12 September, Wednesday	
Location: Hotel Grand Spa Lietuva, Nemunas Hall	
Welcome address - **Vytautas Smirnovas & Silvio Tosatto**	8:50 - 9:00
Session 1, chair: Vytautas Smirnovas	
Invited talk - **Ronald Melki** - Prion-like propagation of alpha-synuclein protein assemblies in Parkinson's disease and related synucleinopathies	9:00 - 9:30
Ludmilla Morozova-Roche - Linking amyloid and neuroinflammatory cascades in neurodegenerative diseases – role of pro-inflammatory S100A9 protein	9:30 - 9:50
Salvador Ventura - A novel small molecule inhibits α-synuclein aggregation, disrupts amyloid fibrils and prevents degeneration of dopaminergic neurons	9:50 - 10:10
Georgios Skretas - Engineering bugs for the discovery of new drugs against diseases caused by protein misfolding and aggregation	10:10 - 10:30
Coffee Break	**10:30 - 11:00**
WG meetings - WG4 and WG2	11:00 - 12:30
Lunch, Hotel restaurant	**12:30 - 14:00**
Session 2, chair: Sandra Macedo-Ribeiro	
Invited talk - **Witold K. Surewicz** – Amyloid fibrils from the N-terminal prion protein fragment as a model for studying the mechanism of prion propagation and transmissibility barriers	14:00 - 14:30
Andrei Kajava - Structural aspects of naturally-occurring amyloids	14:30 - 14:50
Kristaps Jaudzems - Biomimicry of Artificial Spider Silk Spinning Assessed by NMR Spectroscopy	14:50 - 15:10
Flash poster presentations, I session	15:10 - 15:35
Poster Session and Coffee Break	**15:35 - 17:00**
Session 3, chair: Salvador Ventura	
Invited talk – **Meytal Landau** – Extreme Polymorphism in Amyloids from Pathogenic Microbes	17:00 - 17:30
Gunilla Westermark - Cross-seeding of amyloid as a link between diabetes and Alzheimer's disease	17:30 - 17:50
Gil Rosenman - Nanophotonics of Amyloid Peptide Nanostructures	17:50 - 18:10
Edoardo Salladini - Liquid-liquid phase separation by a viral protein as an additional molecular mechanism of virus-induced cell toxicity	18:10 - 18:30
Dinner, Rooftop Hotel restaurant	**19:30 -**

13 September, Thursday	
Location: Hotel Grand Spa Lietuva, Nemunas Hall	
Session 1, chair: Andrei Lupas	
Invited talk - **Kjetill Sigurd Jakobsen** - Hypervariable coding short tandem repeats as modulators of protein function	9:00 - 9:30
Fabio Parmeggiani - Custom Protein Design from Modular Building Blocks	9:30 - 9:50
Flash poster presentations, II session	9:50 - 10:30
Coffee Break	**10:30 - 11:00**
WG meetings – WG1 and WG3	11:00 - 12:30
Lunch, Hotel restaurant	**12:30 - 14:00**
Session 2, chair: Andrey Kajava	
Invited talk - **Drew Thomson** - Computational Design of Alpha-Helical Barrels	14:00 - 14:30
Invited talk - **Maria Anisimova** - A census of protein tandem repeats and their relationship with intrinsic disorder	14:30 - 15:00
Andrei Lupas - Hexads and nonads specify a new type of coiled-coil fiber	15:00 - 15:20
Aitziber L. Cortajarena - Design of artificial antennas and electron transfer chains on repeated scaffolds	15:20 - 15:40
Coffee Break	**15:40 - 16:10**
MC meeting and student event (parallel sessions)	16:10 - 18:30
Dinner, Rooftop Hotel restaurant	**19:30 -**

14 September, Friday	
Location: Hotel Grand Spa Lietuva, Nemunas Hall	
Session 1, chair: Jurij Lah	
Invited talk - **Motonori Ota** - Property of intrinsically disordered "hub" proteins	9:00 - 9:30
Petri Kursula - Intrinsic Disorder and Myelin Protein Complexes	9:30 - 9:50
Julia Marchetti - Towards biologically meaningful ensembles in disordered proteins	9:50 - 10:10
Pedro Pereira - Specific modulation of the anticoagulant activity of madanin-1 by tyrosine sulfation	10:10 - 10:30
Coffee Break	**10:30 - 11:00**
Session 2, chair: Kristina Djinovic-Carugo	
Invited talk - **Attila Remenyi** - Tales of tails: the role of disordered protein kinase regions in the regulation and assembly of kinase domain cores	11:00 - 11:30
Laura Weidmann - The Power of Reuse in the Evolution of natural Proteins	11:30 - 11:50
Pietro Sormanni - Rational design of antibodies targeting specific epitopes within disordered proteins: applications in the study of neurodegeneration	11:50 - 12:10
Flash poster presentations, III session	12:10 - 12:30
Lunch, Hotel restaurant	**12:30 - 14:00**
Session 3, chair: Oxana Galzitskaya	
Invited talk - **Isabelle Callebaut** - Looking at disorder and disorder-to-order transitions using Hydrophobic Cluster Analysis	14:00 - 14:30
Wim Vranken - Exploring the Sequence-based Prediction of Folding Initiation Sites in Proteins	14:30 - 14:50
Marcin Grynberg - ClusterLCR: a new tool to compare low complexity regions in unlike proteins	14:50 - 15:10
Flash poster presentations, IV session	15:10 - 15:30
Poster Session and Coffee Break	**15:30 - 17:00**
Session 4, chair: Sonia Longhi	
Invited talk - **Richard W. Kriwacki** - Phase Separation in Biology and Disease	17:00 - 17:30
Jordan Chill - Mechanistic Aspects of IDP Binding and Signaling Studied by NMR	17:30 - 17:50
Lukáš Žídek - Structure-function relationship in disordered proteins: knowledge acquired by NMR studies of microtubule associated proteins	17:50 - 18:10

San Hadži - Promiscuous intrinsically disordered tail from bacterial repressor mediates binding to the operator	18:10 - 18:30
Closing address	18:30 - 18:40
Dinner, Rooftop Hotel restaurant	**19:30 -**

Flash poster presentations:

I session, day 1	Z. Bednarikova	Effect of Curcumin Derivatives on Aβ40 Fibrils: A Structure – Activity Relationship Study
	O. Galzitskaya	Should the treatment of amyloidosis be personified?
	M. Žiaunys	Emergence of Visible Light Optical Properties of Phenylalanine Aggregates
	F. Figueiredo	Dissecting the ataxin-3 aggregation pathway
	K. Mihalovicova	Structure of intrinsically disordered protein tau in its pathological form
II session, day 2	S. Bliven	Phylogentics of Tandem Repeats with Circular HMMs: A Case Study on Armadillo Repeat Proteins
	L. Hirsh	Unit annotation of tandem repeat proteins, RepeatsDB-Lite
	V. Iglesias	AMYCO: Evaluation of mutational impact on prion-like proteins aggregation propensity
	P. Jane	Quantitative evaluation of the PDZ-mediated human interactome and its hijacking by viral proteins
	D. Linke	Bacterial secretion of repetitive surface fibers: inverse autotransport
	J. Manso	Crystal structure of an Ataxin-3 binding nanobody
	M. Merski	Self-Analysis of Repeat Proteins Reveals Conserved Sequence Patterns
	E. Villain	Large scale analysis of amyloidogenic regions in proteins from evolutionary diverse organisms
III session, day 3	P. Jarnot	Efficiency of custom protein sequence comparison tools to align low complexity regions
	J. Lah	Application of helix-coil theory for determination of the IDP folding energetics
	K. Melkova	Capturing transient conformations and local dynamics of intrinsically disordered microtubule-associated protein 2c
	V. Brandi	Study of the Huntingtin interactome
IV session, day 3	P. Pujols	Disulfide driven folding for a conditionally disordered protein
	J Čalyševa	Exploring internal PDZ-binding motifs
	B. Lang	Interaction predictions enable a global view of the human protein–RNA interactome and its involvement in genetic disorders
	E. Maiani	Defining a new intrinsically disordered BH3 motif

Even numbers present on Wednesday (12th of September)

Odd numbers present on Friday (14th of September) *

*odd number stands will be free for your posters by Wednesday evening

Name	Surname	Poster number
Zuzana	Bednarikova	P22
Spencer	Bliven	P01
Stanislav	Bondarev	P02
Valentina	Brandi	P23
Rima	Budvytytė	P48
Jelena	Čalyševa	P03
Ondrej	Cehlar	P06
Emanuele	Celauro	P40
Kristina	Djinovic-Carugo	P05
Diana	Fedunova	P26
Francisco	Figueiredo	P28
Jēkabs	Fridmanis	P08
Oxana	Galzitskaya	P46
Miroslav	Gančár	P30
Milana	Grbić	P07
Borbála	Hajdu-Soltész	P47
András	Hatos	P09
Layla	Hirsh Martinez	P11
Igor	Iashchishyn	P44
Valentin	Iglesias	P31
Pau	Jané	P13
Patryk	Jarnot	P15
Kristine	Kitoka	P10
Jurij	Lah	P17
Benjamin	Lang	P19
Christofer	Lendel	P42
Daria	Likholetova	P04
Dirk	Linke	P21
Sonia	Longhi	P24
Emiliano	Maiani	P25
José A.	Manso	P33
Katerina	Melkova	P27
Matthew	Merski	P29
Pablo	Mier	P32
Klaudia	Mihalovicova	P12
Marco	Necci	P34
Lisanna	Paladin	P36
Jonathan	Pansieri	P16
Elena	Papaleo	P38
Vladimir	Perovic	P37
Vasilis	Promponas	P39
Jordi	Pujols	P41
Rostislav	Skrabana	P20
Nevena	Veljkovic	P43
Etienne	Villain	P35
Mantas	Žiaunys	P14
Joanna	Ziemska	P45

* Speaker Abstracts *

O01

Prion-like propagation of alpha-synuclein protein assemblies in Parkinson's disease and related synucleinopathies

Ronald Melki
Centre National de la Recherche Scientifique (CNRS)
Paris-Saclay Institute of Neurosciences
1, Avenue de la Terrasse 91190 Gif-sur-Yvette, FRANCE

Background: Protein intracellular inclusions within the central nervous system are hallmarks of several progressive neurodegenerative disorders in man. The protein constituents of those deposits and the affected regions within the brain differ from one neurodegenerative disorder to another. Until recently, the vicious circle consisting of spread, seeded assembly and accumulation over time within the central nervous system of misfolded proteins aggregates was thought to be restricted to the prion protein PrP. Recent reports suggest that other protein aggregates spread and amplify within the central nervous system leading to distinct diseases.

Questions addressed: How alpha-synuclein protein assemblies traffic between cells, amplify by recruiting endogenous monomeric alpha-synuclein and cause distinct synucleinopathies is unclear.

Results and discussion: I will present data illustrating the propagation propensities of alpha-synuclein assemblies. I will discuss the nature of protein assemblies that are "Infectious", how they bind to the cell membranes, what they bind to and the cellular consequences of binding. I will present a quantitative assessment of their uptake, transport and export. I will show data demonstrating that pathogenic protein assemblies disrupt the endo-lysosomal membranes to reach the cytosol where they amplify. Finally, I will describe how and why different alpha-synuclein polymorphs cause distinct diseases. Strategies targeting the propagation of protein assemblies involved in age-related dementias will be presented and discussed.

References:

1. Pieri L et al. (2012) Biophys J. 102: 2894-905.
2. Bousset L et al. (2013) Nat Commun. 4:2575
3. Peelaerts W et al. (2015) Nature 522:340-4.
4. Shrivastava AN et al. (2015) EMBO J. 34 :2408-23.
5. Brahic M et al. (2016) Acta Neuropathol. 131:539-48.
6. Pieri L et al. (2016) Sci Rep. 6:24526.
7. Makky A et al. (2016) Sci Rep. 6:37970.
8. Flavin W et al. (2017) Acta Neuropathol. 134:629-653.
9. Shrivastava et al. (2017) Neuron 95:33-50.
10. Melki R (2018) Neurobiol Dis. 109 :201-208

O02

Linking amyloid and neuroinflammatory cascades in neurodegenerative diseases – role of pro-inflammatory S100A9 protein

Ludmilla A. Morozova-Roche
Department of Medical Biochemistry and Biophysics, Umeå University, Umeå, Sweden

Background: Chronic neuroinflammation is a hallmark of neurodegenerative diseases, associated with increased levels of pro-inflammatory factors in the brain tissues.

Questions addressed: We have studied the role of S100A9 protein in neurodegenerative diseases, which acts both as pro-inflammatory mediator and highly amyloidogenic protein, linking the amyloid and neuroinflammatory cascades.

Methods: Immunohistochemistry, immunofluorescence, fluorescence, circular dichroism, AFM and SEM imaging, ELISA and SPR techniques were used.

Results and discussion: We have shown that the pro-inflammatory mediator and highly amyloidogenic protein S100A9 is involved in the amyloid-neuroinflammatory cascades in Alzheimer's disease, where it co-aggregates with Aβ peptide, contributing to amyloid plaque formation and intracellular aggregation [1], in Parkinson's disease – it forms co-aggregates with α-synuclein and involves in Lewy body formation [2], and in traumatic brain injury – S100A9 is highly abundant in the brain tissues and forms numerous precursor-plaques. In wild-type mouse model the intranasal administration of S100A9 amyloids induces wide-spread cellular stress responses in the brain tissues and Alzheimer's-like behavioural impairment in passive avoidance test [3]. In vitro S100A9 easily aggregates under native pH 7.4 and 37 °C conditions and this process is well quantitatively described by generic Finke-Watzky two-step nucleation-autocatalytic growth model [4]. The co-aggregation of S100A9 with Aβ peptide or α-synuclein occurs significantly faster and leads to formation of larger amyloid aggregates than the self-assembly of individual proteins. S100A9 amyloid oligomers are more toxic than those of Aβ or α-synuclein, while the co-aggregation with Aβ/ α-synuclein mitigates the cytotoxicity of S100A9 oligomers. The levels of S100A9 follow those of Aβ in cerebrospinal fluid during the development of Alzheimer's disease and S100A9 together with Aβ can serve as a biomarker for early stages of Alzheimer's disease, starting from mild cognitive impairment [5]. The finding of S100A9 involvement in neurodegenerative diseases may open a new avenue for therapeutic interventions targeting pro-inflammatory S100A9 and preventing the amyloid self-assembly in affected brain tissues.

References:

1. Wang, C., Klechikov, A.G., Gharibyan, A.L., Wärmländer, S.K., Jarvet, J., Zhao, L., Jia, X., Shankar, SK., Olofsson, A., Brännström, T., Mu, Y., Gräslund, A., Morozova-Roche, L.A. The role of pro-inflammatory S100A9 in Alzheimer's disease amyloid-neuroinflammatory cascade. *Acta Neuropathol*., 127, 507-522, 2014.
2. Horvath, I., Yashchishyn, I.A., Wang, C., Moskalenko, R.A., Wärmländer, S.K.T.S., Wallin, C., Gräslund, A., Kovacs, G.G., Morozova-Roche, L.A. Co-aggregation of pro-inflammatory S100A9 with α-synuclein in Parkinson's disease: *ex vivo* and *in vitro* studies. *J. Neuroimmun.,*2018.
3. Iashchishyn, I.A., Gruden, M., Moskalenko, R.A., Davydova, T.V., Wang, C., Sewell, R.D.E., Morozova-Roche, L.A. *ACS Chem Neurosci.,* 2018.
4. Iashchishyn. I.A., Sulskis, D., Nguyen Ngoc, M., Smirnovas, V., Morozova-Roche, L.A. Finke-Watzky two-step nucleation-autocatalysis model of S100A9 amyloid formation: protein misfolding as "nucleation" event. *ACS Chem Neurosci.,* 8, 2152-2158, 2017.
5. Horvath, I., Jia, X., Johansson, P., Wang, C., Moskalenko, R., Steinau, A., Forsgren, L., Wågberg, T., Svensson, J., Zetterberg, H., Morozova-Roche, L.A. Pro-inflammatory S100A9 protein as a robust biomarker differentiating early stages of cognitive impairment in Alzheimer's disease. *ACS Chem. Neurosci.* 7, 34-39, 2016.

.

O03

A novel small molecule inhibits α-synuclein aggregation, disrupts amyloid fibrils and prevents degeneration of dopaminergic neurons

Ventura, Salvador[1], Pujols, Jordi[1], Peña-Díaz, Samuel[1], Salvatella, Xavier[2], Sancho, Javier[3], Sodupee, Mariona[4], Outeiro, Tiago F.[5], Dalfó, Esther[6]

[1] Institut de Biotecnologia i Biomedicina. Universitat Autònoma de Barcelona. [2]Institute for Research in Biomedicine (IRB Barcelona), [3]Department of Biochemistry and Molecular and Cell Biology, Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza. [4]Departament de Química, Universitat Autònoma de Barcelona. [5]Deparment of Experimental Neurodegeneration, Center for Biostructural Imaging of Neurodegeneration, Center for Nanoscale Microscopy and Molecular Physiology of the Brain, University Medical Center Göttingen. [6]Faculty of Medicine. University of Vic-Central University of Catalonia (UVic-UCC).

Background: Parkinson's disease (PD) is characterized by a progressive loss of dopaminergic neurons, a process that current therapeutic approaches cannot prevent. In PD, the typical pathological hallmark is the accumulation of intracellular protein inclusions, known as Lewy bodies and Lewy neurites, which are mainly composed of α-synuclein.

Questions addressed: Interfering with α-Syn aggregation has been envisioned as a promising disease-modifying approach for the treatment of PD. However, the disordered nature of α-Syn precludes the use of structure-based drug design and optimization for the discovery of novel molecules able to modulate α-Syn aggregation.

Methods: Here, we exploited a high-throughput screening methodology to identify a small molecule (SynuClean-D) able to inhibit α-synuclein aggregation.

Results and discussion: SynuClean-D significantly reduces the *in vitro* aggregation of *wild-type* α-synuclein and the familiar A30P and H50Q variants in a substochiometric molar ratio. This compound prevents fibril propagation in protein misfolding cyclic amplification assays and decreases the number of α-synuclein inclusions in human neuroglioma cells. Computational analysis suggests that SynuClean-D can bind to cavities in mature α-synuclein fibrils and, indeed, it displays a strong fibril disaggregation activity. The treatment with SynuClean-D of two PD *Caenorhabditis elegans* models, expressing α-synuclein either in muscle or in dopaminergic neurons, significantly reduces the toxicity exerted by α-synuclein. SynuClean-D treated worms show decreased α-synuclein aggregation in muscle and a concomitant motility recovery. More importantly, this compound is able to rescue dopaminergic neurons from α-synuclein induced degeneration. Overall, SynuClean-D appears as a promising molecule for therapeutic intervention in Parkinson's Disease.

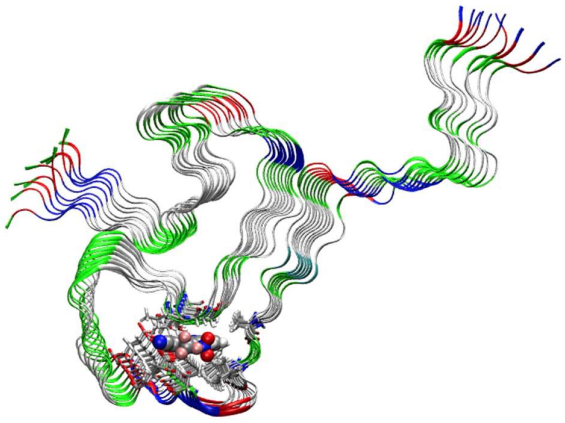


Figure 1. Docking of SynuClean-D into an inner cavity of α-synuclein amyloid fibrils

O04

Engineering bugs for the discovery of new drugs against diseases caused by protein misfolding and aggregation

Skretas, Georgios [1]*, Matis, Ilias[1,2], Delivoria, Dafni Chrysanthi[1,2], Mavroidi, Barbara[3], Papaevgeniou, Nikoletta[1,4], Panoutsou, Stefania[1,5], Bellou, Stamatia[1], Pelecanou, Maria[3], Chondrogianni, Niki[1]

[1]Institute of Biology, Medicinal Chemistry & Biotechnology, National Hellenic Research Foundation, Athens, 11635, Greece; [2]School of Chemical Engineering, National Technical University of Athens, Athens, 157 80, Greece; [3]Institute of Biosciences & Applications, National Center for Scientific Research "Demokritos", Athens, 15341, Greece; [4]Faculty of Biology and Pharmacy, Institute of Nutrition, Friedrich Schiller University of Jena, Jena, 07743, Germany

Protein misfolding and aggregation is a common pathological feature for many human diseases, such as Alzheimer's disease (AD), Parkinson's disease, type 2 diabetes and others. We will describe the development and application of an integrated and generalizable bacterial platform for facile discovery of chemical rescuers of disease-associated protein misfolding [1]. In this system, large combinatorial libraries of macrocyclic molecules 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 β-amyloid 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 (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.

References:

1. Matis, I., Delivoria, D.C., Mavroidi, B., Papaevgeniou, N., Panoutsou, S., Bellou, S., Papavasileiou, K.D., Linardaki, Z., Stavropoulou, A.V., Vekrellis, K., Boukos, N., Kolisis, F.N., Gonos, E.S., Margarity, M., Papadopoulos, M.G., Efthimiopoulos, S., Pelecanou, M., Chondrogianni, N., Skretas, G*. 2017. An integrated and generalizable bacterial discovery platform for chemical rescuers of disease-associated protein misfolding. *Nature Biomedical Engineering*. 1, 838–852.

O05

Amyloid fibrils from the N-terminal prion protein fragment as a model for studying the mechanism of prion propagation and transmissibility barriers

Witold K. Surewicz[1], Jin-Kyu Choi[1], Ignazio Cali[2], Theint Theint[3], Qiuye Li[1], Krystyna Surewicz[1], Qingzhong Kong[2], Darryl Aucoin[3], Philippe Nadaud[3], Pierluigi Gambetti[2], Christopher Jaroniec[3]

[1]Departments of Physiology and Biophysics and [2]Pathology, Case Western Reserve University, Cleveland OH 44106, USA and [3]Department of Chemistry, Ohio State University, Columbus, OH 43210, USA

Background: Prions are infectious amyloids that cause a group of transmissible neurodegenerative diseases (known as transmissible spongiform encephalopathies) that include Creutzfeldt-Jakob disease in humans, scrapie in sheep, bovine spongiform encephalopathy in cattle, and chronic wasting disease in deer and elk. These unusual pathogens are believed to propagate by the templated (seeded) conversion of the normal prion protein (PrP) to the amyloid form (1).

Questions addressed: The mechanistic and structural aspects of prion protein conversion to the infectious form are poorly understood. Little is also known about the structural basis of prion strains and transmissibility barriers.

Results and discussion: Recently we have shown that amyloid fibrils from the N-terminal part of the recombinant mouse prion protein (PrP23-144) are infectious, causing transmissible prion disease in mice (2). Thus, these fibrils which, unlike other mammalian prions, are amenable to detailed biophysical and structural analysis, are of unique value as a model for exploring the mechanism of prion protein conformational conversion as well as the molecular and structural basis of prion strains and transmissibility barriers. Here we will describe a series of low- and high-resolution biophysical experiments, including solid-state NMR spectroscopy, with PrP23-144 amyloids from three species: human, mouse and Syrian hamster. Our data show that these fibrils adopt different structures, and these species-dependent structural differences are controlled by only two residues within the amyloid core region. Importantly, species-specific structural differences correlate with seeding specificities of PrP23-144 amyloid fibrils. The role of these critical amino acid residues as conformational switches can be rationalized based on the structural model for human PrP23-144 amyloid fibrils, providing a foundation for understanding cross-seeding specificity.

References:

1. Cobb, N.J., Surewicz, W.K., Prion diseases and their biochemical mechanisms, *Biochemistry* 48, 2574-2585 (2009)

2. Choi, JK, Cali, Surewicz, K, Kong, Q, Gambetti, P, Surewicz, WK, Amyloid fibrils from the N-terminal prion protein fragment are infectious, *Proc Natl Acad Sci USA* 113, 13851-56 (2016)

O06

Structural aspects of naturally-occurring amyloids

Kajava Andrey[1]

[1]*Centre de Recherche en Biologie cellulaire de Montpellier CNRS, Institut de Biologie Computationnelle, Montpellier, France*

Background: In most of the cases, the exact structure of protein aggregates are unknown due to the fact that methods of high resolution structure determination (x-ray crystallography and NMR spectroscopy) cannot be used because of the insolubility of the aggregates. However, over the past decade, substantial progress has been made in understanding the structure of an important type of the aggregates called “cross-beta amyloid fibrils”. Progress has stemmed largely from the application of solid state NMR, scanning transmission electron microscopy mass measurements and cryo-electron microscopy.

Questions addressed: The number of structural data about naturally-occurring and disease-related cross-beta amyloids is now large enough to support a systematic analysis. In this presentation, I will survey and classify the known structures of these amyloids. The analysis yielded an enhanced understanding of their “pathogenic fold” determinants. I will demonstrate how this structural information can be used to successfully predict amyloidogenicity of proteins based on their amino acid sequence [1,2,3] . Evaluation of the effect of globular domains located in close vicinity of the amyloid-forming regions will be also a theme of this presentation [4].

References:

1. Ahmed A.B., Znassi N., Château M.T., Kajava A.V. - A structure-based approach to predict predisposition to amyloidosis. (2015) Alzheimers Dement. 11(6):681–690.
2. D. B. Roche, E. Villain and A. V. Kajava - Usage of a dataset of NMR resolved protein structures to test aggregation versus solubility prediction algorithms (2017) Protein Science 26(9), 1864-1869
3. S. A. Bondarev, O. V. Bondareva, G. A. Zhouravleva and A. V. Kajava - BetaSerpentine: a bioinformatics tool for reconstruction of amyloid structures (2018) Bioinformatics 34(4):599-608.
4. R. A. Azizyana, A. Garro, Z. Radkova, A. Anikeenko, A. Bakulina, C. Dumas and A. V. Kajava Establishment of constraints on amyloid formation imposed by steric exclusion of globular domains (2018) J. Mol. Biol (in revision).

O07

Biomimicry of Artificial Spider Silk Spinning Assessed by NMR Spectroscopy

Otikovs, Martins[1], Andersson, Marlene[2], Jia, Qiupin[3,4], Nordling, Kerstin[3], Meng, Qing[4], Andreas, Loren B.[5], Pintacuda, Guido[5], Johansson, Jan[3], Rising, Anna[2,3], Jaudzems, Kristaps[1]

[1]Latvian Institute of Organic Synthesis, Riga, Latvia [2]Department of Anatomy, Physiology and Biochemistry, Swedish University of Agricultural Sciences, Uppsala, Sweden [3]Department of Neurobiology, Care Sciences and Society (NVS), Division for Neurogeriatrics, Karolinska Institutet, Stockholm, Sweden [4]Institute of Biological Sciences and Biotechnology, Donghua University, Shanghai, China [5]Institut des Sciences Analytiques, Universite de Lyon, Villeurbanne, France

Background: Spider silk, one of the toughest biomaterials known, is produced through the assembly of large proteins (spidroins) that consist of three units: a central repetitive region which accounts for spider silk's exceptional mechanical properties and two terminal domains (NT and CT) implicated in the silk formation process. Despite many efforts, the mechanical properties of current artificial spider silks lag behind their natural counterparts. The main reason for this is the inability to reproduce the complex molecular mechanisms of native silk spinning. Recently, a novel minispidroin that recapitulates the solubility and pH responsivity of native spidroins was developed by one of our groups [1]. This construct, termed NT2RepCT, is composed of an NT and two repeat regions (2Rep) from *E. australis* major ampullate spidroin 1 and a CT from *A. ventricosus* minor ampullate spidroin, and allows biomimetic spinning using solely aqueous buffers, shear forces, and lowered pH yielding the toughest as-spun artificial spider silk fiber to date.

Questions addressed: Biomimetic spinning of artificial spider silk requires that the terminal domains of designed minispidroins undergo specific structural changes in concert with the β-sheet conversion of the repetitive region in response to altered pH, ion composition, and shear forces [2, 3]. Herein, we analyze the domain-specific structural changes of NT2RepCT upon silk fiber formation.

Methods: We combine solution and solid-state NMR methods with amino acid-type specific labeling to probe domain-specific structural changes that take place during fiber spinning from the designed NT2RepCT minispidroin, which allows us to assess the degree of biomimicry of artificial silk spinning.

Results and discussion: Our approach [4] allowed evaluation of structural changes experienced by the different domains of NT2RepCT upon fiber spinning and drying. Solution NMR spectroscopy of the NT2RepCT minispidroin allowed probing of inter-domain interactions and structural changes that occur before fiber formation in response to low pH. Solid-state NMR spectroscopy of the NT2RepCT fibers revealed the degree of conversion to β-sheet structure and the conformation of the terminal domains that relates to the mimicry of the mechanism of native silk spinning. In addition, we show that the structural effects of post-spinning procedures can be examined. By studying the impact of NT2RepCT fiber drying, we observed a reversible beta-to-alpha conversion. We think that this approach will be useful for guiding the optimization of artificial spider silk fibers.

References:

1. Andersson, M.; Jia, Q.; Abella, A.; Lee, X.Y.; Landreh, M.; Purhonen, P.; Hebert, H.; Tenje, M.; Robinson, C.V.; Meng, Q.; Plaza, G.R.; Johansson, J.; Rising, A. *Nat. Chem. Biol.* **2017**, 13: 262-264.

2. Andersson, M.; Chen, G.; Otikovs, M.; Landreh, M.; Nordling, K.; Kronqvist, N.; Westermark, P.; Jörnvall, H.; Knight, S.; Ridderstråle, Y.; Holm, L.; Meng, Q.; Jaudzems, K.; Chesler, M.; Johansson, J.; Rising, A. *PLoS Biol.* **2014**, 12, e1001921.

3. Kronqvist, N.; Otikovs, M.; Chmyrov, V.; Chen, G.; Andersson, M.; Nordling, K.; Landreh, M.; Sarr, M.; Jörnvall, H.; Wennmalm, S.; Widengren, J.; Meng, Q.; Rising, A.; Otzen, D.; Knight, S.D.; Jaudzems, K.; Johansson, J. *Nat. Commun.* **2014**, 5, 3254.

4. Otikovs, M.; Andersson, M.; Jia, Q.; Nordling, K.; Meng, Q.; Andreas, L.B.; Pintacuda, G.; Johansson, J.; Rising, A.; Jaudzems, K. *Angew. Chem. Int. Ed.* **2017**, 56, 12571–12575.

.

O08

Extreme Polymorphism in Amyloids from Pathogenic Microbes

Meytal Landau[1], Einav Tayeb-Fligelman[1], Nir Salinas[1], Sergei Perov[1], Ofir Lidor[1], Nimrod Golan[1], Itzik Engelberg[1]

[1]Department of Biology, Technion – Israel Institute of Technology, Haifa, Israel

Background: Microbial functional amyloids are structured protein aggregates that serve diverse functions, including biofilm structuring and interactions with the host. Despite their role as key virulence factors and antimicrobial drug targets, amyloids are mostly known for their involvement in fatal human aggregation diseases, and their structures have been studied mostly in eukaryotes.

Questions addressed: The structural hallmarks of functional amyloids – if any – and how they can be distinguished from disease-associated amyloids remain unclear. We investigated the structure-function-fibrillation relationships of microbial functional amyloids, their interactions with host amyloids and receptors and explore routes to modulate their activities.

Methods: We leverage unique methodologies of X-ray microcrystallography for the structural studies of polypeptides and segments of bacterial functional amyloids. The structural and biophysical properties of the amyloid fibrils are correlated with the functions they encode.

Results and discussion: We discovered unique amyloid-like structures in polypeptides and segments of microbial amyloids, including, to our surprise, a structure of a full-length bacterial cross-alpha amyloid-like fibril revealing surprising departure from pathological amyloids in which beta-rich structures are central. The fibrils, of the PSMα3 peptide secreted by *Staphylococcus aureus*, are toxic to human cells, clarifying their involvement in pathogenicity [1]. In contrast, amyloidogenic peptides involved in biofilm structuring share similar atomic structures with pathological amyloids, forming highly stable cross-beta structures that stabilize the biofilm matrix [2]. Surprisingly, three fibrillating antibacterial peptides secreted by different organisms (bacteria, amphibian and human) exposed extremely polymorphic fibrous architectures, all markedly different from the cross-beta fibril. Given our results we predict that the structural and functional repertoire of functional amyloids is far more diverse than previously anticipated, providing a rich source of targets for antimicrobial drug discovery.

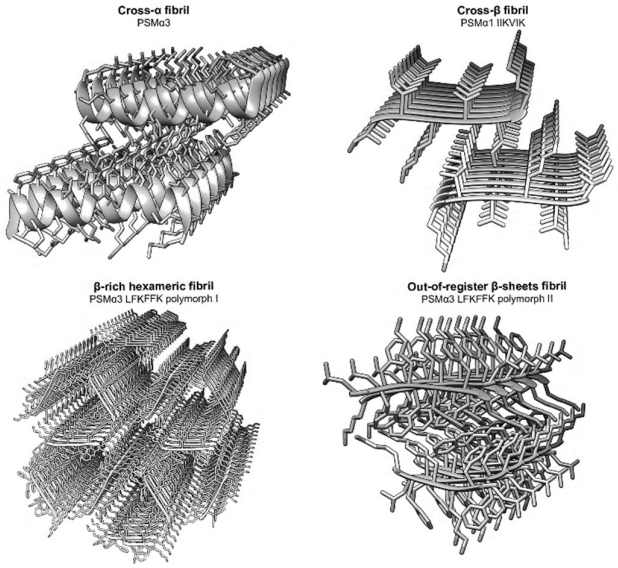


Figure 1. Highly polymorphic structures of *S. aureus* PSMα peptides.

