3rd NGP-Net SYMPOSIUM ON NON-GLOBULAR PROTEINS

August 28 – September 1, 2017

Košice, Slovakia

Book of Abstracts

cost
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY
NGP-net
Non-globular proteins in molecular physiopathology
ÚEF SAV

Scientific Committee

Kristina Djinovic-Carugo | Austria

Zsuzsanna Dosztanyi | Hungary

Arne Elofsson | Sweden

Andrey Kajava | France

Sonia Longhi | France

Sandra Macedo Ribeiro | Portugal

Vytautas Smirnovas | Lithuania

Peter Tompa | Belgium

Silvio Tosatto | Italy

Nevena Veljkovic | Serbia

Salvador Ventura | Spain

Local Organizing Committee

Zuzana Gazova | IEP SAS

Diana Fedunova | IEP SAS

Rostislav Skrabana | IoN SAS

Ondrej Cehlar | IoN SAS

Zuzana Bednarikova | IEP SAS

Igor Hrmo | IEP SAS

Katarina Ulicna | IEP SAS

Miroslav Gancar | IEP SAS

Jana Kubackova | IEP SAS

Invited Speakers

Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome

YLVA IVARSSON

Department of Chemistry - BMC, Uppsala University, Sweden

The disordered regions of the human proteome are enriched in short motifs serving as docking sites for peptide binding domains. Domain-motif interactions are crucial for the wiring of signaling pathways. These interactions are typically transient and difficult to capture through most conventional high-throughput methods. We therefore developed a novel approach for the large-scale profiling of domain-motifs interactions called Proteomic Peptide Phage Display (ProP-PD). In ProP-PD we combine bioinformatics, oligonucleotide arrays, peptide phage display and next-generation sequencing. This allows the interrogation of domain-motif interactions on a proteome-wide scale and the *de novo* motif discovery.

In our pilot experiment we generated two distinct phage libraries, one displaying all human C-terminal sequences and one displaying C-termini of known virus proteins. We used the ProP-PD libraries to identify interactions of human postsynaptic density *95/discs large/zonula occludens-1* (PDZ) domains. We successfully identified novel PDZ domain interactions of potential relevance to cellular signaling pathways and validated a subset of interactions with a high success rate. More recently, we created a ProP-PD library that displays peptides representing the disordered regions of the human proteome, which allows searching for domain-motif interactions in a unprecedented way. We validate our disorderome library against a range of peptide binding domains, which provides novel insights into their binding preferences and suggest interactions of potential biological relevance. A few key cases will be presented here. ProP-PD can be used to uncover protein-protein interactions of potential biological relevance in high-throughput experiments and provides information that is complementary to other methods. ProP-PD is scalable and can be developed to any target proteome of interest.

Evolution and function of disordered regions and linear motifs in viral proteomes

LUCÍA B. CHEMES

Institute for Biotechnological Investigations (IIB), University of San Martín, Argentina

Linear motifs are key elements of regulatory protein interaction networks. Viruses take advantage of this by evolving linear motifs (LMs) that hijack and disrupt the host network. Therefore, identifying relevant linear motifs and their binding strength is central to understanding host regulation and viral pathogenesis. However, the degenerate nature of linear motifs makes their identification challenging. Moreover, the identification of disordered domains and flexible regions within folded domains still remains a challenge and is an essential step for defining regions where functional LMs are located. First, we will discuss how the sequence evolution of Retinoblastoma (Rb) binding motifs reveals fine-tuning of binding affinity that establishes the ranking of binding for host and viral motifs. We also present an example where a single viral protein uses several binding motifs for achieving high affinity Rb-association. Secondly, we present a strategy for the identification of intrinsically disordered regions (IDRs) in Flavivirus proteomes. Here, a combined sequence and structural analysis reveals the presence of multiple flexible regions where functional LMs could be located, despite the low average levels of disorder in this viral family. Evolutionary analysis reveals strong conservation of disorder within the Flavivirus family, and also group-specific IDR regions.

Disordered CDK substrates act as multi-input signal processors to control the key decision points in the cell cycle

MART LOOG

Institute of Technology, University of Tartu, Tartu 50411, Estonia

The decision points between different cell fates involve systems that process alternative signals into binary choices. At the beginning of the cell cycle the decision to commit to division is made at the point where the cyclin-dependent kinase (Cdk) signal overrides its antagonists. In *Saccharomyces cerevisiae* the cues controlling the commitment to the *Start* decision are the mating pheromone-induced signal via MAPK Fus3 and the *Start* signals via G1- and S-CDK complexes. These three signals are integrated by a multi-phosphorylated and disordered Cdk1 inhibitor protein Far1 that serves as a signal processor molecule to calculate the thresholds for the commitment switch. The kinase signals are processed by a network of phosphorylation and docking motifs to form a double-negative feedback loop. At the next decision point of the cell cycle, the G1/S transition, another disordered Cdk1 inhibitor protein Sic1 inhibits the S-phase specific Clb5-Cdk1 complex, while Clb5-Cdk1 also targets Sic1 for destruction by phosphorylation. We show that the timing and the shape of Sic1 degradation switch is controlled by a similar principle of a multi-branched processor. Rewiring the multisite phosphorylation networks in Far1 and Sic1 can alter the timing and shape of both *Start* and G1/S switches. In conclusion, intrinsically disordered phospho-regulated proteins can serve as sophisticated signal processors controlling key cell cycle transitions. Moreover, the idea that integration and processing of multiple input signals can be performed by a single unstructured protein is opposed to the current thought that such signal processing requires entire genetic circuits or pathways. Similar principles of sequential signal processing via multisite phosphorylation networks can be applied to synthetic circuit design.

Accurate structural modeling of peptide-protein interactions using Rosetta FlexPepDock

[a] NAWSAD ALAM, [b] DIMA KOZAKOV, [a] ORA SCHUELER-FURMAN

[a] *Department of Microbiology and Molecular Genetics, Institute of Medical Research Israel-Canada, The Hebrew University of Jerusalem, Israel*

[b] *Department of Applied Mathematics and Statistics, SUNY Stony Brook, NY, USA*

Many partly or fully intrinsically unstructured proteins use short linear stretches to interact with their partners [1]. Such peptide-mediated interactions play major regulatory roles in the cell, and their accurate modeling is of primary importance for the structure-based characterization and modulation of protein communication. Moreover, such models can improve our understanding of the basic principles that underlie the process of protein-protein recognition.

We have previously developed Rosetta FlexPepDock [2,3], a protocol for the accurate refinement and modeling of peptides into receptor binding sites, starting from an approximate peptide structure or location, and used the structural models to identify new binding partners for given peptide binding receptors and enzymes.

I will present recent significant advances to the protocol that allow its effective application to the challenging problem of *global* peptide docking, where only the receptor structure and the peptide sequence are given, but no information about the receptor binding site nor the peptide structure (or sequence binding motif) is available. Our PIPER-FlexPepDock implementation draws from the observation that the bound peptide conformation can be sampled by extracting fragments of similar (predicted) structure and sequence from protein monomer structures, similar to the fragment libraries used for *ab initio* monomer structure prediction in Rosetta [4]. By combining an efficient, exhaustive rigid body search, of the fragment set (using PIPER FFT-based docking [5]) with subsequent detailed flexible refinement (using Rosetta FlexPepDock refinement), we are able to sample the full energy landscape of peptide-protein association and generate models of unprecedented accuracy.

By relying on structural rather than sequence information, binding of novel or unknown sequence motifs is possible, considerably extending the range of application. The success of this approach suggests that the peptides sample the bound conformation – at least in the protein-like encounter complex regime, similar to globular protein-protein association.

Acknowledgement: Financial support from the European Research Council (ERCstG 310873 to O.S F), the USA–Israel Binational Science Foundation (BSF 2009418,2015207 to O.S-F and DK), and NSF (DBI 1458509, AF 1527292, to DK) are acknowledged.

References

1. Petsalaki, E. & Russell, R.B. Peptide-mediated interactions in biological systems: new discoveries and applications. Curr Opin Biotechnol 19, 344-350 (2008).
2. Raveh, B., London, N. & Schueler-Furman, O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins 78, 2029-2040 (2010).
3. Raveh, B., London, N., Zimmerman, L. & Schueler-Furman, O. Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. PLoS One 6, e18934 (2011).
4. Gront, D., Kulp, D.W., Vernon, R.M., Strauss, C.E. & Baker, D. Generalized fragment picking in Rosetta: design, protocols and applications. PLoS One 6, e23294 (2011).
5. Kozakov, D., Brenke, R., Comeau, S.R. & Vajda, S. PIPER: an FFT-based protein docking program `with pairwise potentials. Proteins 65, 392-406 (2006).

Understanding amyloid: from structure to function

[a] KAREN MARSHALL, [a] DEVKEE VADUKUL, [a] KYLE MORRIS,
[a] ZAHRAA AL-GARAWIYOUSSRA AL-HILALY, [a] LENZIE FORD,
[a] MAHMOUD BUKAR MAINA, [a] LOUISE SERPELL

[a] School of Life Sciences, University of Sussex, Falmer, Brighton, BN1 9QG, UK
L.C.Serpell@sussex.ac.uk

Many proteins and peptides with different primary sequences share the ability to self-assemble to form amyloid fibrils. A large number of these have been implicated in protein misfolding diseases but many perform functional roles in living organisms. Amyloid fibrils are also being utilised to form functional materials, highlighting the need to better understand the structure-function relationship.

It is clear that the sequence of amyloidogenic peptides determines the ability to self-assemble and in turn, this relates to the formation of toxic oligomers and mature amyloid fibrils. Our work reveals insights into the mechanism of amyloid-induced toxicity, providing a platform to better understand deleterious effects of oligomeric proteins in disease and how amyloid fibrils may be controlled for functional and non-toxic roles.

References

1. Marshall KE, Marchante R, Xue WF, Serpell LC: The relationship between amyloid structure and cytotoxicity. Prion 2014, 8(2):28860.
2. Marshall KE, Vadukul DM, Dahal L, Theisen A, Fowler MW, Al-Hilaly Y, Ford L, Kemenes G, Day IJ, Staras K et al: A critical role for the self-assembly of Amyloid-beta1-42 in neurodegeneration. Scientific reports 2016, 6:30182.
3. Al-Hilaly YK, Williams TL, Stewart-Parker M, Ford L, Skaria E, Cole M, Bucher WG, Morris KL, Sada AA, Thorpe JR et al: A central role for dityrosine crosslinking of Amyloid-beta in Alzheimer's disease. Acta neuropathologica communications 2013, 1(1):83.
4. Morris KL, Rodger A, Hicks MR, Debulpaep M, Schymkowitz J, Rousseau F, Serpell LC: Exploring the sequence-structure relationship for amyloid peptides. The Biochemical journal 2013, 450(2):275-283.

Multivalency in Transcriptional Regulation of the Hub Protein LC8

[a] SARAH A. CLARK, [b] RADOVAN FIALA, [b] JIRI NOVACEK, [a] ELISAR BARBAR

[a] *Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331*

[b] *Central European Institute of Technology, Masaryk University, Kamenice 753/5, 62500 Brno, Czech Republic*

Hub proteins bind a large number of partners to facilitate structural changes and downstream protein interactions. LC8, a highly conserved 20.6 kDa protein homodimer (10.3 kDa monomer), is a unique hub that acts as a chaperone by dimerizing its primarily disordered partners [1,2]. LC8 binding is associated with a range of cellular processes, from cell division to apoptosis, underscoring LC8's essential role as a regulatory hub. In recent years, the number of experimentally verified LC8 partners has risen to more than 40, and prediction methods indicate that dozens more may specifically bind LC8 [3]. Therefore, gaining insight into how LC8 interacts with partner proteins, and how LC8 levels in the cell are balanced, is paramount to understanding the regulation of many cellular processes. ASCIZ, an 88 kDa protein initially discovered as a component of the DNA damage response has recently been identified as a transcription factor for LC8. Mice with mutations in ASCIZ that prevent LC8 transcription die in late embryogenesis and exhibit serious developmental defects in kidneys and lungs. Here we show that the hub protein LC8 and its transcription factor, ASCIZ, participate in a unique method of feedback regulation that relies on binding to multiple recognition motifs in ASCIZ's intrinsically disordered domain. Even though intrinsic disorder is prevalent among eukaryotic transcription factors, its functional relevance in these systems remains unclear. Using a combination of structural and cell biology, we propose a model that explains the functions relevance of disorder. We describe how disorder in the transcription factor ASCIZ presents multiple short linear motifs to regulate transcription of the hub protein LC8. Using isothermal titration calorimetry, analytical ultracentrifugation, and NMR, we identify seven LC8 recognition motifs in Drosophila ASCIZ that bind LC8 with cooperativity modulated by the position of motifs in the sequence rather than their sequence specificity. ASCIZ and LC8 form a dynamic mixture of complexes which, instead of acting as simple on/off switch, facilitate gradual activation of varying levels of LC8 transcription. These data along with cellular transcription assays of human ASCIZ support a model by which complexes with different LC8 occupancies can sense LC8 concentration and efficiently tune transcription levels to meet cellular needs.

Acknowledgement: This work was supported by National Institutes of Health Grant GM 084276 to E.B. Support to facilities includes the Oregon State University NMR Facility funded in part by the National Institutes of Health, HEI Grant 1S10OD018518, and by the M. J. Murdock Charitable Trust grant # 2014162. Access to the NMR facility of CEITEC Masaryk University was provided by iNEXT, project number 653706, funded by the Horizon 2020 programme of the European Union.

References

1. Barbar, E. (2008). Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. Biochemistry, 47(2), 503-508.
2. Clark, S. A., Jespersen, N., Woodward, C., & Barbar, E. (2015). Multivalent IDP assemblies: Unique properties of LC8-associated, IDP duplex scaffolds. FEBS Lett, 589(19 Pt A), 2543-2551.
3. Rapali, P., Radnai, L., Suveges, D., Harmat, V., Tolgyesi, F., Wahlgren, W. Y.,… Pal, G. (2011). Directed evolution reveals the binding motif preference of the LC8/DYNLL hub protein and predicts large numbers of novel binders in the human proteome. PLoS One, 6(4), e18818.

Repetitive bacterial surface proteins

JENNIFER R. POTTS

Department of Biology, University of York, UK

Cell wall-attached proteins of Gram positive bacteria play important roles in infection including the formation of biofilms on medical devices, an important clinical problem. Several of these proteins have a common organisation that includes a highly repetitive central region. We are using a variety of biophysical and structural techniques including NMR spectroscopy, X-ray crystallography, small angle X-ray scattering and calorimetry [1,2] to characterise these repetitive proteins.

References

1. Gruszka, D. T., Wojdyla, J. Bingham, R. J., Turkenburg, J.P., Manfield, I., Steward A., Leech, A.P., Geoghegan, J.A., Foster, T.J., Clarke, J. and Potts, J.R.* A staphylococcal biofilm-forming protein has a contiguous rod-like structure. Proc. Natl. Acad. Sci. (USA) 109, E1011-1018 (2012).
2. Gruszka, D. T., Whelan, F., Farrance, O. E., Fung, H.K.H., Paci, E., Jeffries, C.M., Svergun, D.I., Baldock, C., Baumann, C.G., Brockwell, D.J., Potts, J.R., Clarke, J. Cooperative folding of unstable domains drives formation of a long, strong protein. Nature Commun. 6, 7271 (2015).

Principles of protein structural ensemble determination

MICHELE VENDRUSCOLO

Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

The biological functions of protein molecules are intimately dependent on their conformational dynamics. This aspect is particularly evident for disordered proteins, which constitute perhaps one-third of the human proteome. Therefore, structural ensembles often offer more useful representations of proteins than individual conformations. Here, we describe how the well-established principles of protein structure determination should be extended to the case of protein structural ensembles determination. These principles concern primarily how to deal with conformationally heterogeneous states, and with experimental measurements that are averaged over such states and affected by a variety of errors. We first review the growing literature of recent methods that combine experimental and computational information to model structural ensembles, highlighting their similarities and differences. We then address some conceptual problems in the determination of structural ensembles and define future goals towards the establishment of objective criteria for the comparison, validation, visualization and dissemination of such ensembles.

References

1. M. Bonomi, G. T. Heller, C. Camilloni and M. Vendruscolo. Principles of protein structural ensemble determination. Curr. Op. Struct. Biol. 42, 106-116 (2017).
2. P. Sormanni, D. Piovesan, G. T. Heller, M. Bonomi, P. Kukic, C. Camilloni, M. Fuxreiter, Z. Dosztanyi, R. Pappu, M. M. Babu, S. Longhi, P. Tompa, A. K. Dunker, V. N. Uversky, S. C. E. Tosatto and M. Vendruscolo. Simultaneous quantification of order and disorder in proteins.Nat. Chem. Biol. 13, 339-342 (2017)

Binding without folding: Extreme disorder and dynamics in a high-affinity protein complex

BENJAMIN SCHULER

University of Zurich, Switzerland

The functions of proteins have traditionally been linked to their well-defined three-dimensional, folded structures. It is becoming increasingly clear, however, that many proteins perform essential functions without being folded. Single-molecule spectroscopy provides new opportunities for investigating the structure and dynamics of such unfolded or 'intrinsically disordered' proteins (IDPs). The combination of single-molecule Förster resonance energy transfer (FRET) with nanosecond correlation spectroscopy, microfluidic mixing, and related methods can be used to probe intra- and intermolecular distance distributions, reconfiguration dynamics, and interactions on a wide range of timescales, and even in heterogeneous environments, including live cells.

Selected Talks

At the cross-road between structured and disordered binding motifs: the Ubiquitin Interacting Motives

[a,b] MATTEO LAMBRUGHI, [a] BURCU AYKAC FAS, [c] GARY S. SHAW,
[b] GAETANO INVERNIZZI, [a,d] ELENA PAPALEO

[a] *Computational Biology Laboratory, Strandboulevarden 49, 2100, Copenhagen, Denmark*
[b] *University of Milano-Bicocca, Department of Biotechnology and Biosciences, Milano, Italy*
[c] *Department of Biochemistry, Schulich School of Medicine and Dentistry, The University of Western Ontario, London, Ontario, Canada, N6A 5C1*
[d] *Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Copenhagen, Denmark*

Proteins containing intrinsically disordered segments (IDSs) are now recognized as key effectors in the biological processes of the cell [1]. They are central elements in signaling hubs, mediating the dynamic interaction with multiple macromolecules and orchestrating cellular pathways [2]. Notably binding motives are frequently embedded in IDSs [3] and, together with their heterogeneous structural biology, make them crucial to modulate the functions of proteins. We are using cutting-edge Molecular Dynamics simulations together with enhanced sampling approaches and integrated by NMR experiments to characterize the structural biology of the disordered regions of the ataxin-3 (AT3). AT3 is a multidomain deubiquitinating enzyme that is involved in aggregation processes and development of neurodegenerative disease and in cancer [4]. The active site of AT3 is located in the folded N-terminal Josephin domain, followed by a long C-terminal disordered segment contains multiple Ubiquitin (Ub) binding motives (UIMs). The IDS of AT3 is proposed to act in the recognition and recruitment of the substrates to the catalytic site of the enzyme and to play important roles in disease associated mechanisms. Despite this, the molecular mechanisms of the interplay between folded and disordered segments in AT3 are largely unknown. We are investigating the IDS of AT3 both in free and in complex with Ub and other binding partners as well as in mutated versions to understand how they affect their structural biology and how this regulates the activity of the protein.

In this contribution, we redefine the signature and structural properties of the UIM motifs for interaction with Ub. Despite the structural portrait provided by X-crystallography in which these motifs have been always ascribed to stable helical structures, we here show that UIM motifs can be classified in canonical and helical motifs and non-canonical and intrinsically disordered ones. This new class of disordered UIM-like proteins, as exemplified by the Cterminal IDSs of AT3, is particularly intriguingly since it is expected to binding Ub with low affinity or not to bind Ub at all. These motifs in IDSs can thus provide versatility and new functions to the proteins that contain them and they open the venue for new research of their interactome. Our results show interesting molecular mechanisms and their alterations, paving the way for future project aimed to use disordered segments as therapeutic targets for the design of new drugs.

References

1. Wright P.E. and Dyson H.J., Intrinsically Disordered Proteins in Cellular Signaling and Regulation Nat. Rev. Mol. Cell Biol. 2015, 16: 18–29.
2. Csizmok V., Follis A.V., Kriwacki R.W., Forman-Kay J.D., Dynamic Protein Interaction Networks and New Structural Paradigms in Signaling. Chem. Rev. 2016, 116:6424-62.
3. Tompa P., Davey N.E., Gibson T.J., Babu M.M. A million peptide motifs for the molecular biologist. Mol. Cell. 2014, 55:161–169.
4. Sacco J.J., Yau T.Y., Darling S., Patel V., Liu H., Urbé S., Clague M.J., Coulson J.M., The deubiquitylase Ataxin-3 restricts PTEN transcription in lung cancer cells. Oncogene. 2014, 33:4265-72.

The relationship between folding and activity in UreG, an intrinsically disordered enzyme

[a] MARTA PALOMBO, [b] ALESSIO BONUCCI, [b] EMILIEN ETIENNE, [a] STEFANO CIURLI, [c] VLADIMIR N. UVERSKY, [b] BRUNO GUIGLIARELLI, [b] VALÉRIE BELLE, [b] ELISABETTA MILEO, [a] BARBARA ZAMBELLI

[a] *Laboratory of Bioinorganic Chemistry, Department of Pharmacy and Biotechnology, University of Bologna*
[b] *Aix-Marseille Univ, CNRS, IMM (FR 3479), BIP (UMR 7281), 31 chemin Joseph Aiguier 13402 Marseille*
[c] *Department of Molecular Medicine, University of South Florida, 12901 Bruce B. Downs Blvd., MDC07, Tampa*

Nickel delivery into the active site of urease, an essential enzyme for plants, fungi and bacteria, requires the presence of four accessory proteins, named UreD, UreF, UreG and UreE [1]. UreG, responsible for hydrolysis of GTP, which regulates the complex interaction network, presents a flexible fold that makes this protein the first discovered intrinsically disordered enzyme [2]. This protein likely uses disorder to couple its catalytic function to a non-enzymatic role as a molecular chaperone, undergoing disorder-to-order transition upon interaction with its partners. Isolated UreG proteins from different biological sources were recognized as collapsed disordered proteins under native conditions, existing as ensembles of interconverting conformations with significant degree of secondary and tertiary structure [3, 4, 5, 6].

