Institute in Dresden aiming not only to introduce students to the basics of breeding and working with the model organism D. melanogaster and to teach the method of genomic DNA isolation from this organism, but also an indication of the need and method of planning a CRISPR/CAS9 experiment. Using the knowledge gained during my stay at the Max-Planck Institute in Dresden, I also prepared an outline of a new lecture for the second level students.

‘Molecular Basics of Developmental Biology’ - the lecture is awaiting introduction to the program in next years.

- “Natural products chemistry” laboratory prepared and conducted 30h/semester, years 1999-2002.

- Fom many years conductred laboratories: Biochemistry, Enzymology, Genetic engineering, and seminar: Molecular biology.

6.2 Supervision of students (research and diploma theses)

Total number of the first stage diploma theses (Bachelor's /Engineering): 28

Piotr Kruszyński, Dawid Zapart, Karolina Sekutowicz, 2023;

Julia Witek, 2022;

Julia Cicha, Dominika Chrapek, Hanna Kończak, Martyna Stępkowska, 2021;

Dominika Dobrowolska, Karolina Mróz, 2019;

Aleksandra Brzóstowicz, Alicja Siewko, 2017;

Zuzanna Gogół, Wioletta Kazana, Joanna Kumejko, Marcin Tomczyszyn, Kamil Dąbrowski, 2016;

Alicja Pakiet, Beata Wilgan, 2015;

Szymon Kordon, Katarzyna Ożga, 2014;

Jakub J. Jerczyński, Karolina Kołaczek, Agnieszka Olechwier, 2013;

Patryk Duraj, Patrycja Polowczyk, 2012;

Natalia Banaś, Daria Sadowska, 2011;

Anna Bukowska (Majcher), 2010;

I would like to point out that the diploma theses I supervised at the first degree of studies were mostly experimental.

Total number of the second stage diploma theses (Master’s Theses): 28

Adrian Balecki, Application of Pichia pastoris yeast system for mammalian protein expression, 2022.

Mateusz Marciniak, Preparation of vectors containing SARS-CoV2 M and E proteins cDNA for expression of these proteins in bacterial cells, 2022.

Julia Cicha, Assessment of relationship between Calmodulin and NPAS4 protein expression, 2022.

Marta Rapińska, Assessment of the proline-rich sequence effect on the NPAS4 protein expression, 2022.

Joanna Kaczyńska, Presentation of the current issues of the animal extinct species restoration, 2022.

Dominika Szczęsna, Preparing of cDNA constructs for interaction studies of NPAS4 with selected proteins in mammalian cells, 2021.

Katarzyna Jurkowska, Adjusting of conditions for expression of full length NPAS4 protein in E. coli cells using pCold-TF vector, 2020.

Kinga Wyzujak, Preparing of cDNA vector for overexpression of mutant of N-terminal part of NPAS4 protein with deletion of rich in prolines sequence in bacterial cells, 2020.

Alicja Siewko, Adjusting of conditions for research of NPAS4 protein interaction with ARNT2, 2019.

Aleksandra Brzóstowicz, Adjusting of conditions for research of NPAS4 protein interaction with 14-3-3 proteins, 2019.

Oliwia Kozieł, Preparing of cDNA constructs for NDRF protein from Mus musculus expression in bacterial cells, 2018.

Natalia Ponikowska, Preparing of cDNA constructs for expression of NeuroD protein from Mus musculus in bacterial cells, 2018.

Wioletta Kazana, Adjusting of conditions for expression and purification of C-terminal region from Rattus norvegicus NPAS4 protein, 2017.

Zuzanna Dennstedt, Adjusting of conditions for expression and purification of N-terminal region from Rattus norvegicus NPAS4 protein, 2017.

Joanna Kumejko, Searching for subcellular localization signals of NeuroD1 from Mus musculus, 2017.

Marcin Tomczyszyn, Searching for subcellular localization signals of NDRF from Mus musculus, 2017;

Katarzyna Ożga, Adjusting of conditions for expression and purification of C-terminal region from Drosophila melanogaster MET protein, 2015.

Szymon Kordon, Adjusting of conditions for expression and purification of C-terminal region from Drosophila melanogaster GCE protein, 2015.

Daria Sadowska, Preparing DNA vectors for research of subcellular localization in mammalian cells and overexpression in bacterial cells of the NDRF protein from *Mus musculus*, 2013.

Natalia Banaś, Preparing DNA vectors for research of subcellular localization in mammalian cells and overexpression in bacterial cells of the NeuroD protein from *Mus musculus*, 2013.

Anna Bukowska, Preparing DNA vectors for subcellular localization and overexpression research of the NPAS4 (Neuronal PAS domain containing protein 4) from *Rattus norvegicus*, 2012.

Natalia Bielak, Preparing DNA vectors for subcellular localization and overexpression research of the TCF4 protein (Transcription factor 4) from *Mus musculus*, 2012, in English;

Mirosława Różycka, Adjusting of conditions for expression and purification of PAS domains from *Drosophila melanogaster*, 2010.

