UNIVERSITY OF THESSALY & NETHERLANDS CANCER INSTITUTE

ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ & ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ:

«Εφαρμογές Μοριακής Βιολογίας-Μοριακή Γενετική-Διαγνωστικοί Δείκτες»

Expression, purification and crystallization of a Vibrio cholerae putative lipase

---

Έκφραση, καθαρισμός και κρυστάλλωση μιας θεωρούμενης λιπάσης από το Vibrio cholerae

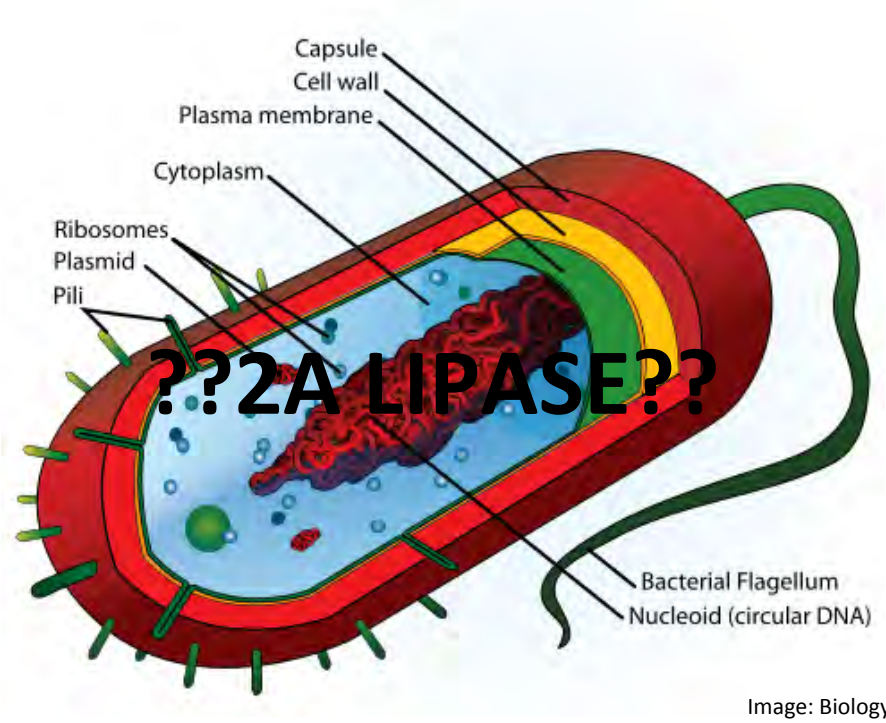


Image: BiologyGuide website

Vasiliki Kolokotroni

October 2014- July 2015

Expression, purification and crystallization of a *Vibrio cholerae* putative lipase

Έκφραση, καθαρισμός και κρυστάλλωση μιας θεωρούμενης λιπάσης από το *Vibrio cholerae*

Τριμελής επιτροπή

- Δρ. Λεωνίδας Δ. Δημήτριος

Αναπληρωτής καθηγητής Βιοχημείας του τμήματος Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας

- Περράκης Αναστάσιος

Ερευνητής-Group leader στο Netherlands Cancer Institute, Amsterdam, Ολλανδία

- Σκαμνάκη Βασιλική

Λέκτορας Βιοχημείας και Μεταβολισμού του τμήματος Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας

Περίληψη

Οι Hrev107 βρέθηκαν σε ευκαρυώτες, picorna ιούς και στο βακτήριο *vibrio cholera.* Οι δομές αυτών των πρωτεϊνών είναι άγνωστες. Κρυσταλλογραφικά πειράματα διεξήχθησαν για το δομικό προσδιορισμό της πρωτεΐνης στο vibrio. Οι πρωτεΐνες που χρησιμοποιήθηκαν ήταν είτε πλήρους μήκους είτε truncated. Αρχικά, οι πρωτεΐνες εκφράστηκαν στο *E.coli* και στη συνέχεια καθαρίστηκαν. Το τελικό βήμα περιελάμβανε την κρυσταλλογραφική ανάλυση. Οι διαφορές που παρατηρήθηκαν στη διαλυτότητα των truncated constructs οφείλονται στην υδροφοβική φύση του καρβοξυτελικού άκρου της πρωτεΐνης. Οι διαφορές αυτές αντανακλούν αυξημένη απόδοση στον καθαρισμό της πρωτεΐνης, γεγονός που επιβεβαιώνεται από το σχηματισμό spherulites και needle crystals.

Abstract

Hrev107-like proteins were found in eukaryotes, picorna viruses and the bacterium *vibrio cholerae*, being the exception. The structures of these proteins are unknown. Crystallization experiments were performed for the structural determination of the vibrio protein. The constructs were both Full Length and truncated. Firstly, the proteins were expressed in *E.coli* and then they were purified. The last step involved the crystallization analysis. The differences observed in the solubility of the short constructs might be attributed to the absence of the C-terminal. The C-terminal part of the protein seems to be hydrophobic which means that in its presence the entire protein becomes insoluble. These differences lead to high purified protein yield, a fact confirmed by the growth of spherulites and needle crystals.

Table of Contents

INTRODUCTION 7
AIM 10
EXPERIMENTAL 11
RESULTS & DISCUSSION 21
CONCLUSION 66
FURTHER RESEARCH 67
SUPPLEMENTAL 68
REFERENCES 71
ACKNOLEDGEMENTS 72

Table of abbreviations

A.I.: Auto-Induction
Amp: Ampicillin

BTP: Bis-Tris Propane

DNA: Deoxyribonucleic acid
DOPC: 1,2-dioleoyl-sn-glycero-3-phosphocholine
DOPG: 1,2-dioleoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]

FL: Full Length
F.T.: Flow Through

G.F.: Gel Filtration
GST: Glutathione S-transferase

H_N_T_ : Hepes_ NaCl_ TCEP_ (_: concentration, expressed into mM)
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IEX: Ion Exchange
IPTG: Isopropyl β- D -1-thiogalactopyranoside

Kan: Kanamycin

LB: Luria Bertani
LRAT: Lecithin Retinol Acyltransferase

MALLS: Multi-Angle Laser Light Scattering
MES: 2-(*N*-morpholino)ethanesulfonic acid

O/N: Over Night

Rcf: relative centrifugal force (G-force or force per gravity)
RNA: Ribonucleic acid
Rpm: revolutions per min

SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC: Size Exclusion Chromatography
SENP: Sentrin-specific protease
SN: supernatant
SUMO: Small Ubiquitin-Like Modifier

TCEP: tris(2-carboxyethyl)phosphine

T_N_T_ : TRIS_ NaCl_ TCEP_ (_: concentration, expressed into mM)

Vol: volume

1. Introduction

Proteomics

Proteomics is the large-scale study of proteins devoted to define their structures and understand their function. A vast number of studies and experiments have been reported to explain the "world of the proteins". The determination of function combined with the determination of structure can reveal the protein profile in detail. Since both the function and the structure are unknown, the evolutionary history and the primary sequence analysis could provide some information about the protein profile. The proteins of interest could come from eukaryotes, virus, bacteria and archaea. The interest for studying these proteins may be motivated by the fact that they are involved in several disorders, the unknown function or structure, the possible interactions with proteins of catalytic role, their homology with essential proteins of unknown function.

In this study, we sought to understand the evolutionary history of the 2a lipases into different genera. The sequence, length and infection-specific roles of 2a proteins vary considerably among picornavirus genera. The high degree of conservation and maintenance of this protein across the different genera suggests that it performs an important function. The studies on the structure and on the function of this protein family could throw light and help in uncovering unknown mechanisms related with this lipase.

Picorna viruses

The picornaviruses are positive sense RNA viruses which include several human and animal pathogens(Hughes & Stanway 2000). More than 200 serotypes have been identified and they have categorised into nine genera (Oberste et al. 2001): Enterovirus (89 serotypes), Rhinovirus (103 serotypes), Cardiovirus (3 serotypes), Apthovirus (8 serotypes), Hepatovirus (2 serotypes), Parechovirus (2 serotypes), Teschovirus (10 serotypes), Kobuvirus (1 serotype) and Erbovirus (1 serotype). The genome organization of picornaviruses is canonical, highly conserved, and a defining characteristic of the family. Their RNA is single stranded, positive sense, infectious, monocistronic and highly structured. The genome length could be from 7032 bases (Avian Encephalomyelitis virus, a Tremovirus) to 8828 bases (Erbovirus), with an average of 7600 bases (Palmenberg et al. 2010). Their genome encodes a poly-protein, which undergoes a cleavage cascade performed by virus-encoded activities, to give the final virus proteins. The genomes of the various genera encode final proteins, but intermediates in the cascade can have a significant half-life and may have distinct functions (Patton & Spencer 2000). All the picornaviruses have the same genome organization. Picornavirus proteins and their precursors take their name from sequential location in the polyprotein ORF. The capsid proteins (1A, 1B, 1C and 1D, known as VP4, VP2, VP3 and VP1, respectively) are encoded towards the N terminus of the polyprotein and the non-structural proteins (2A, 2B, 2C, 3A, 3B, 3C and 3D) are encoded downstream of these. VP4 and VP2 are assembled into the particle in the form of precursor, VP0, the final step of assembly being the cleavage of this precursor which appears to be required for virion infectivity and stability. However, in the human parechoviruses and in Aichi virus, this cleavage appears not to occur and consequently these particles have only three proteins, VP0, VP3 and VP1 (Hyypiä et al. 1992)(Hughes & Stanway 2000) (Yamashita et al. 1998). The majority of these picornavirus proteins show an homology between the protein and the encoding region, though. Two exceptions have already mentioned. The protein encoded

at the N terminus of the polyprotein in the apthovirus, cardioviruses, ERBV and Aichi virus. The other exception is the region that encodes 2A protein. Pamela J. Hughes, Glyn Stanway et al, found that the 2A protein of several diverse picornaviruses, the human parechoviruses, Aichi virus and avian encephalomyelitis virus, possess conserved motifs and seems to have a common function. These motifs, a conserved histidine, an asparagine-cysteine dipeptide and a putative transmembrane domain, are characteristic of a superfamily of cellular proteins (Hughes & Stanway 2000).

Members and features of NIpC/P60 superfamily

NIpC/P60 is a large superfamily encompassing several diverse groups of proteins. In addition to the well characterized P60-like proteins, this superfamily includes the AcmB/LytN and YaeF/YiiX families of bacterial proteins, the amidase domain of bacterial and kinetoplastid glutathionylspermidine synthases (GSPSs), and several proteins from eukaryotes, phages, poxviruses, positive-strand RNA viruses, and certain archaea. The eukaryotic members include lecithin retinol acyltransferase (LRAT), nematode developmental regulator EgI-26, and candidate tumor suppressor Hrev-107. These eukaryotic proteins, along with the bacterial YaeF/poxiviral G6R family, show a circular permutation of the catalytic domain. Three conserved residues, namely a cysteine, a histidine and a polar residue are all found to be involved in the catalytic activities of this superfamily. Evolutionary analysis of this superfamily shows that it comprises four major families, with diverse domain architectures in each of them (Anantharaman & Aravind 2003). Multiple paralogous proteins with NIpC/P60 catalytic domains are represent in most bacteria, suggesting that the NIpC/P60 family is a peptidase family with a widespread role in the dynamics of the bacterial cell wall. Although, as it is described before, several members of this superfamily exist outside the bacterial superkingdom, in eukaryotes, large DNA viruses, positive-strand RNA viruses and certain archaea. In eukaryotes, one of the members of this family has been studied experimentally, and possesses LRAT activity rather than peptidase activity. These eukaryotic versions, along with certain bacterial forms, show a circular permutation of the domain which results in a swapping of the positions of the catalytic cysteine and histidine residues in the sequence(Anantharaman & Aravind 2003).

Lecithin Retinol Acyltransferase (LRAT) family

This family is the only biochemically characterized eukaryotic representative of the superfamily. LRAT transfers an acyl group from the sn-1 position of phosphatidylcholine (lecithin) to retinal, to form retinyl esters. The general base of the enzyme (cysteine) accepts the acyl group from lecithin, and this is followed by retinol being directed to attack the acyl-enzyme by the catalytic histidine followed by the release of the retinyl ester. The phylogenetic relationship between the LRAT family of enzymes and thiol peptidases shows that this catalytic triad was derived only once in evolution and re-used for different mechanistically similar reactions such as acyl-transfer, transacylation or peptide-bond hydrolysis. The LRAT family is thus far found only in eukaryotes and animal viruses, with the sole exception of a single member from the proteomes of *Vibrio cholerae* and *Anabaena*. The LRAT and its obvious orthologs are found, in addition to the vertebrates, in other animals and plants. In vertebrates the enzyme has an important role in the storage and mobilization of retinol (vitamin A) as esters in peripheral tissues and

for generating an intermediate in the synthesis of the rhodopsin chromophore in the visual tissues (Anantharaman & Aravind 2003).

H-rev107, a subfamily of the LRAT

H-rev 107, also called HRASLS3 or PLA2G16, was firstly identified in H-RAS-resistant murine fibro-blasts. H-rev 107 is a member of the HREV107 type II tumor suppressor gene family, which includes H-REV107, retinoid-inducible gene 1 (RIG1), HRASLS2, HRLP5, and HRASLS (A-C1), the last of which is present in humans, rats and mice. In this family, the protein contains a NC (asparagine -cysteine) domain at the N-terminus and a hydrophobic membrane anchoring domain at the C-terminus. In general, the proteins in this family exhibit activities that regulate cellular growth, differentiation and apoptosis. Some research results provide information about the molecular mechanisms underlying the biological functions of HREV107 family proteins(Wang et al. 2014). It has been shown that H-rev 107 could potentially act as an acyl-transferase that might modify membrane components. This fact is supported by the observation that Hrev is associated with the cell membrane (Anantharaman & Aravind 2003).

Vibrio cholera and an evolutionary approach

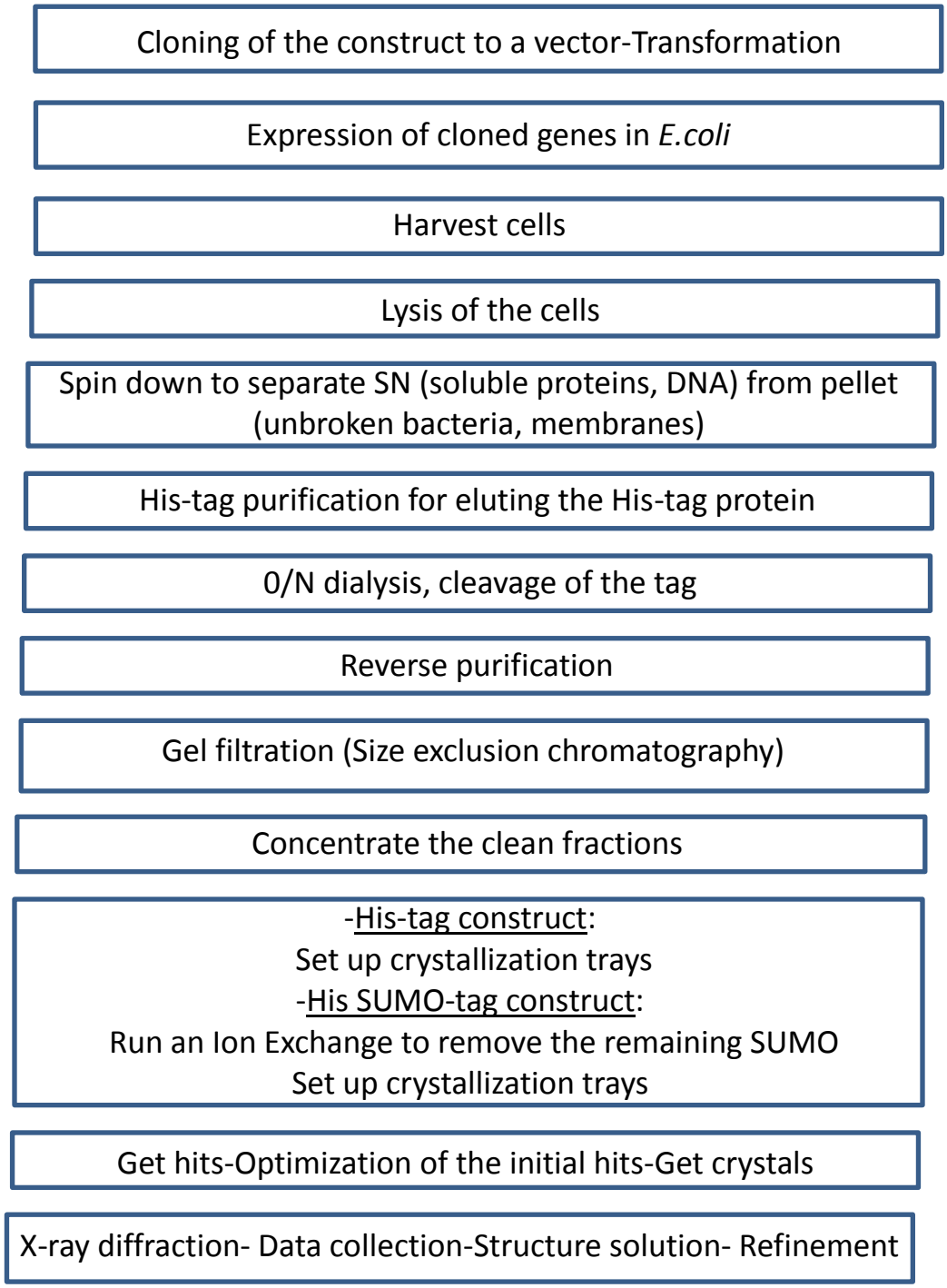
The viral homologs of the H-rev 107 were observed in picorna viruses such as human parechoviruses, Aichi virus and avian encephalomyelitis virus. These proteins could potentially function as a second protease in the processing of viral polyproteins or they could function as a viral enzyme that could modify some membrane component, like their cellular homologs. The widespread distribution of the LRAT family in eukaryotes suggests that they possessed the enzyme from an early stage in their evolution. The NIpC/p60 superfamily is largely absent from most archaeal proteomes. However, it is very widespread in bacteria and includes forms with circular permutations similar to the eukaryotic LRAT-like proteins. Hence, the LRAT family was probably acquired by the eukaryotes, by the lateral transfer of a bacterial precursor similar to the YaeE/YiiX family. This could have either occurred during the primary γ- proteobacterial endosymbiosis, which gave rise to the mitochondrion or as a result of a subsequent transfer from some bacterial source. The currently available sequences do not allow us to distinguish between these possibilities. A homolog of Hrev 107 is detectable in the pathogenic bacterium *v.cholerae* and phylogenetic analysis suggests that it has been acquired through lateral transfer from a vertebrate source. Further experimental analysis of this bacterial protein could be of considerable interest, as it could potentially have a role in interactions with its vertebrate host. The drastic change in biochemical function might have resulted in an accelerated evolution of these eukayotic proteins, resulting in considerable divergence from the bacterial YaeF/YiiX-like forms(Anantharaman & Aravind 2003).

2. Aim of work

The aim of this work was the determination of the structure of a putative lipase from *vibrio cholerae*. This protein is found only in eukaryotes and animal viruses but also at the *vibrio cholerae*, being the exception. The *vibrio cholerae* lipase is an ortholog of the viral and the eukaryotes lipases with 38-40% homology. As a result, the studies on the *vibrio cholera* lipase may offer some insights on the lipase evolutionary history. It may also provide an answer as to its existence at this bacteria. Experimental work on this lipase was facilitated by the fact that this protein is stable. Therefore, a variety of biochemical methods can be applied for its study.

3. Experimental

3.1 Work flow



3.2 Constructs

The His-tag full length Hrev *vibrio cholera* lipase, the His-tag Hrev vibrio 1-130 lipase and the His-tag Hrev vibrio 1-127 lipase were amplified and cloned through ligation independent cloning in the vectors Lic 1,Lic 1 and Lic 1.1, respectively. Additionally, the His-SUMO vibrio full length lipase, the His-SUMO vibrio 1-130 lipase and the His-SUMO vibrio 1-118 lipase were amplified and cloned through ligation independent cloning in Lic 1.10 vectors. Moreover, the GST-tag vibrio full length lipase and the GST-tag vibrio 1-130 lipase were amplified and cloned through ligation independent cloning in Lic 8 vectors (Luna-Vargas et al. 2011).

Constructs	
His-tag F.L.	T7 · 6xHis · 3C · Lic 1 · Col E1 · Kan
His-SUMO tag F.L.	T7 · His-SUMO2 · SENP2 · Lic 1.10 · Col E1 · Kan
GST-tag F.L.	T7 · GST · 3C · Lic 8 · Col E1 · Amp
His-tag 1-130	T7 · 6xHis · 3C · Lic 1 · Col E1 · Kan
His-SUMO tag 1-130	T7 · His-SUMO2 · SENP2 · Lic 1.10 · Col E1 · Kan
GST-tag 1-130	T7 · GST · 3C · Lic 8 · Col E1 · Amp
His-tag 1-127	T7 · 6xHis · 3C · Lic 1.1 · Col E1 · Kan
His-SUMO tag 1-118	T7 · His-SUMO2 · SENP2 · Lic 1.10 · Col E1 · Kan

The table 1 depicts the constructs used for the structural studies of vibrio lipase.

3.3 Mini-prep protocol

The constructs were transformed in DH5alpha cells and the next day one colony of each, was inoculated into 3mL preculture (Lamitina Lab Protocol 2007) and grown for 18hours at 37^{o}C. The cells were harvested at 6000x g for 30sec. The pellet was resuspended in 250 µL suspension buffer/RNAse. 250 µL of lysis buffer was also added in the tube. After the room temperature incubation for 5min, 350 µL of chilled binding buffer was added. The sample was centrifuged for 10min at max speed. The supernatant was spun down in a high pure filter tube at max speed for 30-60sec. The following steps were for eluting the plasmid DNA from the filter (High Pure Plasmid Isolation Kit-Life science, Roche). The expression constructs were sequenced and transformed in T1 phage resistant cells BL21 (DE3) *E.coli* and grown in LB Broth Miller Luria Bertoni 1x (Amresco) under the selection of 30µg/mL Kanamycin (Sigma).

3.4 Test expression protocol

A test expression protocol was carried out to confirm that the construct was expressed. For this reason, the bacteria were transferred to grow in 50-100mL culture and they were incubated for 2h30min or until the OD600 was equal to 0,4 at 140rpm in a 37^{o}C shaker. The protein expression induced by 0,33mM IPTG at 16^{o}C, at 140 rpm. Under these conditions, bacteria were grown O/N and the next morning were harvested by centrifugation (high speed J-26 XP centrifuge, Beckman Coulter). The pellets were resupended in 1mL each of B-PER lysis buffer (Thermo Scientific) and the cells lysed on ice for 30-45 min. The lysates were spun down by centrifugation (Eppendorf centrifuge 5424) at max speed for 20min at 4^{o}C. The SN was transferred into fresh tube and the pellet was resuspended in 1mL 8M Urea. From both SN and pellet, an aliquot was taken for SDS-PAGE gel analysis. The next step was the addition of 50µL $NiCl_2$ charged chelating Sepharose slurry in the SN and the rotation of the samples for 15-30min in the cold room. After tumbling, samples were centrifuged (Eppendorf centrifuge 5424) at max speed. The pellet was resuspended in 1mL 20mM Hepes 300mM NaCl 1mM TCEP pH 7.4 lysis buffer and the sample was spun down at max speed (Eppendorf centrifuge 5424). This step was repeated once more and an aliquot of the SN (=elution) was taken for an SDS-PAGE analysis.

3.5 Large scale protein expression

After the transformation into BL21, the bacteria were transferred to grow in 1L culture (100mL LB 10x, 900mL water and 1mL Kan) and they were incubated/grown for ~2.5h or alternatively, until the OD600 value was ~0.4, at 140rpm in a 37^{o}C shaker. Protein expression was induced by using 0,33mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) (MP Biomedicals) in a precooled shaker (at 16^{o}C, at 140rpm) and when the OD600 value was approx. 0.75 (Manuscript 2012). Under these conditions the bacteria were grown O/N and the next morning, were harvested by centrifugation (RC12BP centrifuge, Thermo Scientific) at 4000xg for 15 min at 4^{o}C. The pellet was resuspended in the lysis buffer 20mM Hepes 300mM NaCl 1mM TCEP pH 7.4 and was frozen down at -20^{o}C. By freezing down the pellet, the lysis of the cells is facilitated and the protein yield increases.