With the aim to address the importance of dynamics for catalysis, we investigated the relationship between folding and activity in *Sporosarcina pasteurii* UreG (*Sp*UreG). The effect of denaturants and osmolytes on protein structure and activity was analyzed using circular dichroism (CD), Site-Directed Spin Labeling (SDSL) coupled to EPR spectroscopy, and enzymatic assays. Results show that the activity of UreG is not directly correlated to the degree of protein rigidity. Rather, they suggest that UreG needs a "flexibility window" to be catalytically competent, with both too low or to high mobility being detrimental for its activity. These findings constitute a step forward toward the understanding of the role of dynamics for enzyme function and evolution, showing potential applications for enzyme design.

References

1. Zambelli, B., Musiani, F., Benini, S. & Ciurli, S., *Acc. Chem. Res.* 44, 520-530 (2011)
2. Zambelli, B. et al. *J. Biol. Chem.* 280, 4684-4695 (2005)
3. Zambelli, B., Musiani, F., Savini, M., Tucker, P. & Ciurli, S. *Biochemistry* 46, 3171-3182 (2007)
4. Zambelli, B., Turano, P., Musiani, F., Neyroz, P. & Ciurli, S. *Proteins* 74, 222-239 (2009)
5. Real-Guerra, R. et al. *Plant Mol. Biol.* 78, 461-475 (2012)
6. Miraula, M., Ciurli, S., Zambelli, B. *J. Biol. Inorg. Chem.* 20, 739-755 (2015)

Cross-talk between redox regulation and protein homeostasis

[a] ODED RIMON, [a] OHAD SUSS, [a] RAVIT MASIKA, [a] RESHEF MINZ, [a] ROSI GILLIN, [a] MEYTAL RADZINSKI, [a] OHAD YOGEV, [a] DANA REICHMANN

[a] *Department of Biological Chemistry, The Hebrew University of Jerusalem, Jerusalem, Israel*

Why do we age? How do we cope with environmental and physical changes? What makes pathogens smarter than their host? One of the biological processes linking all these questions relates to the amazing ability of cells to respond to changes in the environment, such as changes in oxidant levels, and protect its proteome against damage.

Cellular defense depends heavily on redox-regulated proteins, redox switch proteins, which play fundamental roles not only during cellular defense but also in signaling, and protein biogenesis.

Here I will speak about conditionally disordered chaperone, Hsp33 (in bacteria) and TrypOx (in T.brucei pathogen) that use their redox status to protect cellular proteome during oxidative stress. These chaperones need to lose their structure in a redox-dependent manner to become active as chaperones. By modifying Hsp33 sequence we reveal that the metastable region of Hsp33 has evolved to abolish redox-dependent chaperone activity, rather than enhance binding affinity for client proteins. The intrinsically disordered region of Hsp33 serves as an anchor for the reduced, inactive state of Hsp33, and dramatically affects the crosstalk with the synergetic chaperone system, DnaK/J. By using mass spectrometry, we describe the role that the metastable region plays in determining client specificity during normal and oxidative stress condition in the cell. Our results suggest that Hsp33 might serve as a member of the house-keeping proteostasis machinery, tasked with maintaining a “healthy” proteome during normal conditions, and that this function does not depend on the metastable linker region.

Moreover, I will describe a methodology we have developed to characterize interactome of thermobile proteins in cells. We had applied in-vivo crosslinking methodology coupled with mass spectrometry to define interactome of thermobile and stable variants of Luciferase protein in E.coli.

References

1. O. Rimon et al., Antioxid Redox Signal, 2017
2. O. Suss and D. Reichmann, *Front. Mol. Bios*, 2015
3. D. Reichmann et al., *Cell*, *2012*

Comparision of methods for searching for similar protein sequences in the case of low complexity regions

[a] ALEKSANDRA GRUCA, [a] PATRYK JARNOT, [b] DARIUSZ PLEWCZYNSKI, [c] MARCIN GRYNBERG

[a] *Institute of Informatics, Silesian University of Technology, Gliwice, Poland*
[b] *Centre of New Technologies, University of Warsaw, Warsaw, Poland*
[c] *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland*

There are many well-established methods, known from the literature, that are suitable for searching for similar protein sequences in the dataset of interest. Usually, on the basis of the results of similarity searching we can conclude about the properties and functions of a protein. However, the statistical models on which these methods are based are typically designed for "normal" protein sequences that are characterized by high complexity of amino acids composition. Therefore, it is questionable if these methods can be used to find similar low complexity regions (LCRs) coming from very different proteins in an attempt to gain new insights into LCR characteristics.

In the present study we analyze if and how these well-known methods of searching for similar sequences could be applied to the dataset that include only protein low complexity fragments. In the first step we identify all low complexity regions in the proteins from the SwissProt database using the SEG [1] method. Then we search for pairs of similar LCR fragments using three different methods: CD-HIT-2D [2], BLAST [3] and HHblits [4]. We compare the results obtained with the selected methods, the number of similar pairs that were found and the overlap among similar LCR pairs identified by different methods. As we are especially interested in such similar LCRs that come from dissimilar proteins, for each identified pair we also compare the similarity among their corresponding high complexity fragments. In this work we show that both HHblits and CH-HIT-2D outperform BLAST in quality of comparisons.

References

1. Wootton, J. C., and Federhen, S. (1993) Statistics of local complexity in amino acid sequences and sequence databases. Computers & Chemistry, 17(2), 149–163.
2. Li, W. and Godzik, A. (2006). CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22:1658-1659.
3. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden TL. (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421
4. Remmert M., Biegert A., Hauser A., Söding J. (2011) HHblits: Lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods. 9(2):173-175.

Thermodynamic origin of fuzzy interaction between intrinsically disordered antitoxin and ordered toxin

[a,b] SAN HADŽI, [a] ANDREJ MERNIK, [a] ČRTOMIR PODLIPNIK, [b] REMY LORIS, [a] JURIJ LAH

[a] *Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia*
[b] *Department of Bioengineering Sciences, Vrije Universiteit Brussel and Structural Biology Research Center, Vlaams Instituut voor Biotechnologie, Brussels, Belgium*

Intrinsically disordered proteins (IDPs) are often involved in protein-protein association that is accompanied by their folding. In the unbound state IDPs may exhibit transient secondary structure while in the bound state some of them retain a considerable dynamics and conformational heterogeneity that is typical for fuzzy complexes [1]. Both, the level of transient secondary structure in the unbound state and the conformational heterogeneity in the bound state, could be varied by changing the IDP amino acid sequence. Here we will show how these changes affect thermodynamics of IDP folding and binding.

Thermodynamic analysis of interactions was performed on a model system consisting of the antitoxin CcdA (IDP) and its folded target toxin CcdB [2, 3]. Such interactions are crucial for the functionality of bacterial genetic systems called toxin-antitoxin modules [4]. Effect of IDP mutations on the IDP-target association was monitored by ITC and CD-spectroscopy and analyzed by molecular modeling. We will demonstrate how the accompanying changes of energetics may be dissected into folding and binding contributions using structural information. We will show that the differences in IDP folding propensity manifest themselves mainly through the energetics of the bound-state structures, where unfavorable folding leads to a less ordered bound-state, and suggest the thermodynamic origin of this increased 'fuzzines'.

Acknowledgement: This work was supported by Slovenian Research Agency [Grant P1–0201].

References

1. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci.* 2008, 33, 2.
2. Drobnak I, De Jonge N, Haesaerts S, Vesnaver G, Loris R & Lah J. Energetic basis of uncoupling folding from binding for an intrinsically disordered protein. *J Am Chem Soc.* 2013, 135, 1288.
3. Vandervelde A, Drobnak I, Hadži S, Sterckx YG, Welte T, De Greve H, Charlier D, Efremov R, Loris R & Lah J. Molecular mechanism governing ratio-dependent transcription regulation in the *ccdAB* operon. *Nucleic Acids Res.* 2017, 45, 2937.
4. Buts L, Lah J, Dao-Thi MH, Wyns L & Loris R. Toxin-antitoxin modules as bacterial metabolic stress managers. *Trends Biochem Sci.* 2005, 30, 672.

Amplification of subdomain-sized fragments in the genesis of new, non-globular folded proteins

ANDREI N. LUPAS

Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany; andrei.lupas@tuebingen.mpg.de

Protein folding is a complicated, easily disrupted process, and even after 4 billion years of evolution, protein misfolding and aggregation continue to be major challenges for all living beings. Despite these problems, natural proteins nevertheless represent a best-case group, as in their overwhelming majority polypeptides do not appear to have a folded structure at all. Unsurprisingly, except in viruses, proteins emerged *de novo* are largely or entirely unstructured. New structured proteins do however emerge continually in nature from pieces of existing proteins. The dominant mechanism for this is repetition and the resulting proteins are almost invariably solenoids of fibers. Here we will review examples of this process at work in the surface proteins of bacteria.

A novel approach to enable the structural and dynamics characterisation of homorepeat proteins

[a] A. URBANEK, [a] A. MORATÓ, [a] F. ALLEMAND, [a] A. FOURNET, [a] E. DELAFORGE, [b] A. I. JIMÉNEZ, [b] C. CATIVIELA, [c] S. DELBECQ, [a] N. SIBILLE, [a] P. BERNADÓ

[a] *Centre de Biochimie Structurale, CNRS UMR 5048 – INSERM 1054 – UM, Montpellier, France*
[b] *Department of Organic Chemistry, University of Zaragoza, Zaragoza, Spain*
[c] *Faculté de Pharmacie, Université de Montpellier, Montpellier, France*

While most protein sequences are aperiodic and feature most of the 20 proteinogenic amino acids, many proteins harbour low complexity regions (LCRs). LCRs are strikingly simple sequences with a strong bias in their amino acid composition [1]. In their simplest form, the so-called homorepeats (HRs), the sequences consist of repetitions of a single amino acid. Interestingly, HRs are quite common in the human proteome and ~20% of human proteins contain at least one HR of 5 or more residues [2]. While some authors postulated that HRs lack any particular function, they are highly enriched in specific functional classes of proteins (e.g. transcription factors and developmental proteins) and can directly modulate proteinprotein interactions and transcription. Most importantly, >40% of proteins with multiple HRs are key players in human diseases, including Huntington's disease, several ataxias (poly-Q) and synpolydactyly syndrome (poly-A) [1, 2].

Unfortunately, the underlying disease mechanisms are currently poorly understood. This is mostly due to the technical challenges encountered in high-resolution structure/function studies of HR regions. HRs are often found in intrinsically disordered regions, precluding crystallisation, and the assignment of NMR spectra is complicated by the chemical similarity and thus poor peak dispersion of the residues within the repeats. Our work aims to overcome the current limitations of HR structure characterisation by simplifying the NMR spectra of HR proteins by site-specifically incorporating a single 15N/13Clabelled amino acid into the poly-Q and poly-P regions of the model protein huntingtin. To this end, we supplement our cell-free reaction [3] with loaded nonsense suppressor tRNAs [4] that we prepare using either the respective aminoacyl tRNA synthetase or the de novo ribozyme flexizyme [5]. Following this strategy of labelling one individual residue of the HR tract per sample, we have been able to record greatly simplified NMR spectra that enabled us to unambiguously monitor structural and dynamic features of that residue in the context of the HR tract. By integrating all data obtained on HR tract residues we are now able to decipher local structural propensities, structural cooperativity and possible intra-molecular interactions.

Our approach opens up the large family of HR-containing proteins to high-resolution structural biology, a necessary step to understand the bases of multiple pathologies and to guide drug design.

Acknowledgement: This project has received funding from the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement No 648030) and from the Fondation pour la Recherche Médicale (FRM, SPF20150934061).

References

1. Jorda J., Kajava A.V. *Adv Protein Chem Struct Biol*, 2010, 79, 59–88.
2. Karlin S., Brocchieri L., Bergman A., Mrazek J., Gentles A.J. *Proc Natl Acad Sci USA*, 2002, 99,333–338.
3. Kigawa T., Yabuki T., Yoshida Y., Tsutsui M., Ito Y., Shibata T., Yokoyama S. *FEBS Lett*, 1999, 442, 15–19.
4. Liu C.C., Schultz P.G. *Annu Rev Biochem*, 2010, 79, 413–444.
5. Goto Y., Katoh T., Suga H. *Nat Protoc*, 2011, 6, 779–790.

Establishment of constraints on amyloid formation imposed by steric exclusion of globular domains

[a,b] R.A. AZIZYAN, [c] A. GARRO, [d] A. ANIKEENKO, [d] A. BAKULINA, [e] C. DUMAS, [a,b] A.V. KAJAVA

[a] *CRBM, CNRS, Montpellier, France*

[b] *Institut de Biologie Computationnelle, Université Montpellier, Montpellier, France*

[c] *Universidad Nacional de San Luis IMASL-CONICET, San Luis, Argentina*

[d] *Novosibirsk State University, Novosibirsk, Russia*

[e] *Centre de Biochimie Structurale, CNRS, Montpellier, France*

Amyloid fibrils are the subject of special interest due to their link to a broad range of human diseases. In the majority of cases, within the same protein, the amyloid-forming regions coexist with the other sequence motifs including intrinsically unstructured regions and structured domains. It is apparent that these regions flanking the amyloidogenic regions can affect formation of amyloids. For example, in protein with a very short linker between amyloidogenic region and globular domain, the steric repulsion of the globular structures can prevent formation the amyloid fibrils [1]. However, despite considerable interest and importance this issue is poorly understood. This situation may be attributed to the fact that the examination of these effects by experimental methods is hampered by high polymorphism of amyloid structures which brings an uncertainty in our knowledge about the beginning and end of the amyloidogenic regions and linker in a given experiment. In this situation theoretical modeling approaches are becoming essential to understand these effects. Furthermore, the modelling that uses the steric tension as a criteria can give us highly trustworthy results.

Thus, the objective of this work was to evaluate the steric effect of the neighboring globular domains on amyloid formation by using molecular modeling, molecular dynamics and mesoscopic protein models [2]. In particular, we tested a hybrid protein containing an amyloid-forming fragment of Aβ peptide (17-42) and GFP as a globular domain. To find the shortest possible linker for the infinite amyloid fibril of Aβ-linker-GFP molecule we constructed a model with pseudo-helical arrangements of the densely packed GFPs around the Aβ amyloid core and searched for the stereochemically allowed linkers. We also used Targeted Molecular Dynamics approach to model the fibrillogenesis. As a result, we established the linkers of 7 and more residues allow fibrillogenesis of Aβ-linker-GFP molecules. We also were able to establish a more general relationship between the size of the globular domain and the length of the linker by using mesoscopic protein modeling.

The results of this work can provide the fundamental understanding of the molecular mechanisms of amyloidogenesis and also can be used in the interpretation of the experimental results, improvement of the prediction of amyloidogenic regions in proteins, inhibition of amyloid formation and design of nanostructures.

Acknowledgement: The authors acknowledge Armenian Communities Department of Calouste Gulbenkian Foundation for providing "Global Excellence Scholarship" to R.A. This work was also supported by Erasmus Mundus fellowship to A.G. and made within COST Action BM1405.

References

1. Kajava A. V., Baxa U., Wickner R. B., Steven A. C. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 7885–7890
2. R. Azizyan, C. Dumas, A. Garro, A. Bakulina, A. Anikeenko and A.V. Kajava (2017) (in preparation).

Engineered Bacterial Functional Amyloids

URARTU SEKER

UNAM-Institute of Materials Science and Nanotechnology, Bilkent University, Bilkent Ankara Turkey 06800

Bacterial amyloids are generally formed by more than one protein, in E. coli (CsgA and CsgB), in Pseudomonas (FapB and FapC) or in Bacillus (TasA and TapA) [1]. These protein are interacting each other and forming the matrix materials of the biofilm through molecular self-assembly. Biofilm proteins are offering many possibilities to be used a molecular building blocks to form new generation of material systems. These proteins can be functionalized with chemical groups through protein engineering strategies resulting in functional self-assembled entities. We aimed to program CsgA and CsgB proteins to form strategies for the utilization of these proteins as functional material systems. We followed two different routes, in the first approach we used the purified proteins of CsgA and CsgB (CsgA is the major biofilm protein whereas CsgB is the minor protein). We exploit the assembly properties of CsgA and CsgB proteins to form ordered material systems in vitro. Biofilm proteins offer opportunities as cheap, useful, easy-to-synthesize biomaterials that can be functionalized through protein engineering. In the second approach we programmed cellular circuits to secrete biofilm proteins using recombinase protein based genetic logic gates to control the morphology of the final biofilm structures. Additionally we presented how to genetically control the mechanical properties of these non globular protein assemblies in vivo [2]. These enabled us to control the final material properties of a self-assembled protein based nanowire systems. Functional amyloids of bacteria are promising biomaterial systems for functional applications ranging from bioelectronics to tissue engineering [3].

Acknowledgements: This study was supported by TUBITAK through project 114M163

References

1. Chapman, M. R.; Robinson, L. S.; Pinkner, J. S.; Roth, R.; Heuser, J.; Hammar, M.; Normark, S.; Hultgren, S. J. Role of Escherichia coli curli operons in directing amyloid fiber formation. Science 2002, 295 (5556), 851-855.
2. Seker, Chen, Citorick and Lu, Synthetic Biogenesis of Bacterial Amyloid Nanomaterials with Tunable Inorganic–Organic Interfaces and Electrical Conductivity ACS Synthetic Biology, 2016, DOI: 10.1021/acssynbio.6b00166
3. Abdelwahab, M. T., Kalyoncu, E., Onur, T., Baykara, M. Z., & Seker, U. O. S. (2017). Genetically-Tunable Mechanical Properties of Bacterial Functional Amyloid Nanofibers. Langmuir.

Intrinsically disordered proteins protein-protein interaction prediction method based on various sequence representations

[a] VLADIMIR PEROVIC, [a] NEVEN SUMONJA, [b] SANDRO RADOVANOVIC,
[a] BRANISLAVA GEMOVIC, [b] MILAN VUKICEVIC, [a] NEVENA VELJKOVIC

[a] *Centre for multidisciplinary research, Vinca Institute of Nuclear Sciences, University of Belgrade*
[b] *Centre for business decision making, Faculty of organizational Sciences, University of Belgrade*

Intrinsically disordered proteins (IDPs) are involved in the regulation of numerous crucial biological processes and in the pathogenesis of various human diseases via binding to multiple protein partners. Due to their principal biological significance mapping of protein-protein interactions (PPIs) of IDPs will have far-reaching effects on our understanding of cellular functions and on discovery of modulators the IDPs functionalities. Sequence based protocols for prediction of PPIs that utilize machine learning are powerful approach to speed up discovery of new interactions in the most economical way. We developed a sequence based method utilizing physicochemical features of the interacting and non-interacting protein pairs that are pulled out by means of spectral and pseudo amino acid composition sequence representations. The method performs binary classification and feature selection based on machine learning algorithms and allows for prediction interactions between human IDPs and proteins with disorder regions and any other member of the human proteome. Our goal was to define a model which relies on unique intrinsic features of IDPs complexes which will be capable to predict interactions with partners that were unseen to model during the training process. Human IDP sequences were retrieved from the DisProt 7 database and their interactions from the HIPPIE v2.0. In order to identify the best predictive model, we made a procedure for automatic model building and evaluation of machine learning approaches as well as feature selection and parameter optimization. Procedure included evaluation of ensemble based algorithms such as Random Forests and Gradient Boosted Trees, cutting edge single models such as SVMs and traditional models such as Logistic Regression. In order to perform training and testing on strictly non-similar sequences we designed interaction datasets to have less than 40% of homology. Furthermore, given that IDPs act as hubs and have multiple binding partners or many IDPs are prone to interaction with one same partner we had to keep the equal imbalance ratio per protein in training and test sets. Our proposed prediction method outperforms similar methods on this task and achieves an average 74% accuracy and 79% AUC when evaluated by ten-fold cross-validation. We thus demonstrate that our method efficiently generalize intrinsic features of IDPs interactions and can be considered as powerful approach to foster discoveries of insufficiently understood IDPs ensembles.

Acknowledgement: This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (no. 173001) and Swiss National Science Foundation SNSF (IZ73Z0_152415).

Generation of a framework for community-sourced curation of experimentally validated functional modules in intrinsically disordered regions

NORMAN E. DAVEY

Conway Institute of Biomolecular & Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland.

UCD School of Medicine & Medical Science, University College Dublin, Belfield, Dublin 4, Ireland.

The intrinsically disordered regions of proteins contain a multitude of functional modules known as short, linear motifs (SLiMs) and intrinsically disordered domains (IDDs). These sites mediate a diverse set of functions such as directing ligand binding, providing both peripheral docking sites and specificity for modifying enzymes, controlling protein stability and acting as signals to target proteins to specific subcellular locations. However, to date, the vast majority of these modules that have been experimentally validated remain to be curated. We recently estimated that 2/3 of the articles describing experimental validation of SLiMs are unannotated. The current resource collecting these instances, the ELM database, is limited in its scope ignoring high-throughput data such as phage display or peptide arrays. However, it still cannot keep pace with the growth of the available literature due to the complicated curation process. Strikingly, intrinsically disordered domains are not currently catalogued in any resource. A comprehensive resource of experimentally validated functional modules in intrinsically disordered regions will be an unmatched educational resource for biologists and a gold standard source of data to train future discovery tools to analyse the function of intrinsically disordered regions. We present a prototype resource for such data. The resource, DIDI (Database of Intrinsically Disordered Interfaces), will be a repository for community-curated and data-mined information from low-throughput experiments, peptide array data, classical/proteomic phage display and structural studies. The resource can be considered the functionally-oriented sister resource of the structural information-centric DisProt database. The success of the Non-Globular Proteins COST community-driven annotation of DisProt will hopefully be replicated for the functional aspects of the field.

Armadillo Repeat Proteins: From their evolution to their engineering

ANDREAS PLÜCKTHUN

University of Zurich, Dept. of Biochemistry, Switzerland

Most binding proteins used in research are monoclonal antibodies, made by the >40-year-old hybridoma technology, some with questionable performance [1, 2]. More recently, recombinant antibodies and non-antibody scaffolds, selected from synthetic libraries, have started to provide access to molecularly defined molecules [3, 4]. Nonetheless, all of these approaches require one to treat every target as a completely new project. This is currently unavoidable for folded proteins. We hypothesized that for unfolded proteins or unfolded stretches (western blots, tags, posttranslationally modified tails), termed "peptides" for simplicity, the regularity of the peptide main chain can be exploited, and a modular detection system can be devised, which would ultimately allow to generate a sequence-specific binding protein without experimentation.

