1 Identification and characterization of RNA-binding protein (RBP) substrates of cyclin-dependent kinase 12 (CDK12)

Partners: Lab. Dalibor Blažek (Peter Maník), CEITEC MU and Lab. Falk Butter, IMB Mainz

1.1 SUMMARY

Cyclin-dependent kinase 12 (CDK12) is a transcriptional CDK that was found mutated or deregulated in various types of cancer. Recently, several research groups used genome-wide transcriptomics methods to describe the importance of CDK12 activity in the regulation of RNA polymerase II (RNAPII) processivity on long, poly(A) site-rich genes1–4. However, apart from the RNAPII C-terminal domain (CTD), substrates of CDK12 remain mostly uncharacterized. I propose to use CDK12 analogue-sensitive (AS) cell lines and highly accurate quantitative phosphoproteomics including SILAC to characterize novel CDK12 substrates. I hypothesize that many of these substrates are RNA-binding proteins (RBPs) that may play an important role in CDK12-dependent transcription defect.

1.2 INTRODUCTION

Transcription is a complex process consisting of multiple steps including initiation, elongation, and termination. During mRNA transcription, hundreds of proteins interact either directly or indirectly with mRNA or RNAPII and assist with co-transcriptional splicing, RNA modifications and 3' end mRNA processing. CDK12 belongs to a group of transcriptional CDKs that regulate transcription at different steps by phosphorylating the CTD of RNAPII. It was proposed that CDK12 regulates RNAPII elongation1, processivity3 or splicing5 via phosphorylation of serine 2 (Ser2) in the CTD6,7.

Recently, several studies tried to elucidate the role of CDK12 using novel experimental tools and genome-wide methods1–4. Work from our lab has shown that CDK12 activity is crucial for the transcription of DNA repair and replication genes. Specifically, inhibition of CDK12 caused RNAPII processivity defect, which led to a loss of RNAPII on 3‘ ends of predominantly long, poly(A) site-rich genes. Interestingly, CDK12 inhibition did not affect global P-Ser2 levels but led to a shift of P-Ser2 peak into the gene body of CDK12-sensitive genes. The position of this peak coincided with the site where the RNAPII signal was lost1. Altogether, work of our and other labs have shown that CDK12 affects the processivity and elongation rate of RNAPII on the subset of genes. Despite increased interest in CDK12 in recent years and advances in the understanding of cellular functions of CDK12, the mechanism by which CDK12 activity affects subset of genes remains unclear. Besides the CTD, transcriptional CDKs phosphorylate other proteins to fulfil their roles during the transcription. However, since RNAPII CTD remains the only well-studied substrate of CDK12 the characterization of other interacting partners and substrates of CDK12 is essential to get insight into the molecular mechanism of CDK12-directed gene expression.

During my diploma thesis, I used a technique called BioID for the identification of CDK12-associated proteins. This method allowed me to identify proteins whose interaction with

CDK12 is only transient. LC-MS/MS analysis identified over a hundred candidate proteins. Interestingly, GO analysis has shown a high abundance of proteins that are involved in the regulation of mRNA processing and splicing. These data are consistent with the results from previous proteomic studies, that used affinity-purification MS to study CDK12 interacting partners8–10. Many of these proteins have RNA-binding properties, which led us to ask whether CDK12 kinase activity might regulate their association with mRNA. Aberrant association of RBPs to mRNA could result in defects during mRNA synthesis or processing, leading to defect during CDK12-dependent transcription elongation on a subset of long, exon- and poly(A) site-rich genes.

1.3 AIM

Identification and characterization of CDK12 substrates

To address this question I propose to employ quantitative phosphoproteomics techniques such as SILAC, to study the change in the phosphoproteome after CDK12 inhibition11. This would be done in collaboration with Dr Butter's lab that has considerable expertise in quantitative proteomics. I propose to use a cell line carrying alleles with AS CDK12. AS mutation does not affect kinase activity but allows fast and specific inhibition of kinase by ATP analogues. The data obtained from this experiment would be compared with the results of previous BioID experiments to identify high confidence candidate substrates. We expect a significant overlap between these datasets. Specific candidates would be further evaluated as CDK12 substrates using additional biochemical techniques such as in vitro kinase assays (IVKA). Specific Ser/Thr residues will be mutated to verify specific phosphorylation site. For putative CDK12 substrates that are known to directly bind mRNA, their association with RNA after CDK12 inhibition will be tested by iCLIP or complementary methods commonly used in our lab. In summary, these experiments will identify novel CDK12 substrates that will be starting point for our understanding of molecular mechanisms of CDK12-dependent gene expression.