3.6 Cell lysis and His-tagged protein purification

Before the cell lysis by sonication, DNase I (Roche) was added into the resuspended pellet in order to remove unwanted DNA from cell lysates and thus, improve the protein extraction efficiency. The cells were lysed by using sonication (Q-sonica LA Biosyst) at 70% amplitude and the process time was for 3min (pulse on: 10s, pulse off: 30s). The lysate was spun down by centrifugation (high speed J-26 XP centrifuge, Beckman Coulter) at 20,000 rpm for 30min to separate the supernatant (soluble proteins) from the insoluble part-pellet (membranes, unbroken bacteria). Imidazole (20mM) was added into the SN for preventing the unspecific protein binding. As it was described before, the protein constructs are His tag or His-SUMO tag. Expressed His tagged proteins can be purified easily because the string of histidine residues binds to several types of immobilized metal ions, including nickel, cobalt and copper under specific buffer conditions. Both imidazole and His can bind to the affinity medium. Although, since His-tag tends to bind more strongly to the affinity medium, a higher concentration of eluent (Imidazole in this case) can be used during the washing step before elution. This can aid the removal of

contaminants which may otherwise be co-purified with a His tag fusion protein. Chelating Sepharose, when charged with Ni^{2+} ions, selectively binds proteins of complex-forming amino acid residues, in particular Histidine, are exposed on the protein surface. His tag fusion proteins can be easily bound and then eluted with buffers containing Imidazole. For this reason, the following step involved head-over rotation of the SN with charged (with $NiCl_2$) and prewashed (with MQ and lysis buffer) chelating Sepharose beads (Fast Flow) for 1h. The recommended amount of beads is proportional to the protein yield, to be more detailed 1ml of beads is required when the protein yield is approximately 60 mg. The clarified sample/SN was incubated with the affinity support to allow the target molecule in the sample to bind to the immobilized ligand. After 1 h, the sample was poured into a 20mL manual gravity column (GE Healthcare). The flow through was checked onto an SDS-PAGE gel for unbound sample. Afterwards, the beads were washed to get rid of the non-bound sample components by using ~50mL of the lysis buffer (wash 1) and 10mL lysis buffer with 1 M NaCl (wash 2), to remove ionically bound proteins/DNA. The beads were also washed with 10mL lysis buffer containing 40 mM Imidazole (wash 3) and 10mL lysis buffer with 80 mM Imidazole (wash 4). The target protein was eluted by lysis buffer with 400mM Imidazole (elution 1). For better protein elution, the beads incubated with 5mL elution buffer for 5 min. The last step of His-tag protein purification is the passage of another 5mL elution buffer to conducive the complete protein elution (elution 2). Ultimately, all the sample fractions (pellet, sn, ft, w3, w4, e1, e2) were checked on an SDS-PAGE gel. It is remarkable that a single pass of a sample through an affinity column can achieve greater than 1000-fold purification of a specific protein so that only a single band is detected after SDS-PAGE gel analysis (GE_Healthcare 2009). Proteins are globular in secondary and tertiary structure due to disulfide bonds, hydrophobic interactions, and hydrophilic interactions with their aqueous environment. Therefore, Sodium Dodecyl Sulfate (SDS) is a detergent possessing both an hydrophobic end (dodecyl group) and an hydrophilic end (sulfate group). The tertiary structure of most proteins often relies upon hydrophobic interactions at the core of the protein. The hydrophobic end of SDS breaks these interactions through interactions with the hydrophobic side chains of the amino acids. Similarly, a sulfate group can disrupt hydrogen bonding in secondary protein structure. Disulfide bonds can be cut by using reducing agents such as beta-mercaptoethanol (BME) or dithiothreitol (DTT). In addition to denaturing the protein, SDS also possesses an additional purpose. Because each protein is coated with SDS molecules and the sulfate group has a negative charge, SDS also serves to give each protein molecule a net negative charge. This means that when an electrical field is applied to the gel in buffer each protein will move toward the positive electrode. This allows the acrylamide to separate the proteins based on size. As the SDS-coated protein molecules move through the gel matrix smaller molecular weight proteins are able to navigate through the pores in the matrix more quickly than larger ones. As a result, the proteins in the sample are separated by size and relative molecular weight (Bay Area Biotechnology Education- BABEC booklet). After the SDS-PAGE analysis, the gel was stained by Coomassie Blue stain. The absorbance of the fractions w3, w4, e1, e2 was measured on nanodrop, at wavelength 280 nm. The optical absorbance of protein is measured at a wavelength of 280 nm. At this wavelength, the absorbance of protein is mainly due to amino acids tryptophan, tyrosine and cysteine (aromatic amino acids) with their molar absorption coefficients decreasing in that order. Although, it is obvious that the molar absorption coefficient of the protein itself at 280 nm will depend upon the relative concentrations of these three amino acids, by measuring the 280 nm absorbance, the amount of the protein can be determined. Therefore, by this way the amount of protease required to cleave the

tag can be simply calculated. It is also notable to be referred that Imidazole absorbs at this wavelength. As follows, the blind solutions/blank should include Imidazole. As a result/thus, the measurements will not be increased wrongly. The next step was the O/N cleavage of the tag using 3C protease or SENP2 protease. The protease used depends on the tag of the construct, more specifically on the recognition site that the construct possesses for the protease. The 3C protease was used to cleave the His tags and the SENP2 was specific for cleaving the His SUMO tags. 3C protease is a recombinant cysteine protease from the human rhinovirus 3C (HRV) expressed in and purified from E.coli. This protease retains high activity at 4^{o}C and it is compatible with various buffers. 3C protease recognizes Leu-Glu-Val-Leu-Phe-Gln↓Gly-Pro or LEVLFQ↓GP sequence and cleaves between the Gln and Gly, as indicated by the arrow. SENP 2 is a highly active and robust recombinant protease that cleaves the SUMO from the recombinant fusion proteins. SENP2 recognizes the tertiary structure of SUMO, consequently SENP2 will not cleave within the fused protein of interest (BPS Bioscience). The O/N cleavage of the tag was carrying out in dialysis membrane with mwco: 6-8000 (Spectra/Por molecular porous membrane tubing) against 20mM Hepes 100mM NaCl and 2 mM TCEP pH 7.4 in order to eliminate, in the meantime, the amount of Imidazole.

3.7 Reverse purification, size exclusion chromatography and ion exchange chromatography

The day after, a reverse purification over Ni-beads completed (GE_Healthcare 2009). The manual gravity column was washed with MQ water and dialysis buffer for equilibration. The following step was the head-over rotation of the sample with the appropriate amount of beads, for 30 min. Afterwards, the sample was passed through the column. In this case, the His-tag and a part of the protease recognition site supposed to bind onto the Nickel charged beads. As far as the protein concerned, it is passed through the column (Flow through). The reverse purification was progressed by an elution step. The elution buffer was 5mL GF (H20N100T2 pH 7.4) buffer with 400mM Imidazole, pH 7.4. In order to check the reverse purification success, an SDS-PAGE analysis was performed. The sample was concentrated to 2mL (Concentrator: GE Healthcare, 5000 MWCO) and thereupon it was injected to the FPLC ÄKTA system (software: Unicorn 5.31) for a gel filtration run on the Superdex75 column (GE Healthcare). The Superdex 75 column (prep grade) is recommended for the purification of recombinant proteins with molecular weight between 3000-70000Da. The Gel filtration buffer was 20mM Hepes 100 mM NaCl 2 mM TCEP pH 7.4. The Superdex 75 column was firstly washed and equilibrated with the H20N100T2 (pH 7.4) buffer. The flow rate used was 1.4 mL/min (Superdex S75 16/60) or 0.5mL/min (Superdex 75 10/300)(Healthcare & Sciences n.d.). Peak fractions from the Gel Filtration column were pooled and concentrated (concentrator GE Healthcare, MWCO 5000) to ~1.5mL volume. The clean fractions were diluted with H20T5 (pH 7.4), with dilution ration 1:4 in order to decrease the salt concentration. The protein was concentrated up to a concentration of 25 mg/mL.

The His-SUMO tagged constructs

Despite the fact that the His-SUMO tags were cleaved by a specific protease and were removed by applying reverse purification, there were still traces of SUMO. For this reason, we purified the protein sample further with the use of ion exchange chromatography with a Resource Q column (GE

Healthcare), capacity of 1mL. The IEX chromatography was conducted manually by applying flow rate approximately 1mL/min. The column washed with 10 mL MQ water, 5 mL buffer: 20 mM Hepes 1 mM TCEP pH 7.4, 5mL buffer: 20 mM Hepes 1 M NaCl 1 mM TCEP pH 7.4 and again 5 mL of 20 mM Hepes 1 mM TCEP pH 7.4 buffer. The flow through was collected. As in the most other modes of chromatography, a protein sample is injected onto the column under conditions where it will be strongly retained. A gradient of linearly increasing salt concentration is then applied to elute the sample components from the column. Consequently, the following steps were developed to elute the protein by using low salt concentrations and to remove the SUMO at a higher salt concentration. By increasing the salt concentration (generally by using a linear salt gradient) the molecules with the weakest ionic interactions start to elute firstly, from the column (in this case the protein). Although, SUMO requires a significant higher concentration of NaCl in order to be removed (150 mM or 250 mM NaCl). Molecules that have a stronger ionic interaction require a higher salt concentration and elute later in the gradient (in this case SUMO). This happens due to the weak interaction between the solid support and the target-protein, whilst the strong binding of the SUMO on the solid support. The buffers with the increasing concentration of salt were

20 mM Hepes 25 mM NaCl 1 mM TCEP pH 7.4

20 mM Hepes 50 mM NaCl 1 mM TCEP pH 7.4

20 mM Hepes 150 mM NaCl 1 mM TCEP pH 7.4

20 mM Hepes 250 mM NaCl 1 mM TCEP pH 7.4

After passing all, the flow through and the elutions were analyzed on an SDS-PAGE gel to confirm the SUMO removal and to check the protein levels after this purification step. The column washed with 5mL of high salted buffer 1 M NaCl and 10 mL MQ. It was stored at 20% v/v ethanol, at 4^{o}C (GeHealthcare n.d.).

Accordingly to the SDS-PAGE gel, the elution fractions were concentrated up to a protein concentration of 25 mg/mL.

3.8 Crystallization

The aim of all these experimental procedures is to obtain pure protein. Both His tagged and His SUMO tagged constructs were conversed to pure protein, protein free of the tag and other contaminations/impurities coming from the *E.coli* proteome. The final step was to set up crystallization experiments. Crystallization of macromolecules uses a set of experimental techniques aimed at producing crystals suitable for structure determination. Crystallization methods have improved over the years through the introduction of standard screens and the use of robotics which allows the screening of a large number of crystallization conditions in a miniaturized format, reducing the amount of protein needed. This crystallogenesis experiment was carried out using the vapor-diffusion method at 4^{o}C and ambient temperature. Automated crystallization screen was performed with Mosquito (TTP Lab Tech) crystallization robot in a sitting-drop format (96 MRC_2well sitting drop crystallization plate-Molecular Dimensions) plate. The reservoirs were pipetted/dispensed semi-automatically by the CyBi-Selma robot. The volume of the reservoir was 50 μL. The crystallization droplets were composed of 0,1:0,1 μL and 0,2:0,1 μL protein: reservoir. Screening and crystallogenesis experiments can be easily performed by using the standard screening kits. In this study, several sets of the standard screens were tried. To be more specific, Classics I,II Suite (QIAGEN), ComPas Suite (QIAGEN), Index (Hampton), JCSG+ (QIAGEN), MIDAS (Molecular dimensions), MORPHEUS (Molecular dimensions), PACT Suite (QIAGEN), PGA-LM (Molecular Dimensions), pH Clear Suite (QIAGEN), Protein Complex Suite (QIAGEN) standard screens were applied.

3.9 OPTIM

An Optim analysis was carried out with a view to find the conditions/buffer that the lipase His tag 1-127 is more stable and more soluble. In order to determine the structural stability of a protein molecule, its ability to maintain native structure is tested while it is exposed to different conditions (pH, buffer, temperature, additives, salt concentration). The different conditions/buffers are described on the table below (Anon n.d.).

OPTIM BUFFERS

	Concentration mM	Ph	NaCl mM	Additive
Na Acetate	100	4.5	150	
Na Acetate	100	5.5	150	
MES	100	6.5	150	
Hepes	100	7.5	150	
Tris	100	8.5	150	
CHES	100	9.5	150	
MES	100	6.5	0	
Hepes	100	7.5	0	
Tris	100	8.5	0	
MES	100	6.5	500	
Hepes	100	7.5	500	
Tris	100	8.5	500	
Hepes	100	7.5	150	5mM MgCl2
Hepes	100	7.5	150	10%glycerol
Hepes	100	7.5	150	0.1% Tween

The table 2 describes the buffers that were used for the OPTIM analysis of the His-tag vibrio 1-127. The composition of the buffers varies as far as the chemical, the pH, the salt concentration and the presence of some additives.

The conditions screening was varied to different buffers with a range of pH 4.5-9.5. Another variable was the salt concentration (NaCl). The last factor was the presence or not of some additives. 2 different protein concentrations were applied, 1mg/mL and 0,5mg/mL.

3.10 Multi-Angle Laser Light Scattering

This is the preferred analytical technique for determining absolute molar masses. MALLS measurements work by calculating the amount of light scattered at each angle detected. This process overcomes the problems associated with low angle detectors and allows a reliable and accurate measure of the light scattered. MALLS could be characterized as an absolute detection method, when it is coupled with another chromatographic system. MALLS provides an absolute means for measuring the molar mass, size, and distribution of polymers of all sorts.

In this experiment, the injection volume of the protein was 100 µL and the protein concentration 1mg/mL. The protein was injected to FPLC ÄKTA system/MALLS (software: Unicorn 5.31). The column was S75 10/300 and the buffer H20N100T2 pH 7.4. Before injecting the pump was washed and equilibrated into H20N100T2 (pH 7.4) with a flow rate of 0.5 mL/min.

3.11 PLA2G16 Activity assay

The lipase activity assay utilizes lipid analogs (DOPC,DOPG) as a lipase substrate. When uncleaved, the substrate remains in a non-fluorescent, quenched state. Although, upon hydrolysis by the lipase a fluorescent product is produced and it can be measured in a fluorescence microplate reader (485/520 nm).

Reagents:

Red/Green BODIPY ® PC-A2 (Life Technologies)
DOPC-10 mM in $EtOH^{abs}$, stored at -80°C (Sigma Aldrich)
DOPG -10mM in $EtOH^{abs}$, stored at -80°C (Sigma Aldrich)
L-Glutathione reduced-150mM in Reaction buffer pH: 8.5 (Sigma Aldrich)

Preparation of reagents

Reaction buffer (pH 8.5)

- 50mM Tris-HCl pH 8.5
- 100mM NaCl

2x Substrate Mix
Firstly, BODIPY Red/Green Substrate was dissolved in DMSO (1mM stock solution), afterwards 30 µL of each, DOPC and DOPG were added into the substrate. The substrate mix was stirred constantly into 8mL reaction buffer.

Enzyme dilution buffer

The enzyme dilution buffer was consisted of reaction buffer (pH 8.5), glutathione with a final concentration of 5mM GSH.

For this assay, the final enzyme concentration was 15µM and the enzyme was diluted into Enzyme dilution buffer (described above). Each well had 50% 2x substrate mix and 50% of the diluted enzyme. As positive control, the FL lipase was used. The negative control was the FL lipase incubated with the inhibitor. The assay buffer with 2x substrate mix was used as a blank. The fluorescence was measured immediately on PHERAstar FS (BMG LABTECH).

4. Results and discussion

As it was referred before, the experimental goal was the determination of the structure of the vibrio lipase. Different features of the vectors allow different affinity purification strategies better induction mediated or tag mediated solubility. Therefore several constructs with different amino-acid sequence lengths and different tags were tested to accomplish these results. Firstly, all the protein samples were tested for the expression and for the solubility (3.4 part, experimental).

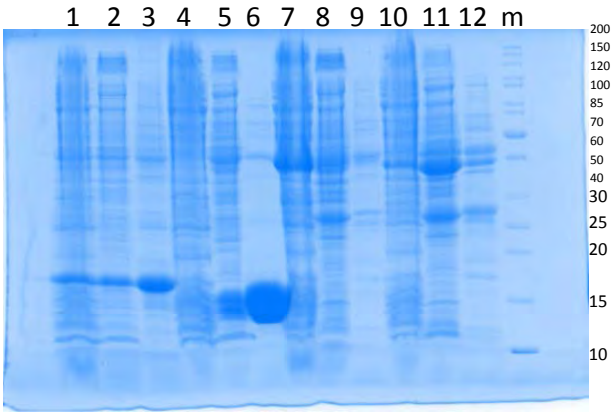
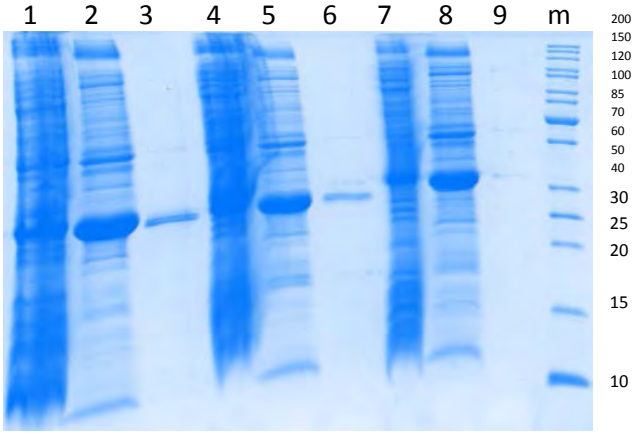


Fig. 1: Test expression protocol was performed for the Hrev vibrio cholera His-tag FL, His-tag 1-130, Hrev vibrio cholera GST-tag FL and GST-tag 1-130. Pellet, SN and elution were analyzed on the gel.

Test expression protocol
Hrev vibrio cholerae His-tag FL
1: pellet
2: SN
3: elution
Hrev vibrio cholerae His-tag 1-130
4: pellet
5: SN
6: elution
Hrev vibrio cholerae GST-tag FL
7: pellet
8: SN
9: elution
Hrev vibrio cholerae GST-tag 1-130
10: pellet
11: SN
12: elution
M: marker

After carrying out the test expression protocol (part 3.4 experimental), pellet, SN and elution fractions were analyzed on an SDS-PAGE gel to confirm the expression of the protein and to check the quantity and purity of the soluble protein. The gel revealed that the His-tag constructs were more soluble than the GST tag constructs. In specific, both His-tag and GST-tag FL expressed, however most of the GST tag FL protein remained at the pellet. Additionally, the fractions corresponding to this construct were not as pure as the fractions of the His tag FL protein. GST-tag 1-130 v. lipase fractions showed that this protein was relatively more soluble but not pure. Thus, based on this SDS-PAGE gel the GST-tag constructs were not analyzed further.



Test expression protocol
SUMO vibrio 1-118
1: pellet
2: SN
3: elution
SUMO vibrio 1-130
4: pellet
5: SN
6: elution
SUMO vibrio 1-159
7: pellet
8: SN
9: elution
M: marker

Fig. 2: Test expression protocol was performed for the SUMO vibrio 1-118, SUMO vibrio 1-130 and SUMO vibrio FL 1-159. Pellet, SN and elution were analyzed on an SDS-PAGE gel.

All these three constructs expressed and were relatively soluble. To describe more, both the SUMO vibrio 1-118 and SUMO vibrio 1-130 proteins were eluted (lanes 3,6 respectively), however the FL SUMO vibrio was not eluted (lane 9).

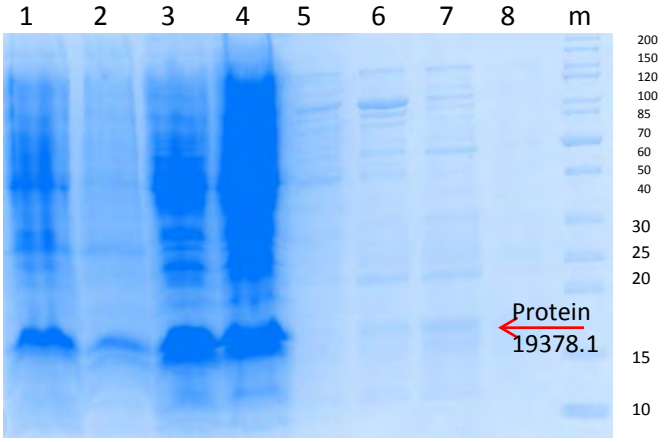
Finishing, the Hrev vibrio 1-127 His-tag was tested before by applying the same protocol. It had expressed and it was soluble.

Full length His-tag 1-161

strain	Medium	lysis buffer	dialysis buffer	vol	Prot conc	yield (mg/lt)	cryst. Buffer	cryst trays
BL21	LB (3L)	H20N300T1 pH 7.4						
Rosetta	LB (2L)	H20N300T1 pH 7.4	H20N100T2 pH 7.4	100µL	3,36mg/mL	0.168 mg/lt	H20N25T5 pH 7.4	Classics I 4C
BL21	A.I. (1L)	H20N500T1 pH 7.4						

The table 3 shows the preps for the His-tag 1-161 construct.

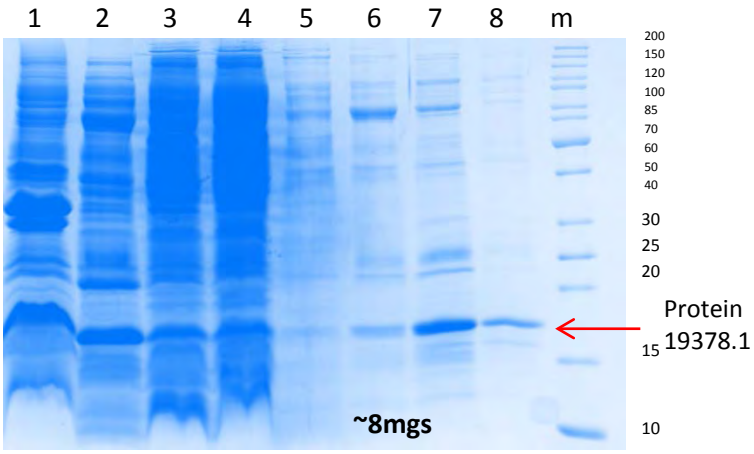
The initial approach was the purification of the FL vibrio lipase and potentially the crystallization of the protein. Hrev vibrio cholera His-tag FL and Hrev *vibrio cholera* GST-tag FL are both full length constructs (gel 1, samples 1,3). Comparing both these constructs for their solubility on an SDS-PAGE gel (1st and 3rd), we concluded that Hrev vibrio His-tag FL was more soluble than the Hrev vibrio GST-tag FL. Although, we cannot claim that Hrev vibrio His-tag FL is definitely soluble since after spinning down the sample shown on figure …, a part/portion of the protein remained to the pellet (see SDS-PAGE gel, lane 1). When the conditions reported at the experimental part were applied to the His-tag FL construct, the yield of the protein was significant reduced. As a result, a number of variables changed to optimize the expression conditions, increase the protein solubility and thus getting a higher protein yield (Pang et al. 2012).



His purification
His-tag vibrio FL (3L pellet)
1: pellet 1
2: pellet 2
3: SN
4: FT
5: wash 3
6: wash 4
7: elution 1
8: elution 2
M: marker

Fig. 3: His-tag vibrio FL expressed into BL21 cells and the cells were harvested into H20N300T1 pH 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography (see gel above). During the His-tag affinity chromatography, aliquots were taken after every step to check the purification success on an SDS-PAGE gel.

***Rosetta* strain**: *E.coli* BL21 strain was replaced by *E.coli* Rosetta strain. Rosetta strain is a BL21 derivative and is designed to enhance the expression of eukaryotic proteins that contain rare codons. These strains supply tRNAs for AGG, AGA, AUA, CUA, CCC, GGA codons on a compatible chloramphenicol-resistant plasmid named pRARE. The Rosetta strains provide the translation of some rare codons that would otherwise be limited if the strain used was *E.coli* (Biocompare). NanoDrop was used as an additional method for the approximate quantification of the protein in the fractions w3, w4, e1, e2. Furthermore, an SDS-PAGE gel was performed to check the purity of the fractions.



His-purification
His-tag vibrio FL (Rosetta, 2L pellet)
1: pellet
2: beads (after F.T.)
3: SN
4: FT
5: w3
6: w4
7: e1
8: e2
M: marker

Fig. 4: His-tag vibrio FL expressed into Rosetta strain, the cells were harvested into H20N300T1 pH 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography. In order to be confirmed that the purification was succeeded, an SDS-PAGE gel was carried out. The amount of the protein in the e1 fraction permits to continue the prep (lane 7, gel).

Gel filtration profile of vibrio His-tag FL (Rosetta strain)

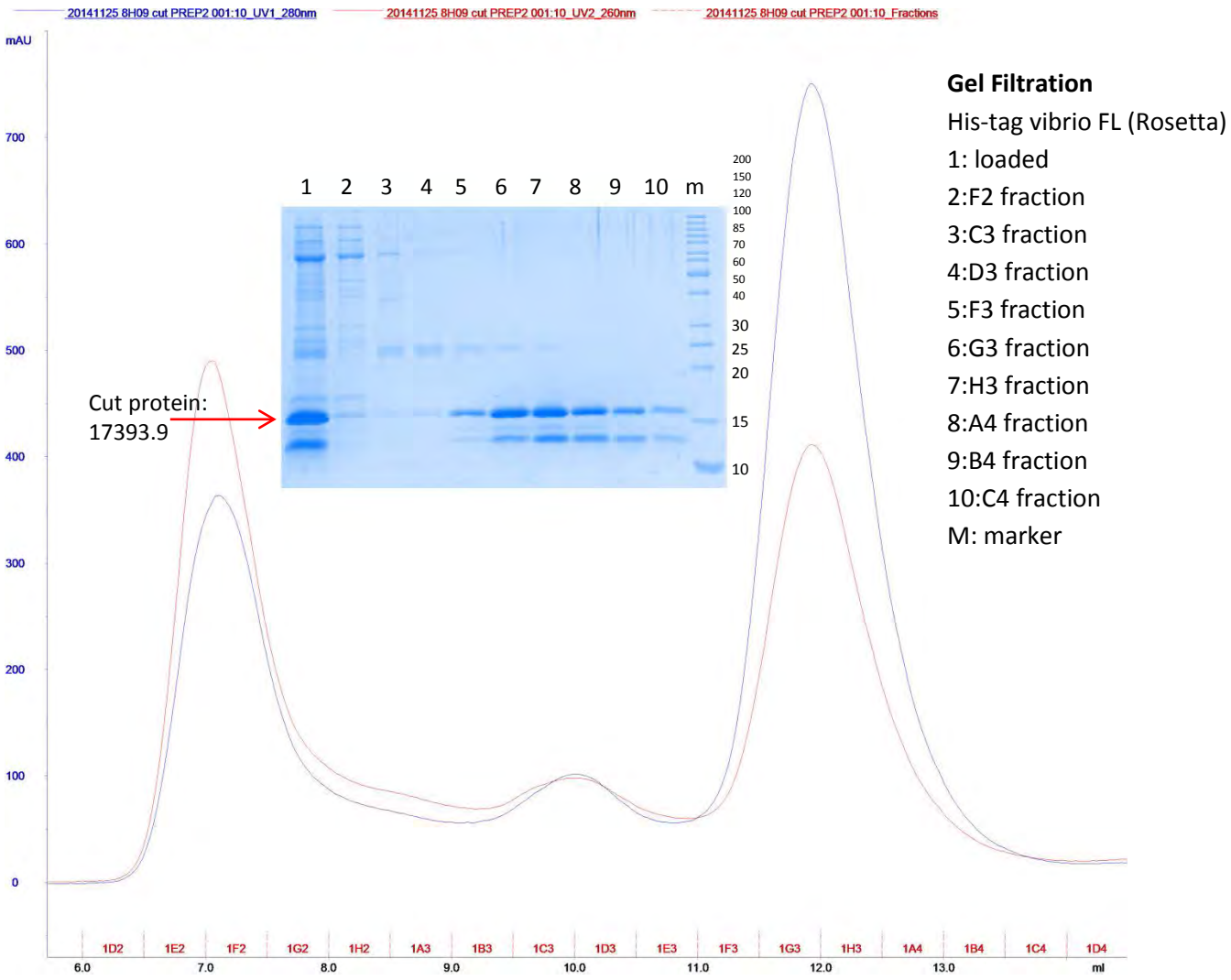
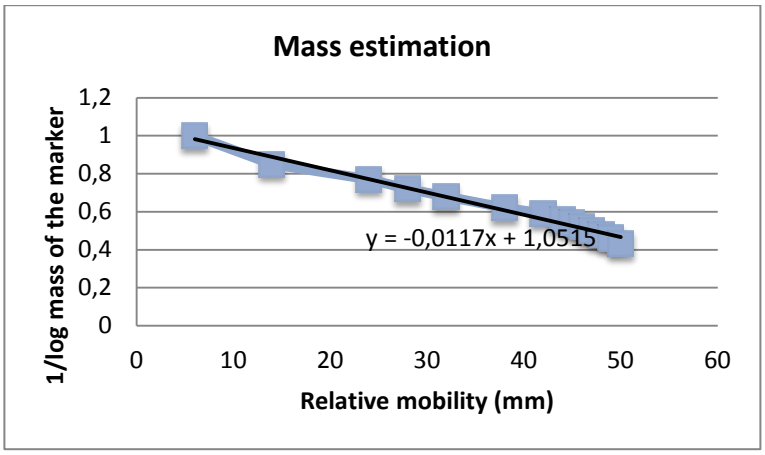


Fig. 5: The figure above is the Gel filtration profile of the vibrio His-tag FL, when it expressed into Rosetta strain. The GF buffer was H20N100T2 pH 7.4.

