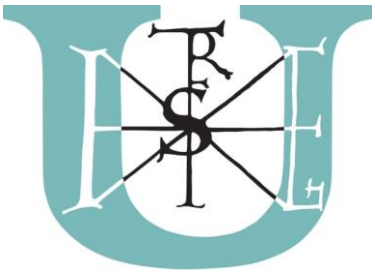

SZENT ISTVÁN UNIVERSITY

Ph.D. thesis

Investigation of *Thermoplasma acidophilum* quinone droplets by lipid and protein analitical methods

Sándor Varga

Gödöllő

2016

Ph.D. School

Name: Ph.D. School of Environmental Sciences

Discipline: Environmental Sciences

Leader of School: Csákiné Dr. Erika Michéli, Ph.D.
Professor, head of department
Szent István University
Faculty of Agricultural and Environmental Sciences
Institute of Environmental Sciences
Department of Soil sciences and Agrochemistry

Supervisor: Dr. István Nagy
Scientist
Max Planck Institute of Biochemistry
Department of Molecular Structural Biology, Planegg, Germany

Dr. Balázs Kriszt
Associate professor, head of department
Szent István University
Faculty of Agricultural and Environmental Sciences
Institute of Aquaculture and Environmental Safety
Department of Environmental Safety and Ecotoxicology

....

Approval of School Leader Approval of Supervisors

Introduction and aims

Microorganisms have influenced the development of biology and biotechnology always in a decisive way. By now the technological methods designed to monitor and manipulate their biochemical and genetic processes provide solution in numerous cases to uncover and utilize the properties inherent in microorganisms that are potentially beneficial for mankind. Carl Woese discovered the archaeas in 1977 and based on the available ribosomal RNA sequences and molecular biological evidences he proposed the partition of living organisms into three domains (Archaea, Bacteria and Eucarya).

Recently the archaeas become the favorite modell organisms of modern structural biology formed in the wake of technological development of imaging methods. Our model organism, *T. acidophilum* is an extremophile, with small cell size which makes it useful for cry-electron tomography (cryo-ET). Cryo-electron microscopic (cryo-EM) analysis of this organism revealed spherical intracellular droplets with 50 nm diameter. Further experiments showed that the spherical droplets unically consist of menaquinones, gulopyranosyl-(β1-1)-caldarchaeol and proteins.

The primary goal of the Ph.D. project was to identify the components of the droplets using biochemical methods, to identify the structure of the droplets by electronmicroscopy (EM) and to determine the environmental conditions (e.g. temperature, pH, aeration) influencing the amount/expression level of the main protein, Ta0547.

During experimental work the isolated *T. acidophilum* proteins proved to be prone to precipitation thus as a secondary objective we tried to achieve the following goals: (1) Screening potential small archaeal modifier protein (SAMP) solubility protein partners. (2) Comparing the solubilization efficiency of SAMPs and to compare it to commercially available, widely used

solubility fusion partners. (3) To investigate the possibility of removing the SAMPs from target proteins using JAMM metalloproteases of the chosen species.

Materials and methods

Isolation of Thermoplasma quinone droplets

To purify Thermoplasma quinone droplets, (TaQDs) we loaded *T. acidophilum* cell extract onto a home made 20 ml hydroxyapatite column (Macro-Prep Ceramic Hydroxiapatite Type I 80 μm, Bio-Rad). Eluted fractions were subsequently purified on a Superose 6 10/300 GL (GE Healthcare) column and proteins of high molecular weight (MW) fractions were loaded onto a 15% SDS PAGE (polyacrylamide gel electrophoresis) to check protein components. The Coomassie stained proteins of interest were cut from gel and analized by MALDI-TOF-MS.

Identification of apolar components of TaQDs

To identify lipid components, purified TaQDs were subjected to thin layer chromatography (TLC), ultraviolet (UV) spectral recording and liquid chromatography/mass spectrometry (LC/MS) analyses. Lipids from both lyophilized *T. acidophilum* cells and TaQD fractions were extracted with a mixture of chloroform, methanol and water for TLC separation. TLC analysis of gulopyranosyl-(β1-1)-caldarchaeol (GuC) and isolated TaQDs was based on standard protocols used for polar lipid based prokaryotic classification at DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen). Extracts were applied to an ALUGRAM SIL G/UV silica gel (Macherey-Nagel) and developed with the same solvent system. UV sensitive spots were visualized by UV light and after that all lipid spots were visualized by

charring. The TLC plate was scanned after spot development and spot intensities were analyzed by the AIDA 2D densitometry software (Raytest Isotopenmassgerate Gmbh, Germany) and standard deviations were calculated.