References:

1. Tayeb-Fligelman *et. al., Science* 355(6327): 831-833; 2017

2. Salinas *et.al. Under revisions*; https://www.biorxiv.org/content/early/2018/06/18/309062

O09

Cross-seeding of amyloid as a link between diabetes and Alzheimer's disease

Westermark Gunilla T[1], Westermark Per[2], Hyman Bradley T[3], Ingelsson Martin[4]

[1]Department of Medical Cell Biology, [2]Immunology, Genetics and Pathology, [4]Public Health and Caring Science and Uppsala University, SE-75123, Uppsala, Sweden.

[3]Departments of Neurology and Radiology; Massachusetts General Hospital, Charlestown, USA and Neuroscience Program, Harvard Medical School, Boston, USA.

Background Epidemiological studies have identified type 2 diabetes (T2D) as a risk factor for Alzheimer's disease (AD). T2D and AD are multifactorial diseases but have in common the formation of local amyloid. We have shown that a single intravenous injection of preformed Aβ-fibrils in hIAPP transgenic mice increases deposition in IAPP amyloid. Also, we have identified co-deposition of IAPP and Aβ in amyloid-deposits in brain, a finding that suggests that IAPP can cross-seed Aβ [1]. There is evidence supporting inflammatory processes in connection to amyloid but, although presence of inflammation has been established, it is not clear if inflammation can be beneficial, harmful or even a trigger of the disease.

Questions addressed: Are there any differences in the inflammatory process in individuals diagnosed with T2D and AD compared to individuals diagnosed with AD only?

Material and methods: Temporal cortex biopsies from patients with T2D+AD (9) and AD (9) were used for mRNA extraction and expression of IL6, IL1β, IL18, CCL2, CXCL1, NLRP3, Pycard, Casp1, CD163, CD14, GFAP, CD4, AIF1, and ADAM10 were analysed with qPCR.

Results and discussion: The inflammasome, a component of the innate immune system, is regulating the immune response. Activated NLRP3 inflammasome triggers the immune system by cleaving pro-caspase1 to caspase1, which in turn activates pro-inflammatory cytokines such as IL1-β and IL-18, inducing tissue damage. It is still unclear which cells that produce IL1-β, but mRNA analysis performed on the human neuronal cell line SHSY5Y shows the expression of IL1-β. The results of the qPCR analysis on brain tissue showed almost a seven-fold higher expression of NLRP3 and a four-fold higher expression of IL-1β in patients with T2D + AD compared to expression levels determined in AD patients without T2D. On the contrary a significantly lower expression of AIF1, a marker for microglia, was detected in patients with T2D+AD. IL1-β is important for recruitment of microglia [2] and an increased inflammasome activation with a concomitant decrease in microglial suggest impairment of the normal pathway leading to a decrease in cortical phagocytic activity in patients with T2D and AD. How this is linked to the presence of more compact neuritic (Aβ1-40) plaques in the cerebral cortex of T2D and AD will be discussed.

References:

1. Oskarsson ME, Paulsson JF, Schultz SW, et al. In vivo seeding and cross-seeding of localized amyloidosis: A molecular link between type 2 diabetes and Alzheimer's disease. Am J Path. 2015, 3:834-846.
2. Halle A, Hornung V, Petzold GC, et al. The NALP3 inflammasome is involved in the innate immune response to amyloid-beta. Nat Immunol. 2008, 8:857-65.

O10

Nanophotonics of Amyloid Peptide Nanostructures

Prof. Gil Rosenman
School of Electrical Engineering, Tel Aviv University

Background: Encoding human genome enabled development of a new generation of bionanomaterials self-assembled from chemically synthesized biomolecules. These bioinspired peptide/protein nanostructures opened the avenue for wide application fields such as tissue engineering, new generation of drugs and nanotechnology. In this work, we propose a new paradigm of peptide nanophotonics and present a new concept of conversion of the originally non-fluorescent peptide/protein nanodots and nanostructures into visible ones by reconformation of their peptide secondary structure.

Questions addressed: Bioimaging is a principal method used in biomedical research and medical trails. It is mostly based on visible fluorescent agents, which combine exceptional optical properties with biocompatibility and biodegradability. Original electronic structure of elementary biomolecules does not enable generate visible fluorescence (FL) except the unique protein of the jellyfish *Aequorea Victoria* and its homologs. In this work we report on a new physics and new technology of visible FL found in non-disease amyloid-like bioinspired nanodots and nanostructures.

Methods: This work covers peptide nanotechnology and fabrication process of peptide nanodots and peptide integrated nanophotonic circuits. We show that optical properties of this new generation of biophotonic materials are governed by fundamental biological processes and could be strongly changed by reconformation of their peptide secondary structure leading to visible FL.

Results and discussion: We found a deep modification of basic physical properties in these materials of biological origin such as elementary symmetry, piezoelectric, linear and non-linear optical and waveguiding effects by thermally-mediated refolding peptide secondary structure from native phase to □-sheet architecture. This biological phase transformation is followed by the appearance of visible tunable fluorescence which has the same physical origin as visible FL found in amyloid fibrils associated with neurodegenerative Alzheimer and Parkinson diseases. The physical origin of the found new biofluorescence phenomenon is ascribed to reassembling of intermolecular hydrogen bonds, stabilizing antiparallel β-sheets structure and reconstruction of their electronic structure. Proposed new concept of peptide secondary structure refolding has been applied for development of new visible multicolour bionanodots with quantum yield reaching ~30%. Another application of this biophotonic phenomenon is peptide integrated optical biochips toward precise medicine for diagnosis, light-induced and activated therapy, optogenetics, health monitoring and more.

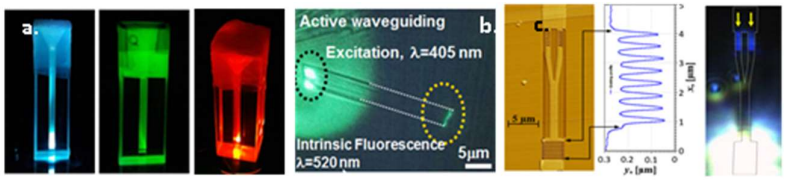


Figure 1. a. Multicolor visible tri-phenylalanine (FFF)-nanodots; **b**. Active waveguiding in FFF-peptide tape; **c.** Peptide integrated chip: FFF -based 1x2-power Y-optical splitter with grating couplers at all three ports.

References:

1. A.Handelman, N. Lapshina, B. Apter, B. Fainberg, G. Rosenman, Peptide Integrated Optics, Adv. Mater. 30, 1705776, 2018.

2. B. Apter, N. Lapshina, A. Handelman, B. D. Fainberg, G. Rosenman,
Peptide Nanophotonics: From Optical Waveguiding to Precise Medicine and Implantable Biochips, Invited Review, Small, accepted May 10, 2018.

O11

Liquid-liquid phase separation by a viral protein as an additional molecular mechanism of virus-induced cell toxicity

Edoardo Salladini[1], Christophe Bignon[1], Vincent Delauzun[1], Claire Debarnot[1], Silvia Spinelli[1], Maria Grazia Murrali[2], Roberta Pierattelli[2] and Sonia Longhi[1]

[1]Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 7257, Centre National de la Recherche Scientifique (CNRS) & Aix-Marseille University, Marseille, France; [2] CERM University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy.

Background: Henipaviruses are severe human pathogens within the *Paramyxoviridae* family. They found their natural reservoir in fruit-eating bats of the *Pteropus* genus and are responsible for severe encephalitis in humans [1]. Those viruses are capable to counteract the host innate immune response thanks to three different accessory proteins (V, W and C) that are all encoded by the P gene [2]. These proteins are therefore promising targets for antiviral strategies. We have previously reported a biophysical characterization of *Henipavirus* V proteins and shown that they are mainly disordered and interact with DDB1 [3], a cellular protein that is a component of the ubiquitin ligase E3 complex. In the course of an in-depth characterization of *Henipavirus* accessory proteins we serendipitously discovered that one of them undergoes a liquid-hydrogel phase transition and a liquid-liquid phase separation.

Questions addressed: We aimed at investigating this liquid-liquid phase separation from a biophysical and structural point of view, as well as at unravelling its biological significance.

Methods: Thioflavin T (ThT) and Congo red (CR) binding assays, circular dichroism, small angle X-ray scattering, NMR, optical and electron microscopy, cell viability assays.

Results and discussion: By combining experimental and bioinformatics approaches, we have identified the region responsible for liquid-liquid phase separation and have further investigated it using a combination of biophysical and structural approaches. ThT and CR binding assays, together with negative-staining electron microscopy studies show that this region forms amyloid-like fibrils according to the model shown in Figure 1. Finally, we show that mammal cells transfected to express this region bind CR and exhibit an increased sensitivity towards a stress agent suggesting a new mechanism of virus-induced cell toxicity.

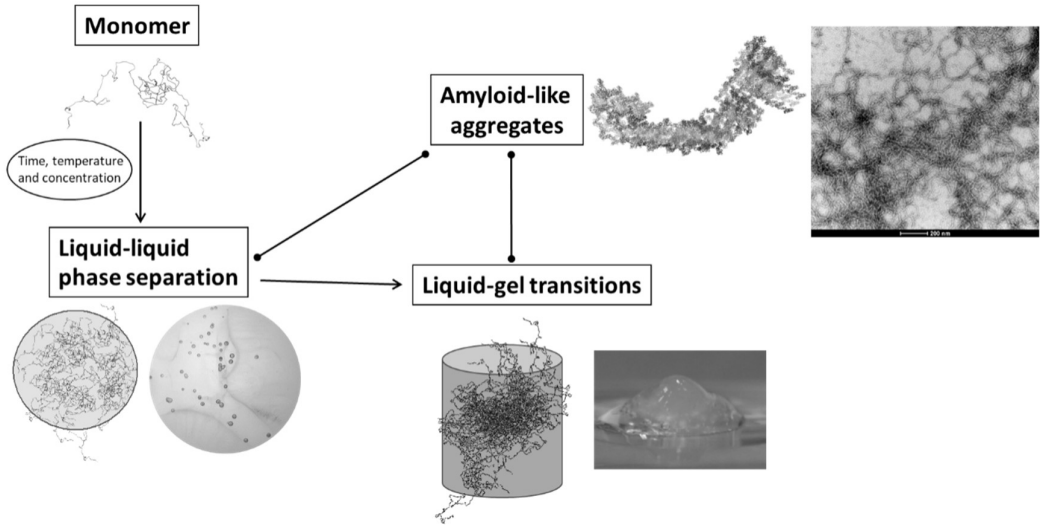


Figure 1. Schematic representation of the different states and conformation that the protein can adopt.

References:

1. M. Aljofan. Hendra and Nipah infection: Emerging paramyxoviruses. *Virus Res.* 2013, 177(2):119-26.
2. J. M. Fontana, B. Bankamp and P. A. Rota. Inhibition of interferon induction and signaling by paramyxoviruses. *Immunol. Rev*., 2008, 225:46–67.
3. E. Salladini, V. Delauzun and S. Longhi. The Henipavirus V protein is a prevalently unfolded protein with a zincfinger domain involved in binding to DDB1. *Mol. BioSyst*. 2017, 13(11):2254-2267.

O12

Hypervariable coding short tandem repeats as modulators of protein function

Jakobsen, Kjetill S[1], Reinar, William B[1], Knudsen, Jonfinn B[2], Tørresen, Ole K[1], Greulich Anne[1,2], Nederbragt, Alexander J[1,3], Stø, Ida M[2], Jentoft, Sissel[1], Butenko, Melinka A[2]

[1]Centre for Ecological and Evolutionary Synthesis, Dept. of Biosciences, University of Oslo, Norway, [2] Section for Genetics and Evolutionary Biology, Dept. of Biosciences, University of Oslo, Norway, [3]Biomedical Informatics Research Group, Dept. of Informatics, University of Oslo

Background: Short tandem repeats (STRs), here defined as units of 1-6 base pairs repeated in tandem (also termed microsatellites), are present in genes and in intergenic regions throughout most genomes. STR mutations arising due to replication slippage either lead to a reduction or an increase in the number of repeated units. STRs have estimated mutation rates of 1×10^{-4} to 1×10^{-3} per locus[1] - orders of magnitude greater than point mutations. Thus, length variations in STRs either at the population level or even at the individual level (heterozygosity) are frequently observed. For instance, the number of repeated CAG units in a STR within the coding part of the *huntingtin* gene correlates with the age of onset of Huntington's disease[2]. However, such dramatic phenotypic effects are likely exceptions. Studies of STRs in other organisms have shown that length variations in certain STRs can modulate gene and protein function without inducing deleterious effects[3,4].

Questions addressed: Despite the increased interest in genic short tandem repeats, the biological importance of genetic variation caused by length variation in STRs has been more or less overlooked. This is partly due to technical challenges associated with DNA sequencing and partly due to bioinformatic limitations. We have investigated the presence and length variation in STRs at the whole genome level in two very different systems, bony fishes[5,6] and in the plant *Arabidopsis thaliana* (representing aquatic and terrestrial life). If STRs provide adaptive benefits, correlations between STR length variants and the environment should be evident. For addressing this question, we analysed whole-genome data from more than a thousand *A. thaliana* natural specimens (accessions), sequenced by the 1001 Genomes Consortium[7,8].

Methods: We used whole genome analysis (WGA) of selected fish genomes sequenced with both short-read and long-read technologies, and WGA of over 1000 *A. thaliana* wild specimens representing the global distribution of the species. Newly developed software tools such as hipSTR and Tandem Repeats Finder were used. Models for environmental variables and STR variations were developed. Altogether, this represents a versatile pipeline for analyzing whole genome variation in STRs and repeat length variation in relation to phenotypes and environmental variables[8]. STR-containing genes were grouped according to functions/protein categories and analyzed *in silico* for structural features (including intrinsically disordered regions).

Results and discussion: We found that both fish and plant genomes harbor substantial amounts of coding STRs. Certain gene (protein) classes seem to be particularly enriched including genes involved in transcriptional regulation, chromatin remodeling and morphogenesis. In some fish species (Atlantic cod and haddock) a substantially higher proportion of STR-containing genes compared to other fishes was observed – to a large extent due to high numbers of STRs in genes involved in signal transduction. In Arabidopsis, STR loci differed in length throughout its geographical distribution, and that the length variation in most cases was significantly correlated to specific environmental conditions. Of these, almost one third of the variants are located in the vicinity or within genes. Strikingly, close to 80 % of STRs located within protein coding sequence had length variation significantly associated with bioclimatic variables. We show that the coding STRs tend to overlap with putative protein binding sites, indicating a functional role for STRs in protein-protein and protein-DNA interactions. Moreover, we found that STR lengths co-varied with gene expression levels, and with variation in key adaptive phenotypic traits, such as the timing of flowering. Taken together, our results suggest that particular STR length variants provide advantages under certain biotic or abiotic conditions through fine-tuning of protein interactions, structure, regulation and thus, play a major role in facilitating adaptation.

References:

1. Sun, J. X. *et al.* A direct characterization of human mutation based on microsatellites. *Nat. Genet.* **44,** 1161–1165 (2012).
2. Stine, O. C. *et al.* Correlation between the onset age of Huntington's disease and length of the trinucleotide repeat in IT-15. *Hum. Mol. Genet.* **2,** 1547–1549 (1993).
3. Gemayel, R. *et al.* Variable Glutamine-Rich Repeats Modulate Transcription Factor Activity. *Mol. Cell* **59,** 615–627 (2015).
4. Undurraga, S. F. *et al.* Background-dependent effects of polyglutamine variation in the Arabidopsis thaliana gene ELF3. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 19363–19367 (2012).

5. Tørresen O.K. et al. An improved genome assembly uncovers prolific tandem repeats in Atlantic cod. *BMC Genomics* **18**:95 (2017)
6. Tørresen O.K. et al. S Genomic architecture of haddock (*Melanogrammus aeglefinus*) shows expansions of innate immune genes and short tandem repeats. *BMC Genomics* **19**:240 (2018)
7. 1001 Genomes Consortium. 1,135 Genomes Reveal the Global Pattern of Polymorphism in Arabidopsis thaliana. *Cell* **166,**481–491 (2016).
8. Reinar WB, Knudsen JB, Jentoft S, Tørresen OK, Butenko MA, Jakobsen KS (2018) Evolutionary adaptability linked to length variations in short genic tandem repeats. *BioRxiv* https://doi.org/10.1101/310045

O13

Custom Protein Design from Modular Building Blocks

Parmeggiani, Fabio[1], Pedrelli, Giacomo[1], Yeh, Chun-Ting[1], Ozture, Mithat Berk[1], Harrison, Laura[1]
[1]School of Chemistry, School of Biochemistry, University of Bristol, UK

Background: Advances in computational methods allow, nowadays, the design of novel proteins not observed in Nature, in particular repeat proteins with extraordinary stability[1] and for which is possible to control the local and overall shape[2]. However, design methods are usually limited to small proteins and symmetric systems, without the possibility to define large-scale structures with custom shapes. Therefore, we developed ELFIN[3], a novel computational approach for design of custom proteins. It enables users to build protein structures with specific shapes, using experimentally validated repeat protein units as modular, compatible and rigid building blocks.

Questions addressed: We wanted to explore the potential of designed repeat proteins as a versatile tool to develop stable custom scaffolds. The goal is to provide a computational and experimental platform for reliable design of novel protein nanoparticle and nanostructures. We are particularly interested in controlling receptor signaling by spatially organize ligands through designed scaffolds.

Methods: Similar to DNA nanostructure design tools, we define a three-dimensional target shape and find the combination of structural building blocks that matches the target most closely. We then characterize experimentally the designs with circular dichroism to assess stability and with small angle X-ray scattering to compare the overall shape to the computational models.

Results and discussion: We designed and characterized modular structures with custom shapes, as well as new building blocks (from designed ankyrin repeat proteins) to enable incorporation of specific binders and assembly of multi-chain structures. We designed modules that embed growth factors within the architectures, allowing precise position of signals to trigger cell surface receptor clustering. The results indicate that modular repeat proteins represent a platform for the design of custom nanostructures with potential applications as tools to control spatial organization of cellular structures and cell behavior.

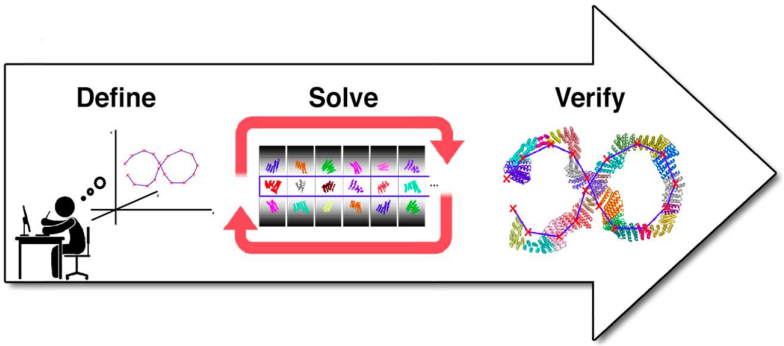


Figure 1. Modular protein design concept. Compatible structural building blocks are combined to satisfy user-defined requirements and assembled into the desired protein architecture.

References:

1. Brunette, Parmeggiani, Huang, Bhabha, Ekiert, Tsutakawa, Hura, Tainer, Baker. *Nature* **2015**, *528* (7583), 580.
2. Park, Shen, Parmeggiani, Huang, Stoddard, Baker. *Nat. Struct. Mol. Biol.* **2015**, *22* (2), 167.
3. Yeh, Brunette, Baker, McIntosh-Smith, Parmeggiani. *J. Struct. Biol.* 2017 S1047-8477(17)30141-7

O14

Computational Design of Alpha Helical Barrels

Dr Andrew Thomson
University of Glasgow, United Kingdom

α-Helical barrels (AHBs) are a subset of coiled-coil proteins with interesting properties. These structures are formed from five of more α-helices arranged like the staves of a barrel. A key feature of this protein fold is a central channel that suggests functions as pores, channels or receptors. These functions are observed in the very scarce examples of AHBs so far observed in nature. Because this protein fold is so scarce it is difficult to extract sequence-to-structure rules for the design of new examples. We have solved this problem by using the assumption that the interface between any two helices making up the barrel is locally approximate to that in a dimeric coiled coil. This, and the fact that AHBs can be strictly parameterized, allows us to make use of information and tools relating to the much more numerous dimers to design interfaces that 'build up' to form AHBs. We have used these tools, as well as modelling studies, to design many new potential examples of this fold. These sequences have been chemically synthesised and analysed through solution-phase biophysical techniques and X-ray crystallography. The result of this is a set of de novo AHBs in the pentamer to heptamer size range. These protein folds are novel, remarkably stable, and fit our design models to a high degree of accuracy. These new protein folds will be useful scaffolds for further functionalization.

O15

A census of protein tandem repeats and their relationship with intrinsic disorder

Anisimova, Maria[1,2], Spencer Bliven[1,2], Schaper, Elke[2,3], Oxana Sachenkova[4], Arne Elofsson[4]
[1]Zürich University of Applied Sciences, ZHAW Wädenswil, Switzerland, [2]Swiss institute of Bioinformatics, Lausanne, Switzerland, [3]Currently at: Carbon Delta AG, Switzerland, [4] Stockholm University, Stockholm, Sweden

Background: Protein tandem repeats (TRs) display an impressive variability of sizes and structures, contributing to diverse functions. Despite much interest, the commonly cited census of protein TRs dates to nearly two decades ago [1]. Further, proteins with TRs appear to be enriched with intrinsic protein disorder (IDP) and vice versa. While the relationship between these non-globular protein features has been observed, this has not been documented through a systematic analysis of TRs and IDP regions. Moreover, the biological reasons for a possible association are not well understood.

Questions addressed: Today the curated protein databank UniProtKB/Swiss-Prot has grown more than 7-fold compared to when the TR census was first conducted in 1999. In order to systematically characterize and explore the enigmatic connection between TRs with IDP, we annotate each protein with IDP regions and summarize the distribution of the overlap of TR and IDP regions over all kingdoms of life.

Methods: Till now over 50 different TR predictors were developed, but they exhibit striking differences in TR predictor properties. Thus, a new statistical framework and a meta-prediction approach was proposed that allows to increase the accuracy and power of the TR annotation [2,3]. It is based on both de novo and profile-based methods followed by filtering of false positives and redundancies. We used the TRAL library [4] that implements these methods to characterize the distribution of protein TRs in approximately 550'000 UniProtKB/Swiss-Prot entries, annotating the TR region start, end, minimal repeated unit length, among unit divergence and TR unit alignments. This information describes the universe of protein TRs in an unprecedented detail. IDP regions were annotated using MobiDB based on a consensus of several different predictors [5].

Results and discussion: An impressive number of TR regions were annotated across proteins of all domains of life. Overall, 59% of all UniProtKB/Swiss-Prot eukaryotic proteins contained at least one TR region. Interestingly, 52% of viral proteins contained TRs, almost as frequently as in eukaryotes. These estimates are much higher than previous estimates [1], presumably due to better annotation. By far the most frequent TRs were short repeats (i.e., unit length 4-15 amino acids). In general, differences in TR distributions observed between kingdoms can be largely attributed to sequence length, with eukaryotes having on average longer proteins. Domain TRs were typically centered around the middle of a protein, while shorter TRs displayed stronger preferences towards the flanks.
Most TRs appeared to be disordered, particularly for short TR units. TR and disorder overlap could be mostly explained by skewed amino acid frequencies. Functional annotations allow to pinpoint some interesting examples. But in general the results should not be overgeneralized since different biological processes may significantly contribute to TR origin, fixation and evolutionary mode. Therefore, we always observe exceptions from the general trend.

References:
1. Marcotte, E. M., Pellegrini, M., Yeates, T. O., Eisenberg, D. (1999). A census of protein repeats. Journal of Molecular Biology, 293(1), 151–160.
2. Schaper, E., Kajava, A.V., Hauser, A., Anisimova, M. (2012). Repeat or not repeat? – Statistical validation of tandem repeat prediction in genomic sequences. Nucl Acids Res 40: 10005-10017.
3. Anisimova, M., Pecerska, J ., Schaper, E. (2015). Statistical approaches to detecting and analyzing tandem repeats in genomic sequences. Frontiers Bioeng Biotech 3: 31.
4. Schaper, E., Korsunsky, A., Pečerska, J., Messina, A., Murri, R., Stockinger, H., Zoller, S., Xenarios, I., Anisimova, M. (2015). TRAL: Tandem repeat annotation library. Bioinformatics 31(18): 3051-3053.
5. Piovesan, D., Tabaro, F., Paladin, L., Necci, M., Mičetić, I., Camilloni, C., Davey, N., Dosztányi, Z., Meszaros, B., Monzon, A.M., Parisi, G., Schad, E., Sormanni, P., Tompa, P., Vendruscolo, M., Vranken, W.F., Tosatto S.C.E. (2017). MobiDB 3.0: more annotations for intrinsic disorder, conformational diversity and interactions in proteins. Nucleic acids research 46.D1: D471-D476.

O16

Hexads and nonads specify a new type of coiled-coil fiber

Hartmann, Marcus D., Hernandez Alvarez, Birte, Bassler, Jens, Lupas, Andrei N.
Department of Protein Evolution, Max Planck Institute for Developmental Biology, Tübingen, Germany

Background: Coiled coils are formed by α-helices winding around each other into superhelical bundles [1,2]. They are characterized by a specific geometry of interaction, called knobs-into-holes, in which core residues contributed by alternating segments of 3 and 4 residues mesh regularly along a seam that runs the length of the helices. This alternation yields a periodicity of 7 residues over 2 helical turns – the heptad repeat – but variant coiled coils can be formed by other periodicities resulting from the combination of 3- and 4-residue segments, such as decads (3+4+3), hendecads (3+4+4), or pentadecads (3+4+4+4), but not with consecutive segments of 3 residues (3+3 or 3+3+3).

Questions addressed: If coiled-coil periodicities are incompatible with successive 3-residue segments, how can these occur in natural coiled coils and what are their structural consequences?

Methods: We identified such segments in the coiled-coil stalks of trimeric autotransporters [3] and inn several other bacterial and archaeal protein families [4]. Modelling predicted that they would increase the supercoil strain sufficiently to break the helices in the bundle. We therefore studied them by X-ray crystallography in natural and designed coiled coils. Results and discussion: Structures of coiled coils containing hexads (3+3) and nonads (3+3+3) show that these segments indeed strain the supercoil to the breaking point, leading to the local formation of short β-strands [4]. These cross to form a triangular plane, which moves the path of each chain by 120˚counterclockwise around the trimer axis. Within this plane, the central residues of the three β-strands form backbone hydrogen bonds whose geometry deviates substantially from that seen in β-sheets. We have named them β-layers and show that they can be brought about in a straightforward way by the deletion of 1 or the insertion of 2 residues into a heptad background. The result is an α/β coiled coil, a substantially novel backbone structure within the allowed regions of the Ramachandran space that can be obtained with only minor changes to a known fold**.**

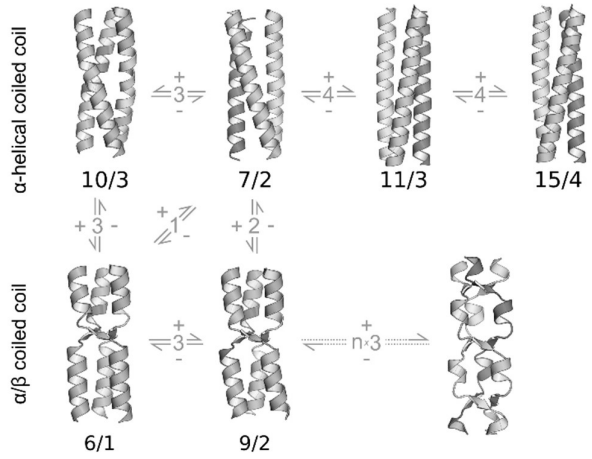


Figure 1. Transitions in Periodicity and Supercoiling Caused by the Insertion or Deletion of Residues in a Heptad Background. (Top) Insertion of a 3-residue segment causes overwinding and the formation of a local 3_{10}-helix. Insertion of 4 residues relaxes the supercoil and leads to essentially straight helices. Further insertion of 4 residues causes the transition to a righthanded supercoil. (Bottom) Insertions of 2 or 2x3 residues strain the helix beyond the breaking point, leading to the formation of β-layers. Inserted repeatedly along the fiber, these result in an α/β coiled coil.

References:
1. Lupas AN, Bassler J, Dunin-Horkawicz S. (2017). The Structure and Topology of α-Helical Coiled Coils. Subcell Biochem. 82:95-129.
2. Lupas AN, Bassler J. (2017). Coiled Coils - A Model System for the 21st Century. Trends Biochem Sci. 42:130-140.
3. Szczesny P, Lupas A. (2008). Domain annotation of trimeric autotransporter adhesins – daTAA. Bioinformatics 24:1251-6.
4. Hartmann MD, Mendler CT, Bassler J, Karamichali I, Ridderbusch O, Lupas AN, Hernandez Alvarez B. (2016). α/β coiled coils. eLife 5: e11861.

O17

Design of artificial antennas and electron transfer chains on repeated scaffolds

Cortajarena, Aitziber L.[1,2,3]
[1]*CIC biomaGUNE, 20014 Donostia-San Sebastián, Spain*
[2] *Ikerbasque, Basque Foundation for Science, Mª Díaz de Haro 3, 48013 Bilbao, Spain*
[3]*IMDEA Nanociencia and Nanobiotechnology Unit (CNB-CSIC), Campus de Cantoblanco, 28049 Madrid, Spain*

Background: Utilization of solar light energy is crucial for life on Earth, and is also increasingly important economically as a source of renewable energy. Given the efficiency of natural light harvesting antennas, inspiration for the design of artificial light harvesting materials has been increasingly drawn from biological systems. Yet, despite the intense research focus on biological light harvesting, the underlying mechanism that enables the efficient capture and transduction of light energy is still not fully understood.

Questions addressed: We aim to design a highly flexible and customisable model light harvesting system for the study of the fundamental principles of biological light harvesting, emphasising some of the features of biological systems: energy transduction along an energy gradient, precision of chromophore placement and system homogeneity. In addition, we aim to design redox centers along the scaffold. These hybrid arrays will enable to study the coupling between the antenna systems and a reaction center and the efficiency of the energy transfer process.

Methods: We have applied protein engineering methods for the design of the proteins scaffolds and combined with bioconjugation techniques to obtain the hybrid systems. Finally, we have applied different characterization methodologies to define the stoichiometry of the conjugates and their functional properties, including their photoconductivity and redox activity.

Results and discussion: On the one hand, the model system that we have generated is comprised of porphyrin chromophores immobilised on idealised modular protein nanofibres. The protein backbone is based on tetratricopeptide repeat (TPR) modules that provide a regular ordered surface for precise chromophore display. Individual TPR modules are functionalised with porphyrin chromophores and subsequently assembled into permanent supramolecular structures through unnatural amino acid-mediated cross-linking. The stepwise assembly of TPR modules ensures homogeneity of assemblies and allows the precise placement of differing chromophores, thus facilitating the establishment of energy gradients. The TPR repeats also provide a natural frame for a model oxidoreduction chain. We have created a TPR module that is design to coordinate [4Fe4S] clusters as redox centers. Using TPR proteins of different lengths we are able to easily coordinate a large number of clusters in ordered chains, mimicking the natural multicluster proteins in the electron transport chains. We have preliminary results on the integration of the systems in a complex multifunctional system by combining the antenna and the redox centers.

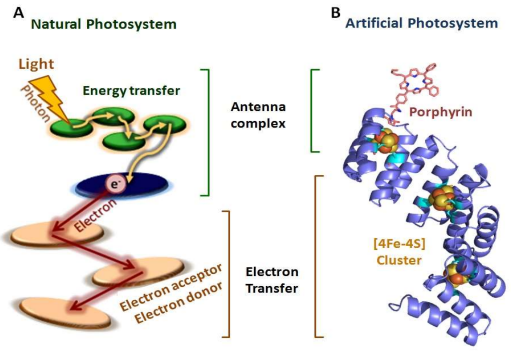


Figure 1. Schematic representation of a natural photosystem and an artificial simplified system.

References:
1. Mejías, S. H., et al., Chem. Sci. **2016**, 7, 4842.
2. López-Andarias, J., et al., Adv. Funct. Materials **2017**, DOI: 10.1002/adfm.201704031

O18

Property of intrinsically disordered "hub" proteins

Ota Motonori.
Nagoya University, Nagoya, Japan

Background: Protein-protein interactions are fundamental for all biological phenomena, and protein-protein interaction networks provide a global view of the interactions. The hub proteins, with many interaction partners, play vital roles in the networks. Many characters concerning to hub proteins have been investigated, but clear relationships with the subcellular localizations, an important property for protein function, were unknown.

Questions addressed: Intrinsically disordered proteins are considered to be hub proteins, and tend to be localized in nucleus. Does it mean hub proteins are likely to be nucleus proteins?

Methods: We investigated the subcellular localizations of proteins in the human network, using HPRD and Uniprot, with allowing multiple localization.

Results and discussion: We found that proteins localized in multiple subcellular compartments, especially the nucleus/cytoplasm proteins (NCP), the cytoplasm/cell membrane proteins (CMP), and the nucleus/cytoplasm/cell membrane proteins (NCMP), tend to be hubs. Examinations of keywords suggested that among NCP, those related to post-translational modifications (PTM) and transcription functions are the major contributors to the large number of interactions. These types of proteins are characterized by a multi-domain architecture and intrinsic disorder. A survey of the typical hub proteins with prominent numbers of interaction partners in the type revealed that most are either transcription factors or co-regulators involved in signaling pathways. They translocate from the cytoplasm to the nucleus, triggered by the phosphorylation and/or ubiquitination of intrinsically disordered regions. Schematically speaking, our analysis indicates (1) proteins localized only in nucleus are not hub proteins; (2) proteins localized in nucleus and cytoplasm (NCP) are hub proteins; (3) among them, the ones related to PTM and transcriptions are intrinsically disordered. We suggest that multiple-localization is a crucial concept to characterize groups of hub proteins and their biological functions in cellular information processing.