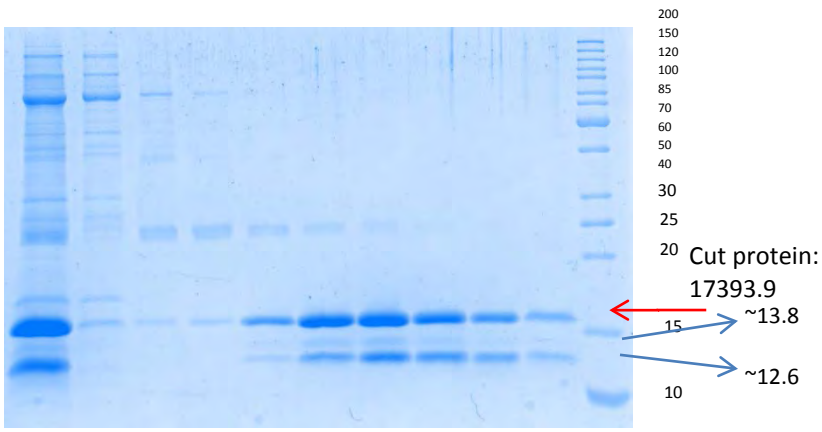
The gel (fig. 4) depicts that Rosetta strain gives significantly more amount of protein in comparison with the BL21. Thus, a crystallization tray (Classics I, 4^{o}C) was set up. The protein concentration was 3.36 mg/mL into buffer: H20N25T5 pH 7.4. This plate was checked through the Rock Imager robot and under the microscope for a period of 8 months. It is important to be referred that even during the 8th month, the majority of the drops was clear, implying that the protein concentration was not high enough. Therefore, for the future trials, a higher protein concentration will be used.

As it is observed on the gel, after GF, except from the expected band (protein), 2 other bands appeared. In order to determine the molecular weight of the proteins from these bands, a known relation from Biochemistry was used (Biochemistry, Stryer). (see also suppl. Table 1)



The electrophoretic mobility of many proteins in SDS-PAGE gels is inversely proportional to the logarithm of their mass.

Fig. 6: The diagram indicates the relation between the inverse logarithm of the marker's mass and the relative mobility of the marker expressed into mm. The determination of the mass will be facilitated by the equation coming of the diagram. The equation is illustrated on the chart area.

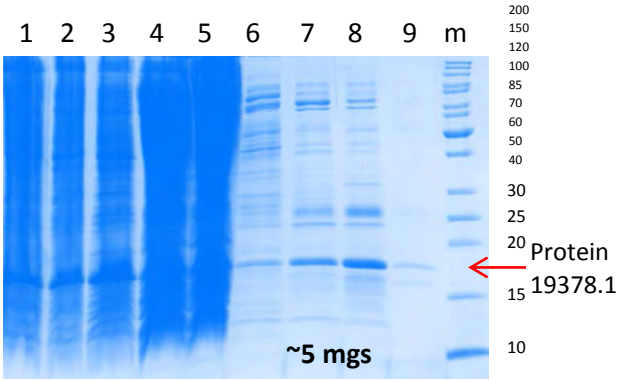


Gel Filtration
His-tag vibrio FL (Rosetta)
1: loaded
2:F2 fraction
3:C3 fraction
4:D3 fraction
5:F3 fraction
6:G3 fraction
7:H3 fraction
8:A4 fraction
9:B4 fraction
10:C4 fraction
M: marker

Fig. 7: The red arrow indicates the His-tag vibrio FL expressed into Rosetta strain. The other 2 unexpected bands were approximately estimated to be at 13.8 kDa and 12.6kDa, respectively. They are pointed on the SDS-PAGE gel by blue arrows.

Both bands were cut and analyzed by Mass-spectrometry (MS) for the precise determination of the mass of these proteins. No relevant hits were found.

Auto-Induction media: Auto-induction media improves the protein yield in comparison with the other media and induction methods. Considering/taking into consideration that A.I. media contains glucose and lactose, the increase/raise of the protein yield can be due to the abundant carbon source in the media. Ai media is formulated with quality biochemical to ensure consistent growth and protein expression using methodology developed by Studier at Brookhaven National Laboratory. ZYP-5052,Auto-induction media activates recombinant protein expression in *E.coli* through diauxic growth with the first growth phase producing a high density bacterial culture. During the 2nd phase, lac promoters are activated and result in induction of prolific transcription and translation of the cloned DNA. The Auto – Induction media is more efficient than conventional methods for high throughput routine protein expression, as it is indicated on the gel, because it eliminates the need for cell density and the manual addition of IPTG for induction (Manuscript 2012). The major part of the protein is at the pellet, despite the fact that the sample was spun down for approximately 2h.



His-purification

His-tag vibrio FL (Auto-induction, 1L pellet)

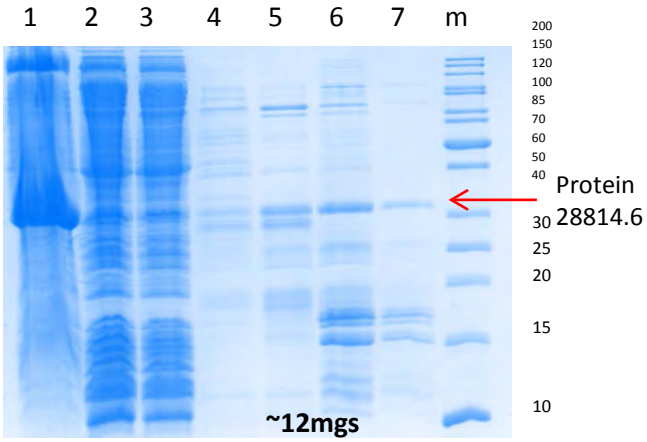
1: pellet 1
2: pellet 2
3: pellet 3
4: SN
5:F.T.
6:w3
7:w4
8:e1
9:e2
M: marker

Fig.8: His-tag vibrio FL expressed into BL21, the medium used was Auto-induction, the cells were harvested into H20N500T1 pH: 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography. In order to be confirmed that the purification was succeeded, an SDS-PAGE gel was carried out. From the gel, it is obvious that the major part of the protein is at the pellet, despite the 3 round centrifugation (total duration of spinning down: 2h).

Different tag + sequence length: Except from the full length His tag, the full length His-SUMO tag was also tried to be expressed. This construct is 3 amino acids shorter than the His-tag FL construct (full length His-SUMO tag 1-159).

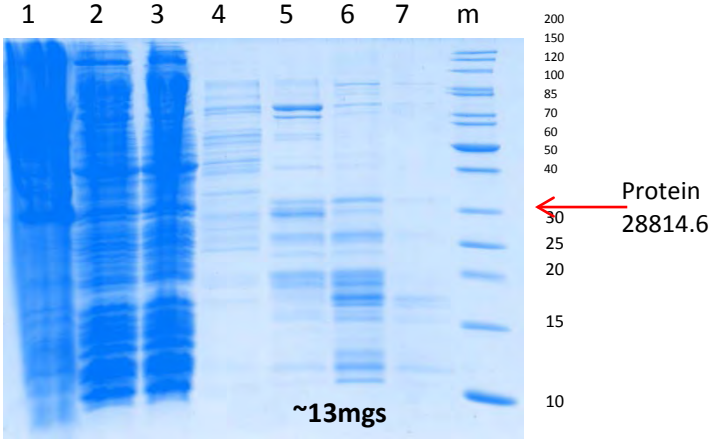
medium	vol	prot conc	Cryst buffer	yield mg/lt
L.B. (4L)	80μL	1.31 mg/mL	H20N30T1 pH 7.4	0.0262mg/lt

The table 4 shows the conditions of the prep for the His-tag 1-159 construct.



His-purification
His- SUMO tag vibrio FL (4L pellet)
1: pellet
2: SN
3: FT
4: w3
5: w4
6: e1
7: e2
M: marker

Fig. 9: His SUMO-tag vibrio FL expressed into BL21 cells and the cells were harvested into H20N300T1 pH: 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography. During the His-tag affinity chromatography, aliquots were taken after every step to check the purification success on an SDS-PAGE gel. On this gel, it is shown that the major part of the protein is at the pellet.



His-purification
His- SUMO tag vibrio FL (4L pellet)
1: pellet
2: SN
3: FT
4:w3
5:w4
6:e1
7:e2
M: marker

Fig. 10: A second prep was applied and the sample was spun down for ~2h. No significant differences were mentioned as it is shown on the gel.

The full length vibrio lipases with His-tag, His-SUMO tag, GST tag weren't soluble enough. Thus, shorter constructs were tried to be purified.

vibrio 1-130 His-tag

His-tag vibrio 1-130 expressed into BL21 cells and the cells were harvested into H20N300T1 pH 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography (see fig.11). During the His-tag affinity chromatography, aliquots were taken after every step to check the purification success on an SDS-PAGE gel.

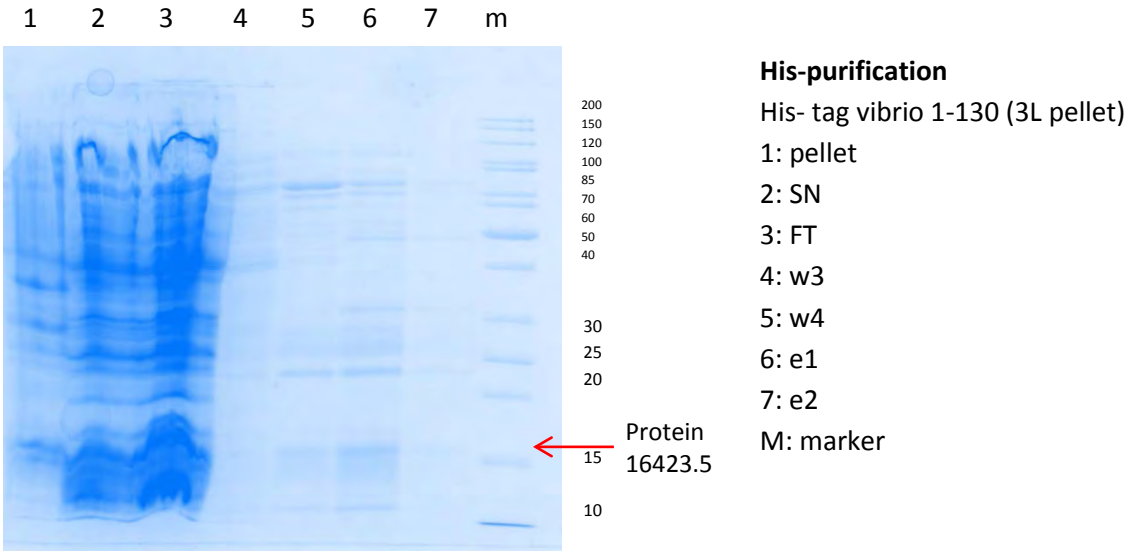


Fig.11: It is obvious on this gel, that the purification wasn't successful. Thus, the process didn't continue.

vibrio 1-130 His-SUMO tag

Medium	Vol	conc of prot	yield mg/lt	cryst buffer	cryst trays
L.B.(4L)	300 µL	24.8 mg/mL	1,86 mg/lt	H20N30T1 pH 7.4	ComPas RT,4C-Index RT,4C
L.B.(4L)	280 µL	24.07 mg/mL	1,68 mg/lt	H20N30T1 pH 7.4	pH clear 4C,RT-PACT R.T.-proComplex R.T.
L.B.(4L)	400 µL	25.34 mg/mL	2,53 mg/lt	T20N30T1 pH 8.5	index RT,4C-Compas RT,4C- proComplex RT,4C
L.B.(4L)	420 µL	~25.7 mg/mL	2,69 mg/lt	H20N30T1 pH 7.4	index RT,4C-MORPHEUS RT,4C- MIDAS RT,4C

The table 5 shows the details of the preps of the His- SUMO tag vibrio 1-130 construct.

The buffer initially used for the 1-130 His-SUMO tag was H20N300T1 pH 7.4. The protein yield was high. That's why Compas RT 4C, Index RT 4C, pH Clear RT 4C, PACT RT, ProComplex RT, MORPHEUS RT 4C MIDAS RT 4C were set up (mentioned on the table, above). No hits were observed besides the high protein yield. This leads us to change the pH of the buffer (from 7.4 to 8.5) and replace Hepes with Tris. The samples of the gels below show either the protein into Hepes either into Tris. When the protein was into the Tris buffer, Index RT 4C, Compas RT 4C, ProComplex RT 4C were set up.

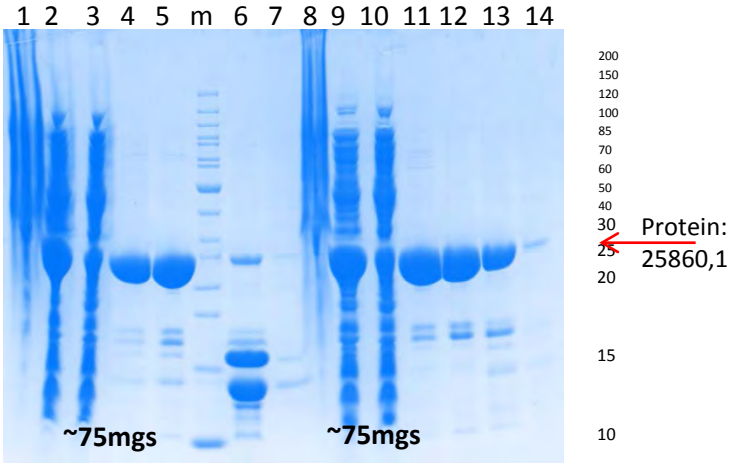


Fig.12: The His-SUMO tag vibrio 1-130 expressed into BL21 cells and the cells were harvested into H20N300T1 pH 7.4 and into T20N300T1 pH 8.5 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography (see gel above). During the His-tag affinity chromatography, aliquots were taken from every step to compare the purification yield with the different buffers on an SDS-PAGE gel. The purification is equally successful with both buffers.

His-purification
His-SUMO tag vibrio 1-130
(4L pellet into H20N300T1 pH 7.4 &T20N300T1 pH 8.5)
into H20N300T1 pH 7.4
1: pellet
2: SN
3: FT
4: w3
5: w4
M: marker
6:e1
7:e2
into T20N300T1 pH 8.5
8: pellet
9: SN
10: FT
11: w3
12: w4
13: e1
14: e2

Gel filtration profile of His-SUMO tag vibrio 1-130 into H20N100T2 pH 7.4

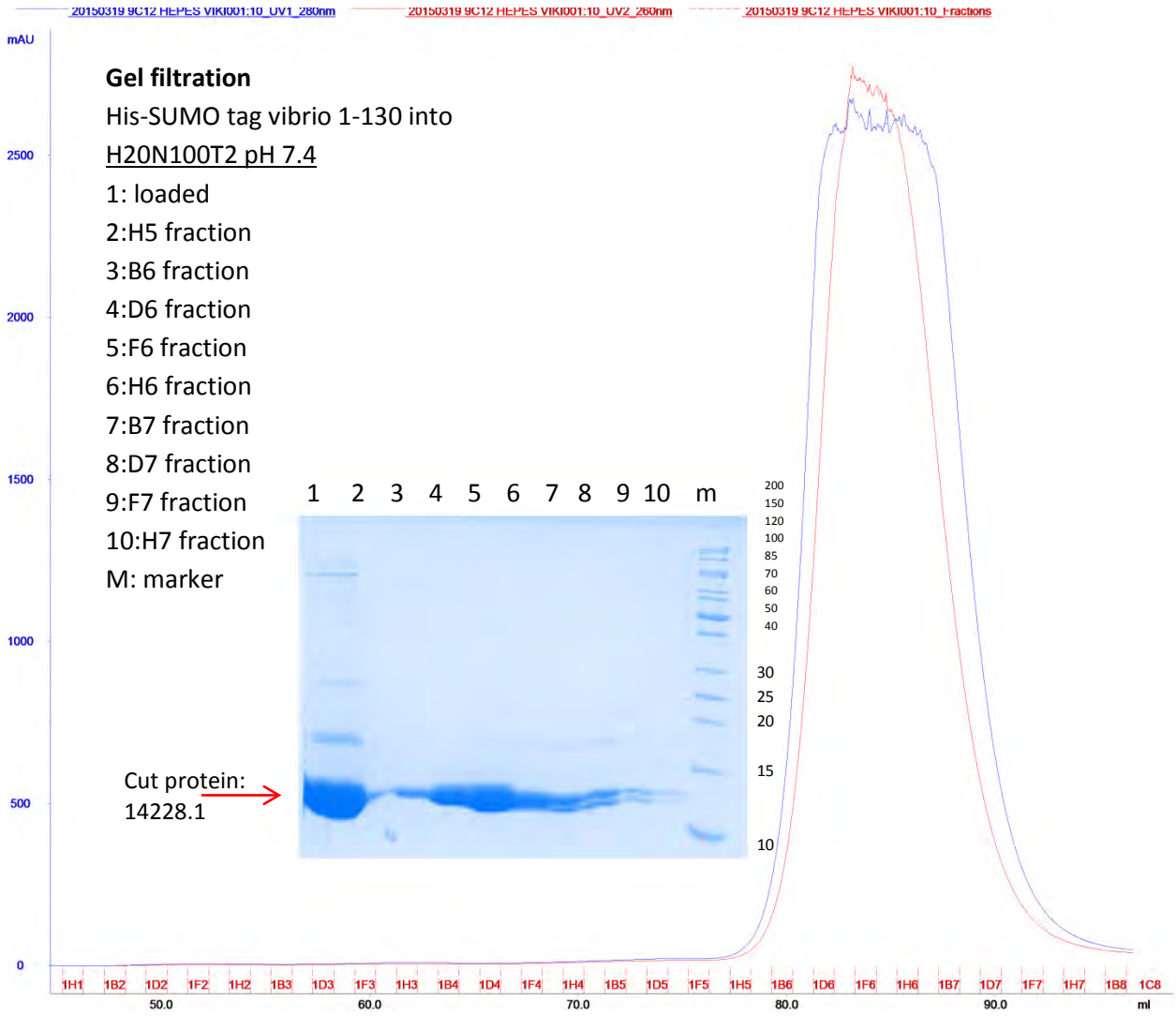


Fig. 13: The figure above is the Gel filtration profile of the His-SUMO tag vibrio 1-130. The GF buffer was H20N100T2 pH 7.4.

Gel filtration profile of His-SUMO tag vibrio 1-130 into T20N100T2 pH 8.5

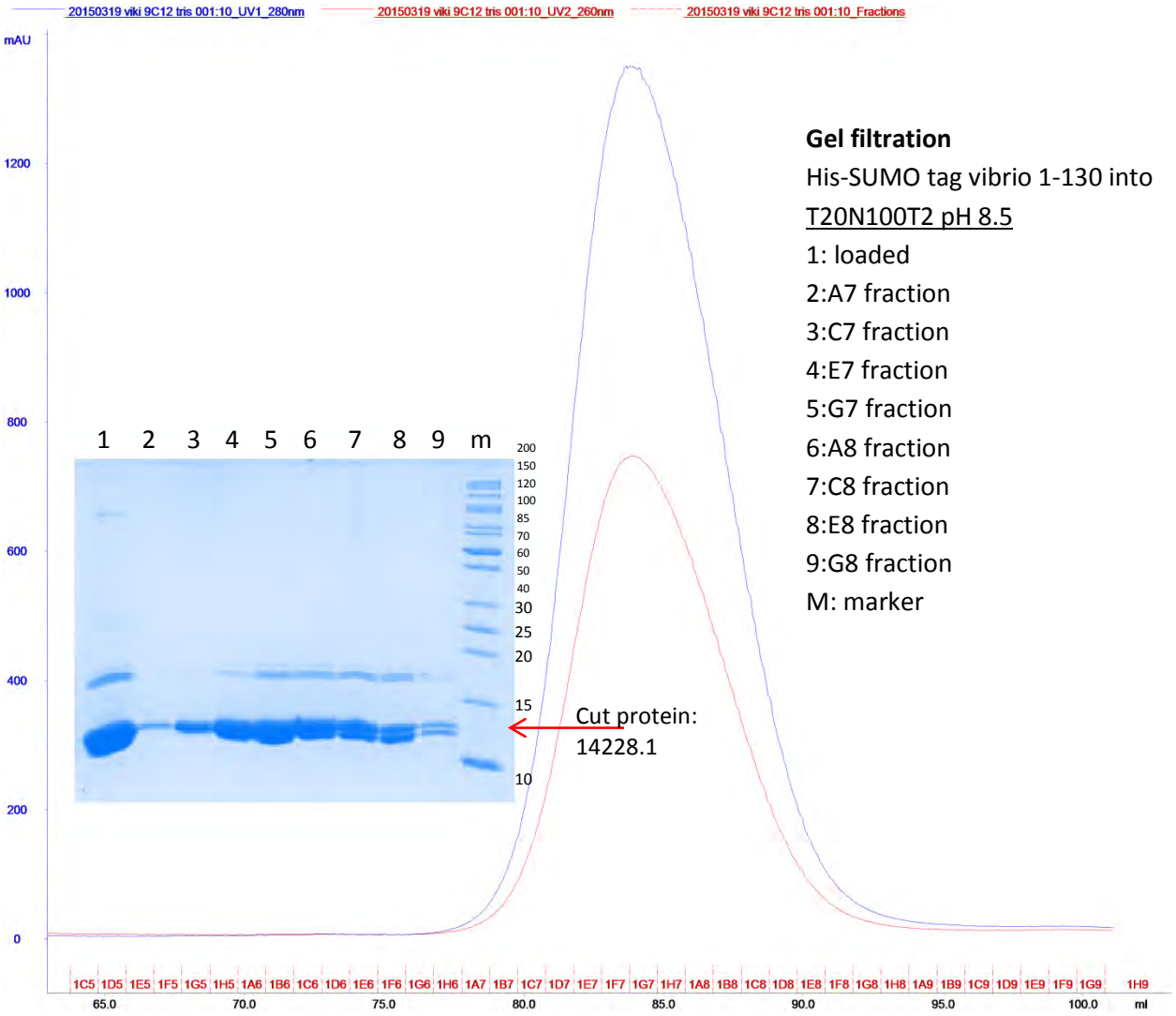


Fig. 14: The figure above is the Gel filtration profile of the His-SUMO tag vibrio 1-130. The GF buffer was T20N100T2 pH 8.5.

As it was observed from the Gel Filtration profile, SUMO traces were remained in the sample despite the reverse purification. For this reason, an ion exchange chromatography was performed in order to remove the remained SUMO.

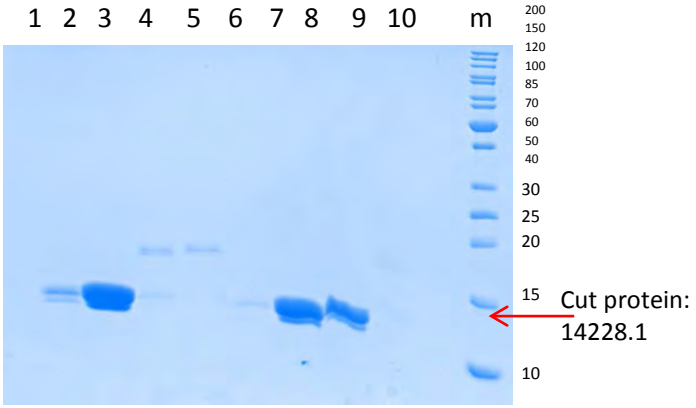


Fig.15: The F.T. and all the elutions were analyzed on an SDS-PAGE gel for comparing the purification success, after IEX, when the buffer used is different.

Ion exchange
His-SUMO tag vibrio 1-130 into
T20N100T2 pH 8.5
1: F.T. sample
2: elution T20N25T1 pH8.5
3: elution T20N50T1 pH8.5
4: elution T20N150T1 pH8.5
5: elution T20N250T1 pH8.5
6: F.T. sample
7: elution H20N25T1 pH7.4
8: elution H20N50T1 pH7.4
9: elution H20N150T1 pH7.4
10: elution H20N250T1 pH7.4
M: marker

Based on the SDS-PAGE gel, the elutions T20N25T1, T20N50T1 pH 8.5 and the elutions H20N25T1, H20N50T1 pH 7.4 were concentrated. When the protein concentration was ~25mg/mL (1,76mM), crystallization trays were set up. The final composition of the buffers was T20N~30T1 pH 8.5 and H20N~30T1 pH 7.4. Both buffers were tested in order to figure out if a higher pH value will play an important role and give significant different results.

