Short summary of the proposal

The overall goal of this proposal was to study the mechanism of sarcomeric thin filament assembly with outstanding spatial resolution. We established a single-molecule localization microscopy-based approach and combined it with structure averaging to achieve nanoscale localization precision. To efficiently and reliably perform measurements and the following analysis, we developed the necessary software tools and optimized a pipeline. We performed a comprehensive nanoscopic analysis on the flight muscle myofibrils of adult *Drosophila* and compiled a localization atlas. This analysis revealed the intricate organization of the sarcomeric H-zone and I-band and by using the localization data and template-based protein structure modeling, we assembled a refined sarcomere model with unparalleled scope and resolution. We continue to improve multiple aspects of the approach by identifying pitfalls and potential artifacts. We started to generalize our approach to study myofibrils isolated from mice and developing flight muscles of *Drosophila*. We are also working on multiple projects aiming to better understand how sarcomeres and thin filaments assemble during development.

Subsequently, I will summarize our results according to these four major topics:

1. Establishing structure averaging combined SMLM-based nanoscopy to reconstruct the molecular organization of myofibrils

2. Compiling a comprehensive nanoscale localization atlas of sarcomeric proteins

3. Optimizing nanoscopy to study myofibril organization in vertebrates

4. Analyzing sarcomere and thin filament assembly

1. Establishing structure averaging combined SMLM-based nanoscopy to reconstruct the molecular organization of myofibrils

The structure of the sarcomeres has been studied in detail by X-ray crystallography and with various EM methods leading to quasi-atomic models of the thin and thick filaments from numerous species. However, although these examinations resulted in a remarkably good understanding of the actin-myosin overlap region, the spatial arrangement of the I-band and H-zone complexes remained largely unknown. Advances in fluorescent super-resolution microscopy (also called nanoscopy) provide spatial resolutions that are well below the diffraction limit. Therefore, we decided to utilize these methods to reveal the organization of sarcomeric protein complexes of I-bands and H-zones. During our preliminary studies, we compared the resolution and adaptability of all the major nanoscopic techniques and as a proof of concept experiment, we studied the localization of the KIg16 epitope of D-Titin. The Klg16 epitope forms two stripes along the Z-disc, which are clearly resolvable with immunogold EM but it cannot be resolved as individual structures with confocal laser scanning microscopes. Among the various super-resolution microscopic techniques, STORM provided the highest lateral resolution. We used isolated individual muscle fibers of the flight muscles and to rule out the possibility of sample preparation artifacts, we compared the nanoscopic structures of isolated individual myofibrils to those of dissected intact flight muscles, which revealed nearly

identical patterns, confirming that the individual myofibril preparations are perfectly suitable for our goal. Furthermore, we compared multiple fixation protocols based on different fixatives (formaldehyde, glutaraldehyde, and glyoxal) to achieve reliable and reproducible fine structure preservation. Since organic fluorophores exhibit certain advantageous characteristics, i.e., high brightness and photostability, we used conventional immunofluorescent labeling. We also optimized the dilution for numerous (>40) primary and secondary antibodies to achieve ideal fluorophore blinking density. After establishing sample preparation and immunolabelling procedures optimal for the indirect flight muscle (IFM) myofibrils, we collected super-resolution images of the sarcomeres. As the focal plane is parallel to the myofibrils we acquired two-dimensional projections of the sarcomeres. These revealed the lateral distribution of proteins at the H-zones and I-bands. Typically, an SMLM measurement generates tens of thousands of microscopic images with millions of 'single-molecule events'. To evaluate these large datasets, we developed a standardized data-flow protocol. Since flight muscle sarcomeres are extremely regular structures, we reasoned that it should be possible to further improve the localization precision by combining numerous measurements of individual sarcomeres. Using rotation and translation we spatially aligned the localization information from hundreds of sarcomeres to generate averaged images. By applying structure averaging we can determine the average position of a fluorescent label with a quasi-molecular, ~5-10 nm localization precision. For processing and quantification of the raw localization datasets, we developed a user friendly and freely available software tool (IFM Analyzer) with a detailed user guide, including a sample dataset. (IFM Analyzer software is available at http://titan.physx.u-szeged.hu/~adoptim/?page_id=1246).