1.4 WORK PROGRAMME

1.4.1 The experimental approach including key techniques or technologies

SILAC is a quantitative proteomics technique that is based on the direct addition of specific stable amino acids into the cell culture medium and is characteristic for its high quantitative accuracy and reproducibility in comparison to label-free quantification strategy (employed at CEITEC proteomics facility). Usage of distinct stable isotopes allows samples to be combined early during the first steps of experimental workflow and thus minimizes experimental error. The samples are analysed with LC-MS/MS and the quantification is based on the ratio of incorporated "heavy" isotope-labelled peptides to "light" unlabeled peptides[11].

1.4.2 What will be done by which partner and the expected synergies

During my stay in Dr Butter's lab, I will design and perform SILAC experiment or similar method such as dimethyl labelling using cell line carrying AS CDK12 alleles. In addition, I will learn the proper data analysis of proteomics datasets. Follow-up experiments such as mutagenesis, IVKA or iCLIP will be done in Blazek's laboratory at CEITEC MU.

1.4.3 The benefit of this research project to RNA research at CEITEC

RBPs are diverse protein class that regulate various processes including RNA splicing, transcription initiation, elongation and termination. RNA-protein interactions play vital roles in these processes and dysregulation of these interactions has been linked to cancers12 and other diseases. However, the functions of many RBPs remain largely unknown. Several proteomic studies of CDK12 including BioID from our lab has indicated enrichment of numerous RNA-processing factors, many of which can directly bind mRNA8–10. This project is focused on the identification of CDK12 substrates and since CDK12 is associated with various RNA-processing factors, some are likely direct substrates of CDK12. Identification of CDK12 substrates among RBPs would help to elucidate molecular mechanisms of CDK12 function in cellular processes. Additionally, this work would contribute to understanding how RBPs modulate the fate of RNA molecules in the cell. Furthermore, CEITEC RNA research community would obtain knowledge of quantitative proteomics data analyses.

1.4.4 Future collaborations or joint funding expected to result from work proposed

Dr Butter's lab has expertise in a wide range of quantitative proteomics techniques and quantitative data analyses. Such knowledge is missing in Blazek`s lab that can however offer extensive knowledge and experimental expertise in the field of transcriptional CDKs. Thus the collaboration is mutually beneficial and a chance to apply for mutual funding, for instance in GACR scheme is realistic once some preliminary data are generated.

1.5 BUDGET

Estimated expenses including the price of services and necessary materials for experiments are approximately 5000 €. These would cover materials required for sample preparation and the cost of chemicals needed for LC-MS/MS experiments as well as expenses associated with cell cultures such as plastics, special growth mediums and mass isotopes of amino acids.

In addition, we will apply for a secondment that will cover travel expenses and approximately one-month stay in IMB, Mainz, Germany.

1.6 REFERENCES

1. Chirackal Manavalan, A. P.; Pilarova, K.; Kluge, M.; Bartholomeeusen, K.; Rajecky, M.; Oppelt, J.; Khirsariya, P.; Paruch, K.; Krejci, L.; Friedel, C. C.; Blazek, D. CDK12 Controls G1/S Progression by Regulating RNAPII Processivity at Core DNA Replication Genes. EMBO Rep. 2019, 20 (9), e47592. https://doi.org/10.15252/embr.201847592.
2. Dubbury, S. J.; Boutz, P. L.; Sharp, P. A. CDK12 Regulates DNA Repair Genes by Suppressing Intronic Polyadenylation. Nature 2018, 564 (7734), 141–145. https://doi.org/10.1038/s41586-018-0758-y.
3. Krajewska, M.; Dries, R.; Grassetti, A. V.; Dust, S.; Gao, Y.; Huang, H.; Sharma, B.; Day, D. S.; Kwiatkowski, N.; Pomaville, M.; Dodd, O.; Chipumuro, E.; Zhang, T.; Greenleaf, A. L.; Yuan, G.-C.; Gray, N. S.; Young, R. A.; Geyer, M.; Gerber, S. A.; George, R. E. CDK12 Loss in Cancer Cells Affects DNA Damage Response Genes

through Premature Cleavage and Polyadenylation. Nat. Commun. 2019, 10 (1), 1757. https://doi.org/10.1038/s41467-019-09703-y.