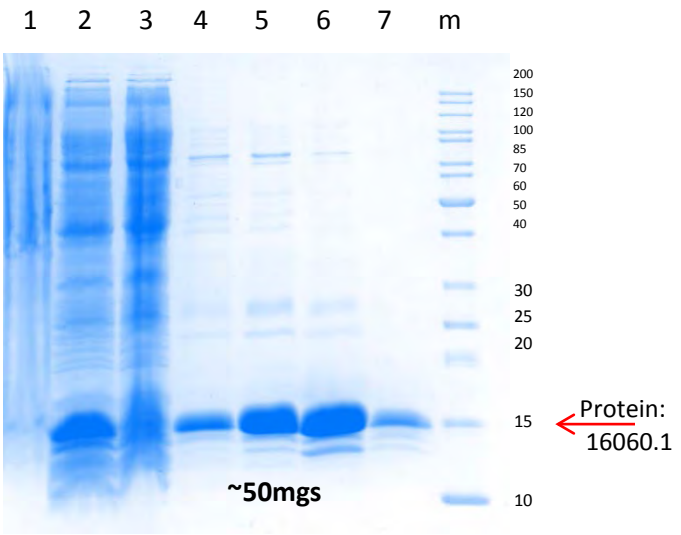
Index RT 4C, ComPas RT 4C, ProComplex RT 4C were tested. The protein concentration was 25.34mg/ml into T20N30T1 pH 8.5 buffer (as it is also written on the …table). Index RT 4C, MORPHEUS RT 4C, MIDAS RT 4C were also checked when the protein concentration was 25.7 mg/mL into H20N30T1 pH 7.4 buffer. Comparing the two different pH buffers, as far as the crystallization hits concerned, we could easily declare that both buffers didn't change the protein behavior significantly. Although, T20N30T1 pH 8.5 gave impulse for more drops with crystalline-like appearance.

vibrio 1-127 His-tag

medium	Vol	prot conc	yield mg/lt	cryst buffer	cryst trays
L.B. (2L)	600 µL	24.8 mg/mL	7,44 mg/lt	H20N25T5 pH 7.4	pH clear RT,4C-ComPas RT,4C-Index RT,4C-JCSG+ RT,4C
L.B. (4L)	500 µL	25.6 mg/mL	3,2 mg/lt	H20N25T5 pH 7.4	Classics RT,4C-Morpheus RT,4C-PACT RT,4C-Procomp RT,4C
L.B. (2L)	600 µL	25.35 mg/mL	7,605 mg/lt	H20N?T5 pH 7.4	MIDAS RT,4C-Classics 2 RT,4C-PGA-LM RT,4C-pH Clear 4C
L.B. (4L)	700 µL	24.95 mg/mL	4,37 mg/lt	T20N25T5 pH 8.5	ProComplex RT,4C-PACT RT,4C-ComPas RT,4C-Index RT,4C- Classics RT,4C-MIDAS RT,4C-pH Clear RT,4C
L.B. (4L)	300 µL	23.99 mg/mL	1,8 mg/lt	T20N25T5 pH 8.5	PGA-LM RT,4C-MORPHEUS RT,4C-Classics II RT,4C-JCSG+ RT,4C

The table 6 shows the details of the preps of the His- tag vibrio 1-127 construct.

The default conditions/conditions that have already described, were applied for the prep of Hrev vibrio 1-127 construct (see table 6). The purity of the protein and the protein yield are illustrated on the SDS-PAGE gel below.



His-purification
His-tag vibrio 1-127 (4L pellet)
1:pellet
2:SN
3:FT
4:w3
5:w4
6:e1
7:e2
M: marker

Fig.16: The His-tag vibrio 1-127 expressed into BL21 cells and the cells were harvested into H20N300T1 pH 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography. During the His-tag affinity chromatography, aliquots were taken from every step to check the purification success on an SDS-PAGE gel. The purification yield is significant high, as it is shown on the gel.

Gel filtration profile of vibrio His-tag vibrio 1-127

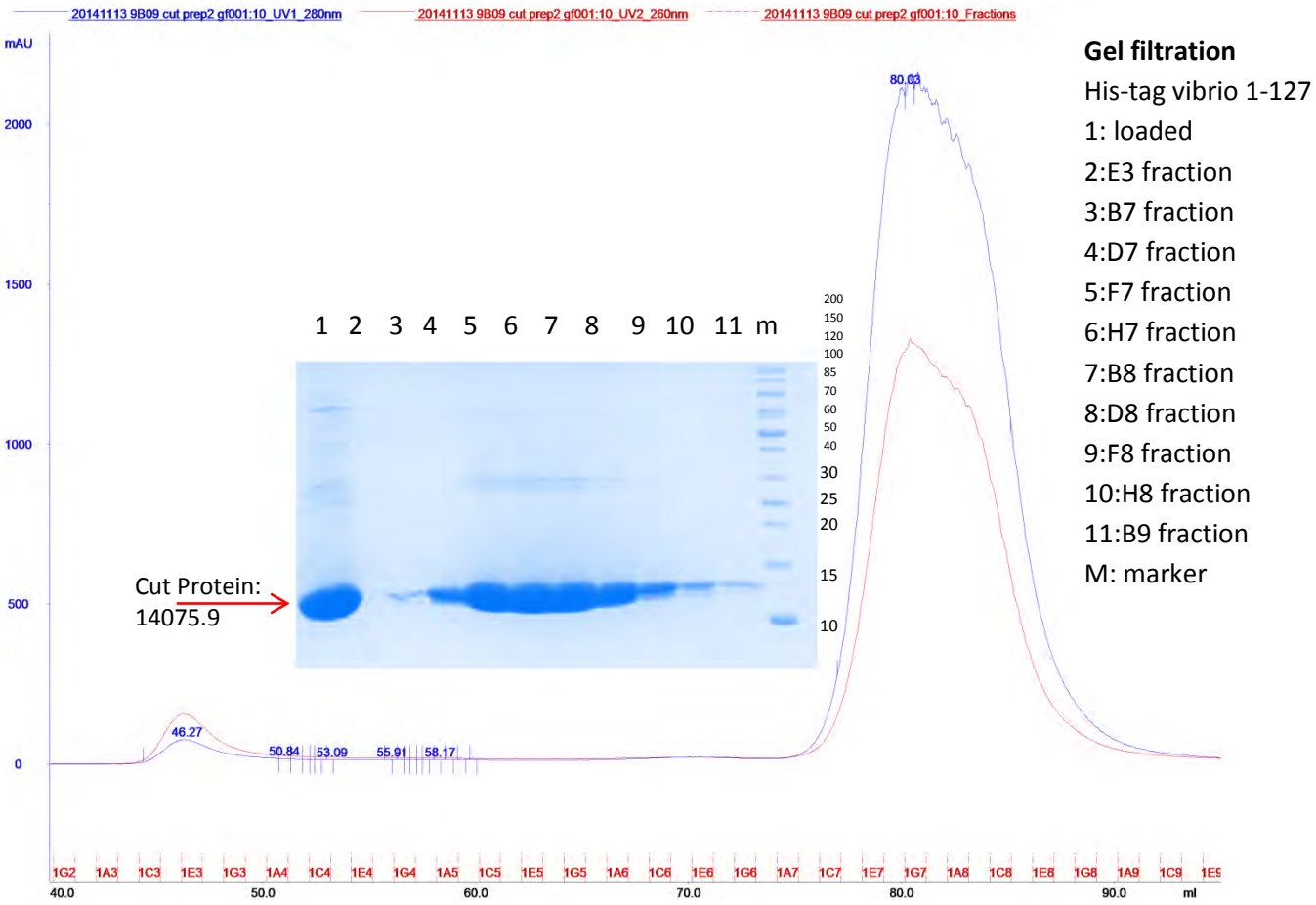


Fig.17: The Gel filtration profile of the His-tag vibrio 1-127 is illustrated above. The GF buffer was H20N100T2 pH 7.4.

After the purification of this protein, since it was efficient (high yield), several crystallizations trays were set up. To be more detailed, JCSG+, pH Clear Suite, Protein Complex Suite, Index, ComPas Suite, Classics I,II Suite, PACT Suite, MIDAS, MORPHEUS, PGA-LM trays were set up at both 4^{o} C and Room temperature. All the plates were checked through the Rock Imager monitoring system and under the optic microscope. No hits were found.

For this reason, an Optim analysis was carried out with a view to the conditions/buffer that the protein is more stable and more soluble. In order to determine the structural stability of a protein molecule, its ability to maintain native structure is tested while it is exposed to different conditions (pH, buffer, temperature, additives, salt concentration). The different conditions/buffers are described on the table 2 (experimental).

Optim diagrams

The first measurement involved the Fluo Intensity of the protein into the different buffers. Fluo intensity is a parameter which can reveal which buffers keep the protein in a soluble form/not precipitated. The buffers with the lowest pH values were excluded (see suppl. figure….) for the rest of the analysis because both buffers induce the precipitation of the protein, beginning at low temperature. Also, the buffer with the highest pH value wasn't considered further because it also provokes protein precipitation very early during the analysis (low temperature). Scatter 266 and Barycen Fluo were measured in order to find the buffer that maintains the protein in a soluble form. Scatter 266 and Barycen Fluo diagrams were constructed for comparing the different pH buffers, MES pH 6.5, Hepes pH 7.5 and Tris pH 8.5, which were not excluded after the Fluo Intensity measurement. The 2 diagrams below point out that Hepes pH: 7.5 and Tris pH 8.5 keep the protein at a similar situation. On the other hand, MES pH 6.5 is less compatible with this protein. The temperature at which the protein begins to aggregate is often, but not always, correlated with the temperature at which the protein begins to unfold and with the rate at which the proteins aggregate (typically a high Tagg is preferred). The magnitude of aggregation after onset gives an indication of the propensity of unfolded proteins to aggregate (typically lower levels of aggregation are preferred). Thus, when the protein is into MES pH 6.5 the Scatter 266 value is significant higher than the Scatter 266 values from the 2 other buffers. Additionally, the protein aggregation starts at a lower temperature into MES pH 6.5. The Barycen Fluo curves indicate a swift at the protein conformation. This swift could be due to aggregation or due to unfolding. At this point, it is remarkable to be referred that the aggregation level depends on the protein concentration. That's why 2 protein concentrations were analyzed by OPTIM. Therefore, in order to distinguish, if it is due to aggregation or due to unfolding the diagram 21 was plotted. Except from the comparison based on the different pH values of the buffers, several more diagrams concerning other parameters (additives and salt concentration) were constructed (see suppl. Figures 2,3,4,5).

FluoIntensity

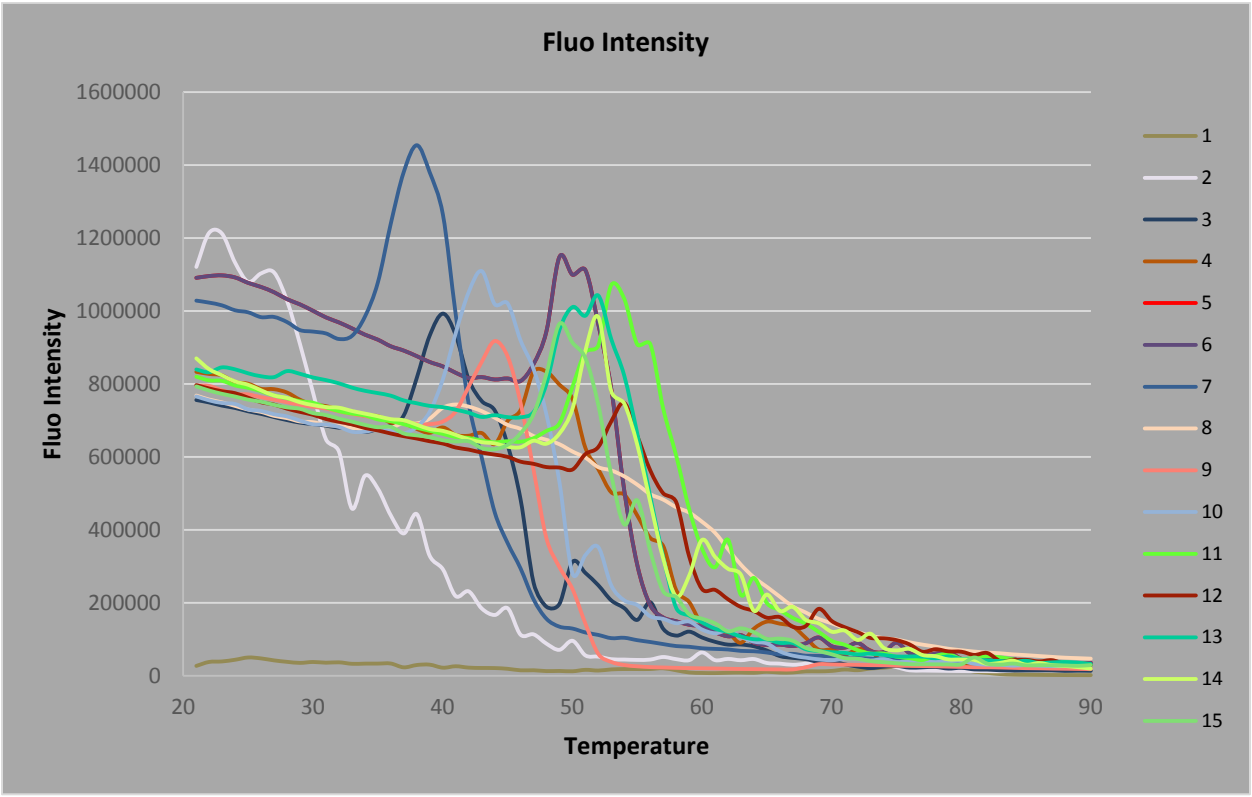


Fig.18: The chart depicts the FluoIntensity of the protein with concentration 1mg/mL at the different buffers. The colors correspond to the buffers which are written on the table below.

		Concentration mM	pH	NaCl mM	Additive
1	Na Acetate	100	4.5	150	
2	Na Acetate	100	5.5	150	
3	MES	100	6.5	150	
4	Hepes	100	7.5	150	
5	Tris	100	8.5	150	
6	CHES	100	9.5	150	
7	MES	100	6.5	0	
8	Hepes	100	7.5	0	
9	Tris	100	8.5	0	
10	MES	100	6.5	500	
11	Hepes	100	7.5	500	
12	Tris	100	8.5	500	
13	Hepes	100	7.5	150	5mM MgCl2
14	Hepes	100	7.5	150	10%glycerol
15	Hepes	100	7.5	150	0.1% Tween

The table 7 describes the buffers used at the Optim analysis. The buffers are numbered from 1 - 15 to be facilitated the labelling of the fig. 18.

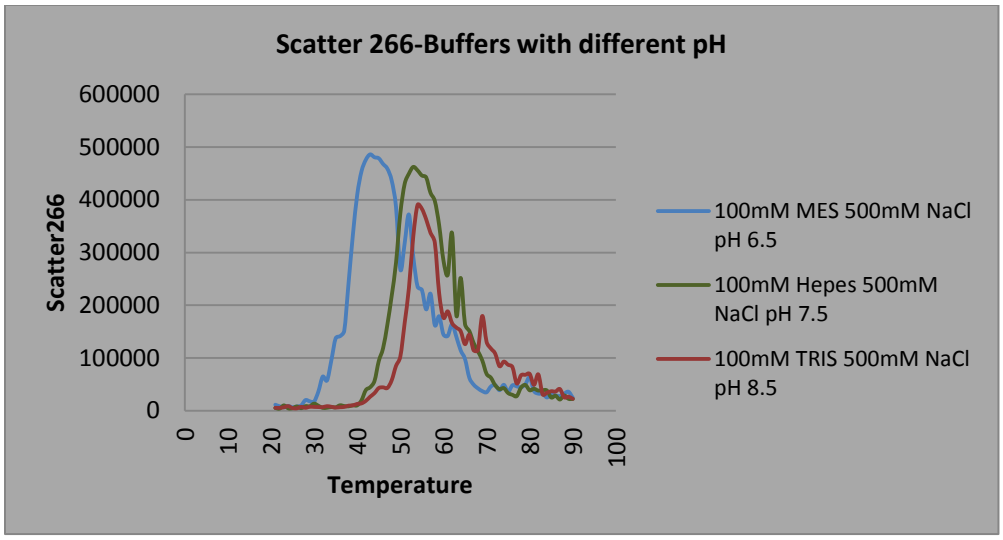


Fig.19: The chart depicts the Scatter 266 measurements of the protein (1mg/mL) into MES pH 6.5, Hepes pH 7.5 and Tris pH 8.5 when a range of temperatures was applied.

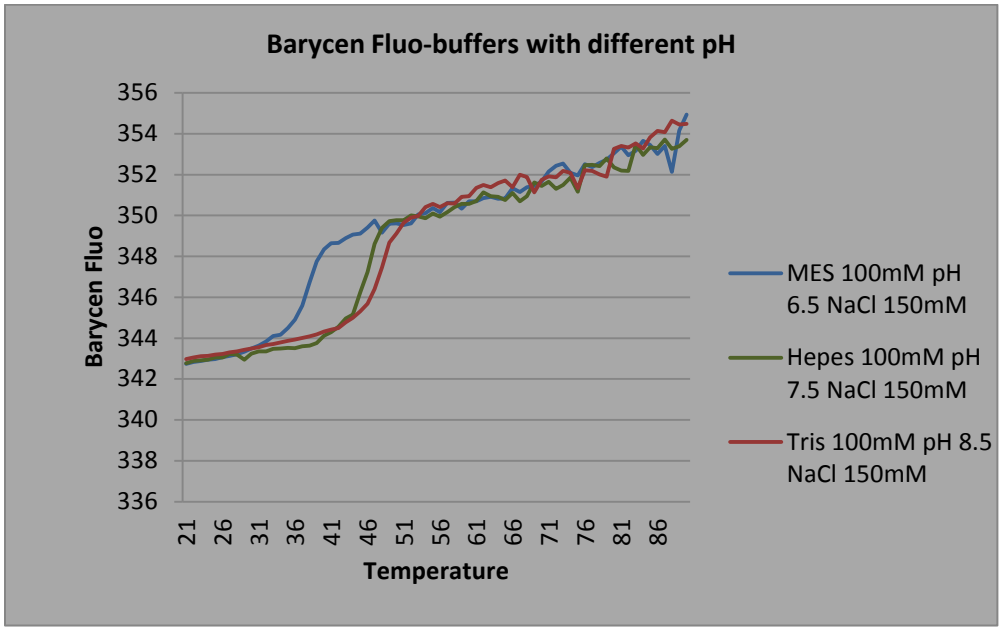


Fig. 20: The chart shows the Barycen Fluo measurements of the protein (1mg/mL) into MES pH 6.5, Hepes pH 7.5 and Tris pH 8.5 when a range of temperatures was applied.

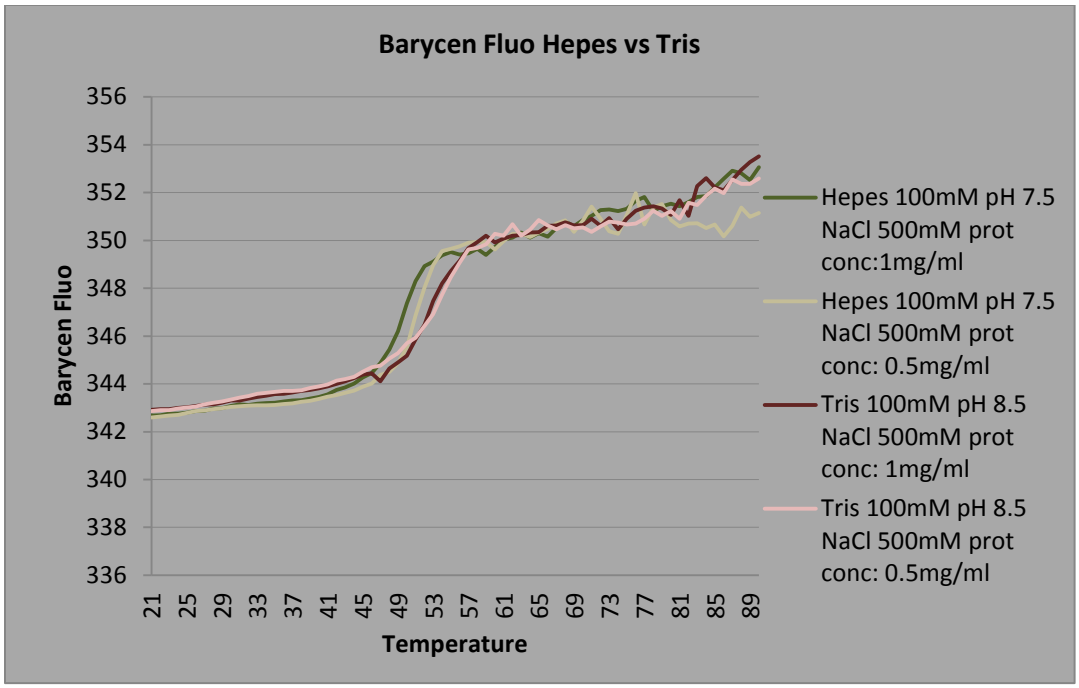


Fig.21: The chart shows the Barycen Fluo measurements of the protein (0.5 mg/mL and 1 mg/mL) into Hepes pH 7.5 and Tris pH 8.5 when a range of temperatures was applied.

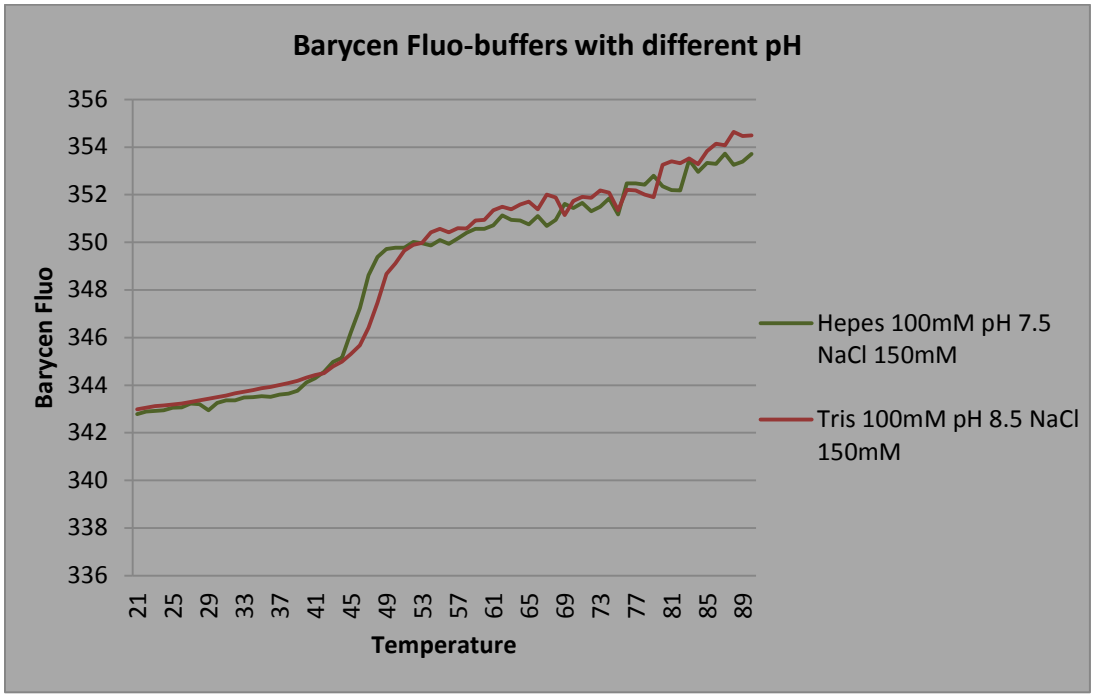


Fig.22: The chart shows the Barycen Fluo measurements of the protein (1mg/mL) into Hepes pH 7.5 and Tris pH 8.5 when a range of temperatures was applied.

A widely used measure of the thermal structure stability of a protein is the temperature at which it unfolds from the native, folded state to a denatured, unfolded state. For many proteins this unfolding process occurs over a narrow temperature range and the midpoint of the transition is often termed the melting temp or Tm. This value reflects the thermal structural stability of the protein and is related to the Gibbs free energy of the protein. Tm can be defined as the temperature at which there is an equal population of folded and unfolded proteins in solution (Anon n.d.).

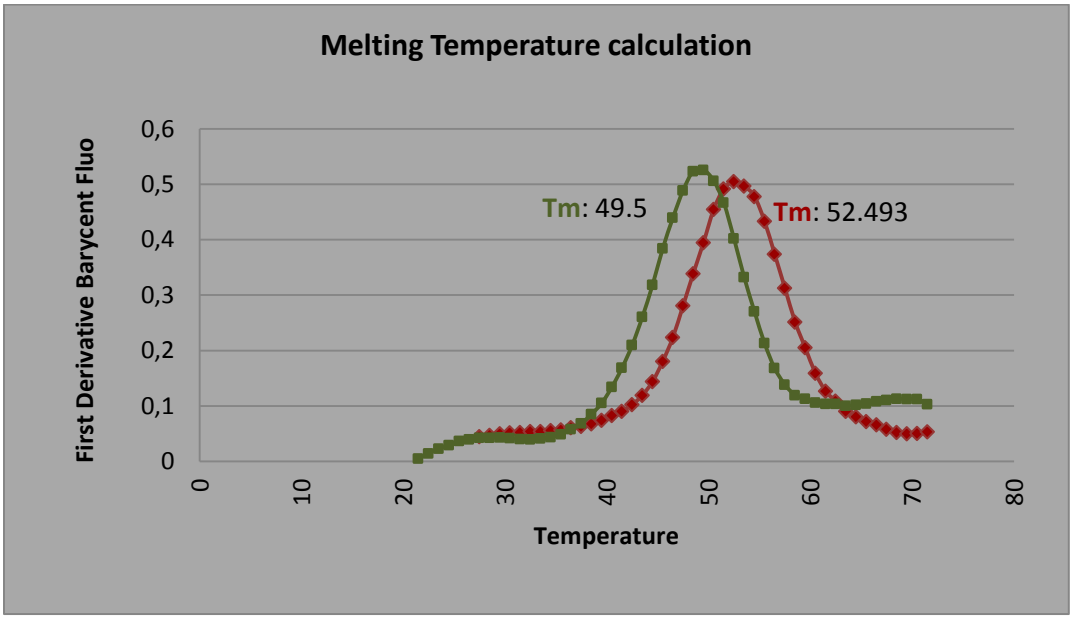
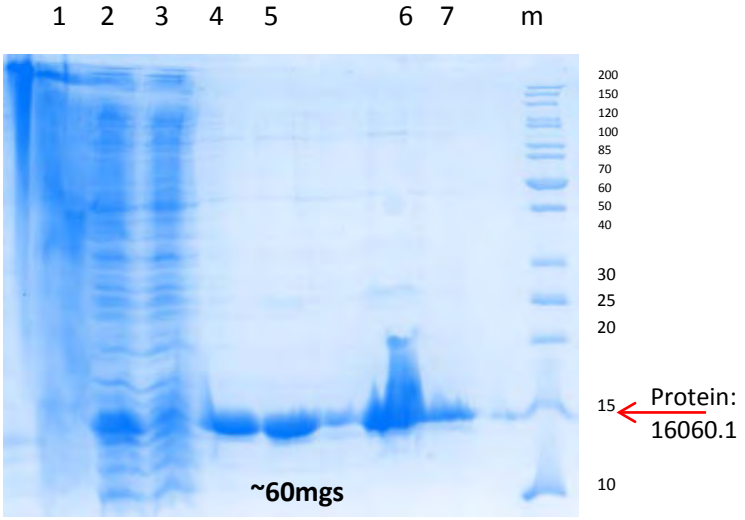


Fig. 23: This diagram depicts the Tm of the protein (1 mg/mL) when it is into Hepes pH 7.5 (green) and Tris pH 8.5 (red).