-We published a detailed technical paper in Bio-Protocol:
Szikora, Szilard; Novak, Tibor ; Gajdos, Tamas ; Erdelyi, Miklos ; Mihaly, Jozsef: Superresolution Microscopy of *Drosophila* Indirect Flight Muscle Sarcomeres, BIO-PROTOCOL 10 : 12 Paper: e3654, 2020

-We also published a review paper summarizing the major nanoscopic techniques and illustrated how these approaches were used to revisit a series of well-known cellular complexes: Szikora, Szilárd; Görög, Péter ; Kozma, Csaba ; Mihály, József: Drosophila Models Rediscovered with Super-Resolution Microscopy, CELLS 10 : 8 p. 1924, 2021

2. Compiling a comprehensive nanoscale localization atlas of sarcomeric proteins

To better understand the organization of the subdiffraction sized and densely packed H-zone and I-band, we performed a comprehensive analysis, and eventually, we applied the myofibril staining procedure and the SMLM analysis pipeline to more than 10000 sarcomeres of flight muscle myofibrils. Ultimately we detected the precise localization of 27 different epitopes in the flight muscle sarcomeres. The proteins involved were actin, myosin (Mhc), Tropomyosin, Troponin T and C, the actin (+) end (a.ka. barbed end) capping Cpa, the actin (-) end (a.ka. pointed end) capping Tropomodulin (Tmod), the Z-disk organizing α-Actinin, Zasp52 (Z band alternatively spliced PDZ-motif protein 52) and Filamin, the M-line associated Obscurin and

Zormin, the elastic proteins Kettin, Projectin and Sls700, the actin monomer binding Profilin, the formin type of actin assembly factors DAAM (Dishevelled associated activator of morphogenesis) and Fhos, and the actin binding SALS (Sarcomere Length Short) and FliI (Flightless I) proteins. Most of the studied proteins enrich either at the H-zone or the I-band, while a few are present at both places (and along the entire thin or thick filament). From the localization data, we have assembled models of the H-zone and I-band complexes by using the coordinates from the Protein Data Bank and maps from the electron microscopy database, if they were available. In the absence of PDB data, structural information was generated by template based modeling, using the Alphafold or RaptorX servers and annotated protein sequences from Flybase. Aside from generating a high resolution and refined model of the refined I-band and H-zone, our analysis successfully established the position of the key sarcomeric landmarks, including the (+) and (-) ends of the thin filaments, ends of the thick filaments and thus width of the H-zone, the I-band and that of the Z-disk. In addition, we determined the location and orientation of the Tropomyosin-Troponin complex; revealed that α-Actinin forms only two layers in the Z-disk; and placed Filamin into a perfect position to crosslink the parallel running thin filaments of the Z-disk. Moreover, we found that actin regulatory proteins of the H-zone are organized into two distinct spatial domains.

- We published these results in 2020:
Szikora, Szilárd ; Gajdos, Tamás ; Novák, Tibor ; Farkas, Dávid ; Földi, István ; Lenart, Peter ; Erdélyi, Miklós ; Mihály, József: Nanoscopy reveals the layered organization of the sarcomeric H-zone and I-band complexes, JOURNAL OF CELL BIOLOGY 219 : 1 Paper: e201907026, 2020

3. Optimizing nanoscopy to study myofibril organization in vertebrates

Fluorescent nanoscopic methods could offer a significant step forward to answer many of the open questions regarding thin filament elongation and sarcomere formation in vertebrates. Since sarcomeres exhibit evolutionary highly conserved structural features, we started a collaboration with the laboratory of Aniko Keller-Pinter in order to adapt our nanoscopic approach to mammalian models. We initially tried to optimize a protocol by testing the localization of sarcomeric proteins in primary myoblasts (C2C12) with confocal microscopy. However, we found that the reproducibility of these experiments is limited and the selection of the correct cells is time consuming. As an alternative approach, we isolated individual myofibrils from the EDL muscles from six weeks old mice. Isolation of intact myofibrils is relatively easy and yields reproducible structures. We already tested twelve antibodies on these preparations with confocal microscopy, and we found that seven of these (myomesin, α-actinin, tmod1, daam1, profilin, actin, DBP) are suitable for high quality staining. These antibodies label the main sarcomeric reference points, and as a next step, we started to characterize them with dSTORM microscopy. This approach has been optimized for single antibody staining, however, we noticed that size of the H-zones somewhat varies in the vertebrate muscles and for precise STORM measurements we need an independent reference point in each experiment. It implies that we need to apply two-color STORM imaging where phalloidin is used to mark the thin filaments. We decided to use spectrally overlapping fluorophores (Alexa Fluor 647 and

680) for labeling which can be excited and recorded simultaneously, which helps the registration of the two channels. However, the 'unmixing' of the two signals requires hardware and software modifications as well. Optimization of this approach is on the way.

