Veera Soikkeli

Histone Acetyltransferase Complexes — expression and purification of ATAC HAT

Metropolia University of Applied Sciences

Bachelor of Laboratory Sciences (UAS)

Laboratory Sciences Degree Programme

Bachelor's Thesis

15.10.2015

Metropolia

Author(s) Title Number of Pages Date	Veera Soikkeli Histone Acetyltransferase Complexes — expression and purification of ATAC HAT 46 pages + 6 appendices 15th of October 2015
Degree	Bachelor of Laboratory Sciences (UAS)
Degree Programme	Laboratory Sciences
Instructor(s)	Imre Berger, professor Tiina Soininen, phil. lic, Senior Lecturer

The Bachelor's thesis project was conducted to gather new information about a histone acetyltransferase (HAT) complex named Ada-Two-A-Containing complex (ATAC). It is a transcriptional activator that modifies the chromatin structure so that it is more available for transcription. The purpose for the work was to produce and purify the HAT module of ATAC as a prerequisite for structural analysis of ATAC HAT nucleosome complexes by native mass-spectrometry.

The laboratory work was done at the European Molecular Biology Laboratory outstation in Grenoble; France, as a part of Berger group's research. The project consists of two main sections: the expression of the three modifications of ATAC HAT protein complexes in insect cells and the purification of these complexes by immobilized metal affinity chromatography and size exclusion chromatography. The steps realised in the laboratory include a great variety of different techniques such as SDS-page, agarose gel electrophoresis, Western blotting, microscoping and fluorescence measurements.

The results obtained show that these three complexes are well producible in insect cells and the purification procedure works as it is. Nevertheless, some modifications could be made for the amounts of cultures prepared. The project provides a full protocol for the production of ATAC HAT complexes in a laboratory. The concrete results of the project are the viruses, plasmids and protein ready to use in the next steps of the scientific research. Nevertheless, the final amount of protein was less than 0.5 mg per protein which remains quite low.

Keywords	ATAC HAT, protein, nucleosome, transcription

Tekijä(t) Otsikko Sivumäärä Aika	Veera Soikkeli ATAC HAT-kompleksin ekspressio ja puhdistaminen 46 sivua + 6 liitettä 15.10.2015
Tutkinto	Laboratorioanalyytikko (AMK)
Koulutusohjelma	Laboratorioala
Ohjaaja(t)	Imre Berger, professori Tiina Soininen, FL, lehtori

Opinnäytetyön aiheena oleva proteiinikompleksi Ada-Two-A-Containing Histone Acetyltransferase, lyhyemmin ATAC HAT, vaikuttaa solun nukleosomissa DNA-juosteen transkription aktiivisuuteen. Kompleksi muokkaa juosteen rakennetta ja pakkautumista niin, että DNA-polymeraasit voivat toimia vapaammin. ATAC HAT kompleksin rakenne ei ole vielä tarkasti selvillä, joten projektin tarkoituksena oli tuottaa ja puhdistaa kyseistä proteiinia myöhempää massaspektrometrista tutkimusta varten.

Työ toteutettiin Euroopan Molekyylibiologian laboratorion (EMBL), Grenoblen yksikössä Ranskassa, osana Bergerin tutkimusryhmän projektia. Opinnäytetyö jakaantuu kahteen suurempaan osaan. Ensimmäinen on kolmen eri ATAC HAT modifikaation tuottaminen hyönteissoluissa, ja toinen kyseisten proteiinien puhdistaminen käyttämällä affiniteettiin ja kokoon perustuvia kromatografisia tekniikoita.

Proteiinin tuottaminen oli valmiin protokollan avulla yksinkertaista ja toimivaa niin ekspression kuin puhdistamisenkin osalta. Soluviljelmän kokoa tosin voisi kasvattaa, jotta proteiinia jäisi enemmän jäljelle kaikkien puhdistusvaiheiden jälkeen.

Työn tuloksena jokaista proteiinia saatiin alle 0,5 mg, joten saanto on melko alhainen. Projektin ansiosta saatiin kuitenkin tarvittavat plasmidit ja virukset ATAC HAT kompleksien tuottamiseksi sekä hieman proteiinia tuleviin tutkimuksiin. Lisäksi opinnäytetyö tarjoaa kokonaisvaltaisen protokollan kyseisten proteiinien valmistukseen alusta loppuun.

Avainsanat	ATAC HAT, proteiini, nukleosomi, transkriptio

Contents

Appendices

Appendix 1 **Plasmid maps for CL39, CL40 and CL50**

Appendix 2 **Eukaryotic Expression Facility brochure**

Appendix 3 **QIAGEN miniprep kit manual**

Appendix 4 **YFP measurement data**

Appendix 5 **SEC chromatograms of ATAC HAT complexes**

Appendix 6 **AKTA Calibration curve for S200 10/300 column**

Abbreviations

ADA2a	Transcriptional adapter 2a
ADA3	Transcriptional adapter 3
AGE	Agarose gel electrophoresis
ATAC	Ada-two-A-containing
BamHI	restriction enzyme from *Bacillus amyloliquefaciens* with a post-fixe for strain identification
BSA	bovine serum albumin
cDNA	complementary DNA
Cre	Cyclisation recombination protein, also a shortening for *causes recombination*
ctop10	chemically competent *E.coli* strain
DPA	Day after proliferation arrest
EMBL	European Molecular Biology Laboratory
GCN5	General control of amino-acid synthesis 5, one type of acetyltransferases
HAT	Histone acetyltransferase
HindIII	restriction enzyme from *Haemophilus influenzae*
IMAC	Immobilized Metal Affinity Chromatography
IPTG	Iso-propyl thiogalactoside
lacZ	Gene that results in production of β-D-galactosidase
LB	Luria-Bertani medium
LoxP	Locus of X-over P1, where P1 is a bacteriophage
O/N	over night
p300	E1A binding protein, a protein involved in cell growth and division
PAGE	Poly acrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCAF	Human p300/cbp-associated factor, one type of acetyltransferases
R&D	Research and development
SAGA	Spt-Ada-Gcn5-acetyltransferase
SDS	Sodium dodecyl sulphate
SEC	Size Exclusion Chromatography
SGF29	SAGA-associated Factor 29
VLP	Virus-like particle
WB	Western blot
YFP	Yellow fluorescent protein
ZnF	Zinc finger

1 Introduction

Proteins are central to all biological processes in the human body. They catalyse essential reactions in our cells. The nucleus of every cell contains genetic information, DNA. A key process in the nucleus is transcription, which is the synthesis of a messenger RNA (mRNA) using DNA as a template. mRNA; on the other hand, sets the template for the synthesis of an amino acid chain which gives rise to a protein in a process called translation. In this way, a protein produced is based on the DNA sequence.

A protein itself is characterised by several levels of structural organization: primary (amino-acid sequence), secondary (α-helices, β-sheets) and tertiary structure (folded molecule). Several proteins together assemble into multi-unit arrangements: quaternary structures. The three-dimensional structures of proteins and their amino acid sequence determine their function in humans and any other eukaryotic organism. For instance, the DNA polymerase enzyme builds DNA molecules from nucleotides, collagen strengthens the bones and skin and opsin makes the eye detect light. These are just a few examples of all the functions proteins are involved in. (Harvey, *et al.*, 2011, 14–23; Learn. Genetics, 2015.).

A further important concept is the nucleosome. It is the structure in the cell nucleus where the genetic information is stored, the genome. The DNA double-helix, is a very intricate, fragile structure, and the nucleosomes protect it from physical damage. The genome is packed into a compact form in the nucleosome and the nucleosomal units consisting of histone proteins and DNA form chromatin, the basic material of the chromosomes.

Briefly, the genetic information of DNA is packed compactly so that 147 base-pairs (bp) of the DNA strand is wrapped around a core histone octamer consisting of histone proteins H2A, H2B, H3 and H4, two of each. Every histone protein contains an N-terminal tail that comprises 15–35 amino acids. The tails are involved in the internucleosomal interaction. The histone tails are important targets for post-translational modifications such as acetylation of lysine residues or phosphorylation of serine and threonine residues. (An, 2007, pp.35–352.)

A great amount of research has been done to clarify the operations of nucleosomes but with varying results. The way they protect the genome and, at the same time, allow the

cells use the information of the DNA strands is not completely understood. It has been presumed that the DNA unfolds partially from around the nucleosome so that the information in the DNA can be read.

The modification of the genome activity is mostly caused by the histones. Histones are not completely like other proteins: they have long tails that reach to the neighboring nucleosomes and bind them strongly together. The nucleus of the cell excretes regulatory enzymes that modify these tails and weaken their interactions with other parts of the complex. In this way, some of the genes are more accessible to polymerases and their information is available for copying. (Goodsell, 2000.)

This bachelor's thesis was conducted in the Berger group in the European Molecular Biology Laboratory (EMBL) outstation in Grenoble, France. The project was realised to gather new information of a few multi-unit protein complexes that effect the transcription of the genome. The ones of interest were three modifications of Ada-Two-A-containing complex Histone Acetyltransferase (ATAC HAT). The main goal for the project was to express and purify these protein complexes for structural characterisation by native mass-spectrometric analysis. The characterisation will be done later by the laboratory of Frank Sobbot in the Netherlands. The analysis is used to determine the structure of ATAC HAT complex, which is not yet well known. The new mass-spectrometric method for the structural analysis can provide more information about the composure of these transcriptional activators.

2 Theoretical background

HATs (histone acetyltransferases) are essential enzymes responsible for acetylation of histone tails required in the control of the chromatin structure. In histone acetylation (Figure 1), an acetyl group is added to lysine residues in the N-terminal tail or on the surface of histone protein's nucleosomal core. Acetylated or deacetylated histones can relax or tighten, respectively, the chromatin structure, making the gene more or less available for transcription. In that way they play a significant role in eukaryotes by controlling chromatin architecture and locus-specific transcription. Hence, ATAC HAT (Ada-Two-A-Containing Histone Acetyltransferase) belongs to the compounds that make the chromatin more available for the transcription. (Verdone, *et al*., 2005; Nagy, *et al*., 2010.)

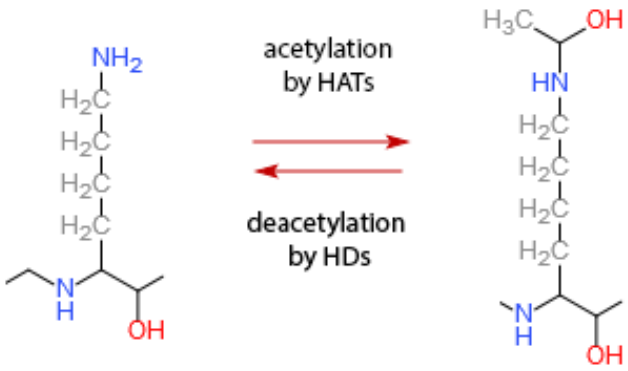


Figure 1. Acetylation of lysine residues at the N-terminus by histone acetyltranferases (HATs). This reaction removes positive charges and reduces the affinity between DNA and histones, which makes the promoter region more easily accessible for transcription factors. Histone deacetylation by histone deacetylases (HDs), on the other hand, decreases transcription on the area. (Molecular Biology Web Book)

One of the best-known HATs is GCN5 (general control of amino-acid synthesis 5), a protein that is conserved from *S.cerevisiae* to humans. The HAT module of ATAC consists either of GCN5 or PCAF (Human p300/cbp-associated factor, a homology 73 % identical with GCN5) and ADA3 (Transcriptional adapter 3), ADA2a (Transcriptional adapter 2a) and SGF29 (SAGA-associated Factor 29). The function of ADA2a and ADA3 is to regulate the HAT activity of GCN5 (Riss, 2012, p.45). The different units of ATAC complexes are shown in the Table 1.

Table 1. ATAC complexes and their components in drosophila and human (Riss, 2012).

	ATAC complex	
	dATAC *Drosophila melanogaster*	hATAC *Homo sapiens*
HAT module	dGCN5	hGCN5/hPCAF
	dADA2a	hADA2a
	dADA3	hADA3
	dSGF29	hSGF29
Others	dATAC1	hZZZ3
	dATAC2	hATAC2
	dATAC3	–
	HCF	–
	WDS	WDR5
	D12	YEATS2
	CHRAC14	–
	NC2β	NC2β
	CG10238	MBIP

2.1 Effects of histone acetyltransferases

The deletion of the SGF29 unit results in decreased levels of histone H3K9, K14 and K23 acetylation in both yeast and human cells. Yeast and human GCN5 acetylate mainly histone H3K14 *in vitro*. However, ATAC from *Drosophila* acetylates also histone H4. The activity with histone H4 has been interpreted to be caused by the presence of the second HAT unit in ATAC complex, ATAC2. (Schram, *et al*., 2013; Nagy, *et al*., 2010.)

Both GCN5 and PCAF are transcriptional activators. A verified trace between H3 and H4 acetylation and p300-mediated (E1A binding protein-mediated) transcription has been established. HATs neutralize the charge of histone tails. When this happens, the tails and the negatively charged DNA backbone get more separated from each other and the chromatin is then more available for transcription. Histone acetylation also seems to create a signal for the binding of the bromodomain which has been found in GCN5, PCAF and p300, for example. The bromodomain is often found in the proteins that regulate and modify the chromatin. (An, 2007, pp.354–357; Riss, 2012, p.44.)

The ADA unit is in form of ADA2a in ATAC (Figure 2). It consists of a zinc finger (ZnF) and two other domains, SANT (**S**wi3, **A**da2, **N**-Cor and **T**FIIIB) and SWIRM (**Swi**3, **R**CS8 and **M**OIRA) (Riss, 2012, p.64). The zinc finger is proteins' means to form stable and compact structures. They are particularly useful in small molecules when there are not enough amino acids to form many hydrogen bonds or charge pairs. When two cysteine molecules and two histidine molecules are close together, they can grab a zinc ion and fold tightly around it. With these structures, just 20–30 amino acids are enough to form a strong structure. Also, the zinc fingers can bind to a DNA strand controlling the transcription. (Goodsell, 2007.)

Figure 2. Protein sequence for hADA2a (Riss, 2012, p.64)

2.2 Histone acetyltransferases in humans

ATAC HAT is not the only HAT compound found in humans. Another similar complex is SAGA HAT (Spt-Ada-Gcn5-acetyltransferase Histone Acetyltransferase) that resembles ATAC HAT in its structure and function. They are composed of same kind of subunits; the only difference is that in SAGA ADA2a is replaced with ADA2b and the ATAC2 unit is missing (Figure 3).

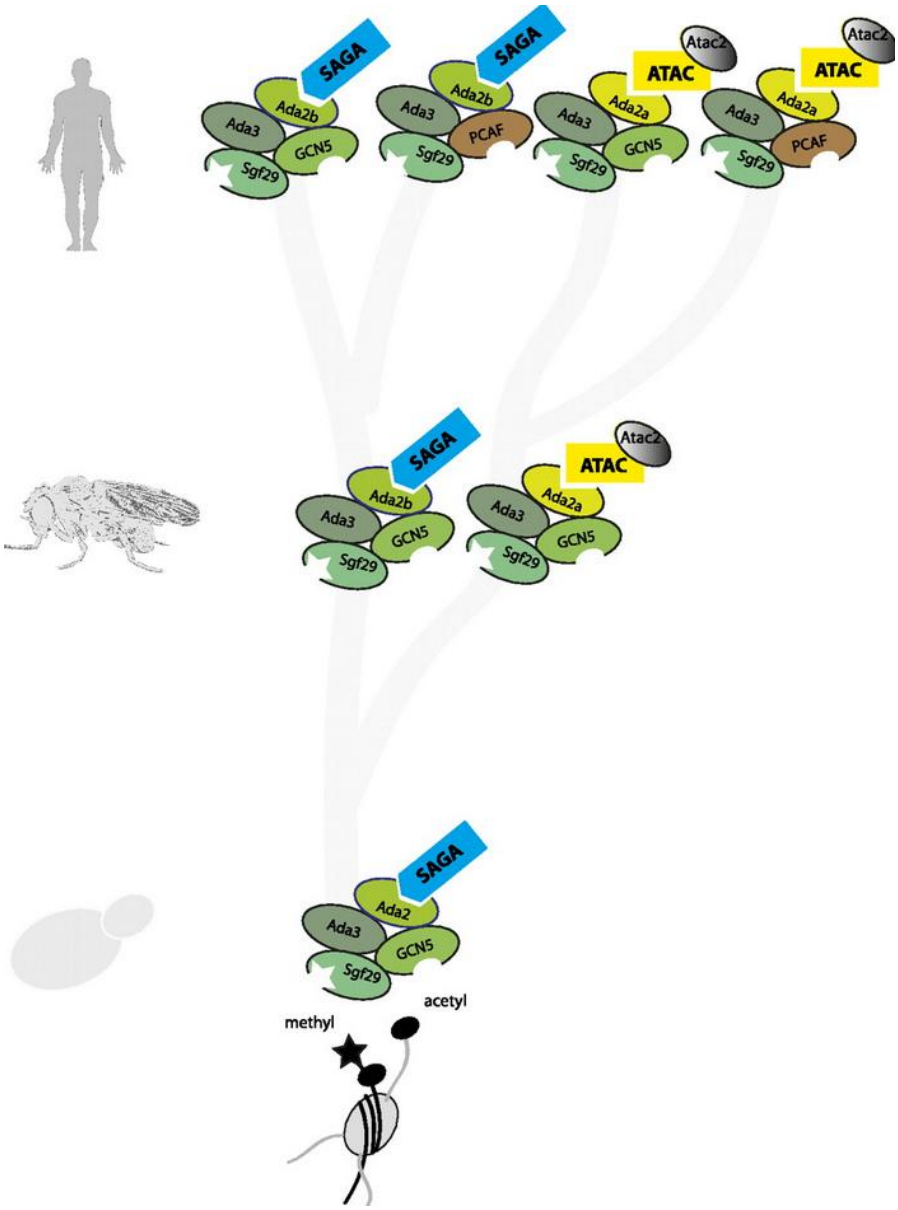


Figure 3. Origins of complexes SAGA and ATAC (Spedale, *et al*., 2012).

The fact that ATAC is present only in multicellular eukaryotes shows that it was developed later during evolution than SAGA. The function of ATAC in mammalian cells is not completely clear at present. It is indistinct why the two, almost same-structured complexes exist. Discussion has arisen if the two types of GCN5 containing complexes have specific functions compared to each other. (Guelman, *et al*., p.208.) At least one difference has already been found: Krebs *et al*. (2011) have showed that ATAC HAT complex is both a gene enhancer and promoter when SAGA HAT complex only binds to promoters (Figure 4).

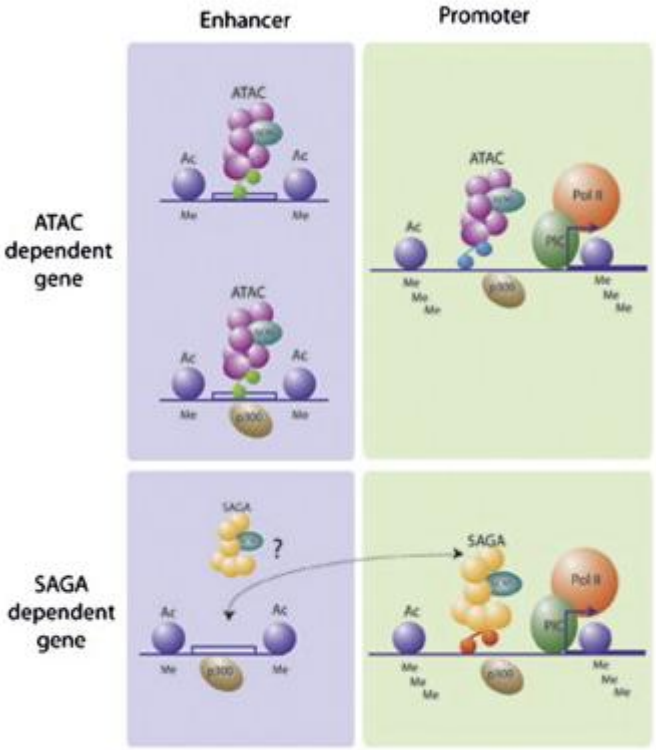


Figure 4. ATAC and SAGA binding to DNA (Krebs, *et al*. 2011).