An additional prep of this protein was performed based on the OPTIM analysis. The lysis buffer H20N300T1 pH 7.4 was replaced with T20N300T1 pH 8.5. The tricky point of this protein purification was the Tris-temperature dependence.



His-purification
His-tag vibrio 1-127 (4L pellet) into T20N300T1 pH 8.5
1: pellet
2: SN
3: FT
4:w3
5:w4
6:e1
7:e2
M: marker

Fig.24: According to the Optim analysis, another His-tag purification was carried out, into T20N300T1 pH 8.5 lysis buffer. On the SDS-PAGE gel, the purification success is illustrated.

Gel filtration profile of vibrio His-tag vibrio 1-127 into T20N100T2 pH 8.5

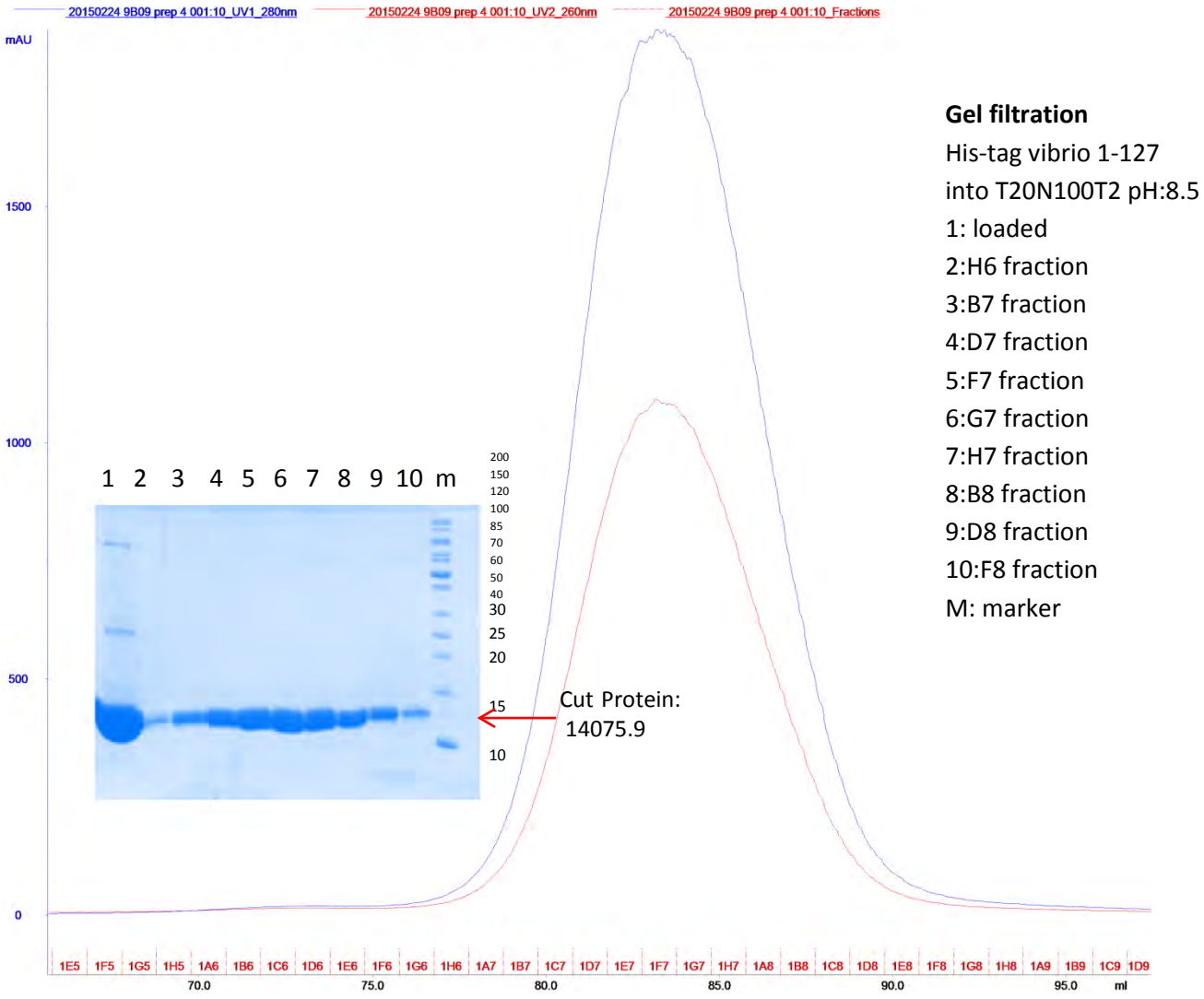


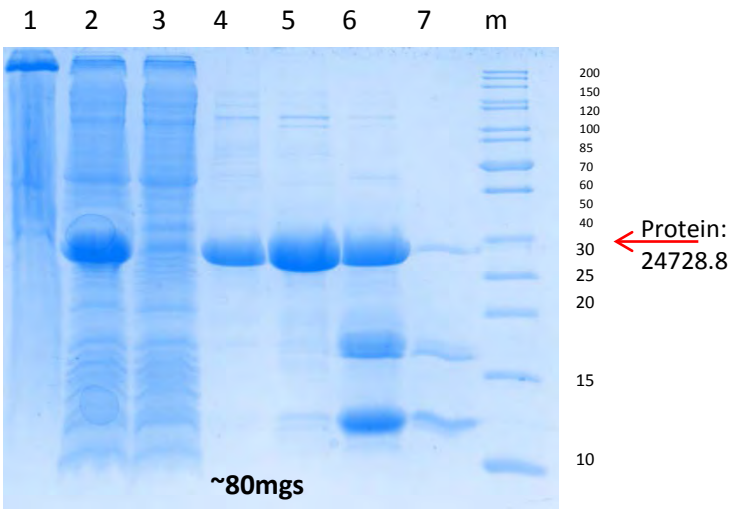
Fig. 25: The Gel filtration profile of the His-tag vibrio 1-127 is illustrated above. The GF buffer was T20N100T2 pH 8.5.

When this protein was purified, JCSG+, pH Clear Suite, Protein Complex Suite, Index, ComPas Suite, Classics I,II Suite, PACT Suite, MIDAS, MORPHEUS, PGA-LM trays were set up at both 4^{o} C and Room temperature. The protein concentration was 24.95 mg/mL and the protein buffer was T20N25T5 pH 8.5. All plates were checked through the rock imager screening and under the microscope. No hits were found despite the high yield of the protein purification.

vibrio 1-118 His-SUMO tag

Medium	vol	Prot conc	yield mg/lt	cryst buffer	cryst trays
L.B. (3L)	300 µL	18.23 mg/ mL	1,823 mg/lt	H20N30T1 pH 7.4	pH clear RT,4C-ComPas RT,4C-Index RT,4C
L.B. (3L)	220 µL	20.87 mg/ mL	1,53 mg/lt	H20N30T1 pH 7.4	Classics I RT, 4C-PACT RT,4C
L.B. (2L)	230 µL	23.97 mg/ mL	2,76 mg/lt	H20N30T1 pH 7.4	ProComplex RT,4C-PGA-LM RT-JCSG+ RT,4C
L.B. (4L)	200 µL	18.45 mg/ mL	0,92 mg/lt	H20N30T1 pH 7.4	MIDAS RT,4C-MORPHEUS RT,4C
L.B. (4L)	150 µL	18.2 mg/ mL	0,7 mg/lt	H20N30T1 pH 7.4	PGA-LM 4C- CLASSICS II 4C,RT

The table 8 shows the details of the preps of the His- SUMO tag vibrio 1-118 construct.



His-purification
His-tag vibrio 1-118 (2L pellet)
1: pellet
2: SN
3: FT
4:w3
5:w4
6:e1
7:e2
M: marker

Fig. 26: The His-tag vibrio 1-118 expressed into BL21 cells and the cells were harvested into H20N300T1 pH 7.4 lysis buffer. After the lysis of the cells, the protein was purified by His-tag affinity chromatography. During the His-tag affinity chromatography, aliquots were taken from every step to check the purification success on an SDS-PAGE gel. The purification yield is significant high, as it is shown on the gel.

Gel filtration profile of vibrio His-tag vibrio 1-118

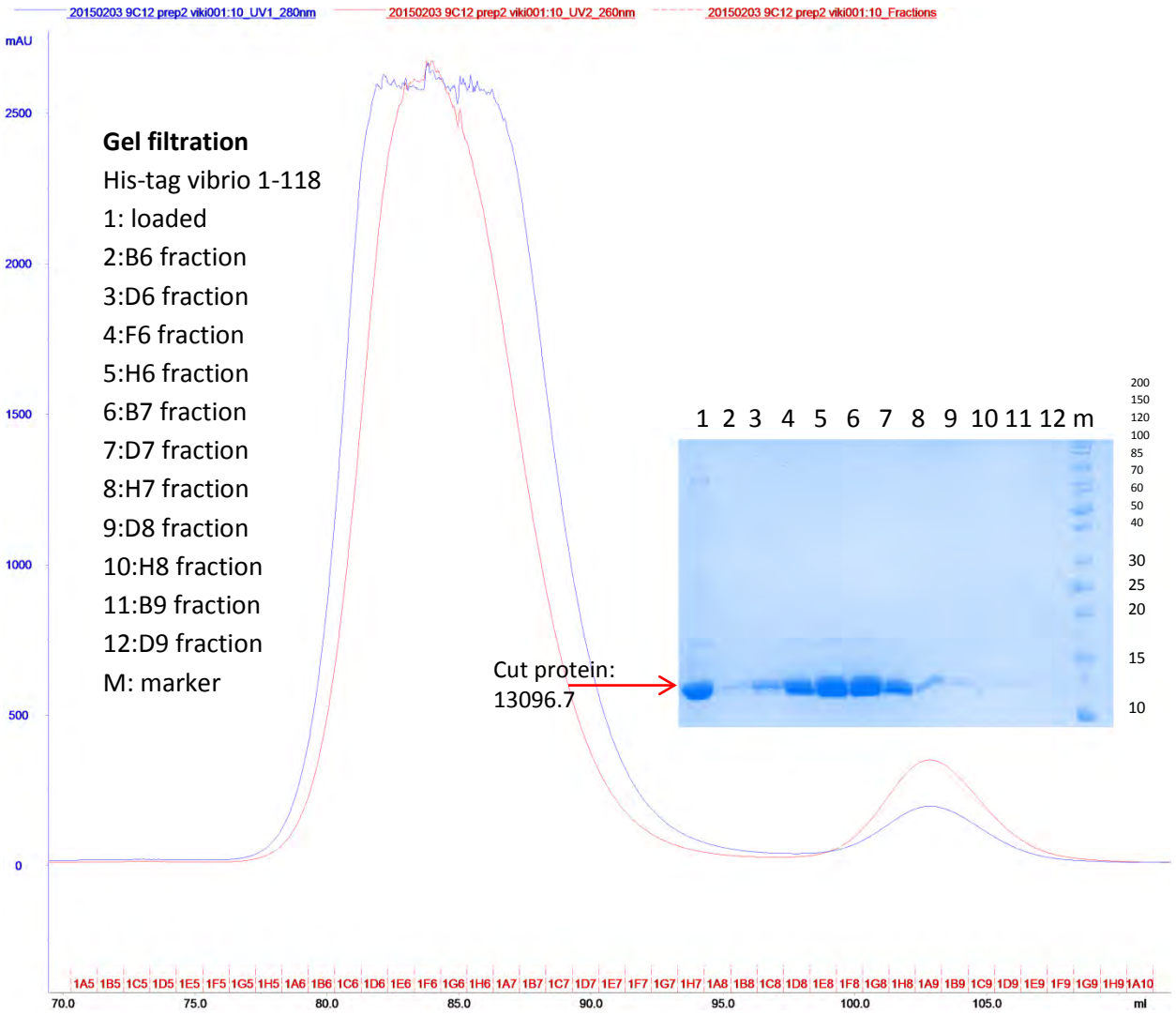
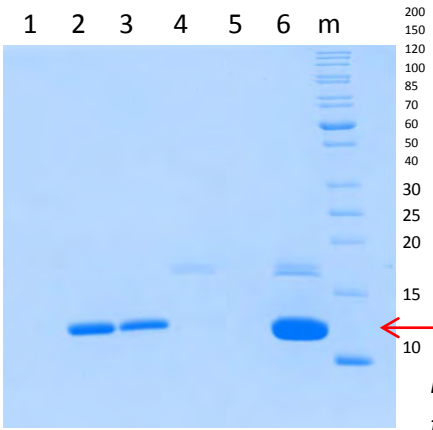


Fig.27: The Gel filtration profile of the His-tag vibrio 1-127 is illustrated above. The GF buffer was H20N100T2 pH 7.4.



Ion exchange

His-tag vibrio 1-118
1: F.T. sample
2: elution T20N25T1 pH8.5
3: elution T20N50T1 pH8.5
4: elution T20N150T1 pH8.5
5: elution T20N250T1 pH8.5
6: loaded
M: marker

Fig. 28: Due to the SUMO traces after the GF, an IEX was performed. On the SDS-PAGE gel, it is obvious that the bands from the lanes 2, 3 correspond to the vibrio 1-118 lipase.

MALLS

An additional method, Multi-Angle Laser Light Scattering (MALLS), was performed for this protein. The mass estimation by MALLS is indeed in accordance with the expected protein mass (fig. 29). The same analysis also allows the characterization of the molecule/protein as a monomer, dimer or polymer. At this study, according to the MALLS run the vibrio 1-118 is a monomeric in solution (fig. 29). After SEC, the fractions corresponding to the peak were collected and were analyzed on an SDS-PAGE gel to be confirmed that the applied protein on the SEC/MALLS was pure (Suppl. fig.6).

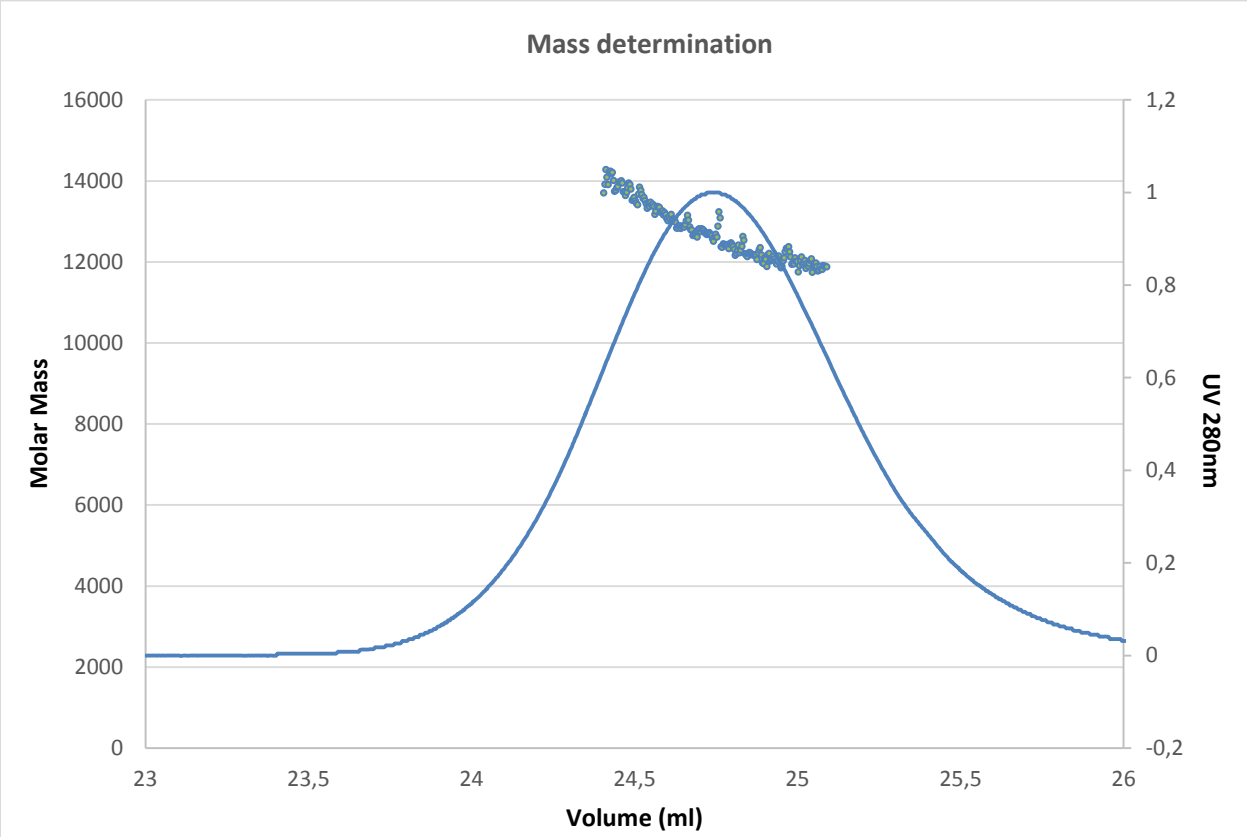
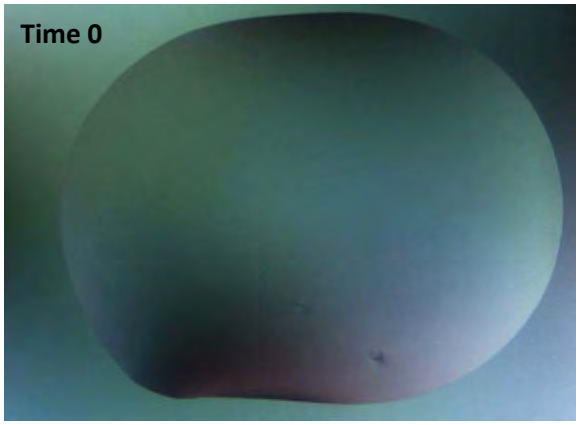


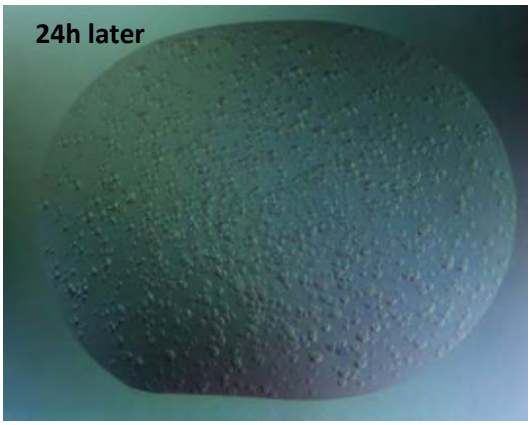
Fig. 29: Normalized UV profiles and MALLS signal from size exclusion chromatography of the vibrio 1-118 indicate molecular weight of 13kDa as a monomer in solution.

After the purification of this protein, JCSG+, pH Clear Suite, Protein Complex Suite, Index, ComPas Suite, Classics I,II Suite, PACT Suite, MIDAS, MORPHEUS, PGA-LM trays were set up at both 4^{o}C and room temperature. All these plates were checked under the optic microscope and through the rock imager robot. Crystallization screens of this protein produced needle crystals and spherulites in several conditions. The hits coming from the initial screening conditions are shown at the images below. These images are from the Rock Imager robot. All the hits appeared at room temperature at INDEX, pH Clear, ProComplex.

Vibrio protein 1-118

Index B7





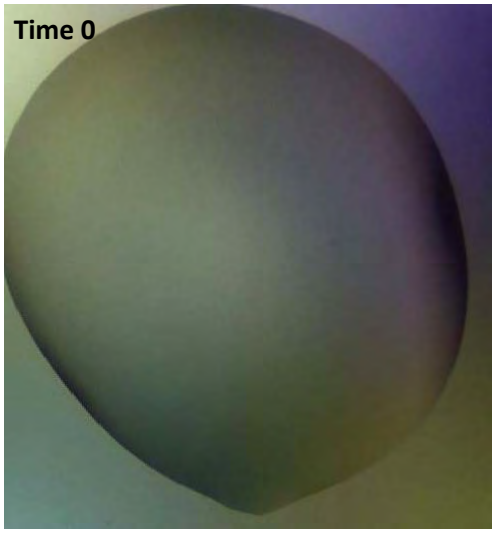
Temperature: 20°C

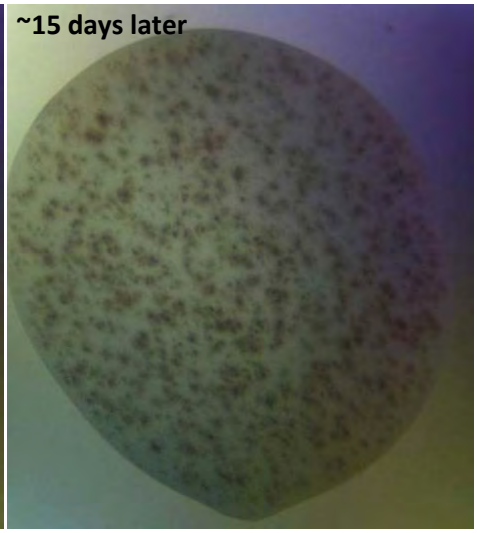
Prot. Buffer: H20N~30T1 pH 7.4
Prot. Conc: 18.23 mg/mL
Yield(mg/lt): 1,823 mg/lt

INDEX (B7):
0,06M NaH2PO4&1,34M K2HPO4

Droplet:
200nl protein:100nl reservoir

pH Clear E10





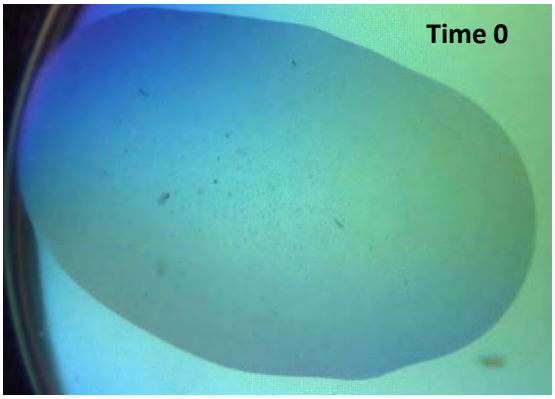
Temperature: 20°C

Prot. Buffer: H20N~30T1 pH 7.4
Prot. Conc: 18.23mg/mL
Yield(mg/lt): 1,823 mg/lt

pH Clear (E10):
0,1M Hepes pH 7
1,6M $(NH_4)_2SO_4$

Droplet:
200nl protein:100nl reservoir

ProComplex F12



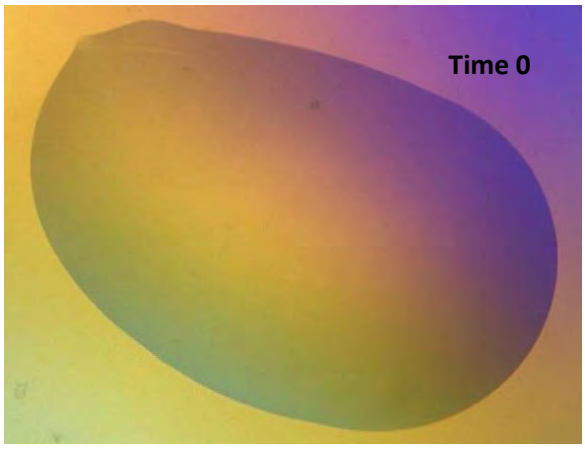
Temperature: 20°C

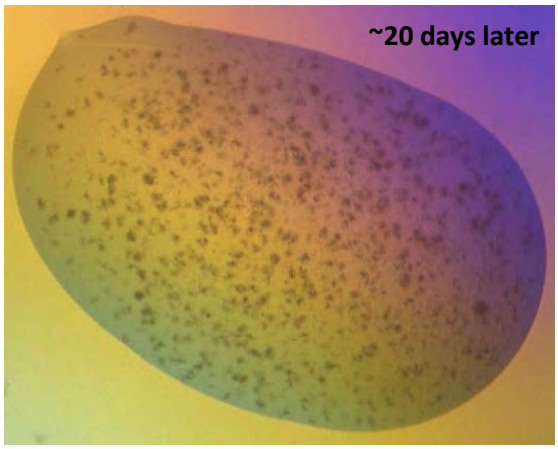
Prot. Buffer:
H20N~30T1 pH 7.4
Prot. Conc: 23.97 mg/mL
Yield(mg/lt): 2,76 mg/lt

ProComplex F12:
0,1M Hepes pH 7
1,5M $(NH_4)_2SO_4$

Droplet:
200nl protein:100nl reservoir

ProComplex G10





Temperature: 20°C

Prot. Buffer:
H20N~30T1 pH 7.4
Prot. Conc: 23,97 mg/mL
Yield(mg/lt): 2,76 mg/lt

ProComplex G10:
1,6M $(Na/K)HPO_4$

Droplet:
200nl protein:100nl reservoir

The conditions required for these hits are mentioned on the right side of the Rock Imager pictures. The protein concentration, the buffer of the protein, the droplet ratio are also reported.