The basis of our approach are Armadillo Repeat Proteins [5-15], which bind peptides in a completely extended way, providing a pocket for each side chain, and thus access to a modular approach. Combining evolutionary engineering, NMR, X-ray crystallography and structure-based computation, we have now achieved well crystallizing ArmRPs with bound peptides, picomolar affinities, and a well functioning selection and evolution technology, as well as a portfolio of biochemical and biophysical analysis technologies for the engineered ArmRPs. Progress in the various aspects will be summarized.

References

1. Bradbury, A., and Plückthun, A. (2015). *Standardize antibodies used in research.* Nature 518, 27-29.
2. Bradbury, A. R., and Plückthun, A. (2015). *Getting to reproducible antibodies: the rationale for sequenced recombinant characterized reagents.* Protein Eng. Des. Sel. 28, 303-305.
3. Binz, H. K.et al. (2005). *Engineering novel binding proteins from nonimmunoglobulin domains.* Nat. Biotechnol. 23, 1257-1268.
4. Plückthun, A. (2015). *Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy.* Annu. Rev. Pharmacol. Toxicol. 55, 489-511.
5. Reichen, C. et al. (2016). *Computationally designed armadillo repeat proteins for modular peptide recognition.* J. Mol. Biol. 428, 4467-4489
6. Hansen, S. et al. (2016). *Structure and energetic contributions of a designed modular peptide-binding protein with picomolar affinity.* J. Am. Chem. Soc. 138, 3526-3532.
7. Reichen, C. et al. (2016). *Structures of designed armadillo-repeat proteins show propagation of inter-repeat interface effects.* Acta Crystallogr. D72, 168-175.
8. Ewald, C. et al. (2015). *A combined NMR and computational approach to investigate peptide binding to a designed armadillo repeat protein.* J. Mol. Biol. 427, 1919-1933.

9. Reichen, C. et al. (2014). *Crystal structures of designed armadillo repeat proteins: Implications of construct design and crystallization conditions on overall structure.* Protein Sci. 23, 1572-1583.
10. Watson, R. P. et al. (2014). *Spontaneous self-assembly of engineered armadillo repeat protein fragments into a folded structure.* Structure 22, 985-995.
11. Reichen, C. et al. (2014). *Modular peptide binding: From a comparison of natural binders to designed armadillo repeat proteins.* J. Struct. Biol. 185, 147-162.
12. Varadamsetty, G. et al. (2012). *Designed armadillo repeat proteins: library generation, characterization and selection of peptide binders with high specificity.* J. Mol. Biol. 424, 68 - 87.
13. Alfarano, P. et al. (2012). *Optimization of designed armadillo repeat proteins by molecular dynamics simulations and NMR spectroscopy.* Protein Science 21, 1298-1314.
14. Madhurantakam, C. et al. (2012). *Structure-based optimization of designed Armadillo-repeat proteins.* Protein Science 21, 1015-1028.
15. Parmeggiani, F. et al. (2008). *Designed armadillo repeat proteins as general peptide-binding scaffolds: consensus design and computational optimization of the hydrophobic core.* J Mol Biol 376, 1282-1304.

Evolution of repeat proteins by exon duplication

[a] LISANNA PALADIN, [a] LAYLA HIRSH, [a] DAMIANO PIOVESAN, [b] PABLO MIER,
[b] MIGUEL A. ANDRADE-NAVARRO, [a,c] SILVIO C.E. TOSATTO

[a] Dept. Of Biomedical Sciences, University of Padova, viale G.Colombo 3, Padova, Italy
[b] Institute of Molecular Biology, Faculty of Biology, Johannes Gutember University of Mainz
[c] CNR Institute of Neuroscience, G. Colombo 3, Padova, Italy

Tandem repeat proteins (TRPs) are a puzzling class of proteins whose 3D architecture consists in the repetition of a simple structural module [1]. Being these structural units stabilized from the inter-unit interactions, they confer to repeat proteins a unique folding pathway and properties. TRPs are central in cell signaling and regulation, in fact they are usually characterized by a large surface free to specialize for a variety of functions, including protein-protein interaction. TRP folds emerged several times across different lineages and are widely distributed across functional pathways. Repeat regions arise from the multiple duplication of a segment in a coding sequence. For Eukaryotes, it has been suggested that repeated segments could correspond to exons, thus being easily recombined and duplicated thanks to the modular intron/exon structure [2]. To test this hypothesis, we extended the evolution analysis of TPRs to all known repeated structures by exploiting RepeatsDB data [3]. We analyzed the exon structure of the repeated regions in order to measure the correspondence with the structural modularity of the repeats. The analysis shows that the exon/unit patterns are family-specific, i.e. conserved among Pfam domains and linked to the evolutionary history of the repeat domain [4]. Ancient repeat domains usually do not show a well defined exon/unit matching on the contrary, TRPs mainly observed in Eukaryotes do. A possible explanation is that repeat domains originated in Eukaryotes took advantage of exon/intron mechanisms to facilitate unit expansion. This scenario is observed in all repeat classes, from elongated to closed to beads-on-a-string repeats, indicating that this tendency is intrinsic in the modularity of these genes and may have contributed to the greater number of TRPs in Eukaryotes than Prokaryotes. Finally, thanks to this detailed analysis it was possible to establish the correct evolutive phase of the repeat families, contributing to the curation of RepeatsDB structural annotation. The analysis focuses on these families by providing and discussing representative examples. By taking repeat proteins as manageable study case, this work sheds light on the mechanisms of protein evolution by duplication of genetic material.

References

1. V. Kajava, «Tandem repeats in proteins: from sequence to structure», J Struct Biol, vol. 179, pagg. 279 – 288, 2012.
2. T. O. Street, G. D. Rose, e D. Barrick, «The Role of Introns in Repeat Protein Gene Formation», J. Mol. Biol., vol. 360, n. 2, pagg. 258–266, lug. 2006.
3. L. Paladin, L. Hirsh, D. Piovesan, M. A. Andrade-Navarro, A. V. Kajava, e S. C. E. Tosatto, «RepeatsDB 2.0: improved annotation, classification, search and visualization of repeat protein structures», Nucleic Acids Res., pag. gkw1136, nov. 2016.
4. L. Paladin e S. C. E. Tosatto, «Comparison of protein repeat classifications based on structure and sequence families», Biochem. Soc. Trans., vol. 43, n. 5, pagg. 832–837, ott. 2015.

Study of amino acid biased regions in proteins

[a,b] MIGUEL A. ANDRADE-NAVARRO, [a,b] PABLO MIER, [b] DANIEL BRÜNE, [c] NORMAN DAVEY

[a] *Faculty of Biology, Johannes-Gutenberg University of Mainz, Gresemundweg 2, 55128 Mainz, Germany.*

[b] *Institute for Molecular Biology gGmbH, Ackermannweg 4, 55128 Mainz, Germany.*

[c] *UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Dublin 4, Ireland.*

Biased regions in proteins, regions of unusual amino acid composition, appear in many sequences. Different to regions that fold into globular domains, these regions have less understood functional and structural properties. Previously these regions were considered mere linkers between domains, but it is becoming more obvious that they have functions, for example holding sites for post-translational modifications, and motifs for the interaction with proteins, RNA and DNA [1]. Biased regions have fast rates of evolution and therefore low conservation in homologous sequences, which difficults their study by sequence comparison.

Here, we present our approaches for the study of biased regions. Since the sequences between globular domains tend to be compositionally biased, we study these connectors, characterizing their amino acid biases, and how those associate with certain functions and locations in respect to the particular types of protein domains they connect to. In addition, we are studying compositionally biased regions composed of very short repeats, since these are likely to have different structural properties (and likely different functions) than biased regions with amino acids in random order.

These analyses follow our previous work using motif discovery [2], and our use of context to characterize homorepeats [3], and take further steps towards the complete evolutionary and functional analysis of large regions of proteins yet unexplored.

Acknowledgement: The authors thank all members of the CBDM group for helpful and constructive discussions. We gratefully acknowledge the support of the COST Association (Cost Action BM1405).

References

1. Mier P, Andrade-Navarro MA. dAPE: a web server to detect homorepeats and follow their evolution. Bioinformatics. 2017 Apr 15;33(8):1221-1223. doi: 10.1093/bioinformatics/btw790.
2. Krystkowiak I, Davey NE. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. Nucleic Acids Res. 2017 Apr 6. doi: 10.1093/nar/gkx238
3. Mier P, Alanis-Lobato G, Andrade-Navarro MA. Context characterization of amino acid homorepeats using evolution, position, and order. Proteins. 2017 Apr;85(4):709-719. doi: 10.1002/prot.25250. Epub 2017 Feb 6.

Repeat proteins: scaffolds for artificial bioelectronics materials

[a] SARA H. MEJÍAS, [b] JAVIER LÓPEZ-ANDARIAS, [c] MANTAS LIUTKUS, [b] CARMEN ATIENZA, [a,b] NAZARIO MARTÍN, [a,c,d] AITZIBER L. CORTAJARENA

[a] IMDEA-Nanociencia, C/ Faraday, 9, Campus de Cantoblanco, E-28049 Madrid, Spain

[b] Departamento de Química Orgánica I, Facultad de Química, Universidad Complutense, E-28040 Madrid, Spain

[c] CIC biomaGUNE, Paseo de Miramón 182, E-20014 Donostia-San Sebastian, Spain

[d] Ikerbasque, Basque Foundation for Science, Mª Díaz de Haro 3, E-48013 Bilbao, Spain

The development of sophisticated ordered functional materials is one of the important challenges in current science. One of the keys to enhance the properties of these materials is the control of the organization and morphology at different scales. In this work, we present a bioinspired methodology to achieve highly ordered donor/acceptor bio-nanohybrids using designed repeat proteins as scaffold, in particular, tetratricopeptide repeats (TPRs) [1,2]. The repetitive nature of the TPR proteins has been shown to provide a regular ordered surface suitable for the display of porphyrin chromophores with desired spacing and density with atomic-resolution precision, forming a highly regular conjugated array of porphyrins for optimal energy capture [3]. The protein scaffold is endowed with photoactive and electron donating porphyrin (P) units, to efficiently wrap around electron accepting Single wall carbon nanotubes (SWCNT). A systematic experimental study reveals that designed proteins wrap around the SWCNT in an efficient manner due to the strong supramolecular interaction existing between the inner concave surface of the protein and the convex surface of the SWCNT. Interestingly, spectroscopy and X-ray diffraction data further confirm that the so-called protein-P-SWCNT donor-acceptor bio-nanohybrids retain the original protein structure. Finally, the new bio-nanohybrid show a remarkable enhancement on the photoconductivity values measured by FP-TRMC technique.

References

1. S. H. Mejias, A. Aires, P. Couleaud, A. L. Cortajarena, *Adv. Exp. Med. Biol.* 2016, *940*, 61-81
2. D. Romera, P. Couleaud, S. H. Mejias, A. Aires, A. L. Cortajarena, *Biochem. Soc. Trans.* 2015, *34*, 825-831
3. S. H. Mejías, J. López-Andarias, T. Sakurai, S. Yoneda, K. P. Erazo, S. Seki, C. Atienza, N. Martín, A. L. Cortajarena, *Chem. Sci.* 2016, *7*, 4842-4847

A heuristic implementation of the CAST algorithm for detecting compositional bias in protein sequences

[a,b] KIRMITZOGLOU I, [a] IOANNIDES A-N, [a] PROMPONAS VJ

[a] *Bioinformatics Research Laboratory, Department of Biological Sciences, University of Cyprus, PO Box 20537, CY 1678, Nicosia, Cyprus*

[b] *Present address: Department of Life Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.*

CAST [1] is an algorithm that detects compositionally biased regions in protein sequences. It is originally implemented as an iterative variant of the Smith-Waterman dynamic programming pairwise sequence alignment algorithm [2], comparing a query sequence to degenerate homopolymers corresponding to the 20 different commonly occurring proteinogenic amino acids. CAST was shown to be a superior choice for masking low complexity regions prior to database searches [1, 3, 4], however it is much slower than competing tools (e.g. the widely used SEG [5]).

In this work, we report a new implementation of the CAST algorithm which (in addition to bug fixes and new output format options) introduces novel options for heuristic detection of compositionally biased regions with a significant speed-up, especially in cases of heavily compositionally biased datasets. We demonstrate, that with careful selection of the novel heuristic parameters, users can maintain the speed-up with only marginal loss of specificity in the detection of compositionally biased regions. We will also describe novel web accessible tools related to the detection of compositionally biased regions in protein sequences.

Acknowledgement: We wish to thank Prof. Christos Ouzounis (CERTH, Greece), Agathoklis Papadopoulos and Theocharis Theocharides (KIOS Center/University of Cyprus, Cyprus) for helpful discussions on this work.

References

1. Promponas VJ, Enright AJ, Tsoka S, Kreil DP, Leroy C, Hamodrakas S, Sander C, Ouzounis CA. CAST: an iterative algorithm for the complexity analysis of sequence tracts. Bioinformatics. 2000 Oct;16(10):915-22.
2. Smith TF, Waterman MS. Identification of common molecular subsequences. J Mol Biol. 1981 Mar 25;147(1):195-7.
3. Kreil DP, Ouzounis CA. Comparison of sequence masking algorithms and the detection of biased protein sequence regions. Bioinformatics. 2003 Sep 1;19(13):1672-81.
4. Kirmitzoglou I, Development of algorithms and software for unravelling the biological role of low complexity regions in protein sequences. PhD thesis. 2014; (available online at http://lekythos.library.ucy.ac.cy/handle/10797/14177).
5. Wootton JC, Federhen S. Analysis of compositionally biased regions in sequence databases. Methods Enzymol. 1996;266:554-71.

Seeding barrier between hamster and human prion protein amyloid-like fibrils

[a] DARIUS ŠULSKIS, [a] GRETA MUSTEIKYTĖ, [a] VYTAUTAS SMIRNOVAS

[a] *Vilnius University, Institute of Biotechnology, Vilnius, Lithuania*

Prions are infectious particles which play the main role in a group of fatal neurodegenerative disorders, also known as the transmissible spongiform encephalopathies. Prion diseases propagate by self-replication of a pathogenic prion isoform using cellular prion protein (PrP) as a substrate. Transmission of prion diseases between different species is possible, but in some cases species barrier can be observed.

Here we studied cross-seeding between human and Syrian hamster PrP amyloid-like fibrils. Our kinetic data shows the barrier between these species to be one-sided – self-replication of hamster PrP fibrils using human PrP as a substrate is much slower than using hamster PrP, while human PrP can self-replicate equally fast using either human or hamster PrP as a substrate. Structural studies using FTIR spectroscopy revealed differences in structures of human and hamster PrP fibrils and inability of human PrP to adopt structural features of hamster PrP fibrils. Comparison of PrP mutants revealed that differences in amino acid sequence at positions 166 and 168 are the main reason for this species barrier.

Molecular tools to tackle protein aggregation in Machado-Joseph disease

[a,b,*] ANA ALMEIDA, [a,b,*] ALEXANDRA SILVA, [a,b] ZSUZSA SÁRKÁNY, [a,b,#] BRUNO ALMEIDA, [a,b] SANDRA MACEDO-RIBEIRO

[a] *IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, 4200- 135 Porto, Portugal*

[b] *i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal*

**Shared first authorship*

[#]*Present address - Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, 4710-057, Portugal.*

Medicine, University of Minho, Braga, 4710-057, Portugal. Machado-Joseph Disease (MJD) is a neurodegenerative disorder, included in the group of polyglutamine (polyQ) diseases, caused by a mutation resulting in the expansion of the polyglutamine tract in the protein ataxin-3 (Atx3). This protein functions as a deubiquitinase and turns pathogenic whenever its polyQ tract exceeds a threshold of 55 glutamines. Enzymatic activity is ensured by the globular Josephin domain (JD), which is followed by a flexible C-terminus tail containing two or three ubiquitin-interacting motifs (UIMs) and the polyQ tract. The JD carries aggregationprone regions required for the first stage of aggregation, independent of polyQ tract. Macromolecular interaction of Atx3 partners within this region proved to modulate Atx3 aggregation rates. Nanobodies, derived from camelid heavy-chain antibodies, are emerging as promising tools in biotherapeutics and as probes for protein aggregation due to their high affinity, specificity and stability. Therefore, nanobodies targeting JD have been produced and tested as tools to interfere with Atx3 self-assembly. One of these nanobodies, NB01, proved to interact both with non-expanded (13Q) and expanded (77Q) Atx3 in thenanomolar range, although with higher affinity towards Atx3 77Q. Furthermore, the molecular interaction between NB01 and Atx3 has an impact on Atx3 self-assembly rates and impairs maturation of pathogenic Atx3 fibrils. Therefore, this molecule exhibits a great potential for further studies and development of future therapies tackling MJD pathologies.

Acknowledgement: This work was financed by FEDER - Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 - Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT – Fundação para a Ciência e a Tecnologia/ Ministério da Ciência, Tecnologia e Ensino Superior in the framework of the project "Institute for Research and Innovation in Health Sciences" (POCI-01- 0145-FEDER-007274). The authors acknowledge the support and the use of resources of Instruct, a Landmark ESFRI project. J. Steyaert and E. Pardon (Nanobodies4Instruct, Vrije Universiteit Brussel) are gratefully acknowledged for support and collaboration in the nanobody production.

Interference of small molecules with amyloid aggregation of Abeta peptide: Structure – activity relationship

[a] ZUZANA GAZOVA, [a] ZUZANA BEDNARIKOVA, [a,b] KATARINA ULICNA, [a] JANA KUBACKOVA, [a] MIROSLAV GANCAR, [a] DIANA FEDUNOVA, [a] JOZEF MAREK, [a] ANDREA ANTOSOVA

[a] *Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, Kosice, Slovakia*

[b] *Institute of Biology and Ecology, Faculty of Science, Pavol Jozef Safarik University, Kosice, Slovakia*

General consensus supports findings that accumulation of Abeta peptide (Aβ) aggregates as soluble oligomers and senile plaques in specific areas of human brain is a key landmark of Alzheimer's disease (AD) [1]. The precise mechanism of toxicity of amyloid aggregates is not fully elucidated; however, there are evidences that prevention or reversion of the amyloid aggregation is beneficial. Therefore, the increasing attention is focused on searching for agents able to interfere with Aβ amyloid aggregation leading to inhibition and/or reduction of amyloid aggregates [2].

For that purpose our aim is to search for inhibitors of amyloid aggregation among small molecules. We have found several derivatives of acridine, curcumin and tacrine-coumarin hybrid molecules, compounds extracted from traditional Chinese herbs as well as short peptides with significant inhibiting abilities [3, 4]. Our results indicate important relationship between compound structure and anti-amyloid activity. The special attention was devoted to fullerenol as novel agents able to affect amyloid aggregation of poly/peptides. The obtained findings favor the application of the active small molecules as therapeutic agents targeting Alzheimer's diseases.

Acknowledgement: Supported by grants VEGA 2/0145/17, 2/0176/14, MVTS COST 083/14 action BM1405 and SAS-MOST JRP 2015/5.

References

1. J. Hardy, The amyloid hypothesis for Alzheimer's disease: a critical reappraisal, J. Neurochem. 2009, 110 (4) 1129–1134
2. K. A. DaSilva, et al., Amyloid-β fibrillogenesis: structural insight and therapeutic intervention, Exp. Neurol. 223 (2010) 311–321
3. S. Hamulakova et al., Targeting copper(II)-induced oxidative stress and the acetylcholinesterase systemin Alzheimer's disease usingmultifunctional tacrine-coumarin hybrid molecules, Journal of Inorganic Biochemistry 161 (2016) 52–62
4. Z. Bednarikova et al. Fullerenol C60(OH)16 prevents amyloid fibrillization of Ab40 – in vitro and in silico approach, Phys. Chem. Chem. Phys., 2016, 18, 18855-18867

DC8E8 inhibits disordered tau protein oligomerization by a combination of pi-electron interaction, flexible recognition and steric hindrance

[a,b] ROSTISLAV SKRABANA, [a,b] PETER FILIPCIK, [a,b] BRANISLAV KOVACECH, [a,b] NORBERT ZILKA, [a,b] EVA KONTSEKOVA, [a,b] ONDREJ CEHLAR, [a,c] MICHAL NOVAK

[a] *Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia*
[b] *AXON Neuroscience R&D Services SE, Bratislava, Slovakia*
[c] *AXON Neuroscience SE, Larnaca, Cyprus*

Tau protein oligomers are likely the toxic agents underlying the neurodegeneration in Alzheimer's disease, therefore, an immunotherapy targeting pathogenic tau epitopes holds promise for a disease modifying treatment. Strong evidences exist that tau interaction is nucleated via hexapeptides VQIINK and VQIVYK in the microtubule-binding tau region. DC8E8 is a mouse anti-tau monoclonal antibody recognizing four homologous epitopes HxPGGG, two of them in the immediate vicinity of the aggregation-promoting hexapeptides. Strikingly, DC8E8 is able to effectively inhibit tau-tau oligomerization in vitro and represents therefore a promising candidate for tau immunotherapy. DC8E8 tau active vaccine AADvac1 derived from an epitope component of DC8E8 is under clinical development. We have determined several X-ray structures of DC8E8 Fab and its complexes with tau peptides and deduced steps of complex formation. Partially flexible antibody loop CDRH3 participates on specific tau recognition. Tau protein chain may adopt the type I beta-turn by residues GGGS and strongly interacts with the antibody aromatic rings. The bulky CDRL1 loop of the antibody remains conformationally unstable in the complex and forms an entropic screen of the beta-structure forming hexapeptides adjoining the epitope. The results allowed a mechanistic insight into DC8E8 anti-aggregation activity and showed a strategy used by immune system to recognize an intrinsically flexible target.