Nagy, *et al*. (2010) have been able to show that there is another difference between the function on SAGA and ATAC. Drosophila was exposed to a toxic agent TPA (12-O-tetra-decanoyl phorbol-13-acetate) which activates the signal canal for PKC kinase (protein kinase C) and also causes the immediate expression of the genes. It was noted that the dATAC complex was recruited immediately to the loci actively transcribed. With dSAGA, this did not happen. Without stimuli the signal of ATAC and SAGA on the chromosomes was weak. The same experiment was executed with human HeLa cells with according results.

To summarise the basics of ATAC HAT, acetylation modifies the chromatin fibre and its physical and chemical properties. That not being the only quality, the process also provides interaction sites for multitude of binding proteins. Histone acetylation has attracted interest from a medical research: drugs inhibiting the histone deacetylases from working have been shown to be effective against particular types of lymphomas. (Spedale, *et al*., 2012.)

2.3 Baculoviruses

Baculoviridae is a virus family the members of which contain double-stranded, circular DNA. It is both occlusion derived and a budding virus that infects mainly anthropods, especially insects (Figure 5). The viruses can usually be hosted by one species only. One of the most important qualities of the baculoviruses for research is their ability to produce polyhedra. They are large particles that appear in the nuclei of infected cells at the end of the infectious cycle, during viral infections.

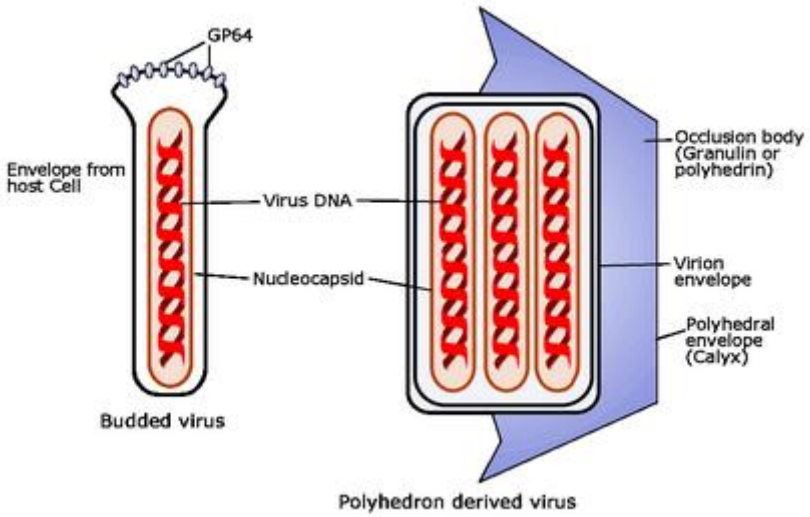


Figure 5. Structure of a baculovirus (Kalmakoff, J. Microbiologybytes).

Virus's ability to produce large amount of polyhedra is one of the factors that made them vectors for foreign protein production. An advantage for this was that baculoviruses do not need polyhedrin to replicate in cultured insect cells. One could simply include the foreign protein-encoding DNA sequence to the viral DNA and harness the recombinant, baculovirus-infected insect cells for protein production. Infection leads to very abundant transcription of foreign cDNA (complementary DNA) during the late phase of infection. The production efficiency is, nevertheless, dependable of the protein that is being produced.

Another advantage with baculoviruses is their eukaryotic protein processing capability. (Jarvis, 2009). Baculovirus-infected insect cells have been proved to be efficient producers of multiprotein subunit complexes like the ATAC miniHAT of interest. Eukaryotic protein complexes can contain over ten subunits, each with a total size of several hundred kDa. Thereby it can be noted that the capacity of *Escherichia coli* is not sufficient for their expression. With baculoviruses, large DNA insertions in the double stranded viral genome can be effectuated. (Berger, *et al*., 2004.)

Even if *E.coli* might be able to produce large protein molecules, they are not processed further properly. For example, the proteins produced in *E.coli* are not glycosylated, in other words sugar groups are not attached to them after translation. The proteins are not folded properly because the bacteria cannot synthesize the disulphide bonds that are present in eukaryotic proteins. Usually this means that the protein is insoluble and inactive. These post-translational modifications are, however, possible with insect cells. (Brown, 2010, p.236.)

The developments in the baculoviral research have noted that the viruses can generate virus-like particles (VLPs) made of multiple virion components that have huge possibilities as vaccine candidates. They do not contain the genetic information of viruses but only the viral envelope. Previously, multiple protein complexes of polio, papilloma viruses and hepatitis C, for example, have been produced in baculovirus-infected insect cells. (Jarvis, 2009.)

2.4 Insect cell cultures and protein expression

The technique, also called MultiBac, involves baculovirus infected insect cells that start to produce certain protein after transfection. The DNA sequence encoding this protein is inserted to the baculoviral genome.

Insect cells can produce bigger complexes than *E.coli,* and in living creatures proteins usually work in such units rather than alone. Some proteins are also produced in low quantities or only for a short amount of time in their native environment which makes their analysis demanding. With MultiBac system large amounts of protein can be produced for the R&D (Research and Development) purposes in the fields of structure, medical development and bioengineering, for example.

MultiBac technique needs certain things to work: an acceptor and a donor vector. Every acceptor has two Tn7 sites as well as one LoxP (Locus of X-over P1) site.

One or more donor vectors can be recombined to an acceptor via their LoxP sites (Graig and Berger, 2011). The site was first discovered in bacteriophages, it consists of two perfectly inverted repeats separated by an 8-bp spacer. The loxP site recombination is catalysed by a single phage-encoded protein, Cre (Causes recombination/cyclisation recombination protein). Firstly, the LoxP site is saturated with two Cre molecules. The Cre-lox complex is then united with a second LoxP site on the same or another molecule and the strands are exchanged. (Sambrook and Russell, 2001, ch.4.82.) After this, the acceptor contains all the coding sites of both plasmids combined. The mechanism is presented in the Figure 6.

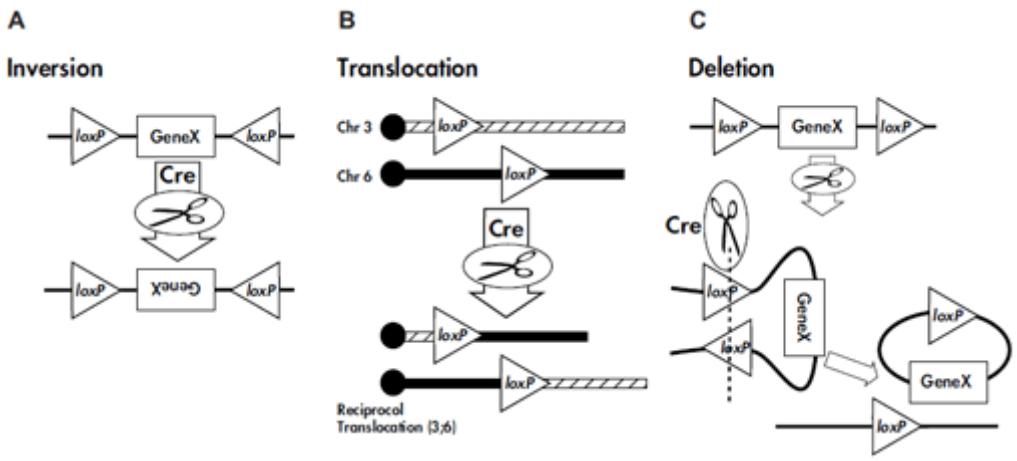


Figure 6. Types of Cre-Lox recombination (The Jackson Laboratory, 2015).

Via the Tn7 (Transposon 7) sites plasmids can be integrated to the baculoviral genome. It is a site that can move between the sites that lack homology. The transposases enable the movement of Tn7, and they are self-encoded by this site. The Tn7 site is composed of Tn7-Left (Tn7-L) and Tn7-Right (Tn7-R) sites and the sequences coding five different transposases. Approximately 150 bp sized Tn7-L and 90 bp Tn7-R make sure that the inserted fragments are oriented preferentially. Transposases TnsA and TnsB recognise the ends of the transposon. They cut the Tn7 from the donor backbone and join the ends to the target DNA. Only the 3' ends are joined and 5' ends stay flanking. These gaps are later repaired to form duplicated 5-bp target sites. The process is presented in the Figure 7. (Peters and Craig, 2001.)

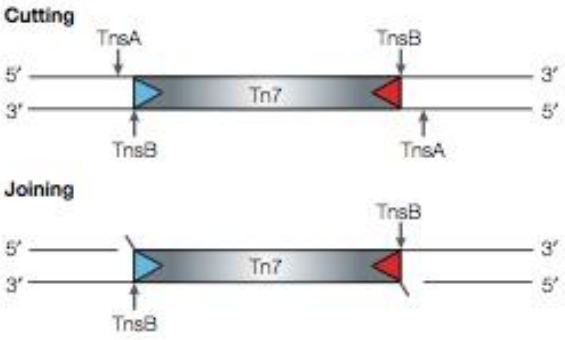


Figure 7. Tn7 transposition (Peters and Craig, 2001).

The cells used in the Multibac technique contain all the qualities described before. These DH10EMBacY cells host the baculoviral genome that is specifically engineered for the technique. In addition, there is an YFP (yellow fluorescent protein) coding sequence integrated to the genome that provides a simple factor for detecting the protein production during cell culturing. The DH10EMBacY cells also contain a helper plasmid for Tn7 transposon enzyme. The virus is both protease and chitinase deficient, the cell viability is increased and the proteolysis is reduced. The complete genome map is provided in the Figure 8. (Craig and Berger, 2011.)

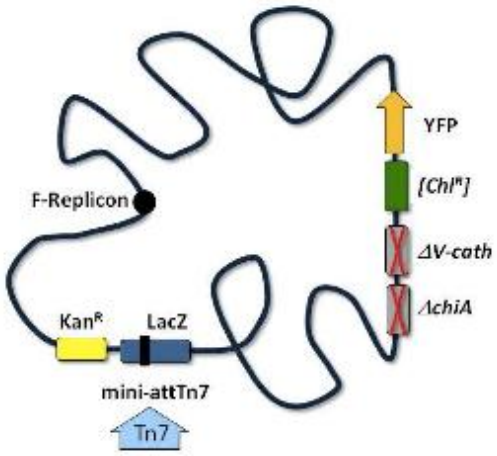


Figure 8. EMBacY baculoviral genome (Craig and Berger, 2011).

Bacmid containing colonies are determined by blue/white screening which is possible due to the lacZ peptide. The location of Tn7 site is within the lacZ gene. If the plasmid DNA is inserted to the viral genome, the operation of lacZ stops and the colonies appear as white. If insertion does not occur, the lacZ continues to produce β-galactosidase, which hydrolyses Bluo-Gal on the plates and produces a blue compound (Juers, *et al*., 2012). Bluo-Gal is an alternative substrate for β-galactosidase that forms darker blue colour than the traditionally used X-gal (Life Technologies). IPTG (iso-propyl thiogalactoside) is an inducer used to trigger the transcription of a lac operon.

The insect cells used in the Multibac technique are Sf21 cells. They are originally from the ovaries of the Fall Army worm, *Spodoptera frugiperda*, which is a moth species. Aseptic work routines are extremely important with insect cell cultures. The cells are grown in a medium that does not contain antibiotics and the optimal temperature for their growth is 27 °C. The conditions make the cultures ideal for yeasts, fungi and other bacteria as well. All the steps included to the cell culturing are effectuated in a laminar hood with proper equipment. UV light is used inside the hood before and after every work to kill any cells, viruses or bacteria that might have got onto the laminar hood surfaces.

3 Materials and methods

The methods in this thesis project are the ones routinely used in a protein laboratory. The project methodology can be divided in three sections: plasmid verification, protein expression in insect cells and purification by immobilized metal affinity chromatography (IMAC) and size exclusion chromatography (SEC). The essential reagents and recipes are listed in brackets during the work description.

3.1 Plasmid verification

The ATAC HAT plasmids engineered by Simon via cre-lox recombination (Table 2) existed from previous experiments. The constructs are presented in the Appendix 1. The difference between ATAC HAT and ATAC miniHAT of interest is the SGF29 group: in ATAC miniHAT the unit is not present (Simon Trowitzsch, personal communication).

Table 2. Plasmid cre-loxes used in the experiments

Name and the plasmid composition	**Produced complex**	Antibiotic resistances within the plasmid sequence
CL39 [pFL_HT-hGCN5 x pIDC_hADA2a x pIDK_mADA3 x pIDS_hSGF29]	**ATAC HAT**, full-length protein	Amp/(Gent)/Cam/Kan/Strep
CL40 [pFL_HT-GCN5$^{481–837}$ x pIDC_ADA2a x pIDK_ADA3$^{85–432}$]	**ATAC miniHAT**, truncated GCN5 and ADA3 units and no SGF29 unit	Amp/(Gent)/Cam/Kan
CL50 [pFL_HT-GCN5$^{481–837}$ x pIDK_mADA3 x pIDS_SGF29 x pIDC_ADA2a]	**ΔN-GCN5 ATAC HAT**, protein derived from limited proteolysis	Amp/(Gent)/Cam/Kan/Strep

The quality of the plasmids had to be checked before proceeding further. The plasmids were digested first with restriction enzyme *BamHI* HF (New England Biolabs) to verify their size and quality. As a control, all the same reagents, except for the restriction enzyme, were pipetted and incubated as the sample. The profile obtained was used again after the amplification.

The quantity of the plasmid DNA was amplified using chemically competent *E.coli* Top10 cells. Thus, 1 µg of plasmid and 50 µl of *E.coli* cells were mixed together and then incubated on ice for 15 min. The cells were transformed via heat shock at 42 °C for 45 s, and the samples were incubated on ice for 2 min. Next, 400 µl of LB (Luria-Bertani medium) broth was added, and the tubes were incubated at 37 °C for 2 h. After that, 50 µl of the mixture was plated on agar containing chloramphenicol. The same triangle without flaming was used for streaking of second plates, in case there would be too many colonies on the first plate. Agar plates were incubated O/N (overnight) at 37 °C.

The following morning, bacterial colonies could be seen. Two clones per construct (A and B) from the second plates were selected and they were transferred into separate culture flasks containing 25 ml of LB broth. The flasks were moved to a shaking incubator at 37 °C to amplify the cells.

The DNA obtained was extracted and purified with Qiagen MiniPrep kit. The accompanying kit protocol was followed (Appendix 3) except that the volumes were doubled for B1, B2 and N3 and the last elution step was repeated twice for better yield.

After the DNA extraction, a restriction digestion analysis by AGE (agarose gel electrophoresis) was run to verify that the amplification was successful. Digested DNA samples with loading buffer (6X) were loaded on agarose gel containing ethidium bromide (around 100 ng of DNA per lane). 1 kb size marker (New England Biolabs, 1 kb DNA ladder) and uncut plasmids as controls were applied and the run was performed with 110 V approximately for 2 h.

3.2 Insect cell culturing and protein expression

To maintain the cell stock, Sf21 insect cells were grown in a serum free medium (Gibco, Life Technologies, Sf-900 II SFM 1X). The healthy range for the cells is 0.5–2.0 · 10^6 cells/ml. The cultures were split every day because of the cell division once every 24 h.

When it comes to the cell culturing, the EMBL's EEF protocol (Appendix 2) was followed. The transfection was performed using the X-tremeGene HP DNA Transfection Reagent reagent (Roche) and the efficacy of the transfection was monitored with a fluorescence spectrometer every 24 hours.

The quality of work was checked with SDS-page paired with Coomassie brilliant blue staining and Western blot with Monoclonal Anti-polyHistidine Phosphatase conjugated Antibody (Sigma-Aldrich, lot. 014M4787V). After the antibody solution, the membrane was soaked in BSA (bovine serum albumin)/milk powder blocking solution and then in the BCIP/NBT (5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium) phosphatase photoreaction solution, which makes the colorimetric detection of the proteins possible.

3.3 Purification

The small-scale production provided sufficient amount of protein to be purified. The function of the protocol was first tested with the small-scale cell pellets and later with larger amount of cells.

The components of the buffers (Table 3) are chosen for certain reasons. Imidazole is added to minimize the unspecific binding of protein. The high salt concentration makes the protein soluble; in lower salt the protein seems to stay insoluble, according to the previous experiments. Tris-HCl performs as a buffer to stabilise the pH. All pHs were adjusted to 7.6 at 4 °C, after the addition of all the chemicals. The protease inhibitors were added as a cocktail of 100X leupeptin, 100X pepstatin and a protease inhibitor tablet (cOmplete EDTA-free, Roche).

Table 3. Purification buffers

Name	Components
Lysis buffer	50 mM Tris-HCl, 500 mM NaCl, 10 mM imidazole, 10 % (v/v) glycerol, protease inhibitors
Wash buffer	50 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole, 5 % (v/v) glycerol, protease inhibitors
Elution buffer	50 mM Tris-HCl, 500 mM NaCl, 200 mM imidazole, 5 % (v/v) glycerol
Dialysis buffer	25 mM Tris-HCl, 500 mM NaCl, 5 % (v/v) glycerol, protease inhibitors

Firstly, the cell pellets were taken out of -80 °C freezer and put on ice. Then, 40 ml of lysis buffer was added, and the cells were thawed. The cells were lysed by sonicator with a macro tip (60 %, 1 min program, 5 s ON and 10 s OFF) and an SNP (supernatant and pellet) sample was taken. The cell lysate was centrifuged in a JA 25.50 rotor (Beckman Coulter) at 20 000 rpm for 45 min at 4 °C, and an SN (supernatant) sample was taken.

As much as 1 ml of the Ni^{2+}-NTA resin was equilibrated with 2 x 30 ml of PBS (phosphate-buffered saline). The IMAC was performed as batch purification so that the resin was in a 50 ml falcon tube and the supernatant from the cell lysis was added on it. The tubes were incubated for 3 hours at 4 °C to bind the his-tagged protein to the resin.

The tube was centrifuged gently (800 rpm, 5 min) to separate the resin and the supernatant, the flow through (FT). A FT sample was taken to see if the protein stayed attached to the resin or if some of it remained in the supernatant. The flow through fraction was transferred to a fresh 50 ml falcon tube and stored at 4 °C. It could be passed on the resin again if a good amount of protein remained in the flow through.

The resin was washed with 50 ml of wash buffer followed by incubation for 20 min at 4 °C. The tube was centrifuged again (800 rpm, 5 min), and a W (wash) sample was taken from the supernatant. The whole wash fraction was stored at 4 °C in case the protein eluted with the washing buffer. The resin was suspended with 1 ml of washing buffer and transferred to a 2 ml eppendorf tube. The buffer was separated from the resin with 5 min centrifugation at 1000 rpm and the supernatant was discarded.

The protein was eluted six times with 1 ml of elution buffer. First, the elution buffer was added, followed by a 20-minute incubation at 4 °C. The tubes were then centrifuged with an eppendorf centrifuge, 5 min at 13 200 rpm. The supernatant was transferred to a fresh tube and centrifuged again for 10 min at 13 200 rpm to remove small amounts of resin. Finally, the supernatant was transferred to a fresh tube and stored at 4 °C until SDS-page analysis. The elution steps were repeated six times as previously.

All the samples were loaded on an 11 % Tris-Tricine gel, and the fractions containing protein of interest were pooled.

The concentration of the sample was measured, and a PRE (before TEV protease) sample was taken aside. 1:50 (mass ratio) volume of TEV protease (1 mg/ml) was added to the pooled fractions, and it was dialysed against dialysis buffer O/N. A POST (after TEV protease) sample was taken the following morning to note the cleavage of his-tag.

The sample was concentrated using ultrafiltration concentrator (30,000 MW cut-off filter, Merck Millipore) to a final volume of approximately 550 µl. The ÄKTApurifier 10 (GE Healthcare) with an automatic fraction collector was used for the SEC analysis. Method information is presented in the Table 4.