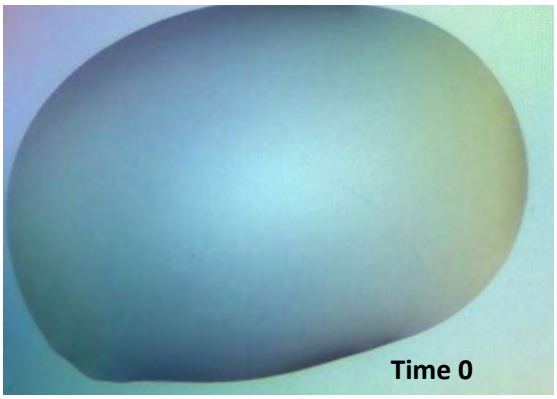
When the hits appeared, the same and slightly different conditions were applied to the vibrio 1-118 protein for reproducing the initial hits. The first optim plate was set up with protein concentration 18.2 mg/mL and protein buffer H20N~30T1 pH 7.4 at ambient temperature. The droplet ratio was 200nL protein: 100nL reservoir. The conditions are described in details below.

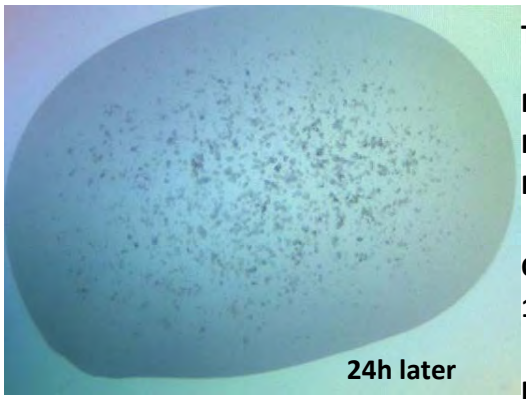
	1	2	3
A	0,1M Hepes pH: 6.5 1,25M [$AmSO_4$]	0,1M Hepes pH:6.5 1,5M [$AmSO_4$]	0,1M Hepes pH:6.5 1,75M [$AmSO_4$]
B	0,1M Hepes pH: 7 1,25M [$AmSO_4$]	0,1M Hepes pH:7 1,5M [$AmSO_4$]	0,1M Hepes pH:7 1,75M [$AmSO_4$]
C	0,1M Hepes pH: 7.5 1,25M [$AmSO_4$]	0,1M Hepes pH: 7.5 1,5M [$AmSO_4$]	0,1M Hepes pH: 7.5 1,75M [$AmSO_4$]
D	1,2M Na/K pH:8.2	1,4M Na/K pH:8.2	1,6M Na/K pH:8.2
E	1,4M Na/K pH: 7.8	1,4M Na/K pH 8	1,4M Na/K pH 8.2

(Na/K: NaH_2PO_4/K_2HPO_4)

The table 9 represents the conditions corresponding to the 1st optim plate.

Neither spherulites either needle crystals were formed at this plate. The most likely parameter for this, seems to be the protein concentration. Therefore, the second optim plate was designed with a view to determine the proper protein concentration. The conditions used, were almost the same with the first plate/conditions giving the initial hits. A slim difference at the Na/K concentration was only mentioned. This lies on the fact, that the protein with the highest concentration of Na/K, formed crystalline-like structures (see pictures below from the Rock Imager). Thus, a higher concentration of Na/K was used at the 2nd plate.





Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 18.2 mg/mL

Optim plate 1 (D3):
1,6M Na/K pH8.2

Droplet:
200nlprotein:100nl reservoir

The 2nd optim plate was set up with four different protein concentrations 19 mg/mL, 24 mg/mL, 26 mg/mL and 28 mg/mL. The protein buffer was H20N30T1 pH 7.4 at 20°C. The droplet was 100nl protein: 100 nl reservoir. The chemicals used for setting up this plate are written below.

	1	2	3
A	0,1M Hepes pH: 6.5 1,25M [$AmSO_4$]	0,1M Hepes pH:6.5 1,5M [$AmSO_4$]	0,1M Hepes pH:6.5 1,75M [$AmSO_4$]
B	0,1M Hepes pH: 7 1,25M [$AmSO_4$]	0,1M Hepes pH:7 1,5M [$AmSO_4$]	0,1M Hepes pH:7 1,75M [$AmSO_4$]
C	0,1M Hepes pH: 7.5 1,25M [$AmSO_4$]	0,1M Hepes pH: 7.5 1,5M [$AmSO_4$]	0,1M Hepes pH: 7.5 1,75M [$AmSO_4$]
D	1,4M Na/K pH: 7.8	1,4M Na/K pH:8	1,4M Na/K pH:8.2
E	1,6M Na/K pH: 7.8	1,6M Na/K pH:8	1,6M Na/K pH:8.2
F	1,8M Na/K pH: 7.8	1,8M Na/K pH: 8	1,8M Na/K pH: 8.2

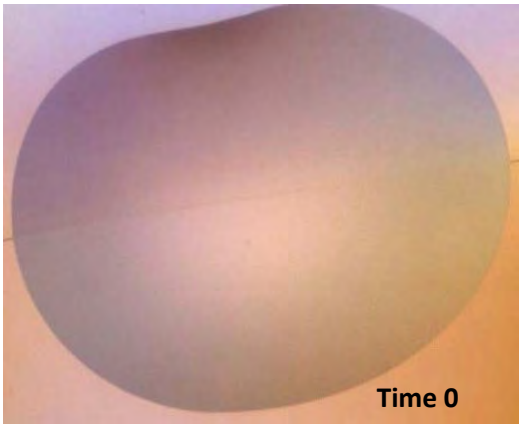
(Na/K: NaH_2PO_4/K_2HPO_4)

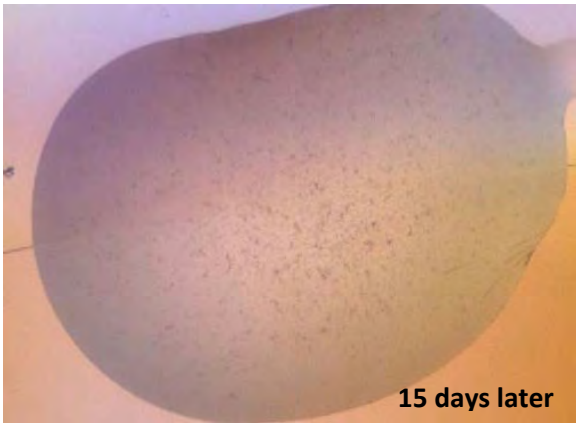
The table 10 describes the conditions corresponding to the 2nd optim plate.

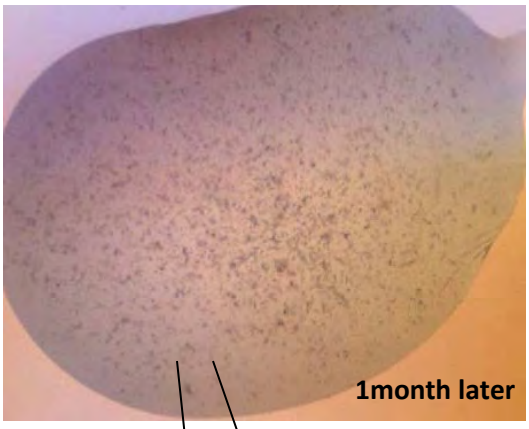
The conditions described on the table above (table 10) were tried for all the four protein concentrations (19 mg/mL, 24 mg/mL, 26 mg/mL and 28 mg/mL). The screening from this plate showed that the protein concentrations of 19mg/ml and 24mg/ml aid the formation of the most well shaped and distinguished needle crystals, whilst the other two concentrations (26 mg/mL and 28 mg/mL) give more clustered needle crystals with a characteristic star-like shape.

The images below come from the Rock Imager system. Although for better image resolution, Visi camera 5.0 was also used for capturing images manually.

OPTIM2 **Protein concentration: 19 mg/mL**





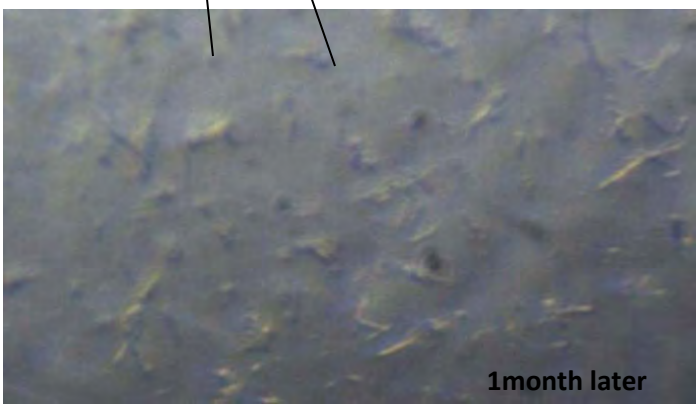


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 19 mg/mL
Prot. yield: 1,7mg/lt

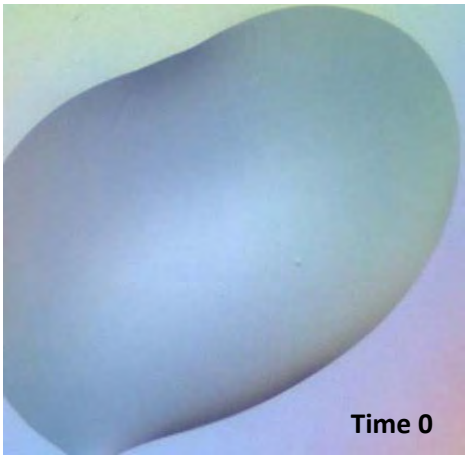
Optim plate 2 (A2):
0,1M Hepes pH6.5
1,5M [$AmSO_4$]

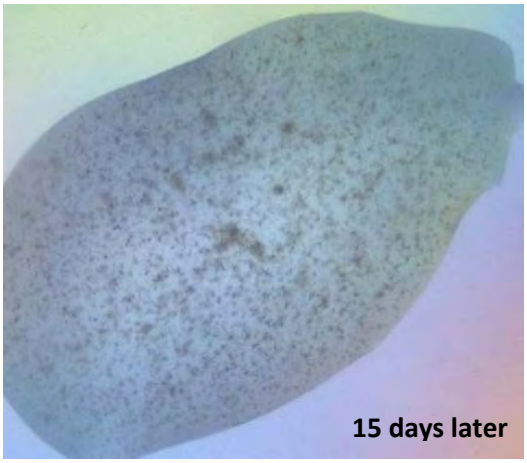
Droplet:
100nlprotein:100nl reservoir

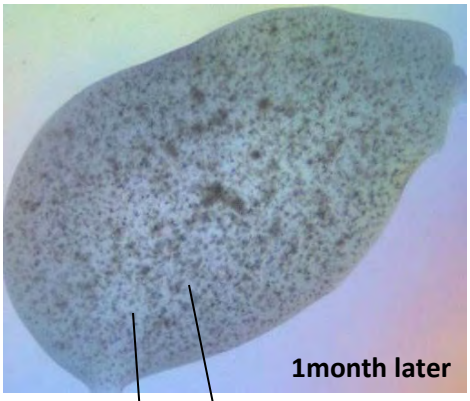


This image was captured by the Visi Cam 5.0 for better resolution.

OPTIM2 **Protein concentration: 24 mg/mL**





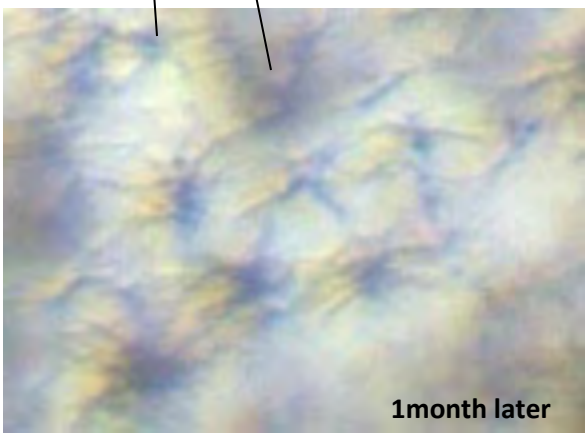


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 24 mg/mL
Prot. Yield: 1,7 mg/lt

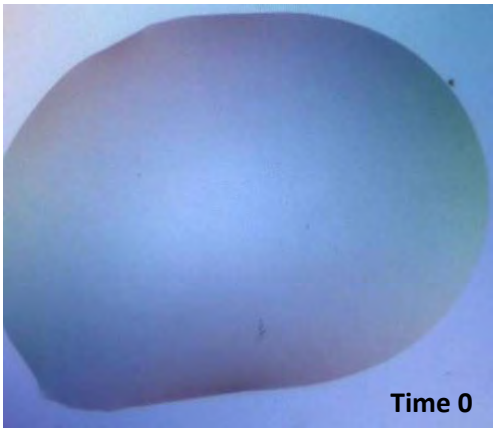
Optim plate 2 (A5):
0,1M Hepes pH6.5
1,5M [$AmSO_4$]

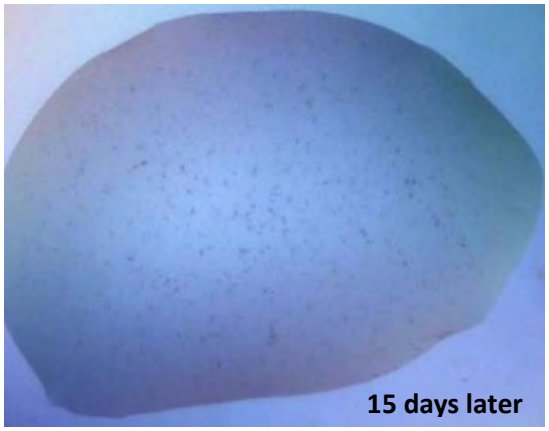
Droplet:
100nlprotein:100nl reservoir

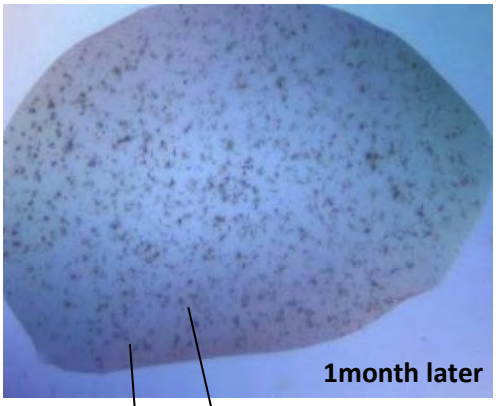


This image was captured by the Visi Cam 5.0 for better resolution.

OPTIM2 **Protein concentration: 26 mg/mL**





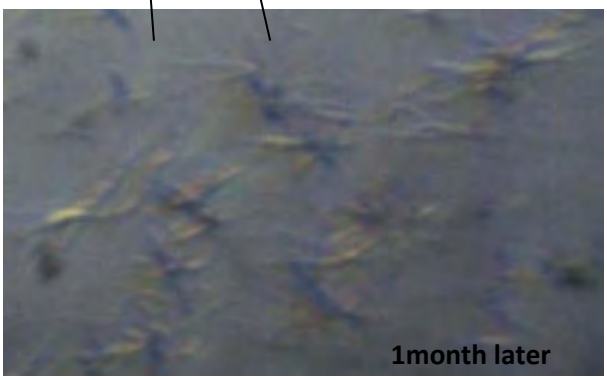


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 26 mg/mL
Prot. Yield: 1,7 mg/lt

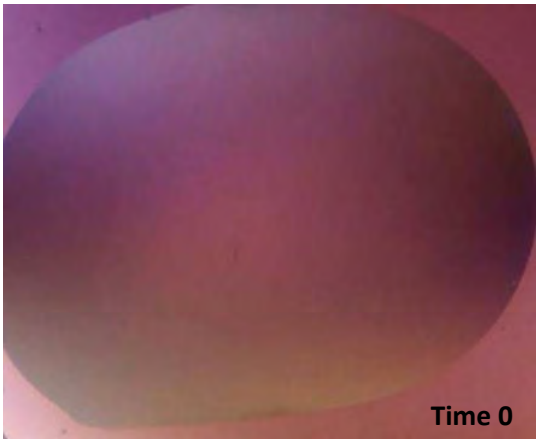
Optim plate 2 (B8):
0,1M Hepes pH7
1,5M [$AmSO_4$]

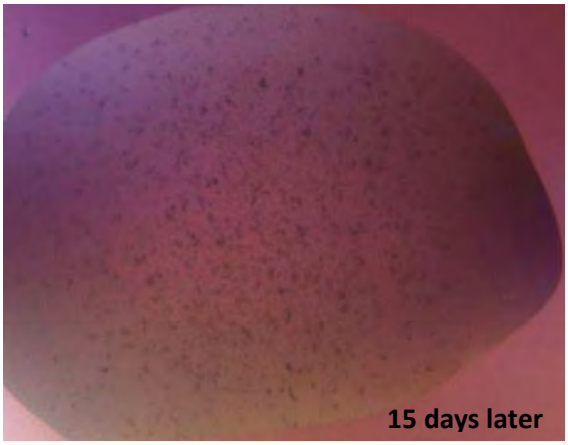
Droplet:
100nlprotein:100nl reservoir

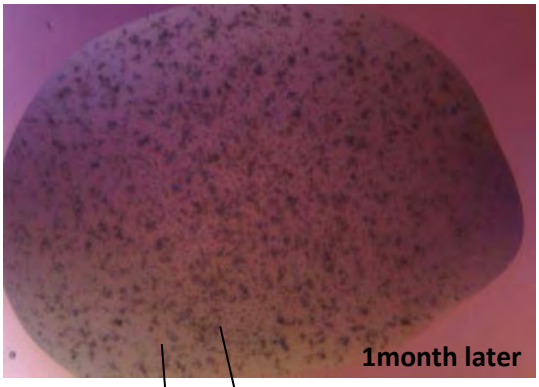


This image was captured by the Visi Cam 5.0 for better resolution.

OPTIM2 **Protein concentration: 28 mg/mL**







Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 28 mg/mL
Prot. Yield: 1,7 mg/lt

Optim plate 2 (B11):
0,1M Hepes
pH 7
1,5M [$AmSO_4$]

Droplet:
100nlprotein:100nl reservoir

This image was captured by the Visi Cam 5.0 for better resolution.

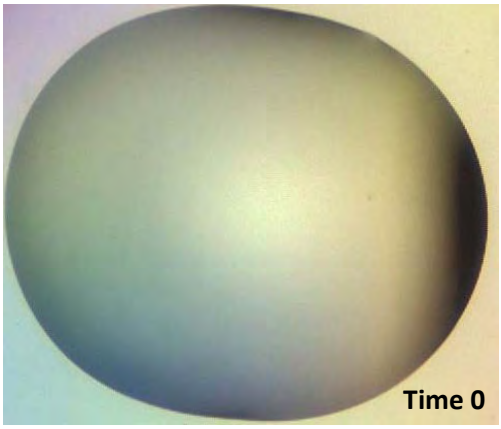
The screening of the optim 2 plate facilitated the design of the optim 3 plate. All the hits from the optim 2 plate had reservoirs with $AmSO_4$ concentration 1,5M which directed us to set the optim 3 plate with $AmSO_4$ concentration closed to 1,5M (1,4M, 1,5M, 1,6M). Another parameter that was also checked on the optim 3 plate was the effect of a different buffer. The pH range of the buffers was chosen to be between 6.5 and 7.5. The used buffers were: TRIS pH 6.5 and 7.2, MES pH 7, BTP pH 6.5, HEPES pH 6.5 and 7.5. The last variable was the addition of NaCl (salt) and/or glycerol. As it was described before, the chosen protein concentrations were: 19 mg/mL and 24 mg/mL.

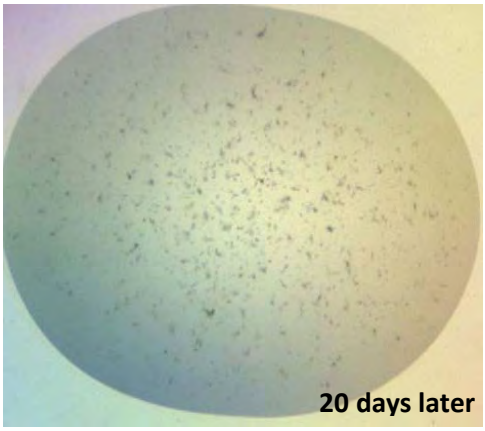
	1	2	3	4	5	6	7	8	9	10	11	12
A	0,1M TRIS pH: 6.5 1,4M [$AmSO_4$]	0,1M TRIS pH: 6.5 1,5M [$AmSO_4$]	0,1M TRIS pH: 6.5 1,6M [$AmSO_4$]	0,1M TRIS pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M TRIS pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M TRIS pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M TRIS pH: 6.5 1,4M [$AmSO_4$] 0,05% glycerol	0,1M TRIS pH: 6.5 1,5M [$AmSO_4$] 0,05% glycerol	0,1M TRIS pH: 6.5 1,6M [$AmSO_4$] 0,05% glycerol	0,1M TRIS pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M TRIS pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M TRIS pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
B	0,1M TRIS pH: 7.2 1,4M [$AmSO_4$]	0,1M TRIS pH: 7.2 1,5M [$AmSO_4$]	0,1M TRIS pH: 7.2 1,6M [$AmSO_4$]	0,1M TRIS pH: 7.2 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M TRIS pH: 7.2 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M TRIS pH: 7.2 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M TRIS pH: 7.2 1,4M [$AmSO_4$] 0,05% glycerol	0,1M TRIS pH: 7.2 1,5M [$AmSO_4$] 0,05% glycerol	0,1M TRIS pH: 7.2 1,6M [$AmSO_4$] 0,05% glycerol	0,1M TRIS pH: 7.2 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M TRIS pH: 7.2 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M TRIS pH: 7.2 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
C	0,1M MES pH: 7 1,4M [$AmSO_4$]	0,1M MES pH: 7 1,5M [$AmSO_4$]	0,1M MES pH: 7 1,6M [$AmSO_4$]	0,1M MES pH: 7 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M MES pH: 7 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M MES pH: 7 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M MES pH: 7 1,4M [$AmSO_4$] 0,05% glycerol	0,1M MES pH: 7 1,5M [$AmSO_4$] 0,05% glycerol	0,1M MES pH: 7 1,6M [$AmSO_4$] 0,05% glycerol	0,1M MES pH: 7 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M MES pH: 7 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M MES pH: 7 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
D	0,1M BTP pH: 6.5 1,4M [$AmSO_4$]	0,1M BTP pH: 6.5 1,5M [$AmSO_4$]	0,1M BTP pH: 6.5 1,6M [$AmSO_4$]	0,1M BTP pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M BTP pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M BTP pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M BTP pH: 6.5 1,4M [$AmSO_4$] 0,05% glycerol	0,1M BTP pH: 6.5 1,5M [$AmSO_4$] 0,05% glycerol	0,1M BTP pH: 6.5 1,6M [$AmSO_4$] 0,05% glycerol	0,1M BTP pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M BTP pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M BTP pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
E	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$]	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$]	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$]	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
F	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol

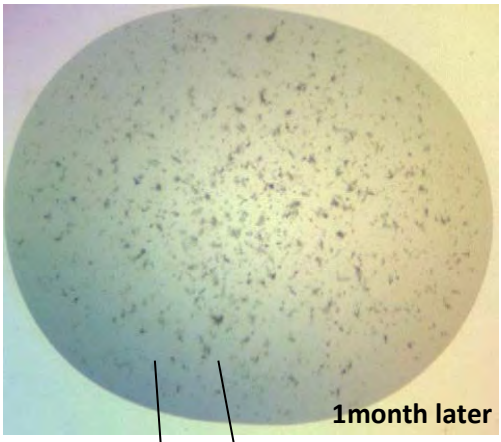
The table 11 describes in details the used chemicals for the 3rd and 4th optim plates.

For testing both 19 mg/mL (OPTIM3) and 24 mg/mL (optim 4) protein concentrations, two plates were set up with the conditions showing at the table above. The images below come from the Rock Imager system. Although for better image resolution, Visi camera 5.0 was also used for capturing images manually (scale 10mm).