Poster Presentations

Partitioning biological networks in *k*-plex subnetworks with maximum edge weights

[a] MILANA GRBIĆ, [b] ALEKSANDAR KARTELJ, [a] DRAGAN MATIĆ, [b] VLADIMIR FILIPOVIĆ

[a] *Faculty of Natural Science and Mathematics, University of Banja Luka*

[b] *Faculty of Mathematics, University of Belgrade*

Partitioning large biological networks into smaller clusters can help in discovering new properties and functionalities of a particular structure. In this work we deal with the partitioning of the edge-weighted networks into *k*-plex components, where a set of some *n* vertices in a network is a *k*-plex if the degree of each vertex in the induced subnetwork is at least *n*-*k*. The aim of the maximum edge-weight *k*-plex partitioning problem (Max-E*k*PP) is to find the *k*-plex partition with the maximum total weight. The problem arises from analyzing the sparse metabolic networks built on a principle that two metabolites are adjacent if they figure in at least one common reaction. Another application of this problem belongs to the area of the protein structure prediction. More concrete, protein threading problem can be reduced to the maximum edge-weight clique problem, which is a special case of the Max-E*k*PP for *k*=1. We present a heuristic approach for solving this NP hard problem and graphically represent the solutions obtained on some large metabolic networks.

References

1. Martins, Pedro. "Modeling the maximum edge-weight k-plex partitioning problem." arXiv preprint arXiv:1612.06243 (2016).
2. Förster, Jochen, et al. "Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network." Genome research 13.2 (2003): 244-253.
3. Akutsu, Tatsuya, et al. "Dynamic programming and clique based approaches for protein threading with profiles and constraints." IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences 89.5 (2006): 1215-1222.
4. Pattillo, Jeffrey, Nataly Youssef, and Sergiy Butenko. "On clique relaxation models in network analysis." European Journal of Operational Research 226.1 (2013): 9-18.
5. Gouveia, Luis, and Pedro Martins. "Solving the maximum edge-weight clique problem in sparse graphs with compact formulations." EURO Journal on Computational Optimization 3.1 (2015): 1-30.
6. Pullan, Wayne. "Approximating the maximum vertex/edge weighted clique using local search." Journal of Heuristics 14.2 (2008): 117-134.

Conformational ensembles inside alpha-solenoids proteins

[a] HIRSH L., [b] PIOVESAN D., [b] PALADIN L., [b,c] TOSATTO S.E.

[a] *Department of Engineering Pontificia Universidad Católica del Perú, Lima, Perú.*
[b] *Department of Biomedical Sciences University of Padua, Padua, Italy.*
[c] *CNR Institute of Neuroscience, Padua, Italy.*

Tandem repeat proteins (TRPs) are modular structures that defy the classical definition of globular proteins and domains, being ensembles of interdependent structural units rather than assemblies of structurally autonomous part. Their modularity originated from the tandem duplication of a sequence motif, and implies a continuum of structural arrangement from unit to unit, as well as a regularity in the contacts pattern that stabilizes each unit via the interaction with the flanking ones [1]. The structural and functional properties of these modular proteins are often preserved despite the high tendency to divergence of sequence repeats. This makes the identification of their sequence periodicity an extremely hard task. To overcome this issue, a number of structure-based methods have been developed. ReUPred [2], a predictor for automatic detection of tandem repeat units in protein structures, was used to populate RepeatsDB 2.0 [3], a resource for high-quality repeat structure annotation. The database includes entries annotated at unit level and classified according to a structural classification based on structural features.

Here we present a manual refinement of solenoidal unit position (repeat phase), together with a new strategy for a finer classification, beyond subclass level. We used a network analysis based on the structural similarity of the repeat units to identify clusters inside subclasses. The structural ensembles (clusters) are characterized by different sequence and structural conformations corresponding to different protein families. We refined the unit definition (phase) inside each cluster iteratively by evaluating the multiple structural alignments. Finally, to prove the efficiency of the new unit definition, we derived a Hidden Markov Model (HMM) for each ensemble and performed a HMMER [4] search against SwissProt. The results were compared to Pfam annotation. A comparable number of sequences are found even if our models are built on much smaller seeds compared to Pfam.

By applying this approach to all repeat subclasses, we will be able to characterize all repeat families in RepeatsDB, standardize the unit phase and derive the corresponding HMM models for large-scale annotation.

References

1. Parmeggiani, Fabio, and Po-Ssu Huang. 2017. "Designing Repeat Proteins: A Modular Approach to Protein Design." *Current Opinion in Structural Biology* 45 (August): 116–23. doi:10.1016/j.sbi.2017.02.001.
2. Hirsh, Layla, Damiano Piovesan, Lisanna Paladin, and Silvio C. E. Tosatto. 2016. "Identification of Repetitive Units in Protein Structures with ReUPred." *Amino Acids* 48 (6): 1391–1400. doi:10.1007/s00726-016-2187-2.
3. Paladin, Lisanna, Layla Hirsh, Damiano Piovesan, Miguel A. Andrade-Navarro, Andrey V. Kajava, and Silvio C.E. Tosatto. 2016. "RepeatsDB 2.0: Improved Annotation, Classification, Search and Visualization of Repeat Protein Structures." *Nucleic Acids Research*, November. doi:10.1093/nar/gkw1136.
4. Finn, R. D., J. Clements, and S. R. Eddy. 2011. "HMMER Web Server: Interactive Sequence Similarity Searching." *Nucleic Acids Research* 39 (suppl): W29–37. doi:10.1093/nar/gkr367.

Novel examples of long conserved intrinsically disordered regions

[a,b,c] TOMAS DI DOMENICO, [d,e] SILVIO TOSATTO, [f] MARCO PUNTA

[a] *Gurdon Institute, University of Cambridge, Cambridge, UK*

[b] *Department of Genetics, University of Cambridge, Cambridge, UK*

[c] *Wellcome Trust Sanger Institute, Cambridge, UK*

[d] *Department of Biomedical Sciences, University of Padua, Italy*

[e] *CNR Institute of Neuroscience, Padova, Italy*

[f] *Center for Evolution and Cancer, The Institute of Cancer Research, UK*

Functional elements of intrinsically disordered regions in proteins are often encoded by short (typically < 10 aa-long) linear motifs [1]. Longer evolutionary divergent conserved regions, however, are also known (sometimes called 'intrinsically disordered domains" or IDDs) [2]. Compared to short linear motifs, IDDs have been shown to span longer interaction interfaces, regions of tightly clustered (sometimes overlapping) linear motifs and regions that fold upon post-translational modification (e.g. phosphorylation [3]). Here, we performed an explorative study in search of novel IDDs in the UniProt database [4], outside of previously annotated Pfam [5] regions. Among the novel long conserved disordered regions that we generated, we describe those for which we could find at least some functional annotation in the literature. These include *Rabenosyn-5 - EH-domain*, Caskin1-*CASK*, *HIC1-CtBP, TNRC6C-PABC, Ataxin-3-VCP, Intersectin-Clathrin adaptor complex AP2 and USP19-SIAH1* interaction regions along with *P47-phox auto-inhibition and CYDL serine-rich* regions. As a result of this study, some of these regions have already been included into the Pfam database. Most (but not all) are restricted to the phylum *Chordata*. In some cases (e.g. the *Rabenosyn 5 - EH-domain interaction region and* the *Intersectin-Clathrin adaptor complex AP2 interaction region*), a link between the existence of an extended region of amino acid conservation and function could be established with some degree of confidence (e.g. presence of multiple cooperating or competing linear motifs). For most of these regions, however, we could only identify overlap to a single known short linear motif. Additional experimental evidence will hence be needed in order to ascertain the functional relevance of the regions of amino acid conservation that extend beyond such isolated linear motifs.

References

1. Davey et al. Mol Biosyst. (2012)
2. Tompa et al. Bioessays. (2009)
3. Bah et al. Nature (2015)
4. The Uniprot Consortium NAR (2017)
5. Finn et al. NAR (2016)

Redox-regulation and chaperone function of the Hsp33 homologue in T. Brucei

[a] SAMAR ARAMIN, [a] ROSI FASSLER, [a] MOR GOLDENBERG, [b] VAIBHAV CHIKNE, [b] SHULAMIT MICHAELI, [a] DANA REICHMANN

[a] *Department of Biological Chemistry, The Alexander Silberman Institute of Life Science, Hebrew University of Jerusalem, Jerusalem, Israel*

[b] *The Institute of Nanotechnology and Advanced Materials, Bar Ilan University, Tel-aviv, Israel*

Hsp33 is an ATP-independent chaperone with a zinc-binding domain that operates as a redox-switch of the Hsp33 chaperone function. Hsp33 binds misfolded client proteins upon oxidative-stress, and transfers them to the ATP-depending chaperone machinery, DnaK/J, for further refolding upon stress relief. Hsp33 was originally characterized in bacteria (E.coli and B.sabtilis). Here, we describe a novel eukaryote homologue of HSP33 in Trypanosoma brucei parasite causing “sleeping sickness” disease in humans, responsible for the death of thousands in Africa every year.

Bioinformatic analysis of the Hsp33 family suggested that eukaryotic parasites, such as Leishmania, Trypansoma and others had presevered Hsp33 homologues in their genomes. Intriguingly, mammal lacks the Hsp33 homologue making pathogenic Hsp33 homologues be a potential drug-target with small chance for toxicity in the host. Here we focused on a functional characterization of the Hsp33 homologue, named TrypOx in Trypanosoma brucei. During pathogenic life cycle Trypanosoma parasites need to cope with elevated temperatures and oxidative stress. Therefore, we tested role of TrypOx in survival of trypanosome during mild oxidative and heat shock stresses. RNAi silencing of TrypOx led to significant decrease in survival of T.brucei in mild stress conditions implying protective role of TrypOx during Trypanosoma life cycle.

To investigate redox activation of TrypOx, we replaced bacterial redox-sensitive domain that is responsible for the chaperone activation by its homologue from T.brucei. Analysis of chaperone activity, substrate binding and release showed functional conservation between the bacterial and T. brucei domains, suggesting that these homologues have similar activation mechanisms in bacteria and protozoa.

Moreover, we tested the role of oxidative stress and silencing of Trypox on the proteome of T.brucei. Proteomics-driven approach (Mass-spectrometry) were opted and a number of a new Heat-shock proteins and chaperons were identified.

High GC Content Causes Orphan Proteins to be Intrinsically Disordered

[a,b] WALTER BASILE, [a,b] OXANA SACHENKOVA, [a,b,c] SARA LIGHT, [a,b,d] ARNE ELOFSSON

[a] *Science for Life Laboratory, Stockholm University, Solna, Sweden*
[b] *Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden*
[c] *Bioinformatics Infrastructure for Life Sciences (BILS), Linköping University, Linköping, Sweden*
[d] *Swedish e-Science Research Center (SeRC), Kungliga Tekniska Högskolan, Stockholm, Sweden*

De novo creation of protein coding genes involves the formation of short ORFs from noncoding regions; some of these ORFs might then become fixed in the population [1]. These orphan proteins need to, at the bare minimum, not cause serious harm to the organism,
meaning that they should for instance not aggregate. Therefore, although the creation of short ORFs could be truly random, the fixation should be subjected to some selective
pressure. The selective forces acting on orphan proteins have been elusive, and contradictory results have been reported. In Drosophila young proteins are more disordered than ancient ones [2], while the opposite trend is present in yeast [3]. To the best of our knowledge no valid explanation for this difference has been proposed.

To solve this riddle we studied structural properties and age of proteins in 187 eukaryotic organisms. We find that, with the exception of length, there are only small differences in the properties between proteins of different ages. However, when we take the GC content into account we noted that it could explain the opposite trends observed for orphans in yeast (low GC) and Drosophila (high GC). GC content is correlated with codons coding for disorder promoting amino acids. This leads us to propose that intrinsic disorder is not a strong determining factor for fixation of orphan proteins. Instead these proteins largely resemble random proteins given a particular GC level. During evolution the properties of a protein change faster than the GC level causing the relationship between disorder and GC to gradually weaken.

Acknowledgements: This work was supported by grants from the Swedish Research Council (VR-NT 2012-5046, VR-M 2010-3555) and the Swedish E-science Research Center. Computational resources were provided by SNIC. SL was financed by Bioinformatics Infrastructure for Life Science (BILS).

References

1. Light S, Basile W, Elofsson A. Orphans and new gene origination, a structuraland evolutionary perspective. Curr Opin Struct Biol. 2014 Jun;26:73–83.
2. Palmieri N, Kosiol C, Schlotterer C. The life cycle of Drosophila orphan genes. Elife. 2014;3:e01311.
3. Carvunis AR, Rolland T, Wapinski I, Calderwood MA, Yildirim MA, Simonis N, et al. Proto-genes and de novo gene birth. Nature. 2012 Jul;487(7407):370–374.

The conformation of proline rich segment of neuronal protein tau studied by the X-ray crystallography, molecular dynamics simulations and biophysical methods

[a,b] ONDREJ CEHLÁR, [a,b] ROSTISLAV ŠKRABANA, [c] RADOVAN DVORSKÝ, [a,b] MICHAL NOVÁK

[a] *Institute of Neuroimmunology, Slovak Academy of Sciences, Dubravska cesta 9, 845 10 Bratislava, Slovakia*

[b] *Axon Neuroscience R&D Services SE, Dvorakovo Nabrezie 10, 811 02 Bratislava, Slovakia*

[c] *Institute of Biochemistry and Molecular Biology II, Heinrich-Heine University, Düsseldorf, Germany*

Introduction: Protein tau, which is implicated in Alzheimer's disease and other tauopathies, is an intrinsically disordered protein (IDP). In contrast to folded globular proteins, IDPs use for their interactions short segments called linear motifs, instead of tertiary structures. Despite their short length and lack of stable structure, linear motifs may have considerable structural propensities, which often resemble bound-state conformations [1].

Method: The monoclonal antibody Tau5 was crystallized with tau peptide from its proline rich region [2]. The dynamics of the X-ray observed conformation was probed by 1 μs unrestrained MD simulations of tau peptide ^{Ace}-^{218}PPTREPKKV^{226}-NH_2 and its T220A mutant. The simulation results were compared with the biophysical measurements.

Results: In the structure of Tau5 Fab fragment with tau peptide we were able to observe 16 amino acid long tau peptide. Its conformation is stabilized by an intrachain hydrogen bond that creates a ST-turn motif. The mutated peptide, which is not able to form this hydrogen bond, has shown decreased affinity to Tau5 in biophysical measurements (SPR, ELISA, ITC).

During the simulation time, the wild type tau peptide have occupied the bound-like conformation four times more than the T220A mutated peptide.

Conclusion: The intrachain hydrogen bond is stabilizing a shot structural motif in the proline rich region of protein tau and its loss leads to 2-fold affinity reduction in affinity to the Tau5 antibody. This supports a suggestion, that local and global conformation of IDP tau chain can be efficiently regulated by local structural propensities.

Acknowledgement: This work was supported by the Slovak Research and Development Agency under the contract No. LPP-0038-09 and by VEGA grant 2/0177/15. The simulations were performed thanks to the access to the Aurel supercomputer of the Slovak Infrastructure for High Performance Computing.

References:

1. Cino, E.A., W.Y. Choy, and M. Karttunen, *Characterization of the Free State Ensemble of the CoRNR Box Motif by Molecular Dynamics Simulations.* Journal of Physical Chemistry B, 2016. 120(6): p. 1060-1068.
2. Cehlar, O., et al., *Crystallization and preliminary X-ray diffraction analysis of tau protein microtubule-binding motifs in complex with Tau5 and DC25 antibody Fab fragments.* Acta Crystallographica Section F-Structural Biology and Crystallization Communications, 2012. 68: p. 1181-1185.

Unraveling TANC2-CDKL5-PP1 interactions: intrinsically disordered regions mediating novel pathways in neurodevelopment

[a,b] ALESSANDRA GASPARINI, [b,c,*] SILVIO C.E. TOSATTO, [a,*] EMANUELA LEONARDI

[a] *Department of Woman and Child's health, University of Padova, Italy;*
[b] *Department of Biomedical Sciences and CRIBI Biotechnology Center, University of Padova, Italy;*
[c] *CNR Institute of Neuroscience Padova, Italy.*
*silvio.tosatto@unipd.it, emanuela.leonardi@unipd.it

Neurodevelopmental disorders (NDDs) are a group of genetically heterogeneous conditions, which include Autism Spectrum Disorders (ASD) and Intellectual Disability (ID)[1]. These clinical manifestations are often caused by mutations in scaffold proteins of excitatory post synaptic densities, affecting neuronal plasticity and synaptic strength [1, 2]. In this study, we focused our attention on the TANC2 protein, a recently emerging candidate for NDDs [3]. TANC2 is a large scaffold protein expressed in the hippocampus within post-synaptic densities, and mutations are associated with several forms of NDDs, including ASD and ID [4-6]. Although it appears to play a critical role in glutamatergic neurotransmission, very little is known about TANC2 protein function in brain cells. From sequence analysis, TANC2 is predicted to contain several protein-protein interaction domains and long intrinsically disordered N- and C-terminal tails. The latter contain several conserved linear motifs for protein binding, suggesting a relevant role in integrating multiple incoming signals. Our work aims to validate experimentally the predicted protein interactions and their involvement in molecular pathways impaired in NDDs. Two highly conserved PP1 docking motifs in the N-terminus suggest that TANC2 could bind the phosphatase and direct its substrate specificity to CDKL5 [7]. Combinations of binary interaction between TANC2, PP1 and CDKL5 were validated by yeast two-hybrid assays and confirmed by co-localization in neuroblastoma cells and primary hippocampal neurons. TANC2 was found to directly bind PP1 through its N-terminus, which seems necessary for protein complex formation. Conversely, the CDKL5 C-terminus binds TANC2 at the C-terminal tail. Interestingly, the TANC2 disordered C-terminus displays a regularly spaced tyrosine pattern of unknown function, which may be involved in its interaction with partners (e.g. CDKL5). These findings suggest that TANC2 could function as a scaffold linking the phosphatase PP1 to its substrate CDKL5, allowing de-phosphorylation and subsequent degradation.

Acknowledgement: This work is supported by Italian Ministry of Health Young Researchers grant GR-2011-02347754.

References

1. Pinto, D. *et al.* Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* 94, 677–694 (2014).
2. Ting, J. T., Peça, J. & Feng, G. Functional Consequences of Mutations in Postsynaptic Scaffolding Proteins and Relevance to Psychiatric Disorders. *Annu. Rev. Neurosci.* 35, 49–71 (2012).
3. Han, S. *et al.* Regulation of Dendritic Spines, Spatial Memory, and Embryonic Development by the TANC Family of PSD-95-Interacting Proteins. *J. Neurosci.* 30, 15102–15112 (2010).
4. de Ligt, J. *et al.* Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N. Engl. J. Med.* 367, 1921–1929 (2012).
5. Fromer, M. *et al.* De novo mutations in schizophrenia implicate synaptic networks. *Nature* 506, 179–184 (2014).
6. Iossifov, I. *et al.* De Novo Gene Disruptions in Children on the Autistic Spectrum. *Neuron* 74, 285–299 (2012).
7. La Montanara, P. *et al.* Synaptic Synthesis, Dephosphorylation, and Degradation: a novel paradigm for an activity-dependent neuronal control of CDKL5. *J. Biol. Chem.* 290, 4512–4527 (2015).

Intrinsically disordered protein induces unusual b-strand dynamics in the ribosomal mRNAse

[a,b] SAN HADŽI, [a] JURIJ LAH, [b] REMY LORIS

[a] *Faculty of Chemistry and Chemical Technology, University of Ljubljana, Slovenia*
[b] *Department of Bioengineering Sciences, Vrije Universiteit Brussel and Structural Biology Research Center, Vlaams Instituut voor Biotechnologie, Brussels, Belgium*

Bacterial toxins from toxin-antitoxin modules are structurally diverse group of proteins. They inactivate important cellular targets (ribosome, gyrase, tRNA...) leading to an induction of dormant cellular state tolerant to antibiotic treatment [1]. One example is HigB2 mRNAse which degrades translating mRNA and thereby halts protein synthesis. Activity of toxin is regulated by formation of a tight complex with an intrinsically disordered antitoxin protein. New crystals structures point to a mechanism where binding of antitoxin induces a one-residue shift in the b-strand register [2]. This leads to an inversion of toxin's palindromic amino acid sequence: residues above the b-strand plan, which are part of the active site, are in now shifted below the plane rendering the toxin inactive. This highly unusually dynamics of the b-strand is induced by binding of an intrinsically disordered antitoxin by jet unknown mechanism. Observed b-strand sliding mechanism points to an important question: how dynamic are b-strands in proteins?

Acknowledgement: This work was supported by Slovenian Research Agency [Grant P1–0201], Fonds voor Wetenschappelijk Onderzoek Vlaanderen [G.0135.15N, G0C1213N, G.0090.11N] and Onderzoeksraad of the Vrije Universiteit Brussel [OZR2232]

References

1. Buts L, Lah J, Dao-Thi MH, Wyns L & Loris R. Toxin-antitoxin modules as bacterial metabolic stress managers. *Trends Biochem Sci.* 2005, 30, 672.
2. Hadzi S, Garcia-Pino A, Haesaerts S, Dukas J, Gerdes K, Lah J, Loris R. Ribosome-dependent Vibrio cholerae mRNAse HigB2 is regulated by a b-strand sliding mechanism . *Nucleic Acid Research* 2017, 45, 2937.

Molecular flexibility of myelin proteins enables tight packing of lipid multilayers

[a] ARNE RAASAKKA, [a] PETRI KURSULA

[a] *Department of Biomedicine, University of Bergen, Norway*

The correct functioning of the vertebrate nervous system requires the rapid conduction of nerve impulses over large distances. This saltatory conduction is enabled by the myelin sheath, a tighly packed, multilayered membrane wrapped around selected axons. Myelin is formed of dozens of lipid bilayers, containing specific integral and peripheral membrane proteins that function in stacking of membrane multilayers. Although many myelin proteins were initially isolated >50 years ago, little is still known about their 3D structure or the molecular mechanisms of their function in myelin formation. Most myelin-specific proteins are involved in neurological disease, and the disease mechanisms involve a breakdown of protein-lipid interactions at the membrane surface, leading to myelin degradation and chronic disability.