Table 4. Parameters of the SEC analysis with AKTApurifier Superdex 200 10/300

Time	Function
0.00	Base volume
0.00	Flow 0.500 {ml/min}
0.00	Alarm_Pressure Enabled 1.50 {MPa} 0.00 {MPa}
0.00	AutoZeroUV
2.00	InjectionValve Inject
2.00	AutoZeroUV
6.00	InjectionValve Load
6.00	Fractionation_900 0.400 {ml/min}
30.00	Fractionation_Stop_900
38.00	End_Method

Finally, 500 µl of the concentrated sample was injected onto a Superdex 200 10/300 column (GE Healthcare) previously equilibrated with dialysis buffer (table 3) via 1 ml loop. The column is ideal for separating protein molecules between 10 000 and 60 000 kDa. Rest of the concentrated sample was kept as an IN (input) sample.

The fractions containing protein were selected based on the spectrum. All the fractions from the peak volumes were analysed on an 11 % Tris-Tricine gel including the PRE, POST and IN samples. Again, the protein-containing fractions were pooled and the sample was concentrated with a 30 MWCO cut-off filter. The product was divided to 50 µl aliquots, and they were flash frozen in liquid nitrogen.

4 Results and discussion

4.1 Plasmid verification

Top10 *E.coli* cells were selected as hosts for the plasmids because of their good transformation efficiency ($1 \cdot 10^9$ cfu/µg plasmid DNA) and their stable production of replicates.

The plasmids were checked in case of degradation with agarose gel. The results were understandable, even though the bands were in a different place and there was a different amount of them on the gel than on our reference gel. Some of the bands were higher in size, but that might tell that some fragments were not digested after all. That could also explain the missing bands. The same also works in to the other direction: too many bands mean more restriction sites than expected. To conclude, the plasmid sequence might be a little different than the one determined by sequencing.

The amplification of clones in LB broth containing chloramphenicol was not successful with one CL39B clone. The colony selected might have been a contaminant, and this sample was discarded. The other samples (CL39A, CL40A, CL40B, CL50A and CL50B) proceeded to DNA extraction. The extraction succeeded well, and the concentrations of the samples can be found in the Table 5. The concentration values were measured with a NanoDrop spectrophotometer.

Table 5. Sample concentrations after DNA extraction. The letters A and B describe the DNA extracted from two clones of the same plasmid-containing bacteria.

Sample	Concentration
CL39A	577.6 ng/µl
CL40A	512.0 ng/µl
CL40B	520.0 ng/µl
CL50A	540.2 ng/µl
CL50B	440.0 ng/µl

Another digestion with *BamHI* HF was performed after this part. The samples were diluted and the amount pipetted to a single gel well was approximately 100 ng of DNA. AGE showed that the plasmid received from another researcher had been successfully amplified. The bands corresponded very well the ones seen in the previous run. The results are in the Figure 9.

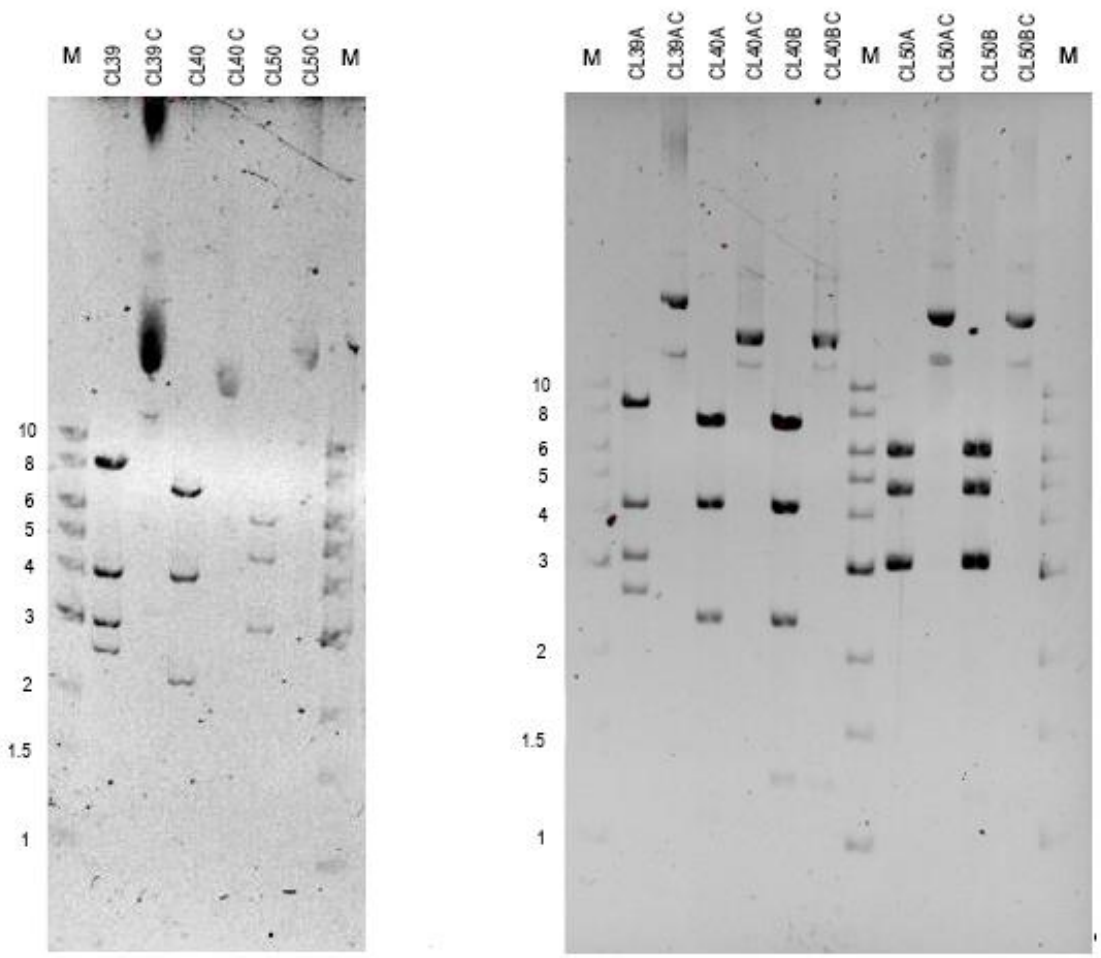


Figure 9. Restriction digestion analysis of plasmids CL39, CL40 and CL50. Old plasmids on the left and amplified ones on the right. M = 1 kb marker (New England Biolabs), C = uncut control plasmid

4.2 Insect cell culturing

4.2.1 Blue-white screening

The blue-white screening was satisfying and the colonies were selected. The biggest problems at this point were that the plates were quite full of blue colonies but they had just few white ones amongst them. The picking of white colonies was done very carefully, not touching the surrounding blue colonies.

After streaking the white colonies and incubation, the genotype confirmation could be made. Some colonies that were thought to be white were actually blue. Sometimes the colour cannot be interpreted very certainly on the first blue-white plate because the colonies might just have very pale blue tone or they might be blue only from the middle. The other explanation could be that regardless of all the carefulness, the inoculation loop might have touched a blue colony at the picking step.

With the construct CL40, not a sufficient number of white colonies could be found. The transformation was done again, and it was found mandatory that the addition of a plasmid was done just after taking the DH10EMBacY cells out of -80 °C, without letting the cells thaw. With this condition much more white colonies appeared. Maybe the colder conditions make the plasmid stick to the cell membrane stronger which provides more efficient transformation. The rest of the protocol was followed as in the first two paragraphs.

4.2.2 Transfection

The quality of the DNA was checked with AGE before transfection. At the point of transfection, the cells were examined under a microscope. Normal cells appear small, round and symmetrical and when they are transfected by the baculoviral genome they grow bigger, a little asymmetrical. This was noticed after 72 h of transfection. Some of the cells appear sausage-shaped and some just bigger in diameter than the cell control. The shape and size differences can be noted in the Figure 10.

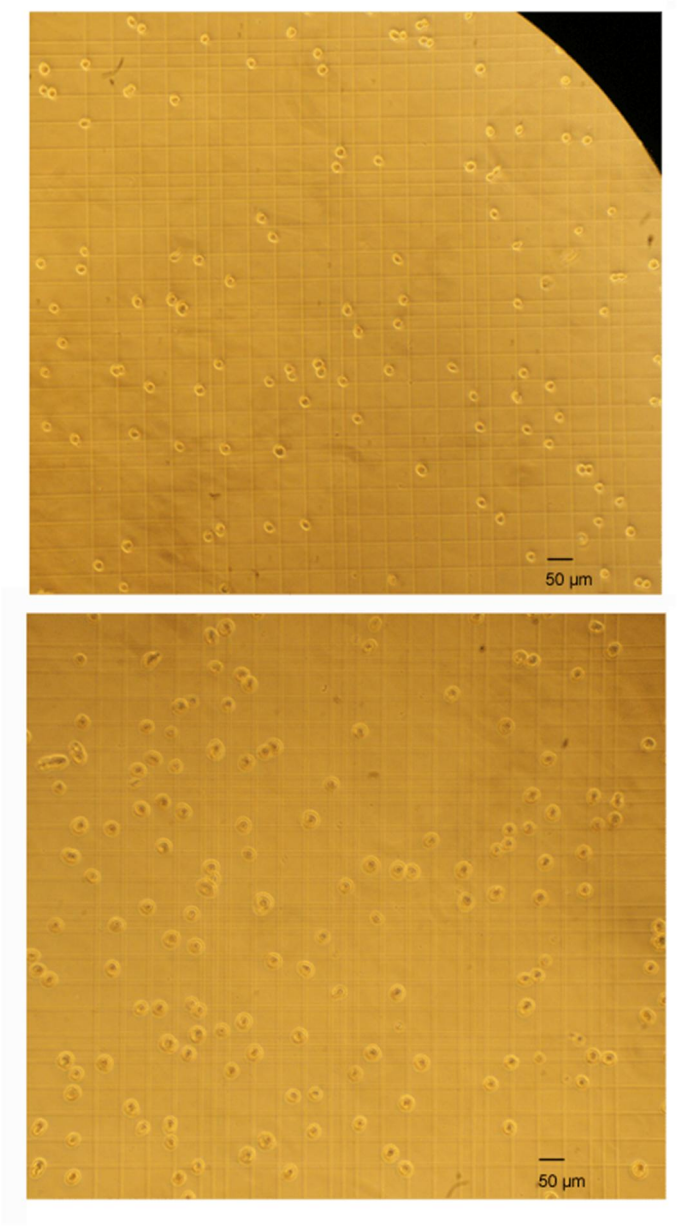


Figure 10. Pictures of uninfected cells compared to the infected ones through a 10X objective. The size difference can be noted and the shape variation can especially be seen in the left corner of the infected cell picture.

When the 6-well plate cells were harvested, the YFP fluorescence values for the cells were low. In this thesis project, the concentration of the bacmid was low (around 100 ng/µl) and it is also the reason why the YFP values measured later told the virus was weak. The virus is used to infect cell culture flasks at the next step of the protocol. The weak virus makes the transfection less efficient, but all cells get infected eventually.

By monitoring the YFP values, the highpoint can be noted, and the cells harvested at the right moment; thus, the virus concentration does not condemn the experiment.

The virus amplification in cell culture flasks was done at the same time on parallel. The cells in these flasks were count done every day and the values told when the DPA (day after proliferation arrest) was reached. Under the microscope all the cells looked infected and bigger than normal cells. After the DPA, the cells were still counted every day to obtain the same amount of cells for YFP measurement every time.

4.2.3 Spectrometric measurement

The YFP values grew steadily after DPA. The maximum obtained was around 28 000. The charts for the YFP data are in the Figures 11, 12 and 13 and the full results in more detail are presented in the Appendix 4.

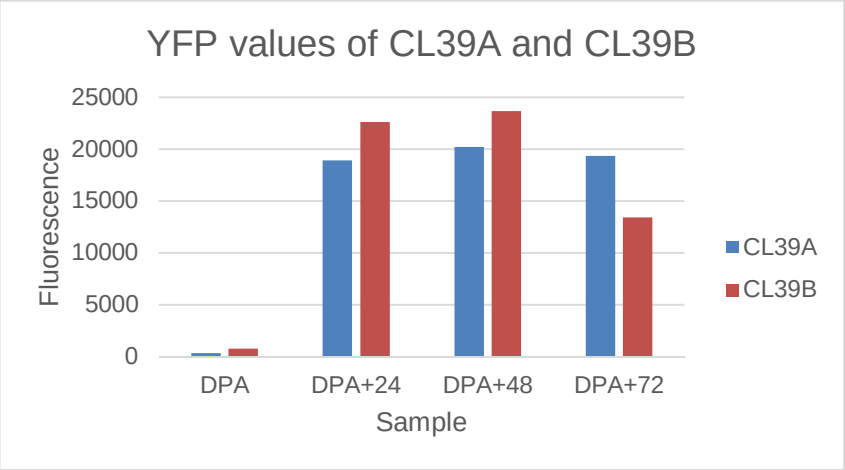


Figure 11. YFP chart of the CL39 samples. The blue bar represents the duplicate A and the red the duplicate B.

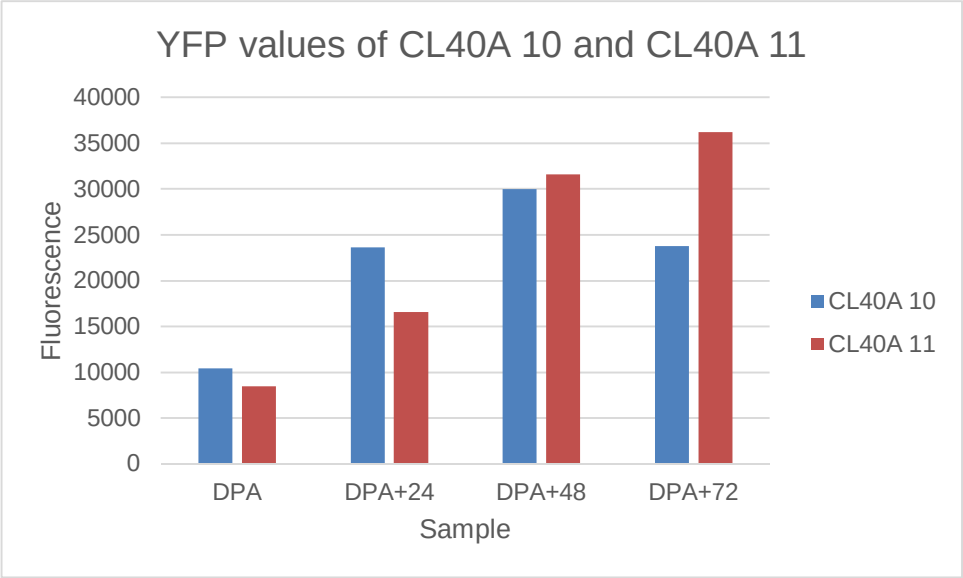


Figure 12. YFP chart of the CL40A samples. The CL40A 10 and 11 are duplicates of the same construct.

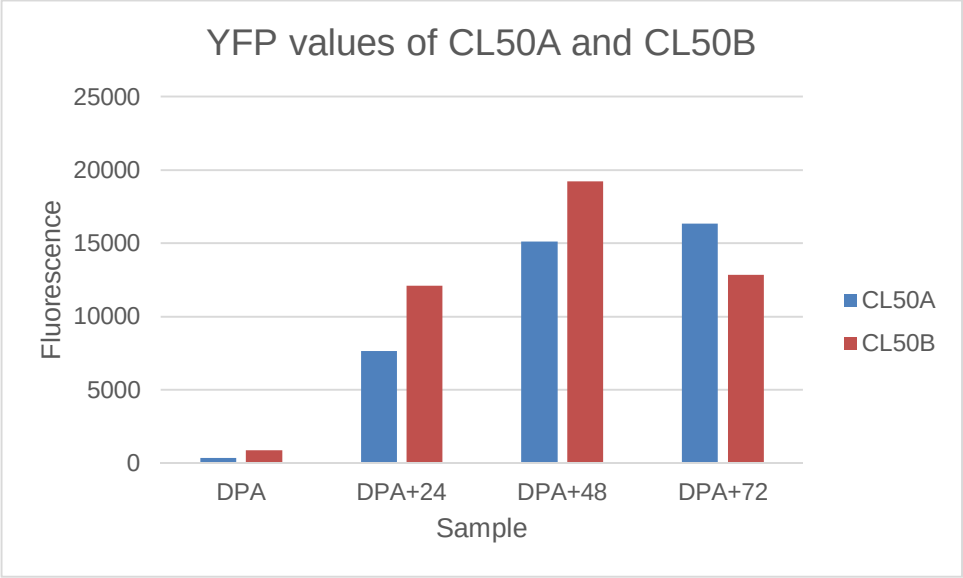


Figure 13. YFP chart of the CL50 samples.

When the plateau phase was reached, the YFP values started to diminish. Sometimes it could happen that the plateau was reached somewhere in the middle of two cell counts. That is the case with the results of CL50A where one can notice a significant drop in the YFP value between DPA+48 and DPA+72 samples. The YFP measurement could also have been done every 12 h after DPA to make sure not to miss the plateau phase.

It can be interpreted based on the results that the production of YFP protein reached the maximum 1–3 days after DPA depending on the construct. The result correlates with the transcription of the baculoviral genome and at some rate the expression of the ATAC complexes also. Nevertheless, at this part of the experiment it can only be observed if the YFP protein is produced or not.

The plateau is important so that one knows when to harvest the cells. It must, however, to be kept in mind that the cells are then no longer in YFP maximum but in a lower expression state. That means that the amount of the protein in the cell pellet is lower than it would be at the maximum samples.

4.2.4 Qualitative analysis of expression by SDS-Page and Western blot

The SNP and SN samples correlating to each YFP maximum were analysed on SDS-page gels. The Coomassie stained gel (Figure 14) shows that the different subunits of ATAC complexes are produced. For CL39 the units are GCN5 at 100 kDa, ADA2a/ADA3 at 50 kDa and SGF29 at 36 kDa and for CL50 ADA2a/ADA3 at 50 kDa, $GCN5^{481–837}$ at 36 kDa and SGF29 at 35 kDa. The units ADA2a and ADA3 are similarly sized; thus, they cannot be separated by the SDS-page gel. The correct bands are marked as an example of the SDS-page gel figure.

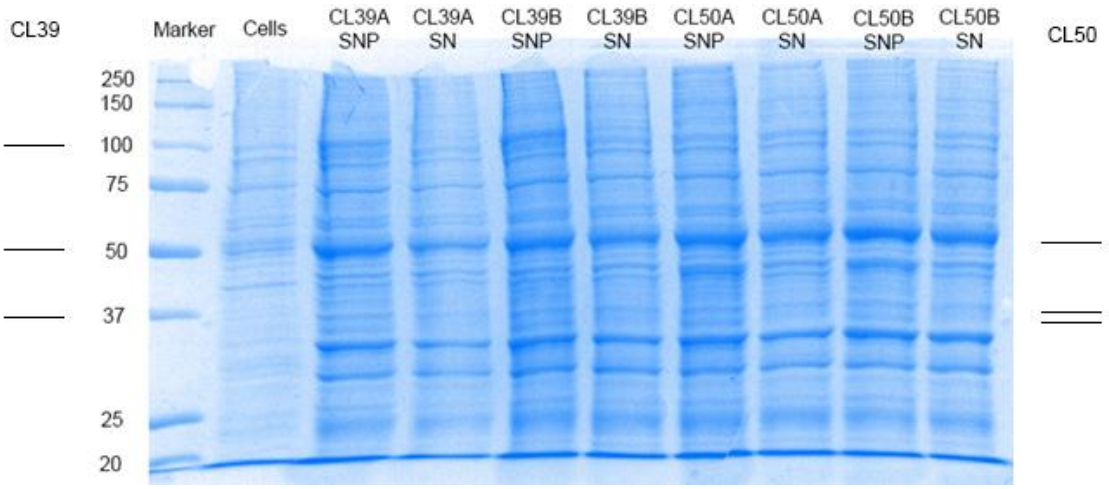


Figure 14. Photo of a SDS-page gel with supernatant and pellet (SNP) and supernatant (SN) samples for CL39A, CL39B, CL50A and CL50B. Samples are from YFP maximums that were DPA+24 h for CL39A and B duplicates and DPA+48 h for CL50A and B duplicates.