OPTIM3 **0,1M MES pH 7 & 1,5M [$AmSO_4$]**





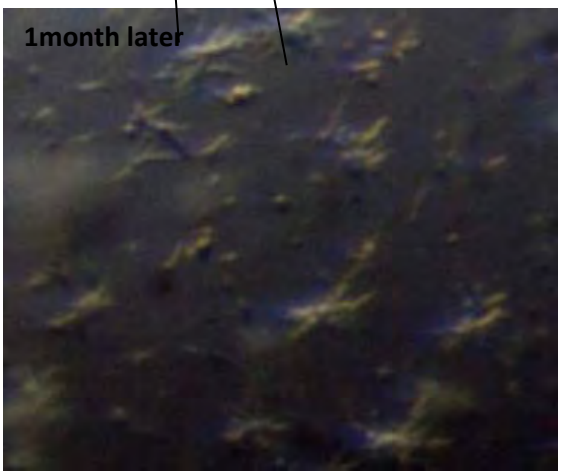


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 19 mg/mL
Prot. Yield: 1,7 mg/lt

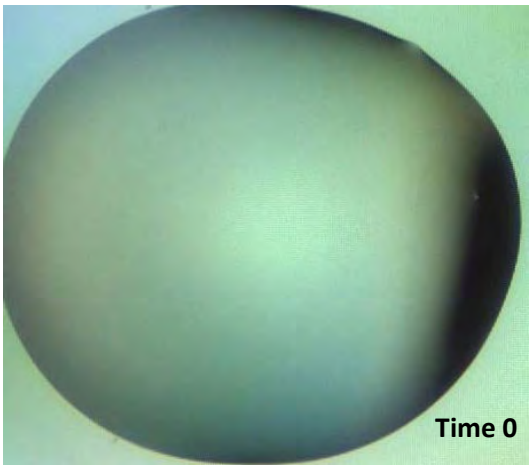
Optim plate 3 (C2):
0,1M MES
pH 7
1,5M [$AmSO_4$]

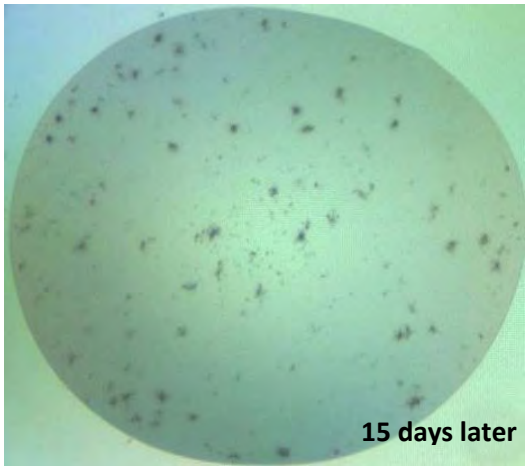
Droplet:
200nlprotein:100nl reservoir

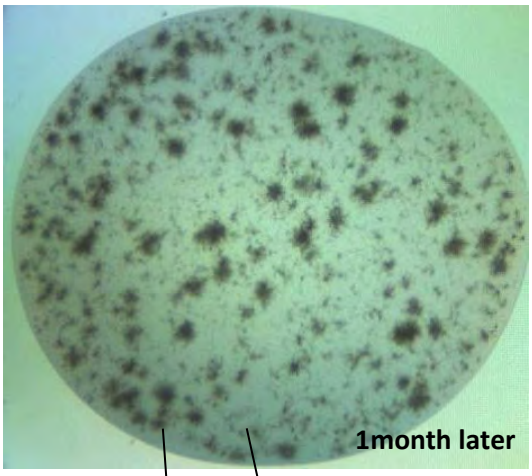


This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

OPTIM3 **0,1M MES pH: 7 & 1,6M [AmSO4]**





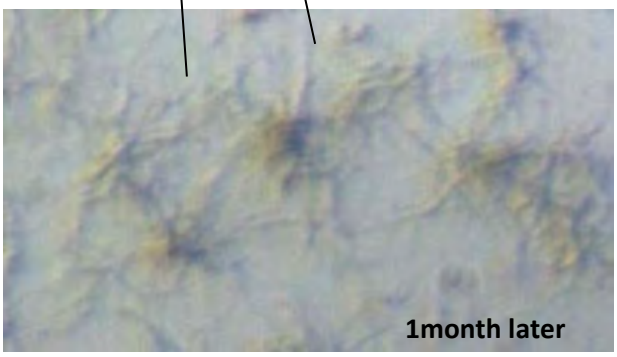


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 19 mg/mL
Prot. Yield: 1,7 mg/lt

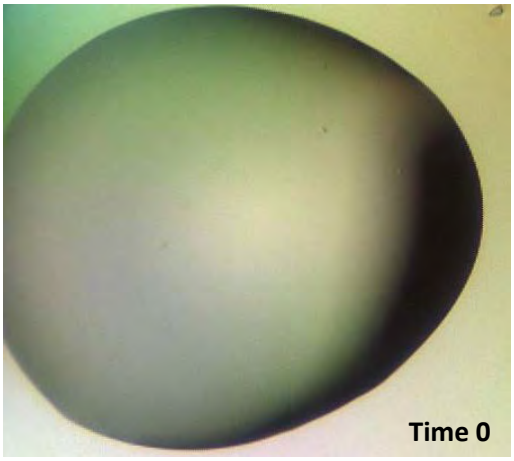
Optim plate 3 (C3):
0,1M MES
pH 7
1,6M [$AmSO_4$]

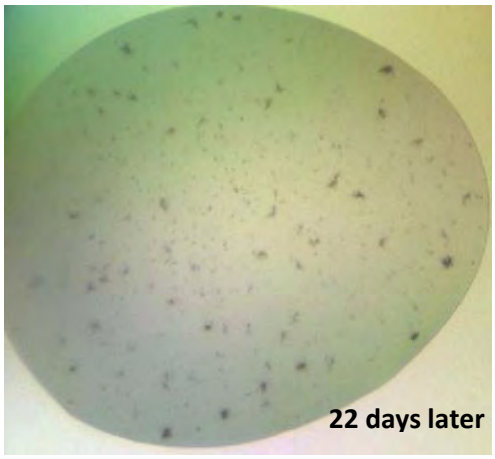
Droplet:
200nlprotein:100nl reservoir

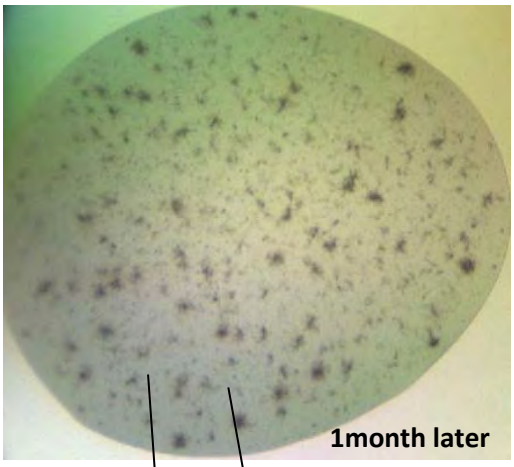


This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

OPTIM3 **0,1M MES pH: 7 & 1,6M [$AmSO_4$] & 0,1M NaCl**





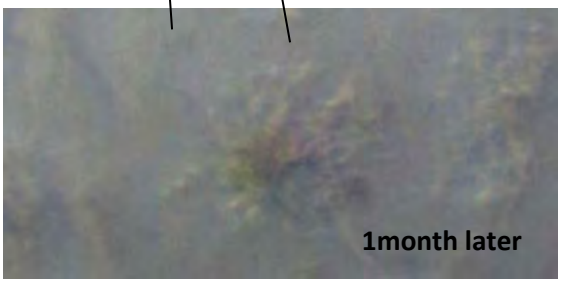


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 19 mg/mL
Prot. Yield: 1,7 mg/lt

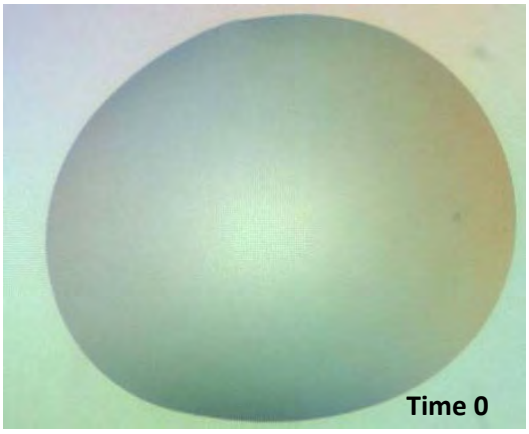
Optim plate 3 (C6):
0,1M MES pH 7
1,6M [$AmSO_4$]
0,1 M NaCl

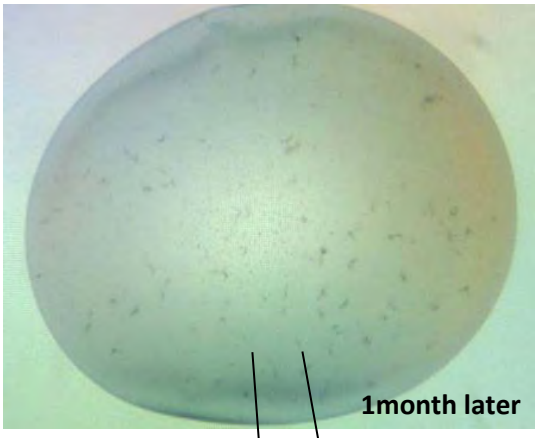
Droplet:
200nlprotein:100nl reservoir



This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

OPTIM3 0,1M HEPES pH: 6.5 & 1,6M [$AmSO_4$]



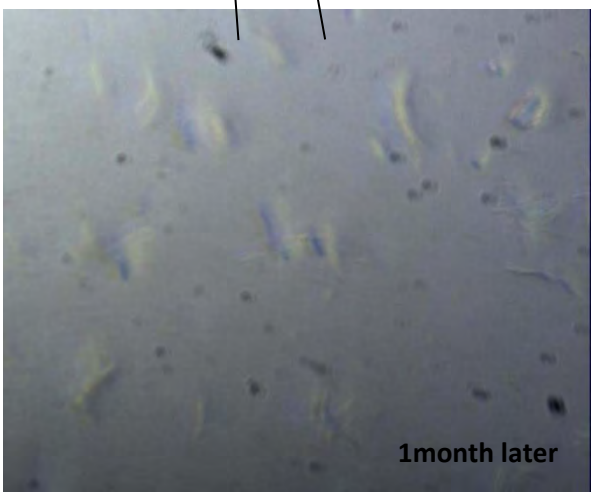


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 19 mg/mL
Prot. Yield: 1,7 mg/lt

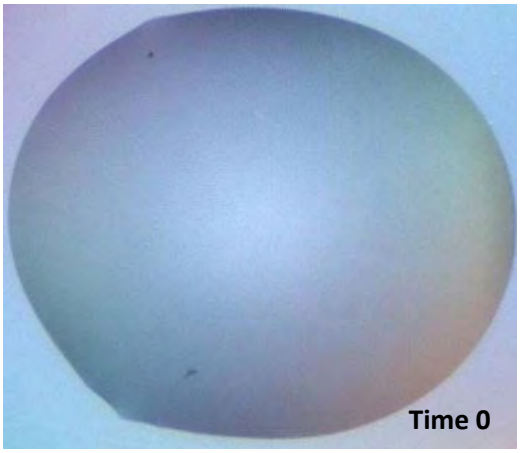
Optim plate 3 (E3):
0,1M HEPES
pH 6.5
1,6M [$AmSO_4$]

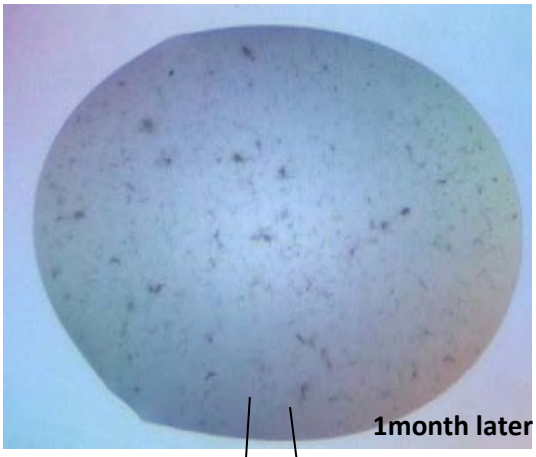
Droplet:
100nlprotein:100nl reservoir



This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

OPTIM4 0,1M MES pH 7 & 1,6M [$AmSO_4$]



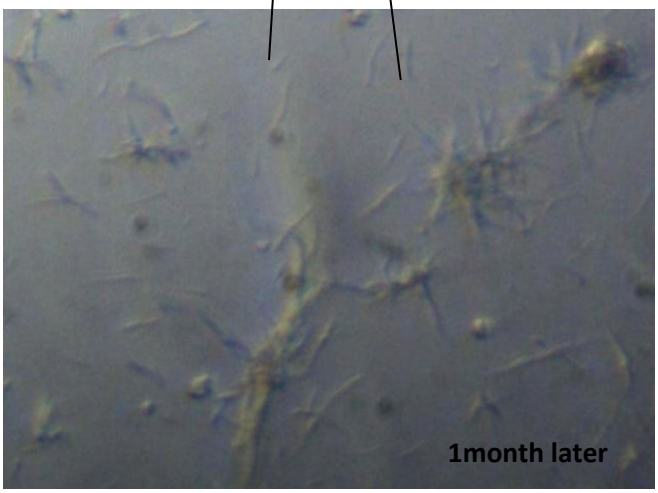


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 24 mg/mL
Prot. Yield: 1,7 mg/lt

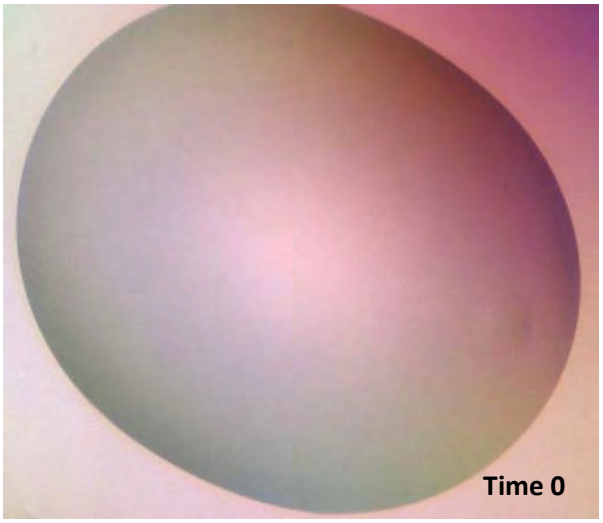
Optim plate 4 (C3):
0,1M MES
pH 7
1,6M [$AmSO_4$]

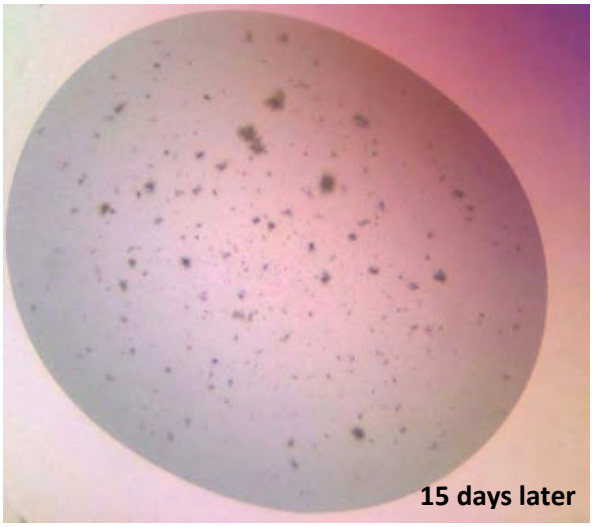
Droplet:
100nlprotein:100nl reservoir

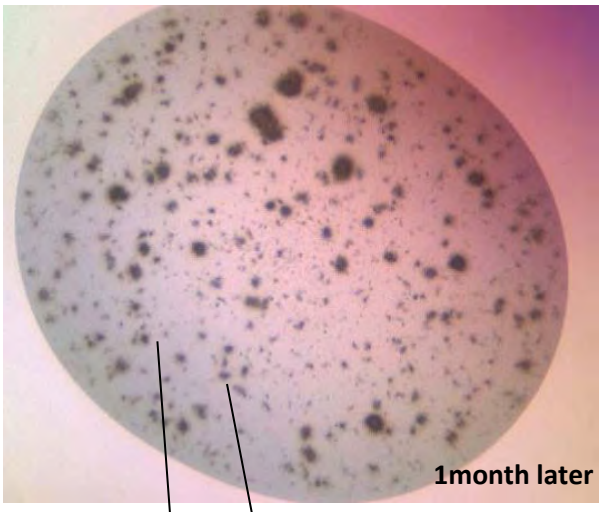


This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

OPTIM4 **0,1M HEPES pH: 6.5 & 1,6M [$AmSO_4$] & 0,1M NaCl**





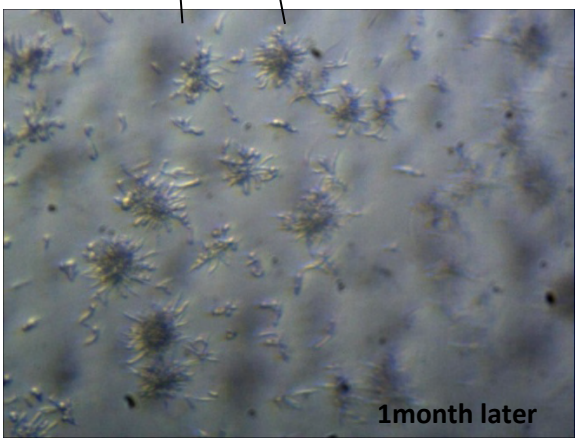


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 24 mg/mL
Prot. Yield: 1,7 mg/lt

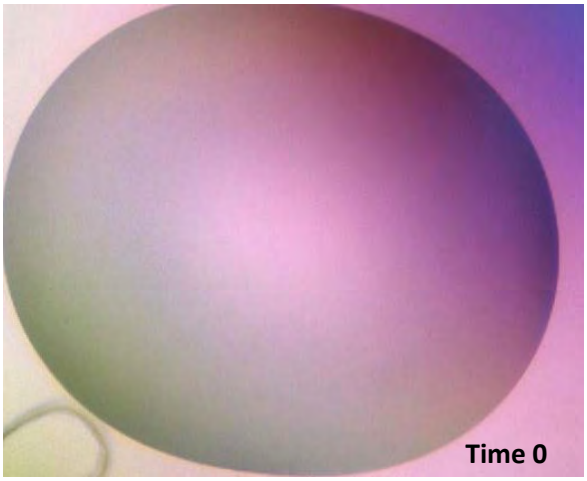
Optim plate 4 (E4):
0,1M HEPES
pH 6.5
1,6M [$AmSO_4$]
0,1M NaCl

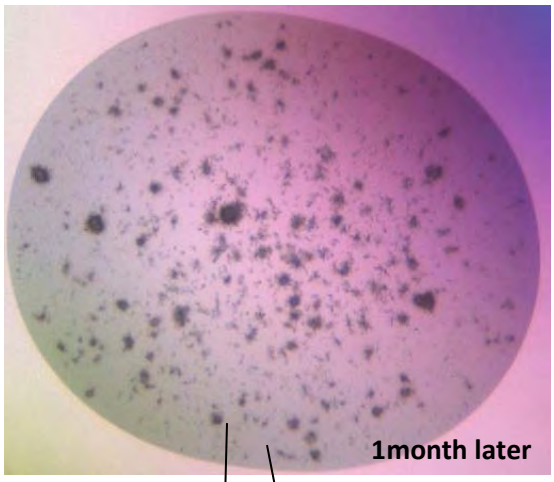
Droplet:
200nlprotein:100nl reservoir



This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

OPTIM4 **0,1M HEPES pH: 7.5 & 1,6M [$AmSO_4$] & 0,1M NaCl**



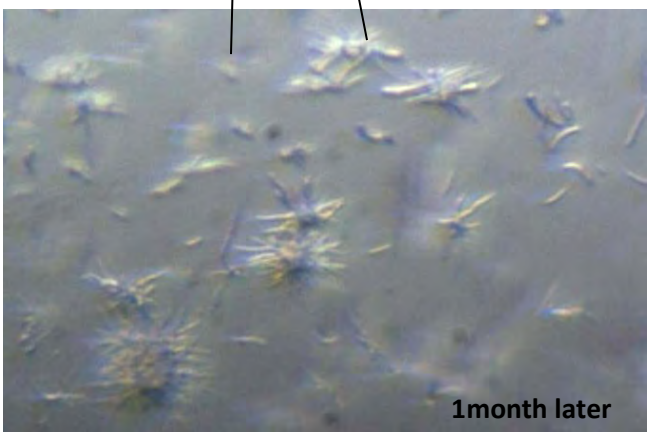


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 24 mg/mL
Prot. Yield: 1,7 mg/lt

Optim plate 4 (F4):
0,1M HEPES
pH 7.5
1,6M [$AmSO_4$]
0,1M NaCl

Droplet:
200nlprotein:100nl reservoir



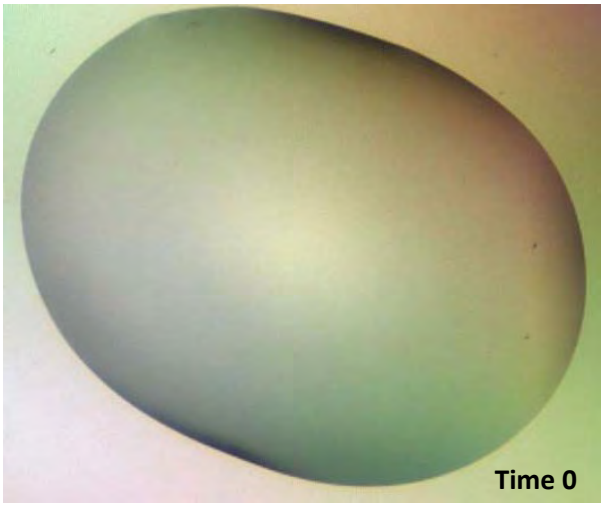
This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

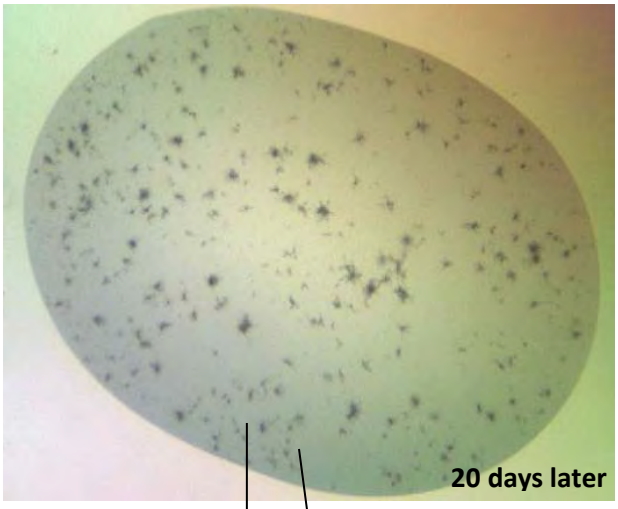
After the hits coming from the optim3 and optim4 plates, another plate was set up. The optim5 includes the conditions that gave hits at the optim3 and optim4 plates. The table 12 describes these conditions in details. 19 mg/mL and 24 mg/mL protein concentrations were tried.

	1	2	3	4	5	6	7	8	9	10	11	12
A	0,1M MES pH: 7 1,4M [$AmSO_4$]	0,1M MES pH: 7 1,5M [$AmSO_4$]	0,1M MES pH: 7 1,6M [$AmSO_4$]	0,1M MES pH: 7 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M MES pH: 7 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M MES pH: 7 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M MES pH: 7 1,4M [$AmSO_4$] 0,05% glycerol	0,1M MES pH: 7 1,5M [$AmSO_4$] 0,05% glycerol	0,1M MES pH: 7 1,6M [$AmSO_4$] 0,05% glycerol	0,1M MES pH: 7 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M MES pH: 7 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M MES pH: 7 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
B	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$]	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$]	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$]	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 6.5 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 6.5 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
C	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$] 0,1 M NaCl	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$] 0,05% glycerol	0,1M HEPES pH: 7.5 1,4M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol	0,1M HEPES pH: 7.5 1,6M [$AmSO_4$] 0,1 M NaCl 0,05% glycerol
D	0,1M HEPES pH: 6.5 1,5M [$AmSO_4$]	0,1M HEPES pH: 6.8 1,5M [$AmSO_4$]	0,1M HEPES pH: 7 1,5M [$AmSO_4$]	0,1M HEPES pH: 7.4 1,5M [$AmSO_4$]	0,1M HEPES pH: 7.5 1,5M [$AmSO_4$]	0,1M HEPES pH: 7.8 1,5M [$AmSO_4$]	0,1M HEPES pH: 8.2 1,5M [$AmSO_4$]					

The table 12 shows the used conditions for the 5th optim plate.

OPTIM5 0,1M HEPES pH 7.4 & 1,5M [AmSO4]



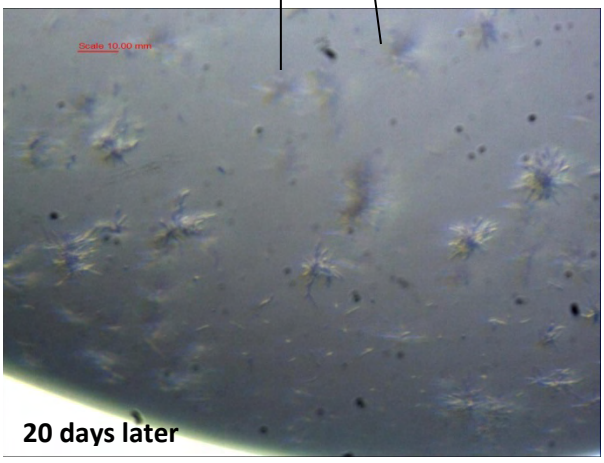


Temperature: 20°C

Prot. Buffer:
H20N30T1 pH 7.4
Prot. Conc: 19 mg/mL
Prot. Yield: 1,7 mg/lt

Optim plate 5 (D4):
0,1M HEPES
pH 7.4
1,5M [$AmSO_4$]

Droplet:
200nlprotein:100nl reservoir



This image was captured by the Visi Cam 5.0 for better resolution (scale 10mm).

The optim5 plate was checked last time after 20days from the date that it was set up, as it is also indicated on the pictures above. This could be the reason that no more hits were found.

All these five optim plates direct the further screening. First of all, the optim protein concentration is 24 mg/mL and it is proven of the hits coming from the optim3 and optim4 plates. When the protein concentration was 24 mg/mL, well-shaped needle crystals were formed. On the other hand, when the protein concentration was 19mg/mL, smaller and thinner clusters of needle crystals appeared. The buffers that seem to be more compatible to this protein for creating crystals, were HEPES and MES with a range of pH between 6.5-7.5. The $AmSO_4$ concentration that gave the most well-shaped crystals was 1,5M or 1,6M. The addition of NaCl could give different and better crystal images. Finishing, more conditions should be tried for the crystallization of this protein. To be more detailed, different salts, additives, more buffers are some of them.

PLA2G16 Activity assay

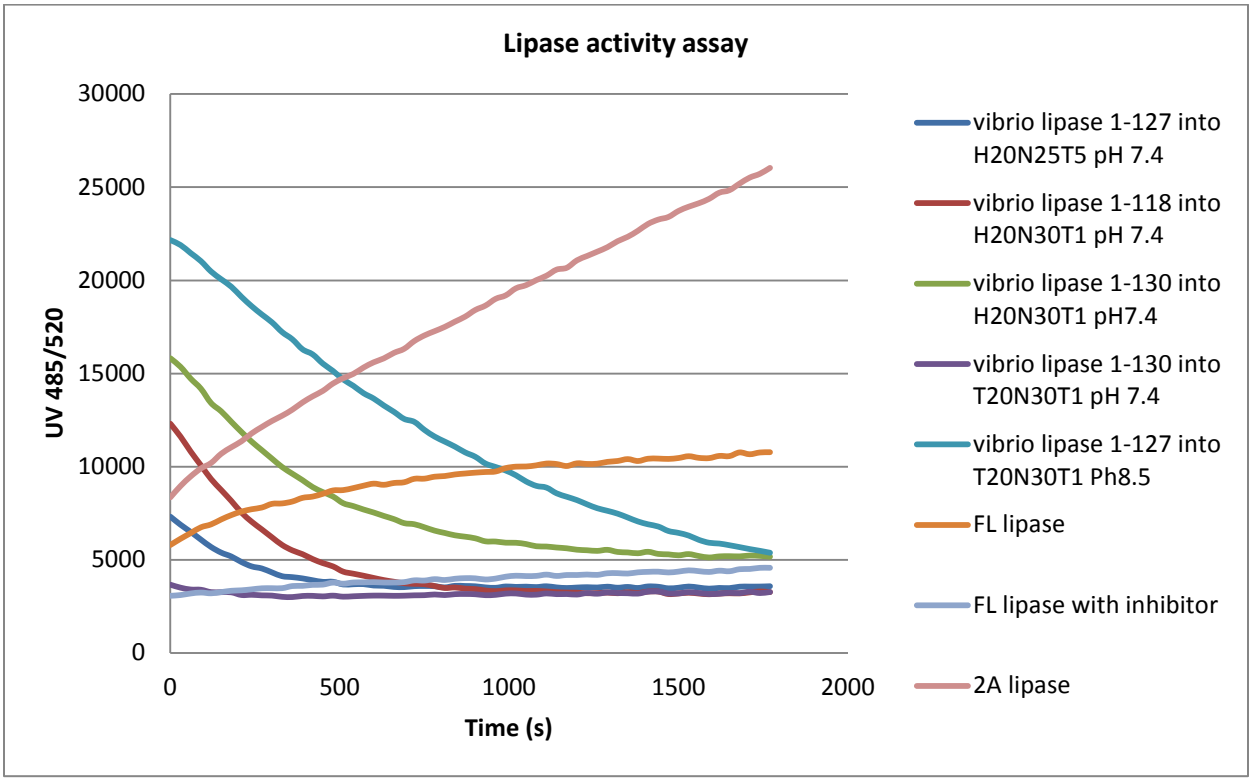


Fig.30: The diagram above shows the results from the lipase activity assay.