- We summarized the current understanding of how thin filament length is regulated in the form of a review paper, which is currently under peer-review at the International Journal of Molecular Sciences:
Szikora, Szilárd; Görög, Péter; Mihály, József: The mechanisms of thin filament assembly and length regulation in muscles

4. Analyzing sarcomere and thin filament assembly

Molecular mechanisms of actin assembly in developing myofibrils

To describe the assembly/growth of sarcomeres and thin filaments, we performed a careful analysis with confocal microscopy. IFM development takes place during the pupal and early adult stages in multiple phases. The immature myofibrils are assembled simultaneously by 30 hours after puparium formation (APF) and during myofibrillogenesis, the short sarcomeres grow in length and thickness in the following ~80-90 hours to reach their mature muscle mass 24 hours after eclosion (AE). To encompass the whole process, we isolated myofibrils from ten developmental timepoints and systematically quantified sarcomere length and myofibril width. Using thousands of measurements, we compiled a comprehensive model of sarcomere growth. Based on this model we established that sarcomeres grow in a biphasic manner: In the early pupal stages, after the myofibrils are assembled, there is a fast elongation phase, where sarcomere elongation dominates, and later a steady growth phase, where the diameter also increases. Furthermore, combining our measurements with previous results, we were able to assemble the models of the smallest sarcomeres that form during myofibril assembly and a mature sarcomere that is present in the adult fly. These models reveal that the mass of actin filaments grows ~40-fold during myofibrillogenesis. Throughout this process, the length of existing thin filaments increases, and new filaments are added to the periphery of sarcomeres. To directly follow the growth of thin filaments we designed 'pulse chase' experiments: we expressed GFP-labeled actin monomers in a temporally regulated manner and followed their incorporation into the thin filament array. When we expressed actin-GFP temporarily before the onset of myofibrillogenesis with a myoblast/myotube specific driver the labeled actin was only present in the core of the sarcomeres that are assembled right before myofibrillogenesis. Interestingly this labeled core remained stable 3 days after the actin-GFP 'pulse', which suggests that myofibrillar thin filaments are significantly less dynamic than previously suggested. When we expressed actin-GFP after the onset of myofibrillogenesis we observed two distinct patterns: (i) In some cases the whole sarcomere was labeled which presumably corresponds to the sarcomeres that were assembled *de novo* after the onset of myofibrillogenesis. (ii) In the rest of the cases we observed a labeled 'frame' around the unlabeled core. The core presumably corresponds to the initial thin-filament array assembled during myofibril assembly. Conversely, the frame-like structure is the newly synthesized F-actin corresponding to the extension of the initial fibers at their pointed ends, and addition of

new filaments at the circumference of the sarcomere. These observations support the previously proposed pointed end elongation mechanism of thin filament growth. The F-actin based thin filaments are assembled from a polymerization competent G-actin pool. Classically, the G-actin pool has been thought of as homogenous. However, recent work has shown that actin monomers can exist in multiple groups with regulated and distinct subcellular localization. To better understand the growth of thin filaments, we set out to characterize the properties of the myofibrillar G-actin pool(s). We tested multiple G-actin specific probes and in line with the pointed end elongation hypothesis, we found a strong signal in the H-zone of growing sarcomeres which almost completely disappears in the mature sarcomeres. While cytoplasmic G-actin is usually sensitive to detergent extraction, we found that the H-zone actin pool is Triton-X 100 insensitive, which suggests that it is associated with a cytoskeletal component. We also tested its localization with dSTORM microscopy and found that it is localized in a narrow band in the middle of the H-zone completely colocalizing with the formin proteins DAAM and Fhos. To identify the factors involved in the turnover of the sarcomeric actin monomer pool we designed a targeted RNAi screen and currently, we are validating the potential candidates.