The bands for the units can be seen, except that the SGF29 band around 35–36 kDa is a little hard to interpret. The interpretation is made easier with the Sf21 cell control that shows solely the proteins originally present in the insect cells. The lanes are comparable because the cells were counted with a Neubauer chamber before sampling, so every sample contains one million cells.

The solubility of the complexes is not very good as the comparison of the SNP and SN lanes suggest: the bands are stronger on the SNP lanes. The band at 39 kDa corresponds to the VP39, a structural protein present in the baculovirus capsids. The gel could have been run longer to make the bands after 20 kDa separate although all the units of interest are above 25 kDa in size.

For clearer visualisation, also all the YFP samples from DPA+24 to DPA+96 of the construct CL40A were run to see the protein expression enhancement during the insect cell culturing (Figure 15). The bands on the gel represent the subunits of ATAC miniHAT which are ADA2a (around 52 kDa), ADA3$^{85–432}$ (45 kDa) and GCN5$^{481–837}$ (40 kDa). The bands are represented on the right side of the SDS-page gel figure.

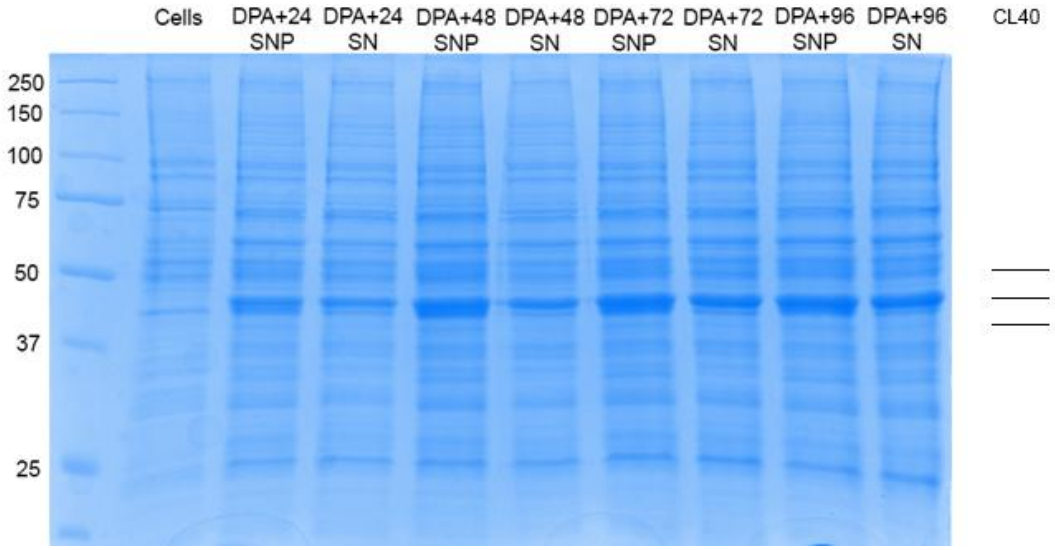


Figure 15. SDS-page gel of sample CL40A, from DPA+24 to DPA+72.

The band for the truncated ADA3$^{85–432}$ is very thick and visible. The bands of ADA2a and GCN5$^{481–837}$ are not similarly clear, but they can be seen especially when compared to the lysed cells. The intensity of the bands grows along with the maturing culture that is what we expect.

At DPA+72 the amounts of protein seem to be at their maximum and that was also when the YFP values were the highest. At approximately 39 kDa, one can see the VP39 protein band.

Comparing to the constructs CL39 and CL50, the amount of soluble protein is greater with the construct CL40. The bands on the SN lanes are weaker than the ones on SNP lanes, but the difference is not as clear as with the other gel. The destaining of the gel could have been longer.

Western blot analysis (Figure 16) shows the protein of interest in more detail. The proteins are his-tagged so the antibody sticks to them and makes them visible. Because the his-tag is located at the C-terminal end of the protein, it is attached to the GCN5 unit of all complexes.

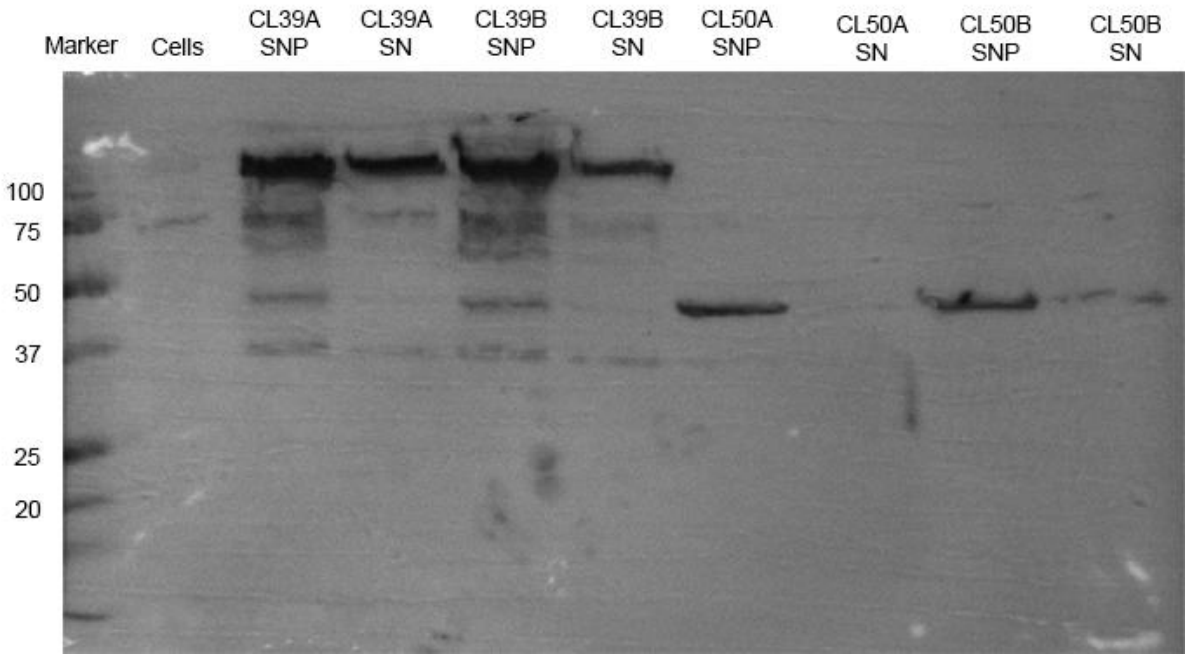


Figure 16. Photo of WB membrane: CL39A, CL39B, CL50A and CL50B. The antibody used was Monoclonal Anti-polyHistidine Phosphatase conjugated Antibody. The WB was performed with the gel from the same run as Figure 14.

From the membrane (samples CL39 and CL50A) one can interpret that the GCN5 unit in CL39 is clearly seen around 100 kDa. Also some unspecific binding can be seen with regard to the weaker bands at 75 kDa, 50 kDa and 37 kDa. The 75-kDa band is not a unit of the ATAC protein, but it is visible on the cell control lane as well; thus, it is originated from the cells. The membrane might have been in the photoreaction solution for a little too long, which showed the other bands.

For CL50 (Figure 17) the results are more varying: the band at the level of ADA2a/ADA3 (50 kDa) is clear on the SNP lanes and weaker on the SN lanes. The band for GCN5 should be at 36 kDa that is not convergent with the WB membrane. It was assumed that the GCN5$^{481–837}$ unit has not cleaved from the complex and it might migrate slower than expected.

The solubility of the protein might possess some problems because the bands on the SNP lanes are stronger than the ones on the SN lanes. According to the previous experiments, the protein was not well soluble in low salt concentration. When the expression and purification is done in larger scale, all the buffers should contain 500 mM of NaCl.

When it comes to CL40A, the results are in accord with CL50. Only the band at the level of the truncated ADA3$^{85–432}$ is seen, more intensively on the SNP than on the SN lanes. The unit GCN5 is maybe bind to some other unit and migrates at higher size level than it should.

The gel was cracked at the DPA+96 lanes, which explains the notch in the bands.

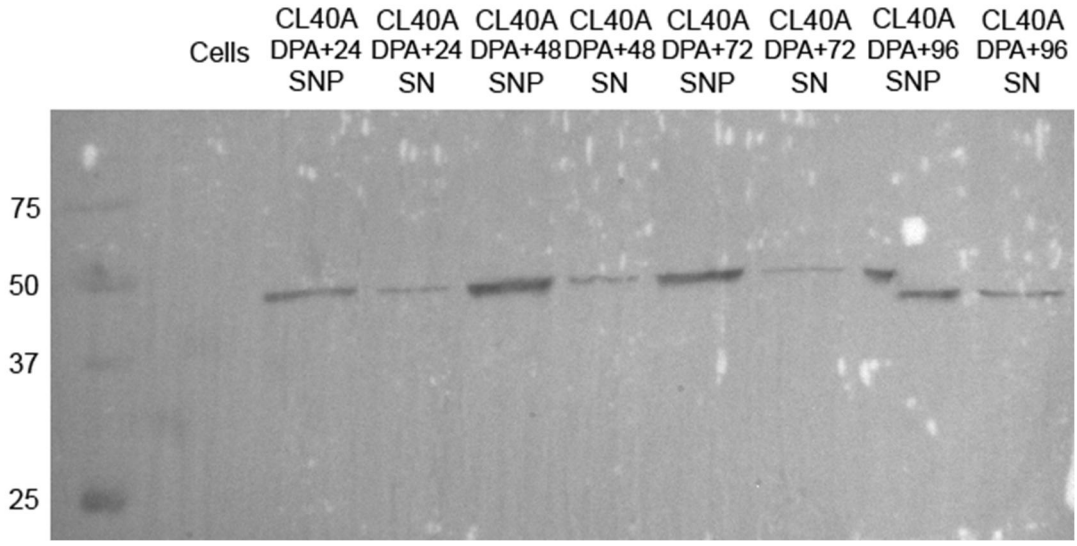


Figure 17. WB analysis for CL40A samples (DPA+24 to DPA+96).

4.2.5 Large scale expression

Following the addition of the second passage of virus (V_1), the cells stopped dividing after 24 h. This is a little too fast because in the ideal case the cells would have multiplied once and then stopped dividing. This would have given a larger amount of cells and thus more protein to work with.

On the other hand, 24 h post-infection all the cells looked well infected and the YFP measurements were started. The values went up quickly, but, also as a downside of the fast infection, the YFP values diminished rapidly after a plateau. The cells were harvested 72 h after infection which led some of the cells at YFP value around 15 000

and the others at 20 000. It was interpreted that the amount of protein would be noticeable.

SDS-page gels of this section correspond with the previous SDS-page figures and show the production of the ATAC HAT proteins. The same subunits can be found as in the small-scale expression. However, the expression does not seem to be as strong on these gels as on the previous ones; the bands are not so thick. The amount of protein was surprisingly at its most intense level in the small-scale expression pellets. The V_1 used to infect the 400 ml cell cultures might have lost some of its power during the storage in 4 °C although that should not happen in a few weeks. It might also be that the cells were not very healthy when they were infected which might have had an effect on the expression levels. One good option would have been to grow larger volume of the cell cultures to increase the protein amount.

The CL50 construct cultures were contaminated at some state of the experiment and due to this fact, it was decided that the pellets of small-scale expression would be purified instead of the large scale pellets of CL50. There are not as many cells as in large scale, but at least the protein is correct and well-expressed.

4.3 Purification

4.3.1 Immobilized Metal Affinity Chromatography

The IMAC purification with Ni2+-NTA resin was successful. On the basis of the Coomassie-brilliant-blue-stained SDS-page gels (Figures 18, 19 and 20) it can be noted that the protein eluted well during the first three elution steps. Some of the protein remained bonded on the resin; thus, the concentration of imidazole could be increased in the elution buffer. Nevertheless, the adsorption and desorption of the protein is based on an equilibrium, so the reaction happens all the time. The FT sample shows that some of the protein does not bind to the resin during the first three-hour incubation. The binding step could be repeated to see if it affects the binding. Because of the time management reasons this could not be tested during the thesis project.

All subunits of the complexes are well visible on the SDS-page gels. The elution fractions containing protein were pooled (E_1–E_4), and the purification was continued with the total volume of sample.

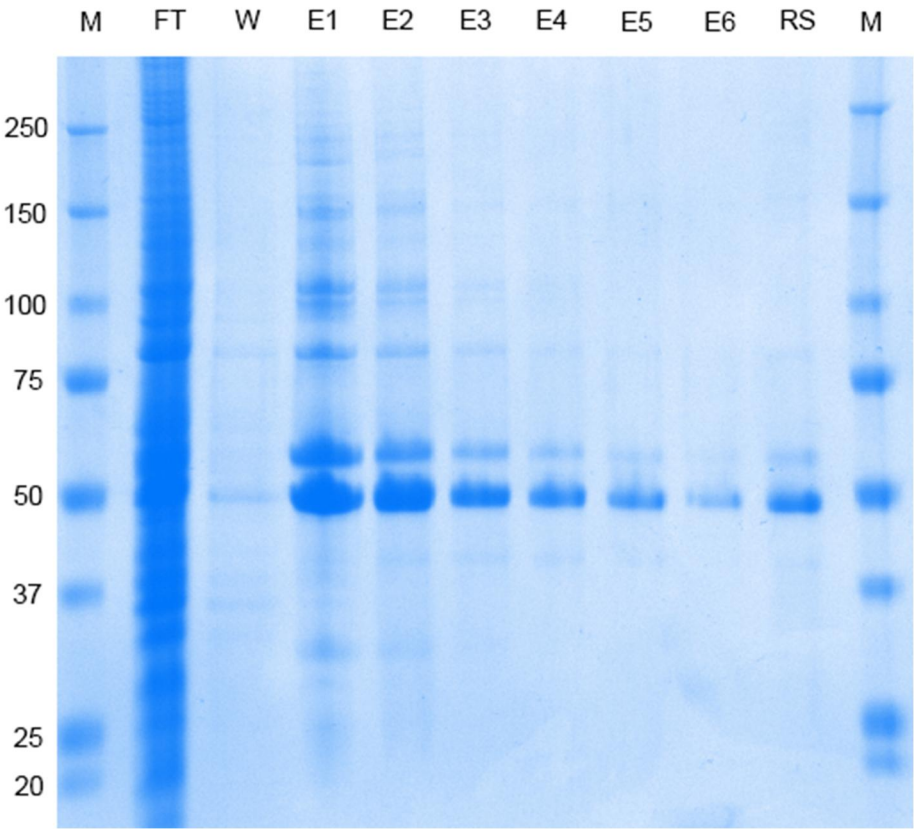


Figure 18. IMAC purification of CL39A. (E=elution, FT=flow through, W=wash and RS=resin).

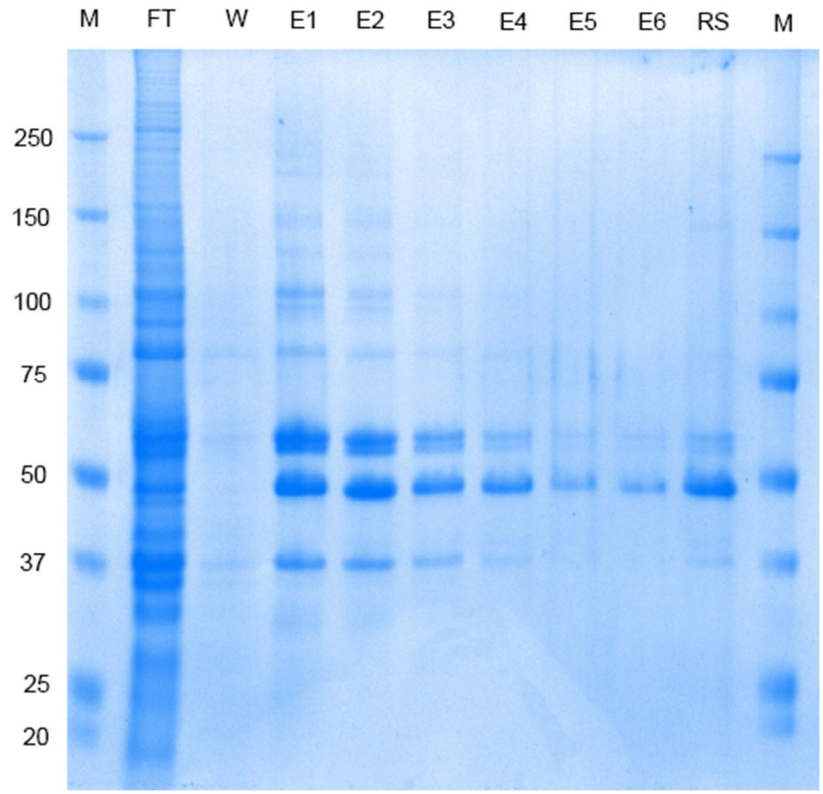


Figure 19. IMAC purification of CL40A I

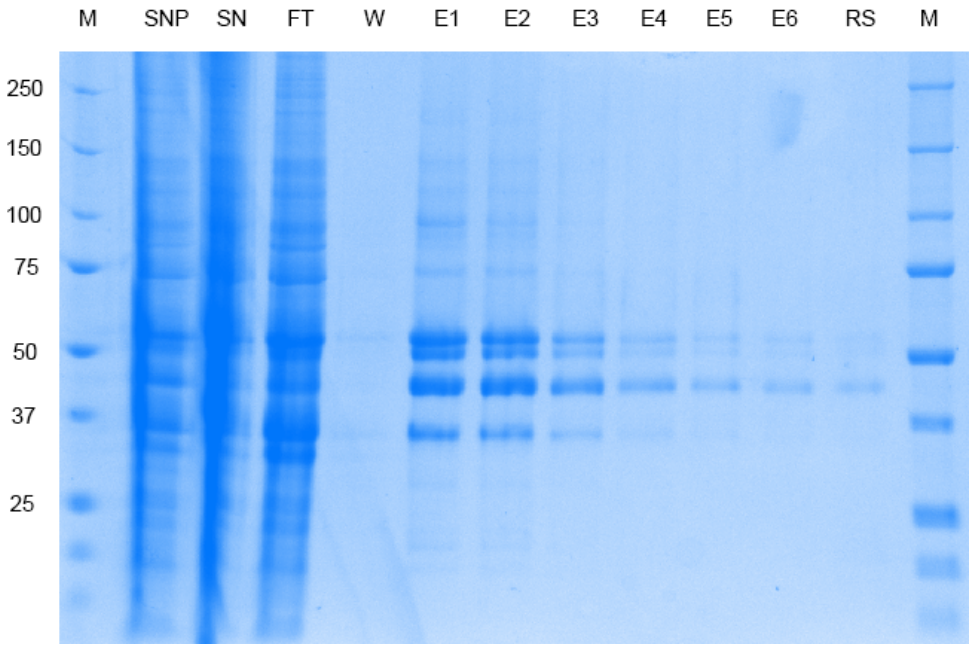


Figure 20. IMAC purification of CL50A

4.3.2 Size Exclusion Chromatography

The size exclusion chromatography (SEC) purification was run quite successfully with all the samples as well. The amount of protein was not very large, but the elution of the target protein could be visually noted by SDS-page analysis (Figures 21, 22 and 23). Some subunits are present in lower quantity than others which may explain the differences in the band intensities. Alternatively, it is known that the stain which was used to visualize the bands, Coomassie Brilliant Blue, sometimes stains proteins stronger or weaker, depending on their lysine and arginine (i.e. positively charged) amino acid contents to which the dye binds. The concentration of protein is very low as well so the bands are somewhat weak in general.