The lipase activity assay indicates that the C-terminal domain is required for the lipase to act as an enzyme. It was easily proven when the activity assay was applied to the truncated constructs, since the lack of the C-terminal reflects the formation of an inactive enzyme. For this assay, the FL lipase was used as a positive control. By comparing the activity of the FL lipase (positive control) with the truncated constructs, it is obvious that the lipase is active only in the presence of the C-terminal. Negative control was the FL lipase incubated with the inhibitor. The results of this assay shows clearly the importance of the C-terminal domain for the lipase, acting as an enzyme.

Our results show that the presence of the C-terminal hydrophobic domain adversely affects the solubility of the bacterially expressed vibrio lipase. Additionally this domain is undoubtedly important for the activity of the lipase.

5. Conclusion

The final goal of this study was the structural determination of the *vibrio cholerae* lipase. The initial approach involved the trial of the full length lipase purification. For this reason, a lot of parameters were tested for their possible effect on the purification yield of the FL lipase. The main problem was that the FL lipase seems to be expressed in bacteria in an insoluble form. Therefore, except from media, strain, Lysis buffer composition, time of spinning down the sample, different tags were also tested. As it was shown, His-tag, GST-tag and His-SUMO tag didn't change significantly the protein solubility. Therefore, constructs with a shorter amino acid sequence were expressed and purified. The new truncated constructs displayed improved degree of solubility especially those with a His-tag or a His-SUMO tag. Comparing these two tags, the His-SUMO tag increased the protein solubility/purification yield, a fact confirmed by the growth of spherulites and needle crystals. Although, numerous different conditions should be further tried for the optimization of these hits to grow crystals suitable for X-ray diffraction experiments. To analyze the conditions that help the protein to be in a soluble form, several different buffers were tested by an OPTIM analysis.

The differences observed in the solubility of the short constructs might be attributed to the absence of the C-terminal. The C-terminal part of the protein seems to be hydrophobic which means that in its presence the entire protein becomes insoluble. This finding implies that the C-terminal Domain may be associated with the membranes.

6. Further research

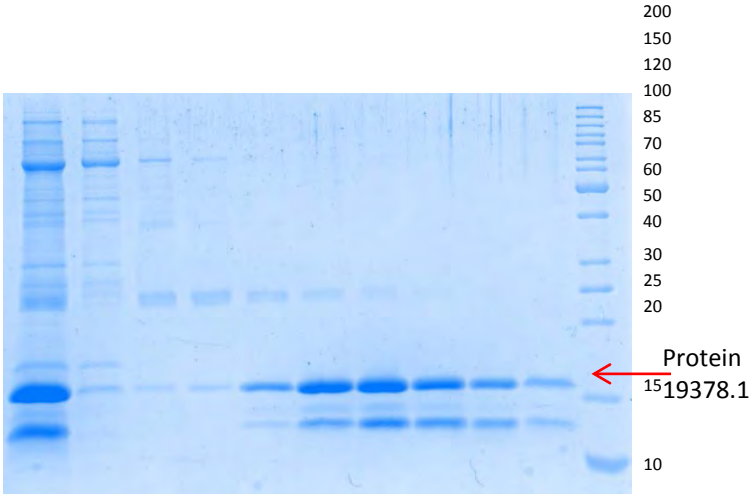
The aim of this project was the structural determination of the vibrio lipase. Although, the purification yield of the FL vibrio lipase was inefficient for setting up crystallization trays. This might happen due to folding problems or most probably due to the C-terminal hydrophobic part of the protein. An additional approach was the expression and the purification of truncated constructs for avoiding the putative /possible effects of the C-terminal.

Other tags (for instance: trigger factor) should be also tried. By this way, the full length protein might be more soluble and therefore its purification will be possible. An alternative option is the replacement of the *E.coli* cells with insect cells. This should be combined with the adjustment of the lysis buffer composition. Based on the previous preps (major part of the protein at the pellet), it is recommended to change the buffer composition, by adding for example a detergent. By this way, during its solubilization the possible membrane protein environment switches from a lipid environment to that of a non-native environment where lipid molecules are replaced by detergent molecules. In case that the yield of the protein is increased, a range of potential buffers could be checked by an optim analysis. The pH, the salt concentration, the replacement of the TCEP, additives are all variables that can effect on the protein stability and solubility.

7. *Supplementary tables & figures*

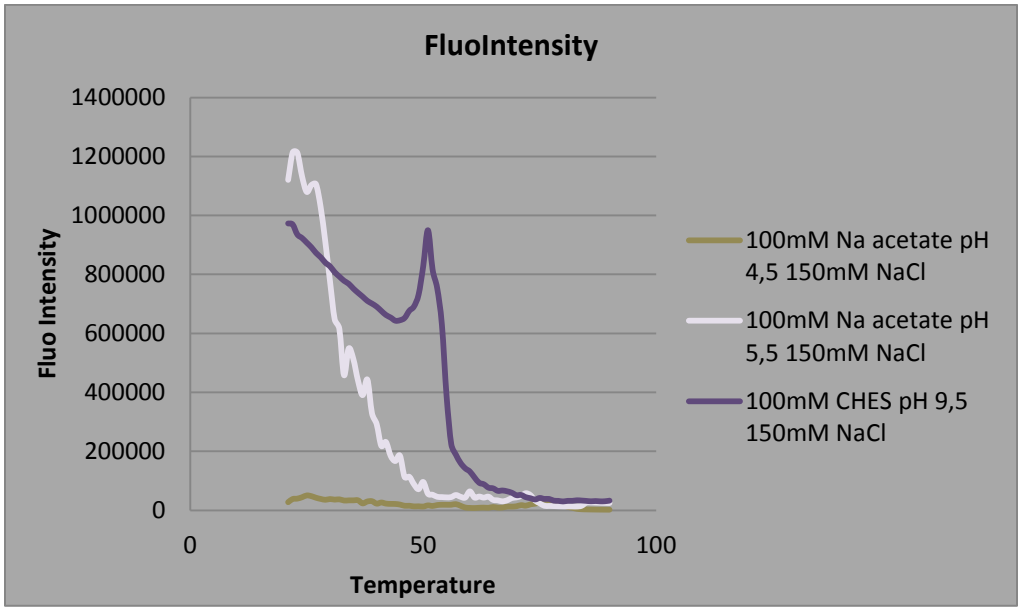
Relevant estimation of Molar Mass (see also Fig.6)

Mobility (mm)	1/(log of mass)
6	1
14	0,850274154
24	0,768621787
28	0,715338279
32	0,676992493
38	0,624196351
42	0,58859191
44	0,562381856
45	0,54197662
46	0,518290759
47	0,5
48	0,480958551
49	0,459539551
50	0,43458799

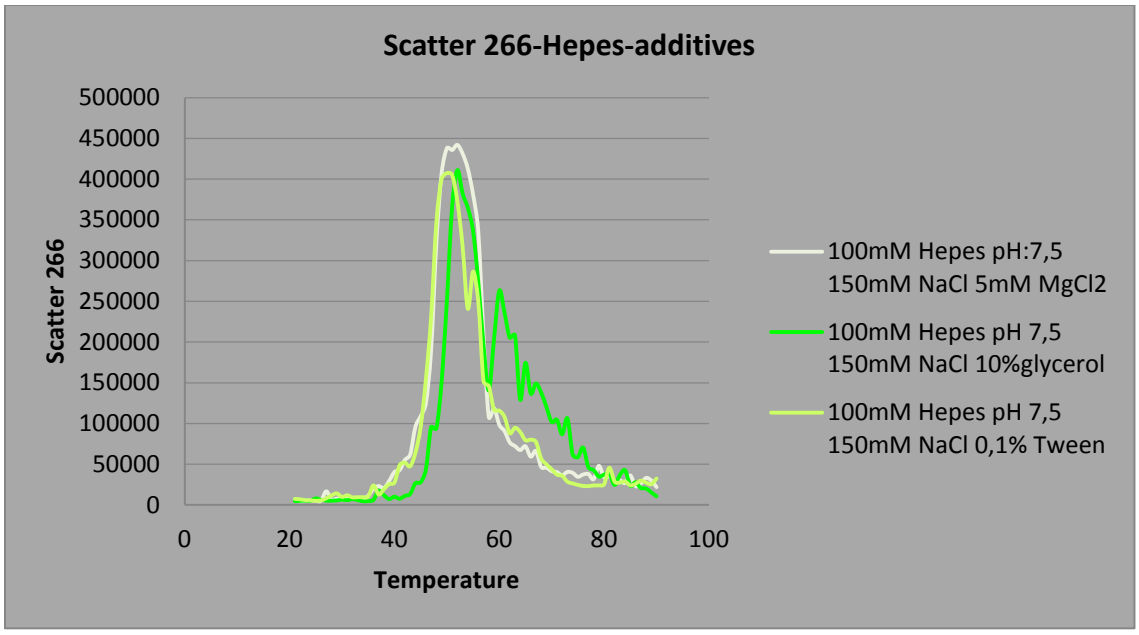


Suppl.Table1: The table shows the mobility of the marker expressed into mm accompanied by the calculations of the inverse logarithm of the marker's mass. This can lead to the relative estimation of the molecular mass corresponding to the unknown bands appeared on the gel (Biochemistry Stryer).

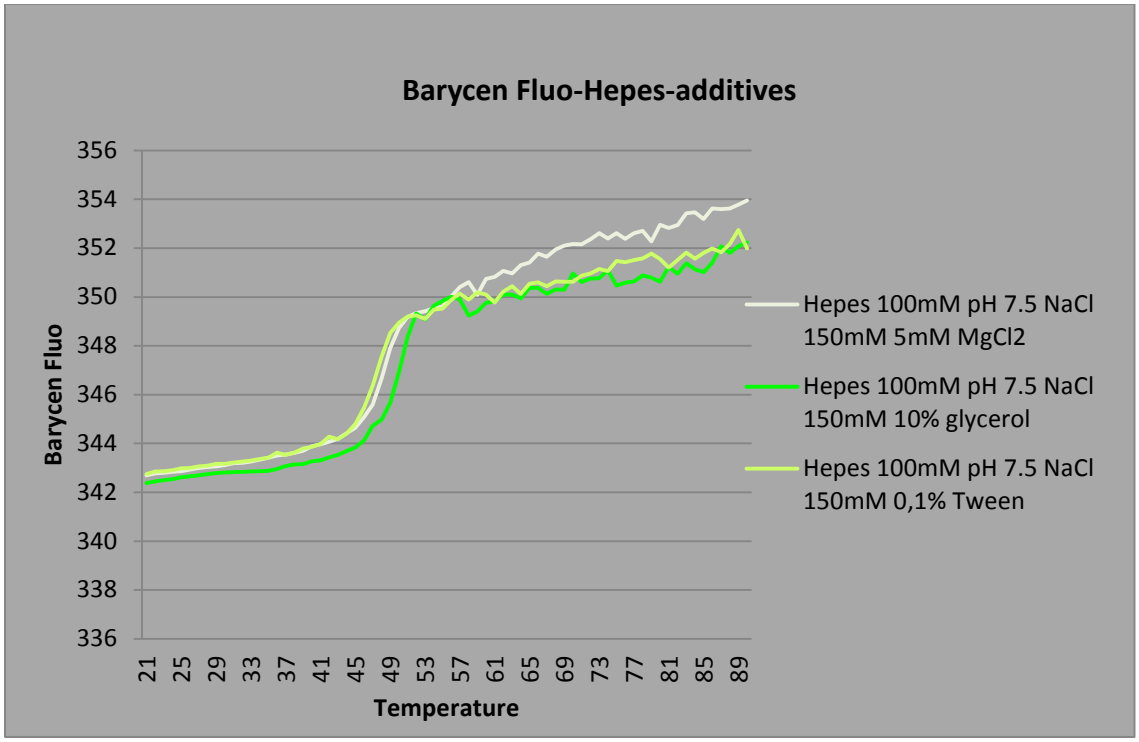
OPTIM



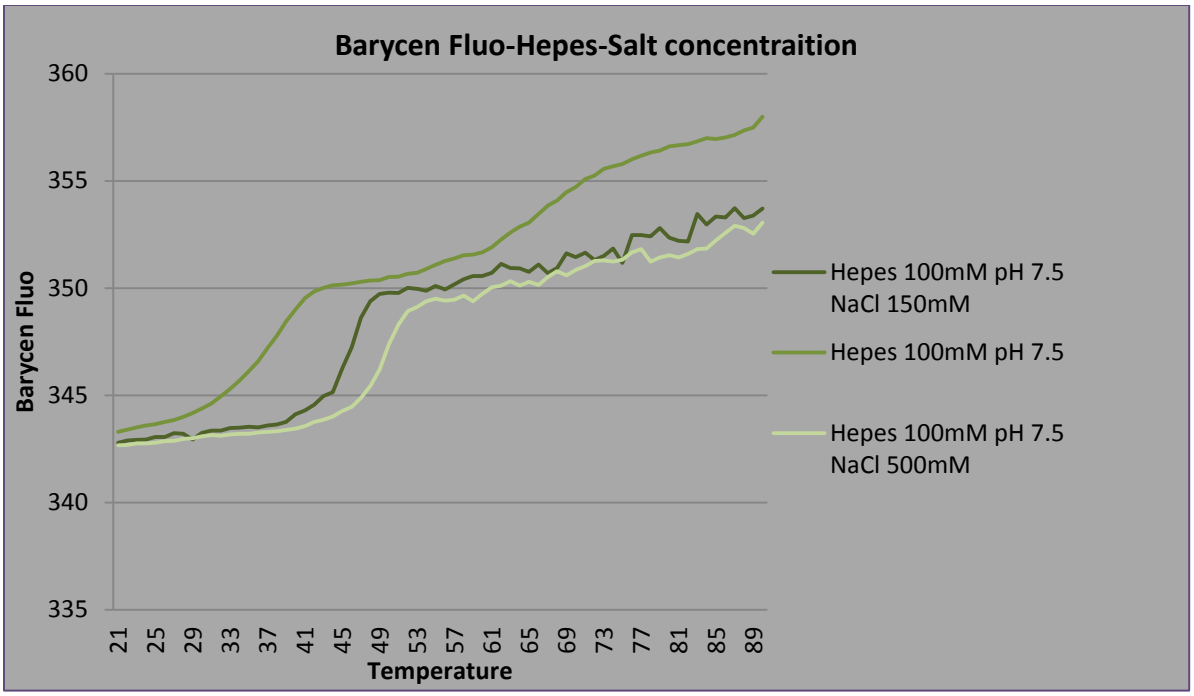
Suppl. Figure1: The FluoIntensity of the protein were tested in all the buffers. The buffers with the lowest pH values were excluded. As it is indicated on the diagram, both buffers provoke the precipitation of the protein beginning at low temperature. Also, the buffer with the highest pH value was excluded. On the figure, it is shown that the protein into CHES pH 9.5, starts precipitating very early.



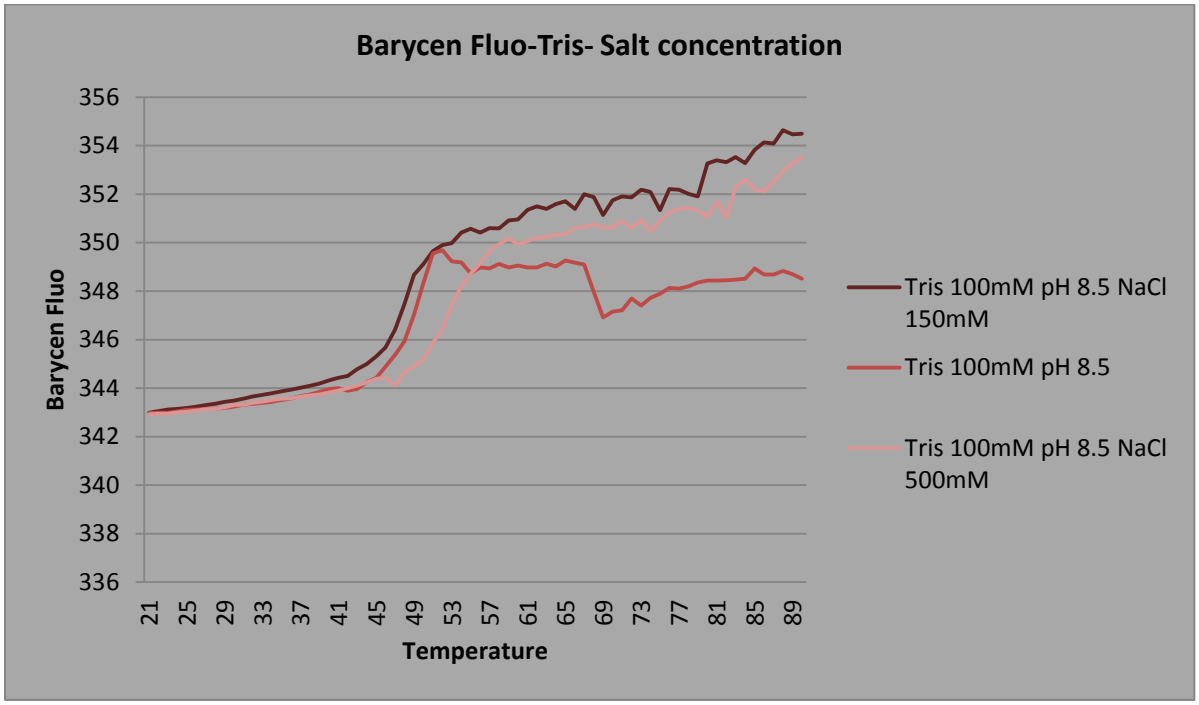
Suppl. Fig.2: The chart depicts the Scatter 266 measurements of the protein (1mg/mL) and how MgCl2,glycerol and Tween effect when a range of temperature is applied.



Suppl. Fig.3: The chart depicts the Barycen Fluo measurements of the protein (1mg/mL) and how MgCl2,glycerol and Tween effect when a range of temperature is applied.

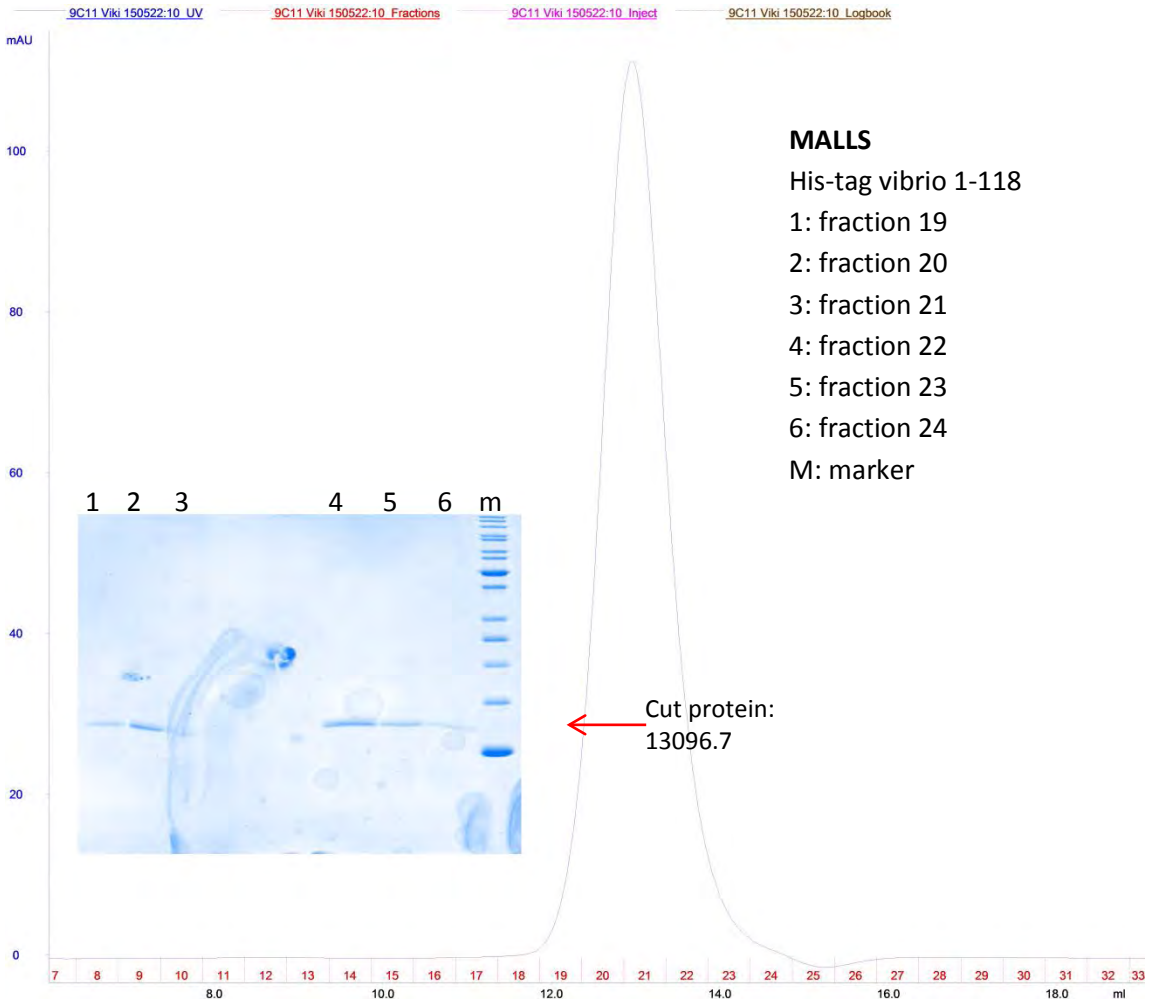


Suppl. Fig.4: The chart depicts the Barycen Fluo measurements of the protein (1mg/mL) into Hepes 7.5 and the effect of the salt concentration when a range of temperature is applied.



Suppl. Fig.5: The chart depicts the Barycen Fluo measurements of the protein (1mg/mL) into Tris pH 8.5 and the effect of the salt concentration when a range of temperature is applied

MALLS (see also fig. 29)



Suppl. Fig.6: The chromatogram above corresponds to a MALLS analysis coupled with a SEC method. The SDS-PAGE gel was carried out in order to be confirmed that the fractions contributing the peak include the protein of interest and no other impurities.

8. References

1. Anantharaman, V. & Aravind, L., 2003. Evolutionary history, structural features and biochemical diversity of the NlpC/P60 superfamily of enzymes. *Genome biology*, 4(2), p.R11.
2. Anon, O ptim 1000 ®.
3. GE_Healthcare, 2009. Recombinant Protein Purification Handbook.
4. GeHealthcare, Ion-exchange chromatography & Chromatofocusing. Principles and Methods. *Chromatographia*, 13(2).
5. Healthcare, G.E. & Sciences, L., Size Exclusion Chromatography For more information refer to www.gelifesciences.com/handbooks.
6. Hughes, P.J. & Stanway, G., 2000. The 2A proteins of three diverse picornaviruses are related to each other and to the H-rev107 family of proteins involved in the control of cell proliferation. *Journal of General Virology*, 81(1), pp.201–207.
7. Hyypiä, T. et al., 1992. A distinct picornavirus group identified by sequence analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 89(18), pp.8847–8851.
8. Luna-Vargas, M.P. a et al., 2011. Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases. *Journal of Structural Biology*, 175(2), pp.113–119. Available at: http://dx.doi.org/10.1016/j.jsb.2011.03.017.
9. Manuscript, A., 2012. NIH Public Access. *Changes*, 29(6), pp.997–1003.
10. Oberste, M.S. et al., 2001. Molecular identification of new picornaviruses and characterization of a proposed enterovirus 73 serotype. *Journal of General Virology*, 82(2), pp.409–416.
11. Palmenberg, a, Neubauer, D. & Skern, T., 2010. Genome Organization and Encoded Proteins. *The Picornaviruses*, pp.3–17.
12. Pang, X.Y. et al., 2012. Structure/function relationships of adipose phospholipase A2 containing a Cys-His-His catalytic triad. *Journal of Biological Chemistry*, 287(42), pp.35260–35274.
13. Patton, J.T. & Spencer, E., 2000. Genome replication and packaging of segmented double-stranded RNA viruses. *Virology*, 277(2), pp.217–225.
14. Wang, C.-H. et al., 2014. Phospholipase A/Acyltransferase enzyme activity of H-rev107 inhibits the H-RAS signaling pathway. *Journal of biomedical science*, 21(1), p.36. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4012743&tool=pmcentrez&rendertype=abstract.
15. Yamashita, T. et al., 1998. Complete nucleotide sequence and genetic organization of Aichi virus, a distinct member of the Picornaviridae associated with acute gastroenteritis in humans. *Journal of virology*, 72(10), pp.8408–8412.

Acknowledgements

I would like to express my greatest gratitude to the people who have helped & supported me throughout my project. I am grateful to Tassos Perrakis and of course, to Eleonore von Castelmur for their continuous support for the project, from initial advice & contacts in the early stages of conceptual inception & through ongoing advice & encouragement to this day.

I am equally thankful to Dimitris Leonidas for his support. He had been very kind and patient while correcting my doubts.

A special thank of mine goes to Tatjana Heidebrecht and of course to all the members of the lab who helped me in completing the project & they exchanged their interesting ideas, thoughts & made this project easy and accurate.