Nanoscopic analysis of the localization of SALS during IFM myofibrillogenesis

During IFM myofibrillogenesis sarcomeres grow significantly both in length and in radius by elongation of the existing filaments and addition of new filaments to the periphery of the sarcomeres. The long isoform of the tandem WH2 domain containing Sarcomere Length Short (SALS) protein is expressed in the IFM and required during myofibrillogenesis. Loss of SALS causes shortening of thin filaments due to the lack of actin incorporation at their pointed ends. Reduction of SALS level enhanced the effect caused by the overexpression of the thin filament capping Tmod, as demonstrated by greater thin-filament shortening in cultured myocytes, which suggests that Tmod and SALS antagonize each other functionally during thin-filament assembly at the filament pointed-ends. To assess the physical interaction between Tmod and SALS, we previously performed a nanoscopic analysis in the IFM of adult flies (~24h after eclosion (AE); Szikora et al., 2020) and found that they practically colocalized at the ends of thin filaments. However, to address the developmental roles it was necessary to study the nanoscale distribution of SALS during pupal development. We tested the nanoscopic localizations in growing and mature phases of sarcomerogenesis. While in the mature sarcomeres Tmod and SALS practically 'colocalize', in the earlier, growing phase we found a substantial difference, as the central region of SALS localized significantly closer to the M-line, suggesting a thin filament independent localization. To get a more comprehensive picture on the nanoscopic organization of SALS, we also determined the localization of its N- and C-terminus. For these studies, we used a pair of full-length SALS transgenes FLAG-tagged either on the N- or C-termini and we mapped the position of the FLAG-tag in growing (72 hours APF) and mature (24 hours AE) sarcomeres. We first determined the position of the N-terminal FLAG-tag and found that, as compared to the central WH2 domain containing region, it is located closer to the pointed ends in growing sarcomeres, while in mature sarcomeres it is closer to the center of the H-zone. Contrasting to that, the C-terminal FLAG-tag maintains its localization in close proximity of the pointed ends both in growing and mature sarcomeres, suggesting a continuous interaction between the pointed ends and that of C-term SALS. Given

that the central WH2 domain containing region is associated with the pointed ends only in mature sarcomeres, it might remain free in the growing sarcomeres, possibly to fulfill essential functions in thin filament elongation. These measurements also demonstrated that the conformation of SALS changes significantly as thin filament elongation halts at the end of myofibrillogenesis.

-Collectively, the detailed STORM analysis of SALS revealed a number of interesting aspects of the mechanisms of SALS, that will be summarized in a paper that is already under preparation.

Nanoscopic organization of the Oxoglutarate dehydrogenase complex in the I-band of IFM sarcomeres.

The Zasp/Enigma proteins and α-actinin are essential Z-disc assembly factors. They are required for the first steps of Z-disc formation during myofibril assembly and they localize to nascent Z-discs. Zasp is a scaffold protein that coordinates protein recruitment, while α-actinin cross-links actin filaments from opposite sides of the Z-discs. To identify new proteins recruited to the Z-disc by Zasp and actinin, we took advantage of the observation that interacting proteins often have correlated evolutionary substitution rates, in other words, they coevolve. Our collaborating partners in the Laboratory of Frieder Schöck (McGill University, Montreal, Canada) screened the *Drosophila* proteome for proteins with evolutionary correlated substitution rates with Zasp and actinin. They made GFP-tagged versions of the candidates and assessed their localization at the Z-disc. One of the positive candidates is Oxoglutarate dehydrogenase (OGDH), a key enzyme in the TCA cycle. Their results show that the OGDH mutants have small Z-discs, suggesting a role in sarcomere growth regulation. The OGDH complex is composed of three components: oxoglutarate dehydrogenase (OGDH), dihydrolipoyl succinyltransferase (DLST), and dihydrolipoyl dehydrogenase (DLD). To better understand their molecular function, we set out to determine the nanoscopic organization of all three components and their relative position to the core Z-disk proteins. We used GFP-tagged lines to assess their sarcomeric localization, visualized by anti-GFP immunostaining. We performed a large number of dSTORM measurements and we developed a new software tool that can process sparse localization data without introducing artifacts. Based on these measurements, we found that the OGDH protein fills the I-band region (~200 nm), while DLD and DLST show a double line distribution with a peak intensity close to the edge of the Z-disc (~120 nm) or inside the Z-disk (~60 nm), respectively. Together, these measurements indicate an asymmetric OGDH complex organization, which is similar to the structure of the recently published pyruvate dehydrogenase complex. The mutant analysis also revealed that the function of the sarcomeric OGDH complex is independent of the TCA cycle, as it directly regulates the amino acid metabolism that is essential during the assembly steps of myofibrils.

-The preprint is posted on bioRxiv and under peer-review at eLife:
Nicanor González Morales, Océane Marescal, Szilárd Szikora, Miklos Erdelyi, Péter Bíró, Tuana Mesquita, József Mihály and Frieder Schöck: Oxoglutarate dehydrogenase coordinates myofibril growth by maintaining amino acid homeostasis, doi: https://doi.org/10.1101/2021.12.13.472149