References:

1. M. Ota, H. Gonja, R. Koike, S. Fukuchi, Multiple-localization and hub proteins, *PLoS ONE* **11** (2016) e0156455

O19

Intrinsic Disorder and Myelin Protein Complexes

Raasakka, Arne[1], Myllykoski, Matti[2], Ruskamo, Salla[2], Kursula, Petri[1,2]
[1]Department of Biomedicine, University of Bergen, Norway, [2]Faculty of Biochemistry & Molecular Medicine, University of Oulu, Finland

Background: Myelin is a crucial structure for ensuring rapid nerve impulse transmission in vertebrates. The myelin sheath is a differentiated plasma membrane of a glial cell, wrapped dozens of times around a neuronal axon, and its biochemical composition is unique. The compact multilayered myelin proteolipid membrane contains specific proteins and very little aqueous solvent. Myelin proteins present large degrees of intrinsic disorder, and they have in general little homology to other proteins [1]. However, myelin proteins specifically interact with binding partners, including membranes and other proteins, and the disordered domains of these proteins play central roles in the interactions. Understanding these processes is crucial to understand the process of myelination as well as demyelinating neurodegenerative diseases, such as multiple sclerosis.

Questions addressed: Our work focuses on the fine details of myelin protein structure, their interactions with lipid membranes, and the mechanisms, through which they induce membrane multilayer formation. The research questions include both conformation and dynamics of flexible myelin proteins, their folding upon lipid membrane binding, as well as specific protein-protein complex formation involving myelin protein disordered segments.

Methods: We use a variety of biophysical, spectroscopic, computational, and structural biology techniques to study intrinsic disorder in myelin proteins. The main techniques involve the use of synchrotron radiation, including both far-UV and X-rays, and neutrons.

Results and discussion: We will present recent published and unpublished data on myelin proteins, their disorder, and their specific molecular interactions with both lipid membranes as well as other proteins. For example, the membrane binding and stacking properties of myelin basic protein [2] and P0 are presented. The identification of two novel protein-protein complexes involving binding of a disordered myelin protein domain to a folded partner is also discussed.
Overall, our results indicate the variety of interactions and conformations that disordered proteins in myelin present. Despite no genetic similarity between the myelin proteins, often, they show surprisingly similar biophysical and structural properties.

References:

1. Han H., Myllykoski M., Ruskamo S., Wang C. & Kursula P. (2013) Myelin-specific proteins: a structurally diverse group of membrane-binding molecules. *BioFactors* 39: 233-241.
2. Raasakka A., Ruskamo S., Kowal J., Barker R., Baumann A., Martel A., Tuusa J., Myllykoski M., Bürck J., Ulrich A.S., Stahlberg H. & Kursula P. (2017) Membrane association landscape of myelin basic protein portrays formation of the myelin major dense line. *Sci. Rep.* 7:4974.

O20

Towards biologically meaningful ensembles in disordered proteins

Julia Marchetti[1], Alexander Miguel Monzon[1], María Silvina Fornasari[1], Silvio Tosatto[2] and Gustavo Parisi[1]
[1]Structural Bioinformatics Group. Universidad Nacional de Quilmes. Argentina, [2]Biocomputing UP. University of Padova. Italy.

Background: Ensembles represent the native state of proteins as a collection of different conformational states. Due to their structural differences, conformers show different biological activities constrained by specific affinity constants, kinetical parameters, stability and so on. Since the early crystallization of hemoglobin in their two conformational states, it has been well established that all the conformers in the ensemble are required to maintain biological function. However, as protein flexibility increases, the number of conformers in the ensemble increases as well, giving rise to very complex ensembles as in the case of intrinsically disordered proteins (IDPs). For IDPs, it is still difficult to obtain a statistical description of the relative abundance of each conformer in the ensemble to assess their specific biological contribution [2]. However, in a previous work, we were able to describe the specific conformer contribution to the sequence conservation pattern of each particular protein. Structural differences in each conformer constrain sequence divergence in a specific way, making possible to detect individual contributions of each conformer in a given ensemble using a evolutionary analysis[3].

Questions addressed: In this work we are interested in a biological description of the ensembles studying the specific conformational contributions on sequence divergence more than use a statistical population description of the ensemble. Are all the conformers equally contributing to the sequence divergence pattern? Does a particular conformer contribute more than the rest?

Methods: The studied dataset consisted in 76 proteins with known NMR structure with more than 50% of disordered positions estimated with ESpritz [3]. We estimated the constrained substitution pattern using SCPE model [4]. This model generates a whole set of site-specific substitution matrices that take into account the constraints imposed by the structure on the sequence divergence. Structurally constrained sites (SCS) were estimated using a maximum likelihood analysis with HyPhy [5]. Basically, for each protein, we compared the performance of SCPE matrices with those evolutionary models lacking structural information (such as JTT, Dayhoff and WAG) to explain the substitution pattern observed in a set of homologous proteins. Akaike information criteria was used for statistical purposes to assess the sites where SCPE outperforms JTT, Dayhoff and WAG models (SC sites).

Results and discussion: Structurally constrained sites (SCS) are positions evolving under structural constraints, mainly inter-residue contacts. Tanking into account the complex nature of IDP ensembles and the extreme conformational diversity they have, ~100% of the positions per protein have at least one inter-residue contact in at least one conformer. According with this result the great majority of the positions would evolve under structural constraints since most of them (~100%) have inter-residue contacts. However, we have found that in average the maximum number of SCS per conformer is ~35%, meaning that most of the inter-residue contacts in the ensemble could be artifacts or just not biologically important. Accumulating the SCSs for all the conformers in a given ensemble increases only to ~40% the total number SCS per protein. These results indicate that in average conformers share most of the SCS, differing among them at most ~5%. The maximum and minimum average number of SCS per protein are ~35 and ~27% showing a narrow scope of SCSs along the ensemble. This result suggests that IDP ensembles of the observed protein are mostly “flat” showing no bias towards the existence of conformers particularly constrained or particularly relaxed. Our results suggest that the use of evolutionary information could be useful to characterized conformational distributions in IDPs towards a biologically meaningful comprehension of ensembles.

References:

1. Sormanni P, Piovesan D, Heller GT, Bonomi M, Kukic P, Camilloni C, et al. Simultaneous quantification of protein order and disorder. Nat Chem Biol. 2017;13: 339–342.
2. Juritz E, Palopoli N, Fornasari MS, Fernandez-Alberti S, Parisi G. Protein conformational diversity modulates sequence divergence. Mol Biol Evol. 2013;30: 79–87.
3. Walsh I, Martin AJM, Di Domenico T, Tosatto SCE. ESpritz: accurate and fast prediction of protein disorder. Bioinformatics. 2011;28: 503–509.
4. Parisi G, Echave J. Generality of the Structurally Constrained Protein Evolution model: assessment on representatives of the four main fold classes. Gene. 2005;345: 45–53.

O21

Specific modulation of the anticoagulant activity of madanin-1 by tyrosine sulfation

Ripoll-Rozada, Jorge[1], Payne, Richard J.[2], Pereira, Pedro J. B.[1]

[1]IBMC - Instituto de Biologia Molecular e Celular and Instituto de Investigação e Inovação em Saúde, 4200-135 Porto, Portuga, I [2]School of Chemistry, The University of Sydney, Sydney, NSW 2006, Australia

Background: Blood-feeding animals rely on a complex arsenal of bioactive molecules to counter the defense mechanisms of their hosts, in particular the blood clotting cascade for which they produce highly specific inhibitors. Many of these anticoagulants target thrombin, a central proteinase of the coagulation cascade, and over the last decades the molecular mechanisms of inhibition of several of these molecules have been unveiled. An important but elusive class of natural thrombin inhibitors is composed by a diverse group of small cysteine-free proteins, which are devoid of discrete structure in solution, yet can recognize and inhibit thrombin with great specificity and affinity. Madanin-1 is one of such small cysteine-free thrombin inhibitors that facilitates blood feeding in the tick *Haemaphysalis longicornis*.

Questions addressed: The aim of this work was elucidating the role of a specific post-translational modification – tyrosine sulfation – identified in recombinant madanin-1 produced in baculovirus-infected insect cells in the modulation of the anticoagulant activity of this molecule.

Methods: Homogeneously sulfated variants of madanin-1 were produced using one-pot ligation-desulfurization chemistry. The anticoagulant activity of the different variants was assessed *in vitro*. The three-dimensional structure of the thrombin:madanin-1 complex was determined using X-ray crystallography.

Results and discussion: We unveiled the modulatory effect of tyrosine sulfation on the inhibitory activity of madanin-1, which amounts to three orders of magnitude [1]. We also uncovered its mechanism of thrombin recognition, revealing that madanin-1 is an exosite II-targeting anticoagulant [1]. The importance of tyrosine sulfation within this family of thrombin inhibitors, together with the now uncovered unique binding mode, sets an experimental foundation for the development of anticoagulant drug leads based on this structural scaffold.

This work was funded 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 form of project "Institute for Research and Innovation in Health Sciences" (POCI-01-0145-FEDER-007274) and of postdoctoral fellowship SFRH/BPD/108004/2015 (to J. R.-R.).

References:

1. Thompson, R.E.; Liu, X.; Ripoll-Rozada, J.; Alonso-García, N.; Parker, B.L.; Pereira, P.J.B.; Payne, R.J. (2017) Tyrosine sulfation modulates activity of tick-derived thrombin inhibitors. *Nature Chemistry* **9**, 909-917.

O22

Tales of tails: the role of disordered protein kinase regions in the regulation and assembly of kinase domain cores

Reményi, Attila
Institute of Enzymology, Research Center for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

Background: Since publication of the crystal structure of PKA nearly three decades ago, a structural portrait of the conserved kinase core has been filled in. The next challenge is to elucidate structures of full-length kinases and to address the intrinsically disordered regions (IDRs) that typically flank the core as well as the Small Linear Motifs (SLiMs) that embedded in the IDRs. It is increasingly apparent that unstructured regions integrate the kinase catalytic chassis into multi-enzyme based regulatory networks, which allow protein kinase cascades to operate as complex switches that contribute to non-linear signaling network properties.

Questions addressed: We have been studying the role of disordered kinase regions facilitating the formation of specific protein kinase-substrate pairs within the mitogen activated protein kinase (MAPK) signaling pathways [1]. In particular, we are exploring how the N- and C-terminal extensions to the globular kinase core, which is shared basically by all eukaryotic kinases, enable the common catalytic chassis to play specific roles in different signaling cascades [2].

Methods: The ERK-RSK-PDK cell growth promoting multi-enzyme complex will be used as an example to demonstrate how IDRs and SLiMs govern communication between four different kinase catalytic cores to mediate activation and how in molecular terms these promote the formation of kinase heterodimers in a context dependent fashion**.** We use structural biology techniques (X-ray, SAXS, MD) as well as cell-based assays to monitor the assembly and signaling activity of various MAPK based signaling complexes.

Results and discussion: Phosphorylation-based regulation is possibly the most widespread regulatory mechanism in cellular signaling. Protein kinases also regulate each other this way. Kinase cascades are an intrinsic feature of these signaling networks. This regulation apparently is not only determined by the common structural elements of the conserved kinase domain chassis, but for many kinases it also involves evolutionarily more divergent IDRs. Interestingly, IDR-mediated regulation mostly orchestrate the protein-protein interactions of the kinase cores. Due to lack of insights into full-length kinase structures, as most mechanistic work on kinases have focused on their structured kinase cores, the role of IDRs is understudied and thus not well-appreciated. Fortunately, for some kinase families with relatively short IDRs (such as p90RSK), it is becoming apparent that the full functionality of a given protein kinase core can only be understood in the context of the full-length kinase and in its complexes, in particular, with other proteins.

References:

1. Zeke A, Bastys T, Alexa A, Garai A, Meszaros B, Kirsch K, Dosztanyi Z, Kalinina OV, Reményi A (2015) Systematic discovery of linear binding motifs targeting an ancient protein interaction surface on MAP kinases. MOLECULAR SYSTEMS BIOLOGY 11, 837

2. Alexa A, Gogl G, Glatz G, Garai A, Zeke A, Varga J, Dudas E, Jeszenoi N, Bodor A, Hetenyi C, Reményi A (2015) Structural assembly of the signaling competent ERK2-RSK1 heterodimeric protein kinase complex. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 112, 2711-2716

O23

The Power of Reuse in the Evolution of natural Proteins

Weidmann, Laura[1], Dijkstra, Tjeerd[2], Kohlbacher, Oliver[2], Lupas, Andrei[1]
[1]*Department of Protein Evolution, Max Planck Institute for Developmental Biology, Tübingen, Germany,*
[2]*Applied Bioinformatics Group, Department of Computer Science, University of Tübingen, Tübingen, Germany*

Background: The probability of a foldable and functional protein sequence to emerge *de novo* is extremely small. The evolution of natural proteins therefore often proceeds through the amplification of already existing sequences or their integration into different genomic contexts. Copies of the same protein sequence will diversify over time, leading to the co-existence of similar sequences in nature.

Questions addressed: This scenario is not the only explanation for the presence of similar sequences in present-day proteomes: Sequence similarities can arise through common descent, convergence and random chance. We address the individual contributions of these phenomena by analyzing a non-redundant set of bacterial genomes from a statistical point of view.

Methods: In an all-against-all approach, we separate randomly expected from unexpected similarities between natural sequence fragments of the same length. We therefore use an unbiased definition of sequence similarity, simply the position-specific sequence identity, without allowing for insertions, deletions or a substitution matrix. Sampling the pairwise similarities between fragments of equal length, reveals the frequencies of certain similarities. These frequencies are then compared to an expected value derived form a null model, which is biased by the underlying natural amino acid composition. We model the expected values using a binomial distribution that estimates the frequency of fragment pairs with a certain similarity in random sequence data. The ratio between natural and randomly expected frequencies results into the over-representation of natural recurrence, the difference in the possible amount of naturally driven reuse.

Results and discussion: The majority of natural similarities can be described by the amino acid biased null model. This is reflecting the fact, that fragments are in general similar to a small subset of other fragments but not to the majority of all natural fragments. Hence, most natural fragments are scattered randomly across sequence space with additional local agglomerations in regions, where sequences are reused in nature.
We are able to capture the magnitude of these local agglomerations by a trend line of similarity frequencies. It indicates the extent of over-represented highly similar fragments in nature and thereby accounts for homology. This confirms that reuse of existing protein sequences is a major mechanism in protein evolution.
When subtracting these presumably homologous similarities from the overall natural similarities, there is still a significant difference to the null model. Especially in the 20-40% sequence identity region, an increase of natural distances can be observed. We assume this increase to be caused by convergence and are currently investigating this hypothesis.

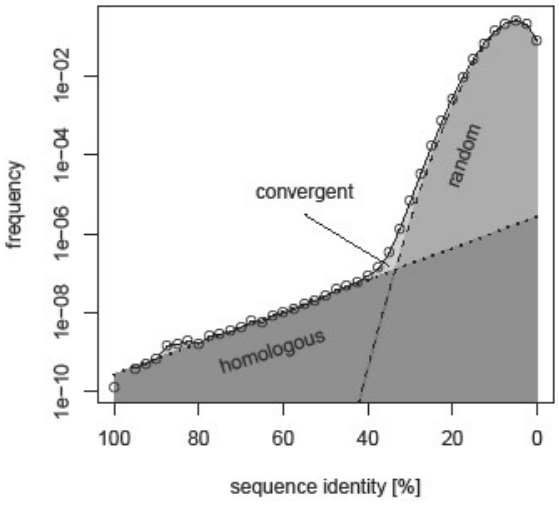


Figure 1. Similarity frequencies between natural protein sequences comprised of 40 amino acids. Sequence similarities can arise through common descent, convergence and random chance. The majority of pairwise fragment similarities can be described by random chance using an amino acid biased null model (dashed). Together with a strong signal of pairwise similar fragments caused by homology (dotted) the natural distribution of similarities (circles) can be reasonably well approximated. Still, there is a significant amount of pairwise similarities that cannot be described by either of these mechanisms. They are possibly a result of sequence convergence.

O24

Rational design of antibodies targeting specific epitopes within disordered proteins: applications in the study of neurodegeneration.

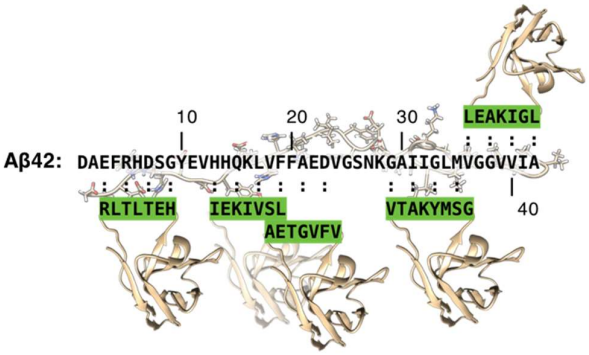
Pietro Sormanni
Centre for Misfolding Diseases, University of Cambridge

Background: immune responses against IDPs have been reported to have specific characteristics distinct from those against structured proteins, and importantly may be more restricted towards fewer immunodominant epitopes within disordered antigens. However, different regions of IDPs typically mediate different interactions regulating different processes, including pathological ones such as neurotoxic aggregation[1].

Questions addressed: It is therefore important to be able to obtain antibodies targeting specific epitopes within disordered regions, ideally with single-residue resolution.

Methods: To address this requirement, we developed the Modular design strategy that combines different methods[2,3,4] to enable the design of novel antibodies targeting linear epitopes of interest, while controlling for stability and solubility.

Results and discussion: all eleven antibodies tested so far, which target different epitopes within three unrelated antigens, were stable and active[2,5]. The possibility of quickly obtaining antibodies with single-residue resolution in the choice of epitope has facilitated novel types of experiments. For example, we designed a panel of antibodies targeting different epitopes systematically covering the sequence of the amyloid-β peptide, whose aggregation is a hallmark of Alzheimer's disease. All the designed antibodies (DesAbs) were effective in inhibiting the Aβ-aggregation process *in vitro*, which they affected in completely distinct manners fully compatible with the corresponding suppression of toxicity *in vivo* in a *C. elegans* model of Aβ-overexpression. Taken together, our results demonstrate that computational approaches are becoming sufficiently mature to be highly competitive for some applications, thus offering novel opportunities to streamline antibody development and accelerate discoveries in biomedical sciences.



References:
1.Sormanni, P. *et al.* Simultaneous quantification of protein order and disorder. *Nat. Chem. Biol.* **13,** 339–342 (2017).
2.Sormanni, P., Aprile, F. A. & Vendruscolo, M. Rational design of antibodies targeting specific epitopes within intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 9902–9907 (2015).
3.Sormanni, P., Aprile, F. A. & Vendruscolo, M. The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **427,** 478–490 (2015).
4.Sormanni, P., Amery, L., Ekizoglou, S., Vendruscolo, M. & Popovic, B. Rapid and accurate in silico solubility screening of a monoclonal antibody library. *Sci Rep* **7,** 8200 (2017).

O25

Looking at disorder and disorder-to-order transitions using Hydrophobic Cluster Analysis

Bitard-Feildel, Tristan[1], Lamiable, Alexis[2], Mornon, Jean-Paul[2], Callebaut, Isabelle[2].

[1] Sorbonne Université, CNRS, IBPS, Laboratoire de Biologie Computationnelle et Quantitative (LCQB), 75005 Paris, France, [2] Sorbonne Université, Muséum National d'Histoire Naturelle, UMR CNRS 7590, IRD, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, 75005 Paris, France

Background: Hydrophobic Cluster Analysis (HCA) is an original approach for protein sequence analysis, allowing to get information about regular secondary structures from the only information of a single amino acid sequence [1,2]. As regular secondary structures are more conserved than sequences, HCA has been widely used for highlighting remote relationships between sequences of well-folded domains [e.g. 3]. Further developments have also been made to get insight into the whole foldable repertoire of the protein universe, which includes a large number of unannotated sequences and also comprises sequences undergoing disorder to order transitions [3,4].

Questions addressed: SEG-HCA is the tool allowing to delineate foldable regions within protein sequences. It is based on the detection of segments with high density in hydrophobic clusters, which mainly correspond to regular secondary structures. We examined the small segments of known 3D structures delineated using SEG-HCA within the UniProt database, in order to evaluate to what extent these correspond to regions undergoing disorder to order transitions and what kind of information could be deduced on these segments using HCA.

Methods: A search of the UniProt database was performed using SEG-HCA for identifying, short foldable segments of lengths ≤ 30 amino acids. We then further analyzed segments matching sequences of known 3D structures, as included in the Protein Data Base (PDB).

Results and discussion: A shown with the example reported in Figure 1, short foldable segments often contain small linear motifs, which become ordered upon contact with partners and typically match hydrophobic clusters, whose secondary structure affinities generally correspond to those observed in the 3D structures within the complexes. From several examples, we suggest that the limits of such small foldable segments can be considered to clarify the boundaries of the functional IDRs, including flanking segments critical for interaction specificity or regulation. We also present cases of larger conditionally disordered domains, with lower density in hydrophobic clusters than well-folded globular domains or with exposed hydrophobic patches, which are stabilized by interaction with partners. HCA thus provide a qualitative analysis of disorder, which can be explored for a better understanding of the dark proteome.

O26

Exploring the Sequence-based Prediction of Folding Initiation Sites in Proteins

Vranken, Wim[1,2], Daniele Raimondi[1,2,3], Gabriele Orlando[1,2,3], Taushif Khan[1,2].
[1]Interuniversity Institute of Bioinformatics in Brussels, ULB/VUB, Triomflaan, BC building, 6th floor, CP 263, 1050 Brussels, Belgium; [2]Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium; [3]Machine Learning Group, Université Libre de Bruxelles, Boulevard du Triomphe, CP 212, 1050 Brussels, Belgium

Background: The very early stages of protein folding, defined by intrinsic local interactions between amino acids close to each other in the protein sequence, are poorly understood. We created the Start2Fold database that identifies these residues from pulsed labelling and related Hydrogen Deuterium eXchange (HDX) experiments[1]. Based on an additional analysis in relation to protein backbone rigidity predictions[2], we developed EFoldMine[3], a predictor of early folding residues based on data from high-quality experimental NMR-based HDX studies. EFoldMine predicts from the primary sequence of a protein where early folding is likely to take place, and show that these early folding predictions give insights into the folding process. The predicted early folding residues are correlated with aggregation sites, and become the residues that are more surrounded in the final fold. These connections suggest that amino acids close to each other in the sequence that interact favorably help shape the folding landscape, and that this has a lasting effect on the subsequent states of the protein.

Questions addressed: Where do protein sequences start to fold, and how is this related to the subsequent behavior of, and states adapted by, the protein?

Methods: From 30 Start2Fold sequences totaling 3398 residues[2], 482 were designated 'early folding', as residues experimentally determined to be the first to form sufficient local structure to protect their backbone HN from solvent. We used 5 features computed from the protein sequence: DynaMine backbone dynamics[4,5], and a new set of predictions for side-chain dynamics and secondary structure propensity. We used a window of flanking residues between i-2 and i+2, resulting in a 25-dimensional feature vector. Performances were evaluated in strict cross-validation settings, using BLASTCLUST to stratify the 30 sequences in function of a 25% sequence identity (SI) cutoff at 90% coverage.

Results and discussion: The EFoldMine performance reaches an MCC of 35.4, and an AUC of 80.8. The EFoldMine predictions for two protein pairs that have a very similar topology but different folding pathways show very different early folding profiles between the pairs of proteins. The predictions also relate very well to independent HDX-MS experiments. On a proteome scale, they consistently encompass many of the residues that form the most extensive contacts in the folded protein, and not only for the typical (hydrophobic) structure-forming residues. We also observe that residues with evolutionary covariance signals tend to be early folding residues, and that there is a link with residues involved in aggregation.

EFoldMine identifies the amino acid residues in proteins that are inclined to form structural elements unaided at the very first stage of the folding process. The connection of the early folding predictions to both folding pathway data and the folded protein structure suggests that the initial statistical behaviour of the protein chain has a lasting effect on its subsequent states.

References:

1. Pancsa, R., Varadi, M., Tompa, P. & Vranken, W. F. *Nucleic Acids Res.* **44,** D429–D434 (2016).
2. Pancsa, R., Raimondi, D., Cilia, E. & Vranken, W. F. *Biophys J* **110,** 572–583 (2016).
3. Raimondi, D., Orlando, G., Pancsa, R., Khan, T. & Vranken, W.F. *Scientific Reports* **7**, 8826 (2017).
4. Cilia, E., Pancsa, R., Tompa, P., Lenaerts, T. & Vranken, W. F. *Nat Commun* **4,** 2741 (2013).
5. Cilia, E., Pancsa, R., Tompa, P., Lenaerts, T. & Vranken, W. F. *Nucleic Acids Res.* **42,** W264–W270 (2014).

O27

ClusterLCR: a new tool to compare low complexity regions in unlike proteins

Jarnot Patryk[1], Ziemska Joanna[2], Grynberg Marcin[3], Gruca Aleksandra[1]
[1]Institute of Informatics, Silesian University of Technology, Poland , [2]Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland , [3]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Poland

Background: Low complexity regions (LCRs) are common among proteins. About 14% of proteins contain LCRs [1] and some articles prove they play crucial roles, e.g. in binding proteins, DNA or RNA [2][3]. One can find a plethora of different methods for searching for similar LCRs of proteins, however there is a serious scarcity of programmes to compare LCRs [4][5]. In order to overcome this problem, we propose a new method designed specifically for searching for similar repeats of low complexity regions in the diverse world of protein sequences.

Methods: We present the tool for searching for similar LCRs. This tool is a combination of three methods. Each of them uses a different approach to search for similarities of LCR. The first one is the Graph Based Sequence Clustering (GBSC) which uses graph theorem for clustering similar LCRs. The second method is the MotifLCR which uses patterns and PSSMs in order to cluster sequences. As the third method we modified BLAST in order to deal with LCRs. ClusterLCR contains a database with pre-calculated results which are ready to query by a user. We use the SEG programme in order to retrieve low complexity regions from protein sequences [5]. This database is programmed for automatic updates. In this study we compare methods by execution time.

Results and discussion: ClusterLCR is a user friendly web application that allows users to search for similar sequences using low complexity fragments or full sequences in the fasta format. Users can also search for repeats in sequences. In case of repeats both GBSC or the MotifLCR are suitable. Modified version of the BLAST method is able to search for similarities using common parts of sequences only. The tool can be downloaded to a local machine and run on a custom database of LCRs.

References:

1. Marcotte E., Pellegrini M., Yeates T., Eisenberg D. A census of protein repeats. J Mol Biol. 1999;293:151–160.
2. O'Rourke T.W., Loya T.J., Head P.E., Horton J.R., Reines D. Amyloid-like assembly of the low complexity domain of yeast Nab3. Prion. 2015;9(1):34-47.
3. Kumari B., Kumar R., Kumar M. Low complexity and disordered regions of proteins have different structural and amino acid preferences. Mol Biosyst. 2015 Feb;11(2):585-94.
4. Wise MJ. 0j.py: a software tool for low complexity proteins and protein domains. Bioinformatics. 2001;17 Suppl 1:S288-95.
5. Wootton JC, Federhen S. Analysis of compositionally biased regions in sequence databases. Methods Enzymol. 1996;266:554-71.

O28

Phase Separation in Biology and Disease

Bappaditya Chandra[1], Mylene Ferrolino[1], Eric Gibbs[1], Diana M. Mitrea[1], Cheon-Gil Park[1], Aaron Phillips[1], Ramiz Somjee[1,2], Michele Tolbert[1], Michael R. White[1], Christopher Stanley[3], Maxime Cuypers[1] and Richard Kriwacki[1,2]

[1]Department of Structural Biology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105 USA; [2]SummerPlus Fellowship Program, Rhodes College, Memphis, TN, USA; and [3]Biology and Biomedical Sciences Group, Biology and Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge TN, USA.

Phase separation mediates formation of various membraneless organelles, as well as assembly of other cellular structures that control diverse biological processes, including membrane receptor signaling, endocytosis, stress sensing, and both activation and silencing of gene transcription. Phase separation organizes chemically diverse biopolymers, including proteins, nucleic acids, and lipids, to perform complex biochemical and mechanical processes. Consequently, alteration of the process of phase separation, yielding aberrant molecular organization and condensed phases with atypical material properties, is associated with a range of human diseases such as neurodegeneration and cancer. For example, the properties of the nucleolus, the largest membraneless organelle in human cells and a center for ribosome biogenesis and stress signaling, are altered through different mechanisms in ALS and certain cancers. Further, regulation of gene expression, hypothesized to be controlled through condensation of genes, transcription factors and RNA polymerase II, may be altered by phase separation-prone oncoproteins in certain leukemias. We will discuss our efforts to understand how phase separation contributes to biological processes in normal cells and how these processes are altered in disease, incuding neurodegeneration and cancer.

O29

Mechanistic Aspects of IDP Binding and Signaling Studied by NMR

Chill, Jordan[1], Sokolik, Chana[1], Halle-Bikovski, Adi[1], Lepechkin-Zilbermintz, Veronica[1], Baluom, Saja[1], Qasim, Nasrin[1], Shaked, Hadassa[1]

[1]Department of Chemistry, Bar Ilan University, Ramat Gan 52900, ISRAEL

Background: The plasticity of intrinsically disordered proteins (IDPs) is a major factor in cellular signaling and regulation [1]. A case in point is the fully disordered WASp-interacting protein (WIP), a 503-residue polypeptide ubiquitously found in various cellular settings, known to chaperone Wiskott-Aldrich syndrome protein (WASp), and bind actin and additional key players in cytoskeleton rearrangement [2]. Since WIP-mediated cellular events are central to homeostasis, malignancy and cell signaling, understanding the molecular mechanisms underlying WIP activity is of great interest.

Questions addressed: IDP behavior can be explained by an ensemble of rapidly interchanging conformations that determine structural characteristics. Thus, it is challenging to offer a structural context for the observed interactions of WIP with its binding partners. We are interested in (1) How do binding partners influence the conformational ensembles of WIP? (2) Is there a fundamental difference between a 'pre-formed' polyproline epitope in one case and an 'induced' amphipathic helix - do we observe folding-then-binding or binding-then-folding, and can this be generalized? (3) How does phosphorylation affect conformational ensembles, and does this hint at their biological role?

.

Methods: NMR is a powerful method for studying IDPs, offering a local residue-by-residue view of transient structure and backbone motions. Chemical shifts and couplings between spins report on backbone dihedrals, and dipolar/paramagnetic interactions follow long-range compaction of structure [3]. Backed by an ensemble-based computational analysis, conformational distributions can thus be simulated and trajectories predicted.

Results and discussion:
Three WIP interaction fragments were determined and investigated, N-terminal, central, and C-terminal domains binding to actin, cortactin and the WASp-EVH1 domain, respectively. The first two are weak (micromolar) complexes, whereas the third forms a tight interaction spanning four binding epitopes (Figure 1)[4]. The free and bound average conformations in the weak complexes were determined and compared by following backbone dynamics, affording an important contrast between binding modes. The structure of the tight WIP/WASp complex was determined by NMR and the relative importance of the binding epitopes established. Phosphorylation-mediated dissociation of this latter interaction was investigated by phospho-mimicking mutations introduced in the WIP sequence. These suggest an unexpected phospho-switching of WASp-shielding from degradation by WIP. Overall NMR demonstrates its ability to follow the implications of IDP conformational flexibility upon affinity to globular binding partners, affording an all-important mechanistic understanding of protein-protein networks in cellular signaling.

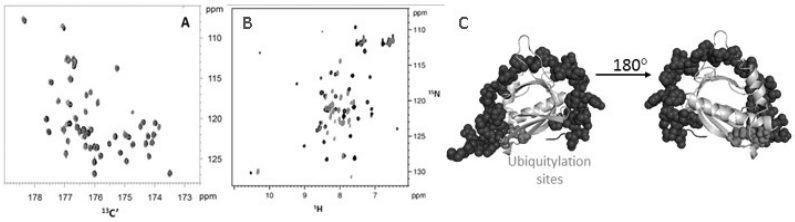


Figure 1. Binding of **(A)** actin to WIP(2-65), **(B)** WASp to WIP(442-492) as shown by changes in their $^{15}N,^{1}H$-HSQC spectra. Black and grey represent free and bound WIP, respectively. **(C)** Structure of WIP(442-492) (black) wrapped around the WASp EVH1 β-sandwich domain (grey). Dark grey spheres indicate the putative ubiquitylation sites.

References:
1. Uversky, V.N. (2018) *Adv. Prot. Chem. Struct. Biol.* **110**, 85-121.
2. Matalon, O., Reicher, B. & Barda-Saad, M. (2013). *Immunol Rev* **256**, 10-29.
3. Nasrollah, R.-G., Blackledge, M. & Zweckstetter, M. (2012). *ChemBioChem* **13**, 930-950.
4. Halle-Bikovski, A., Fried, S., et al, Barda-Saad, M., Chill, J.H. (2018.