Myelin basic protein (MBP) is one of the most central proteins that mediate lipid bilayer stacking in compact myelin. This positively charged intrinsically disordered protein manifests as many isoforms and post-translationally modified variants, and it is an essential component in the formation of myelin. MBP is a peripheral membrane protein that changes its properties upon membrane binding. We characterized the membrane binding mechanisms of MBP using model lipid membranes. Using an array of complementary methods, we investigated the folding of MBP upon membrane binding, which is mainly mediated by electrostatic interactions. Subsequent charge neutralization allows MBP to insert into the membrane, whereby the hydrocarbon tails also still affect the folding of MBP, but have little effect on affinity. Calorimetric data show that MBP can alter the transition temperature of the lipid tails, but only after membrane insertion. We verified the membrane insertion mode using neutron reflectometry, and we determined the membrane stacking mode of MBP using atomic force microscopy, electron microscopy, and small-angle X-ray diffraction. MBP presents itself as an amorphous dense layer spread along the membrane surface during the process of myelin formation, and above a critical concentration, this disordered protein triggers spontaneous lipid bilayer stacking. Our work allows the construction of a detailed model for the structural foundation of MBP-mediated myelination. The results will help in understanding myelin development as well as disease etiology in demyelinating diseases, such as multiple sclerosis.

Partner-mediated polymorphism of an intrinsically disordered protein

[a] CHRISTOPHE BIGNON, [a,b] FRANCESCA TROILO, [b] STEFANO GIANNI, [a] SONIA LONGHI

[a] *Aix-Marseille Univ, CNRS, Architecture et Fonction des Macromolécules Biologiques (AFMB) UMR 7257, Marseille, France.*

[b] *Istituto Pasteur Italia - Fondazione Cenci Bolognetti, Istituto di Biologia e Patologia Molecolari del CNR, Dipartimento di Scienze Biochimiche "A. Rossi Fanelli", Sapienza Università di Roma, 00185, Rome, Italy*

Many intrinsically disordered proteins (IDPs) undergo folding upon binding to their partner(s), with the disorder-to-order transition often concerning only short molecular recognition elements (MoREs). Some IDPs display a high extent of plasticity with respect to the partner, a behavior that challenges the role of preconfigured MoREs in the recognition process. The C-terminal, intrinsically disordered N_{TAIL} domain of the measles virus nucleoprotein [1] undergoes folding upon binding to the X domain (XD) of the viral phosphoprotein, with α-helical folding taking place within a MoRE that encompasses residues 486-502 and that is partly pre-configured as an α-helix in the free form [2-5]. Beyond XD, N_{TAIL} also binds the major inducible heat shock protein 70 (hsp70). Although the major hsp70-binding site has been already mapped to the MoRE [6], no structural information is available for the N_{TAIL}/hsp70 complex. Here, using mutational studies combined with a protein complementation assay based on green fluorescent protein (GFP) reconstitution [7], we have investigated the N_{TAIL}-hsp70 interaction. Results indicate that although N_{TAIL} uses the same binding region to bind the two partners, the binding mechanisms are not the same. Hsp70 is much more tolerant of N_{TAIL} substitutions than XD, with this higher tolerance likely coming at the cost of a lower affinity compared to the N_{TAIL}/XD pair. The N_{TAIL}/hsp70 interaction is more evolvable, i.e. the majority of substitutions lead to increased interaction strength. Furthermore, while an increased and a decreased α-helicity of the MoRE lead to enhanced and reduced interaction strength with XD, respectively, the extent of helicity of the free form of the MoRE has a negligible impact on hsp70 binding, suggesting that the MoRE does not adopt an α-helical conformation once bound to hsp70. Here, by showing that the conformation sampled by the free form of the MoRE of N_{TAIL} does not necessarily commit N_{TAIL} to adopt an α-helical conformation in the bound form, we provide an additional example of partner-mediated polymorphism and of the relative insensitiveness of the bound structure to the pre-recognition state. The present results therefore contribute to shed light on the molecular mechanisms by which IDPs recognize their partners.

References

1. S. Longhi, V. Receveur-Brechot, D. Karlin, K. Johansson, H. Darbon, D. Bhella, R. Yeo, S. Finet, B. Canard. *J. Biol. Chem.*, 278 (2003) 18638-18648.
2. K. Johansson, J.M. Bourhis, V. Campanacci, C. Cambillau, B. Canard, S. Longhi. *J Biol Chem*, 278 (2003) 44567-44573.
3. S. Gely, D.F. Lowry, C. Bernard, M. Ringkjobing-Jensen, M. Blackledge, S. Costanzo, H. Darbon, G.W. Daughdrill, S. Longhi. *J Mol Recognit*, 23 (2010) 435-447.
4. V. Belle, S. Rouger, S. Costanzo, E. Liquiere, J. Strancar, B. Guigliarelli, A. Fournel, S. Longhi. *Proteins: Structure, Function and Bioinformatics*, 73 (2008) 973-988.
5. M. Ringkjøbing Jensen, G. Communie, E.D. Ribeiro, Jr.,, N. Martinez, A. Desfosses, L. Salmon, L. Mollica, F. Gabel, M. Jamin, S. Longhi, R.W. Ruigrok, M. Blackledge. *Proc Natl Acad Sci U S A*, 108 (2011) 9839-9844.
6. M. Couturier, M. Buccellato, S. Costanzo, J.M. Bourhis, Y. Shu, M. Nicaise, M. Desmadril, C. Flaudrops, S. Longhi, M. Oglesbee. *J Mol Recognit*, 23 (2010) 301-315.
7. A. Gruet, M. Dosnon, D. Blocquel, J. Brunel, D. Gerlier, R.K. Das, D. Bonetti, S. Gianni, M. Fuxreiter, S. Longhi, C. Bignon. *FEBS J*, 283 (2016) 576-594.

Study of domain interactions of intrinsically disordered Microtubule-Associated Protein 2c (MAP2c) with its binding partners via computational methods and nuclear magnetic resonance

[a,b] KATEŘINA MELKOVÁ, [a,b] VOJTĚCH ZAPLETAL, [a] SÉVERINE JANSEN, [a] JOZEF HRITZ, [a,b] LUKÁŠ ŽÍDEK

[a] *Central European Institute of Technology, Masaryk University, Kamenice 5, CZ-625 00, Brno, Czech Republic.*
[b] *National Centre for Biomolecular Research, Faculty of Science, Masaryk, Žerotínovo nám. 617/9, CZ-601 77, Brno, Czech Republic.*

Intrinsically disordered proteins (IDPs) are macromolecules without stable spatial order whose conformation can flexibly and quickly change. Given the disordered nature of IDPs, NMR is the key experimental method for studying IDPs since it provides direct information on local structure and internal motions in the molecule. IDPs play very important role in wealth of cell processes including nervous system development and function, and many of them are involved in the onset of neurodegenerative diseases.

The main focus of the presented study is microtubule-associated protein 2c (MAP2c), an IDP regulating structure and dynamics (polymerization and degradation) of microtubules (MTs) and actin crosslinking essential for correct function of neurons and other cells. MAP2c is a 49 kDa protein consisting of several structural and functional regions. The N-terminal part contains two important regions: The N-terminal region with a high content of negatively charged amino acids and the proline-rich region. The former segment contains a proposed binding site for steroids, while the latter one interacts with SH3 domain of plectin [1] which act as a cytolinker and regulates actin dynamics. The second important part of MAP2c is a highly-conserved C-terminal domain that binds to MTs.

We described experimentally structural features, dynamics, phosphorylation, and interactions of MAP2c with the atomic resolution. Using multidimensional non-uniformly sampled NMR experiments, we were able to study the full-length protein with sufficient resolutions [2]. In addition, we applied the "divide-and-conquer" approach, partitioning the MAP2c protein into smaller fragments differing the presence or absence of individual functional domains. In order to compare the experimental data with molecular dynamics simulations, optimized for disordered proteins.

Acknowledgement: This work was supported by the program VES17 INTER-EXCELLENCE, subprogram INTER-COST.

References

1. R. G. Valencia *et al.*, "Intermediate filament-associated cytolinker plectin 1c destabilizes microtubules in keratinocytes," *Mol. Biol. Cell*, vol. 24, no. 6, pp. 768–784, Mar. 2013.
2. S. Jansen *et al.*, "Quantitative mapping of microtubule-associated protein 2c (MAP2c) phosphorylation and regulatory protein 14-3-3ζ-binding sites reveals key differences between MAP2c and its homolog Tau," *J. Biol. Chem.*, vol. 292, no. 16, pp. 6715–6727, Apr. 2017.

Exploring functional plasticity of protein proteostasis network in bacteria using in-vivo cross linking coupled with mass spectrometry

[a] RAVIT MESIKA , [a] RESHEF MINTZ, [a] DANA REICHMANN

[a] *Department of Biological Chemistry, The Hebrew University of Jerusalem, Jerusalem, Israel*

In this project, our lab established an efficient state-of-the-art technology coupling *in vivo* crosslinking with mass spectrometry to reveal direct protein interaction and map specific cellular pathways. This method is based on the incorporation of UV-inducible non-native amino acid BPA into a target protein.

In order to characterize stress-specificity of proteostasis network in bacteria, we applied this new methodology on the thermolabile *Photinus pyralis* Luciferase (PP), which undergoes unfolding during stress conditions, and serve as a „chaperone prey" protein. We designed few variants of PP-Luciferase protein, each contains BPA at different location along the protein sequence leading to variable thermostability. Then we mapped interactome of these variants in-vivo during cold and heat conditions in *E.coli*.

This analysis reveled the stress-related interactome of the thermolabile protein, revealing known and novel members of the proteostasis in *E.coli*. We believe, that this study will allow us to characterize substrate and stress specificity of components of the proteostasis network. We aim to identify key bacterial proteins that assist bacteria in adaptation and stress response, which is a crucial factor in mammalian associated bacteria.

Therefore, outcomes of this project might lay the foundations to a new generation of safe and efficient probiotic and antibiotics.

Conformational diversity and disorder transitions reveals adaptations in different kingdoms

[a] ALEXANDER MIGUEL MONZON, [a] GUSTAVO PARISI, [a] MARÍA SILVINA FORNASARI

[a] *Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Buenos Aires, Argentina*

Protein motions are a key feature to understand biological function. Following how disorder/order transitions are distributed among conformers we found that intrinsically disordered proteins (IDPs) could be split in partially disordered proteins (PDPs), that have on average 67% of their conformers with disordered regions and an average RMSD=1.1 Å, while the so-called malleable proteins (MP) have on average only 25% of disordered conformers and RMSD=1.3 Å [1]. Moreover, disordered regions tend to be mostly disordered in all the conformers in PDPs and mostly ordered in MPs. As it is well established that different kingdoms of life and viruses show different content of disordered [2], in this work we studied how the percentage of disordered residues associated with PDPs and MPs reveals taxonomic adaptations. To this end, we collected 2100 proteins with different degrees of conformational diversity which show at least one disordered region in at least one of their known conformers extracted from CoDNaS database [3]. We found that the conformational diversity distribution is almost identical among the organisms as well as the proportion of PDPs and MPs (about 70% and 30% respectively). However, proteins from different organisms show a differential behavior in terms of disorder transitions among their conformers. Eukaryota proteins show the maximum percentage of disordered residues followed by Archaea and Bacteria and lastly Viruses. These results in part contradict the established taxonomic distribution of disorder among organisms, because Eukaryota and Viruses are the organisms showing higher proportions of disorder. Furthermore, Viruses and Bacteria show similar percentages of disordered residues associated with PDPs (mostly disordered among conformers) followed by Archaea proteins. Also, in terms of MPs associated disorder (mostly ordered among conformers), we found that Viruses are the organisms with the minor fraction followed by Archaea and Bacteria. Again, Eukaryota show the higher proportion of disordered residues associated to MPs. These differential proportions in the content of disordered regions classified accordingly with the conformational distributions per protein could contribute in our understanding of functional adaptations or differential roles of disorder in the different organisms.

References

1. Monzon AM, Zea DJ, Fornasari MS, Saldaño TE, Fernandez-Alberti S, Tosatto SCE, et al. Conformational diversity analysis reveals three functional mechanisms in proteins. PLoS Comput Biol. 2017;13: e1005398.
2. Uversky VN. A decade and a half of protein intrinsic disorder: biology still waits for physics. Protein Sci. 2013;22: 693–724.
3. Monzon AM, Rohr CO, Fornasari MS, Parisi G. CoDNaS 2.0: a comprehensive database of protein conformational diversity in the native state. Database. 2016; 2016. doi:10.1093/database/baw038

Homology modeling in a dynamical world

[a] ALEXANDER MIGUEL MONZON, [b] DIEGO JAVIER ZEA, [b] CRISTINA MARINO-BUSLJE, [a] GUSTAVO PARISI

[a] *Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, CONICET, Bernal, Argentina*
[b] *Structural Bioinformatics Unit, Fundación Instituto Leloir, CONICET, C1405BWE, Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina*

A key concept in Template-Based Modeling (TBM) is the high correlation between sequence and structural divergence. The main practical consequence of this correlation is that homologous proteins that are similar at the sequence level will also be similar at the structural level allowing the selection of a proper template for a target sequence. Pioneering work by Chothia and Lesk [1] found a non-linear and well correlated relationship between sequence and structural divergence. However, a given protein sequence could exists in different structures (conformers) where their structural differences describe their conformational diversity (CD). In this work, we explored the impact that CD has on the relationship between structural and sequence divergence.

CoDNaS database [2] was used to recruit proteins exhibiting conformational diversity. Maximum conformational diversity for each protein is the maximum C-alpha RMSD derived from all conformers pairwise comparisons. Using this set, we ran BLASTClust to obtain all available clusters at 30% of local sequence identity. The final dataset contains 2024 different protein chains with a total of 37755 conformers. These proteins are grouped in 524 families.

To estimate the structural divergence for each homologous protein pairs in a cluster, we calculated the C-alpha RMSD for all possible pairs of conformers belonging to the proteins being compared. Additionally, we calculated the percent of sequence identity for each homologous protein pairs using a global sequence alignment. The total comparisons among all vs all conformers for each homologous protein pairs and structures of the same protein give an amount of ~4 millions of pair.

We found that the use of a highly redundant sequence dataset (that is, considering the CD) blurs the well-established relationship between sequence and structure divergence more than shown in previous studies. It is also evident that the extent of conformational diversity can be as high as the maximum structural divergence among families reached by accumulation of nonsynonymous substitutions. Also, the presence of CD impairs the well-established correlation between sequence and structural divergence, which is more complex than previously suggested due to the existence of different structures (CD) for the same sequence. However, we found that this noise can be resolved using a priori information from the structure-function relationship. We showed that protein families with low CD, which we called “rigid” [3], show a well-correlated relationship between sequence and structural divergence (Spearman’s rank correlation rho of -0.83), which is severely reduced in proteins with larger CD (Spearman’s rho = -0.51).

This lack of correlation could impair template-based modeling (TBM) results in highly dynamical proteins due to the uncertainty to select a proper target structure. Finally, as proteins with disordered regions show higher extensions of CD, we also found that the presence of order/disorder can provide useful beforehand information for better template selection and TBM performance.

References

1. Chothia C, Lesk AM. EMBO J. 1986.
2. Monzon AM, Rohr CO, Fornasari MS, Parisi G. Database. 2016.
3. Monzon AM, Zea DJ, Fornasari MS, Saldaño TE, Fernandez-Alberti S, Tosatto SCE, et al. PLoS Comput Biol. 2017.

Coloring protein darkness

[a] NECCI M., [a] PIOVESAN D., [a,b,*] TOSATTO S.C.E.

[a] *Department of Biomedical Sciences, University of Padua, Viale G. Colombo 3, 35121 Padova, Italy.*

[b] *CNR Institute of Neuroscience, Viale G. Colombo 3, 35121 Padova, Italy.*

[*] *To whom correspondence should be addressed: silvio.tosatto@unipd.it*

Since dark regions in proteomes have been spotted a few years ago (1), effective approaches to improve sequence annotation are still missing. Dark residues are defined as regions inaccessible to homology modeling and residues annotated by Pfam cover only ca. 50% of eukaryotic proteomes. The remainder mostly corresponds to intrinsically disordered regions [1]. While unknown folds and transmembrane regions require ab initio methods or experimental determination, ID regions can be predicted based on amino acid composition. Here, we provide an extensive analysis of ID in the protein universe based on the UniProt database from sequence-based predictions in MobiDB [2]. The charged residue composition was used to classify ID proteins by structural propensities. We present a statistical analysis of ID in the protein universe considering features associated with different ID flavors, content, region length and position, number of ID regions, taxonomic distribution and GO enrichment [3]. The normalized distribution of long ID regions among proteins in different domains of life shows an increase from bacteria, archaea, viruses, towards eukaryotes. Bacteria, despite being the most represented class in UniProt, display the lowest ID content. Archaea show slightly more, but ID content is quite variable among phyla and species as ID correlates with peculiarities of archaea species prospering at extreme conditions. As previously noted, eukaryotes are the most disordered domain of life. This is often speculated to be due to the increased need for control and regulation. Viruses use ID for multiple reasons, for example, to exploit ID advantages in one-to-many and many-to-one interactions. Adapting quickly to hostile environmental changes and effective use of a compact genome have been also postulated to favor ID in viruses. The four main classes defined by Pappu [4], with Swollen Coils split into highly negative and highly positive instances, were functionally characterized by means of a GO term enrichment to highlight molecular functions specific to one class among ID functions. While this classification itself is a step towards the identification of different ID flavors, we tried to characterize them further by calculating six additional sequence features typically associated with ID. Interestingly, a similar pattern is observed between Coils & Hairpins and Swollen Coils and between Globules & Tadpoles and Undefined classes. This suggests that undefined regions are more similar to Globules than Coils but have a bias towards low complexity compared to Globules. Notably many class members escape this sequence feature classification, with the exception of Swollen Coils, indicating that ID sequence diversity is a continuum that cannot be represented by extreme cases only. The MobiDB-lite pipeline used to annotate ID has been recently included in the InterPro database, significantly reducing the number of dark residues without annotation [5].

References

1. Perdigão,N., *et al.* (2015) *Proc. Natl. Acad. Sci.*, 112(52):15898-903
2. Necci, M., *et al* (2017). *Bioinformatics.*
3. Necci, M., *et al*. (2016). *Protein Sci.* 25(12):2164-2174
4. Das, R.K. and Pappu, R.V. (2013). *Proc. Natl. Acad. Sci.*, 110:13392–13397.
5. Finn, R.D., *et al.* (2017). *Nucleic Acids Res.*, 45:D190–D199.

Analysis of cell-cycle regulatory linear motifs bound by the pRb retinoblastoma tumor suppressor

[a,b] NICOLAS PALOPOLI, [c] LUCÍA B. CHEMES, [d] BENJAMIN LANG, [d] MALVIKA SHARAN, [e] M. MADAN BABU, [d] TOBY J. GIBSON

[a] *Structural Bioinformatics Group, National University of Quilmes - CONICET, Buenos Aires, Argentina.*

[b] *Structural Bioinformatics Unit, Leloir Institute and IIBBA-CONICET, Buenos Aires, Argentina.*

[c] *Protein Structure Function and Engineering Laboratory, Leloir Institute and IIBBA-CONICET, Buenos Aires, Argentina.*

[d] *Structural and Computational Biology Unit, EMBL, Heidelberg, Germany.*

[e] *MRC Laboratory of Molecular Biology, Cambridge, United Kingdom.*

Linear motifs are key elements of regulatory protein interaction networks. Viruses take advantage of this by evolving linear motifs that hijack and disrupt the host network. Therefore, identifying relevant linear motifs and their binding strength is central to understanding host regulation and viral pathogenesis. However, the degenerate nature of linear motifs makes their identification challenging. The pRb retinoblastoma tumor suppressor protein is a central cell-cycle regulator inactivated in human cancer and targeted by carcinogenic viruses including human papillomavirus and Merkel cell polyomavirus. Many pRb interactions are mediated by the LxCxE motif, present in host and viral proteins.

In this study, we undertook a comprehensive analysis of the pRb-binding LxCxE motif in order to improve our ability to identify binders and our understanding of the pRb interaction network. We found that the LxCxE motif was associated with a diversity of protein functions and architectures (Fig. 1). Analysis of sequence alignments representing 1,150 motifs from 13 host proteins and 577 motifs from 6 viral families allowed us to expand the [LI].C.[DE] definition with novel features that account for all relevant contact sites. The positioning of these features differed between viral and cellular motifs, and appeared more constrained in viruses. Host L/C/E core residues were strongly conserved compared to the disordered regions within which they were embedded ($p=7.3*10-7$, Wilcoxon-ranked test) and were under similar constraint than structural residues ($p=0.72$).

We have identified novel features that expand the definition of the LxCxE motif and will improve its proteome-wide identification. Host LxCxE motifs are highly conserved within vertebrates, suggesting that they play a key role in the pRb network. The central role of the LxCxE motif in viral pathogenesis is underscored by its striking conservation across plant and animal viruses.

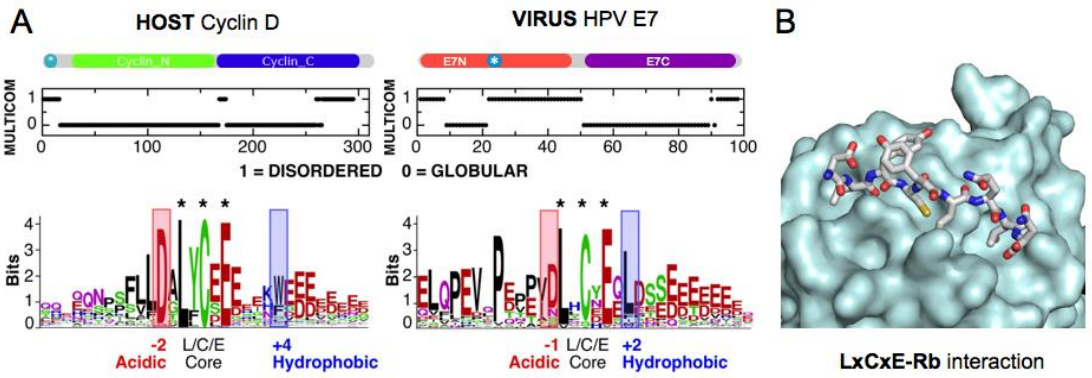

Figure 1. Structural and sequence features of the LxCxE motif. A) Top: Schematic representation of LxCxE-containing proteins. Domains are depicted as colored boxes and LxCxE motifs as cyan circles. Disorder predictions used MULTICOM. Bottom: Sequence logos for the Cyclin D and HPV E7 motifs showing the positioning of features (colored boxes) with respect to the L/C/E core (black asterisks). B) Structure of the pRb binding site showing the LxCxE peptide in stick representation.