Katarzyna Dzierzba, Preparing DNA vectors for expression of PAS domains from the MET protein in bacterial cells using pQE-80L plasmid, 2008.

Agnieszka Bronowicka, Preparing DNA vectors for expression of PAS domains from the MET protein in bacterial cells using pGEX-2T plasmid, 2008.

Marta Groblica, Preparing DNA vectors for subcellular localization studies of MET protein in Schneider cells, 2007.

Zyta Bart, Chromatographic separation of metabolites isolated from entomopathogenic fungus *Paecilomyces s*p. Cat. 3444b (Bałazy), 2002.

Marta Gerelecka, Chromatographic separation of metabolites isolated from entomopathogenic fungus *Hirsutella thompsoni*i, 2002.

Scientific supervision of foreign students

Ibrahim Duman, research laboratory, summer 2018/2019.

Beytullah Ecevit, Cloning of the ARNT2 cDNA into pQE80L vector and the NPAS4 cDNA into pETM-11 vector, winter, Erasmus student, bachelor, 2017/2018.

Ezgi Berfin Ceper, research laboratory, summer 2016/2017.

Ayca Aslan, research laboratory, winter 2015/2016.

It is worth to note that most of the graduates I supervised decided to continue their scientific work, gaining further academic degrees, and achieving successes.

Additionally, I was the reviewer of several BEng and Master Theses.

6.3 Scientific supervision of doctoral students as an assistant supervisor.

- Assistant supervisor of the PhD student M. Kolonko-Adamska. The dissertation entitled 'Molecular analysis of C-terminal fragments of receptors: Methoprene tolerant and Germ cell-expressed from Drosophila melanogaster' was realized by Dr. M. Kolonko-Adamska in 2015-2020 and defended in 2020.
- Assistant supervisor of the PhD first year student Dominika Szczęsna 2021/2022.

6.4 Other achievements in teaching and in science popularization

6.4.1 Supervision of the student science club

Since 2017, I have been the supervisor of the biotechnology student's science club „BioTop". This includes active involvement in the selection of the scientific projects as well as active daily scientific supervision. I introduced the members of the club to model organisms

Drosophila melanogaster and *Galleria mellonella* by acquiring these organisms and teaching students the basics of working with them. Similarly, by establishing cooperation, I obtained strains of entomopathogenic fungi and presented to students the basics of work with these organisms.

Recently, as the project manager, I actively supervised preparation and sent an application to the MEiN for funding research of the student scientific club BioTop as part of the program "Student Scientific Clubs create innovations". The project "Study of the interaction between entomopathogenic fungi and cultivated plants" was approved by the MEiN. Duration of the project was 2.07.2021-1.12.2022. Importantly, results of the project were presented at the International Conference and published in the Journal with IF [A16].

As the supervisor of the 'BioTop', I also initiated the organization of virtual information meetings on the SARS-CoV2 virus using the ZOOM platform by the scientific club in the first months of the pandemic (2020). Experts in the field employed at Wroclaw University of Science and Technology were invited to deliver information, and then the topic was discussed.

6.4.2 Participation in the Lower Silesian Science Festival

For many years I have been actively involved in popularizing science and actively participate in the Lower Silesian Science Festival at the Wroclaw University of Technology from the first festival in 2007, each year (documented by many diplomas). During the first festival, I prepared laboratory presentations myself, later with the help of my colleagues from the Department of Biochemistry. Every year I prepare lectures on various topics, for example: 'Introduction to Circadian rhythm', 'Entomopathogenic fungi contra insects - military race', 'Model organisms'. I also encouraged representatives of the scientific club BioTop to participate in the DFN BioTop.

6.4.3 Other activities to popularize science

- I was invited to present a lecture (subject: circadian clock, date of presentation: 6.03.2023) for the University of the Third Age hosted by Wroclaw University.
- I prepared and presented lectures during Open Days and LabDay organized by the Faculty of Chemistry Wroclaw University of Science and Technology (years 2018-2022)

- I took part in the promotion of the Faculty of Chemistry, Wroclaw University of Science and Technology during the 10th Lower Silesian Educational Presentations - Tared 2007.

6.5 Organizational achievements

- Award of the Rector of the WUST for a significant contribution to the functioning of the University in 2006, 2011 and 2016.
- Preparation of the Department's GMO applications (the first application at the University in 2006, later in 2011, 2016). Supervision of work with GMOs in the Laboratory. Participation in the inspection of the Laboratory by authorized state institutions as the Department
- Scheduling of individual employees/PhD students' responsibilities in the Department, also supervision of persons responsible for equipment/ reagents.
representative.
- Chair of the GMO Commission in the Faculty of Chemistry.
- The member of the Disciplinary Committee of the Wroclaw University of Technology for employees (2016-2020).
- Elected to be a member of the Council of the Faculty of Chemistry in the years 2012-2015, 2016-2019 and 2021-2024.
- The Member of the Faculty Advisory Board in the years 2020-2021.
- The secretary of the Electoral Committee of the Faculty of Chemistry from 2016 to date.
- I am the author of idea and founder (talks with J.M Rector, announcements about the recruitment of students and employees, searching a conductor, organization) and organizing the activity (arranging the headquarters, running the FB page) of the Orchestra of Wroclaw University of Technology in 2011/2012 - existing to date. Also, long-time concertmaster of the orchestra.
- Obtaining ISO 9001 Certificate in February 2023, Polish Center of Certification (PCC).