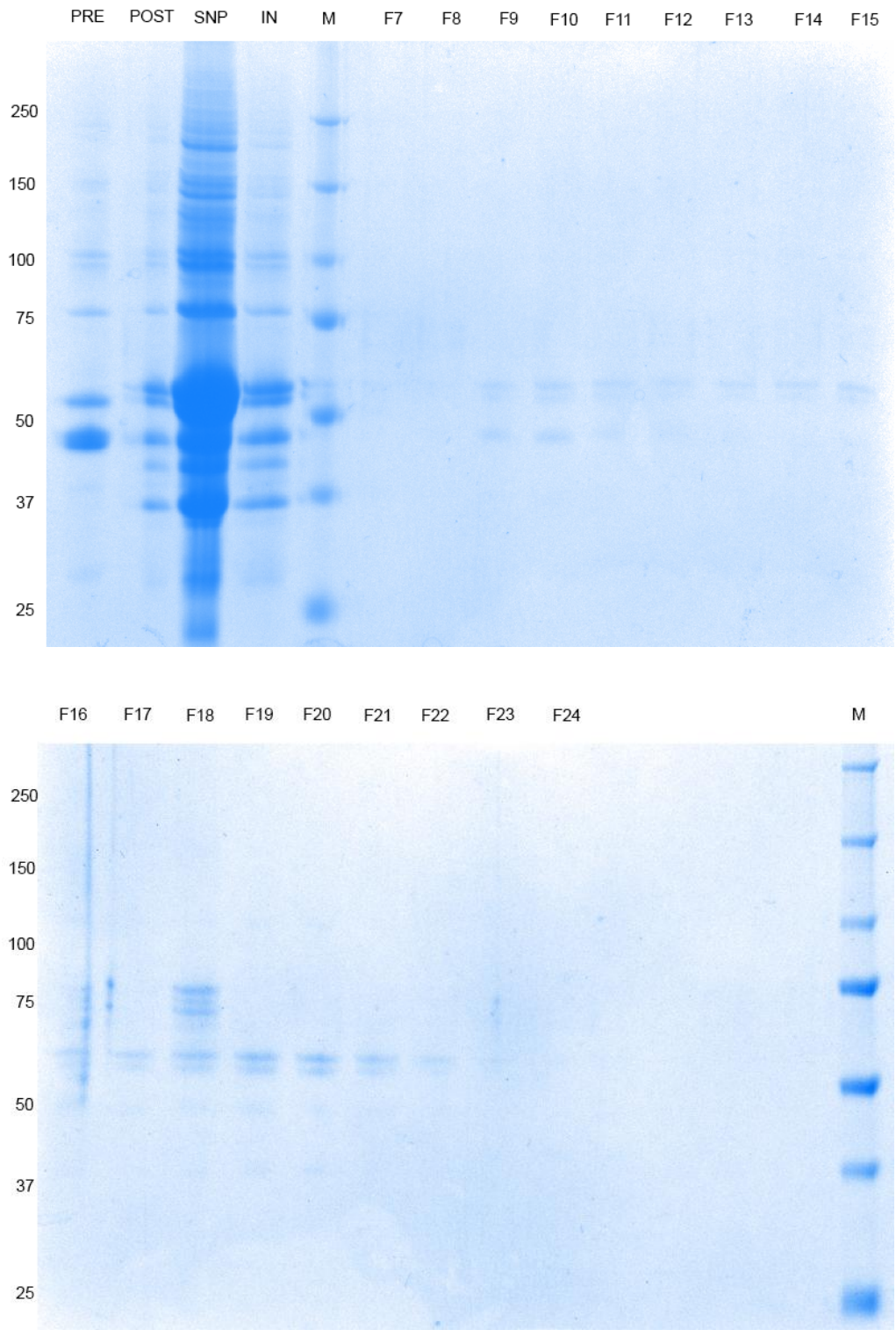


Figure 21. SDS-page gels of SEC elution fractions (CL39A).

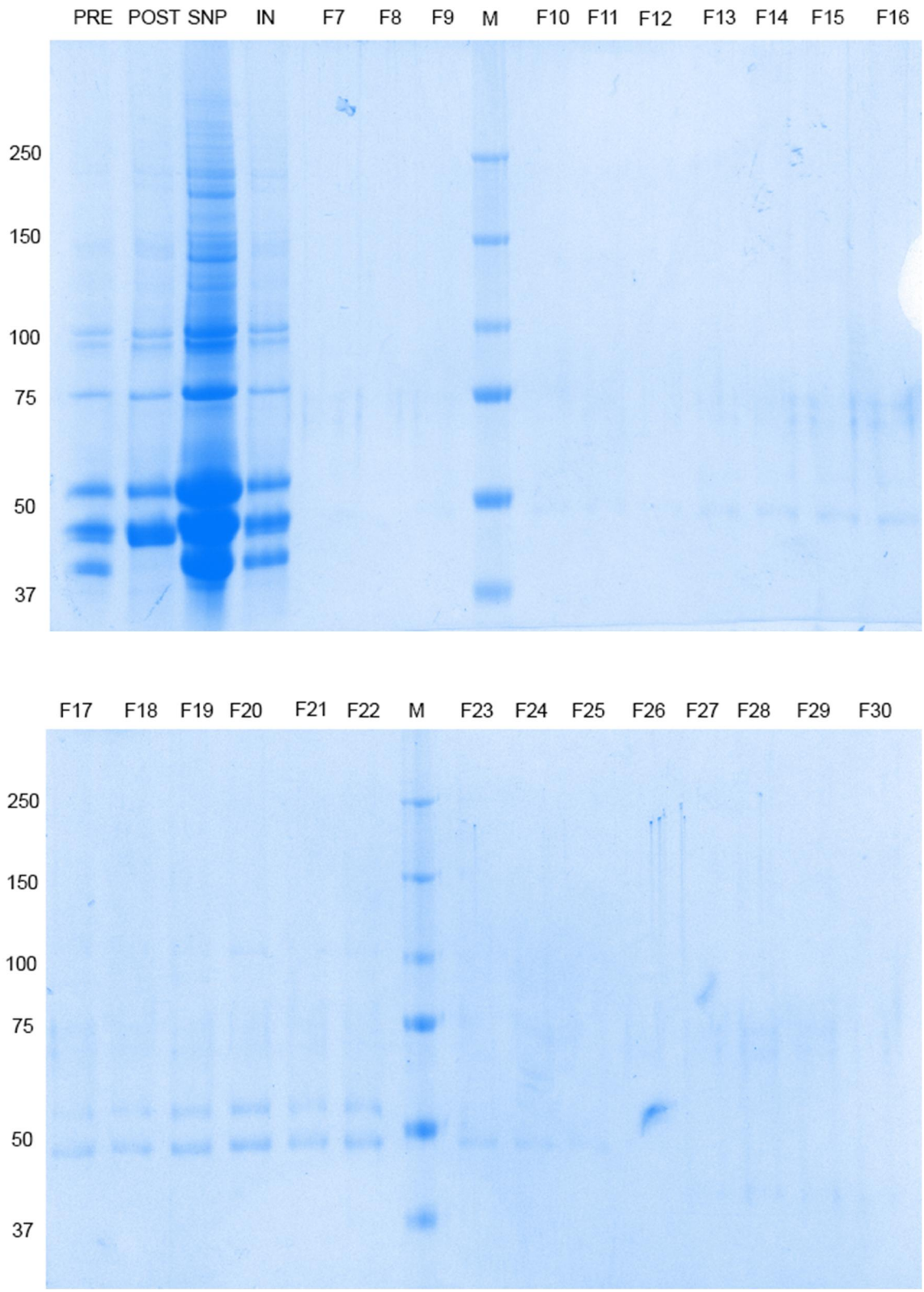


Figure 22. SDS-page gels of SEC elution fractions (CL40A 11)

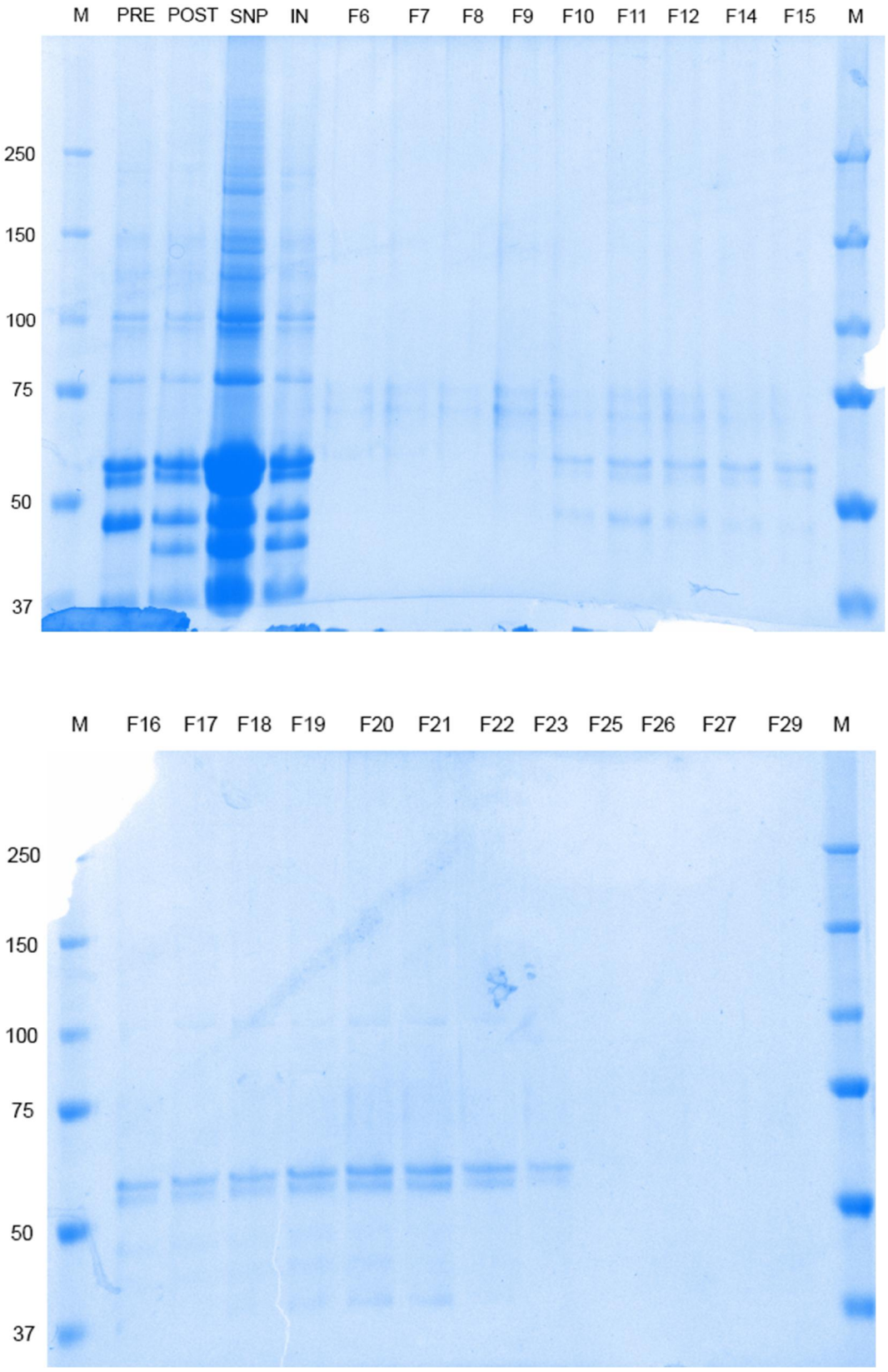


Figure 23. SDS-page gels of SEC elution fractions (CL50A)

The ATAC complexes could be stated to elute at the volume of approximately 12 ml. The fact correlates with the calibration table made for S200 10/300 column by the laboratory (Appendix 6) and the approximate calculated molecular weight of the complex, which is 216 kDa. The peak profiles remained the same with all the samples, which tells that the contaminants and the protein remain similar with different complexes. One SEC chromatogram is presented as an example (Figure 24) and the rest can be found in the Appendix 5.

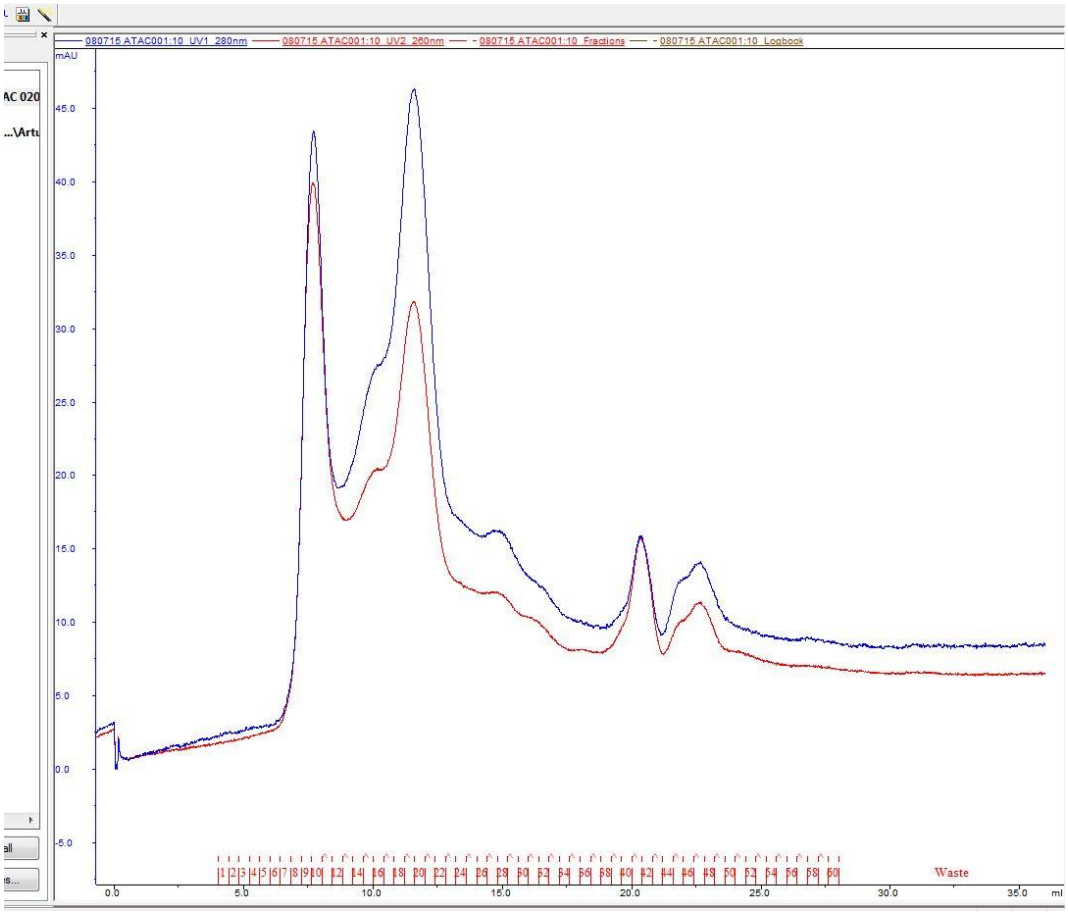


Figure 24. SEC chromatogram of CL39

For CL39 fractions 15–21, CL40A 17–22 and CL50A 10–23 were pooled and concentrated. The final amount of protein was around 0.1 mg/ml which is low and needs to be improved.

5 Conclusion

All in all, the project of expressing and purifying ATAC HAT complexes was successful when it comes to expression and purification by IMAC and SEC. The final amount of pure protein stayed around 0.5 mg per construct; thus, the yield was not overly high. This can be ameliorated by using larger expression cultures.

The protocol from the previous researcher on the topic provided a profound basis to the project. A great amount of effort had been devoted to it and and my results indicate that with some more practice I could reach similar yields and quality of proteins. The work period of 14 weeks with this topic was rewarding and also very challenging. With the time limit I could not reach all the goals in their entirety I would have liked to have achieved.

The thesis project provided more information about the purification of ATAC HAT. It also provides the full expression protocol from the beginning until the extraction of the protein. Mainly, the project was a project of learning. I have deepened my understanding of proteomics and the techniques used in this field as well as my skills of working as a part of an international research group. The experiences I got in EMBL are invaluable. As a result of the project the plasmids, viruses and some protein exist for further use.

The research work could be continued by a next researcher joining the project by improving the SEC purification step. The proteins might have been aggregated because no clear peak but multiple peaks could be perceived. The complexes elute just at the beginning of the elution; therefore, a column with a better separation on that area could improve the peak separation. A greater volume of cell culture would also improve the purification because there would be more protein to be purified and the yields would be higher.

The results of the thesis project will be used to analyse the structure of the ATAC HAT complexes. More protein will be produced with the virus obtained, and all the purified proteins, alone and in complex with nucleosomes, will be analysed by an MS spectrometer.

References

An, W., 2007. Histone acetylation and methylation. In: Kundu, T.K, Dasgupta, D., ed. 2007, *Subcellular Biochemistry - Chromatin and disease.* Springer, New York, USA. Volume 41. Ch.16.

Berger, I., Fitzgerald, D.J., Richmond, T.J., 2004. Baculovirus expression system for heterologous multiprotein complexes. *Nature biotechnology,* [online] Available at: <http://www.nature.com/nbt/journal/v22/n12/full/nbt1036.html> [Accessed 17 April 2015].

Brown, T.A., 2010, Gene Cloning and DNA Analysis, 6th edition. [e-book] Wiley-Blackwell, UK. Available through: Metropolia library website http://libguides.metropolia.fi/> <http://site.ebrary.com.ezproxy.metropolia.fi/lib/metropolia/reader.action?docID=10387072> [Accessed 9 June 2015].

Cube Biotech. 2013. NTA vs. IDA: a tale of two ligands [online] Available at: <http://www.cube-biotech.com/files/AppNotes/AN_NTAvsIDA_0913.pdf> [Accessed 14 August 2015].

Goodsell, D., 2000. Nucleosome. *PDB Protein Data Bank,* [online] Available at: <http://www.rcsb.org/pdb/101/motm.do?momID=7> [Accessed 23 June 2015].

Goodsell, D., 2007. Zinc fingers. *PDB Protein Data Bank,* [online] Available at: <http://www.rcsb.org/pdb/101/motm.do?momID=87> [Accessed 21 May 2015].

Graig, A., Berger, I., 2011. ACEMBL Expression System Series MultiBacTurbo – User Manual, [online] Available at: <http://www.embl.fr/multibac/multiexpression_technologies/multibac/> [Accessed 22 April 2015].

Guelman, S., Kozuka, K., Mao, Y., Pham, V., Solloway, M.J., Wang, J., Wu, J., Lill, J.R., Zha, J., 2008. The Double-Histone-Acetyltransferase Complex ATAC Is Essential for Mammalian Development. *American Society for Microbiology – Molecular and Cellular Biology*, [online] Available at: <http://mcb.asm.org/content/29/5/1176.full> [Accessed 16 April 2015].

Harvey, R.A., Ferrier, D.R., Cooper, M., 2011. *Biochemistry.* Lippincott Williams & Wilkins, USA.

Jarvis, D.L., 2009. Baculovirus-Insect Cell Expression Systems. *Methods in Enzymology*, [online] Available at: <http://www.sciencedirect.com/science/article/pii/S0076687909630147> [Accessed 17 April 2015].

Juers, D.H., Matthews, B.W., Huber, R.E., 2012. LacZ β-galactosidase: structure and function of an enzyme of historical and molecular biological importance. *Protein Sci,* [online] Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23011886> [Accessed 24 April 2015].

Life Technologies. Bluo-Gal manual. Available at: <http://tools.lifetechnologies.com/content/sfs/manuals/15519028.pdf> [Accessed 14 August 2015].

Learn.Genetics, 2015. University of Utah, Health Sciences. [website] Available at: <learn.genetics.utah.edu/content/molecules/proteintypes/> [Accessed 6 May 2015].

Kalmakoff, J. *Microbiologybytes*, [online] Available at: <http://www.microbiologybytes.com/virology/kalmakoff/baculo/baculo.html> [Accessed 21 May 2015].

Krebs, A.R., Karmodiya, K., Lindahl-Allen, M., Struhl, K., Tora, L., 2011. Saga and ATAC Histone Acetyl Transferase Complexes Regulate Distinct Sets of Genes and ATAC Defines a Class of p300-Independent Enhancers. *Molecular Cell,* [online] Available at: <http://www.sciencedirect.com/science/article/pii/S1097276511007891> [Accessed 28 April 2015].

Nagy, Z., Riss, A., Fujiyama, S., Krebs, A., Orpinell, M., Jansen, P., Cohen, A., Stunnenberg, H.G., Kato, S., Tora, L., 2010. The metazoan ATAC and SAGA coactivator HAT complexes regulate different sets of inducible target genes. *Cellular and Molecular Life Sciences,* [online] Available at: <http://link.springer.com/article/10.1007%2Fs00018-009-0199-8#page-1> [Accessed 15 April 2015].