Identification of intrinsically disordered regions and linear motif mimicry across Flavivirus proteomes

[a,*] JULIANA GLAVINA, [b,c,*] NICOLÁS PALOPOLI, [d] LUCÍA B. CHEMES

[a] *Protein Physiology Laboratory, Department of Biological Chemistry, Faculty of Exact and Natural Sciences, University of Buenos Aires, Buenos Aires, Argentina.*

[b] *Structural Bioinformatics Group, National University of Quilmes - CONICET, Buenos Aires, Argentina.*

[c] *Structural Bioinformatics Unit, Leloir Institute and IIBBA-CONICET, Buenos Aires, Argentina.*

[d] *Protein Structure Function and Engineering Laboratory, Leloir Institute and IIBBA-CONICET, Buenos Aires, Argentina.*

* *contributed equally*

Flaviviruses are a major threat for public health due to their high epidemic potential and the presence of multiple recent outbreaks in Latin America. Despite their clinical relevance, specific antiviral treatments are still lacking. Short Linear Motifs (SLiMs) mediating protein-protein interactions are appealing candidates as novel therapeutic targets, since these interactions rely on defined sequences and sufaces that could be targeted by specific inhibitory drugs. Viruses commonly manipulate the cell machinery by mimicking host SLiMs. However, while SLiMs are mostly present in intrinsically disordered regions (IDRs) of proteins, SLiMs and disordered regions in Flavivirus proteins are still poorly studied.

Our goal was to define IDRs and identify putative SLiMs comprehensively across Flaviviridae groups including Dengue, Yellow Fever and Zika. We intend to infer the conservation of IDRs within and outside groups, identify SLiMs in these regions and assess their potential for binding cellular proteins by inspecting their sequence and structure properties.

A comprehensive set of 128 sequences from Flaviviridae displayed intrinsic disorder ranging from 5-15% in different groups. Analysis of 13 proteins from 72 Flavivirus species and 361 corresponding structures revealed many conserved IDRs in several viral proteins including Capsid, NS3, NS2B and NS5, with small (5-30 residues) intrinsically disordered domains and structured, yet flexible regions within globular domains. While some IDRs were conserved across all groups, certain viral strains displayed unique disorder characteristics. A preliminary identification of SLiMs showed strongly conserved candidates, including those within the N-terminus of NS3, a region required for Dengue viral assembly and budding.

Our analysis of Flavivirus proteomes reveals the presence of IDRs and several putative SLiMs that could be key for understanding poorly described aspects of these pathogens.

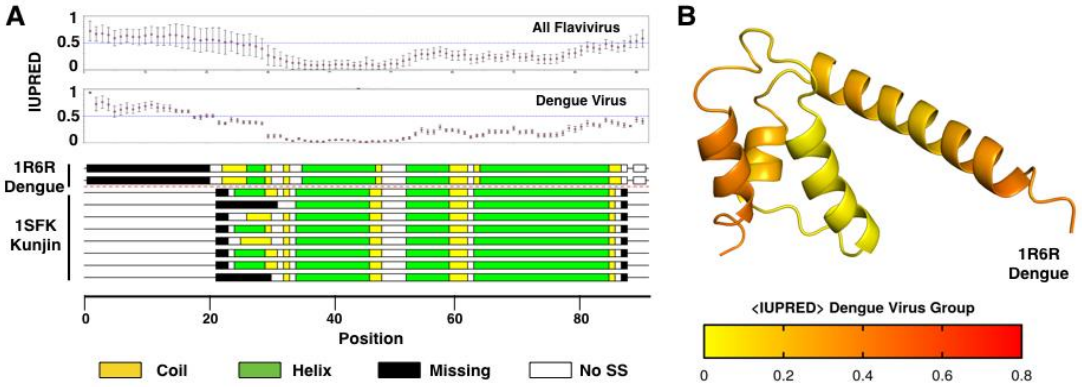

Figure 1. – Example analysis of Capsid protein. A) Top panel: Intrinsic disorder per position across all Flavivirus and for Dengue only. Bottom panel: Secondary structure regions observed in representative structures of Dengue and Kunjin virus. B) Mapping of IUPred disorder scores in a representative Dengue Capsid protein structure.

Conformational properties of Short Linear Motifs

[a,b] NICOLAS PALOPOLI, [b] CRISTINA MARINO-BUSLJE, [a] GUSTAVO PARISI

[q] *Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes; CONICET. Roque Saenz Peña 352, Bernal, (B1876BXD), Buenos Aires, Argentina.*

[b] *Fundación Instituto Leloir, Instituto de Investigaciones Bioquímicas de Buenos Aires; CONICET. Av. Patricias Argentinas 435, (C1405BWE), Buenos Aires, Argentina.*

Short Linear Motifs (SLiMs) are small peptides mediating protein-protein interactions that follow a recurrent pattern of fixed and variable positions. They have generally been studied on their sequence properties and assuming they occur in intrinsically disordered regions. Recent experimental and computational evidence indicate that many SLiMs display characteristic structural features which may impact on the formation and regulation of domain-motif interactions.

We have built a dataset of available SLiM structures in order to distinguish their conformational preferences and infer their functional implications. We extended the number of structural descriptions in the Eukaryotic Linear Motif database (ELM; elm.eu.org) [1] by searching the Protein Data Bank (PDB; www.rcsb.org) [2] for all exact occurrences of known motif definitions and re-mapping them to the corresponding Uniprot sequences. Amino acid composition patterns, secondary structure features and accessible surface area were calculated for all motif positions and their flanking regions. Disordered residues in SLiM instances were identified as missing coordinates in the PDB files. The extended dataset greatly increases the number of SLiM instances with known structural information and provides a higher coverage in the represented ELM classes and Uniprot sequences. We observed differences in the structural properties of fixed, degenerate and more variable positions of the SLiM and also in comparison with their flanking regions. We found a 5-fold increase in PDB diversity and in the number of SLiM entries with 2 or more PDBs, allowing us to inspect the conformational diversity of these motifs that highlight alternative structures in the native state.

Overall, our analysis is allowing us to understand structural characteristics of SLiMs that differentiate them from similar peptide sequences not involved in protein interactions.

Acknowledgement: This work was supported by National Scientific and Technical Research Council in Argentina (CONICET), Universidad Nacional de Quilmes and Fundación Instituto Leloir.

References

1. Dinkel H, Van Roey K, Michael S, Kumar M, Uyar B, Altenberg B, Milchevskaya V, Schneider M, Kühn H, Behrendt A, Dahl SL, Damerell V, Diebel S, Kalman S, Klein S, Knudsen AC, Mäder C, Merrill S, Staudt A, Thiel V, Welti L, Davey NE, Diella F, Gibson TJ. ELM 2016--data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Res*. 2016 Jan 4;44(D1):D294-300.
2. Berman HM1, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res*. 2000 Jan 1;28(1):235-42.

Assessing structural constraints in intrinsically disordered proteins using evolutionary methods

[a] JULIA MARCHETTI, [a] MARÍA SILVINA FORNASARI, [a] GUSTAVO PARISI

[a] *Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes (CONICET), Bernal, Buenos Aires, Argentina.*

Intrinsically disordered regions (IDRs) or proteins (IDPs) are characterized by their high flexibility and mobility but clearly related with well-established disorder-based biological functions [1]. These proteins lack a well defined three-dimensional structure under physiological conditions. Intrinsic disorder extension is further complicated by the presence of conditionally disordered regions and/or partially disordered regions. These regions alternate their state between disordered and structured or ordered dependending on the conditions and environment (i. e. binding of ligands, pH, oligomerization state) [2]. Moreover, these order/disorder transitions are difficult to estimate because they depend on the conformers being studied [3]. In this work we estimated the fraction of structural information content of IDPs using evolutionary methods. For this purpose we used a model of protein evolution that takes into account the structure/s of a protein to estimate substitution matrices, SCPE, from structurally constrained model of protein evolution [4]. SCPE derives substitution matrices that describe the pattern of amino acid substitution in a given position of a protein. Using maximum likelihood (ML) estimations, it is possible to compare the performance of SCPE matrices with unconstrained models of protein evolution (i.e. JTT). Using structural ensembles of IDPs derived from PED database [5] we found that IDPs proteins have in average between 30 to 40% of structurally constrained positions. Importantly ML estimations are not correlated with the number of contacts derived from the ensembles, evidencing that the evolutionary analysis (using homologous proteins and phylogenetic topology) is more sensitive than the simple inspections of structural ensembles. We also found that the inclusion of structural information in the substitution matrices provides a more reliable description of the substitution pattern in IDPs proteins.

References

1. Lee *et al* “Classification of Intrinsically Disordered Regions and Proteins.” *Chemical Reviews* 114 (13): 6589–6631.
2. DeForte *et al* “Resolving the Ambiguity: Making Sense of Intrinsic Disorder When PDB Structures Disagree.” *Protein Science: A Publication of the Protein Society* 25 (3): 676–88.
3. Monzon *et al* “Conformational Diversity Analysis Reveals Three Functional Mechanisms in Proteins” 435: 1–29.
4. Parisi *et al* “Structural Constraints and Emergence of Sequence Patterns in Protein Evolution.” *Molecular Biology and Evolution* 18 (5): 750–56.
5. Varadi *et al* 'pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins' Nucleic Acids Res. 2014 Jan (Database issue); Epub 2013 Oct 29.

The role of CTCF and cohesin complex in chromatin looping and higher-order organization of Human genome

[a,b,c,*] DARIUSZ PLEWCZYNSKI, [a,c] PRZEMYSŁAW SZAŁAJ, [a] MICHAŁ SADOWSKI, [b] ZHONGHUI TANG, [b] YIJUN RUAN

[a] *Centre of New Technologies, Warsaw University, Warsaw, Poland*
[b] *The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA,*
[c] *Centre for Innovative Research, Medical University of Bialystok, Białystok, Poland*
** Presenter: Dariusz Plewczynski (e-mail: d.plewczynski@cent.uw.edu.pl)*

The recent genomic and bioimaging insights in the higher order chromatin organisation in human nucleus motivated us to propose novel simulation method aimed at the effective prediction of three-dimensional human genome structure. We applied an advanced long-read Chromatin Interaction Analysis by Paired-End Tag Sequencing (ChIA-PET) experimental strategy [1] combined with computational modelling [2] to comprehensively map higher-order chromosome folding and specific chromatin interactions mediated by CCCTC-binding factor (CTCF), cohesin and RNAPII with haplotype specificity and nucleotide resolution in different human cell lineages [3]. We demonstrate the effectiveness of biophysical modeling in building 3D genome models at multiple levels, including the entire genome, individual chromosomes, and specific segments at megabase (Mb) and kilobase (kb) resolutions of single average and ensemble structures.

We find that CTCF/cohesin-mediated interaction anchors serve as structural foci for spatial organization of constitutive genes concordant with CTCF-motif orientation, whereas RNAPII interacts within these structures by selectively drawing cell-type- specific genes towards CTCF-foci for coordinated transcription. Furthermore, we show that haplotype-variants and allelic-interactions have differential effects on chromosome configuration influencing gene expression and may provide mechanistic insights into functions associated with disease susceptibility [1]. 3D-genome simulation suggests a model of chromatin folding around chromosomal axes, where CTCF is involved in defining the interface between condensed and open compartments for structural regulation [2]. Our 3D-genome strategy thus provides unique insights in the topological and thermodynamical mechanisms [4] of human variations and diseases.

Acknowledgement: DP, MS, PS are supported by grants from the Polish National Science Centre (2014/15/B/ST6/05082), the European Cooperation in Science and Technology action (COST BM1405 and BM1408). All authors were supported by grant 1U54DK107967-01 "Nucleome Positioning System for Spatiotemporal Genome Organization and Regulation" within 4DNucleome NIH program.

References

1. "CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription" Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, Szalaj P, Trzaskoma P, Magalska A, Wlodarczyk J, Ruszczycki B, Michalski P, Piecuch E, Wang P, Wang D, Tian SZ, Penrad-Mobayed M, Sachs LM, Ruan X, Wei CL, Liu ET, Wilczynski GM, Plewczynski D, Li G, Ruan Y. Cell. 2015 Dec 17;163(7):1611-27. doi: 10.1016/j.cell.2015.11.024.
2. "An integrated 3-Dimensional Genome Modeling Engine for data-driven simulation of spatial genome organization" Szałaj P, Tang Z, Michalski P, Pietal MJ, Luo OJ, Sadowski M, Li X, Radew K, Ruan Y, Plewczynski D. Genome Res. 2016 Dec;26(12):1697-1709.
3. "3D-GNOME: an integrated web service for structural modeling of the 3D genome" Szalaj P, Michalski PJ, Wróblewski P, Tang Z, Kadlof M, Mazzocco G, Ruan Y, Plewczynski D. Nucleic Acids Res. 2016 Jul 8;44(W1):W288-93. doi: 10.1093/nar/gkw437.
4. "Free energy based high-resolution modeling of CTCF-mediated chromatin loops for human genome" Dawson W and Plewczynski D. [http://biorxiv.org/content/early/2017/03/08/105676] BIORXIV/2017/105676.

Insights into (Un)structure-Function Relationships from the Conditionally Disordered Chaperone Hsp33

[a] ODED RIMON, [a] DANA REICHMANN

[a] *The Hebrew University of Jerusalem, Israel*

Cells regularly encounter environmental challenges that have potentially fatal consequences. Accordingly, evolution has favoured the development of systems that defend cells and their proteomes from stress. A major constituent of these defence systems is a network of interrelated proteins known as molecular chaperones. These proteins work together to prevent protein misfolding and aggregation and maintain the rest of the proteome in a healthy state. One protein that belongs to this network is Hsp33, a highly conserved ATPindependent chaperone that protects organisms ranging from bacteria to unicellular eukaryotes from the toxic effects of oxidative stress. When oxidation occurs, Hsp33 is positioned at the colloquial 'front lines' and undergoes redox-dependent unfolding. Intriguingly, however, this unfolding does not damage the chaperone's activity. On the contrary, Hsp33 gains its function when it loses its fold. It is only under oxidative stress that Hsp33 can bind its clients and protect them from aggregation – normal reducing conditions render it inactive. The unfolding activation mechanism of Hsp33 is triggered by its redoxsensing C-terminal domain, which unfolds in response to stress and destabilizes the adjacent metastable linker region [1]. However, the exact roles of the metastable linker and its orderto- disorder transition in activation, client binding, and client release are still unclear. We reveal that the metastable linker region of Hsp33 is a regulatory inactivator, preserving its monomeric architecture in the reduced state and mediating crosstalk with the synergistic ATP-dependent chaperone system DnaK/J-GrpE [2]. We utilise state of the art structural mass-spectrometry techniques to study the interactions between Hsp33 and its client proteins, and reveal an amazing trait which enables Hsp33 to work as a chaperone even without one of its most important client binding regions. We create chimeric proteins of Hsp33 with the human protein acetylcholinesterase, which exhibits chaperone-like activities in neurons, and find an intriguing difference between Hsp33's interactions with different types of aggregate-forming clients. Finally, we describe a software tool that we developed for quicker and more accurate analysis of chaperone activity assay results [3].

Acknowledgement: To the labs of Hermona Soreq (HUJI), Dina Schneidman (HUJI), Meytal Landau (Technion, Israel), Andrea Sinz (MLU, Germany) and Patrick Griffin (Scripps, FL, USA) for fruitful collaborations.

References

1. Reichmann D, Xu Y, Cremers CM, Ilbert M, Mittelman R, Fitzgerald MC and Jakob U (2012) Order out of Disorder: Working Cycle of an Intrinsically Unfolded Chaperone. *Cell* 148, 947-957
2. Rimon O, Süss O, Goldenberg M, Fassler R, Yogev O, Amartely H, Propper G, Friedler A, Reichmann D (2017) A role of metastable regions and their connectivity in the inactivation of a redox-regulated chaperone and its inter-chaperone crosstalk. *Antioxid. Redox Signal.* DOI: 10.1089/ars.2016.6900
3. Rimon O and Reichmann D (*in review*) Kfits: A software framework for fitting and cleaning noisy kinetic measurements. *Bioinformatics (in review)*

The structural landscape of two new disordered BH3-like proteins and their interaction with the Bcl-2 family members

[a] VALENTINA SORA, [a] MATTEO LAMBRUGHI, [a] MADS NYGAARD,
[b] FLAVIE STRAPPAZZON, [a] DANIELA DE ZIO, [c] BIRTHE KRAGELUND,
[a] FRANCESCO CECCONI, [a] JIRI BARTEK, [a] ELENA PAPALEO

[a] *Danish Cancer Society Research Center, Copenhagen, Denmark*
[b] *IRCCS Santa Lucia Foundation, Rome, Italy*
[c] *Department of Biology, University of Copenhagen, Copenhagen, Denmark*

The delicate balance between promoters and inhibitors in the mitochondrial apoptotic pathway is crucial in order to determine whether the cell will survive or not, responding to different stimuli. The understanding of the fine-tuned pathways leading to or preventing cell death is fundamental in cancer research, because the failure of these systems could eventually lead to the onset of the disease. One of the best-known protein families involved in the initiation of programmed cell death is the Bcl-2 family, whose members share with the parent Bcl-2 one or more sequence motifs (called *B* cl-2 *H* omology motifs and numbered 1-4) involved in protein-protein interactions. One of them, the BH3 motif, has increasingly come to the research forefront both because BH3 mimetics have shown to have therapeutic value [1] and because of the presence of this motif in Bcl-2-unrelated proteins [1], which could subsequently play a role in the apoptotic pathway by interacting with Bcl-2 family members. In this context, our group is currently characterizing the binding mode to Bcl-2 and Bcl-XL of a new class of BH3-like proteins including p14-ARF [2] and AMBRA1 [3], which are intrinsically disordered and have remote similarity to the canonical BH3 motif. In particular, we are using different computational approaches to identify structural propensity in the bound and unbound state of these two proteins and the key residues for the binding to the Bcl-2 family members, as well as to study the impact of germline or somatic cancer mutations.

Acknowledgements: This project is supported by a PRACE-DECI14th grant (CancerBH3), ISCRA-CINECA and a grant from LEO Foundation.

References

1. Aouacheria, C. Combet, P. Tompa, J. M. Hardwick, *Redefining the BH3 death domain as a "Short Linear Motif"*, Trends Biochem. Sci., 40(12):736-748, doi:10.1016/j.tibs.2015.09.007
2. C. Christensen, J. Bartkova, M. Mistrík, A. Hall, M. K. Lange, U. Ralfkiær, J. Bartek, P. Guldberg, *A short acidic motif in ARF guards against mitochondrial dysfunction and melanoma susceptibility* , Nature Comm., 5:5348, doi:10.1038/ncomms6348
3. A. Di Rita, F. Strappazzon, *AMBRA1, a Novel BH3-Like Protein: New Insights Into the AMBRA1-BCL2-Family Proteins Relationship*, Int. Rev. Cell. Mol. Biol., 2017;330:85-113

Conformational study on binding domains of the intrinsically disordered protein MAP2c (microtubule-associated protein 2c)

[a] VOJTĚCH ZAPLETAL, [a] KATEŘINA MELKOVÁ, [a] SÉVERINE JANSEN,
[b] MARTIN BLACKLEDGE, [a] LUKÁŠ ŽÍDEK

[a] *CEITEC-MU, Masaryk University, Kamenice 5, Brno 625 00, Czech Republic*
(vojtis@mail.muni.cz)

[b] *Institut de Biologie Structurale, CEA, CNRS, University Grenoble Alpes, Grenoble 38044, France*

Microtubule associated protein 2c (MAP2c) is an intrinsically disordered protein (IDP) regulating stability and dynamics of microtubules in the cytoskeleton of neurons and other cells. MAP2c is a 49 kDa homolog of tau protein which differs in localization, expression, and binding properties [1, 2]. Important structural and functional regions were revealed in MAP2c.

Typical experimental methods for studies of IDPs are nuclear magnetic resonance (NMR) and small angle X-ray scattering (SAXS). We used 5D NMR experiments to obtain chemical shifts. We also performed paramagnetic relaxation enhancement (PRE) experiments to get the information about intramolecular distances and small angle x-ray scattering experiments to monitor the overall size and shape of the MAP2c in solution. The pools of possible chain conformations were generated by the program Flexible-Meccano, which builds consecutively a polypeptide chain assuming that peptide planes are rigid entities connected through Cα atoms. The ensemble of structures was selected by the program ASTEROIDS which uses a genetic algorithm to produce and ensemble of IDPs consistent with experimental data, measured with a simple quadratic fitness function measuring the deviation of the back-calculated parameters from the measured ones. From the selected ensemble, we prepared three different starting structures (containing the N-terminal domain and microtubule binding domain, microtubule binding domain or N-terminal domain) to generate structural ensembles using molecular dynamics simulations under different force field parameters. The reliability of the obtained ensembles will be checked by the comparison of the corresponding calculated properties with their experimental values.

Acknowledgement: This study was supported by the Ministry of Education, Youth, and Sports (program INTER-EXCELLENCE, subprogram INTER-COST, project LTC17078). The work was supported from European Regional Development Fund-Project "CIISB4HEALTH" (No.CZ.02.1.01/0.0/0.0/16_013/0001776)

Probing role of electrostatic interactions in formations of transient structures of intrinsically disordered C-terminal domain of delta subunit of RNA polymerase from Gram-positive bacteria

[a] LUKÁŠ ŽÍDEK, [a] VOJTĚCH KUBÁŇ, [a] HANA ŠTÉGNEROVÁ, [a] PAVEL SRB, [a] VOJTĚCH ZAPLETAL, [b] HANA ŠANDEROVÁ, [b] LIBOR KRÁSNÝ, [c] MARTIN BLACKLEDGE

[a] *CEITEC-MU, Masaryk University, Kamenice 5, 62500 Brno, Czech Republic (lzidek@chemi.muni.cz)*

[b] *Institut of Microbiology of the Czech Academy of Sciences, Vídeňská 1083, 142 00 Prague, Czech Republic*

[c] *Institut de Biologie Structurale, CEA, CNRS, University Grenoble Alpes, 380044 Grenoble, France*

A unique component of RNA polymerase in Gram-positive bacteria, the delta subunit, contains an intrinsically disordered C-terminal domain. This domain consists mostly of aspertic acid and glutamic acid residues, with an important exception of a short fragment rich in lysines. Electrostatics plays an important role in physico-chemical properties and consequently of physiological function of the intrinsically disordered domain. In order to probe the role of electrostatic interactions directly, we compared biophysical properties and biological functions of the wild-type delta subunit with its mutant having lysines replaced by glutamic acid residues. The physical behavior of the proteins was monitore by small-angle X-ray scattering and nuclear magnetic resonance (analysis of chemical shifts, residueal dipolar couplings, paramagnetic relaxation enhancement). The observed parameters described redistribution of transient intradomain and interdomain electrostatic interactions as a result of the mutation. In vitro experiments revealed a direct impact of the mutation on transcription at promoters controlled by initiating nucleotide triphosphates. In vivo, the mutation substantially decreased viability of the mutated Bacillus subtilis strain in competition with the wild type.