O30

Structure-function relationship in disordered proteins: knowledge acquired by NMR studies of microtubule associated proteins

Zidek, Lukas[1,2], Melkova, Katerina[1,2], Zapletal, Vojtech[1,2], Jansen, Severine[1,2], Hritz, Jozef[1], Jensen, Malene R.[3], Blackledge, Martin[3]

[1]Masaryk University, Central European Institute of Technology, Kamenice 5, 62500 Brno, Czech Republic, [2]Masaryk University, Faculty of Science, National Centre for Biomolecular Research, Kamenice 5, 62500 Brno, Czech Republic, [3]University Grenoble Alps, CEA, CNRS, Institut de Biologie Structurale, Grenoble 38000, France

Background: On one hand, the structure-function relationship is the cornerstone of structural biology. On the other hand, its understanding is a considerable challenge in the case of intrinsically disordered proteins (IDPs), present in multiple conformations. Description of conformational ensembles is not trivial. Although characteristic features of IDP ensembles are encoded in the IDP's sequence and various computational tools translating sequence to structure-related information are available, structural analysis of experimental data remains important. Nuclear magnetic resonance (NMR) is often used to characterize IDPs at atomic resolution. Among methods of NMR data analysis, ASTEROIDS [1,2] represents an approach allowing us to describe IDPs using the *lingua franca* of structural biology, i.e. structural ensembles with specified atomic coordinates. Ensembles selected by ASTEROIDS do not contain all conformers present in real samples, but contain correct populations of all local structural features needed to reproduce the ensemble-averaged experimental data.

Questions addressed: We used analysis of ensemble-averaged NMR parameters to reveal relation between structural features (degree of disorder, populations of specific conformations, intramolecular contacts) and known or proposed function in regions of intrinsically disordered microtubule associated protein 2c (MAP2c), and compared the results with a related protein Tau.

Methods: NMR parameters (chemical shifts, relaxation rates, paramagnetic relaxation enhancement, residual dipolar couplings) were obtained using multidimensional non-uniformly sampled experiments in order to achieve resolution required for the disordered 49kDa protein of interest [3,4]. NMR was also used to monitor MAP2c functions (protein binding and phosphorylation). Relaxation and conformational analyses were based on spectral density mapping and ASTEROIDS, respectively.

Results and discussion: The relaxation rates revealed large differences in flexibility of individual regions of MAP2c. The known and proposed binding and phosphorylation regions [4-6] were significantly more ordered. Specific intramolecular contacts were observed in the more ordered regions. Quantitative analysis of NMR parameters identified several well defined structural motifs. The major advantage of the ensemble-based approach is that population of any structural motif can be evaluated (the analysis is not limited to general trends such as secondary structure propensity). Reliability of the results has been tested by cross-validation (experimental data not used in the analysis were compared with their prediction from the selected ensembles). Comparison with protein Tau [2,7,8] showed that long-range interactions, dynamics, and local conformation motifs in the N-terminal domains of MAP2c and Tau differ significantly. In summary, the results revealed (1) importance of the N-terminal regions for the specificity of regulatory roles of MAP2c and Tau and (2) a close relation between biological functions and conformational behavior of these proteins, as an example of structure-function relationship observable in IDPs. The obtained results provided a new insight into behavior of proteins regulating microtubule dynamics in neurons and may serve as a benchmark for software tools predicting structural features of disordered proteins.

References:

1. Nodet G. et al., *J. Am. Chem. Soc.* **131** (2009), 17908-17918.
2. Schwalbe M. et al., *Structure* **22** (2014) 238-249.
3. Novacek J. et al., *J. Biomol. NMR* **56** (2013) 291-301.
4. Jansen S. et al., *J. Biol. Chem.* **292** (2017) 6715-6727.
5. Malmendal A. et al., *Biophys. Biochem. Res. Commun.* **301** (2003) 136-142.
6. Laurine E at al., *J. Biol. Chem.* **278** (2003) 29979-29986.
7. Mukrasch MD et al., *PLoS Biology* **7** (2009) e1000034.
8. Lippens G et al., *Biomolecules* **6** (2016) 28.

O31

Promiscuous intrinsically disordered tail from bacterial repressor mediates binding to the operator

Hadži, San[1,2,3]

[1]Department of Physical Chemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, 1000 Ljubljana, Slovenia, [2] Structural Biology Brussels, Department of Biotechnology, Vrije Universiteit Brussel, B-1050 Brussel, Belgium, [3]Molecular Recognition Unit, Structural Biology Research Center, Vlaams Instituut voor Biotechnologie, B-1050 Brussel, Belgium

Background Toxin-antitoxin systems are believed to mediate bacterial stress response. They are tightly regulated networks both at the protein and DNA level. Here, we study a bacterial regulator HigA2 which regulates transcription of the *higBA2* operon.

Questions addressed: What mediates high promiscuity of IDP for binding different molecular surfaces such as protein and DNA?

Results and discussion: HigA2 regulator consists of a globular DNA binding domain and intrinsically disordered tail, which was shown to interact with the partner protein HigB2. Surprisingly, we find that the IDP tail not only interacts with the globular partner protein but also interacts with the operator sequence. The full-length HigA2 binds to the operator with nanomolar affinity, while the removal of IDP tail leads to a significant drop in affinity. Using a variety of structural (crystallography, SAXS, CD and NMR), biophysical and biochemical experiments we characterize how intrinsically disordered tail mediates operator binding and shed light on which characteristics of IDPs enable binding promiscuity.

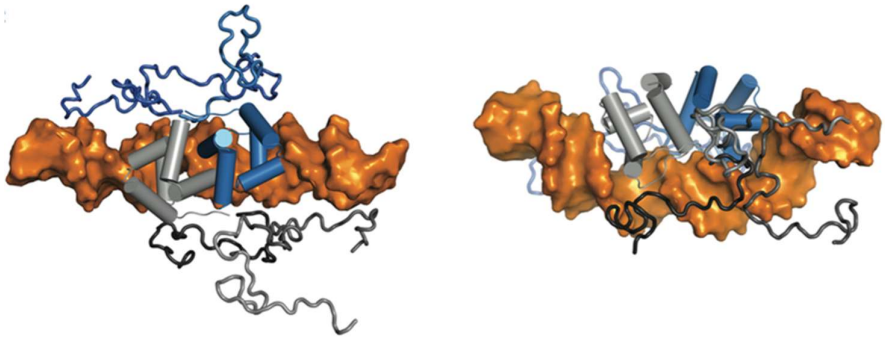


Figure 1. Integration of crystallographic and small-angle scattering data providing a glimpse of IDP-operator interactions that modulate the binding of HigA2 regulator to the operator sequence.

References:

1. To be published.

* Poster Abstracts *

Even numbers present on Wednesday (12th of September)

Odd numbers present on Tuesday (14th of September) *

*odd number stands will be free for your posters by Wednesday evening

P01

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

Bliven, Spencer[1,2], Anisimova, Maria[1,2]
[1]*Institute for Applied Simulation, Zurich University of Applied Sciences,* [2]*Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland*

Background: 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].

Questions addressed: Although tandem repeats are common, precisely identifying the relationships between repeats remains a challenge. We would like to precisely position repeat duplication and deletion events among the complex speciation and gene duplication events that have led to diverse tandem repeat families. The copy number changes are often very distant and thus require powerful phylogenetic methods to elucidate. Focusing on the ArmRP family, we present a comprehensive list of repeats as well as a detailed phylogeny of these repeats. Generalizing the method to a broader set of tandem repeats, we hope to gain insight into how biology has utilized repeats to generate new folds and functions.

Methods: 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 analyze the conservation and diversification patterns through evolution, based on the information about tandem repeat number, order and their distribution on phylogenies.

Results and discussion: TRAL has been run over a nonredundant set of eukaryotic proteomes to identify all Armadillo repeats. The number of ArmRP per species varies widely, indicating significant gene duplication and loss events. We also analyze the conservation patterns in the number of repeats between homologs to identify likely repeat expansions.

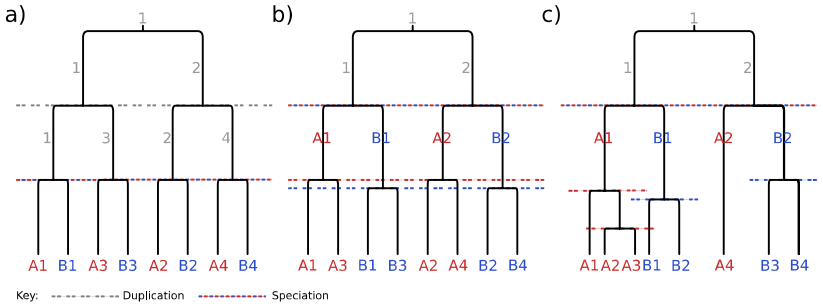


Figure 1. Possible evolutionary histories for two 4-repeat proteins. a) TR expansion occurs before speciation, giving identical TR orders in both species. b) Both species undergo independent whole-gene duplications, giving similar subtree patterns in both species. c) Both species undergo independent TR duplications showing different patterns.

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)..

P02

SDD-AGE as a method for estimation of molecular weight of amyloid aggregates

Bondarev Stanislav A.[1,2], Drozdova Polina B.[1,3], Belousov Mickhail V.[1,4], Zhouravleva Galina A.[1,2]
[1]SPbU, Department of Genetics and Biotechnology St Petersburg, Russia , [2]SPbU, The Laboratory of Amyloid Biology St Petersburg, Russia, [3]Institute of Biology at Irkutsk State University, Irkutsk, Russia, [4]ARRIAM, Laboratory for Proteomics of Supra-Organismal Systems, Pushkin, St Petersburg, Russia

Background: The semi-denaturing detergent agarose gel electrophoresis (SDD-AGE) was proposed as a routine method for a comparison of amyloid aggregate sizes [1]. Currently this approach is widely used for amyloid investigations, but only as a qualitative approach.
The yeast Sup35 protein is one of the well-studied amyloids. Its aggregation in cells leads to the appearance of the [*PSI*$^+$] prion. Unlike mammalian prions, [*PSI*$^+$] does not lead to cell death and may provide beneficial traits in certain environmental conditions [2]. Aggregates of Sup35 have an in-register super-pleated β-structure with one molecule in fibril cross-section [3]; the distance between adjacent molecules is 0.47 nm [4], [5]. These facts make it possible to calculate molecular weight (MW) of Sup35 aggregate from its length.

Questions addressed: We have proposed a method for estimation of the amyloid aggregates size based on SDD-AGE results.

Methods: We used an SDD-AGE protocol developed in the laboratory of Ter-Avanesyan [1] with published modifications [6]. Sup35NM (the fragment of the protein that has been widely used *in vitro* experiments) was purified from *E. coli* with affinity chromatography. MW of the obtained aggregates was estimated with transmission electron microscopy (TEM). We assumed that fibril MW is equal to product of Sup35NM MW and fibril length in nm divided by 0.47 nm (distance between two molecules in an aggregate). The program tool for SDD-AGE analysis and aggregates size evaluation was based on the combination of ImageJ/Fiji software and a custom R shiny application.

Results and discussion: The choice of a MW marker was a critical step of our work, because such reference is essential for estimation of molecule sizes of protein aggregates. For this purpose we used a DNA ladder, since in the literature we can find mathematical model describing mobility of DNA molecules in an agarose gel. We reproduced this model for DNA ladder in SDD-AGE conditions and then developed the protocol and tool for evaluation of protein aggregate sizes based on the results of electrophoresis. To test our approach we compared distributions of aggregates MW counted from TEM and SDD-AGE data and demonstrated their accordance.
This work was supported by RFBR (grants 18-34-00536,17-54-150002, 16-04-00202), the grant of the President of the Russian Federation (MK-512.2017.4 to MVB) and RRC MCT SPbSU.

References:
1. D. S. Kryndushkin, I. M. Alexandrov, M. D. Ter-Avanesyan, and V. V. Kushnirov, “Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104,” J. Biol. Chem., vol. 278, no. 49, pp. 49636–49643, Dec. 2003.
2. S. W. Liebman and Y. O. Chernoff, “Prions in yeast.,” Genetics, vol. 191, no. 4, pp. 1041–1072, Aug. 2012.
3. A. V. Kajava, U. Baxa, R. B. Wickner, and A. C. Steven, “A model for Ure2p prion filaments and other amyloids: the parallel superpleated beta-structure,” Proc. Natl. Acad. Sci. U.S.A., vol. 101, no. 21, pp. 7885–7890, May 2004.
4. F. Shewmaker, R. B. Wickner, and R. Tycko, “Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure,” Proc. Natl. Acad. Sci., vol. 103, no. 52, pp. 19754–19759, 2006.
5. F. Shewmaker, D. Kryndushkin, B. Chen, R. Tycko, and R. B. Wickner, “Two prion variants of Sup35p have in-register parallel beta-sheet structures, independent of hydration,” Biochemistry, vol. 48, pp. 5074–5082, 2009.
6. R. Halfmann and S. Lindquist, “Screening for Amyloid Aggregation by Semi-Denaturing Detergent-Agarose Gel Electrophoresis,” NIH Public Access, vol. 17, no. 2008, pp. 1–4, 2009.

P03

Exploring internal PDZ-binding motifs

Čalyševa, Jelena[1], Lüchow, Susanne[2], Ivarsson, Ylva[2], Gibson, Toby[1].
[1]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstraße 1, 69117, Heidelberg, Germany, [2]Department of Chemistry, Uppsala University , BMC Box 576, SE-75123 Uppsala, Sweden

Background: PDZ domains are highly abundant in Metazoa and are a part of a variety of crucial processes, including cellular signalling. This project is a part of PDZnet, a European Training Network focused on studying PDZ-protein interactions and exploring the possibility of treating PDZ domains as drug targets for various diseases, including brain diseases and cancer. Within PDZnet, several research and industrial institutions are focusing on these questions from different perspectives through individual work and collaborations.

Questions addressed: PDZ domains are most commonly known to bind C-terminal motifs in proteins, but cases of internal protein binding are also known. This project focuses on studying that type of PDZ-protein interaction, identifying internal PDZ-binding motifs and overall patterns behind internal PDZ binding. Upon identifying those patterns, the relevance of internal PDZ binding to crucial biological processes and disease can be assessed.

Methods: To study the internal PDZ binding, various bioinformatics methods are being applied on available data. The data for PDZ internal interaction analysis is obtained from collaboration with Ylva Ivarsson's group in Uppsala University, where they perform proteomic peptide phage display to identify internal binders of all human PDZ domains. To analyse the data, methods such as sequence alignment, regular expression search, protein disorder prediction, secondary structure prediction, structural alignment are being used.

Results and discussion: Several internal PDZ-binding motifs have been identified, highly resembling C-terminal binding. However, to paint a clear picture of internal PDZ binding, more protein-protein interaction analysis tools should be applied. Combining different methods and previous knowledge, a pattern behind internal PDZ binding can be recognised and new potential binders can be predicted and tested experimentally. Cancer and disease mutation data will also be scanned for the internal PDZ-binding motifs to identify their relevance to those diseases.

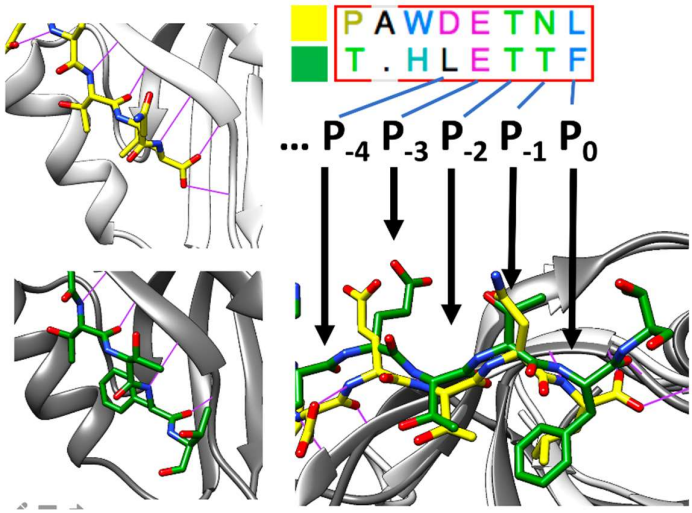


Figure 1. An example of internal PDZ binding and its similarity to C-terminal binding. The internal binding of syntrophin PDZ domain (left bottom, PDB 1QAV, bottom in the alignment on the top right) is very similar to SHANK1 C-terminal PDZ binding (top left, PDB 3L4F, top in the alignment on the top right), as can be seen from the structure superimposition (bottom right). The internal binding motif is similar to known C-terminal PDZ binding motif x[TS]xΦ, where x – any residue, Φ – hydrophobic residue, and residues in the square brackets correspond to possible residues in that position.

P04

Interplay between structures of mature aggregates and "nucleus of aggregation"

Likholetova Daria V.[1], Bondarev Stanislav A.[1,2], Zhouravleva Galina A.[1,2]
[1]SPbU, Department of Genetics and Biotechnology, St Petersburg, Russia, [2]SPbU, The Laboratory of Amyloid Biology, St Petersburg, Russia,

Background: Prions are infectious amyloids. In humans, prions cause such neurodegenerative diseases as Curu, Creutzfeldt–Jakob, Fatal familial insomnia, *etc*. At the same time yeast prions usually do not lead to cell death and can provide beneficial traits to the organism. All prions are able to form various variants or strains which lead in different diseases in mammals or to distinct phenotypic changes in yeasts. For instance, variants of the well-known yeast [*PSI*$^{+}$] prion have different level of nonsense-suppression, that can be detected by colony colour. This phenomenon is based on the polymorphism of amyloid aggregate structure. Despite a long period of structural investigations of amyloids, the principles of amyloid formation are still poorly understood. This process is triggered by accumulation of misfolded protein and appearance of "nucleus of aggregation", which then initiates amyloid aggregate's assembly. According to current data, little is known about the role of "nucleus of aggregation" in formation of mature prion aggregates. Two theoretical possibilities may be proposed: (1) structures of aggregation nucleus and mature aggregate are identical or (2) their arrangements are different. [*PSI*$^{+}$] yeast prion of *Saccharomyces cerevisiae* leads to aggregation of Sup35 protein (a translation termination factor eRF3). Sup35 has three domains: aminoterminal N-domain, middle M-domain and carboxyterminal C-domain. N-domain is Q/N rich and sufficient for prion propagation. Recently in our laboratory a range of *sup35*^{KK} alleles was constructed, all of these mutations lead to substitutions of polar amino acids to charged one in Sup35 prionogenic domain. Charged residues disrupt super-pleated β-structure, which is accepted for Sup35 aggregates. These mutations allow us to restrict and determine region included into Sup35 aggregate.

Questions addressed: In this study we investigate how structure of aggregation nucleus corresponds to the arrangement of mature aggregates.

Methods: New [*PSI*$^{+}$] prion variants were obtained by overproduction of Sup35NM (WT or with Q46K/Q47K, Q70K/Q71K substitutions) in [*psi*$^{-}$][*PIN*$^{+}$] strain. Prion phenotype and its strength was analysed on ¼ YEPD or on adenine omitting media. Mutational analysis was used for identification of Sup35 region forming super-pleated β-structure, as was described earlier [1]. Briefly, [*PSI*$^{+}$] variants with single *SUP35* copy on the *URA3* plasmid were transformed with set of plasmids with *sup35*^{KK} alleles. Transformants were plated on 5'-FOA containing media for counter selection of *URA3* plasmid. After this prion phenotype of cell bearing only *sup35*^{KK} mutation was analysed. According to our assumptions, prion loss should be observed only if mutation is located in amyloid forming region.

Results and discussion: According to hypothesis (1), mature aggregate and nucleus of aggregation share a common structure. To check this, we obtained a set of [*PSI*$^{+}$] variants by overexpression of *sup35*^{KK} alleles. These mutations restrict amyloidogenic region of Sup35 protein, thus we obtained different prion variants with defined region forming super-pleated β-structure. Also investigated mutations have significant impact on amyloidogenic properties of the protein and affect the frequency of [*PSI*$^{+}$] induction. We demonstrated that the reduced rate of [*PSI*$^{+}$] colonies after overproduction of Sup35^{KK} is linked with decreased amyloidogenic potential, predicted by ArchCandy program [2].
To test both hypotheses, we determine the Sup35 amyloid forming region in obtained [*PSI*$^{+}$] strains using mutational analysis. We changed *SUP35* allele maintaining prion to one of *sup35*^{KK} mutations. If substitution in Sup35 protein is located in amyloidogenic region we will observe prion loss after allele shuffling. We found three cases out of 19 in which structure of mature prion aggregate differs from nucleus of aggregation. So we conclude that amyloid structure of mature aggregates is not strictly determined by nucleus of aggregation. This work was supported by RFBR (grants 18-34-00537, 17-54-150002, 16-04-00202) and RRC MCT SPbSU.

References:

1. Bondarev, S. A., Shchepachev, V. V., Kajava, A. V., & Zhouravleva, G. A. (2013). Effect of charged residues in the N-domain of Sup35 protein on prion [PSI+] stability and propagation. *Journal of Biological Chemistry*, *288*(40), 28503-28513.
2. Ahmed, A. B., Znassi, N., Château, M. T., & Kajava, A. V. (2015). A structure-based approach to predict predisposition to amyloidosis. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, *11*(6), 681-690.

P05

Fuzzy Interactome of α-Actinin of Sarcomeric Z-disk

Antonio Sponga, Valeria Stefania, Joan Lopez Arolas, Julius Kostan & Kristina Djinovic-Carugo
Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna,

The ultra-structure of sarcomere, the basic contractile unit in striated muscle cells, is well organized and delimited by Z-disks, which play a central role in the mechanical stabilization and force transmission. α-Actinin isoform 2 (208 kDa) is a key protein in Z-disk assembly as it crosslinks antiparallel actin filaments from adjacent sarcomeres. It is also a binding platform for a number of other Z-disk proteins such as titin, FATZ-1and ZASP, which participate in Z-disk formation and regulation. FATZ-1/myozenin-1/calsarcin-2 and ZASP/cypher are predominantly intrinsically disordered ad believed to act as adaptors linking α-actinin-2 to other Z-disk proteins, while titin, a giant multidomain protein (~3,800 kDa; including folded and disordered parts) is the molecular ruler that spans half the sarcomere and combines elastic, architectural and signaling functions. Both FATZ-1, ZASP and titin (via its Zq domain) are reported to interact with the rod domain of α-actinin-2, while PDZ domain of ZASP interacts with the calmodulin-like domain of α-actinin-2.

In a comprehensively study we addressed the macromolecular assembly of α-actinin-2 with FATZ-1, titin and ZASP by combining complementary techniques such as crosslinking/MS, NMR, SAXS, and crystallography (our unpublished data). Crystal structure analysis of a soluble FATZ-1 construct in complex with half dimer of α-actinin-2 shows that FATZ-1 interacts in an extended conformation with spectrin-like repeats 1, 3 and 4 of α-actinin-2 rod, and modulates the EF hand pairs 1-2 the position found in the previously reported structure of full length α-actinin-2 (1). We subsequently asked whether interactions of FATZ-1 and titin with the α-actinin-2 rod are competitive, synergic or independent. To this end we determined crystal structure of the complex between titin Zq region and α-actinin-2 rod and characterized the binding affinities and stoichiometry for binary (α-actinin-2/titin and α-actinin-2/FATZ-1) and ternary (α-actinin-2/titin/FATZ-1) complexes using ITC and SEC-MALS, which corroborated our structural data. SAXS derived solution structure of binary and ternary assemblies display features of fuzzy complexes, with large segments of α-actinin-2 binders being intrinsically disordered. Finally, we show ZASP forms a stable binary, ternary and also quaternary fuzzy complexes with α-actinin-2, titin and FATZ-1. These findings will be discussed in terms of Z-disk structure and assembly.

References:

1.Ribeiro Ede, A., Jr., Pinotsis, N., Ghisleni, A., Salmazo, A., Konarev, P. V., Kostan, J., Sjoblom, B., Schreiner, C., Polyansky, A. A., Gkougkoulia, E. A., Holt, M. R., Aachmann, F. L., Zagrovic, B., Bordignon, E., Pirker, K. F., Svergun, D. I., Gautel, M., and Djinovic-Carugo, K. (2014) The structure and regulation of human muscle alpha-actinin. *Cell* **159**, 1447-1460

P06

Conformational biases of DC8E8 tetratope peptides

Cehlar Ondrej [1,2], Skrabana Rostislav [1,2], Novak Michal [1]
[1]*Institute of Neuroimmunology, Slovak Academy of Sciences, Dubravska cesta 9, 845 10 Bratislava, Slovakia,*
[2]*Axon Neuroscience R&D Services SE, Dvorakovo Nabrezie 10, 811 02 Bratislava, Slovakia*

Background: Accumulation of intrinsically disordered protein tau in the form of insoluble aggregates is a common feature of neurodegenerative tauopathies. Monoclonal antibody DC8E8 is able to inhibit tau-tau interaction and discriminate between full length disordered tau species and truncated aggregation prone mis-disordered tau species [1]. Therefore it holds promise for the immunotherapy of Alzheimer's disease. The active vaccine based on the DC8E8 epitope peptide has successfully passed the phase 1 clinical trial [2, 3]. The same peptide sequence was incorporated into the Hepatitis B core virus like particle based vaccine tested in transgenic mouse model [4].

Questions addressed: The various affinity of DC8E8 to particular microtubule binding repeats (MTBRs) may be proportional to the percentage of sampled bound-like conformation in the free state.

Methods: We have used molecular dynamics and peptide docking to address the above mentioned question. The MD simulations were run using the NAMD 2.8 software with CHARMM36m force field [5] suitable for simulation of disordered proteins. The conformations were compared to the X-ray structure of DC8E8 with tau peptide from MTBR2 (Skrabana, in preparation, PDB ID: 5MP1). The epitope sequences from each MTBR of tau protein were docked into DC8E8 Fab holo-structure using PIPER-FlexPepDock protocol [6].

Results and discussion: We have performed 300 ns long unrestrained all-atom molecular dynamics simulations of 18-mer tau peptides from all four tau MTBRs (Figure 1) and compared the percentage of sampled bound-like conformations with the antibody affinity. We have also evaluated the formation of main chain hydrogen bond stabilizing the PGGG loop conformation that was most prevalent in peptide from MTBR2. DC8E8 has also the highest affinity to MTBR2, what is consistent with its higher affinity to 4R over 3R tau forms [1]. We have also performed peptide docking to compare the conformations of peptides from other three MTBRs to X-ray conformation of MTBR2 from complex with DC8E8. Unravelling the unique mode of recognition of DC8E8 antibody and conformational biases of tau protein repeat regions can aid to reveal the hindered structural features of tau protein biology and broaden the knowledge about rthe interactions of disordered proteins with their globular interaction partners.

Figure 1. The schema of tau microtubule binding domain composed of repeats and interrepeats, followed by region R'. Peptides used for the MD simulations and DC8E8 tetratope are shown in unfilled and grey boxes respectively.

Acknowledgements: This work was supported by the VEGA grant No. 2/0177/15. Calculations were performed in the Computing Centre of the Slovak Academy of Sciences using the supercomputing infrastructure acquired in project ITMS 26230120002 and 26210120002 (Slovak infrastructure for high-performance computing) supported by the Research & Development Operational Programme funded by the ERDF.

References:

1. Kontsekova, E., et al., Identification of structural determinants on tau protein essential for its pathological function: novel therapeutic target for tau immunotherapy in Alzheimer's disease. Alzheimers Research & Therapy, 2014. **6**(4).
2. Kontsekova, E., et al., First-in-man tau vaccine targeting structural determinants essential for pathological tau-tau interaction reduces tau oligomerisation and neurofibrillary degeneration in an Alzheimer's disease model. Alzheimers Research & Therapy, 2014. **6**(4).
3. Novak, P., et al., Safety and immunogenicity of the tau vaccine AADvac1 in patients with Alzheimer's disease: a randomised, double-blind, placebo-controlled, phase 1 trial. Lancet Neurology, 2017. **16**(2): p. 123-134.
4. Ji, M., et al., Hepatitis B core VLP-based mis-disordered tau vaccine elicits strong immune response and alleviates cognitive deficits and neuropathology progression in Tau.P301S mouse model of Alzheimer's disease and frontotemporal dementia. Alzheimer's research & therapy, 2018. **10**(1): p. 55.
5. Huang, J., et al., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nature Methods, 2017. **14**(1): p. 71-73.
6. Alam, N., et al., High-resolution global peptide-protein docking using fragments-based PIPER-FlexPepDock. Plos Computational Biology, 2017. **13**(12).

P07

A heuristic approach for clustering metabolic networks into highly connected components

Grbić, Milana[1], Kartelj, Aleksandar[2], Matić, Dragan[1], Janković, Savka[1], Filipović, Vladimir[2]
[1]Faculty of Natural Science and Mathematics, University of Banja Luka, Bosnia and Herzegovina, [2]Faculty of Mathematics, University of Belgrade, Serbia

Background: Network clustering is a commonly used approach for obtaining new information from large biological data. A sequence of metabolite reactions can be represented as a network, where two metabolites are adjacent if they coexist in the same reaction. By removing as few edges as possible, we cluster a network into the so called highly connected components. A component is highly connected if all its nodes have a degree greater than $n/2$, where n is the number of nodes in that component. This approach has already been proven as useful in computational biology, for example in finding complexes in protein–protein interaction (PPI) data with similar GO annotations.

Questions addressed:
By clustering a network into dense subnetworks, the dimension of the biological structure is decreased, but useful information about particular biological functionalities can still remain in the dense clusters. The mentioned network clustering problem is called Highly connected deletion problem and it belongs to the class of NP hard problems. Therefore, exact methods cannot be applied on large-scaled networks, so the usage of heuristic methods is justified.

Methods:
In this research we apply our local search based heuristic method to various metabolic networks. The essential part of the algorithm is fast local search procedure that systematically swaps the clusters of pairs of metabolites. Specifically designed objective function takes into account degree of nodes in each partition and penalizes infeasible solutions, directing the search into more promising areas of the solution space.

Results and discussion: We show that proposed method can be used for clustering biological networks into smaller components in a reasonable running time. By applying the algorithm on various metabolite networks, we can analyze similarities and differences between the obtained clusters of different organisms.

References:

1. Hüffner, F. et al. (2014). " Partitioning Biological Networks into Highly Connected Clusters with Maximum Edge Coverage. " IEEE/ACM Trans Comput Biol Bioinform., 11(3):455-67.
2. Förster, J. et al. (2003). "Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network." *Genome research* 13.2, 244-253.
3. Feist, A.M. et al. (2007). "A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information" Molecular Systems Biology 3:121
4. Martins, P. (2016). Modeling the maximum edge-weight k-plex partitioning problem. arXiv preprint arXiv:1612.06243.

P08

NMR assignment and structural characterization of mouse prion protein fibrils

Fridmanis, Jēkabs[1], Toleikis, Zigmantas[2], Šneideris, Tomas[2], Liepiņš, Edvards[1], Smirnovas, Vytautas[2], Jaudzems, Kristaps[1]

[1]Latvian Institute of Organic Synthesis, Aizkraukles 21, Riga LV-1006, Latvia,
[2]Vilnius University, Institute of Biotechnology, Sauletekio 7, Vilnius, LT-10257, Lithuania

Background: Prions are infectious particles that play the main role in a group of fatal neurodegenerative disorders, also known as the transmissible spongiform encephalopathies (TSE's). Prion diseases propagate by self-replication of a pathogenic prion protein isoform (PrP^{Sc}) using cellular prion protein (PrP^{C}) as a substrate [1, 2]. Understanding the structure and possible mechanisms of prion aggregation would be a big step towards curing these diseases. However, the structures of infectious forms of PrP are still only partially defined, and several different models for the PrP fibril structure have been proposed. To get insight into the PrP fibril structure we have used solid-state nuclear magnetic resonance spectroscopy (ssNMR) and studied fibrils formed by the mouse prion protein comprising residues 89-230 MoPrP(89-230).

Questions addressed: Which residues form the rigid core of MoPrP(89-230) fibrils? Which residues are involved in regular secondary structures? What is the overall structure of MoPrP(89-230) fibrils?

Methods: Uniformly ^{13}C and ^{15}N labeled MoPrP(89-230) protein was expressed in E. coli, purified and fibrillized as described previously [3]. Two rounds of seeding were necessary to obtain a structurally homogeneous sample. NMR spectra were recorded on an 800 MHz Bruker Avance III HD spectrometer equipped with a 3.2-mm $^{13}C/^{15}N\{^{1}H\}$ E-free magic-angle spinning (MAS) probe. The MAS frequency was set to 12.5 or 20 kHz and the temperature was regulated using a cooling gas flow at 273 K.

Results and discussion: Analysis of 2D NCA, NCO, DARR and 3D NCACX, NCOCX, CANCO, CONCA MAS NMR spectra allowed a partial assignment of residues in the rigid core of MoPrP(89-230) fibrils. The rigid core is formed by residues 165-223, which corroborates previous findings from ssNMR studies of full-length Syrian hamster PrP fibrils [4] suggesting that the rigid core is comprised primarily of residues in the 173-224 range. Further analysis of the $^{13}C\alpha$ and $^{13}C\beta$ secondary chemical shifts allowed identification of the locations of regular secondary structures. Based on these data we propose a structural model of MoPrP(89-230) fibrils.

References:

1. Prusiner, S. B. Prions. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 13363–13383 (1998).
2. Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24,** 519–550 (2001).
3. Milto, K., Michailova, K., Smirnovas, V. Elongation of mouse prion protein amyloid-like fibrils: effect of temperature and denaturant concentration. *PLoS ONE* **9**, e94469 (2014).
4. Tycko, R., Savtchenko, R., Ostapchenko, V. G., Makarava, N., Baskakov, I. V. The α-helical C-terminal domain of full-length recombinant PrP converts to an in-register parallel β-sheet structure in PrP fibrils: evidence from solid state nuclear magnetic resonance. *Biochemistry* **49**, 9488-9497 (2010).

P09

Text mining and data curation in the DisProt database

András Hatos[1], Damiano Piovesan[1] and Silvio C.E. Tosatto[1].
[1] *Department of Biomedical Sciences, University of Padua, Italy*

Background: The Database of Protein Disorder (DisProt, URL: www.disprot.org) is the major repository for manually curated intrinsic disorder (ID) annotations which are provided by almost 40 curators from 12 different groups in as many country. DisProt annotation uses a controlled vocabulary of functional terms associated with ID, in three main categories (molecular function, structural transition and interaction partner) in addition to a selection of experimental techniques used to determine ID.