Peters, J.E., Craig, N.L. 2001. Tn7: Smarter Than We Thought. *Nature Reviews Molecular Cell Biology* [online] Available at: <http://www.micab.umn.edu/courses/8002/Peters.pdf> [Accessed 14 August 2015].

Porath, J. 1988. IMAC – Immobilized metal ion affinity based chromatography. *Trends in Analytical Chemistry,* [online] Available at: <http://www.sciencedirect.com/science/article/pii/016599368885074X#> [Accessed 26 June 2015].

Riss, A., 2012. Caractérisation fonctionelle de l'activité de l'histone acetyltransferase GCN5 au sein des complexes ATAC et SAGA chez l'homme. [doctoral thesis] Université de Strasbourg. Available at: <http://www.hal.inserm.fr/file/index/docid/868785/filename/RISS_Anne_2012_ED414.pdf> [Accessed 21 May 2015].

Sambrook, J., Russell, D.W., 2001. Molecular Cloning – A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York, USA.

Schram, A.W., Baas, R., Jansen, P.W.T.C., Riss, A., Tora, L., Vermeulen, M., Timmers, T.Th.M., 2013. A Dual Role for SAGA-Associated Factor 29 (SGF29) in ER

Spedale, G., Timmers, M., Pijnappel, P., 2012. ATAC-king the complexity of SAGA during evolution. *Genes & Development*, [online] Available at:<http://genesdev.cshlp.org/content/26/6/527.full> [Accessed 15 April 2015].

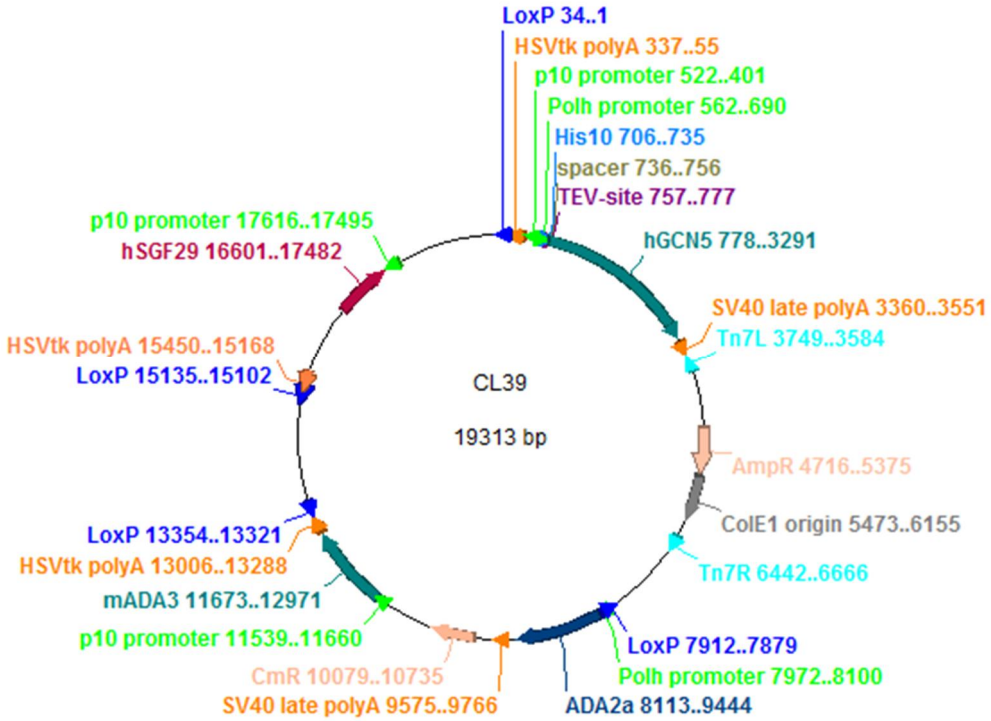
Sulkowski, E. 1989. The Saga of IMAC and MIT. *Bio Essays,* [online] Available at: http://onlinelibrary.wiley.com/doi/10.1002/bies.950100508/pdf [Accessed 26 June 2015].

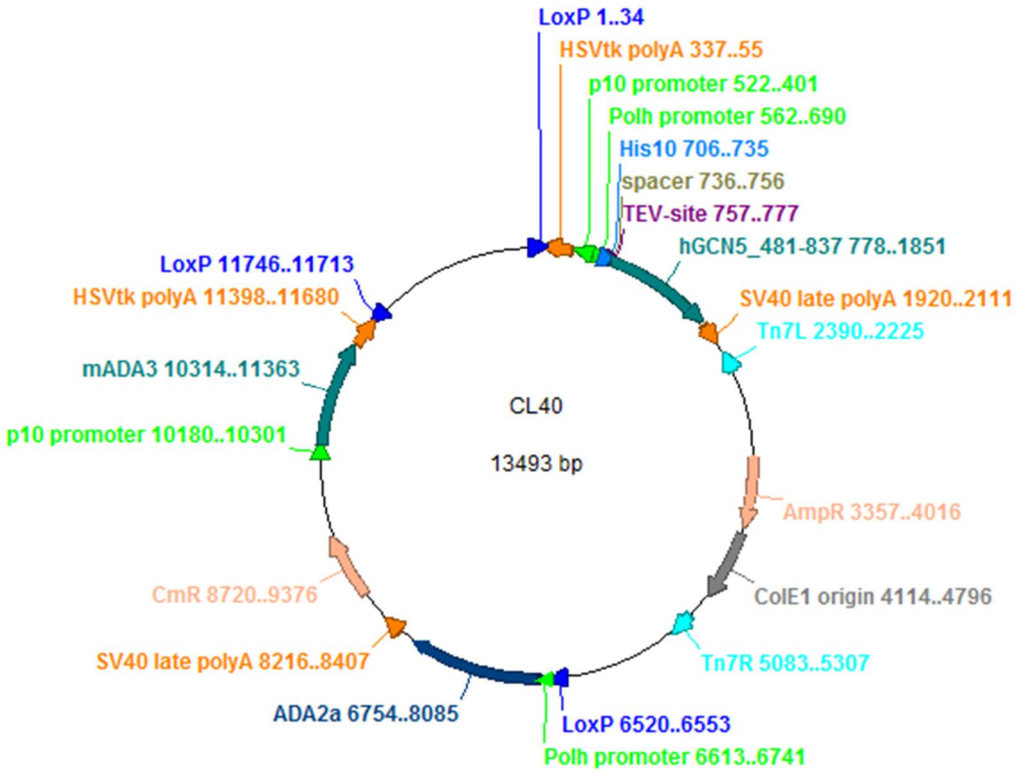
The Jackson Laboratory. 2015. *Introduction,* [online] Available at: <http://cre.jax.org/introduction.html> [Accessed 21 May 2015].

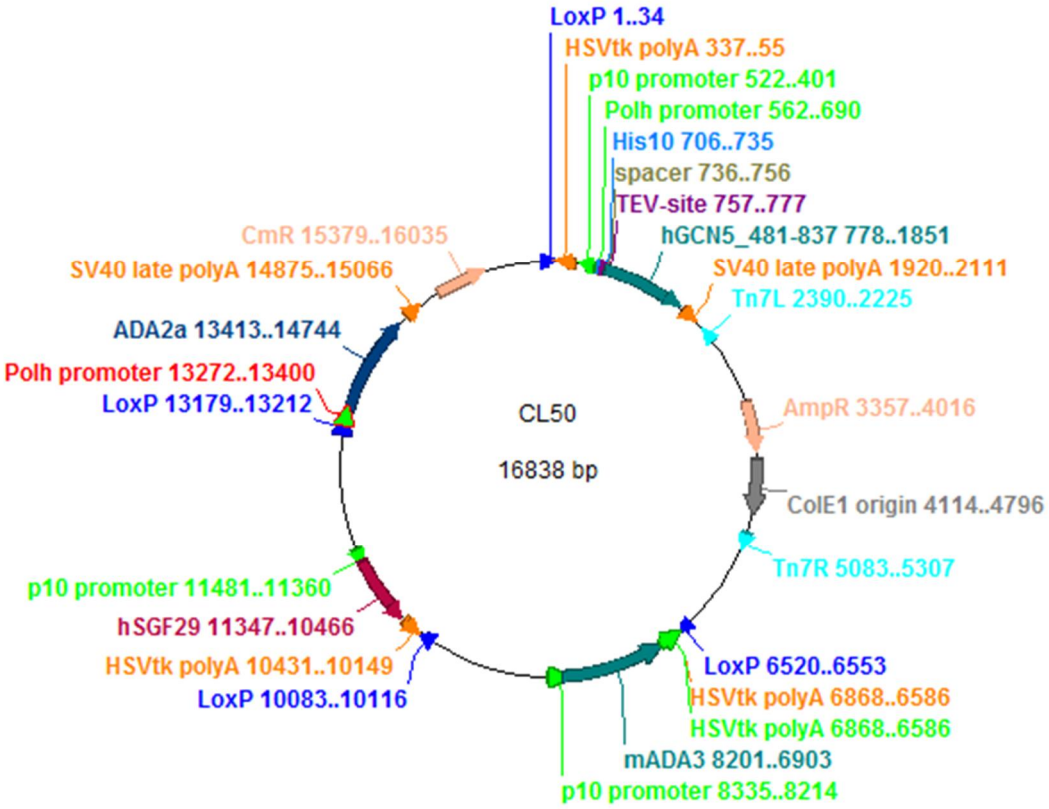
Trowitzsch, S. trowizsch@biochem.uni-frankfurt.de, 2015. *SAGA and ATAC miniHAT*. [email] Message to V.Soikkeli (vsoikkeli@embl.fr). Sent Tuesday 14 April 2015: 22:51. [Accessed 16 April 2015].

Verdone, L., Caserta, M., Di Mauro, E., 2005. Role of histone acetylation in the control of gene expression. *Biochemistry and Cell Biology*, [online] Available at: <http://www.nrcresearchpress.com/doi/abs/10.1139/o05-041?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed&#.VS-nSRYfGWV> [Accessed 15 April 2015].

Appendix 1 Plasmid maps for CL39, CL40 and CL50






LoxP 1..34
HSVtk polyA 337..55
p10 promoter 522..401
Polh promoter 562..690
His10 706..735
spacer 736..756
TEV-site 757..777
hGCN5_481-837 778..1851
SV40 late polyA 1920..2111
Tn7L 2390..2225
AmpR 3357..4016
ColE1 origin 4114..4796
Tn7R 5083..5307
LoxP 6520..6553
HSVtk polyA 6868..6586
HSVtk polyA 6868..6586
mADA3 8201..6903
p10 promoter 8335..8214
LoxP 10083..10116
HSVtk polyA 10431..10149
hSGF29 11347..10466
p10 promoter 11481..11360
LoxP 13179..13212
Polh promoter 13272..13400
ADA2a 13413..14744
SV40 late polyA 14875..15066
CmR 15379..16035
CL50
16838 bp

Appendix 2 Eukaryotic Expression Facility brochure

Eukaryotic Expression Facility at EMBL Grenoble

(June 2013 edition)

GARZONI Frederic, AUBERT Alice and DRNCOVA Petra BERGER Imre

Blue/white screening & transfection

Day 1

Use the competent cells DH10EMBacY. Transform them with the plasmid:

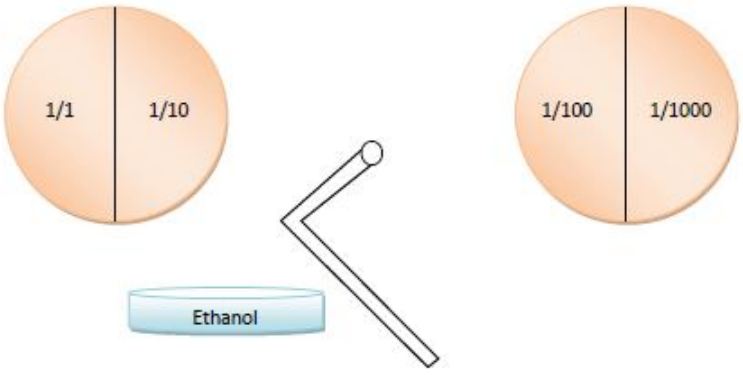
- Take one tube of DH10EMBacY out of -80°C and put immediately on ice!
- Add ~1 µg of plasmid (sterile!)
- Leave on ice 20 min
- Heat-shock 45 sec at 42°C
- Leave on ice 2 min
- Add 400 µL of LB (w/o antibiotic) (sterile!)
- Incubate O/N shaking at 37 °C.

Day 2

Streak out the transformed cells on Kanamycin/Tetracycline/Gentamycin/IPTG/BluOGal plates in dilution series as followed:

1/1 → 1/10 → 1/100 → 1/1000 (135 µL of medium and 15 µL of cells)

Put 135 µL of each dilution on plates as followed (keep the rest of the 1/1 undiluted cells at 4°C).



Incubate the plates O/N at 37°C.

Day 3

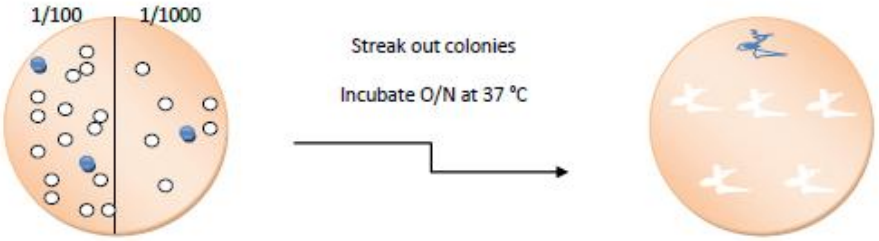
You should obtain white and blue colonies after >24 hrs.

Both types of colonies now have the plasmid and will be resistant to Gentamycin.

The blue colonies are still able to process the BluOGal (via the betaGalactosidase enzyme), thereby producing a blue compound.

The white colonies owe their color to the loss of the betaGalactosidase activity, following the transposition of the gene of interest into their baculoviral genome and the disruption of the lacZ gene (betaGalactosidase gene).

Pick 5 white colonies on Kan/Tet/Gent/BluOGal/IPTG to confirm the phenotype and 1 blue colony as a positive control from the plate.



Day 4

For each plate, choose 2 white clones and launch 2 mL of LB cultures for each with Kan/Tet/Gent shaking at 37°C O/N.

Day 5

ON THE BENCH

1) Centrifuge the 2mL cultures 10 min at 2900 rcf. Discard the supernatant (freeze at -20°C if prep is scheduled later).
2) Add 300 µL of Buffer P 1 (home made or Qiagen MiniPrep Kit) to the pellet, pipet up & down to resuspend and transfer to a 1.5 mL tube.
3) Add 300 µL of Buffer P2 (MiniPrep Kit). If Buffer P1 contains Lysis Blue, your sample will turn blue. GENTLY invert the tube until the blue color is homogenous. Do not incubate > 5 mins.

4) Add 300µL of Buffer N3 (MiniPrep Kit). A white precipitate will now appear. GENTLY invert the tube until the blue colour has completely disappeared.

   Centrifuge 10 mins at max speed.

5) Transfer the supernatant into a new tube in one shot, using P1000.

   Centrifuge 5 mins at max speed.

6) Transfer the supernatant into a new tube as previously.

   Add 700 µL of 100% Isopropanol (ideally you want to use 0.7 volume isopropanol to have roughly 40% final concentration of isopropanol).

   Invert GENTLY until isopropanol and the aqueous portion are homogeneously mixed.

   Centrifuge 10 mins at max speed.

7) Take off most of the supernatant with p1000 and then the last amount with p200. You might not see any pellet (bacmid pellet should be clear). A white pellet might be due to residual white precipitate (eg. RNA).

   Add CAREFULLY (drop by drop on the opposite side of the pellet, holding the tube horizontally) 200 µL of 70% Ethanol.

   Centrifuge 5 mins at max speed.

8) Take off Ethanol with p200.

   Add CAREFULLY (drop by drop on the opposite side of the pellet, holding the tube horizontally) 50 µL of 70% Ethanol. (Should not sit in ethanol longer than 30mins)

UNDER A STERILE HOOD (EEF)

1) Take care of your DNA

Remove all the Ethanol with p200.

Let pellet dry (open eppendorf) during 10 mins (DO NOT OVERDRY).

Meanwhile do step 2.

Add **20 µL of filter sterilized water**.

Resuspend pellet by tapping 10 times on the hood bench.

Add **200 µL of medium** in each tube.

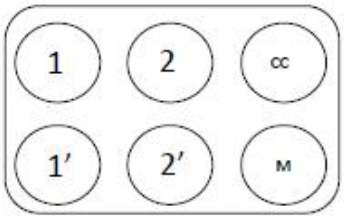
2) Take care of your cells

Count your cells. You need to have between 0.5 and $1*10^6$ cells/well in a total volume of 3 mL.

Note 1: you will transfect with DNA from 2 different clones you have picked from your plate for each plasmid.

Note 2: you will set up duplicates of each clone in case one gets contaminated during the experiment. Therefore for each plasmid (construct) you will set up 2*2 = 4 wells
You will also need one well for cell control and one for media control.

First, pipet the appropriate volume of medium (2mL/well, 3mL for the media control), then the cells, drop by drop (using "S" setting on pipetboy) and evenly (1mL/well).
Leave ~ 15 mins at 27°C (RT).



CC = Cell control (uninfected cells) ; M = Medium only

3) Take care of the transfection reagent

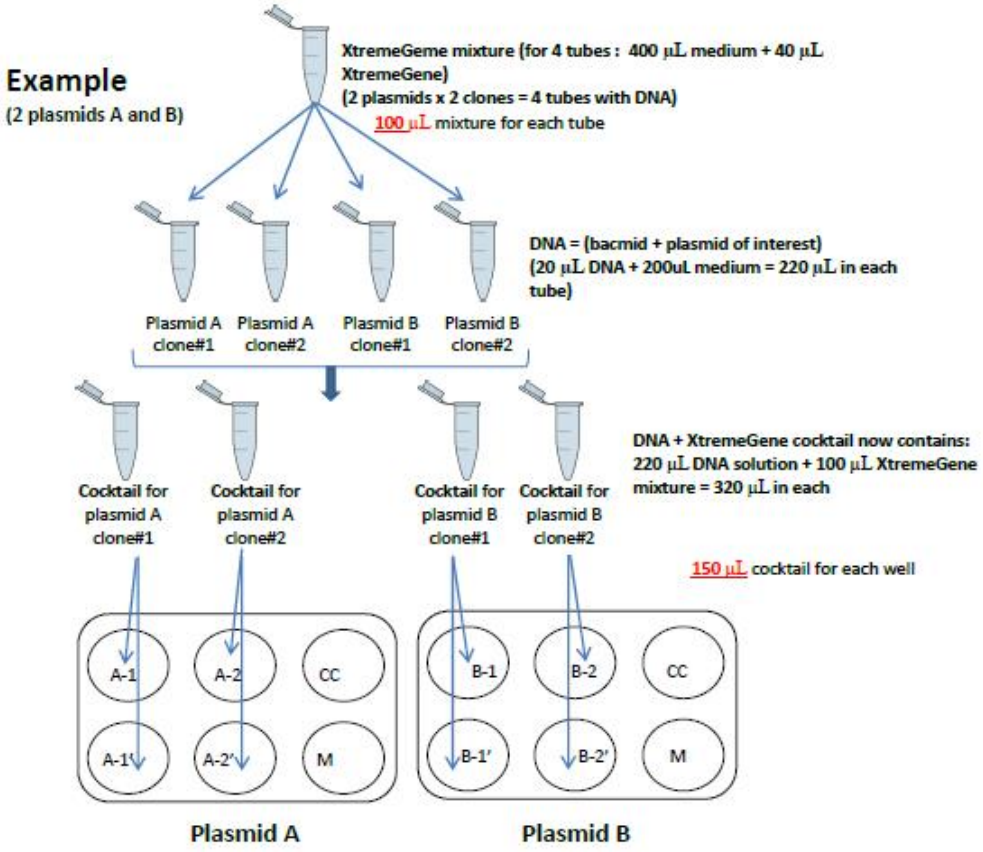
For X tube(s) of samples, prepare ONE tube with:
X00µL of medium
+ X0µL of XtremeGene Transfection Reagent

4) Mix DNA and transfection reagent

Add 100 µL of the XtremeGene mixture to EACH of your DNA containing tubes.