The spectrum of the UV visible lipids was recorded in a Lambda 40 (Perkin Elmer) photometer in the range between 210 nm to 400 nm.

LC/MS analysis

Isolated TaQDs were extracted by organic solvents and the extract was loaded onto a Waters X Terra-MS C8 column with 2.1x100 mm dimensions. Bound materials were eluted with a gradient of 0.05% TFA in H_2O and 0.05% TFA in acetonitril. The m/z ratios of eluted compounds were recorded by ESI-MS (microTOF LC Bruker Daltonics) in the 100-1600 amu (atomic mass unit) mass range.

Cryo-ET of isolated TaQDs

A solution of purified TaQDs treated with colloidal gold was deposited on holey carbon copper EM grids (Quantifoil) and plunged into liquid propane/ethane mixture cooled by liquid nitrogen. Vitrified samples were imaged on a Tecnai G2 Polara (FEI) cryo-electron microscope operated at 300 kV. IMOD software package was used to process images.

Monitoring the expression level of Ta0547 by Western blot

T. acidophilum cells were grown at standard aerobic conditions and 50 ml aliquots were removed daily over 6 days and analyzed. Cells were lysed and SDS was added to a final concentration of 0.1 %, then stored at -20 °C. The protein concentration was measured by BCA assay (Bio-Rad), thereafter the sample was loaded on a 1D-SDS gel and blotted on a Whatman NC membrane. Western hybridization was carried out with Ta0547 specific serum

antibody and anti-rabbit secondary antibody (IgG-peroxidase (Sigma)). For signal detection ECL solution containing Luminol (Sigma Aldrich) was used. Light intensities were recorded in a LAS3000 imager (Fuji), and intensity values were measured/calculated by Aida V.4.15.025 software (Raytest Isotopenmassgerate Gmbh).

LC-MS/MS analysis of TaQD associated proteins

For MS analysis samples were denatured, disulfide bonds were reduced and the resulting free thiol (-SH) groups were subsequently alkylated with iodoacetamide. The solutions were then digested with peptidases and desalted by using 'Stop and Go' extraction tips (STAGE). All digested peptide mixtures were separated by on-line nanoLC and analyzed by electrospray ionization tandem mass spectrometry. The data analysis was performed using the MaxQuant software supported by Mascot (Matrix Science) database search engine.

To determine MW of intact proteins the Bruker Daltonics microTOF mass spectrometer was used in LC/MS mode. To separate proteins, the HPLC Agilent 1100 machine was equipped with a Phenomenex Aeris WIDEPORE C4 3.6μ (100 x 2.1 mm) column. Proteins were separated with a gradient of 0.05% TFA in H_2O and 0.05% TFA in ACN. UV-chromatogram was recorded at 214 nm.

For N-terminal protein sequence identification the generally accepted automated Edman degradation was used. An Applied Biosystems Model 492 cLC Procise Protein Sequencing System was operated according to the supplier's recommendations.

Bioinformatical methods

Protein homology database searches were carried out using the BLAST algorithm at the National Center for Biotechnology Information (NCBI). To

find/identify transmembrane helices of TaQD associated proteins the DAS-TM filter, Tmpred, Phobius and HMMTOP softwares designed for transmembrane (TM) domain recognition were used. Protein annotation search of Ta0547 was carried out using the KEGG (Kyoto Encyclopedia of Genes and Genomes) database. Secondary structure was predicted by using the Quick2D software. The search for homologue domains was done by HMMER software package and vitellogenin-N hmm (hidden Markov model) profile (PF01347) was used as a template. The Ta0547 vitellogenin domain was modeled using the Modeller software. Template matching was performed by TM-align algorithm.