Questions addressed: The ID curation process is complicated and difficult to automatize. Integration of text mining algorithms and the collection of article sentences associated to DisProt annotations will provide a starting point to scale ID curation on a large scale and speed up the process.

Methods: The new DisProt annotation interface was re-designed to includes SciLite, a text mining software maintained by EuropePMC, which automatically extract bio-entities (gene/protein names, functional terms, ...) from articles. The DisProt vocabulary was mapped to other ontologies and enriched with a new set of synonyms which has been used to re-train the NextA5 web server in order to detect ID-related articles associated to specific proteins. The new version of the NextA5 web server has been developed in collaboration with the SIB text-mining laboratory and is available to DisProt curators.

Results and discussion:
The new curation interface of DisProt exploits text mining algorithms to rank the literature and highlight bio entities in manuscripts. Also, curators now are encouraged to report publication sentences corresponding to annotated terms. This will provide a new corpus of statements for the development of a new generation of text mining tools which in turn will improve article curation and processing. In the future it will be possible to completely automate curation and extend the process on a large scale and integrate DisProt annotations into SciLite and EuropePMC.

P10

Incorporation of fluorinated amino acids into Alzheimer's amyloid-beta peptide and studies of its aggregation mechanism

Kitoka, Kristine[1], Biverstal, Henrik[1,2], Jaudzems, Kristaps[1]
[1]*Latvian Institute of Organic Synthesis, Riga, LV-1006, Latvia*
[2]*Karolinska Institutet, NVS, Division for Neurogeriatrics, 141 52 Huddinge, Sweden*

Background: Alzheimer's disease (AD) is a progressive neurodegenerative disease that affects many people worldwide. AD is characterized by the progressive loss of cognitive functions. Amyloid-beta (Aβ) peptides, which are produced by the proteolytic cleavage of the amyloid precursor protein (APP), play the main role in the disease progress. These peptides have tendency to form aggregates in the central nervous system. The aggregates are toxic for the brain cells, and as a result, the cells are damaged or destroyed and cannot work properly. The predominant isoform in the amyloid aggregates is the 42-residue peptide Aβ42. It has an increased tendency to aggregate and is more toxic than other isoforms. Intensive research has been performed to understand the detailed mechanism of Aβ aggregation. Unfortunately, there is little progress in these studies. NMR spectroscopy is an excellent method to study protein structure and dynamics, with the aim to decipher molecular mechanisms. However, NMR requires production of peptides with magnetically active nucleus labels. Fluorine is an attractive NMR label with high sensitivity (83% of that of proton) and 100% natural abundance. The chemical shift of ^{19}F nucleus is extremely sensitive to the surrounding environment, so it can be used to monitor changes in protein conformation, that mostly are not detectable using other techniques. For instance, the fluorine chemical shift may increase or decrease by as much as 8 ppm when protein is denatured [1]. Furthermore, incorporation of fluorinated amino acids by *in vivo* expression in *E. coli* is a widely used straightforward method, that allows modifying proteins for different useful aspects [2].

Questions addressed: The aim of this study is to express and purify the Aβ42 peptide and its fluorinated analogues to study their aggregation by means of NMR.

Methods: Protein expression in *E. coli*, nickel affinity, desalting and size exclusion chromatography and lyophilisation were used to produce and purify proteins and peptides. Proteins and peptides were characterized using NMR, UV-Vis spectroscopy and MALDI-TOF mass spectrometry.

Results and discussion: Investigation of aggregation mechanisms requires the production of large amounts of pure Aβ peptides in monomeric form. In our study, we used a construct (NT*Aβ42), in which the Aβ42 peptide is fused with a solubility enhancing tag derived from a spider silk protein [3]. We present a method for high yield expression and purification of the fusion protein, its fluorinated analogues, and for obtaining the peptides in monomeric form. Obtained proteins and peptides were studied by NMR using 2D [^{15}N-^{1}H]-HSQC, ^{1}H, and ^{19}F experiments.

References:

1. Yoder N, Kumar K: Fluorinated amino acids in protein design and engineering. *Chem Soc Rev,* 2002; 31:335-341.
2. Hou Y, Hu W, Li X, Skinner J, Liu D, Wütrich K: Solvent-accessibility of discrete residue positions in the polypeptide hormone glucagon by ^{19}F-NMR observation of 4-fluorophenylalanine. *J Biomol NMR*, 2017; 68:1-6.
3. Kronqvist N, Sarr M, Lindqvist A, Nordling K, Otikovs M, Venturi L, Pioselli B, Purhonen P, Landreh M, Biverstal H, Toleikis Z, Sjoberg L, Robinson CV, Pelizzi N, Jornvall H, Hebert H, Jaudzems K, Curstedt T, Rising A, Johansson J: Efficient protein production inspired by how spiders make silk. *Nat Commun*, 2017; 8:15504.

P11

Unit annotation of tandem repeat proteins, RepeatsDB-Lite

Hirsh, Layla[12+] , Paladin, Lisanna[1+] , Piovesan, Damiano[1] Tosatto Silvio C. E. [13]
[1] Dept. of Biomedical Sciences, University of Padua, Padua, Italy, [2] Dept. de Ingeniería, Pontificia Universidad Catolica del Peru, Lima, Peru, [3] CNR Institute of Neurosciences, Padua, Italy.

Background: Tandem repeats (TR) in proteins are ubiquitous in genomes and have been demonstrated to be of fundamental importance in many biological processes (1). TR proteins are characterized by a modular structure stabilized by a pattern of local interactions (2), which can be arranged in a wide variety of shapes providing functional diversity. Structural TR modules, called units (5), correspond to repeated segments in the sequence which can be loosely conserved both at the DNA and amino acid level. Subtle differences in the structural conformation of the units give rise to large differences in shape and structural properties of the whole protein, including curvature and twist (3). Protein function is also strongly related to unit types and single units are often recognized as functional determinants of protein families. The largest collection of TR proteins detected by structural features is provided by the RepeatsDB database (5). It relies on computational approaches and expert manual curation to detect TR in the Protein Data Bank (PDB) structures (6). RepeatsDB annotates the unit position and provides a hierarchical classification of TR protein families at the level of classes and sub-classes (5). Structure based methods detect TR periodicity by implementing different strategies, such as the identification of 3D symmetries or the analysis of regularity of structural features. In general, these methods are designed to discriminate between repeat and non-repeat structures. Our method ReUPred (7) identifies structural units by comparing the protein structure against a manually curated library of TR units.

Questions addressed: Is it possible to create a tool that facilitates the annotation process, a web-based graphical user interface that allows a complete visualization of the Repetitive Unit Predictor (ReUPred) results?

Methods: RepeatsDB-lite (8) is a web server designed for the prediction, visualization and analysis of repeated regions in protein structures. It is based on an improved ReUPred algorithm (7) using several checks to minimize errors in the unit detection step and speeding up the calculation. A refactoring of the TR unit library allows it to cover all RepeatsDB classes. Its ability of predicting unit position is evaluated against all manually curated RepeatsDB entries (5). The web interface is designed to provide a complete visualization of the data including structural and sequence alignments of the predicted units. In addition, it includes an intuitive form to manually refine the annotation and visualize the effect on the unit alignments on the fly.

Results and discussion: TR proteins are increasingly studied as evidence for new functions accumulate and new structures become available. The precise identification of the structural repeat modules allows to infer structural properties of the entire protein, its family and function. RepeatsDB-lite (8) allows to identify units and classify the protein exploiting a structural similarity search and the information available in RepeatsDB. It outperforms existing methods and can be applied to all types of TR proteins. The web interface allows to visualize similarity relationships between TR units at both the sequence and structure level. The prediction can be manually refined by the user, visualizing the effects of the edits in real time. Annotations can be submitted to RepeatsDB for reviewed and we hope this will increase the amount of communitycurated entries in the database. RepeatsDB-Lite (8) can be seen as an example of gamification principles to engage a wider community towards database curation.

References: 1. Marcotte,E.M., et al. (1999) A census of protein repeats. J. Mol. Biol., 293, 151–160. 2. Kajava,A.V. (2012) Tandem repeats in proteins: from sequence to structure. J. Struct. Biol., 179, 279–288. 3. Kobe,B. and Kajava,A.V. (2000) When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends Biochem. Sci., 25, 509–515. 5. Paladin,L., Hirsh,L., Piovesan,D., Andrade-Navarro,M.A., Kajava,A.V. and Tosatto,S.C.E. (2017) RepeatsDB 2.0: improved annotation, classification, search and visualization of repeat protein structures. Nucleic Acids Res., 45, D308–D312. 6. Burley,S.K., et al. (2017) Protein data bank (PDB): the single global macromolecular structure archive. Methods Mol. Biol. Clifton NJ, 1607, 627–641. 7. Hirsh,L., Piovesan,D., Paladin,L. and Tosatto,S.C.E. (2016) Identification of repetitive units in protein structures with ReUPred. Amino Acids, 48, 1391–1400. 8. Hirsh,L., et al.(2018) RepeatsDB-lite: a web server for unit annotation of tandem repeat proteins. Nucleic Acids Research, 46, W402–W407.

Structure of intrinsically disordered protein tau in its pathological form

Mihalovicova Klaudia[1], Cehlar Ondrej[1,2], Legenova Andrea[2], Drozdikova Eva[2], Majerova Petra[1,2], Novak Michal[1], Skrabana Rostislav[1,2],

P12

[1] *Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia,* [2]*AXON Neuroscience R&D Services SE, Bratislava, Slovakia*

Background: Tau protein is an intrinsically disordered protein, which plays an important role in stabilization of microtubules [1]. However, in Alzheimer's disease and other tauopathies, tau protein is the constituent of neurofibrillary tangles [2]. Ultrastructurally, tau inclusions are made of paired helical filaments (PHF) and straight filaments. Several monoclonal antibodies specifically recognize pathological fold present on the PHF core [3].

Questions addressed: Structure of tau fibrils has been recently elucidated using cryo-electron microscopy, however, there are still unresolved electron densities at both N and C terminal regions of the filament core [4,5]. Equally, it is not clear, what is the role of tight steric zipper structures described on short tau peptides [6,7], as they were not observed in the mature PHF core fibrils [4].

Methods: We are studying pathological transition of tau by inducing the PHF core structure by conformationally specific monoclonal antibodies. Following analysis of available structures of pathological tau, we have selected specific antibodies. We expressed antibodies in the CHO cells and performed co-crystallization with selected PHF core tau protein variants. After screening co-crystallization conditions and crystal growth optimization we will solve the structure using X-ray.

Results and discussion: Crystallographic and cryo-EM studies show that the interaction of beta sheets is central to the assembly of soluble oligomers as well as mature amyloid fibers, which contain a common cross beta spine. Protein tau contains two six-residue fibril forming segments VQIINK and VQIVYK at the start of repeat 2 and 3, respectively, which drive tau aggregation. However, the hexapeptide steric zippers are not necessarily present in the core of mature fibrils.

We have developed an expression vector driving secretion of recombinant antibodies into culture media of the CHO cells. We achieved a high Fab yield (200 mg/l), with good quality for crystallography. We did screening of co-crystallization conditions between two antibodies Fabs and recombinant truncated tau protein dGAE (297-391), initial crystallization hits will be further optimised.

References:

1. T. Guo, W. Noble, D. P. Hanger. Acta Neuropathol 133, (2017), 665-704
2. C. M. Wischik, M. Novak, H. C. Thogersen, et al. Proc Natl Acad Sci USA 85, (1988), 4506-4510
3. B. Kovacech, M. Novak. Curr Alzheimer Res 7, (2010), 708-16
4. M. Novak, J. Kabat, C. M. Wischik. EMBO J 12, (1993), 365-370
5. A. W.P. Fitzpatrick, B. Falcon, S. He, et al. Nature 547, (2017), 185-190
6. M. R. Sawaya, S. Sambashivan, R. Nelson, et al. Nature 447, (2007), 453-7
7. P. M. Seidler, D. R. Boyer, J. A. Rodriguez, et al. Nature Chem 10, (2018), 170-176.

P13

Quantitative evaluation of the PDZ-mediated human interactome and its hijacking by viral proteins

Jané P[1], Chiron L[2], Vincentelli V[3], Nominé Y[1], Travé G[1].
1Institut de Génétique et de Biologie Moléculaire et Cellulaire – 1 rue Laurent Fries – 67404 Illkirch, France, 2CASC4DE Company, Illkirch, France, 3AFMB UMR Campus de Luminy- 163 Avenue Luminy 13288 Marseille, France

Background: Most of the data collected in large interactomics projects are binary (i.e. "binds" or "doesn't bind"). This is a simplified qualitative view, since proteins are involved in intracellular competitions depending on affinity constants. Interestingly, many protein-protein interactions are mediated by globular domains binding to short linear motifs (domain-motif interactomes).

Questions addressed: Our project aims to quantify all the interactions mediated by PDZ domains which recognize specific linear motifs (PBMs) at the C-termini of partner proteins. The human proteome contains 266 PDZ domains (the PDZome) dispersed over 154 proteins, and a few thousands of putative PBMs. PDZmediated interactions are recurrently perturbed in carcinogenesis. Several tumor virus proteins, such as human papillomavirus E6 oncoprotein, contain "viral mimics" of human PBMs, that target cancer-relevant PDZ-containing cellular proteins.

Methods: To quantitatively address the specificity of such networks, we have developed the highthroughput Hold-Up assay, which measures quantitative in vitro binding affinity values for large numbers of domain-motif complexes (Vincentelli et al., Nature Methods 2015).

Results and discussion: We use bioinformatics to build "binding profiles" extracted from the Hold-Up assay for each tested PBM against the PDZome. Those data will build up a quantitative specificity map of the PDZ-PBM interactome and provide a wealth of biological insights on individual PDZ-mediated interactions. On the other hand, we are also analyzing the sequences of the 266 human PDZ of the potential 5000 human PBMs and of a series of viral PBMs, in order to classify and compare them by means of different parameters (physico-chemistry, conservation, disorder/globularity, gene ontology..). By merging this information with the Hold-Up affinity data, we will generate a clarified and more quantitative map of the global PDZ-PBM human interactome and of its hijack by tumor virus proteins that carry mimics of human PBMs.

P14

Emergence of Visible Light Optical Properties of Phenylalanine Aggregates

Žiaunys, Mantas, Šneideris, Tomas, Smirnovas, Vytautas
Institute of Biotechnology, Life Sciences Center, Vilnius University, Lithuania

Background: The ability of phenylalanine to form fibrillar nanostructures was demonstrated on multiple occasions and such an oligomerization reaction could be the cause of cytotoxicity in patients with phenylketonuria [1]. These findings were quickly followed by claims that these fibrils have amyloid-like properties and can be detected through the use of thioflavin T, an amyloid-specific fluorescent dye[2].

Questions addressed: The previously conducted experiments did not take into consideration the possible phenylalanine molecule aromatic ring interactions, resulting in pi-stacking aggregates, which have been shown to exhibit visible light spectrum fluorescence [3]. These structures would also have the properties of J-aggregates, resulting in a fluorescence emission red-shift. Such an occurrence could give false positive results during a thioflavin-T assay and lead to a false identification of amyloid-like fibrils.

Methods: 100 mM phenylalanine solutions were incubated at 60° C for 15 days and their fluorescence 3D spectra were recorded every 3 days using an excitation range of 200-550 nm, emission range of 250-650 nm and wavelength increment of 5 nm. The resulting aggregate size correlation to fluorescence spectra positions were analysed by sample filtration through 0.22 μm filters and 3 kDa concentrators. Aggregate stability was examined by sample dilution and sonication.
Phenylalanine hydrogel was generated by dissolving phenylalanine at 80° C to a concentration of 300 mM, followed by cooling to 25° C.

Results and discussion: In this work we have shown the emergence of phenylalanine aggregates with visible light optical properties throughout a 15 day incubation period (Figure 1). The larger size aggregate fluorescence excitation-emission maxima positions in the 3D spectra were seen at higher wavelength positions, suggesting a correlation between aggregate size and their optical properties. The formed aggregates were stable both after dilution and sonication. The hydrogel, however, did not possess any unusual optical properties.

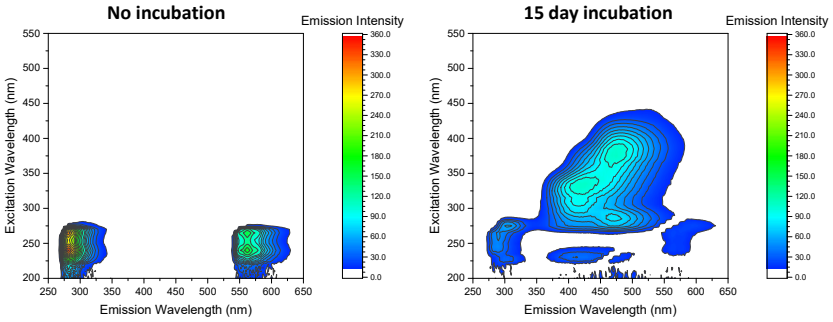


Figure 1. 100 mM phenylalanine solution 3D fluorescence spectra before and after incubation.

References:

1. Adler-Abramovich, L. *et al.* Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. *Nat. Chem. Biol.* **8,** 701–706 (2012).
2. Shaham-Niv, S., Adler-Abramovich, L., Schnaider, L. & Gazit, E. Extension of the generic amyloid hypothesis to nonproteinaceous metabolite assemblies. *Sci. Adv.* **1,** 1–6 (2015).
3. Kuo, A. Fluorescence Resulting from π-stacking in Polystyrene Solutions. *CheM* 1, 80–86 (2011).

P15

Efficiency of custom protein sequence comparison tools to align low complexity regions

Gruca Aleksandra[1], Jarnot Patryk[1], Grynberg Marcin[2]
[1]Institute of Informatics, Silesian University of Technology, Poland, [2]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Poland

Background: There are many methods, known from the literature, that are suitable for searching for similar protein sequences in the given dataset. Usually, on the basis of the results of similarity searching we can conclude about the properties, functions and structure of a protein. However, these methods are based on the statistical models which determine the similarity of proteins by comparison high complexity fragments Therefore, it is questionable if results of searching for low complexity regions (LCRs) generated by these methods are valuable.

Methods: The goal of this study is to check if and how these well-known methods of searching for similar proteins could be applied to the datasets that include either high or low complexity sequences. First of all we retrieve all high and low complexity regions from the SwissProt database using the SEG method [1]. Then we search for similar pairs of LCRs using CD-HIT-2D [2], BLAST [3] and HHblits [4]. In case of BLAST and HHblits we had to create custom databases of HCRs and LCRs. For BLAST we generated two sets of results for LCRs. The first set is based on default parameters for short sequences. In order to generate the second set of the results we adjusted parameters of BLAST specifically for LCR.

Results and discussion: Our results are specifically focused on similar LCR pairs that come from different protein families. Our comparison includes the number of pairs and the overlap obtained using selected methods. As a control of our workflow we compare results obtained from both HCR and LCR datasets.

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

P16

New insight into Aβ42 aggregation implicated in Alzheimer's disease: antagonist effects of S100A9 protein depending on monomeric and fibrillary forms

Jonathan Pansieri[1], Mai Nguyen Ngoc[2], Vytautas Smirnovas[2], Rodolphe Antoine[3], Ludmilla Morozova-Roche[1]
[1]Department of Medical Biochemistry and Biophysics, Umeå University, SE-901 87 Umeå, Sweden, [2]Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Vilnius University, LT-10257 Vilnius, Lithuania, [3]Laboratoire de Spectrométrie Ionique et Moléculaire, UMR CNRS 5579, Claude Bernard Lyon 1 University, France

Background: Amyloidoses are diseases characterized by self-aggregation of misfolded proteins into fibrillary forms. They are involved in many diseases, particularly concerning neurodegenerative diseases as Alzheimer's disease (AD), characterized by amyloid-beta peptide deposition (Aβ42). Thus, understanding of amyloid formation for diagnosis and treatment remains a critical challenge. Recently, numerous studies have shown a contribution of local inflammation to the AD development with a significant role of a small pro-inflammatory calcium-binding protein, so called S100A9 **[1]**. Moreover, high concentration of S100A9 plaques found in brain after traumatic brain injury indicates that this amyloid protein is an important factor in AD progression **[2]**. In this work, we investigate the effects of native and fibrillar S100A9 on Aβ42 fibrillization. This study aims at a better understanding on amyloid mechanisms and their cross-interactions for new prospects on AD management.

Questions addressed: How induced-proinflammatory amyloid products affects Aβ plaques?

Methods: ThT Fluorescence, AFM, mass spectrometry.

Results and discussion: S100A9 protein interacts with the Aβ42 peptide and can form different structures, as shown by mass spectrometry measurements. However, it remains unclear as whether S100A9 promotes or inhibits Aβ42 aggregation. On the one hand, monomeric S100A9 strongly inhibits Aβ42 fibrillization and induces heterogeneous amorphous structures, suggesting a blockage of amyloid nucleus formation. On the other hand, S100A9 fibrils catalyze Aβ42 fibrils formation in new fibrillary structures, by promotion of nucleation, considering the shape of kinetics curves. Finally, co-aggregation of these two proteins confirms coexistence of both effect, depending on time frame and S100A9 concentration. These results demonstrate the specific balance between the concentrations of native and fibrillar S100A9 species, which can affect drastically the Aβ42 aggregation reaction.

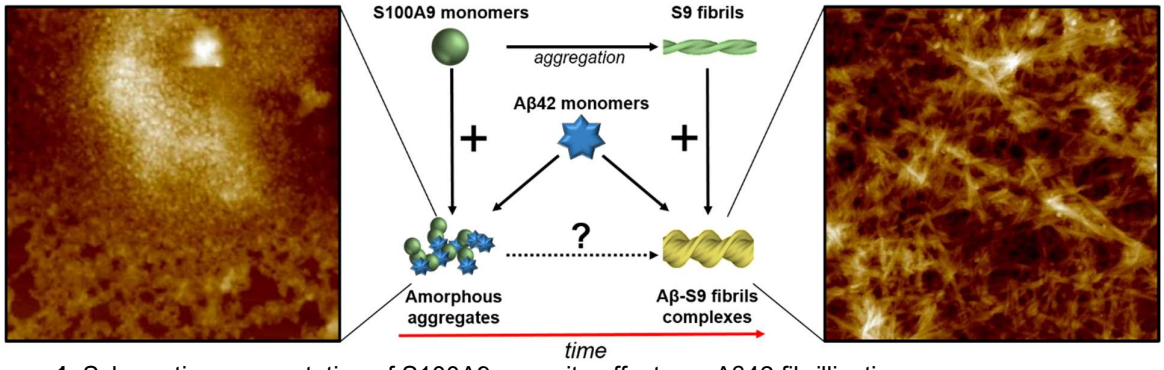


Figure 1. Schematic representation of S100A9 opposite effects on Aβ42 fibrillization

References:

1. Wang, C., et al. "The role of pro-inflammatory S100A9 in Alzheimer's disease amyloid-neuroinflammatory cascade." Acta neuropathologica 127.4 (2014): 507-522.
2. Engel, S., et al. "Dynamics of microglial activation after human traumatic brain injury are revealed by delayed expression of macrophage-related proteins MRP8 and MRP14." Acta neuropathologica 100.3 (2000): 313-322.

First mass and charge determination of entire amyloid co-aggregated populations involved in neurodegenerative diseases

Jonathan Pansieri[1], Hussein Fakhouri[2], Mai Nguyen Ngoc[3], Vytautas Smirnovas[3], Rodolphe Antoine[2], Ludmilla Morozova-Roche[1]

[1]Department of Medical Biochemistry and Biophysics, Umeå University, SE-901 87 Umeå, Sweden.[2]Laboratoire de Spectrométrie Ionique et Moléculaire, UMR CNRS 5579, Claude Bernard Lyon 1 University, France. [3]Department of Biothermodynamics and Drug Design, Institute of Biotechnology, Vilnius University, LT-10257 Vilnius, Lithuania.

Background: Amyloidoses are diseases characterized by self-aggregation of misfolded proteins into fibrillary forms. They are involved in many diseases, among them are neurodegenerative diseases such as Alzheimer's disease (AD), related to amyloid-beta peptide deposition (Aβ). Recently, numerous studies have shown a contribution of local inflammation to the AD development with a significant role played by a small pro-inflammatory calcium-binding protein S100A9 **[1]**. To gain knowledge about interaction and co-aggregation of these two amyloid polypeptides, charge-detection mass spectrometry (CD-MS), developed initially for characterization of large self-assembled systems such as nanoparticles, has been used in a straightforward manner on entire samples. Here we present the first CD-MS study on amyloid fibrillar co-aggregates, which enables us to distinguish an ensemble of the coexisting species with mass and charge distinction on newly acquired structural architectures.

Question addressed: How amyloid aggregation and co-aggregation of polypeptides into fibrillar species are reflected in their mass and charge distributions.

Methods: CD-MS, AFM.

Results and discussion: The following major issues are highlighted here: (1) Aβ42, S100A9 and their amyloid co-aggregates show unique patterns of mass and charge distribution as well as clustering of fibrils upon further incubation. After ca. 24 h incubation, S100A9 fibrils were characterized by lower masses with significantly lower charges, compared to the fibrils of Aβ42 (Fig. 1) and Aβ42-S100A9 mixture. (2) After 7 days, leading to fibrillar maturation and clustering, the fibrils of Aβ42-S100A9 mixture show lower molecular weight species (<150 MDa) compared to Aβ42 and S100A9 fibrils (>300 MDa). This reveals that the amyloid aggregates and S100A9 fibrils in particular undergo significant structural re-arrangements and maturation with aging. (3) After limited sonication, the mixed fibrils displayed population of smaller species (<50 MDa), compared to Aβ42 and S100A9 fibrils. This indicates that the mixed aggregates are likely less stable than individual fibrils. We plan to pursue these CD-MS studies further monitoring the specific fibrillization steps to provide unique insight into the whole ensemble of amyloid species.

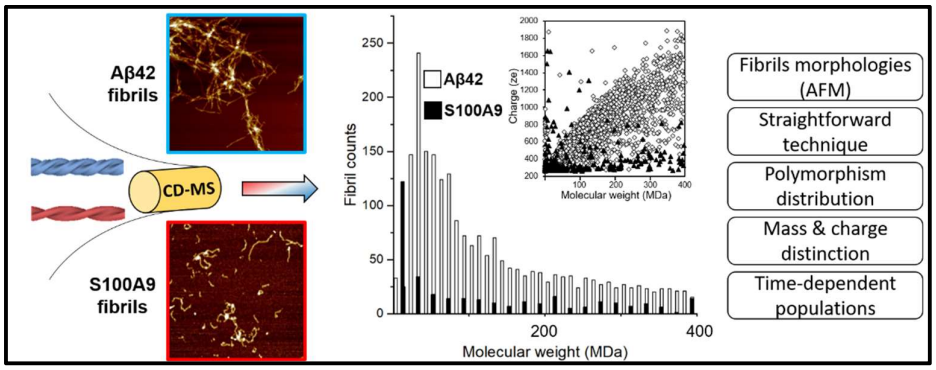


Figure 1. Schematic representation of the application of CD-MS experiments to analyse the mass and charge distributions on amyloid fibrils of S100A9 and Aβ42.

Reference:

1. Wang, C., et al. "The role of pro-inflammatory S100A9 in Alzheimer's disease amyloid neuroinflammatory cascade." Acta neuropathologica 127.4 (2014): 507-522.

P17

Application of helix-coil theory for determination of the IDP folding energetics

Hadži, San[1,2,3], Lah, Jurij[1]

[1]Department of Physical Chemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, 1000 Ljubljana, Slovenia, [2]Structural Biology Brussels, Department of Biotechnology, Vrije Universiteit Brussel, B-1050 Brussel, Belgium, [3]Molecular Recognition Unit, Structural Biology Research Center, Vlaams Instituut voor Biotechnologie, B-1050 Brussel, Belgium

Background: IDP are often observed to fold upon binding their targets. Affinities of such reactions are determined both by the IDP-target interactions (binding energy) as well as folding propensity of IDP.

Questions addressed: How can IDP folding energy contribution be estimated? Here, we present a method for determination of IDP folding energetics allowing delineation of the folding and binding contributions for the IDP-target interaction.

Results and discussion: The method is based on measuring circular-dichroism of the isolated IDP peptide, converting the signal into fractional helicity and calculating the free energy of folding using helix-coil theory. When measurements are preformed at different temperatures enthalpy of folding can be obtained, which can be used to calculate enthalpy of IDP-target binding. Such binding enthalpy (without contribution from folding) reflects specificity of IDP interactions and can be used to asses whether IDPs indeed uncouple binding specificity form the binding affinity.

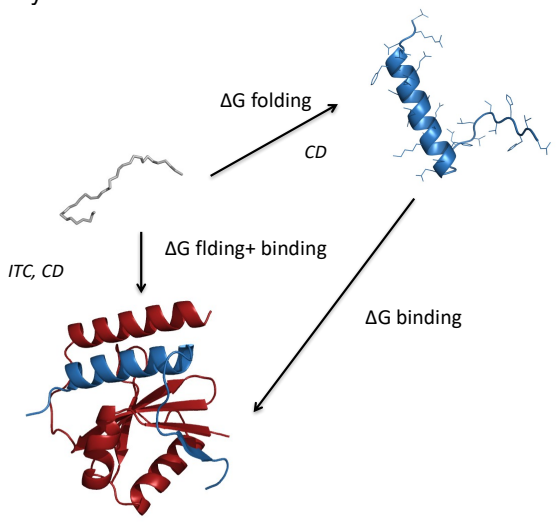


Figure 1. Separation of energetic contributions in folding upon binding reaction.

References:

1. Hadži S, Mernik A, Podlipnik Č, Loris R, Lah J. The Thermodynamic Basis of the Fuzzy Interaction of an Intrinsically Disordered Protein. Angew Chem Int Ed Engl. 2017 Nov 13;56(46):14494-14497.

P19

Interaction predictions enable a global view of the human protein–RNA interactome and its involvement in genetic disorders

Lang, Benjamin[1], Armaos, Alexandros[1], Tartaglia, Gian Gaetano[1]
[1]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, 08003, Spain.

Background: RNA-binding proteins are implicated in a number of neurodegenerative and neuromuscular genetic disorders in humans, with molecular mechanisms ranging from defects in splicing, localisation and translation to the formation of aggregates. Examples include heterogeneous and life-threatening genetic disorders such as amyotrophic lateral sclerosis (ALS), spinocerebellar ataxia and retinitis pigmentosa, among others[1]. Around 1,400 human proteins are known to bind RNA. These proteins contain one or more RNA-binding regions, either in the form of canonical globular domains or of more recently discovered, intrinsically disordered RNA interaction regions. Additionally, protein-protein interaction interfaces and even enzymatic active sites are sometimes employed for RNA binding. Protein–RNA interactions form an intricate network, and RNAs play structural roles in many types of phase-separated biological condensates, such as stress granules[2].

Questions addressed: Our research targets genetic disease predisposition via variants in the human protein–RNA interactome and is enabled by recent data on expression quantitative trait loci (eQTLs) and experimentally determined RNA–protein and RNA–RNA interactions. We are complementing these data with high-quality protein–RNA interaction predictions, expanding the human protein–RNA interactome in a genome-wide manner beyond experimental data, which is currently available for only around 170 of the ~1,400 known human RNA-binding proteins[3,4]. Additionally, we are developing a resource, RNAct (http://rnact.crg.eu), which makes available our novel protein-RNA interaction predictions and combines them with large-scale experimental protein-RNA and RNA-RNA interaction data, natural variation, disease-associated variants, as well as binding domains and regions and their preferred RNA sequence motifs.

Methods: To compute the interaction propensity scores, we used the catRAPID approach[5] with the fragmentation procedure[5,6]. For each protein–RNA pair, the fragment with the maximum interaction propensity score is used to assess overall binding ability.

Results and discussion: RNAct covers the human and mouse genomes and currently contains a total of 3.7 billion pairwise interactions, reflecting 10 years of computation time on the CRG's high-performance computing cluster. It combines experimentally identified interactions from ENCODE[3] with *ab initio* predictions[5-7], enabling full coverage of the known RNA-binding proteome. We stress that the method was trained on X-ray and NMR data and that its performance on the experimental eCLIP data (54,100 high-confidence interactions observed in two human cell lines with two replicates each, against a total of 2 billion protein–RNA pairs, ROC AUC=0.77) reflects its strong predictive power. This enables a global view of the human protein–RNA interactome and its involvement in genetic disorders.

References:

1. Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat Rev Genet *15*, 829–845.
2. Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. Nat Rev Mol Cell Biol *546*, 617.
3. Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods *13*, 508–514.
4. Hu, B., Yang, Y.-C.T., Huang, Y., Zhu, Y., and Lu, Z.J. (2017). POSTAR: a platform for exploring post-transcriptional regulation coordinated by RNA-binding proteins. Nucleic Acids Res *45*, D104–D114.
5. Bellucci, M., Agostini, F., Masin, M. and Tartaglia, G.G. (2011) Predicting protein associations with long noncoding RNAs. Nat. Methods, 8, 444–445.
6. Cirillo, D., Agostini, F., Klus, P., Marchese, D., Rodriguez, S., Bolognesi, B. and Tartaglia, G.G. (2013) Neurodegenerative diseases: quantitative predictions of protein-RNA interactions. RNA, 19(2), 129–40.
7. Agostini, F., Cirillo, D., Bolognesi, B., and Tartaglia, G.G. (2013) X-inactivation: quantitative predictions of protein interactions in the Xist network. Nucleic Acids Res., 41(1):e31.