5) Add the transfection cocktail to your cells

From each DNA+XtremeGene cocktail, take 2x 150µL that you add to each of the two dedicated (duplicate) wells.
Use a p200 and be careful not to pass your hand over the uncovered wells; slowly rotate the plate each time. Pour the cocktail drop by drop.



Wait for 48-60 hours. Check every day to see if everything is clear (no contamination, cells still alive, etc). Infected cells will appear larger, irregularly shaped, and not confluent compared to control cells.

Then, remove supernatant from 6 wells plates. **This is the V_0.**

Tilt the plate so that you can put the pipet against a side and collect the entire medium. You can combine the Vo for each duplicate (eg. 1 and 1'). You will now have ~6 mL virus for each clone; put it in a 15mL Falcon tube, seal with parafilm, wrap in aluminum foil and store at 4°C.

Once you have harvested Vo, add 3 mL of fresh medium to the 6 well plate (gently, drop by drop and on the side in order not to disrupt the cell layer).

Wait for 48-60 hours and harvest the cells (process with the "protein expression test on 6-wellplate" protocol).

Note 3: you may already start the V1 amplification at this stage, if you have some pre-adapted flasks of 25 mL culture at a density of 0.5 to 0.8*10^6 cells/mL. Add 3 mL of V0 per 25 mL of culture.

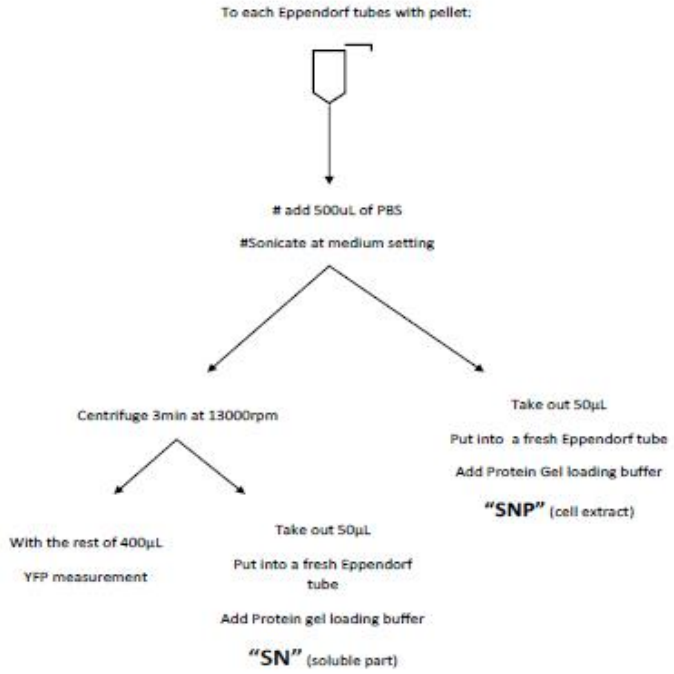
Protein Expression Test experiment on 6-well plate

Since the number of cells in the 6-well plate in not known, one has to be aware that this is a qualitative test.

96-120 hrs after the transfection / 48-60 hrs after the harvesting of V0, remove the supernatant
At this step 6-well plate can be stored at -20 °C.

Resuspend both duplicate wells corresponding to each clone with 500 µL of 1X PBS by pipetting up and down. Transfer in a fresh 1.5 mL tube.
Then, follow the "**Processing the sample for SDS-PAGE and YFP measurement**" protocol, starting from "Sonicate at medium setting" step.

Processing the Sample for SDS-PAGE and YFP measurement



Note well: if your analysis includes several Western blot runs, take > 50 µL SNP and SN.

Freeze the SNP/SN aliquots at -20°C. Load 100 µL of each remaining SN sample onto a black 96 well BD Falcon micro-plate. Also include 100 µL of PBS and YFP standard in two of your wells.

(Biophysical Platform CIBB 001)

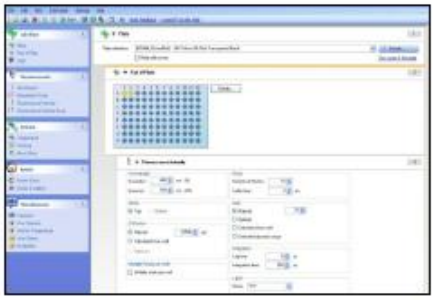
- Turn on the machine with the main switch located at the back.
- Press the *eject* button, on the top right to open the sliding plate holder.
- Insert the plate in the correct orientation and press the *eject* button to close.
- On the desktop, double click on the TECAN icon.
- A window will pop up; select "Default Script" and double click.
- To open a script file of your choice, click on the folder (or file "open").
- You should see the following window: "EEF user only" and 5 script options.

What is the difference between a "detection scan" and a "single read"?

Detection scan will result in data output including the full spectrum graph and will take longer. A detection scan is usually used to check for DPA because the shape of the curve will tell us whether there is YFP or not. Also the gain on the detection scan can be adjusted to your own needs to become more sensitive (i.e. changing from 72 to 90 for example).

Single read will only provide one data point at the peak of the spectrum. This is a quick way of checking your large scale expression for example, or following a sample that already has a detectable YFP value.

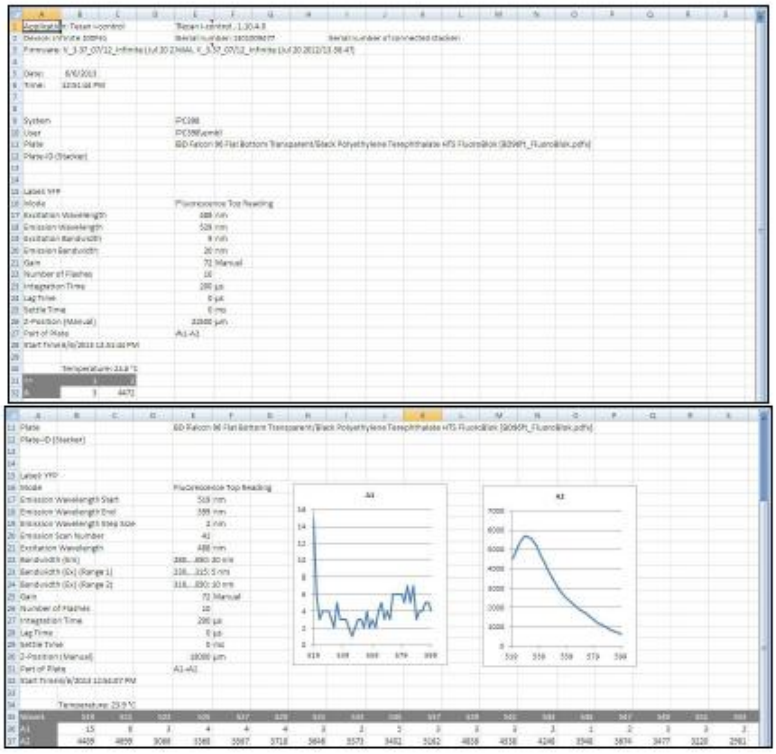
- Double click on the script of your choice. At this point you can either use this script as is and DO NOT MODIFY IT, i.e. do not save any changes. OR save this script under your own name "save as" and then you can modify parameters as you like and use this script in the future.
- When you are ready to measure fluorescence:
  1. **PLATE:** make sure you have selected the correct plate type and that the box labeled "plate with cover" is unchecked.
  2. **PART OF PLATE:** drag the mouse to select the wells you want scanned (they will be yellow instead of blue).
  3. **FLUORESCENCE INTENSITY:** you should not change anything here unless this is your own script.



- To start the measurement, press the *Start* icon or go to *Instrument* → *Start*.

When the measurement is finished, an excel file with the output will automatically appear on the screen. Here are the excel files for a single read and detection scan.

The PBS should give no reading and there should be no curve present. The YFP standard value should be roughly around 4500 to 5000.

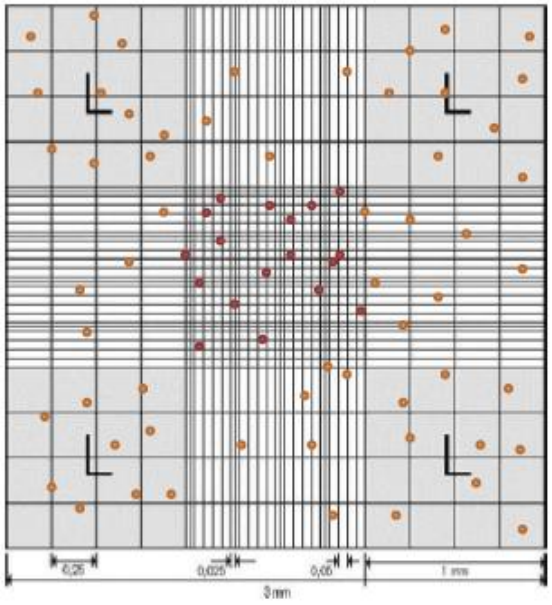


- You can connect to the T:drive and save your excel files under your name folder.
- When you close the TECAN script window, you will be asked: **The script "EEF ....." has been changed, do you want to save it?** Click NO
- Take out your plate and switch off the machine if it's the end of the day (> 6 pm).

How do we count cells?

Fill a Neubauer's cell with 20 µL of suspension.

Count all insect cells present inside the 16 large squares (see red cells below).



Here one can count 18 cells.

This is the number of cells per mL with a 10^4 factor.

Here we have $18*10^4$ cells/mL or $0.18*10^6$ cells per mL.

Information:

The population of cells is doubling every 24 hrs.

The "healthy" concentration of cells is between 0.5 and $2*10^6$ cells/mL. Therefore, one has to split the cells one time a day around the same time.

Virus Amplification & Protein Expression Test experiment

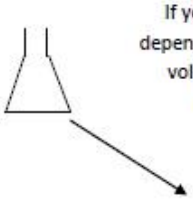
Infect shaker flasks containing 25mL at $0.5*10^6$ cells/mL with 3 mL of V_0. Count cells every day.

If the concentration is < $1*10^6$ cells/mL, then we don't need to dilute because cells have stopped dividing (they are still growing but the population is not doubling in 24 hrs).

If the concentration is > $1*10^6$ cells/mL, then we need to dilute to $0.5*10^6$ cells/mL, because the cells are still in division.

DPA

(Day after Proliferation Arrest)



If your cells have stopped dividing, depending on concentration – retrieve a volume which contains $1*10^6$cells *(i.e. 1.1 mL at density 0.9)*

Transfer cell suspension in Eppendorf tube labelled "DPA"

Centrifuge 2 min at 13000 rpm

Discard supernatant

*

DPA+24

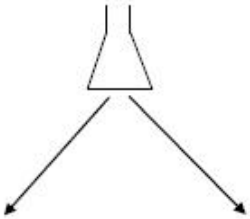
Let's say that you have diluted your culture once and its volume is 50 mL.

Centrifuge culture gently 3 min at 800 rpm in a 50mL Falcon tube.

Take supernatant (V_1 Virus) and put it in a fresh sterile and properly labelled 50mL Falcon tube.

Gently resuspend the cell pellet with 50 ml of fresh medium.

Transfer resuspended cells back into the same shaker flask for expression test.



Count cells and retrieve volume that contains $1*10^6$cells in Eppendorf tube labelled "dpa+24"

Centrifuge 2min at 13000rpm

Discard supernatant

*

* Continue with protocol "Sample Processing" and measure YFP. This sampling must be performed every 24 hrs until the YFP value reaches a plateau. Then harvest cells (800 rpm, 3 min), discard supernatant and keep the pellet at -20 °C.

Small scale expression protocol

Preparation of cells:

Prepare many flasks with 50mL of culture at $0.5*10^6$cells/mL.

Infection:

We need to find the optimal volume of V_1 virus which leads to the proliferation arrest.

If cells stop their proliferation immediately after adding of V_1, restart the experiment by decreasing the volume of V_1.

If cells don't stop dividing, increase the volume of V_1.

Counting the cells:

Count cells every 24 hrs. If they are still dividing, maintain a concentration < $1*10^6$ cells/mL.

Wait for DPA (Day after Proliferation Arrest).

Monitoring YFP:

When cells stop doubling, monitor YFP every 24hrs (keep probes for gel , cf. "Processing the sample for SDS-PAGE and YFP measurement" protocol).

Monitor YFP until it reaches a plateau.

Then, harvest cells, discard supernatant, keep the pellet at -20°C and/or make a protein expression in large volume.

If the expression of protein is good, you may use this V1 pellet to start a small scale purification test.

Freezing and thawing insect cells

1) Freezing cells

- count cells, make sure that you have enough cells for preparing 2-4 vials (see table)
- prepare cryovials on ice
- centrifuge the cells at 400-600g for 10 mins at RT
- remove supernatant

(For High Five cells, keep the conditioned medium for making freezing medium)

- resuspend cells in the given density in the right medium
- transfer 1mL to sterile cryovial
- place at -20 °C for 1h, then store at -80 °C for 24-48h
- transfer in liquid N2 for long term storage.

Cell line	Freezing medium	Density
Sf21	60% Sf900 medium 30% FBS 10% DMSO	$1*10^7$ cells/mL
High Five	42.5% conditioned Express 5 medium 42.5% fresh Express 5 medium 5% FBS 10% DMSO	$3*10^6$ cells/mL

2) Thawing cells from frozen stock

- Remove vials from liquid nitrogen and place it in a water bath at 37°C
- Thaw rapidly with gentle agitation until cells are almost thawed and remove the cells from the water bath **(Leaving cells at 37°C after they have thawed will result in cell death)**
- Quickly decontaminate the outside of the vial by treating with 70% ethanol, dry the vial and place on ice
- Pre-wet a 25 cm2 flask by coating the adherent surface with 4 ml medium
- Transfer the 1 ml cell suspension directly into the 4 ml of medium
- Transfer flask to a 27°C incubator and allow cells to attach for 30-45 minutes
- After the cells are attached, gently remove the medium (as soon as possible to remove the DMSO from the freezing medium)
- feed cells with 5 ml of fresh medium
- after 24h, change the medium
- leave the cells grow until confluence and then start monolayer or suspension culture.

Freezing Baculovirus-infected insect cells (BIIC) stocks

Here is a home-made protocol from the paper in press "The titerless infected-cells preservation and scale-up (TIPS) method for large scale production of NO-sensitive human soluble guanylate cyclase (sGC) form insect cells infected with recombinant baculovirus" from D.J. Wasilko *et al.*

Freezing BIIC (250mL of culture for 25 aliquots of 1mL of BIIC)

1. Grow 250 mL culture of SF21 cells until $1*10^6$ cells/mL
2. Infect cells with Virus (volume should be checked according to "Small Scale Expression" protocol)
3. Maintain a concentration of $1*10^6$ cells/mL until Day after Proliferation Arrest (DPA)
   You will be able to just barely detect YFP on the fluorimeter
4. Centrifuge 800rpm for 5-10 mins. Use 5 falcon tubes to split the volume
5. Prepare a 50mL solution containing:
   45mL = 90% of medium (Hyclone)
   0.5g = 10g/L of BSA
   } Sterile Filtered (0.22 µm)
   5mL = 10% of DMSO – already sterile and for insect cells only
6. When centrifuge is done, remove supernatant
7. Resuspend cells gently to a final density of $1*10^7$ cells/mL (use 25mL of the above solution)
8. Aliquot 1mL into each cryovial
9. Place at -20°C for 1 hr
10. Store at -80°C for 24-48 hrs
11. Store in liquid nitrogen.

Infection after thawing a BIIC (1mL of BIIC for 800mL of expression culture)

1. Pre-adapt 375 mL cultures of uninfected SF21 cells at ~$1*10^6$ cells/mL
2. Quickly thaw one vial in your hands (use paper towels to protect your skin)
3. Dilute the vial in 50 mL of medium (Hyclone)
   You obtain 50 mL at $2*10^5$ cells/mL
4. Add 25 mL of this solution in each of the 375 mL flasks
   You should obtain 400 mL of cells at slightly more than $1*10^6$ cells/mL (the concentration increases by 2.7% only)
5. Maintain the cells at $1*10^6$ cells/mL until you observe the proliferation arrest
6. Monitor the YFP until it reaches a plateau

7. Harvest the cells.

Stock concentrations of reagents in the 103 lab

Antibiotics 1000X:

*Amp: 100mg/mL

*Chl: 30mg/mL (in absolute Ethanol)

*Kan: 50mg/mL

*Tet: 10mg/mL (in absolute Ethanol)

*Gent: 10mg/mL

*Spec: 50mg/mL

Others:

*IPTG: 1M (1000X)

*BluOGal: 100mg/mL (500X) Light sensitive!!!

Cleaning Flasks

Dispose cells and/or media into a glass bottle (bring to autoclave when full)

Add 35-50 ml 10% acetic acid to flask

Brush flask (make sure to remove the cell ring, you can use the same acid for each flasks)

Dispose the 10% acetic acid in an acid waste container (provided by Pierre or Annie)

Rinse flasks with distilled water several times and fill with distilled water to a level above where the cell ring was formed

Bring flasks with your name and room number on autoclave tape to Annie/Virginie/Audrey for EMBL and Pierre/Mireille for CIIB, they will:

~Autoclave once with water inside

~Autoclave a second time without water (dry)

Appendix 3 QIAGEN miniprep kit manual

QIAprep Spin

Protocol: Plasmid DNA Purification using the QIAprep Spin Miniprep Kit and a Microcentrifuge

This protocol is designed for purification of up to 20 µg of high-copy plasmid DNA from 1–5 ml overnight cultures of *E. coli* in LB medium. For purification of low-copy plasmids and cosmids, large plasmids (>10 kb), and DNA prepared using other methods, refer to the recommendations on page 34.

Please read "Important Notes" on pages 12–18 before starting.

Note: All protocol steps should be carried out at room temperature (15–25°C).

Procedure

1. **Resuspend pelleted bacterial cells in 250 µl Buffer P1 and transfer to a microcentrifuge tube.**

   Ensure that RNase A has been added to Buffer P1. No cell clumps should be visible after resuspension of the pellet.

   If LyseBlue reagent has been added to Buffer P1, vigorously shake the buffer bottle to ensure LyseBlue particles are completely dissolved. The bacteria should be resuspended completely by vortexing or pipetting up and down until no cell clumps remain.

2. **Add 250 µl Buffer P2 and mix thoroughly by inverting the tube 4–6 times.**

   Mix gently by inverting the tube. Do not vortex, as this will result in shearing of genomic DNA. If necessary, continue inverting the tube until the solution becomes viscous and slightly clear. Do not allow the lysis reaction to proceed for more than 5 min.

   If LyseBlue has been added to Buffer P1 the cell suspension will turn blue after addition of Buffer P2. Mixing should result in a homogeneously colored suspension. If the suspension contains localized colorless regions or if brownish cell clumps are still visible, continue mixing the solution until a homogeneously colored suspension is achieved.

3. **Add 350 µl Buffer N3 and mix immediately and thoroughly by inverting the tube 4–6 times.**

   To avoid localized precipitation, mix the solution thoroughly, immediately after addition of Buffer N3. Large culture volumes (e.g. ≥ 5 ml) may require inverting up to 10 times. The solution should become cloudy.

   If LyseBlue reagent has been used, the suspension should be mixed until all trace of blue has gone and the suspension is colorless. A homogeneous colorless suspension indicates that the SDS has been effectively precipitated.

4. **Centrifuge for 10 min at 13,000 rpm (~17,900 x *g*) in a table-top microcentrifuge.**

   A compact white pellet will form.