EM analysis

For negative staining and EM analysis of purified TaQDs and *T. acidophilum* cells the samples were applied to carbon-coated copper grids and treated with 2% uranyl acetate. Electron micrographs in case of TaQDs and *T. acidophilum* cells were acquired with a Philips CM200 and CM20 transmission electron microscope (TEM), respectively. For immunogold labeling *T. acidophilum* cells were embedded in LR white resin according to the supplier recommendations (Fluka), sliced, and then incubated with rabbit polyclonal antibodies and finally labelled by anti-rabbit IgG gold conjugate and stained with 2% aqueous uranyl acetate. For cryo-EM analysis *T. acidophilum* cells were plunge-frozen in liquid ethane and stored in liquid nitrogen. Vitrified samples were analysed by a Philips CM300 TEM instrument. We used EM and TOM software packages for image processing.

Generating *Thermoplasma* SAMP expression vectors

Codon optimized, synthetic Ta0895, Ta1019 and Ta1442 genes (putative SAMP homologs) of *T. acidophilum* were cloned into pET28a vector using *Nde*I/*Xho*I sites to have 6His affinity tag on the N-termini. After

transformation into *E. coli* BL21(DE3) cells (New England Biolabs) target proteins were expressed and sample purity was observed by SDS-PAGE after His-tag protein purification.

Generating JAMM protease expression vectors and cleavage assay for expressed proteases

Codon optimized, N-terminally 6His-tagged JAMM protease genes HVO_2505, Mbur_0623, Nmar_1227 and Ta0623 of *H. volcanii*, M. *burtonii*, *N. maritimus* and *T. acidophilum*, respectively, were cloned into pET28a vector and were transformed into *E. coli* BL21(DE3) cells and target proteins were expressed. Cleavage activity of JAMM1 metalloproteases Hvo_2505, Nmar_1227 and Mbur_0623 was tested on SAMP fused substrates in 20 µl phosphate buffer at different temperatures. Reaction mixtures were directly loaded on SDS-PAGE to visualize cleavage products.

Purification and detection of expressed ScRpn11 fusion protein variants to compare their solubility

The 20S proteasome lid particle protein ScRpn11 was fused to different fusion tags, and expressed at standard conditions. Cell disruption was carried out with an Avestin, Emulsiflex-C5 Homogenizer, the cell extracts were loaded onto a HisTrap HP column with a 1 ml bed volume connected to an ÄKTA basic (GE Healthcare) protein purification system and further purified on a Superdex 200 PC 3.2/30 column (GE Healthcare) with 2.4 ml bed volume. Overall, 50 µl of sample was loaded onto the size exclusion chromatography (SEC) column and eluted. Collected fractions of interest were sent for LC/MS analysis to identify proteins. To visualize and compare the expression levels of the tagged proteins at different stages of purification, Agilent 2100 bio-analyzer and Agilent Protein 80 kit (Agilent Technologies) were used. Data were visualized using the 2100 Expert Software (Agilent Technologies).

Results

Cryo-ET of *T. acidophilum* cells

T. acidophilum cell size spans from 0.5 to 3 μm, with an average thickness of 0.2-0.5 μm, which makes this organism sufficiently transparent and amenable for structural studies by means of cryo-ET. Tomograms of anaerobically grown *T. acidophilum* cells revealed a large number of cytoplasmic, globular particles. Image analysis methods showed spherical geometry for the particles and their evenly distribution in the cell, furthermore 35-50% of the *T. acidophilum* cell volume was filled with particles. The particles have a Gaussian size distribution with a mean value of 50 nm and a standard deviation of 15 nm.

Identification of Ta0547 the main protein component of TaQD

Cytosolic proteins of *T. acidophilum* lysate were subjected to column chromatography based purification, and proteins of high molecular weight fractions were identified by MALDI-TOF-MS. Thereby, we were able to identify Ta0547 as dominant protein of void volume fractions. Notably, while other large protein complexes were present only in negligible quantities, negative-stained EM images of these fractions showed many, large, globular particles which were reminiscent in size and shape of inclusion particles seen in cryo-ET.

To confirm this result Ta0547 specific antibodies were raised in rabbit, using recombinant Ta0547. Negative-stained EM analysis of immuno-purified, native Ta0547 showed globular shape inclusions similar in size and shape to the ones isolated with column chromatography. Overall, these results show that Ta0547 is indeed the major protein component of observed *T. acidophilum* inclusion particles.