Acknowledgement: This work was supported by Czech Science Foundation, grant number GA 13-16842S and by the Ministry of Education, Youth and Sports of the Czech Republic (MEYS CR) under the National Sustainability Programme II, project CEITEC 2020 (LQ1601).

Phylogenetics of Tandem Repeats with Circular HMMs: A Case Study on Armadillo Repeat Proteins

[a,b] SPENCER BLIVEN, [a,b] MARIA ANISIMOVA

[a] *Institute for Applied Simulation, Zurich University of Applied Sciences*
[b] *Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland*

Tandem repeat proteins are characterized by multiple sequential copies of repeats with significant structural or sequence similarity. Tandem repeats evolve via repeat expansion, duplication and loss, and many protein families exhibit very diverse repeat counts. Identifying the complex relationships between homologous proteins and between individual repeats is a challenging task. Using tools developed in our group, we present a detailed phylogenetic analysis of the repeats in the Armadillo Repeat Protein (ArmRP) family.

The ArmRP family is very diverse, appearing throughout the eukaryotes and having a wide range of functions. They are well characterized structurally, with ~42 amino acid repeats forming three alpha-helices which assemble into a solenoid structure. ArmRP are exciting candidates for protein design, as they have been shown to bind peptides in a modular manner [1].

Phylogenetic analysis of tandem repeats has several unique challenges. Identifying homologous regions is complicated by the repetitive nature of the sequence, which can cause register shifts when applying standard alignment tools. We surmount this problem using the Tandem Repeat Annotation Library (TRAL), a tool for accurately identifying repeats using circular profile hidden Markov models [2]. After constructing a multiple alignment of the repeats of ArmRP representatives, we infer a phylogenetic tree relating the different ArmRP and use it to analyse the conservation and diversification patterns through evolution, based on the information about tandem repeat number, order and their distribution on phylogenies.

Acknowledgement: This research was supported by COST Action BM1405 and COST Switzerland SEFRI project IZCNZ0-174836.

References

1. Reichen, C. et al. Computationally Designed Armadillo Repeat Proteins for Modular Peptide Recognition. J Mol Biol 428, 4467–4489 (2016).
2. Schaper, E. et al. TRAL: tandem repeat annotation library. Bioinformatics 31, 3051–3053 (2015).

The relationship between the store-operated calcium entry and the optic nerve damage; an electron microscopic study

[a] TUBA DEMİRCİ, [b] METİN UÇAR, [c] NURAY BİLGE, [a] OZLEM OZGUL ABUC, [a] HİLAL ATİLAY

[a] *Ataturk University Faculty of Medicine, Deparment of Histology and Embryology, Erzurum, TURKEY*
[b] *Regional Training and Research Hospital, Department of Ophtalmology, Erzurum, TURKEY*
[c] *Ataturk University Faculty of Medicine, Deparment of Neurology, Erzurum, TURKEY*

Introduction & Objectives: Acute retinal ischemia is a clinical condition that occurred in the various ocular pathologies such as central retinal artery occlusion, acute glaucoma, anterior ischemic optic neuropathy, diabetic retinopathy and other retinal vascular occlusions. It is common and causes vision loss [1, 2]. Most of the patients are over 60 years old and they typically complain of sudden, unilateral visual loss [3]. In this study, we aimed to investigate ultrastructurally the possible protective effect of 2-aminoethoxydiphenyl borate (2-APB), an store operated calsium entry (SOCE) inhibitor, in the optic nerve damage resulting from an experimentally generated acute retinal ischemia-reperfusion (ARIR) injury model.

Materials & Methods: Thirty Wistar-Albino male rats were randomly divided into 3 groups as sham group (n=10), ARIR group (n=10) and ARIR10 group (n=10). In the ischemia-reperfusion groups, under the anesthesia, right eyes were cannulated with a 30-gauge needle and the saline solution was given into the anterior chamber. The bottle was hung at a height of 200 cm from the subject and the intraocular pressure was elevated to 120 mmHg. The retina was examined by using ophthalmoscope and the formation of ischemia was confirmed via the presence of the opacity in the retina. An hour later, reperfusion was realized by withdrawing the neddle. In the rats of sham group, the intraocular pressure was not raised; just needle was applied and removed. ARIR group was not given 2-APB and also ARIR10 group was applied 4 mg/kg 2-APB as i.p 10 minutes before reperfusion. On the third day of the experiment, the rats were sacrificied under anesthesia. Optic nerves were removed and electron microscopy procedure was performed for ultrastructural examination. One µm thickness semithin sections were obtained from the tissues embeded into the araldite blocks by using ultramicrotome and they were examined under the light microscope by staining with toluidine blue. Subsequently, 70–80 nm thickness thin sections were taken on the grids and contrasted using uranyl acetate and lead citrate. And then, they were examined under the Jeol 100 SX electron microscope.

Results: When the electron microscopic sections of the Sham group were examined, it was seen that the optic nerve axons were arranged close to each other and had a compact myelin structure. In the ARIR group, optic nerve axons were swollen and infrequently aligned; and it was obseved that neural connective tissue was increased between axons.

In some areas, vacuoles were seen in different shapes and sizes. The axonal appearance were deteriorated and myelins were broken up into lamellas. When compared with sham group, it was observed that the number of normal myelinated axons count was increased in the ARIR10 group and there was a structure close to the sham group.

Conclusions: According to our findings, it has been concluded that acute retinal ischemia-reperfusion model applied to rats can cause the optic nerve damage and 2-Aminoethoxydiphenyl borate may have protective effect or even therapeutic effect against this damage.

Acknowledgement: This study was support by the 2016/12-numbered Scientific Research Fund of Ataturk University.

References

1. *D. Grozdani Sinisa, S. Sakaguchi Donald, H. Kwon Young, H. Kardon Randy, and M. Sonea Ioana.* Functional Characterization of Retina and OpticNerve after Acute Ocular Ischemia in Rats. OVS, June 2003, Vol. 44, No.6
2. Barron KD, Dentinger MP, Krohel G, Easton SK, Mankes R. Qualitative and quantitative ultrastructural observations on retinal ganglion cell layer of rat after intraorbital optic nerve crush. J Neurocytol. 1986;15:345–362.
3. Brown GC. Retinal arterial obstructive disease. In: Schachat AP, Murphy RB, Patz A, editors. Retina. .2. St.Louis: C.V.Mosby Company; 1989. p. 403-19.

The repeat structure of two paralogous invasin genes from *Yersinia ruckeri* sheds light on the evolution of adhesive capacities of a fish pathogen

[a] AGNIESZKA WROBEL, [a] CLAUDIO OTTONI, [a] JACK C. LEO, [a] DIRK LINKE

[a] *Department of Biosciences, University of Oslo, Norway*

Inverse autotransporters comprise the recently identified type Ve secretion system and are exemplified by intimin from enterohaemorrhagic *Escherichia coli* and invasin from enteropathogenic *Yersiniae*. These proteins share a common domain architecture and promote bacterial adhesion to host cells. Here, we identified and characterized two putative inverse autotransporter genes in the fish pathogen Yersinia ruckeri ATCC29473, namely yrInv (for *Y. ruckeri* invasin) and yrIlm (for *Y. ruckeri* invasin-like molecule). When trying to clone the genes for structural and functional studies, we experienced problems in obtaining PCR products. Based on prior experience with poorly annotated bacterial genomes, and with sequencing errors of highly repetitive DNA sequences, we re-sequenced the genome of *Y. ruckeri* ATCC29473 using PacBio sequencing. According to our new sequencing data, YrIlm is composed of 2603 amino acids and has a molecular mass of 256.4 kDa. This is in contrast to the previously deposited data, where YrIlm is only half the size (1303 amino acids), missing 13 of the 20 almost identical immunoglobulin-like domains identified in the re-sequenced genome. Based on the new genome information, we performed PCR analysis on five non-sequenced *Y. ruckeri* strains and found that the yrInv gene is present in all strains tested except *Y.ruckeri* 1435-95, whereas the yrIlm gene is present in all strains tested except *Y.ruckeri* 1006-94. The internal repeats of the yrInv gene product are highly diverged, but represent the same bacterial Immunoglobulin-like domains as in yrIlm. Using qRT-PCR, we found that yrIlm and yrInv are differentially expressed under conditions relevant for pathogenesis. In addition, we compared the genomic context of both genes in the newly sequenced reference strain to all available *Y. ruckeri* genomes, and found indications of recent events of horizontal gene transfer. Taken together, this studies demonstrates and highlights the power of SMRT technology for sequencing highly repetitive proteins, shows the limitations of deposited genomics data produced from short-read technologies, and sheds light on the genetic events that gave rise to these highly repetitive genes in a commercially important fish pathogen.

Acknowledgement: This work by supported by Research Council of Norway.

References

1. Wrobel *et al*., Journal of Structural Bilogy (special issue on non-globular proteins), *submitted*

Analysis of the Relationships in the Repeat Regions of RAP Domain-Containing Proteins

[a] MATTHEW MERSKI, [a] JAKUB SKRZECZKOWSKI, [a] KRZYSZTOF MŁYNARCZYK, [a] STANISŁAW DUNIN-HORKAWICZ, [a] MARIA GÓRNA

[a] *Biological and Chemical Research Center, University of Warsaw, Warsaw, Poland*

The RAP domain (Rna-binding abundant in APicomplexans) [1] is a small (~65 amino acids), protein domain that is widely dispersed throughout the eukarya with over 1100 unique RAP domain sequences currently recognized in the InterPro database [2]. Despite its name, the function of the RAP domain is still unproven, although the domain is associated with organelle-directed proteins, including the mitochondrial-associated FASTKD proteins in vertebrates [1] which may potentially be the mitochondrial nucleases, the chloroplastassociated RAP protein in vascular plants [3] and the numerous RAP proteins in apicomplexans such as the malaria organism, *P. falciparum*, as well as *Babesia* species which cause disease in livestock and native ruminants. The RAP domain is typically located at the C terminus of a protein with the organelle directing domain at the N terminus. The middle region of the protein is predicted to be largely alpha helical, containing a set of tandem tetratricopeptide-like repeats (TPR) usually with low internal sequence similarity between the tandem repeats [4]. We have attempted to organize and categorize the RAP domain containing proteins by comparing the sequences of the fairly small RAP domains as well as analyzing the conservation patterns within the helical repeats. The overall phylogeny appears highly similar using either method and the tandem repeat helical domains do not appear to be evolving faster than or independently of the RAP domains. Our analysis of the TPR-like repeats within these proteins revealed sets of conserved patterns that correlate with apparently rational sub-groupings (by function or evolutionary heritage or both) among the proteins. By analyzing the patterns of these relationships between the repeats within a given protein, sets of conserved sequence structures could be readily observed. These sequence structures are also readily detectable in other types of repeat proteins and there is some correlation with expected functional and structural aspects of these proteins. The implications and functional application of the detection of these patterns within the RAP domain family of proteins as well as repeat proteins in general are examined.

Acknowledgement: This work was supported by a SONATA 8 (#2014/15/D/NZ1/00968) and Premia na Horyzoncie (#353103/PnH/2017) grants from the National Science Centre of Poland and EMBO Installation Grant from European Molecular Biology Organization and a Marie Skłodowska-Curie Actions grant to Maria Gorna.

References

1. *Trends in Biochemical Sciences* 29, 567-570 (2004)
2. *Nucleic Acids Research* 45, D190-D1997 (2017)
3. *Plant Cell.* 26, 777-87 (2014)
4. *BioEssays* 21, 932–939 (1999)

An efficient cell-free strategy to produce stable huntingtin for structural studies

[a] A. MORATÓ, [a] A. URBANEK, [a] F. ALLEMAND, [a] A. FOURNET, [a] E. DELAFORGE, [b] S. DELBECQ, [a] N. SIBILLE, [a] P. BERNADÓ

[a] *Centre de Biochimie Structurale, CNRS UMR 5048 – INSERM 1054 – UM, Montpellier, France*
[b] *Faculté de Pharmacie, Université de Montpellier, Montpellier, France*

Huntington's disease (HD) is a fatal neurodegenerative disorder caused by an abnormally extended poly-Q tract in the N-terminal region of the huntingtin protein (httex1). While the poly-Q repeat length ranges between 6 and 35 in healthy subjects, patients with HD exhibit lengths of 36 or greater, defined as the pathological threshold.

On the molecular basis it is postulated that the homotypic and uncharged nature of httex1 renders the protein unstable and prone to aggregation, causing the disease. However, contradictory observations have been reported regarding the presence of secondary structural elements and the compactness of different httex1 constructs [1].

We aim to clarify the structural bases of the pathological threshold by conducting high resolution structure and dynamics studies on several httex1 constructs, hosting poly-Q tracts of 16, 35 and 46 glutamines. In order to prepare well characterized protein samples for structural studies we set up a pipeline to produce optimized constructs using a cell-free platform and NMR stability studies. Cell-free coupled transcription-translation systems based on *Escherichia coli* extracts are an attractive alternative to conventional expression systems *in vivo* [2]. In addition to being independent of bacterial growth, the cell-free reaction is an open system that can be easily modified. Cell-free extracts are very stable and tolerant against a wide variety of substances that can be added at any time over the course of the reaction. For example, detergents that can improve protein stability or avoid problems of aggregation can be present right from the beginning [3].

Overall, the cell-free system allowed us to rapidly assess various httex1 fusion protein constructs under defined conditions in a microplate format at the same time. Furthermore, cell-free reactions can be optimized to synthesize substantial amounts of protein and can also be used at preparative scale. These efforts led to the selection and purification of an optimized httex1 (16Q) construct for structural and dynamics studies by NMR.

Acknowledgement: This project has received funding from the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement No 648030) and from the Fondation pour la Recherche Médicale (FRM, SPF20150934061).

References

1. Adegbuyiro A., Sedighi F, Pilkington A.W., Groover S., Legleiter J. *Biochemistry*, 2017, 56, 1199-1217.
2. Apponyi MA, Ozawa K, Dixon NE, Otting G. *Methods Mol Biol*, 2008, 426, 257-268.
3. Rosenblum G, Cooperman BS. *FEBS Letters*, 2014, 588, 261-268.

Calorimetric study of interaction between DARPin and maltose binding protein

[a] MICHAL NEMERGUT, [b] ERIK SEDLÁK

[a] *Department of Biophysics, P. J. Šafárik University, Jesenná 5, 041 54 Košice, Slovakia, michal.nemergut@gmail.com.*

[b] *Center for Interdisciplinary Biosciences, P. J. Šafárik University, Jesenná 5, 041 54 Košice, Slovakia, erik.sedlak@upjs.sk*

DARPins (designed ankyrin repeat proteins) are known as a novel class of binding proteins with non-immunoglobulin scaffolds. DARPins can be rapidly engineered to detect diverse target proteins with high specificity and affinity. From this point of view, they offer an attractive alternative to antibodies. In this work, we study properties of DARPin off7, which was selected by ribosomal display for binding to maltose binding protein (MBP) with high specificity and affinity. The effect of different pH of solvent on stability of the complex MBP-off7 has been studied by isothermal titration calorimetry and differential scanning calorimetry. Our results indicate that affinity of the off7 to MBP decreases with increasing pH of solvent. This behavior probably results from changes in conformational properties of both proteins and particularly of oligomeric state of off7. Our study may lead to better understanding of the mechanism of interaction of off7 with MBP. This is critical point for development of affinity chromatography, which enables isolation/purification of proteins containing MBP tag.

Acknowledgement: We would like to thank to Prof. Dr. Andreas Plückthun for providing us with plasmids encoding both proteins. This work was supported by grants of Slovak grant agency VEGA 1/0423/16, APVV-15-0069 and VVGS-PF-2016-72626.

References

1. S. WETZEL et al., Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins. In J. Mol. Biol., 2008, 376, 241-257.
2. H. K. BINZ et al., High-affinity binders selected from designed ankyrin repeat protein libraries. In Nat. Biotechnol., 2004, 22, 575-582.
3. M. A. SEEGER et al., Design, construction, and characterization of a second-generation DARPin library with reduced hydrophobicity. In Protein Science, 2013, 22, 1239-1257.

Improving a machine-learning approach for determination of genuine tandem repetitive sequences with enclosure of Fourier-based features

[a] VLADIMIR PEROVIC, [a] NEVEN SUMONJA, [b] ANDREY V. KAJAVA, [a] NEVENA VELJKOVIC

[a] *Centre for multidisciplinary research, Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia*

[b] *Centre de Recherche en Biologie cellulaire de Montpellier (CRBM), UMR 5237 CNRS, Universite´ Montpellier 1919 Route de Mende, Cedex 5, Montpellier 34293, France*

Tandem Repeats (TRs) are abundant in proteins, having a variety of fundamental functions. In many cases, evolution has blurred their repetitive patterns. This leads to the problem of distinguishing between sequences that contain highly imperfect TRs, and the sequences without TRs. Recently developed scoring tool denoted Tally [1] is based on a robust machine learning approach and trained and evaluated on an expertly curated benchmarking datasets. Tally outperforms state of the art scoring procedures in separation between sequences with structural TRs and sequences of aperiodic structures. Here, we propose a modification of Tally with enclosure of the Fourier-based features for transforming protein sequences into feature vectors [2]. Fourier transform captures the repetitions of amino acids by considering sequences in the frequency domain using the Fast Fourier Transform (FFT) algorithm. FFT has been widely used to transform proteins of variable length into fixed length vectors after sequence is encoded by the physicochemical properties of amino acids contained in AAindex database [3]. Here, the best predictive model is selected based on evaluation of machine learning models as well as feature selection and parameter optimization. This new model TallyF which combines features stemming from multiple sequence alignments with Fourier-based features outperforms Tally at a level of accuracy 87.0 % and an Area Under the Receiver Operating Characteristic Curve of 0.890 on test set. These findings indicate that Fourier analysis can complement MSA in recognition of structurally relevant and functionally meaningful TRs.

Acknowledgments: This work was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia [grant no. 173001]. The COST Actions BM1405 is kindly acknowledged for supporting the short term scientific mission of NS to AK laboratories.

References

1. Richard FD, Alves R, Kajava AV. Tally: a scoring tool for boundary determination between repetitive and non-repetitive protein sequences. Bioinformatics. 2016 Jul 1;32(13):1952-8.
2. Shepherd AJ, Gorse D, Thornton JM. A novel approach to the recognition of protein architecture from sequence using Fourier analysis and neural networks. Proteins. 2003 Feb 1;50(2):290-302.
3. Nakai K, Kidera A, Kanehisa M. Cluster analysis of amino acid indices for prediction of protein structure and function. Protein Eng 1988;2:93–100.

Destroying effect of extracts from traditional Chinese herbs on Aβ42 amyloid fibrils

[a] Z. BEDNARIKOVA, [a] M. GANCAR, [a] D. FEDUNOVA, [b] RUI WANG, [b] LEI MA, [b] YUN TANG, [b] LIXIA SUN, [b] BU-BING ZENG, [a] Z. GAZOVA

[a] *Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, Kosice, Slovakia*
[b] *Shanghai Key Laboratory of New Drug Design, School of Pharmacy, East China University of Science and Technology, Shanghai, China*

Alzheimer's disease is the most common cause of dementia and current treatment strategies are only symptomatic and temporarily improve quality of life. It was suggested that formation of Aβ peptide amyloid aggregates is one of the central pathological events in AD. Therefore, one of the recent therapeutical strategies is focused on the inhibition of the Aβ peptide amyloid self-assembly or clearance of amyloid plaques. Small compounds have high potential to interfere with amyloid fibrillization or reduce amount of the amyloid aggregates. In our research we focused on the extracts from herbs used in traditional Chinese medicine. Our objective was to find out, if any of 12 studied compounds from different Chinese herbs have substantial impact on Aβ42 fibrils. The effect of compounds on Aβ42 fibrils was studied using Thioflavin T assay in broad concentration range and DC_{50} values were determined. Noticeable destroying effect was observed for four compounds, but the most effective were extracts from the root of *Salvia miltiorrhiza* with DC_{50} values of 10 μM and 0.7 μM respectively. Our results were supported by atomic force microscopy. *In silico* study confirmed the results obtained in *in vitro* experiments as the most effective compounds have considerable affinity to Aβ42 fibrils. Extracts from *Salvia miltiorrhiza* showed reasonable potential and could become future therapeutical agents.

Acknowledgements: This work was supported by VEGA 2/0145/17, 2/0176/14; APVV SK-CN-2015-0023; MVTS COST 083/14-NGP-NET Action BM1405.

Challenges in the computational prediction of peptide amyloidogenicity

[a] MICHAŁ BURDUKIEWICZ, [b] STEFAN RÖDIGER, [c] MARLENA GĄSIOR GŁOGOWSKA, [a] PAWEŁ MACKIEWICZ, [c] MAŁGORZATA KOTULSKA

[a] *University of Wrocław*
[b] *Brandenburg University of Technology*
[c] *Wrocław University of Science and Technology*

Amyloids are self-aggregating proteins that participate in neurodegenerative disorders, such as Alzheimer's or Parkinson's diseases. The computational prediction of amyloidogenic proteins is a great challenge because regions responsible for amyloidogenicity cannot be described by specific amino acid motifs, but rather residues with specific physiochemical properties. Henceforth, we created AmyloGram, a software suitable for the detection of amyloids. The reduced amino acid alphabet, based on the physicochemical properties of amino acids, allows AmyloGram to handle the diversity of amyloid proteins.

We found eight peptides that were annotated in the the AmyLoad database as non-amyloidogenic but assesed by AmyloGram with the high probability of amyloidogenicity but. We analyzed them using the Fourier transform infrared spectroscopy and found out that seven of these eight peptides are in fact amyloidogenic. For three out of seven amyloidogenic peptides, our experimental results were also confirmed independently by other studies. The computational prediction using other amyloid-predicting software, PASTA 2.0, revealed amyloidogenic properties in only two of the seven amyloids that were confirmed experimentally.