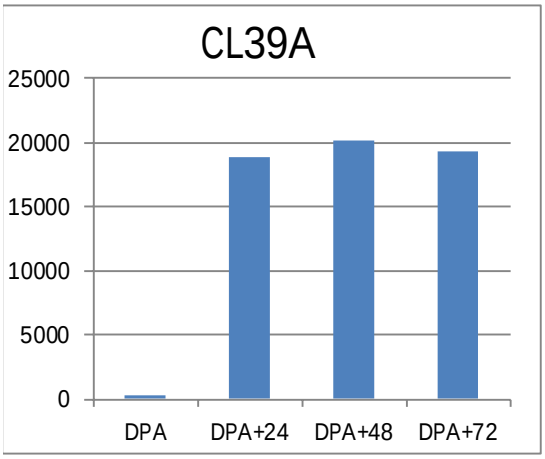
5. **Apply 800 µl of the supernatant from step 4 to the QIAprep 2.0 spin column by pipetting.**
6. **Centrifuge for 30–60 s. Discard the flow-through.**
7. **Recommended: Wash the QIAprep 2.0 spin column by adding 0.5 ml Buffer PB and centrifuging for 30–60 s. Discard the flow-through.**

   This step is necessary to remove trace nuclease activity when using *endA*+ strains such as the JM series, HB101 and its derivatives, or any wild-type strain, which have high levels of nuclease activity or high carbohydrate content. Host strains such as XL-1 Blue and DH5®α do not require this additional wash step.
8. **Wash QIAprep 2.0 spin column by adding 0.75 ml Buffer PE and centrifuging for 30–60 s.**
9. **Discard the flow-through, and centrifuge at full speed for an additional 1 min to remove residual wash buffer.**

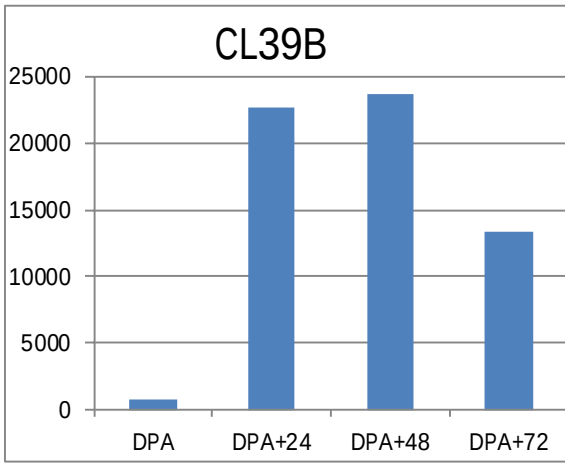
   Important: Residual wash buffer will not be completely removed unless the flow-through is discarded before this additional centrifugation. Residual ethanol from Buffer PE may inhibit subsequent enzymatic reactions.
10. **Place the QIAprep 2.0 column in a clean 1.5 ml microcentrifuge tube. To elute DNA, add 50 µl Buffer EB (10 mM Tris-Cl, pH 8.5) or water to the center of each QIAprep 2.0 spin column, let stand for 1 min, and centrifuge for 1 min.**

Appendix 4 YFP measurement data

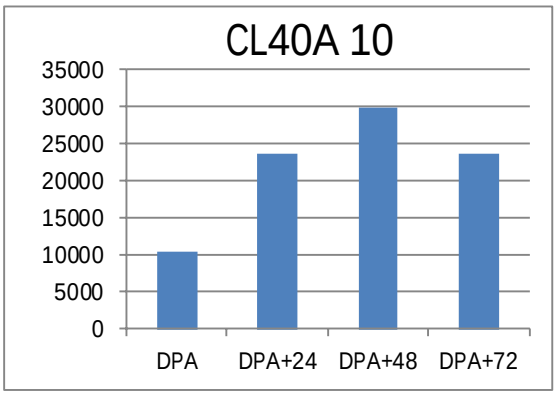
CL39A	yfp
DPA	326
DPA+24	18866
DPA+48	20201
DPA+72	19295



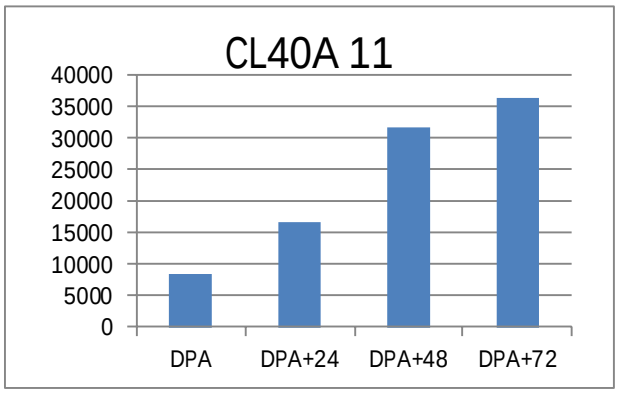
CL39B	yfp
DPA	738
DPA+24	22620
DPA+48	23636
DPA+72	13391



CL40A 10	yfp
DPA	10404
DPA+24	23628
DPA+48	29994
DPA+72	23778

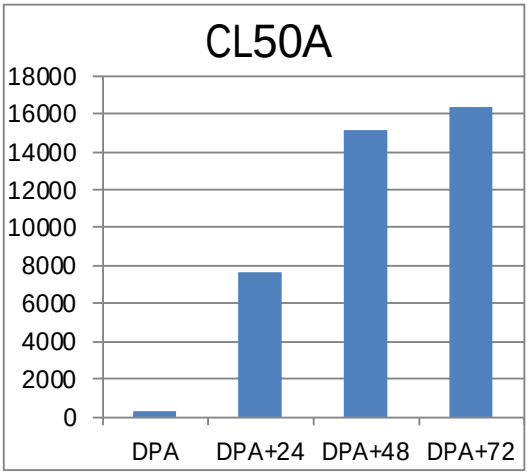


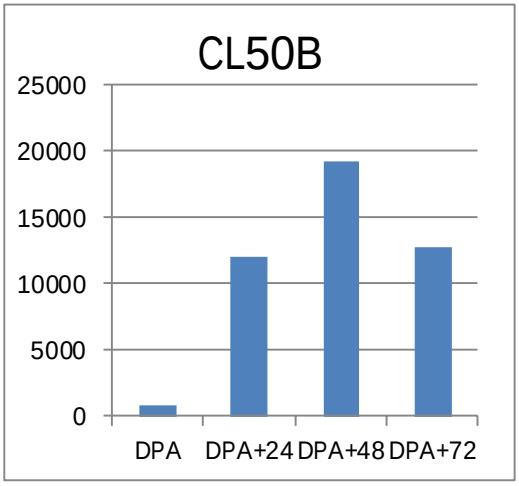
CL40A 11	yfp
DPA	8474
DPA+24	16568
DPA+48	31599
DPA+72	36228



CL50A	yfp
DPA	372
DPA+24	7628
DPA+48	15113
DPA+72	16327

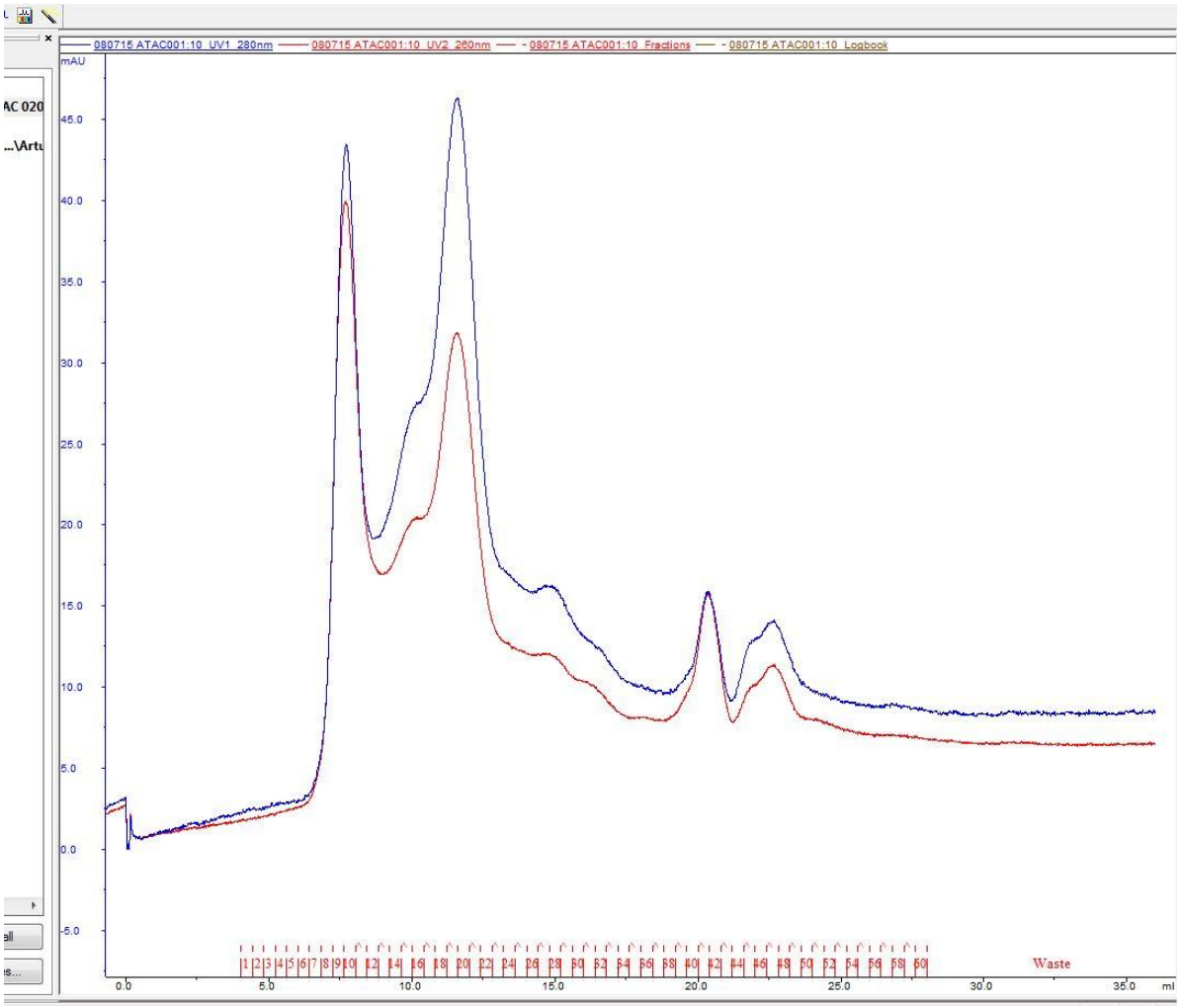
CL50B	yfp
DPA	879
DPA+24	12111
DPA+48	19216
DPA+72	12838



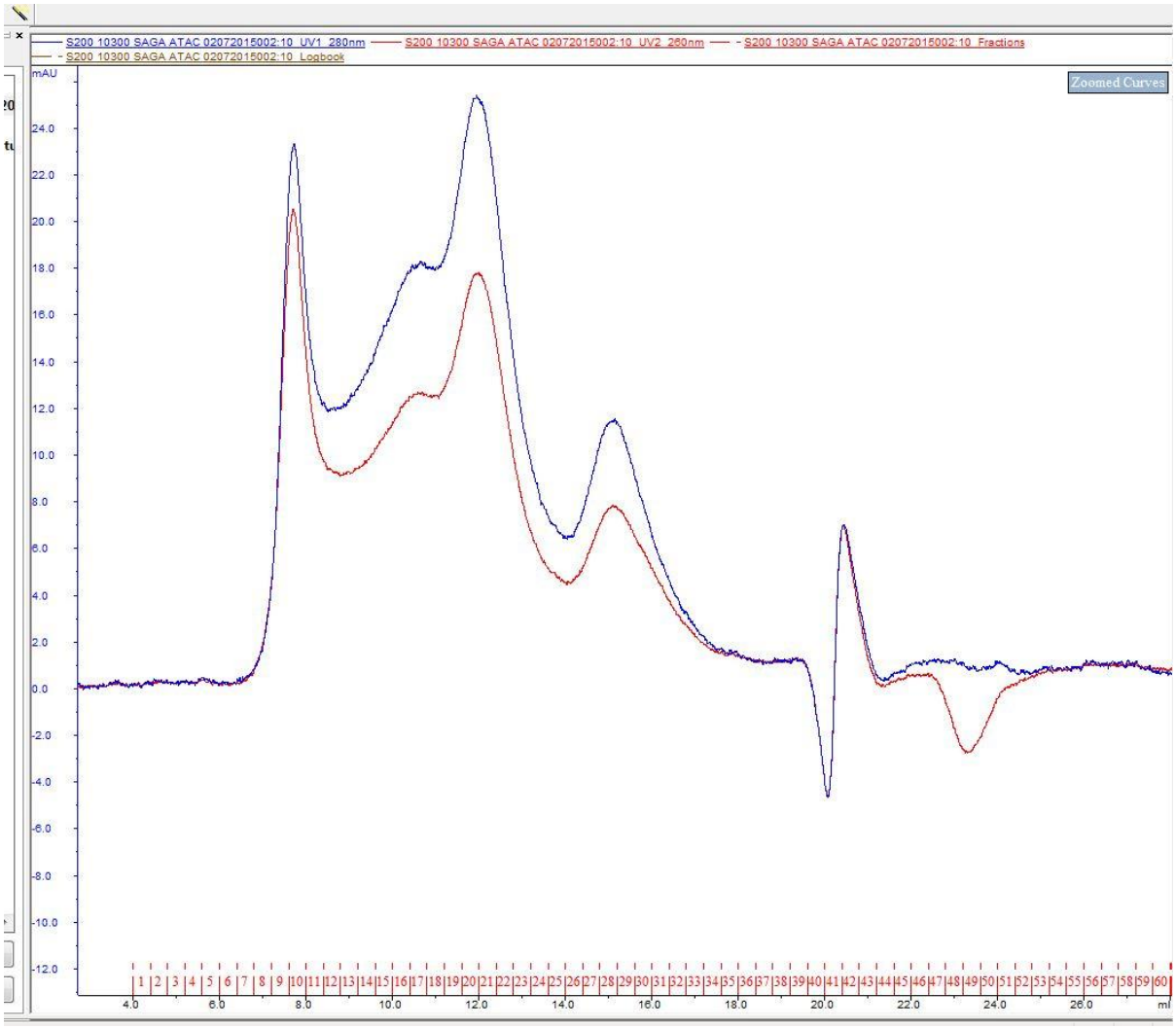


Appendix 5 SEC chromatograms of ATAC HAT complexes

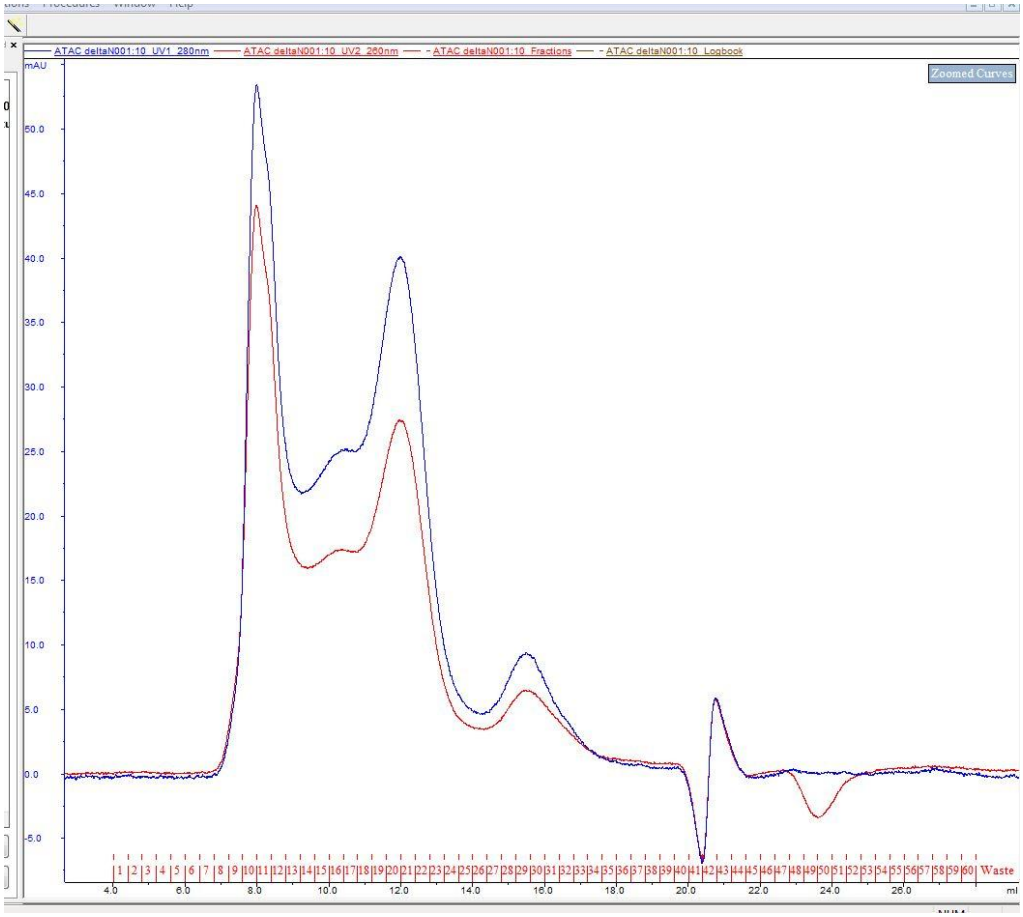
CL39, the full-length ATAC HAT



CL40, the miniHAT of ATAC



CL50, the N-terminal-truncated ATAC



Appendix 6 AKTA Calibration curve for S200 10/300 column

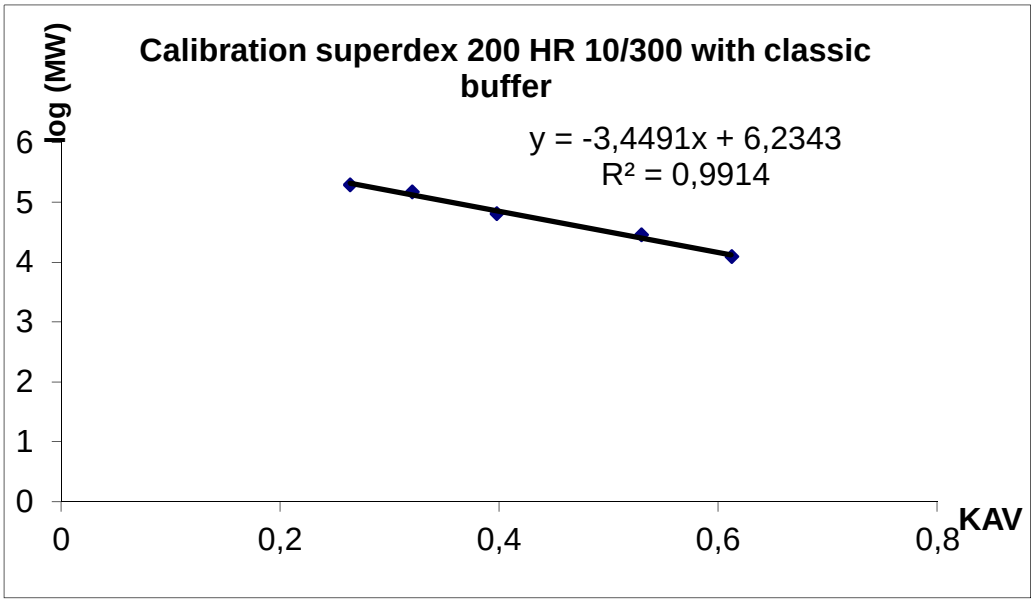
Calibration performed at 4deg on AKTA K, 09/10/09

Dead volume Vo (mL)	8.11
Total volume Vt (mL)	24

KAV= (Ve-Vm) / (Vt-Vm)

Buffer = 25mM tris pH 8.0, 250mM NaCl, 2mM betaMercap-toethanol

calibre	MW (Da)	Elution volume Ve	KAV	log (MW)
Dextran Blue	2000000	8.11	0	6.301029996
Beta amylase	200000	12.3	0.263687854	5.301029996
alcohl deshydrogen-ase	150000	13.2	0.32032725	5.176091259
BSA (bovine serum albumin)	66000	14.42	0.397105098	4.819543936
carbonic anhydrase	29000	16.52	0.529263688	4.462397998
cytochroma C	12400	17.83	0.611705475	4.093421685



y=ax+B a= -3.4491 B= 6.2343

Enter the molecular weight of your protein (Da) 216000 it should be eluted at 12.26 mL

Enter the elution volume of your protein (mL) 10 its molecular weight is 666890.65 Da