Bioinformatic analysis of Ta0547 protein

Databases cite Ta0547 as an uncharacterized protein. Secondary structure feature prediction searches/methods predicted it as soluble protein and neither signal peptide nor transmembrane domain(s) could be detected. The secondary structure of Ta0547 consists mainly of α-helices and small β-sheets, and to a lesser extent of disordered regions at the N-terminus. Ta0547 is annotated as a vitellogenin-N domain protein by the KEGG database, however the Pfam database search with Ta0547 did not retrieve vitellogenin-N domain as hit. To solve this contradictory result, a profile-sequence domain search was performed locally with the vitellogenin-N hmm-profile (Pfam: PF01347) against the *T. acidophilum* proteome. This returned only Ta0547 as hit, Ta0547 matches with a short C-terminal portion of the PF01347 domain. This result triggered us to obtain a preliminary 3D model of Ta0547 using the high-resolution crystal structure of vitellogenin-N domain of lipovitellin (PDB: 1LSH) as template. Superimposition of Ta0547 model against the lipovitellin structure performed by the TM-align server showed almost identical fold (from residues 313 to 435).

Identification of TaQD associated proteins by LC/MS/MS and Edman degradation

To discriminate between genuine inclusion particle associated proteins and contaminations detected by highly sensitive proteome analysis methods, fractions from both immuno-precipitation and column chromatography experiments were subjected to LC- MS/MS analysis. We detected 7 proteins matching the search criterias, namely Ta0547, Ta0337, Ta0437, Ta0438, Ta0182, Ta0093 and Ta1223a, which are the most likely components of T. acidophilum inclusion particles. These proteins except Ta0437 are smaller than 20 kDa. Ta0547, Ta0182, Ta0438 and Ta0337 are without known

functions, Ta1223a showed homology to archaeal Sec-independent twin-arginin translocases, while Ta0437 is homologous to anion transporting and oxyanion translocating ATPases found in many other organisms. From MS based quantification of small proteins the only conclusion could be drawn that Ta0547 protein is the dominant one.

TLC, UV spectroscopy and LC/MS analysis of non protein components of TaQDs

Since the appearance of isolated particles resembled LDs from other organisms we carried out experiments to identify possible lipid constituents by means of thin layer chromatography (TLC). The commonly used Bligh-Dyer total lipid extraction method was used to extract lipid components of both lyophilized *T. acidophilum* cells and isolated inclusion particles and TLC separation and spot visualization experiments were carried out. After TLC plates were developed they were subjected to a UV light source to monitor UV detectable spots whereas non UV detectable spots appeared only after charring with H_2SO_4. The polar lipid pattern of the cell extract corresponded well with published results revealing two major lipids, the gulopyranosyl-(β1-1)-caldarchaeol (GuC/U4) and the gulopyranosyl-(β1-1)-caldarchaetidylglycerol (GuCGp), the latter being dominant. The lipid profile of purified TaQDs was basically different compared to total cell extract. Two spots, the dominating yellowish UV visible spot migrating at the solvent front and a less intense spot the bit slower migrating U4 (GuC) compund were detected only

To identify the UV visible compounds solvent extract of purified *T. acidophilum* inclusion particles were analyzed by ultraviolet absorption spectroscopy. Absorption maxima were recorded at 325 nm, 269 nm, 260 nm, 248 nm and 243 nm reminiscent of menaquinones of which vitamin K2 (MK-

4) is the most prominent group member. On the basis of this result, a liquid chromatography separation method was employed to discriminate UV visible compounds and eluted peaks/compounds were analyzed by mass spectrometry. The main component of the inclusions was found to be menaquinone-7 (MK-7), and other reported quinones thermoplasmaquinone-7 (TPQ-7), methionaquinone-7 (MTK-7) and menaquinone-4 (MK-4) of the *T. acidophilum* HO-62 strain were also found.

Structural characterization of TaQDs by crio-electron tomography

To gain further insights into TaQD structure, TaQDs purified by SEC were imaged by cryo-ET. Consistently with our in-situ observations, TaQDs were roughly spherical and 25-60 nm in diameter. Most TaQDs had an electron dense lumen surrounded by a denser boundary. In some cases, a smaller spherical structure with clear lumen was attached to the larger dense lumen droplets. TaQD boundary was not continuous, as it appeared to be formed by the clustering of discrete globular densities of ~ 4 nm in diameter, suggesting that this boundary may not be of lipidic nature. Filamentous structures anchored to the droplet boundary were also observed projecting outwards.