P20

About the far UV absorption of peptide bond in disordered proteins

Skrabana Rostislav[1], Moravcsikova Jana[1], Cehlar Ondrej[1], Legenova Andrea[2]
[1] Institute of Neuroimmunology, Slovak Academy of Sciences, Bratislava, Slovakia, [2]AXON Neuroscience R&D Services SE, Bratislava, Slovakia

Background: Intramolecular environment of a disordered protein influences spectral properties of present chromophores, absorbing in the UV region. Especially, the most frequent chromophore in proteins – peptide bond – exhibits a profound dependence of its far UV absorption spectra on the conformation and the type of secondary structure present [1]. Recently, a method for determination of protein concentration based on the absorption at 205 nm has been published, using a calculator of the molar absorption coefficient directly from the knowledge of the protein sequence and known absorption coefficients of individual side chain chromophores [2]. For this purpose, the authors determined molar absorption coefficient of the peptide bond using a panel of prevalently globular proteins.

Questions addressed: The value of the previously determined molar absorption coefficient for the peptide bond at 205 nm has to be validated for disordered proteins.

Methods: The value of molar absorption coefficient of the peptide bond at 205 nm in disordered proteins was determined using a panel of various tau protein deletion mutants. The concentration of the calibrator tau proteins was determined using the absorbance at 280 nm and known absorption coefficient of tyrosine. Simultaneously, molar absorption coefficients of individual pure amino acids at 205 nm has been determined to revise the previously published values.

Results and discussion: We revised the value of molar absorption coefficient of peptide bond at 205 nm in fully disordered proteins, together with the absorption coefficients of individual amino acid side chains at the same wavelength. We proposed a modified calculator for molar absorption coefficient of disordered proteins at 205 nm. Determined concentration of disordered tau proteins were validated by using orthogonal biophysical methods.

References:

1. ROSENHECK, K., DOTY, P., 1961. The far ultraviolet absorption spectra of polypeptide and protein solutions and their dependence on conformation. In: Proc Natl Acad Sci 47, 1775–1785
2. ANTHIS, N.J., CLORE, G.M., 2013. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. In: Protein Science, 22, 851-858

P21

Bacterial secretion of repetitive surface fibers: inverse autotransport

Linke, Dirk[1], Wrobel, Agnieszka[1], Ottoni, Claudio[1], Leo, Jack[1]
[1]*Department of Biosciences, University of Oslo*

Background: Inverse autotransporters comprise the recently identified type Ve secretion system and are exemplified by intimin from enterohaemorrhagic Escherichia coli and invasin from enteropathogenic Yersiniae [1]. These proteins promote bacterial adhesion to host cells. They share a common domain architecture, with a lectin-like binding domain, and multiple Ig-like domains connecting this binding domain to the transmembrane domain that anchors the protein on the bacterial cell surface, and that exports the other domains in a process termed “autotransport” [2]. The Ig-like domains (ca. 100 amino acids) can be repeated more than 20 times in different exemplars of the Type Ve secretion systems.

Questions addressed: Our core interest is to understand host specificity in bacterial adhesion. To this end, we compare homologous repetitive adhesins from related species that infect different hosts. 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).

Methods: We use standard molecular biology and microbiology methods in this work. In the future, we hope to obtain high-resolution structures of adhesin molecules in complex with their receptors.

Results and discussion: 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 study [3] 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.

References:

1. The inverse autotransporter family: intimin, invasin and related proteins. Leo JC, Oberhettinger P, Schütz M, Linke D., Int J Med Microbiol. 2015 Feb;305(2):276-82.
2. Secretion of the Intimin Passenger Domain Is Driven by Protein Folding. Leo JC, Oberhettinger P, Yoshimoto S, Udatha DB, Morth JP, Schütz M, Hori K, Linke D., J Biol Chem. 2016 Sep 16;291(38):20096-112.
3. The repeat structure of two paralogous genes, Yersinia ruckeri invasin (yrInv) and a "Y. ruckeri invasin-like molecule", (yrIlm) sheds light on the evolution of adhesive capacities of a fish pathogen. Wrobel A, Ottoni C, Leo JC, Gulla S, Linke D., J Struct Biol. 2018 Feb;201(2):171-183.

P22

Effect of Curcumin Derivatives on $A\beta_{40}$ Fibrils: A Structure – Activity Relationship Study

Bednarikova, Zuzana[1], Orteca, Giulia[2], Ferrari, Erika[2], Asti, Mattia[2], Menziani, Maria Cristina[2], Gazova, Zuzana[1]

[1]Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, SK-041 67 Kosice, Slovakia, [2] Department of Chemical and Geological Sciences, University of Modena, IT-41125 Modena, Italy

Background: The presence of Aβ amyloid fibrils is one of the central pathological hallmarks in Alzheimer's disease. Alzheimer's disease is so far incurable disease and the effective treatments to stop its development are unsatisfactory. Therefore inhibition of aggregation and disruption of fibrils of Aβ peptides may be an effective strategy for combating this progressive disease [1]. One of the recent therapeutical strategies is focused on the inhibition and clearance of amyloid fibrils using small molecules. Mainly, planar compounds with aromatic rings, such as polyphenols, are extensively studied for their anti-amyloid properties [2].

Questions addressed: In our research we focused on curcumin derivatives as curcumin is polyphenol extracted from turmeric and is commonly used in traditional Asian medicine. It possess several beneficial properties such as antioxidant, anti-inflammatory and anti-cancer. Our objective was to explore the destroying activity of novel curcumin derivatives.

Methods: Thioflavin T assay and atomic force microscopy were used to evaluate the destroying activity of 12 curcumin derivatives. In order to gain some insights into binding mechanism of curcumin derivatives to $A\beta_{40}$ fibrils *in silico* methods were used. Cytotoxicity of studied compounds was tested using hippocampal HT-22 mouse cells.

Results and discussion: The results showed that curcumin derivatives are able to interfere with $A\beta_{40}$ fibrils depending on their structure. The structure-activity relationship data identified derivatives modified by adding the methoxy and acetate groups to curcumin scaffold as the most potent inhibitors. The least effective were curcumin derivatives with no substituent. The obtained DC_{50} values for best inhibitors were in low μM range. *In silico* calculation confirmed the *in vitro* results. Our results pointed out the anti-amyloid properties and also suggest high affinity of selected curcumin derivatives for $A\beta_{40}$ fibrils. Based on these results, we believe that radiolabeled curcuminoids could be potential diagnostic agents in means of PET imaging as well as potential therapeutics for AD.

References:

1. A. F. McKoy, J. Chen, T. Schupbach and M. H. Hecht, Chem Biol Drug Des, 84(5), (2014), 505-512.
2. L. M. Young, A. E. Ashcroft and S. E. Radford, Current Opinion in Chemical Biology, 39 (2017), 90–99.

P23

Study of the Huntingtin interactome

Brandi V.[1], Di Lella V.[1], Tortosa V.[1], Macari G.[1], De Simone G.[1], Di Masi A.(1), Marino M.[1], Ascenzi P.[1,2], Polticelli F.[1,3]

[1]Department of Sciences, Roma Tre University, [2]Interdepartmental Laboratory for Electron Microscopy , Roma Tre University, [3]National Institute of Nuclear Physics, Roma Tre Section

Background: An abnormal polyglutamine (polyQ) expansion in the N-terminal fragment of Huntingtin (Htt) protein leads to the devastating neurodegenerative disorder called Huntington's disease (HD). There has been an intense research aimed at understanding the molecular mechanisms underlying the deleterious effects of the presence of the aberrant polyQ expansion in Htt [1]. Through molecular modeling techniques, structural models of Htt-ordered domains have been obtained, predicting the presence of a previously undetected HEAT repeats region in the third ordered domain, dubbed Hunt3. Further, this domain has been predicted to be the putative region of Htt involved in the interaction with GTP-binding proteins [2]. Recently the structure of the full-length human Htt in complex with HAP40 (Htt-associated protein 40) has been determined to an overall resolution of 4 Å by cryo-electron microscopy, confirming the overall fold of Hunt3 [3].

Questions addressed: Is it possible to obtain further functional information on the third ordered domain of Htt? Can the analysis of the Htt interactome provide a global picture of the structure-function relationships of Htt?

Methods: Htt interacting proteins have been identified by PrePPI, IntAct, Hippie and Biogrid. Molecular docking simulations have been performed using ZDOCK. The 2000 complexes predicted by ZDOCK were re-ranked using ZRANK. Hunt3 domain is currently being expressed in recombinant form in *E. coli* for experimental validation of the predicted function.

Results and discussion: The analysis of Htt interactome has highlighted that among Htt-interacting proteins there are Rab proteins, members of the Ras superfamily of monomeric G proteins, and $G\alpha_o1$, a GTP-binding protein whose mutations cause a severe neurodevelopmental disorder, featuring symptoms similar to those observed in HD. Previous molecular docking simulations between Hunt3 and the two GTP-binding proteins showed that Rab11 and $G\alpha_o1$ localize in the same region of Hunt3. These interaction region in the cryo-electron microscopy structure of Htt is masked by an α-helix, stabilised by electrostatic interactions. Moreover, this α-helix contains two phosphorylation sites. Interestingly, molecular docking simulations carried out using the experimental structure of Hunt3 show that when the α-helix is removed from the structure, G proteins bind to Htt replacing the α-helix (**Figure 1**). To validate the hypothesis that the α-helix could act as a molecular switch for the binding of G proteins when phosphorylation occurs, molecular dynamics simulations are currently being performed. At the same time, expression and purification of Hunt3 domain are being carried out in order to verify the interaction site between Htt and $G\alpha_o1$.

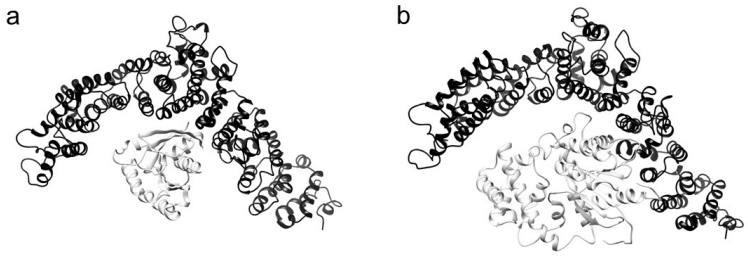


Figure 1. Structural complexes obtained by molecular docking simulations between the experimental structure of Hunt3 (in dark grey) and Rab11a (in light grey) in (**a**) and between the same structure of Hunt3 (in dark grey) and $G\alpha_o1$ subunit (in light grey) in (**b**).

References:

1. Saudou F. and Humbert S., 2016 *Neuron*, **89**(5):910-26;
2. Brandi V., Di Lella V., Marino M., Ascenzi P., Polticelli F. 2017, *J Biomol Struct Dyn* (Article in press)
3. Guo Q, Bin Huang, Cheng J, Seefelder M, Engler T, Pfeifer G, Oeckl P, Otto M, Moser F, Maurer M, Pautsch A, Baumeister W, Fernández-Busnadiego R, Kochanek S, 2018 *Nature,* **555**(7694):117-120.

P24

Towards a better understanding of the molecular mechanisms by which fuzzy regions affect the folding rate of adjacent molecular recognition elements

Longhi, Sonia[1], Troilo, Francesca[1,2], Bignon, Christophe[1], Bonetti, Daniela[2] and Gianni, Stefano[2]
[1]Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR 7257, Centre National de la Recherche Scientifique (CNRS) & Aix-Marseille University, France; [2]Department of Biochemical Sciences, La Sapienza, University of Rome, Italy

Background: In spite of the partial disorder-to-order transition that intrinsically disordered proteins often undergo upon binding to their partners, a considerable amount of residual disorder may be retained in the bound form, thus resulting in a fuzzy complex. Fuzzy regions flanking molecular recognition elements (MoREs) may enable partner fishing through non-specific, transient contacts, thereby facilitating binding, but may also disfavor binding through various mechanisms. The complex between the intrinsically disordered C-terminal domain of the measles virus (MeV) nucleoprotein (N_{TAIL}, aa 401-525) and the C-terminal X domain (XD) of the viral phosphoprotein is an illustrative example of fuzziness. Indeed, although binding to XD triggers α-helical folding within a MoRE encompassing residues 486-502, the N_{TAIL} region preceding the MoRE remains conspicuously disordered. In a previous study we showed that the removal of N-terminal fuzzy region results in increased binding towards XD [1]. Kinetics studies between XD and either a peptide mimicking the MoRE or full-length N_{TAIL}, suggested that the fuzzy appendage reduces the rate of folding of N_{TAIL} [1].

Questions addressed: Which are the molecular mechanisms by which the fuzzy region reduces the rate of folding of the adjacent MoRE? Does the rate of folding vary linearly with increasing removal of the fuzzy appendage? Does this rely on a purely entropic mechanism? Is this sequence-specific?

Methods: Rational design of single-site conservative N_{TAIL} variants; protein expression & purification; fluorescence-based kinetics studies (Temperature-jump experiments).

Results and discussion: Using kinetics, we studied the interaction between XD and either a series of MeV N_{TAIL} truncated variants, or an artificial (full-length) N_{TAIL} variant. By comparing the kinetics obtained with the set of truncated variants, we found that the gradual shortening of the long fuzzy appendage results in an increase in the rate of folding. In particular, the dependence of k_{obs} upon ligand concentration switches from hyperbolic to linear. An artificial N_{TAIL} sequence, in which the MoRE is unaltered but the region preceding the MoRE differs profoundly from the *wild-type* (*wt*) sequence (identity of 6%), yields the same kinetics as *wt* N_{TAIL}, suggesting that the fuzzy appendage dampens the rate of folding through a purely entropic mechanism. To ascertain the sequence-independent nature of this self-inhibition, we have targeted for site-directed mutagenesis the 450-480 region of N_{TAIL}. We have designed 19 single-site, conservative N_{TAIL} variants and have characterized their kinetic behavior towards XD. Results argue for a combination of entropy and enthalpy in the regulation of the rate of folding of N_{TAIL} by the fuzzy appendage.

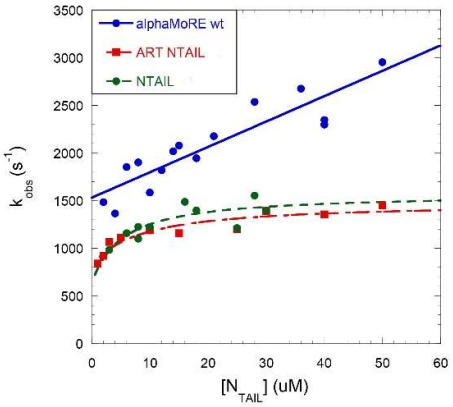


Figure 1. Kinetics of binding between XD and either N_{TAIL}, or artificial N_{TAIL} or a peptide mimicking the MoRE.

References:

1. Gruet A, Dosnon M, Blocquel D, Brunel J, Gerlier D, Das RK, Bonetti D, Gianni S, Fuxreiter M, Longhi S and Bignon C. (2016) Fuzzy regions in an intrinsically disordered protein impair protein–protein interactions***. FEBS J***. 283(4):576-94.

P25

Defining a new intrinsically disordered BH3 motif

Emiliano Maiani[1], Matteo Lambrughi[1], Ágota Toth[1], Mads Nygaard[1], Birthe Brandt Kragelund[2] and Elena Papaleo[1,3].

[1]Computational Biology Laboratory, Danish Cancer Society Research Center, Copenhagen 2100, Denmark.[2]Department of Biology, University of Copenhagen, Copenhagen 2200, Denmark.
[3]Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen 2200, Denmark.

Background: The Bcl-2 family of proteins plays a key role in the control of apoptotic pathways and mitochondrial physiology[1]. This class of proteins is conventionally divided in 2 subgroups, depending on the presence of different BH motifs on their structures: i) prosurvival proteins such as Bcl-2 and Bcl-xL and proapoptotic proteins Bax and Bak share four different BH motifs (BH1-4), ii) BH3-only proteins such as DP5, Puma, Bad, and Bix contain only the BH3 motif and are well known inhibitors of the prosurvival Bcl-2 proteins. BH3 motifs have been traditionally defined as an α helix structure but emerging evidence are pointing at BH3 motifs as Short Linear Motifs (SLiMs)[2], which generally occur in intrinsically disordered proteins. On this context, it has been demonstrated that the highly disordered protein p14ARF interacts with BCLxL through a short motif[3], predicted as intrinsically disordered, which could represent a new unconventional BH3 motif.

Questions addressed: We aim at characterize the structural properties of the p14ARF BH3 motif, in the unbound state and in complex with BCLxL. The discovery of a new and uncanonical BH3 motif will represent a starting point for the identification of a new class of BH3 intrinsically disordered motifs. Through the characterization of p14ARF interaction with BCLxL we will define the contribution of BH3 residues involved in the complex formation. This information will be relevant in the context of the design of new BH3 mimetics.

Methods: To achieve these goals we are employing an orthogonal strategy based on structural and computational biology. We have collected a set of different simulations with different force fields of the free pre-bound state of the p14ARF peptide, which has been shown to interact with BCLxL. In parallel, we are evaluating the single contribution of p14ARF amino acids, and of cancer related selected mutations, in the binding with BCLxL by means of peptide array. We are also using chemical shift perturbation experiments of the unbound versus bound forms of the p14ARF BH3 domain to identify the interaction interface between the two molecules.

Results and discussion: Through molecular simulations we have been able to identify the φ and θ torsional angle distribution of p14ARF peptide in the pre-bound state. Results of this analysis indicate that a portion of this peptide is not structured in α helix. Through peptide array we have defined the contribution of some mutations on p14ARF/BCLxL interaction. As a control, we designed peptides by introducing helical breakers prolines in the sequence of conventional α helix structured BH3 of Bim1. Interestingly, we were able to detect binding between these peptides and BCLxL, supporting the notion that BH3 motifs do not necessary have to be α helix structured to interact with Bcl2 family of prosurvival proteins.

References:

1. Shamas-Din,A. *et al.* (2013) Mechanisms of Action of Bcl-2 Family Proteins. *Cold Spring Harb Perspect Bio.*
2. Aouacheria, A. *et al.* (2015) Redefining the BH3 Death Domain as a “Short Linear Motif.” *Trends Biochem. Sci.*
3. Christensen, C. *et al.* (2014) A short acidic motif in ARF guards against mitochondrial dysfunction and melanoma susceptibility. *Nat. Commun.*

Acknowledgments:
The prokaryotic expression vector for BCL-xL was kindly provided by Richard W. Kriwacki (St. Jude Children’s Research Hospital). This work is supported by LEO Foundation Exploratory grant n. LF17006 and HPC PRACE-DECI14th CancerBH3.

P26

The effect of ionic liquids on poly/peptide fibrillization

Fedunova Diana, Vanik Vladimir, Bednarikova Zuzana, Antosova Andrea, Gazova Zuzana
Institute of Experimental Physics, Slovak Academy of Sciences, Kosice, Slovakia

Background: One of the important area of the study of amyloid aggregation is directed to the finding the strategies for controlling the self-assembly process *in vitro.* Formation of *s*tructurally and morphologically defined oligomers and fibrils can be utilized across diverse fields dealing with amyloid aggregation from understanding its general mechanism to developing drugs against amyloid-related diseases or biotechnological applications. The important roles in controlling the amyloid aggregation *in vitro* play solvent conditions. Ionic liquids (ILs) are new class of media, which could be tailored to solvents with desired properties for applications in multiple fields. ILs consist of organic cations and anions, are liquid bellow 100^{o}C and can be used as neat solutions or dissolved in water or other solvents. Several studies have shown that ILs can effectively promote but also inhibit process of self-assembly of intrinsically disordered as well as globular poly/peptides [1,2]. Characterization of the relationship between physico-chemical properties of ILs and the kinetics of amyloid fibrillization or morphology of mature fibrils could be beneficial in extending our knowledge of the amyloid fibrillization process. In this work we have studied the effect of ILs with imidazolium cation and anions of Hofmeister serie on kinetics and morphology of Aβ_{40} peptide and globular proteins insulin and lysozyme.

Questions addressed: The aim of this work is to identify ILs capable to affect the process of fibrillization of Aβ_{40} peptide, lysozyme and insulin as well as to determine the relation between physico-chemical properties of ILs (the position of anion in Hofmeister serie, hydrophobicity, size, charge density) and the kinetics of fibrillization.

Methods: The kinetics of amyloid fibrilization was measured by thioflavin T (ThT) binding assay. ThT fluorescence intensity is sensitive to the presence of beta-sheet structures organized in amyloid fibrils, the intensity increase is in correlation with the amount of fibrils. For further detection of fibrils we have used Congo red (CR) spectrophotometric assay. The secondary structure content was determined by CD spectroscopy. Amyloid fibrils were visualized by AFM microscopy.

Results and discussion:
The kinetics of fibrillization of Aβ_{40} peptide were examined at neutral conditions (pH 6.9, 25 μM peptide and 37 ^{o}C) leading to formation of mature fibrils within 7 days of incubation. Addition of 100 mM 1-ethyl-3methyl-imidazolium-bis(trifluoro methyl sulfonyl) imide (EMIM TFSI) has led to the acceleration of fibrillization by shortening the lag phase, mature fibrils are already present after 24 h of incubation. This hydrophobic ILs probably facilitates the association of hydrophobic surfaces of Aβ_{40} peptide and thus promote the fibrillization. Next, we have studied the effect of hydrophilic ILs, with kosmotropic anion hydrogen sulfate (EMIM HSO_4) and smaller chaotropic anion tetrafluoroborate (EMIM BF_4). The fibrillization of Aβ_{40} peptide was promoted by both ILs However, presence of chaotropic anion has the weakest effect on Aβ_{40} peptide kinetics as the lag phase was 96 hours whereas kosmotropic anion led to shorter lag phase (48 h). As a comparison, we have studied the effect of these ILs on kinetics of fibrillization of lysozyme at acidic conditions (pH 2.7, 65^{o}C) and insulin (pH 1.6, 50 ^{o}C). We have found that the ILs with hydrophobic TFSI anion enhanced interactions of partially unfolded globular structures of both proteins with exposed hydrophobic core leading to formation of amorphous aggregates. The fibrillization of lysozyme was accelerated in studied hydrophilic ILs even in those with kosmotropic anion. The lag phase and polymeziation rate obeyed Hofmeister series. The kinetics of insulin fibrillization was also accelerated in studied hydrophilic ILs, but the lag phase and polymerization rate did not depend on the type of used ILs at studied concentration. We have found that ILs are able to modulate fibrillization of Aβ_{40} peptide, insulin and lysozyme, reflecting different driving forces responsible for fibrillization of structurally different poly/peptides. The effect of Hofmeister serie of anions in ILs on protein fibrillization is complex and cannot be easily predicted as it is in a case of addressing their effect on protein stability.

Acknowledgement
This work was supported by the research grant from the Slovak Grant Agency Vega No. 2/0030/18, 2/0145/17 and MVTS COST 083/14 action BM1405.

References:
1. Bae, S.Y., Kim, S., Hwang, H., Kim, H.K., Yoon, H.C., Kim, J.H., Lee, S., Kim, T.D. BBRC 400 (2010) 531-536.
2. Kalhor, H.R., Kamizi, M., Akbari, J., Heydari, A. Biomacromolecules. 10 (2009) 2468-2475.

P27

Capturing transient conformations and local dynamics of intrinsically disordered microtubule-associated protein 2c

Melkova, Katerina[1,2], Vojtech, Zapletal[1,2], Jansen, Severine[1,2], Zidek, Lukas[1,2], Hritz, Jozef[1], Jensen, Malene R.[3], Blackledge, Martin[3]

[1]Masaryk University, Central European Institute of Technology, Kamenice 5, 62500 Brno, Czech Republic, [2]Masaryk University, Faculty of Science, National Centre for Biomolecular Research, Kamenice 5, 62500 Brno, Czech Republic, [3]University Grenoble Alps, CEA, CNRS, Institut de Biologie Structurale, Grenoble 38000, France

Background: The main focus of the presented study is microtubule-associated protein 2c (MAP2c), an intrinsically disordered protein (IDP) regulating structure and dynamics (polymerization and depolymerization) of microtubules (MTs), which is essential for growth and development of neurons. MAP2c and its close homologue Tau are expressed in neurons, with MAP2c localized in dendrites, whereas Tau is found mainly in axons. MAP2c is a 49 kDa protein consisting of several structural and functional regions. The N-terminal domain contains two important segments: The 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 interacts with SH3 domains, including a non-canonical SH3 domain of plectin [1]. The second important part of MAP2c is a highly conserved C-terminal domain that binds to MTs. The C-terminal part of Tau and MAP2c is homologous, but the N-terminal domains differ. Interactions of MAP2c with microtubules have been described [2] but, overall, much less is known about MAP2c than about Tau, and the results of various studies are partly contradictory.

Questions addressed: Experimental characterization of the structure and dynamics of individual MAP2c domains, their role in the protein function, and their interactions with potential interaction partners.

Methods: Given their disordered nature, nuclear magnetic resonance (NMR) is a key experimental method for studying IDPs. We used high-resolution NMR and small angle X-ray scattering (SAXS) to acquire atomic-resolution data reflecting structural and dynamic features of MAP2c. The obtained experimental data were converted into a structural model using the ASTEROIDS algorithm. NMR was also used to monitor binding to the interaction partners of MAP2c.

Results: Assigned NMR chemical shifts [3], paramagnetic NMR relaxation enhancement (PRE) data [4], and SAXS data [4] served as an experimental input, reflecting local conformation, long-range contacts, and overall shape of the molecule, respectively. The results allowed us to correlate structural features and dynamics of MAP2c with its known and proposed binding and phosphorylation sites, and to directly compare MAP2c with Tau [3], [4]. We obtained a detailed description of the transient secondary structure of MAP2c. We also identified intramolecular contacts, caused by electrostatic interactions, with a great impact on dynamics of MAP2c. Moreover, NMR relaxation data revealed lower flexibility and increased populations of specific conformations in regions of MAP2c known (or expected) to interact with other proteins. This indicates that the local conformational propensities may identify protein regions that play an important role in the protein partner recognition.

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. L. Dehmelt and S. Halpain, “The MAP2/Tau family of microtubule-associated proteins,” *Genome Biol.*, p. 10, 2004.
3. 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.
4. K. Melková *et al.*, “Functionally specific binding regions of microtubule-associated protein 2c exhibit distinct conformations and dynamics,” *J. Biol. Chem.*, p. jbc.RA118.001769, Jun. 2018.

P28

Dissecting the ataxin-3 aggregation pathway

Silva, Alexandra*[1], Figueiredo, Francisco*[1], Almeida, Ana[1], Martins, Pedro M. [1,2], Macedo-Ribeiro, Sandra[1]
[1]IBMC - Instituto de Biologia Molecular e Celular and i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal, [2]ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal

Background: Ataxin-3 (Atx3) is an intracellular protein with a polyglutamine (polyQ), which is involved in numerous protein quality control pathways. This multimodular protein contains a globular Josephin domain (JD), displaying an ubiquitin hydrolase fold, and a C-terminal tail composed of Ubiquitin Interaction Motifs (UIMs) and the repeated polyQ segment. Machado-Joseph disease is a highly incapacitating neurodegenerative disorder caused by expansion the polyQ stretch [1]. Aggregation of the non-pathological and polyQ-expanded Atx3 is well characterized and is critically dependent on early self-assembly events modulated by the JD. PolyQ expansion beyond a certain threshold elicits a second (polyQ-dependent) aggregation step that is critical for fibril maturation [2,3].

Questions addressed: Biophysical studies unveiled the Atx3 multistep aggregation pathway, but structural and mechanistic details of the aggregation pathways as well as the characteristics of the self-assembly intermediates is still lacking.

Methods: To provide a time-resolved perspective on Atx3 aggregation pathway(s), the self-assembly of Atx3 was monitored by Dynamic Light Scattering (DLS) to follow the size distribution of the different particles in solution, and by Thioflavin-T binding kinetics to track the formation of amyloid species. At selected time-points during the aggregation pathway, the oligomerization state of the soluble species was analyzed by size exclusion chromatography, while the morphological features of the different species were evaluated by Transmission Electron Microscopy (TEM). Those studies were combined with theoretical modelling [4] and experimental determination of equilibrium dissociation constants to provide an unprecedented quantitative perspective of the Atx3 aggregation mechanisms.

Results and discussion: The combined experimental theoretical approaches identify deviations from the simple nucleation-polymerization mechanism and suggest the presence of Atx3 aggregation pathways parallel to amyloid fibrillation, despite the clearly high thermodynamic tendency for amyloid fibril formation. Accordingly, Atx3 oligomers act as a reservoir of soluble protein that can be later dissociated into monomers to promote the formation of insoluble filaments. The oligomerization rate constants estimated from the current aggregation model, allow the estimation of a monomer-to-oligomer dissociation constant in the μM range, in agreement with experimental data obtained for Atx3 and the isolated JD. Based on the current model of aggregation, the relative contribution of different Atx3 protein domains has been evaluated from the analysis of protein aggregation progress curves and kinetic scaling laws. In addition, this knowledge is critical to understand the mechanisms underlying the modulatory effect of different Atx3 binding molecules (nanobodies and small molecules) for on- and off- pathway aggregation steps.

This work was funded 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); and by FEDER through Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement in the framework of Project Norte-01-0145-FEDER-000008 (Porto Neurosciences and Neurologic Disease Research Initiative at I3S). FF is the recipient of a FCT PhD fellowship (SFRH/BD/133009/2017). The support of Bastian Groitl and Nena Matscheko (Dynamic Biosensors GmbH (HQ), Germany) on the analysis of Atx3 interactions is greatly acknowledged. AA acknowledges funding by the Amyloidosis Foundation.

References:

1. Silva A, de Almeida AV, Macedo-Ribeiro S. Polyglutamine expansion diseases: More than simple repeats. J Struct Biol. 2018; 201:139-154. doi: 10.1016/j.jsb.2017.09.006.
2. Scarff CA, Almeida B, Fraga J, Macedo-Ribeiro S, Radford SE, Ashcroft AE. Mol Cell Proteomics. 2015;14:1241-53. doi: 10.1074/mcp.M114.044610.
3. Lupton CJ, Steer DL, Wintrode PL, Bottomley SP, Hughes VA, Ellisdon AM. Enhanced molecular mobility of ordinarily structured regions drives polyglutamine disease. J Biol Chem. 2015; 290:24190-200. doi: 10.1074/jbc.M115.659532

4. Silva A, Almeida B, Fraga JS, Taboada P, Martins PM, Macedo-Ribeiro S. Distribution of Amyloid-Like and Oligomeric Species from Protein Aggregation Kinetics. Angew Chem Int Ed Engl. 2017; 56:14042-14045. doi: 10.1002/anie.201707345.

P29

Self-Analysis of Repeat Proteins Reveals Conserved Sequence Patterns

Merski, Matthew[1], Młynarczyk, Krzysztof [1], Skrzeczkowski, Jakub[1], Dunin-Horkawicz, Stanisław[2] and Górna, Maria W.[1]

[1]Biological and Chemical Research Centre, Department of Chemistry, University of Warsaw, Poland,
[2]Centre of New Technologies, University of Warsaw, Poland [2]

Background: Protein repeats may occur in as many as 25% of all proteins [1] and while they are more common in eukaryotic than in prokaryotic proteins, 81% of archaeal and 96% of bacterial taxa contain at least one repeat protein [2], although differences in functions between eukaryotic and prokaryotic repeat proteins have suggested that they may have had separate evolutionary origins [3]. Furthermore, the repetitive nature and apparent low complexity of their sequence make repeats difficult to analyze. For example, a recent study suggested that 25 out of 95 of a set of proteins previously identified as containing three-helix armadillo repeats actually contained two-helix HEAT repeats, possibly mis-identified due to a common evolutionary origin of the two kinds of repeats [4]. The low-complexity of repeat sequences also adds to the complexity of their analysis [5]. Combined, these issues make it difficult to determine if similar repeats are related or simply the result of convergent evolution.

Questions addressed: We are interested in what kind of information is being conserved in repeat protein sequences.

Methods: We examined repeat protein sequences using the self-comparison program DOTTER to generate dot plots and then compared the distinct patterns of proteins.

Results and discussion: Our analysis showed that the dot plot patterns produced by DOTTER were unique to individual, related proteins (Fig. 1 A&B). These patterns allowed comparisons between different proteins to be made using a simple Jaccard relationship (J_X). Analysis and comparisons were quick and efficient; the entire RepeatsDB set of repeat proteins (>6000 chains) was analyzed in less than an hour. The DOTTER plots were shown to be maintained in related proteins, in some cases even over hundreds of millions of years despite the fact that he patterns decayed rapidly in the absence of selective pressure. In a test set of 79 repeat proteins, J_X values were reduced by half after 8.5% of residues had been mutated, although highly similar DOTTER patterns were observed in proteins in which every type of amino acid had been mutated to a homologous residue maintaining the patterns in the sequence (Fig. 1 A & C). These results suggest that relationships indicated by these dot plots show a sequence “blueprint” for the specific structure and function of the repeat protein.