Our data indicate that AmyloGram is able to detect amyloid proteins that were falsely assigned as non-amyloidogenic. Furthermore, it is able to find more false non-amyloids than other software maintaining the same specificity (in both cases thresholds were adjusted to assure the 0.95 specificity). Nevertheless, the potential peptides should be tested in a wider experimental setting because some proteins may display amyloid properties only under specific conditions.

AmyloGram is available as a web-server (www.smorfland.uni.wroc.pl/amylogram/) and R package (https://cran.r-project.org/package=AmyloGram).

Solid-state NMR structural studies of peptide induced prion protein aggregates

[a] JĒKABS FRIDMANIS, [a] KRISTAPS JAUDZEMS, [b] ZIGMANTAS TOLEIKIS,
[b] TOMAS ŠNEIDERIS, [c] BRETT KUAN-YU CHU, [c] RITA PY CHEN,
[b] VYTAUTAS SMIRNOVAS, [a] EDVARDS LIEPIŅŠ

[a] *Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga LV-1006, Latvia*
[b] *Vilnius University, Institute of Biotechnology, Sauletekio 7, Vilnius, LT-10257, Lithuania*
[c] *Institute of Biological Chemistry, Academia Sinica, 128 Academia Road Sec. 2, Nankang, Taipei 115, Taiwan*

Prion-like spreading may be employed in a number of fatal neurodegenerative disorders, including Alzheimer's and Parkinson's diseases [1]. Understanding the possible mechanisms of such spreading would be a big step towards curing these diseases. Recent work showed that prion protein aggregation can be induced by short peptides. It seems that either structure of peptide-induced prion protein aggregates or the mechanism of their formation is different from the current knowledge in the field [2]. To get insight into the structure of peptide-induced prion protein fibrils we have used solid state nuclear magnetic resonance spectroscopy (ssNMR) and studied mouse prion protein fibrils obtained using different aggregation methods including seeding with short peptides. In the poster, we will report initial results from these studies.

Acknowledgement: the authors thank Prof. Witold Surewicz for sharing the MoPrP(89-230) expression plasmid.

References

1. Frost B, Diamond MI. Prion-like Mechanisms in Neurodegenerative Diseases. *Nature reviews Neuroscience*. 2010;11(3):155-159. doi:10.1038/nrn2786.
2. Chatterjee, B. et al. Amyloid Core Formed of Full-Length Recombinant Mouse Prion Protein Involves Sequence 127-143 but Not Sequence 107-126. PLoS One 8, (2013).

Oligomers are building blocks for fibril formation of fragments of the Alzheimer's disease Aβ peptide

[a] OLGA M. SELIVANOVA, [a] ANNA V. GLYAKINA, [a] ELIZAVETA I. GRIGORASHVILI, [a] MARIYA YU. SUVORINA, [a] NIKITA V. DOVIDCHENKO, [a] ALEXEY K. SURIN, [a] OXANA V. GALZITSKAYA

[a] *Institute of Protein Research, Russian Academy of Sciences, Pushchino, Russia*

A large body of information accumulated to date allows using experimental data, as well as theoretical predictions, studying the amyloid-forming capacity of most proteins, and describing the mechanism of the formation of amyloid fibrils. It seems possible that despite a number of contradictions among some theoretical and experimental data, the process of amyloidogenesis must possess general key features common to all proteins and peptides. To negate the effect of varying contexts in proteins, it is convenient to make use of model peptides, whose sequence mostly comprises well-characterized amyloidogenic regions.

To identify the key stages in the amyloid fibril formation we studied the aggregation of amyloidogenic fragments of Aβ peptide: Aβ(16-25), Aβ(31-40), and Aβ(33-42), using the methods of electron microscopy, ThT fluorescence, X-ray analysis, mass spectrometry, and structural modeling. We have found that the fragments Aβ(31-40) and Aβ(33-42) form amyloid fibrils in the shape of bundles and ribbons, while Aβ(16-25) fragment forms only nanofilms. Combined analysis of the data allows us to speculate that both the fibrils and the films are formed via the association of ring-shaped oligomers, whose external diameter measured about 6-7 nm, the internal diameter 2-3 nm, with the height of about 3 nm. We conclude that such oligomers are the main building block in fibrils of any morphology. The interaction of the ring oligomers with each other in different ways makes it possible to explain their polymorphism.

The data obtained in the recent years, certain accumulated contradictions, as well as some unsolved questions related to the structural organization of the fibrils, combined with our own data on the organization of amyloidogenic peptides and their fragments [1-3], allows us to undertake a revision of the traditional views upon the processes of amyloidogenesis. Correct interpretation of such molecular processes is paramount for understanding, prophylactics, and future treatment of the amyloid-related neurodegenerative conditions in humans. In that connection, we believe that a key effort must currently be concentrated upon studying mechanisms of the transition from the monomeric state to the formation of oligomers.

Acknowledgement: This study was supported by the Russian Science Foundation (14-14-00536).

References

1. Selivanova OM, Surin AK, Marchenkov VV, Dzhus UF, Grigorashvili EI, Suvorina MY, Glyakina AV, Dovidchenko NV, Galzitskaya OV (2016) The Mechanism Underlying Amyloid Polymorphism is Opened for Alzheimer's Disease Amyloid-β Peptide. J. Alzheimers Dis. JAD 54, 821–830.
2. Selivanova OM, Glyakina AV, Gorbunova EY, Mustaeva LG, Suvorina MY, Grigorashvili EI, Nikulin AD, Dovidchenko NV, Rekstina VV, Kalebina TS, Surin AK, Galzitskaya OV (2016) Structural model of amyloid fibrils for amyloidogenic peptide from Bgl2p–glucantransferase of S. cerevisiae cell wall and its modifying analog. New morphology of amyloid fibrils. Biochim. Biophys. Acta BBA - Proteins Proteomics 1864, 1489–1499.
3. Selivanova OV, Suvorina MY, Surin AK, Dovidchenko NV, Galzitskaya O (2017) Insulin and Lispro Insulin: What is Common and Different in Their Behavior? Curr. Protein Pept. Sci. 18(1):57-64.

Objective Structural Analysis of Multicomponent Amyloid Systems by Chemometric SAXS Data Decomposition

[a,b] FÁTIMA HERRANZ-TRILLO, [b] MINNA GROENNING, [b] ANDREAS VAN MAARSCHALKERWEERD, [c] ROMÀ TAULER, [b] BENTE VESTERGAARD, [a] PAU BERNADÓ

[a] *Centre de Biochimie Structurale. INSERM U1054, CNRS UMR 5048, Université de Montpellier. 29, rue de Navacelles, 34090-Montpellier, France.*

[b] *Department of Pharmacy and Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, 2100 Copenhagen, Denmark.*

[c] *Environmental Chemometrics Group, Department of Environmental Chemistry, Institute of Environmental Assessment and Water Diagnostic (IDAEA-CSIC), Barcelona, Spain.*

In amyloid pathologies (e.g. Parkinson's or Alzheimer's diseases) there are indications that oligomeric aggregated precursors of fibrillation, and not mature fibrils, are the main cause of cytotoxicity and neuronal damage. Hence the importance of characterizing early stages in the fibrillation process. The structural analysis of these oligomeric species is a major challenge due to their instability, low relative concentration, the difficulties for isolation, and their interdependence with other species of very different sizes [1]. Mechanistic studies normally monitor individual species of the fibrillation process, such as mature fibrils, whereas the other species remain invisible.

In SAXS studies of fibrillation, the resulting individual scattering pattern measured at different time-points throughout the fibrillation process, is a sum of the contributions from each component of the mixture. This additive nature of SAXS data allows for probing the evolution of these mixtures of oligomeric states [4, 5]. We present an objective SAXS data decomposition method by adapting the Multivariate Curve Resolution Alternating Least Squares (MCR-ALS [2, 3]) chemometric method. The approach enables rigorous and robust decomposition of synchrotron SAXS data by simultaneously introducing these data in different representations that emphasize molecular changes at different time and structural resolution ranges. The approach has allowed the study of the amyloidogenic processes of insulin and the familial mutant E46K of the Parkinson's disease related protein alpha-synuclein, and has allowed the structural characterization of the species present (including the oligomeric species) and the kinetic characterization of their transformations. Our approach is generally applicable to any macromolecular mixture with tunable equilibria that can be probed by SAXS.

Acknowledgements: This work has been sponsored by ANR-SPIN-HD (France), Institut national de la santé et de la recherche médicale (INSERM) – France, University of Copenhagen and Neuromed (SUDOE) and the Danish Research Council Sapere Aude programme SAFIR.

References

1. Vestergaard, B., Groenning, M., Roessle, M., Kastrup, J. S., van de Weert, M., Flink, J. M., Frokjaer, S., Gajhede, M. & Svergun, D. I. PLoS Biol. 5, 1089–1097 (2007).
2. Tauler, R. Chemom. Intell. Lab. Syst. 30, 133–146 (1995).
3. Blobel, J., Bernadó, P., Svergun, D. I., Tauler, R. & Pons, M. J. Am. Chem. Soc. 131, 4378–86 (2009).
4. Langkilde, A. E. & Vestergaard, B. in Amyloid Proteins: Methods and Protocols (eds. Sigurdsson, E. M., Calero, M. & Gasset, M.) 849, 137–155 (Humana Press, 2012).
5. Giehm, L., Svergun, D. I., Otzen, D. E. & Vestergaard, B. Proc. Natl. Acad. Sci. U. S. A. 108, 3246–51 (2011).

Modeling motifs and transmembrane properties of amyloid proteins

[a*] MALGORZATA KOTULSKA, [a] PAWEL P. WOZNIAK, [b] MICHAL BURDUKIEWICZ, [c] JULITA KULBACKA

[a]*Department of Biomedical Engineering, Faculty of Fundamental Problems of Technology, Wroclaw University of Technology, Wroclaw, POLAND,*

E-mail: malgorzata.kotulska@pwr.edu.pl

[b] *Department of Genomics, Faculty of Biotechnology, University of Wroclaw, Wroclaw, POLAND*

[c] *Department of Medical Biochemistry, Medical University of Wroclaw, Chałubińskiego 10, 50-368 Wroclaw, POLAND*

Amyloids are proteins capable of forming fibrils whose intramolecular contact sites form a characteristic zipper pattern. A number of diseases related to amyloid proteins is constantly increasing. They include Alzheimer's disease Parkinson's disease, amyotrophic lateral sclerosis, and many others. There are several hypotheses suggesting potential mechanisms for development of fully symptomatic diseases. Recognition of factors responsible for protein misfolding and subsequent cascade of events can contribute to better understanding of the diseases mechanisms and potential drug design.

Aggregation of amyloid proteins depends on the presence of short amyloidogenic motifs, which typically include 4-10 residues. The fragments can be harmless when they are burried inside a protein. We developed computational methods that recognize amyloidogenic hot-spots. They utilize a variety of machine learning methods [1,2] including those specifically tailored to amyloid proteins, such as site specific co-occurrence patterns [3] and n-gram analysis of physicochemical alphabets of amino acids [4]. Our methods show sequential patterns in the amyloid fragments, strongly correlated with hydrophobicity, a tendency to form beta structures, and rigidity of amino acid residues. Moreover, an extensive database of experimentally tested peptides, supporting classification of amyloids, will be presented [5].

Recent studies indicate that the neurodegenerative processes may correspond to incorporation of amyloid oligomers into the cell and organelle membranes, creating weakly cation-selective ion channels that allow uncontrolled influx of calcium into nerve cells. The excessive influx of calcium into the cytoplasm leads to disruption of intracellular pathways, membrane depolarization, ATP depletion, and mitochondrial membrane depolarization with impairment of mitochondrial function. Our results on modeling hypothetical structures of the amyloid pores will be presented. Additionally, we introduced a new experimental protocol for studying the effect of calcium on neural cells. The effects of increased calcium influx due to non-selective nanopores was experimentally tested on pheochromocytoma cells (PC-12) utilizing electroporation with micro- and nanosecond pulses.

References

1. Stanislawski J, Kotulska M, Unold O: Machine learning methods can replace 3D profile method in classification of amyloidogenic hexapeptides. BMC Bioinformatics 2013, 14 (art. 21):1-19.
2. Kotulska M, Unold O: On the amyloid datasets used for training PAFIG - how (not) to extend the experimental dataset of hexapeptides. BMC Bionformatics 2013, 14 (art. 351):1-8.
3. Gąsior P, Kotulska M: FISH Amyloid - a new method for finding amyloidogenic segments in proteins based on site specific co-occurence of aminoacids. BMC Bioinformatics 2014, 15 (art. 54):1-8.
4. Burdukiewicz MJ, Sobczyk P, Rödiger S, Duda-Madej A, Mackiewicz P, Kotulska M. (2016) Prediction of amyloidogenicity based on the n-gram analysis. PeerJ Preprints 4:e2390v1
5. Woźniak P.P, Kotulska M: AmyLoad: website dedicated to amyloidogenic protein fragments. Bioinformatics 2015, 31(20):3395-3397.
6. Żulpo M, Kotulska M: Comparative modeling of hypothetical amyloid pores based on cylindrin. Journal of Molecular Modeling 2015, 21(151), s. 1-12.

Impact of S100 protein multimers in zinc homeostasis in Alzheimer's disease mice models

[a] MARIANA A. ROMÃO, [b] SIMONE HAGMEYER, [b] ANDREAS M. GRABRUCKER, [a] CLÁUDIO M. GOMES

[a] *Biosystems and Integrative Sciences Institute, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal and Departmento de Química e Bioquímica, Universidade de Lisboa, 1749-016 Lisboa, Portugal*

[b] *WG Molecular Analysis of Synaptopathies, Neurology Department, Neurocenter of Ulm University, Ulm, Germany and Department of Biological Sciences, University of Limerick, Limerick, Ireland*

FCUL Protein Folding and misfolding Laboratory • http://folding.fc.ul.pt • maromao@fc.ul.pt

Age related neurodegenerative diseases, such as Alzheimer's disease (AD), are associated with the occurrence of protein aggregates and activation of inflammatory pathways. The occurrence of dysregulated protein expression is a common feature that are correlated with the misfolding of normally soluble proteins and their subsequent conversion into toxic amyloid aggregates [1] which is accompanied by upregulation of pro-inflammatory cytokines such as S100 proteins. S100s are small (12kDa) Ca2+-binding signaling proteins which occur mostly as homodimers. Ca2+ binding occurs at two EF-hand motifs and some homologues contain additional regulatory Zn2+/Cu2+ binding sites [2]. S100 proteins are involved in numerous intra and extracellular pathophysiological processes and some neuronal S100s are consistently altered in neurodegeneration, including AD. Among these are S100A8, S100A9 and the heterodimer S100A8/A9 (calprotectin, CP) which seem to undergo self-assembly upon zinc and calcium binding, a process which is likely mechanistically linked to the cross-beta forming propensity which we have recently elucidated [3].

In this work, we report an investigation dealing with our hypothesis that is zinc and calcium-binding to neuronal S100s promotes the formation of protein deposits in the AD brain and that this is related to changes in zinc homeostatic levels, which are known to be altered in ageing and neurodegeneration. For the purpose, we analysed brain sections from APP23 AD mice models and different ages (3 and 15months), using immunohistochemical analysis and fluorescence microscopy in combination with biochemical assays. We have analysed the occurrence of S100A8 and S100A9 protein assemblies in correlation to brain zinc levels and the presence of amyloid plaques. Here we present our preliminary results that pave the way to explore new roles for S100s as modifiers of Alzheimer's Disease.

Acknowledgments: Support for this work was provided by grants from FCT/MCTES (PTDC/NEU-NMC/2138/2014 to CMG) and by Bial Foundation (343/14 to CMG). MAR acknowledges fellowship support from FCT/MCTES PD/BD/114390/2016 as a BioSYS PhD program student. This work is partly supported by the COST Action TD1304 The Network for the Biology of Zinc (Zinc-Net) via Short-Term Scientific Mission (STSM) travel grants (to M.A.R. and S.H.)

References

1. Knowles, T., Vendrusculo, M., Dobson, C. The Amyloid state and its association with protein misfolding diseases. *Nature reviews*. 2014, 15
2. Leal, S.S et al. Metal ions as modulators of protein conformation and misfolding in neurodegeneration. *Coordination Chemistry Reviews*. 2012, 256 (19-20), 2253-227
3. Carvalho, S. B., Botelho, H. M., Leal, S. S., Cardoso, I., Fritz, G., Gomes, C. M., Intrinsically Disordered and Aggregation Prone Regions Underlie β-Aggregation in S100 Proteins. *PLOS ONE*. 2013, *8(10)*, e76629

Molecular evolution of macrocyclic rescuers of disease-associated protein misfolding

[a,b] ILIAS MATIS, [a,b] DAFNI CHRYSANTHI DELIVORIA, [c] BARBARA MAVROIDI, [a,d] NIKOLETTA PAPAEVGENIOU, [a,e] STEFANIA PANOUTSOU, [a] STAMATIA BELLOU, [a,f] ZACHAROULA I. LINARDAKI, [e] ALEXANDRA V. STAVROPOULOU, [e] SPIROS EFTHIMIOPOULOS, [c] MARIA PELECANOU, [a] NIKI CHONDROGIANNI, [a]* GEORGIOS SKRETAS

[a] *Institute of Biology, Medicinal Chemistry & Biotechnology, National Hellenic Research Foundation, Athens, 11635, Greece*

[b] *School of Chemical Engineering, National Technical University of Athens, Athens, 15780, Greece*

[c] *Institute of Biosciences & Applications, National Center for Scientific Research "Demokritos", Athens, 15341, Greece*

[d] *Faculty of Biology and Pharmacy, Institute of Nutrition, Friedrich Schiller University of Jena, Jena, 07743, Germany*

[e] *Faculty of Biology, National and Kapodistrian University of Athens, Athens, 15701, Greece*

[f] *Department of Biology, University of Patras, Patras, 26504, Greece*

Protein misfolding is a common pathological feature for many human diseases, such as Alzheimer's disease (AD), Parkinson's disease, type II diabetes and others. We will describe the development and application of an integrated and generalizable bacterial platform for facile discovery of macrocyclic rescuers of disease-associated protein misfolding. In this system, large combinatorial libraries of macrocycles are biosynthesized in *Escherichia coli* cells and simultaneously screened for their ability to rescue pathogenic protein misfolding using a genetic assay based on fluorescence-activated cell sorting. We will first describe the effectiveness of this approach through the identification of drug-like, head-to-tail cyclic peptides that modulate the aggregation of the amyloid β peptide (Aβ) of AD. By using a series of biochemical, biophysical and biological assays using isolated Aβ, primary mammalian neurons and various established AD models in the nematode Caenorhabditis elegans, we have found that the selected macrocycles potently inhibit the formation of neurotoxic Aβ aggregates. Further, to showcase the generality of our approach, we will describe the application of the same platform for the identification of misfolding rescuers of mutant Cu/Zn superoxide dismutase 1 (SOD1), a protein whose misfolding and aggregation is associated with inherited forms of amyotrophic lateral sclerosis. Overall, our approach represents a straightforward strategy for the discovery of molecules that rescue the misfolding of polypeptides known to be associated with disease effectively.

Tacrine derivatives *vs.* amyloid aggregation of Aβ peptide

[a,b] K.ULICNA, [a] J. KUBACKOVA, [a] Z. BEDNARIKOVA, [a] Z. GAZOVA

[a] *Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, Watsonova 47, 040 01 Kosice, Slovakia*

[b] *Institute of Biology and Ecology, Faculty of Science, Pavol Jozef Safarik University, Srobarova 2, 041 54 Kosice, Slovakia*

The conversion of normally soluble protein into fibrillar amyloid aggregates is of central importance for several amyloid-related diseases as the presence of amyloid deposits in different tissues has toxic consequence to various cell types leading to their dysfunction or death [1]. In last few years, a range of low molecular weight compounds have been selected to effectively inhibit amyloid aggregation and clear amyloid fibrils.

We have investigated the ability of distinct tacrine-benzothiazole heterodimers to prevent the formation of Aβ-peptide amyloid aggregates *in vitro*. Tacrine core was functionalized with methoxy- or chloro- functional groups and attached to 2-amino-benzothiazole with aliphatic linker with different length (from 2 to 8 carbons) (Fig.1). Anti-amyloid activities of heterodimer derivatives were examined using Thioflavin T fluorescence assay and atomic force microscopy. It was found that inhibitory efficiency characterized by the half-maximal inhibition concentration (IC_{50} value) was affected mainly by: i) the presence of function group attached to the tacrine core and ii) by the length of the linker. It should be noted that modified tacrines and 2-amino-benzothiazole alone do not exhibit any or very low inhibition of Aβ-peptide fibrillization. The obtained results clearly indicate that connection of tacrine and benzothiazole into the heterodimer significantly increases their inhibitory efficiency.

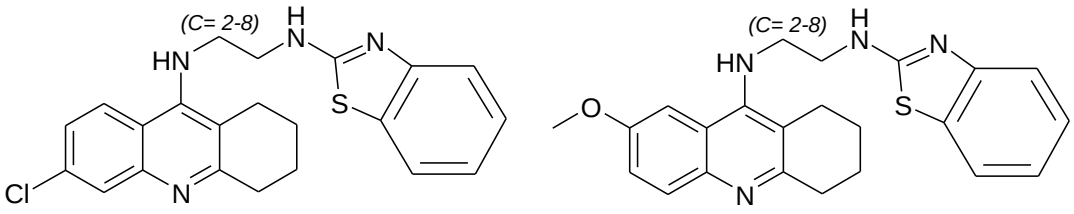

Figure 1. Structure of 6-Cl-tacrine-2-amino-benzothiazole (left) and 7-methoxy-tacrine-2-amino-benzothiazole (right) heterodimers with various length of the linker (number of carbons = 2 – 8).

Acknowledgement: This work was supported by the research grant VEGA 0145 and VEGA 0176, SAS-MOST JRP 2015/5 and MVTS COST 083/14 action BM1405.

References

1. Knowles, T. P. J.; Vendruscolo, M. and Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nature* 15 (2014) 384-396.

Register

Book of Abstracts, 3rd NGP-Net SYMPOSIUM ON NON-GLOBULAR PROTEINS,
August 28 – September 1, 2017, Košice, Slovakia

Editors: Zuzana Gazova, Zuzana Bednarikova, Rostislav Skrabana
Katarina Ulicna and Miroslav Gancar

ISBN: 978-80-89656-20-2

EAN: 9789089656202

ISBN: 978-80-89656-20-2
EAN: 9789089656202