The structural role of Ta0547

In order to investigate the structural role of Ta0547 in particle formation and/or stabilization, we carried out EM experiments on TaQDs treated by destructive compounds like trypsin and detergents sodium dodecyl sulfate (SDS), and n- Dodecyl-β-D-maltoside (DDM). Interestingly, DDM (10%), a detergent used for membrane protein isolation, did not affect particle structure, thus indicating the lack of a lipid monolayer. By contrast, treatment with trypsin led to particle disintegration, which is presumably due to Ta0547 degradation. These results led us to assume a major structural role for Ta0547 in the architecture of *T. acidophilum* inclusion particles. To localize

intracellular Ta0547 thin sections of *T. acidophilum* cells were labeled by immunogold staining. We found that Ta0547 proteins were located mostly on the surface of inclusion particles that is in good agreement with EM analyses, which could detect neither a protein coat nor a lipid monolayer.

Monitoring changes in the expression level of Ta0547

After identifying inclusion particle components we attempted to assess the possible physiological role of the inclusions by monitoring if any change in the expression level of Ta0547 might occur in response to environmental challenges. Therefore, several conditions such as elevated pH, aerobic and anaerobic growth, and growth period were screened by Western blot analysis. Interestingly, we found that none of the tested environmental conditions had effect on Ta0547 expression level, at least in two days old cultures. However, the amount of Ta0547 increased continuously over a 6 day long growth period.

Expression tests of putative *T. acidophilum* SAMPs in *E. coli*

Codon optimized SAMP homologs Ta0895, Ta1019 and Ta1442 of *T. acidophilum* in *E. coli* were highly expressed and the purity of the samples was close to homogeneity already after His affinity purification. Ta0895 was the best candidate to use it as fusion tag as it was expressed at high level in pure form in *E. coli*, and it was the most abundant when its expression level was compared to the expression level of other SAMPs in *T. acidophilum*. For protein tagging purposes the synthetic/codon optimized version of Ta0895 was used.

Expression of ScRpn11 protein fused to commercial solubility tags and SAMPs to compare their impact on solubility

To determine the relative efficiency of Ta0895 considering protein expression level and solubility it was compared to three putative archaeal SAMPs, namely

Ta1019, Hvo_2619 and Mbur_1415, and commercial tag proteins such as 6His, Trx, GST, MBP, SUMO and NusA-tag. ScRpn11 was fused to the C-terminus of these tags and expressed. Yields of expressed proteins were measured after every purification step. Recombinant ScRpn11 solubility was determined from cell extracts based on the percentage of recombinant protein/total protein signal strength. After His-tag purification only MBP, NusA, Ta1019 and Ta0895 tagged Rpn11 samples provided suitable protein amounts for SEC purification. Seemingly the protein yield of best performing NusA-ScRpn11 and MBP-ScRpn11 were higher compared to SAMP samples but we have to take into consideration that the molecular weight of SAMP tagged fusion proteins is significantly lower than NusA-ScRpn11 and MBP-ScRpn11. Therefore we corrected the values with the molecular weight of the fusion partner according to the following calculation: [target protein MW/(target protein MW + tag protein MW)] * (target protein amount/total protein amount).

Screening archaeal metalloproteases for SAMP cleavage activity

Aiming to enzymatically remove SAMP fusion tags in screening experiments JAMM1 metalloproteases Hvo_2505, Nmar_1227 and Mbur_0623 were mixed with SAMP-ScRpn11 and SAMP-Ta0547 substrate variants and incubated at different temperatures. Denaturing SDS PAGE analysis of treated samples indicated cleavage products when Hvo_2505 was mixed with Ta0895_VAGG-Rpn8-Rpn11 and Hvo_2505 mixed with Ta0895_VSGG-Ta0547. Hvo_2505 was active at 24–37–48 °C but no cleavage activity was detected at lower (10 °C) temperature. The other two enzymes did not show activity at given test conditions.

Summary of new scientific results

TaQDs similar in appearance to eukaryotic lipoprotein particles were isolated and characterized for the first time.

Thesis 1: The main protein component of isolated Thermoplasma quinone droplets is Ta0547. Based on the bioinformatical analysis Ta0547 protein is a vitellogenin-N domain homologue.

Thesis 2: The main lipid components of isolated Thermoplasma quinone droplets are menaquinone-7, gulopyranosyl-(β1-1)-caldarchaeol and in smaller quantities they contain methylmenaquinone, methionaquinone-7, thermoplasmaquinone-7 and menaquinone-4 compounds.