7. Apart from information set out in 1-6 above, the applicant may include other information about his/her professional career, which he/she deems important.

7.1 Scientific courses leading to a diploma/certificate

2010/2011 Two semester study 'Management of research project and commercialization of research results', Faculty of Computer Science and Management, Wroclaw University of Science and Technology, Poland.

16-20.11.2015 The course ’Training for persons responsible for the planning and performing procedures and experiments and killing animals’; Certificate holder. PolLASA, Wroclaw Poland.
3-7.04.2017 The course of ‘Bioinformatics of protein-protein interactions for wet lab scientist’, Pasteur Institute, Paris, France.
5-6.10.2017 Training in the ÄKTA explorer system. GE Healthcare's ÄKTA™ User Day, Ołtarzew, Poland.
11-15.01.2021 EMBO Practical Course ‘Drosophila Genetics and Genomics’, Heidelberg, Germany, virtual.

7.2 Participation in conferences for training purposes

22-23.02.2022 6th Ulm Meeting “Biophysics of Amyloid Formation”, Ulm University, Ulm, Germany, virtual.
9-12.05.2022 EMBL Symposium ‘Cellular mechanisms driven by phase separation’, Heidelberg, Germany, virtual.
28-30.09.2022 neuroRNA Conference „RNA regulation in brain function and disease”, Institute of Bioorganic Chemistry of the Polish Academy of Sciences, Poznan, Poland, virtual.

7.3 Scientific activity as a reviewer and editor

I was invited to review multiple manuscripts submitted to the international journals with impact factor according to Journal Citation Report: *Scientific Reports* (IF_{2018} 4,01), *Insect Biochemistry and Molecular Biology* (IF_{2020} 4,714), *Journal of Insects Physiology* (IF_{2018} 2,862); *International Journal of Molecular Sciences* (IF_{2022} 5,6), *Biomolecules* (IF_{2020} 4,879), *Genes* (IF_{2022} 3,331), *Journal of Fungi* (IF_{2021} 5,72), *Clocks and Sleep* (IF_{2022} 3,1), *Journal of Fungi* (IF_{2021} = 5.724), *Applied Microbiology and Biotechnology* (IF_{2020} = 4,813), *Journal of Insect Science* (IF_{2019} = 1,325).

I was a guest editor of Special Issue ‘Protein Disorder and Phase Separation in Transcription’ International Journal of Molecular Sciences, 2021.

I was invited as a guest editor to the Jove Methods Collection journal, to which I invited younger colleagues Dr M. Kolonko-Adamska and Dr A. Tarczewska, who are also experts in the proposed topic 'Methods in Intrinsically Disordered Proteins Research'. The release is currently open for submission.

7.4 Invited publications

In 2019 invitation to Special Issues of International Journal of Molecular Sciences (IF_{2019}=4,18): "Cytoplasmic Sensors: Infection, Inflammation, and Immunity" - published article [P5] and "Functionally Relevant Macromolecular Interactions of Disordered Proteins 2019" - published articles [P8] and [A11].

In 2022 invitation to write an article on the hormonal regulation of insects for the special issue of a journal KOSMOS [A10].

7.5 Activities in the Scientific Committee of the Conference

- Initiator, supervisor of the Organizational Committee and Chair of the Scientific Committee of the International Conference agroBiotechnology for Better Future, 12 March 2022, virtual.
- Participation in the Scientific Committee of the IV National Scientific Conference "Primary and secondary metabolites of plants and fungi" 15 July 2021, virtual.
- Scientific Committee member of ChemBiotIc Conference 2021 and ChemBiotIc Conference 2023, Wroclaw University of Science and Technology, Wroclaw, virtual.
- Scientific Committee member of Student's Science Conferences in years 2015-2020 organized by Wroclaw University of Science and Technology, also Chair of biotechnology session of 14th Student's Science Conference (2016), Wroclaw University of Science and Technology, Wroclaw, stationary.
- Volunteer in the Organizing Committee of the I Polish Congress of Biotechnology, Wroclaw, Poland, 20-25 September 1999, stationary.

7.6 Other information

Winner of the WUST Primus Action for paper published in 2022.

Winner of the WUST Primus Action program 2 for papers published 2019 and 2020.

Member of the Polish Biochemical Society (PTBioch).

Member of the Polish Biophysical Society (PTBF).

Bronze Medal for Long Service.

Beata Greb-Markiewicz
(Applicant’s signature)