4. Tellier, M.; Zaborowska, J.; Caizzi, L.; Mohammad, E.; Velychko, T.; Schwalb, B.; Ferrer-Vicens, I.; Blears, D.; Nojima, T.; Cramer, P.; Murphy, S. CDK12 Globally Stimulates RNA Polymerase II Transcription Elongation and Carboxyl-Terminal Domain Phosphorylation. Nucleic Acids Res. 2020, 48 (14), 7712–7727. https://doi.org/10.1093/nar/gkaa514.
5. Tien, J. F.; Mazloomian, A.; Cheng, S.-W. G.; Hughes, C. S.; Chow, C. C. T.; Canapi, L. T.; Oloumi, A.; Trigo-Gonzalez, G.; Bashashati, A.; Xu, J.; Chang, V. C.-D.; Shah, S. P.; Aparicio, S.; Morin, G. B. CDK12 Regulates Alternative Last Exon MRNA Splicing and Promotes Breast Cancer Cell Invasion. Nucleic Acids Res. 2017, 45 (11), 6698–6716. https://doi.org/10.1093/nar/gkx187.
6. Bösken, C. A.; Farnung, L.; Hintermair, C.; Merzel Schachter, M.; Vogel-Bachmayr, K.; Blazek, D.; Anand, K.; Fisher, R. P.; Eick, D.; Geyer, M. The Structure and Substrate Specificity of Human Cdk12/Cyclin K. Nat. Commun. 2014, 5 (1), 3505. https://doi.org/10.1038/ncomms4505.
7. Cheng, S.-W. G.; Kuzyk, M. A.; Moradian, A.; Ichu, T.-A.; Chang, V. C.-D.; Tien, J. F.; Vollett, S. E.; Griffith, M.; Marra, M. A.; Morin, G. B. Interaction of Cyclin-Dependent Kinase 12/CrkRS with Cyclin K1 Is Required for the Phosphorylation of the C-Terminal Domain of RNA Polymerase II. Mol. Cell. Biol. 2012, 32 (22), 4691–4704. https://doi.org/10.1128/MCB.06267-11.
8. Liang, K.; Gao, X.; Gilmore, J. M.; Florens, L.; Washburn, M. P.; Smith, E.; Shilatifard, A. Characterization of Human Cyclin-Dependent Kinase 12 (CDK12) and CDK13 Complexes in C-Terminal Domain Phosphorylation, Gene Transcription, and RNA Processing. Mol. Cell. Biol. 2015, 35 (6), 928–938. https://doi.org/10.1128/MCB.01426-14.
9. Cyclin-dependent kinase 12 increases 3’ end processing of growth factor-induced c-FOS transcripts - PubMed https://pubmed.ncbi.nlm.nih.gov/25384976/ (accessed Mar 21, 2021).
10. Bartkowiak, B.; Greenleaf, A. L. Expression, Purification, and Identification of Associated Proteins of the Full-length hCDK12/CyclinK Complex https://www.jbc.org/article/S0021-9258(20)57832-0/fulltext (accessed Mar 21, 2021).
11. Chen, X.; Wei, S.; Ji, Y.; Guo, X.; Yang, F. Quantitative Proteomics Using SILAC: Principles, Applications, and Developments. Proteomics 2015, 15 (18), 3175–3192. https://doi.org/10.1002/pmic.201500108.
12. Wang, E.; Lu, S. X.; Pastore, A.; Chen, X.; Imig, J.; Chun-Wei Lee, S.; Hockemeyer, K.; Ghebrechristos, Y. E.; Yoshimi, A.; Inoue, D.; Ki, M.; Cho, H.; Bitner, L.; Kloetgen, A.; Lin, K.-T.; Uehara, T.; Owa, T.; Tibes, R.; Krainer, A. R.; Abdel-Wahab, O.; Aifantis, I. Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell 2019, 35 (3), 369-384.e7. https://doi.org/10.1016/j.ccell.2019.01.010.