Thesis 3: Based on electronmicroscopic imaging the quinone droplets are approximately spherical, their diameter is 25-60 nm and are bordered by discrete globular densities. The Ta0547 protein plays a crucial role in maintaining the proper structure.

Thesis 4: The expression level of Ta0547 increased continously during a 6 days long cell culture growth period.

Thesis 5: Expression tests in *E. coli* proved that SAMPs of *T. acidophilum* are able to solubilize the ScRpn11 target protein.

Thesis 6: Considering the comparative tests Ta0895 SAMP was the most efficient solubility partner amongst the investigated proteins.

Thesis 7: The specific removal of Ta0895 fusion partner is possible using the Hvo_2505 metalloprotease derived from *H. volcanii*.

Conclusions and proposals

The primary aim of the Ph.D. project was the isolation and biochemical/structural characterization of quinone droplets discovered in *T. acidophilum.* TaQDs were purified by two independent methods and sample purity was monitored by SDS-PAGE, TLC and EM analyses. These experiments revealed that TaQDs were purified close to homogeneity, as samples from aerobically and anaerobically grown cells purified by immuno-precipitation and/or column chromatography provided the same protein and lipid profile, respectively, while they were significantly different from total cell lipid and protein profiles.

TaQD associated proteins were identified and cross referenced by several proteome analysis methods. The identified protein profile differs from the composition of any known lipid droplet or lipid based cell components. The LC-MS/MS based protein identification method found Ta0547 as major and Ta1222a, Ta0438, Ta0093, Ta0182, Ta0337, and Ta0437 as minor TaQD associated proteins. Protein TM domain analyses implied that Ta1223a is probably and Ta0337 may be a TM protein, whereas the other proteins do not possess TM domains.

TLC analysis of isolated TaQDs identified two major spots. The more intense spot running at the solvent front was UV detectable and later analyses identified it as a mixture of menaqinones reminiscent to MMK, TPQ-7 and MK-7, the latter one being dominant. This finding was surprising because isoprenoid ubiquinones and menaquinones are lipophilic, non-protein components of prokaryotic electron-transfer chains shuttling electrons between membrane-bound protein complexes. Non cell membrane-embedded quinone accumulation has been demonstrated in plastoglobules of plastids from plants and algae. Based on these findings it was suggested that plastoglobules act as reservoirs to enable a rapid response to environmental

conditions at which the increase or decrease of these compounds are needed to maintain healthy photosynthetic homeostasis. Contrary to plastoglobules, TaQDs are not covered by lipid layer, instead, they have an electron dense lumen surrounded by a denser, discontinuous boundary formed by a cluster of discrete globular densities of ~ 4 nm in diameter, suggesting that this boundary is rather of protein than of lipidic nature.

Observed similarities between TaQDs and lipoprotein particles are limited only to morphology level as major associated proteins (TaQD proteome vs apolipoproteins) and lipids (menaquinones and U4 polar lipid vs TAG, cholesterol and phospholipids) are different. However, one might speculate that TaQDs take part in lipid trafficking similar to lipoprotein particles. This idea is supported by bioinformatical analysis of Ta0547 which is a vitellogenin-N domain protein by profile-sequence domain search with the Pfam PF01347 domain hmm against the *T. acidophilum* proteome. A structural model was built for Ta0547 using the lipovitellin crystal structure (1LSH) as template and we found that this model exhibits the same fold as the lipovitellin helical domain from residues 313 to 435.

We evidenced that TaQDs undergo dynamic changes regarding menaquinone composition and Ta0547 protein amount elevation during a 6 day long growth period. Besides of the biochemically detectable changes we were able to observe visible changes of TaQDs, too. In 7 day old cultures we observed much less TaQDs than in 2 days-old cells, whereas the size of few particles has dramatically increased. Besides, aggregation of TaQDs in immuno-stained cells, and QD merging at *in vitro* conditions could also be detected. Additionally, we detected budding/merging TaQDs which might be indicative for active biochemical processes of a primitive cell organelle.

To define the biochemical role of TaQDs and their associated proteins and lipids future experiments will be needed: KO mutagenesis of genes of interest,

solving tertiary and quaternary protein structures at atomic resolution, identification most of the lipid components by LC-MS supported lipid profiling and measurement of quantitative changes in lipid composition of cell membrane and TaQDs upon changes in environmental factors.