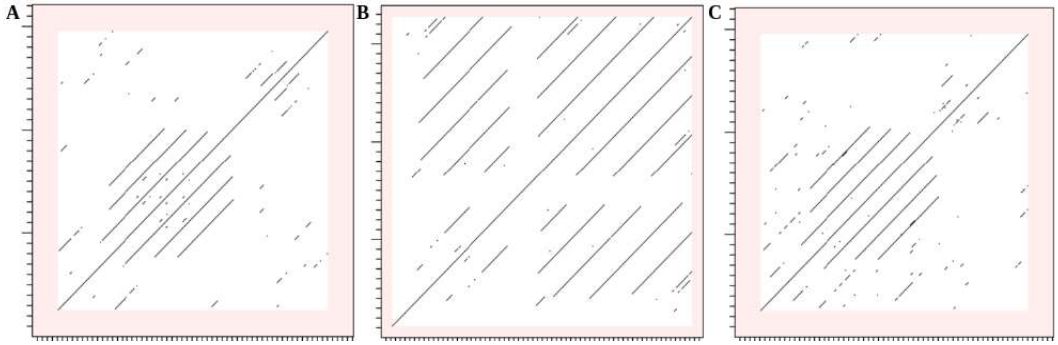


Figure 1. DOTTER plots for A) B. burgdorferi OspA B) M. musculus G-protein modulator 2 C) sequence permuted B. burgdorferi OspA with J_X = 0.55 to the normal protein

References:

1. Pellegrini, Renda, & Vecchio, Ab initio detection of fuzzy amino acid tandem repeats in protein sequences. *BMC Bioinformatics* **13,** S8 (2012)
2. Jernigan & Bordenstein Tandem-repeat protein domains across the tree of life. *Peerj* **3,** e732 (2015)
3. Kajava, Proteins with repeated sequence - Structural prediction and modeling. *J Struct Biol.* **134,** 132 (2001).
4. Gul *et al.* Metazoan evolution of the armadillo repeat superfamily. *Cell Mol Life Sci.* **74,** (2017)
5. D'Andrea & Regan TPR proteins: the versatile helix. *Trends Biochem Sci* **28,** 655 (2003)

P30

Compound CID 9998128 - a potential multi-target drug for Alzheimer's disease

Gancar, Miroslav[1], Thai, Nguyen Quoc[2,3,4], Bednarikova, Zuzana[1], Linh, Huynh Quang[4], Hu, Chin-Kun[5,6,7,8], Li, Mai Suan[9] and Gazova, Zuzana[1]

[1]Department of Biophysics, Institute of Experimental Physics, Slovak Academy of Sciences, Watsonova 47, Kosice, Slovakia, [2]Institute for Computational Sciences and Technology, Ho Chi Minh City, Vietnam, [3]Dong Thap University, 783 Pham Huu Lau Street, Ward 6, Cao Lanh City, Dong Thap, Vietnam, [4]Biomedical Engineering Department, University of Technology -VNU HCM, 268 Ly Thuong Kiet Str., Ho Chi Minh City, Vietnam, [5]Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan, [6]Physics Division, National Center for Theoretical Sciences, Hsinchu 30013, Taiwan, [7]Department of Physics, National Dong Hwa University, Hualien 97401, Taiwan, [8]Department of Systems Science, University of Shanghai for Science and Technology, Shanghai 200093, China
[9]Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

Background: Alzheimer's disease (AD) is an irreversible neurodegenerative disorder that slowly impairs memory and cognitive skills. It is the most common cause of dementia among elders and there is currently no treatment to halt the neurodegenerative process. AD is a very complex and multifactorial disease characterized by the accumulation of amyloid plaques and neurofibrillary tangles in brain accompanied, among others, by the loss of cholinergic function important for cognition and homeostasis dysregulation [1].

Questions addressed: One of the generally favored approaches to address such complex disease is a use of small compounds designed to target specific structures connected to AD pathology. Several studies argued that better approach in contrast to "one drug one target" concept is necessary. Novel multi-target-directed ligand (MTDL) strategy is based on assumption that a single compound is able to interact with multiple targets [2]. Therefore, we have probed capability of the small compound CID 9998128 as a potential multi-target drug for the AD using *in vitro* and *in silico* experiments.

Methods: *In vitro* experiments used thioflavin T (ThT) fluorescence assay, BACE-1 FRET assay and atomic force microscopy (AFM) to determine anti-amyloid properties of CID 9998128 and to visualize quantitative and morphological changes in samples. Using *in silico* methods, the all-atom docking simulation and molecular dynamics simulation, the binding affinity and binding free energy were calculated, and also the role of van der Waals and electrostatic interactions was established.

Results and discussion: *In vitro* experiments have shown that CID 9998128 inhibits the amyloid beta 42 ($A\beta_{42}$) amyloid fibrillization and is capable to clear $A\beta_{42}$ fibrils. Moreover, this compound dose-dependently decreases β-site amyloid precursor protein cleaving enzyme (BACE-1) activity (Fig. 1). By *in silico* methods, we have demonstrated that this compound strongly binds to $A\beta_{42}$ monomer, dimer, fibrils and BACE-1. It was determined that van der Waals interaction is dominating over the electrostatic interaction in binding affinity. Thus, our study has revealed that CID 9998128 is a good candidate for AD treatment by preventing the production of $A\beta_{42}$ peptides and degrading their amyloid aggregates [3].

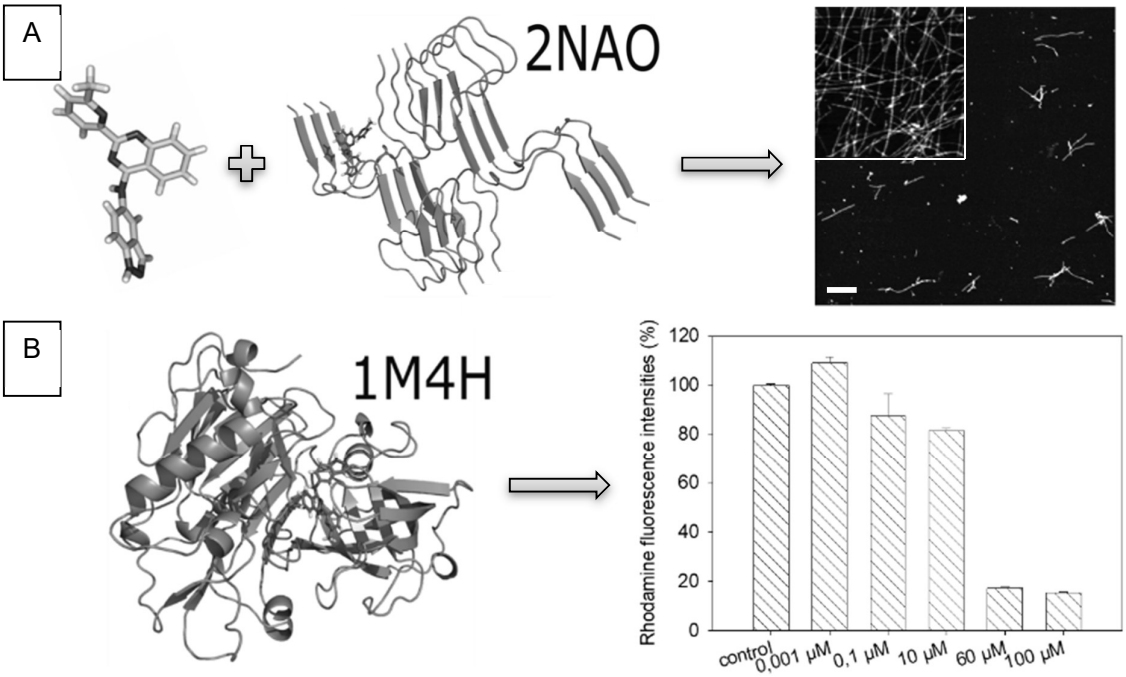


Figure 1. A) Strong binding of the compound CID 9998128 with $A\beta_{42}$ fibrils (2NAO) leading to extensive depolymerization of amyloid fibrils as confirmed by AFM. Representative AFM image of $A\beta_{42}$ fibrils [5 μM] before (inset) and after treatment with CID 9998128 [100 μM]. Bar represents 1 μm. B) A concentration-dependent decrease in rhodamine fluorescence intensity in response to lowered BACE-1 (1M4H) enzymatic activity, due to interaction with CID 9998128.

Acknowledgment: *This work was supported by the research grant projects in a frame of VEGA 2/0145/17, 2/0030/18, MVTS COST 083/14 action BM1405 and SAS-MOST JRP 2015/5. This research was supported by Institute for Computational Science and Technology (grant No. 13/2017/HĐ-KHCNTT), and Department of Science and Technology at Ho Chi Minh City, Vietnam. C.K.H was supported by grants MOST 105-2112-M-001-004 and MOST 106-2112-M-001-027.*

References:

1. Xu, Y., Yan, J., Zhou, P., Li, J., Gao, H., Xia, Y., Wang, Q. (2012) Neurotransmitter receptors and cognitive dysfunction in Alzheimer's disease and Parkinson's disease. *Prog. Neurobiol*. 97, 1-13.

2. Gazova, Z., Soukup, O., Sepsova, V., Siposova, K., Drtinova, L., Jost, P., Spilovska, K., Korabecny, J., Nepovimova, E., Fedunova, D. (2017) Multi-target-directed therapeutic potential of 7-methoxytacrine-adamantylamine heterodimers in the Alzheimer's disease treatment. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1863, 607-619.

3. Thai, N. Q., Bednarikova, Z., Gancar, M., Linh, H. Q., Hu, C. K., Li, M. S., Gazova, Z. (2018). CID 9998128 compound is a potential multi-target drug for Alzheimer's disease. *ACS Chem. Neurosci*.

P31

AMYCO: Evaluation of mutational impact on prion-like proteins aggregation propensity

Valentin Iglesias[1], Oscar Conchillo-Sole[1], Cristina Batlle[1] and Salvador Ventura[1]
[1] *Institut de Biotecnologia i Biomedicina and Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Bellaterra, 08193, Spain*

Background: Around 1% of the human proteome has been predicted to contain a disordered and low complexity prion-like domain (PrLD) [1]. Mutations in PrLDs might promote a transition to an amyloid-like, aggregation-prone state linked to disease [2].

Questions addressed: Could we apply current prion-domain based knowledge to correctly predict mutation impact on the aggregation propensity of human prion-like proteins? Do these predictions match human disease-related mutations?

Methods: Prion predictors rely on either scanning for the biased amino acid composition of their PrLD or instead specific sequential features facilitating their transition to amyloid-like states. We have recently shown that a function that considers both the effects of mutations on PrLDs composition and in localized amyloid propensity can approach their impact on intracellular protein aggregation [3]. We introduce the AMYCO (combined AMYloid and COmposition based prediction of prion-like aggregation propensity) web server which implements this approach to allow automated and fast predictions.

Results and discussion: The AMYCO algorithm outperforms the state-of-the-art predictors [3] for a pool of punctual and multiple mutations or deletions for hnRNPA2 (Figure 1). Moreover, AMYCO classifies correctly mutations promoting the apparition of a *de novo* prion-like behaviors on non-prionic PrLD's of yeast proteins and predicts an increase in aggregation propensity behind disease-linked mutations in different prion-like proteins such as hnRNP D0/AUF1 or hnRNP DL. AMYCO has been developed as a web application to assess the impact of mutations on the aggregation propensity of prion-like proteins, allowing a fast and accurate evaluation of the effect of disease-associated mutations; as well as engineering novel variants with designed aggregation propensities. AMYCO does not require previous registration and is freely available to all users at: http://bioinf.uab.es/amyco/.

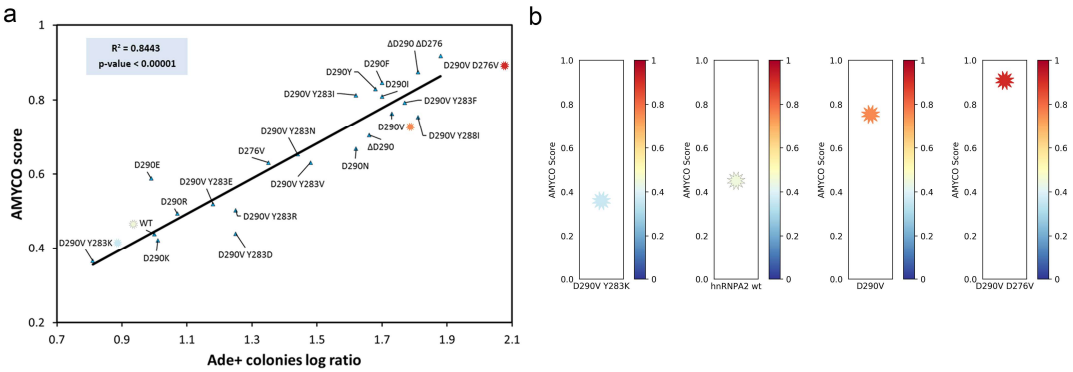


Figure 1. Correlation between AMYCO predictions and the aggregation propensity of human hnRNPA2 prion-like protein. a. Graphic representation of the correlation between the variants AMYCO scores and their ability to form Ade+ colonies when expressed in yeast, a direct reporter for their aggregation propensity [3]. **b.** AMYCO output of hnRNPA2 prion-like protein variants.

References:

1. King, O. D., Gitler, A. D. & Shorter, J. (2012) The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease, *Brain Research.* **1462**, 61-80.
2. Patel, A., Lee, H. O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M. Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T. M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, L. A., Weigert, M., Myers, E. W., Grill, S., Drechsel, D., Hyman, A. A. & Alberti, S. (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation, *Cell.* **162**, 1066-77.
3. Batlle, C., Fernandez, M. R., Iglesias, V. & Ventura, S. (2017) Perfecting prediction of mutational impact on the aggregation propensity of the ALS-associated hnRNPA2 prion-like protein, *FEBS letters*.

P32

Evolutionary study of the evolution of homorepeats in protein sequences

Mier, Pablo[1], Andrade-Navarro, Miguel A.[1]
[1]Computational Biology and Data Mining, Faculty of Biology, Johannes Gutenberg University of Mainz, 55128 Mainz, Germany

Background: Homorepeats (or polyX) in protein sequences are stretches of consecutive repetitions of a single amino acid residue. The length of a repeat determines its function but this length is variable from protein to protein making it hard to identify when a repeat becomes functional. Codon usage is another variable that could be indicative of homorepeat functionality.

Questions addressed: To facilitate the study and detection of homorepeats, we are creating resources for their annotation (dAPE [1]; Figure 1), and we are studying their context and properties within protein families. In particular, we recently studied the evolution of codon usage in polyQ in primates [2].

Results and discussion: Our studies of the use of homorepeats suggest that homorepeats modulate protein-protein interactions in a context dependent manner. In the case of polyQ, CAG triplets are more abundant in short repeats (4-8 residues), and it is particularly high in polyQ of unstable length and in human proteins associated to polyQ associated diseases (caused by CAG triplet expansion). After introducing a series of concepts describing the context of homorepeats, we will focus on the evolutionary study of particular types of homorepeats to understand how they emerge and enlarge over evolutionary time.

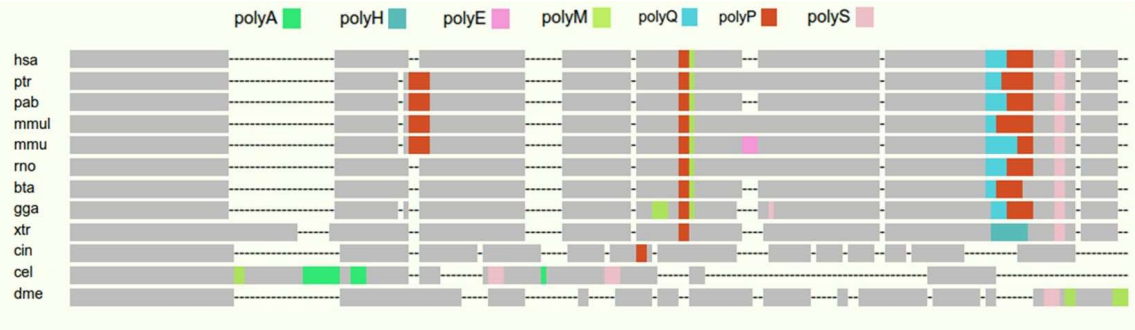


Figure 1. dAPE overview of a protein family. The view shows a simplified multiple sequence alignment where homorepeats (consecutive tracts of the same amino acid) are represented as coloured blocks. dAPE can be used to examine the variation of homorepeats in protein families at http://cbdm-01.zdv.uni-mainz.de/~munoz/polyx/.

References:

1. Mier P and Andrade-Navarro MA. (2017). dAPE: A Web Server to Detect Homorepeats and Follow Their Evolution. Bioinformatics, 33: 1221–1223.
2. Mier P and Andrade-Navarro MA. (2018). Glutamine codon usage and polyQ evolution in primates depend on the Q stretch length. Genome Biology and Evolution. 10: 816-825.

P33

Crystal structure of an Ataxin-3 binding nanobody

Gomes, Marta N.*; Manso, José A.*; Silva, Alexandra; Almeida, Ana; Sárkány, Zsuzsa; Pereira, Pedro J. B.; Macedo-Ribeiro, Sandra

IBMC - Instituto de Biologia Molecular e Celular and i3S - Instituto de Investigação e Inovação em Saúde, 4200-135 Porto, Portugal

Background: Machado-Joseph disease (MJD) is a neurodegenerative disorder included in the group of polyglutamine (polyQ) expansion diseases and is caused by the expansion of the polyQ tract of Ataxin-3 (Atx-3) [1]. Atx-3 possesses a globular Josephin Domain (JD), which contains aggregation-prone regions (APRs) required for the initial steps of aggregation. Recently, Atx-3 binding chaperones called nanobodies (NB) have been developed to target various Atx-3 domains. One of these nanobodies (NB05) interacts with Atx3 with nanomolar affinity, as determined by isothermal titration calorimetry, and significantly reduces the aggregation of the disease-related form of the protein.

Questions addressed: The aim of this work was the determination of the three-dimensional structure of NB05 to obtain information about the loops involved in Atx-3 recognition and binding.

Methods: NB05 was expressed in *Escherichia coli*, purified and crystallized for determination of its three-dimensional structure by X-ray crystallography. Preliminary screenings with commercially available collections of crystallization solutions originated crystals in three different conditions, which were later optimized to yield crystals suitable for X-ray diffraction analysis.

Results and discussion: The structure of NB05 was determined at a resolution of 2.5 Å from crystals belonging to space group $P4_22$ and containing three identical nanobody molecules in the asymmetric unit (Figure 1). The main structural features of the three hypervariable regions (CDRs) were analyzed to understand the properties of the specific Atx-3 interaction sites. CDR3 displays a prominent loop with negative electrostatic potential, while CDR1 and CDR2 exhibit positive electrostatic potential, suggesting that the interaction with Atx-3 may have an electrostatic component.

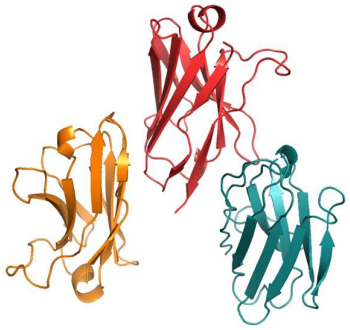


Figure 1. Cartoon representation of the three molecules in the asymmetric unit of NB05 crystals.

This work was funded by a research grant from the Amyloidosis Foundation, 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); and by FEDER through Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement in the framework of Project Norte-01-0145-FEDER-000008 (Porto Neurosciences and Neurologic Disease Research Initiative at I3S). Support of Instruct (Nanobodies4Instruct), ESRF (Grenoble, France) and ALBA (Barcelona, Spain) are greatly acknowledged.

References:

1. A. Silva, A. V. de Almeida, and S. Macedo-Ribeiro, “Polyglutamine expansion diseases: More than simple repeats,” *J. Struct. Biol.*, vol. 201, no. 2, pp. 139–154, 2017.

P34

Perspectives of intrinsically disordered proteins

Necci Marco[1,2,3], Piovesan Damiano[1], Tosatto Silvio[1,4].

[1] Dept. of Biomedical Sciences, University of Padua, Padua, Italy , [2]Dept. of Agricultural Sciences, University of Udine, via Palladio 8, 33100, Udine, Italy , [3]Fondazione Edmund Mach, Via E. Mach 1, 38010, S. Michele all'Adige, Italy , [4]CNR Institute of Neuroscience, Padua, Italy

Background: One of the key problems with intrinsically disordered (ID) proteins is the lack of a clear definition of the phenomenon. Computational methods have contributed to the identification of ID proteins from sequence and recently published curated ID databases provide a starting point for analysis. Different databases focus on different aspects of ID, yet their integration remains very limited. Only 117 proteins (out of 1,710) are annotated in at least two different databases. The recent update of MobiDB, the central resource for ID protein annotations, presented the opportunity to carry out an in-depth comparison of the content of these validated ID collections, namely DIBS [1], DisProt [2], IDEAL [3], MFIB [4], FuzDB [5], ELM [6] and UniProt.

Questions addressed: In order to assess what is specific to different ID flavors, we analyzed relevant sequence-based features, such as amino acid composition, length, taxa and Gene Ontology terms, highlighting differences and similarities among datasets.

Results and discussion: The analyzed sources of manually curated ID present many similarities and some significant differences. Each dataset focuses on a particular set of proteins, with surprisingly little overlap between them except for a fraction of entries from DisProt, IDEAL and DIBS. All datasets appear to draw different samples from the large and pervasive population of ID proteins. Most proteins annotated in all datasets come from eukaryotic organisms, although DisProt and MFIB display an interest in bacterial ID. UniProt focuses on Eukaryotes and viruses, with these latter being covered by all other datasets to some extent. Very few example of Archaean ID proteins are instead collected. The annotated ID regions belong approximately to the same pool of lengths, with the exception of ELM annotating short regions by definition. Instead, some differences are highlighted when analyzing the net charge regions. While most regions have balanced and relatively small net charge, IDEAL, DisProt and DIBS also include regions with high positive or negative net charge. On the other hand, UniProt presents very few of these cases and MFIB is completely devoid of them. All datasets follow the typical compositional bias observed in ID when compared to sequences of globular proteins, but differences specific to particular amino acids reflect the different focus of the datasets. In this study, ID is rarely associated with low sequence complexity, suggesting that the latter is not necessary to achieve the former.

References:

1. Schad E, Fichó E, Pancsa R, Simon I, Dosztányi Z, Mészáros B DIBS: a repository of disordered binding sites mediating interactions with ordered proteins. Bioinformatics [Internet]. Available from: https://academic.oup.com/bioinformatics/article/doi/10.1093/bioinformatics/btx640/4386920/DIBS-a-repository-of- disordered-binding-sites

2. Piovesan D, Tabaro F, Mičetić I, Necci M, Quaglia F, Oldfield CJ, Aspromonte MC, Davey NE, Davidović R, Dosztányi Z, et al. (2017) DisProt 7.0: a major update of the database of disordered proteins. Nucleic Acids Res. 45:D1123–D1124.

3. Fukuchi S, Amemiya T, Sakamoto S, Nobe Y, Hosoda K, Kado Y, Murakami SD, Koike R, Hiroaki H, Ota M (2014) IDEAL in 2014 illustrates interaction networks composed of intrinsically disordered proteins and their binding partners. Nucleic Acids Res. 42:D320-325.

4. Fichó E, Reményi I, Simon I, Mészáros B MFIB: a repository of protein complexes with mutual folding induced by binding. Bioinformatics [Internet]. Available from: https://academic.oup.com/bioinformatics/article/doi/10.1093/bioinformatics/btx486/4061276/MFIB-a-repository-of- protein-complexes-with-mutual

5. Miskei M, Antal C, Fuxreiter M (2017) FuzDB: database of fuzzy complexes, a tool to develop stochastic structure- function relationships for protein complexes and higher-order assemblies. Nucleic Acids Res. 45:D228–D235.

6. Dinkel H, Van Roey K, Michael S, Kumar M, Uyar B, Altenberg B, Milchevskaya V, Schneider M, Kühn H, Behrendt A, et al. (2016) ELM 2016--data update and new functionality of the eukaryotic linear motif resource. Nucleic Acids Res. 44:D294-300.

P35

Large scale analysis of amyloidogenic regions in proteins from evolutionary diverse organisms.

Villain Etienne[1,2], Fort Philippe[1], Kajava Andrey[1,2]
[1]*Centre de recherche de biologie cellulaire de Montpellier, CNRS-UMR5237, université de Montpellier,*
[2]*Institut de Biologie Computationelle de Montpellier*

Background: Various normally innocuous and soluble proteins have a potential to aggregate to form insoluble amyloid fibrils. Amyloid fibrils are the subject of special interest mainly due to their link to a broad range of human diseases (amyloidosis) including neurodegenerative diseases and cancer. In some organisms, amyloid structures can also play important beneficial roles where they are called functional amyloids (1). Over the last decade, numerous studies have demonstrated that the propensity to form amyloids is coded by the amino acid sequence (2,3). Consequently, various computational tools have been developed to predict amyloidogenic regions from sequences.

Questions addressed: Despite the growing interest about amyloids, there is not yet any global picture of the amyloidogenicity: how common they are remains largely debated, and their distribution across the domains of life poorly known. In this context, we aim to develop an annotation pipeline based on state of the art computational methods available to predict amyloidogenic regions and to use it to provide a global picture of the amyloidogenicity on a dataset of proteomes selected to provide a good coverage of the tree of living.

Methods:
As a preliminary for our work, we performed a benchmark of the different amyloidogenic region predictors on sequences available as standalone software (3). Secondly, based on this study we developed, following clear software engineering principles (component oriented architecture, high unit test coverage) our annotation pipeline based on this selected tools. As the ability of an amyloidogenic region to form fibrils largely depend on structural context we complete these predictors with rules to keep only predictions located within intrinsically disorder regions and rules to reject predictions within transmembrane regions. Finally, we applied this pipeline on a dataset of proteomes selected for their phylogenetic diversity and their quality (as defined by the Uniprot Consortium (4)).

Results and discussion:
This large scale analysis shows both differences between amyloidogenicity predictors and important variations in the amyloidogenic propensity of these organisms. We studied proteome features such as the protein length distribution, cellular localization, intrinsically disorder region coverage…etc. and how they are linked with the amyloidogenicity. We found some organisms (*Dictyostelum Discoideum, Plasmodium Falciparum, Drosophila Melanogaster*) exhibit highly amyloidogenic proteomes, which correlate with their enrichment in Q/N rich regions and confirm the literature on this particular topic. In addition, we identified some new variations that we proposed are linked with environmental conditions and especially temperature: thermophilic species tend to present much less amyloidogenic regions than mesophilic species and even less than psychrophilic ones. These results allow us to hypothesize a link between the optimal growing temperature of organisms and the average amyloidogenicity of their proteomes.

References:
1. Otzen, D. and Nielsen, P. H. We find them here, we find them there: Functional bacterial amyloid, Cellular and Molecular Life Sciences, 2008
2. Ventura, S. and Villaverde, A. Protein quality in bacterial inclusion bodies, Trends in Biotechnology, 2006
3. Ahmed, A.B. and Kajava, A.V. Breaking the amyloidogenicity code : Methods to predict amyloids from amino acid sequence, FEBS Letters, 2012
4. Roche, D.B. and Villain, E. and Kajava, A.V. Usage of a dataset of NMR resolved protein structures to test aggregation vs. solubility prediction algorithms, Protein science, 2017
5. The UniProt Consortium. UniProt: the universal protein knowledge database, Nucleic acid research, 2017.

P36

SODA: prediction of protein solubility from disorder and aggregation propensity

Paladin Lisanna [1], Piovesan Damiano [1], Tosatto Silvio [1,2]
[1] *Department of Biomedical Sciences, University of Padua, Viale G. Colombo 3, 35121 Padova, Italy,*
[2] *CNR Institute of Neuroscience, Viale G. Colombo 3, 35121 Padova, Italy*

Background: Solubility is an important, albeit not well understood, feature determining protein behavior. It is of paramount importance in protein engineering, where similar folded proteins may behave in very different ways in solution. The identification and tuning of sequence determinants for protein aggregation has been used as a valuable tool to regulate protein solubility (1). Among the determinants of protein aggregation, intrinsic disorder has also been shown to play a major part (2). The highly dynamical disordered regions of a protein can increase its propensity to aggregate under different conditions. Both aggregation and intrinsic disorder propensity are influenced by the physico-chemical properties of each amino acid in the sequence, such as hydrophobicity, secondary structure propensity and charge (3).

Questions addressed: We wondered if the integration of all these concepts (aggregation and disorder propensity, hydrophobic profile, predicted secondary structure components) could enhance state-of-the-art prediction of protein solubility.

Methods: Here, we describe SODA, a new method to predict the effects of sequence variations on protein solubility. SODA characterizes an input sequence with its intrinsic solubility profile, derived from PASTA aggregation energy (4), ESpritz disorder propensity (5), the negative Kyte–Doolittle hydrophobicity profile (6) and the two secondary structure propensities for α-helix and β-strand calculated with FESS (7). It is able to evaluate the effects on solubility of any sequence variation (including indels) by comparing the original and derived profiles. SODA is trained using 5-fold cross-validation on a filtered version of the PON-Sol dataset (8).

Results and discussion: The comparison to other recently published sequence-based methods shows that SODA has state-of-the-art performance and is particularly well suited to predict mutations decreasing solubility. The method is fast, returning results for single mutations in seconds. The web server has two different operating modes, allowing the user to either target mutations or evaluate the effect of all possible substitutions on the input sequence. A usage example estimating the full repertoire of mutations for a human germline antibody highlights several solubility hotspots on the surface. The web server was designed to allow large-scale annotation through its RESTful web service, while the user interface provides an intuitive form to guide detailed selection of mutations based on sequence solubility plot and, if the protein structure is given, residues accessibility to solvent.
SODA can be accessed from URL: http://protein.bio.unipd.it/soda.

References:
1. Ventura S. Sequence determinants of protein aggregation: tools to increase protein solubility. Microb. Cell Factories. 2005; 4:11.
2. De Simone A., Kitchen C., Kwan A.H., Sunde M., Dobson C.M., Frenkel D. Intrinsic disorder modulates protein self-assembly and aggregation. Proc. Natl. Acad. Sci. U.S.A. 2012; 109: 6951-6956.
3. Chiti F., Stefani M., Taddei N., Ramponi G., Dobson C.M. Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature . 2003; 424:805–808.
4. Walsh I., Seno F., Tosatto S.C.E., Trovato A. PASTA 2.0: an improved server for protein aggregation prediction. Nucleic Acids Res. 2014; 42:W301–W307.
5. Walsh I., Martin A.J.M., Di Domenico T., Tosatto S.C.E. ESpritz: accurate and fast prediction of protein disorder. Bioinformatics . 2012; 28:503–509.
6. Kyte J., Doolittle R.F. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 1982; 157:105–132.
7. Piovesan D., Walsh I., Minervini G., Tosatto S.C.E. FELLS: fast estimator of latent local structure. Bioinformatics . 2017; doi:10.1093/bioinformatics/btx085.
8. Yang Y., Niroula A., Shen B., Vihinen M. PON-Sol: prediction of effects of amino acid substitutions on protein solubility. Bioinformatics . 2016; 32:2032–2034.

P37

Prediction of Human Gene - Phenotype Association for Intrinsically Disordered Proteins

Vladimir Perovic, Radoslav Davidovic, Branislava Gemovic, Neven Sumonja, Nevena Veljkovic
Centre for Multidisciplinary Research and Engineering, Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia

Background: Intrinsically disordered proteins (IDPs) act as hubs in interaction networks and play key roles in signaling pathways. Consequently, dysregulation of IDPs emerges as a critical element for many diseases, including cancer, metabolic disorders, and diseases of protein homeostasis. Human Phenotype Ontology (HPO) organizes human diseases and associated genes into hierarchical classes based on the phenotypes they present. Human IDPs are currently under-represented in HPO, but computational predictions of gene-HPO relationships have great potential to accelerate the pace of discovery and to prioritize candidate disease genes.

Questions addressed: Our goal was to predict IDPs-HPO relationships based on universal, sequence characteristics.

Methods: Human IDP sequences were represented as a 470 dimension vectors by using (1) pseudo amino acid composition (PAAC) based on disorder characteristics propensity scales and (2) dipepdide composition (DC). We denoted this PAACDC encoding. We explored different machine learning algorithms and developed a model for automatic annotation of IDPs in HPO.

Results and discussion: We demonstrate that our method performs comparably well to the PHENOSTRUCT (1) and HEMDAG (2) which are state of the art methods that utilize domain knowledge. On the other hand, a significant advantage of our method is that it is exclusively sequence based and computationally more efficient.

References:

1. Kahanda I, Funk C, Verspoor K and Ben-Hur A. PHENOstruct: Prediction of human phenotype ontology terms using heterogeneous data sources. F1000Research 2015, 4:259
2. Notaro et al. Prediction of Human Phenotype Ontology terms by means of hierarchical ensemble methods. BMC Bioinformatics (2017) 18:449.

P38

Short linear interacting motifs under the lens of structural biology

Papaleo, Elena[1]
[1]Computational Biology Laboratory, Danish Cancer Society Research Center, Copenhagen

Background:
Short linear interacting motifs (SLiMs) are short, ubiquitous and functional protein interaction modules, which are highly degenerated and hard to detect [1]. They are often enriched in disordered regions and often contributing to specificity for the interaction. High-throughput proteomics initiative are nowadays providing a large amount of possible protein-protein interactions that need to be dissected and studied in details to discriminate between false positives and truly important functional interactions. The SLiM paradigm can help in this direction but it is not sufficient on its own.

Questions addressed: My group is pursuing a multi-disciplinary effort to combine different computational methods with in-vitro assays to unveil the role and mode of action of SLiMs at the structural level, as well as the impact of disease related mutations or post-translational modification on SLiMs functions.

Methods: Our approaches range from templated-based/ab-initio prediction of the complexes between SLiMs and the partner of interactions, high-throughput in silico estimate of change in binding free energies upon mutations to all-atom molecular dynamics simulations [2,3]. In our group, we also integrate in our studies with experimental validation using peptide arrays and NMR experiments.

Results and discussion:
I will illustrate the application of our integrative workflow to different cases of studies where we shed new light on the related molecular mechanisms. In particular, I will discuss SLiMs in intrinsically disordered proteins that regulate ubiquitin-related functions, apoptosis and autophagy which are in turn regulated by post-translational modifications.

References:
1. Van Roey K et al., Chem Rev, 2013
2. Papaleo et al., Chem Rev, 2016
3. Papaleo et al., Frontiers Mol Biosciences, 2016.