The secondary goal of the Ph.D. project was to find potential solubility fusion protein partners suitable to purify TaQD proteins. Based on homology searches we have set up expression tests in *E. coli* for putative *T. acidophilum* SAMPs: Ta0895, Ta1019 and Ta1442. To test the theory ScRpn11 was fused to the C-terminus of Ta0895. The expression trials provided evidence that this SAMP could keep the target protein in solution and inspired us to test other SAMPs and commercially available tags as well. A comparative analysis of soluble protein concentration values of ScRpn11 fused to four putative SAMPs Ta0895, Ta1019, Hvo_2619 and Mbur_1415 and six widely used tags 6His, Trx, GST, MBP, SUMO and NusA was carried out. In the first part of the experiment the commercially available tags performed significantly better, however concentration values measured for Sumo-ScRpn11, GST-ScRpn11, 6His-ScRpn11 and Trx-ScRpn11 in the supernatant after centrifugation indicate that these proteins precipitate during the expression and form inclusion bodies. After His purification only four fusion proteins, MBP-ScRpn11, NusA-ScRpn11, Ta1019-ScRpn11 and Ta0895-ScRpn11 provided enough material for SEC purification, since they had high soluble recombinant protein/soluble total protein ratios. However, the correction with the molecular weight of the tag proteins has changed the rank in favor of Ta0895 and Ta1019.

For certain experiments native (tag-less) proteins are required. To remove SAMP solubility tags we attempted to find enzymes capable of cleaving off the tags at specific sites, when necessary, and possibly at low temperature. Four JAB domain metalloproteases and their putative substrates were selected

from four archaeons: *H. volcanii* (Hvo_2619 and Hvo_2505), *N. maritimus* (Nmar_1227), *M. burtonii* (Mbur_1415 and Mbur_0623) and *T. acidophilum* (Ta0895, Ta1019 and Ta0623). These organisms are adapted to different temperature ranges and we were interested especially in the psychrophilic *M. burtonii* which can tolerate even 1-2 ° C. The *Haloferax* Hvo_2505 enzyme was the only one that could cleave some of the linearly fused SAMPs. Although this way SAMP release is possible but the conditions required for enzyme activity are far from ideal (2 M NaCl concentration and temperature 24-48 ° C). Identification of proteases degrading the target protein and selection/generation of host cells without the protease coding gene(s) and a Ta0895 specific protease active at low temperature may be the next steps to further improve the Ta0895 based solubility system.

Publications

Scientific paper:

Nagy, I., Knispel, R. W., Kofler, C., Orsini, M., Boicu, M., Varga, S., Weyher-Stingl, E., Sun, N., Fernandez-Busnadiego, R., Kukolya, J., Nickell, S. and Baumeister, W. (2016) 'Lipoprotein-like particles in a prokaryote: quinone droplets of Thermoplasma acidophilum', *FEMS*, doi: 10.1093/femsle/fnw169

Varga, S., Pathare, G. R., Baka, E., Boicu, M., Kriszt, B., Székács, A., Zinzula, L., Kukolya, J. and Nagy, I. (2015) 'Enhancing recombinant protein solubility with ubiquitin-like small archeal modifying protein fusion partners', *Journal of Microbiological Methods*, 118, pp. 113–122. doi: 10.1016/j.mimet.2015.08.017.

Communications published in congress publications:

Sándor Varga, Ganesh Pathare Ramnath, Erzsébet Baka, Marius Boicu, Balázs Kriszt, András Székács, József Kukolya, István Nagy (2015): ENHANCING RECOMBINANT PROTEIN SOLUBILITY WITH A SAMP FUSION PARTNER Hungarian Molecular Life Sciences 2015: Programme & Book of Abstracts. 304 p. (ISBN:978-615-5270-15-4)

Sándor Varga, Roland W Knispel, Christine Kofler, Marius Boicu, E Weyher-Stingl, Erzsébet Baka, József Kukolya, Stephan Nickell, Wolfgang Baumeister, István Nagy (2014): Blobology: The case of Thermoplasma quinone droplets. ACTA MICROBIOLOGICA ET IMMUNOLOGICA HUNGARICA 62 (Suppl.1), p. 79. 1 p. (Print ISSN: 1217-8950 Online ISSN: 1588-2640)