P39

Intrinsically disordered regions in selective macroautophagy receptors and adaptors

Promponas, Vasilis J.[1], Kalvari, Ioanna[1], Chatzichristofi, Agathangelos[1]
[1]Bioinformatics Research Laboratory, Department of Biological Sciences, University of Cyprus, CY 1678, Nicosia, Cyprus

Background: Macroautophagy is a tightly regulated cellular catabolic process, conserved throughout the eukaryotes, driving cytoplasmic material (either selectively or in bulk) into a dynamic, double-layered membrane compartment (autophagosome), which eventually fuses with a lysosome and leads to the degradation of the isolated material [1]. The most well characterized cases of selective macroautophagy are mediated by receptor or adaptor proteins, which interact with Atg8 homologues via a short linear motif (AIM or LIR motif) [2].

Questions addressed: It was soon hypothesized that AIM/LIR motifs may be functional when appearing within an intrinsically disordered region [3] and this notion soon led to a predictor of functional AIM/LIR motifs [4]. In particular, of all the possible AIM/LIR motifs present in a protein sequence of interest the iLIR method [4] highlights those overlapping with subsequences flanking or overlapping intrinsically disordered regions (as predicted with ANCHOR [5]) with a high potential to be stabilized upon binding to a target molecule (in this case, an Atg8 homolog). In our effort to develop more accurate predictors of functional AIM/LIR motifs, we set to identify in more detail the impact of the existence of intrinsically disordered regions within characterized selective autophagy receptors/adaptors with experimentally determined AIM/LIR-mediated interactions to Atg8 homologs.

Methods: Along these lines we performed extensive literature searches, consulted relevant freely-accessible databases with information related to intrinsically disordered proteins (e.g. mobiDB [6] and DisProt [7]) and are currently benchmarking different predictors of intrinsically disordered regions (IDRs) for their ability to discriminate functional AIM/LIR motifs from noise. We are also using several relevant sequence-derived properties to develop machine learning classifiers that enhance the prediction of functional AIM/LIR motifs.

Results and discussion: We have collected an updated list of selective autophagy receptors and adaptors which are experimentally verified to posses AIM/LIR motifs that mediate interactions with Atg8 homologs in several eukaryotes. Using these sequences as a training dataset we have developed a prototype of a Decision Tree (DT) based classifier, which using a robust 5-fold cross-validation procedure manages to discriminate functional AIM/LIR motifs more accurately than the current state of the art. We are currently examining the decision rules applied by the DT to identify the biological significance of the features used in the prediction model. Moreover, we also currently examining the usefulness of the information provided by different IDR prediction methods towards building more successful classifiers and, hopefully, elucidating those particular features of IDRs that make them important in this specific type of interactions.

References:

1. Klionsky DJ et al., Guidelines for the use and interpretation of assays for monitoring autophagy (**2016)** *Autophagy*, 12(1):1-122.
2. Pankiv S et al., p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy **(2007)** *J Biol Chem*, 282:24131-45.
3. Noda et al., Structural basis of target recognition by Atg8/ LC3 during selective autophagy **(2008)** *Genes Cells*, 13:1211-8.
4. Kalvari et al., iLIR: A web resource for prediction of Atg8-family interacting proteins **(2014)** *Autophagy*, 10(5):913-25.
5. Mészáros B et al., Prediction of protein binding regions in disordered proteins **(2009)** *PLoS Comput Biol*, 5:e1000376.
6. Piovesan D et al., DisProt 7.0: a major update of the database of disordered proteins **(2017)** *Nucleic Acids Res*, 45(D1):D1123-D1124.
7. Piovesan D et al., MobiDB 3.0: more annotations for intrinsic disorder, conformational diversity and interactions in proteins **(2018)** *Nucleic Acids Res,* 46(D1):D471-D476.

P40

CRISPR-mediated genome editing of TDP-43 in SHSY5Y cells to model ALS

Emanuele Celauro, Elin Esbjörner
Department of Biology and Biological engineering, Chalmers University
Email: celauro@chalmers.se

Background: Amyotrophic lateral sclerosis (ALS) is an incurable, strongly debilitating, rapidly progressing, and lethal degenerative disorder that affects motor neurons and voluntary muscle control. A common neuropathological characteristic of ALS is the deposition of the TAR-DNA binding protein (TDP-43) as aggregates in the cytosol, which is found in 97% of patients, in addition to its clearance from the nucleus. Many different approaches have been developed to study the onset of this disease, but most of them rely on cellular models in which the nuclear localization signal (NLS) of TDP-43 is mutated, leading to its accumulation in the cytoplasm, or in which modulation of autophagy, stress granules formation and protein folding is used to trigger and then mimic the pathological progression and formation of protein aggregates[1].

Questions addressed: A possible weakness of such approaches is that they involve transient or stable overexpression of a fluorescently labelled TDP-43, always in addition to the endogenous protein, which brings noise to the cell model system.

Methods: We took advantage of the recently developed CRISPR-mediated genome editing, to stably tag the endogenous TDP-43 by knocking in a fluorophore-PuroR cassette through homology directed repair (HDR)[2]. With the same approach, additional rounds of genome editing can be easily performed to reintroduce or disrupt the NLS mutation, as well as the ubiquitination targets, without having the endogenous TDP-43 mitigating the obtained phenotype.

Results and discussion: Preliminary analysis of the CRISPR-edited SHSY5Y cell line showed an unexpected basal level of cell-to-cell TDP-43:GFP transmission, which will be further studied to assess its role in ALS onset and spreading. At the same time, reintroduction of the NLS1 mutation by ssDNA-mediated knock-in caused the death of all cells; viability was rescued only in presence of an overexpression of TDP-43, casting additional shadows on the most common cellular model for ALS, based on the forced overexpression of the NLS1 mutant.

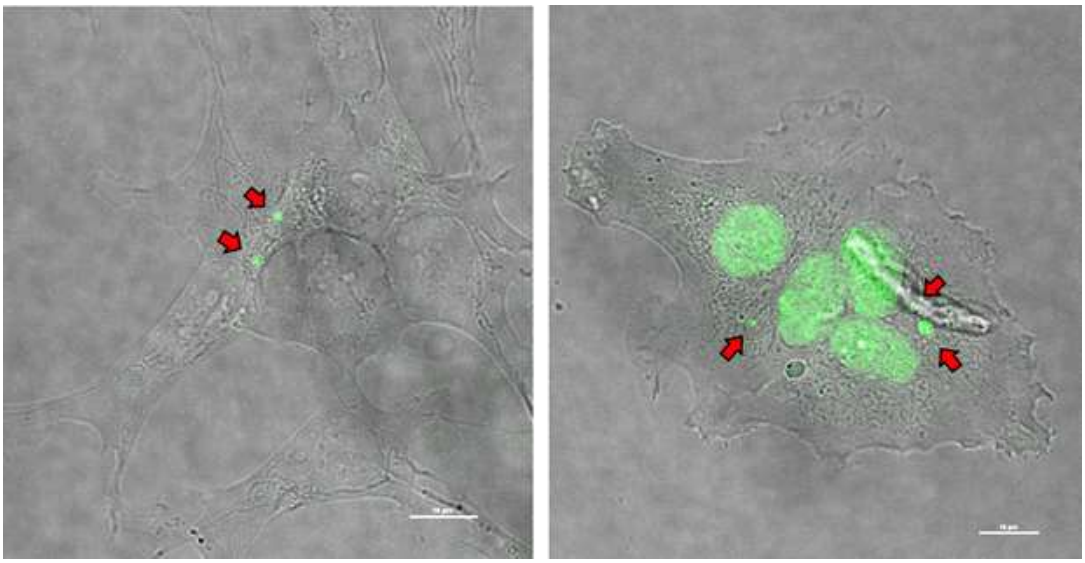


Figure 1. Left image: normal SHSY5Y cells whit uptake of TDP-43:GFP (arrows) after co-culture with TDP-43:GFP CRISPR line (right image).

References:

1. Scotter, E. L., et al. (2015). "TDP-43 Proteinopathy and ALS: Insights into Disease Mechanisms and Therapeutic Targets." Neurotherapeutics 12(2): 352-363.
2. Wang, B., et al. (2015). "Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes." Biotechniques 59(4): 201-202, 204, 206-208.

P41

Disulfide driven folding for a conditionally disordered protein

Jordi Pujols[1,2], Hugo Fraga[1,2], Francesc Canals[3], Pau Bernadó[4], Rita Grandori[5], Oscar Millet6 & Salvador Ventura[1,2]

1Institut de Biotecnologia i Biomedicina. Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain; 2Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, 08193, Bellaterra, Spain, 3Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain,4Centre de Biochimie Structurale, INSERM-U1054, CNRS UMR-5048, Université de Montpellier, 29, rue de Navacelles, 34090, Montpellier, France, 5Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milano, Italy, 6Protein Stability and Inherited Diseases Laboratory, CIC bioGUNE, 48160, Derio, Spain.

Background: Conditionally disordered proteins are either ordered or disordered depending on the environmental context. The substrates of the mitochondrial intermembrane space (IMS) oxidoreductaseMia40 are synthesized on cytosolic ribosomes and diffuse as intrinsically disordered proteins to the IMS, where they fold into their functional conformations; behaving thus as conditionally disordered proteins. It is not clear how the sequences of these polypeptides encode at the same time for their ability to adopt a folded structure and to remain unfolded [1,2]

Questions addressed: The small cysteine-rich protein hCox17 suffers a conformational switch to become functional. In this work we address which are the determinants that drive this phenomenon, driven by reduction and oxidation events both in presence and absence of its specific chaperone Mia40.

Methods: In order to characterize the disorder-to-order transition of hCox17, we use an integrated realtime approach, including RP-HPLC, fluorescence, CD, FTIR, SAXS, NMR, and MS analysis.

Results and discussion: We demonstrate that the the conformational switch between the disordered and folded state of hCox17 is controlled by the formation of a single disulfide bond, both in the presence and in the absence of Mia40. We provide molecular details on how the folding of a conditionally disordered protein is tightly regulated in time and space, in such a way that the same sequence is competent for protein translocation and activity.

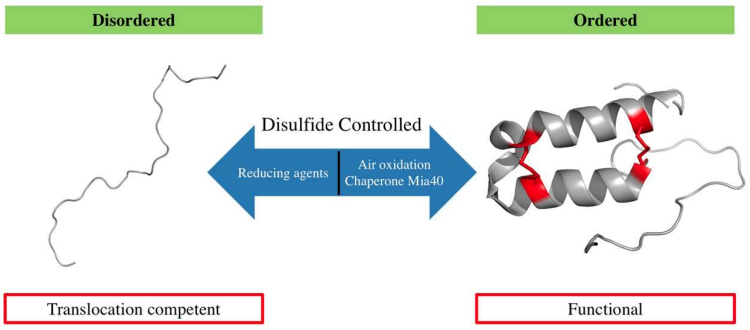


Figure 1. Schematic representation of the two biologically relevant conformations of hCox17, the canonical substrate of hMia40 chaperone. The conversion between the two states is redox controlled by the formation of native disulfide bonds.

References:

1. Mesecke N, et al. A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. Cell. 2005;121:1059–1069. doi: 10.1016/j.cell.2005.04.011.
2. Banci L, et al. MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. Nat. Struct. Mol. Biol. 2009;16:198–206. doi: 10.1038/nsmb.1553.

P42

On the role of peptide hydrolysis for fibrillation kinetics and amyloid fibril morphology[1]

Lendel, Christofer[1], Ye, Xinchen[2], Langton, Maud[3] A, Hedenqvist, Mikael S.[2]

[1] Dept. of Chemistry, KTH Royal Institute of Technology, Stockholm, Sweden. [2] Dept. of Fibre and Polymer Technology, KTH Royal Institute of Technology, Stockholm, Sweden. [3] Dept. of Molecular Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Background: The formation of amyloid-like protein nanofibrils are not only key events in several diseases, such fibrils are also associated with intriguing biological function and constitute promising components for new biobased materials. Still there are many aspects of the molecular processes leading to fibril formation that are not well understood. The bovine whey protein β–lactoglobulin is an important model system for the development of protein nanomaterials as well as general studies of amyloid formation mechanisms. As many other proteins, it has been shown to form fibrils with different morphologies depending on the reaction conditions. An intriguing observation is the sharp switch from long and straight to short and curved fibrils achieved by increasing the initial protein concentration.[2,3]

Questions addressed: We investigate the molecular mechanism behind the concentration-dependent switch in morphology of β–lactoglobulin fibrils using whey protein isolate as stating material.[1]

Methods: Kinetic investigations based on thioflavin T-fluorescence are combined with structural characterization using spectroscopic methods, atomic force microscopy and mass spectrometry.

Results and discussion: We find that peptide hydrolysis is the rate-limiting step for the formation of β–lactoglobulin fibrils under the investigated conditions. Moreover, the presented results suggest that the concentration-dependent morphological switch is related to different structures of the fibril nuclei. The two classes of morphologically distinct fibrils have different compositions of the peptide building blocks, as observed by mass spectrometry. Based on these results, we propose that not only the fibrillation kinetics but also the switch in fibril morphology is connected to the hydrolysis of the protein chain. These findings open new possibilities to control the structural properties of protein nano-materials e.g. by enzymatic processing. The results may also be of relevance for amyloid pathology as they suggest that not only the total concentration of an amyloidogenic protein, but also the processing capacity could determine what type of aggregates that are formed.

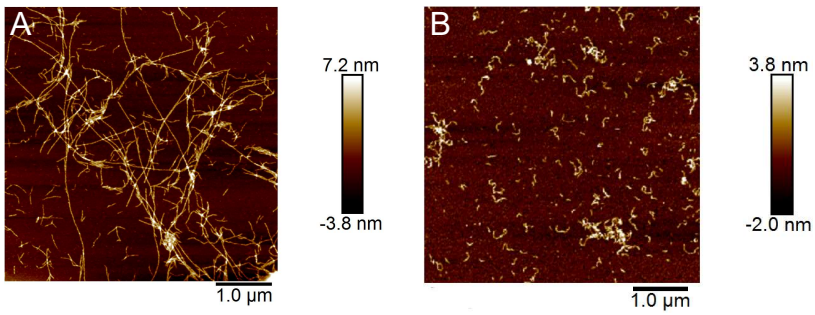


Figure 1. Atomic force microscopy images showing the (A) long and straight fibrils formed at lower protein concentrations and (B) short and curved fibrils formed at higher protein concentration.

References:

1. Ye X, Hedenqvist MS, Langton M, Lendel C. (2018) *RSC Adv.* 8: 6915-6924.
2. Kamada A, Mittal N, Söderberg LD, Ingverud T, Ohm W, Roth SV, Lundell F, Lendel C. (2017) *Proc. Natl. Acad. Sci. USA*. 114: 1232-1237.
3. vandenAkker CC, Engel MF, Velikov KP, Bonn M, Koenderink GH. (2011) *J. Am. Chem. Soc.* 133: 18030-18033.

P43

IDPpi: Protein-Protein Interaction Analyses of Human Intrinsically Disordered Proteins in DisProt

Veljkovic Nevena[1], Sumonja Neven[1], Perovic Vladimir[1]
[1]Centre for Multidisciplinary Research and Engineering, Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia

Background: The human interactome is estimated to encompass several hundred thousand binary protein-protein interactions (PPI). The sub-interactome of intrinsically disordered proteins (IDPs) is particularly complex because a single IDP usually binds a few more partners than an ordered protein. IDPs are involved in key signaling and regulatory processes and therefore, understanding of their protein interactions will enlighten different aspects of cellular physiology and pathophysiology.

Questions addressed: IDPs are distinct due to their compositional bias which influences their binding propensities and selection of partners. This motivated us to develop IDPpi, a method for PPI predictions which utilizes supervised machine learning algorithms and compositional content together with the distribution of features associated with the promotion of structural disorder along the sequence string.

Methods: We capitalized on the collection of manually curated human IDPs in the latest version of the DisProt database and developed a model that utilizes machine learning to classify and predict binary PPIs. The classifier exploits sequences encoded by five amino acid scales that represent different order or disorder endorsing propensities and experimentally confirmed PPIs of the DisProt human subset. In order to translate PPI information into functional insights, we associated predictions with Gene Ontology (GO) and Human Phenotype Ontology (HPO) terms. The list of terms enriched among the predicted interactors is transferred to a query IDP gene.

Results and discussion: Consideration of both sequence determinants specific for conformational organizations and the multiplicity of IDP interactions in the training phase ensured a reliable approach that efficiently predicts PPIs even on the proteome scale. The comparative evaluation performed by cross-validation and on hold-out test sets showed that our model significantly outperforms other sequence-based methods and identifies candidates for populating the sparse zones of the human interactome. IDPpi service is provided as a web tool to expedite the discovery of new interactions and IDP functions with enhanced efficiency.

P44

Model of S100A9 Amyloid Formation:Protein Misfolding as "Nucleation" Event

Iashchishyn, Igor A.[1,3], Sulskis, Darius[2], Nguyen Ngoc, Maja[2], Smirnovas, Vytautas [2], Morozova-Roche, Ludmilla A.[1]

[1]*Department of Medical Biochemistry and Biophysics, Umeå University, Umeå, Sweden,* [2]***Department of Biothermodynamics and Drug Design,*** *Institute of Biotechnology, Vilnius University, Vilnius, Lithuania,* [3]*Department of General Chemistry, Sumy State University, Sumy, Ukraine*

Background: S100A9 is an α-helical, calcium-binding pro-inflammatory alarmin associated with many diseases related to inflammation. It forms amyloid fibrils under physiological conditions (pH 7.4, 37 °C) and plays critical role in the amyloid-neuroinflammatory cascade in Alzheimer's disease [1].

Questions addressed: The rates of S100A9 amyloid formation were estimated from kinetic assay and fibril length distribution. Based on the results the mechanism of amyloid formation is suggested.

Methods: Amyloid kinetics were monitored by Thioflavin T fluorescence, protein secondary structure – by FTIR spectroscopy, fibril length distribution – by atomic force microscopy (AFM). Autocatalytic model, classical nucleation theory, perturbation theory and Becker-Döring model were applied for the analysis.

Results and discussion: Kinetics of S100A9 amyloid formation demonstrate no lag phase and is insensitive to seeding, clearly showing that secondary pathways are not involved in amyloid formation. The fitting of kinetic data by phenomenological autocatalytic model and the model based on perturbation theory give similar estimates of the critical nuclei size which is less than unity [2]. Remarkably, autocatalytic model is also accounting for the halftime-concentration dependence (Fig. 1A). Such results indicate that nucleation process effectively is misfolding. The fraction of the nucleation is approx. three orders of magnitude higher than elongation being rate limiting process. Exponential distribution of the fibril length also confirms the lack of secondary pathways and gives estimate of the rate formation (Fig. 1B).

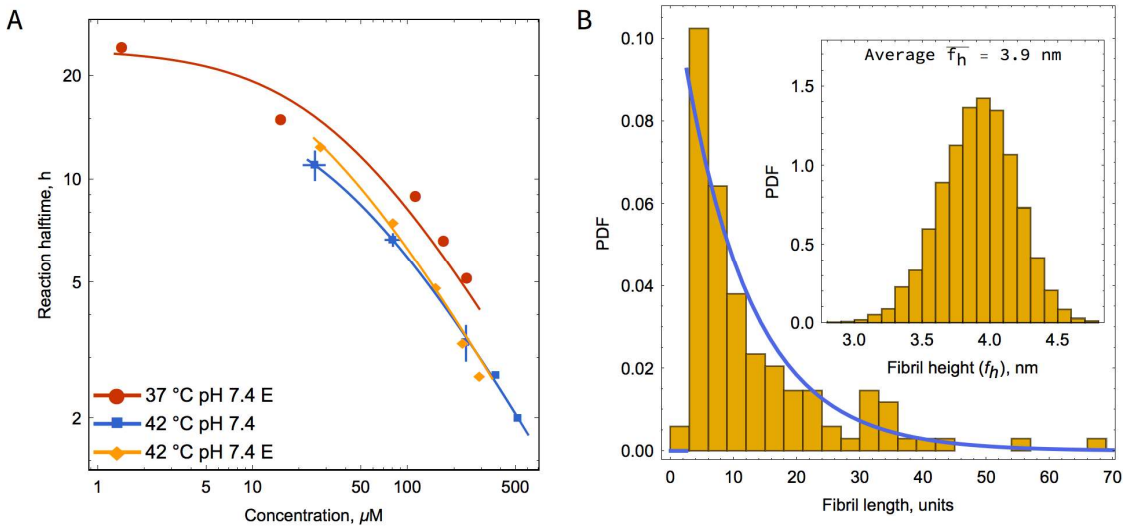


Figure 1. A – fitting of the halftime vs concentration for S100A9 amyloid formation with autocatalytic model (pH 7.4, 42 °C, ThT fluorescence); B – Fibril length distribution of S100A9 fibrils (AFM), insert – fibril height distribution. PDF – probability density function, i.e. number in class.

References:

1.Wang, C., et al., The Role of Pro-Inflammatory S100A9 in Alzheimer's Disease Amyloid Neuroinflammatory Cascade. Acta Neuropathologica, 2014. **127**(4): p. 507-522.

2.Iashchishyn, I.A., et al., Finke-Watzky Two-Step Nucleation-Autocatalysis Model of S100A9 Amyloid Formation: Protein Misfolding as "Nucleation" Event. ACS Chem Neurosci, 2017. **8**(10): p. 2152-2158.

P45

New methods for searching for similar low complexity regions in protein sequence

Gruca, Aleksandra[1], Jarnot, Patryk[1], Ziemska, Joanna[2], Grynberg, Marcin[3]
[1]Institute of Informatics, Silesian University of Technology, Poland , [2]Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland , [3]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Poland

Background: Low complexity regions (LCRs) are abundant in the protein universe. About 14% of proteins contain LCRs [1] and they could play key roles in protein functions and are relevant to protein structure [2-4]. Therefore, the ability to identify and analyse similar LCRs in different protein sequences could provide the answer to the question if similar low complexity fragments share similar biological function. However, current statistical models implemented in state-of-the-art methods for searching for similarities in protein sequences are not designed for analysis of low complexity regions, and some of them even mask LCRs to improve searching for homologous proteins. In order to overcome this problem we propose and compare new methods designed specifically for searching for similar low complexity regions in protein sequences.

Methods: We propose three different methods, which are able to search for similar LCRs. The first one creates graphs from LCR sequences in which nodes are k-mers of the sequence. Subsequently it finds cycles in this graph and finally uses these cycles to assign sequence to the cluster including similar sequences. The second method calculates PSSMs basing on similar LCRs, and then by using these matrices compares sequences. As the third method we use BLAST with specific scoring matrix that treats each amino acid substitution with equal similarity score. In this case we disabled “composition-based statistics” option, which recalculates scoring matrix basing on input sequence in order to mask LCRs.

Results and discussion: Three methods designed to search for similar LCRs are presented. Each of them could be used for different purposes. Graph method searches for similar sequences using repeats and accepts existence of noise in that repeats. Noise tolerance is given as a parameter. PSSM method is very sensitive and does not care about position of the similar part of the sequence. BLAST is able to search for LCRs, which have similar part of sequence. Graph and PSSM methods use repeat patterns in LCRs in order to search for similar sequences while BLAST is based on statistical analyses of amino acid composition in protein sequences. Results obtained with new methods are quantitatively and qualitatively compared with the standard BLAST approach.

References:

1. Marcotte E., Pellegrini M., Yeates T., Eisenberg D. *A* census of protein repeats. J Mol Biol. 1999;293:151–160.
2. Hamm D.C., Bondra E.R., Harrison M.M. Transcriptional activation is a conserved feature of the early embryonic factor zelda that requires a cluster of four zinc fingers for DNA binding and a low-complexity activation domain. The Journal of Biological Chemistry. 2015;290(6):3508-3518.
3. O’Rourke T.W., Loya T.J., Head P.E., Horton J.R., Reines D. Amyloid-like assembly of the low complexity domain of yeast Nab3. Prion. 2015;9(1):34-47.
4. Kumari B., Kumar R., Kumar M. Low complexity and disordered regions of proteins have different structural and amino acid preferences. Mol Biosyst. 2015 Feb;11(2):585-94.

P46

Should the treatment of amyloidosis be personified?

Galzitskaya, Oxana[1], Surin Alexey[1], Glyakina Anna[1], Rogachevsky Vadim[2], Selivanova Olga[1]
[1]Institute of Protein Research, [2]Institute of Cell Biophysics, Pushchino, Moscow Region, Russia

Background: Studies of intermediate states of protein aggregation upon formation of fibrils are in the focus of attention not only for deciphering the mechanism of fibril formation, but also for elaboration of different approaches to cope with oligomerization, which will facilitate proposing novel methods for prophylactics and prevention of harsh amyloidogenic diseases [1,2].

Questions addressed: Should the treatment of amyloidosis be personified?

Methods: Electron microscopy, X-ray analysis, limited proteolysis and mass spectrometry analysis, molecular modeling.

Results and discussion: Polymorphism is a specific feature of amyloid structures. We have studied amyloid structures and the process of their formation using synthetic and recombinant preparations of Aβ peptides and their three fragments. Fibrils of different morphology were obtained for these peptides. We suppose that formation of Aβ peptides and their fragments proceeds according to the simplified scheme: destabilized monomer → ring-like oligomer → mature fibril that consists of ring-like oligomers. We are the first who did 2D reconstruction of amyloid fibrils provided that just a ring-like oligomer is the main building block of an amyloid fibril, like a cell in an organism, it is easy to explain polymorphism of fibrils as well as splitting of mature fibrils under different external actions, branching and inhomogeneity of fibril diameters. Identification of regions in the protein chain that form the backbone of an amyloid fibril is a direction in investigating amyloidogenesis. It has been demonstrated for $A\beta_{42}$ peptide and its fragments that their complete sequence is inaccessible for the action of proteases, which is evidence of different ways of association of ring-like oligomers upon formation of fibrils. Based on the electron microscopy and mass spectrometry data, we have proposed a molecular model of a fibril formed by both Aβ peptide and its fragments. In connection with this, the unified way of formation of fibrils from oligomers, which we have discovered, could facilitate development of relevant fields of medicine of common action. But in our opinion, the basic attention should be focused on physiological, genetic and other reasons leading to destabilization of native molecules of proteins and peptides and triggering the process of fibrillogenesis.

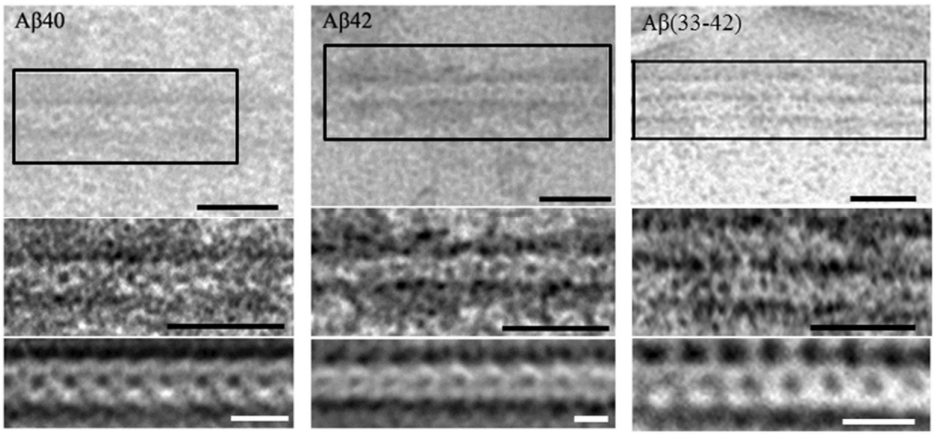


Figure 1. EM images and 2D reconstructions of fragments of amyloid fibrils of Aβ40, Aβ42 peptides and Aβ(33-42) peptide fragment.

References:

1. Selivanova et al. Langmuir. 2018 Feb 13;34(6):2332-2343.
2. Galzitskaya OV, Selivanova OM. J Alzheimers Dis. 2017;59(3):785-795.

P47

Case Studies in DisProt Annotation

Borbála Hajdu-Soltész, Zsuzsanna Dosztányi
MTA-ELTE Momentum Bioinformatics GroupBudapest, Hungary

Intrinsically disordered proteins (IDPs) are increasingly recognized for their importance in a wide range of biological processes, especially in signalling and regulation. Intrinsic disorder generally means the absence of a well-defined 3D structure, but for long its diverse terminology had not been unified in the literature. In 2005 the Database of protein disorder (DisProt) was launched with the aim of creating a manually curated database of experimentally validated IDPs to facilitate efficient management and annotation of IDP-related structural and functional information (Vucetic, 2005; Sickmeier, 2007). With the growing interest in IDPs the DisProt database has been recently upgraded and updated with new examples. In addition, the categories in DisProt's functional ontology have also been re-defined and extended (Piovesan, 2017).
This year (2018) a massive extension of DisProt has been initiated. Several curators have joined to create new entries and add new annotations. Curators can upload their annotations on a user-friendly web-based platform, and they instantly get rewarded with contribution scores for adding new regions, proteins or articles, or fixing annotations already present in DisProt.
Importantly, all entries submitted into the DisProt database have to be supported by literature evidence. Looking for articles on disordered proteins might seem easy as there are thousands of papers in this field. What makes annotation rather difficult, however, is that the interpretation of the experimental results regarding disorder is not always so clear. Sometimes the boundaries of the regions are not marked in the text but can be read from a figure. It also happens that the disordered region cannot be annotated because it fails basic criteria of DisProt, for it contains less than 10 amino acids, or the evidence clearly does not represent native structure. Most studies are carried out on protein fragments, which again can make the annotation process more challenging. In this work I show examples of straightforward and ambiguous cases which I encountered during the annotation of nearly 40 proteins. This might help other curators to interpret what they find in the papers and DisProt users to rationalize the content of this database.

References:

1. Vucetic et al. DisProt: a database of protein disorder. Bioinformatics, 2005.
2. Sickmeier et al. DisProt: the Database of Disordered Proteins. Nucleic Acids Res., 2007.
3. Piovesan et al. DisProt 7.0: a major update of the database of disordered proteins. Nucleic Acids Res., 2017.

P48

Interaction of tethered bilayer membranes with β-amyloid and s100a9 aggregates

Budvytyte Rima[a], Smirnovas Vytautas[b], Lösche Mathias,[c] and Valincius Gintaras[a]
[a]Life Sciences center, Institute of Biochemistry, University of Vilnius, Vilnius, Lithuania; [b] Life Sciences center, Institute of Biotechnology, University of Vilnius, Vilnius, Lithuania; [c]Department of Physics, Carnegie Mellon University, Pittsburgh, USA; rima.budvytyte.@bchi.vu.lt

Background: A central event in pathogenesis of Alzhaimer's diseases are thought to be intracellular and extracellular accumulation, aggregation and misfolding of low molecular mass peptides [1] such as β-amyloid ($A\beta_{1-42}$), tau protein (Tau) and s100A9. Small size aggregates-oligomers were found to be extremely neurotoxic *in vitro* and *in vivo* with the ability to disrupt the major neuron membranes [2,3] and lead to synaptic dysfunction, mithochondrial dysfunction, neuronal apoptosis and brain damage [4].

Questions addressed: In this work different sizes of soluble recombinant $A\beta_{1-42}$ [5,6] and S100A9 aggregates were used to investigate their interaction with tethered phospholipid membranes (tBLM).

Methods: $A\beta_{1-42}$ aggregates were obtained without fluorinated compounds derivatives or solvents, e.g.dimethylsulfoxide, HFIP. The morphology and size of misfolded protein aggregates ($A\beta_{1-42}$ and S100A9) were monitored by dynamic light scattering (DLS) and atomic force microscopy (AFM). These protein aggregates exhibited the membrane damaging properties as probed by the electrochemical impedance spectroscopy (EIS).

Results and discussion: The function and morphology of misfolded proteins was depending on different oligomerisation conditions. Amyloid induced membrane conductance exhibited relatively weak temperature dependence. The effective activation energy of the ion transport through the membrane (which contains sphingomyelin) defects exhibited values below 10 kJ/mol, which is consistent with physical picture of the water-filled pores formed by pore forming toxins. Membrane composition was found to be one of the important factors affecting the interaction of the $A\beta_{1-42}$ and s100A9 oligomers to phospholipid membranes.

References:

1. Taylor J.P. Toxic proteins in neurodegenerative disease. *Science*, **2002**;296:1991-1995.
2. Andreasen M, Lorenzen N, Otzen D. Interactions between misfolded protein oligomers and membranes: A central topic in neurodegenerative diseases? *Biochimica et Biophysica Acta,* **2015**;1848:1897–1907.
3. Marin R. The neuronal membrane as a key factor in neurodegeneration. *Frontiers in Physiology*, **2013**;188(4):1- 3.
4. Lippens S, Timmerman V, Janssens S. Mitochondria-associated membranes as hubs for neurodegeneration. *Acta Neuropathol.*, **2016**;131:505–523
5.Rima Budvytyte *et al.*, Size-dependent neurotoxicity of beta-amyloid oligomers, *Arch. Biochem. Biophys.*, 2010, 496:84.
6. Rima Budvytyte, Mathias Lösche, Gintaras Valincius. Electrochemical study of interaction of phospholipid membranes with inhibitor- free β-amyloid oligomers. *In preparation*

Scientific Committee

Norman Davey / Ireland

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

Vytautas Smirnovas / VU

Gintaras Valinčius / VU

Rima Budvytytė / VU

Tomas Šneideris / VU

Mantas Žiaunys / VU

Greta Musteikytė / VU

Andrius Sakalauskas / VU

Lukas Krasauskas / VU

Katryna Pampuščenko / LUHS

Venue:

Grand SPA Lietuva

V. Kudirkos g. 45

Druskininkai, Lithuania 66120

Sponsors