Hedmark University
of Applied Sciences
Campus Hamar, Faculty of Educational and Natural Sciences

MD Fakhrul Islam Faruque

Protein separation using Design of Experiments and structure prediction in mixed mode chromatography of potato protease inhibitors

Master's degree in Applied and Commercial Biotechnology

Master's Thesis2016

Consent to lending by University College Library YES ☒ NO ☐

Consent to accessibility in digital archive Brage YES ☒ NO ☐

ACKNOWLEDGEMENT

This thesis was written as a part of my master's degree in Applied and Commercial Biotechnology program at the Faculty of Educational and Natural Science at Hedmark University of Applied Science (HUAS), in the period between August 2014 and February 2016.

First and foremost, I would like to express my gratitude to my supervisor Professor Knut Olav Strætkvern for his continuous support, patience, motivation, and enthusiasm. His guidance helped me in all the time for writing of this thesis. A special thanks to Knut Olav for always keeping your door open and tolerating my "Silly" questions.

I really relished working on this thesis and have learned lots of new things, which I did not imagine when I started.

I would like to thank my parents, younger sister and friends back home in Bangladesh for supporting me always during my master's programme.

At last but not at least, I am very grateful to the staff at Campus Hamar and especially I would like to thank senior lab engineer Anne Bergljot Falck-Ytter for the help and guidance in the lab. Moreover, special thanks goes to Hoff Norske Potetindustrier SA (Brumunddal, Norway) for the providing us with raw materials.

Finally, I would like to thank all my classmates and teachers who provided good studying environment and motivation.

MD Fakhrul Islam Faruque

Hamar, February 15, 2016

ABBREVIATIONS

BSA	Bovine Serum Albumin
CAP	Clarified and acid treated potato fruit juice
CETI	Chicken eggwhite trypsin inhbitor
CV	Column volume
DoE	Design of Experiments
EPA	Eluted peak Area
FT	Flow through
IEC	Ion exchange chromatography
L	Ladder/ Molecular weight marker
LFR	Linear flow rate (cm h^{-1})
MLR	Multiple linear regression
MiMo	Mixed mode
MMC	Mixed mode cation
PCI	Potato carboxypeptidase protease inhibitor
PCPI	Potato cysteine protease inhibitor
PFJ	Potato fruit juice
PIs/ PPIs	Protease inhibitors/ Potato protease inhibitors
pI	Isoelectric point
pK	Protein kinase
PKPI	Potato Kunitz-type protease inhibitor
PSPI	Potato serine protease inhibitor
RSD	Relative standard deviation
TCA	Trichloroacetic acid
TIA	Trypsin Inhibitor Activity

Table of Contents

1. ABSTRACT

Mixed-mode chromatography is a chromatographic method in protein purification that utilizes more than one form of interactions between the stationary phase and the solutes. As being advantageous over other types of adsorption chromatography it has better salt-tolerant adsorption, unique selectivity, and facile elution by charge repulsion.

Potato fruit juice (PFJ) is a by-product from industrial starch manufacture and contains several proteins such as patatin, protease inhibitors and others. Several protease inhibitors (Kunitz-type inhibitors, Potato protease inhibitors I and II, potato cysteine etc.) are ubiquitous and they account for 1-10% of the total protein. Potato proteins, especially the protease inhibitors are a potential resource of proteins for potential pharmaceutical and nutraceutical applications. Recent advances in the isolation of major proteins from industrial potato fruit juice (PFJ) using novel chromatographic resins are opening new possibilities to develop value-added proteins. The aim of the work that has been presented in this thesis was to separate potato protease inhibitors using mixed mode chromatography and try to understand the aspect of interactions between a cationic mixed-mode (MMC) ligand and proteins via bioinformatics tools.

The protease inhibitor fraction was isolated from industrial PFJ by acidification and removal of precipitated patatin, before being loaded onto Capto MMC HiTrap-columns (1 ml and 5 ml). The ÄKTA Pure chromatography platform (GE Healthcare) was used in all of the experiments. The effects of pH and conductivity on the adsorption behavior of PIs from clarified and acid treated PFJ (CAP) properties have been investigated. For protein elution pH-shift, different salts (NaCl, NH_4Cl) and mobile phase modifiers (arginine, guanidine, lysine) were used to investigate the binding behavior of protease inhibitors such as hydrophobic interaction, repulsion etc. Responses were measured as eluted peak area (EPA) and by use of a trypsin inhibitor assay (TIA). The statistical methods used in experimental design were one-way and multi linear regression analysis (Design of Experiments). To predict the protein structure and surface characterization different bioinformatics servers (NCBI, PDB) and tools (BLAST, Phyre2) were used.

Maximum binding capacity of protease inhibitors on Capto MMC occurs at CAP sample pH 4.3 and column buffer pH at 5.5-6.0, and at conductivity ranges 10.0-11.5 mS/cm. The molecular weight of adsorbed PIs was in the range 3-22 kDa. Results indicated that optimal binding capacity for PIs on Capto Hi-Trap MMC happens at precipitation pH 4.3. The effects of column pH, sample conductivity, sample load and their interaction were significant.

Of the mobile modifiers, arginine acted to weaken electrostatic interactions by increasing the dielectric constant; this would lead to a uniform decrease in retention. An increased concentration (0.1M) of arginine showed that more recovery of aggregated protein especially protease inhibitors II group and PCPI (20 kDa).

By predicting protein surface revealed that mixed mode resin could be recognizing different hydrophobic surface features exposed on the protease inhibitors.

Last but not least, it has done to separate protease inhibitors by multimode chromatography and different condition effects can understand to binding behaviors of protein and further also suggested chemical interaction of these protease inhibitors should be studied.

2 INTRODUCTION

2.1 Protein separation by mixed mode chromatography

In mixed mode chromatography process, many form of interaction can be achieved between the stationary stage and solutes in a specific feed stream. In contrast with other types of chromatography, for example affinity chromatography as well as ion-exchange chromatography, mixed mode chromatography is advantageous because of its salt independent adsorption and also, unique selectivity (Zhao, Dong, & Sun, 2009).

Dilution and addition of salt, increases buffer consumption and causes the long processing time in the traditional ion-exchange and hydrophobic interaction chromatography whereas the mixed mode chromatography has remarkable advantages considering buffer consumption and processing time issues during the protein separation process (Simon C. Burton & Harding, 2001).

In most of mixed mode chromatography, multiple interactions of proteins simultaneously influence the adsorption of solute molecule (Sasaki et al., 1979; Sasaki et al., 1982). A new form of mixed mode chromatography was proposed in 1988, which is hydrophobic charge induction chromatography (HCIC) (S. C. Burton & Harding, 1998). The ligand is supposed to be uncharged at neutral pH when pKa applied for a specific HCIC is chosen. Therefore, at physiochemical conditions, adsorption is achieved solely with the help of hydrophobic interaction. However, the ligand can take on charges along with same polarity group as the protein by adjustment of the mobile phase pH (Zhao & Sun, 2007).

Ionic and hydrophobic interactions occur at the same time in mixed mode chromatography, which helps to increase the specificity as well as selectivity of the protein. By exploring screening methods that include multiple interaction (ionic, hydrogen bonding, hydrophobic) modes, it is possible to design MMC media for high salt concentrations; capture of specific proteins or pH tunable hydrophobicity (Kallberg, Johansson, & Bulow, 2012). The general mixed mode chromatography principle is presented in Figure 1.

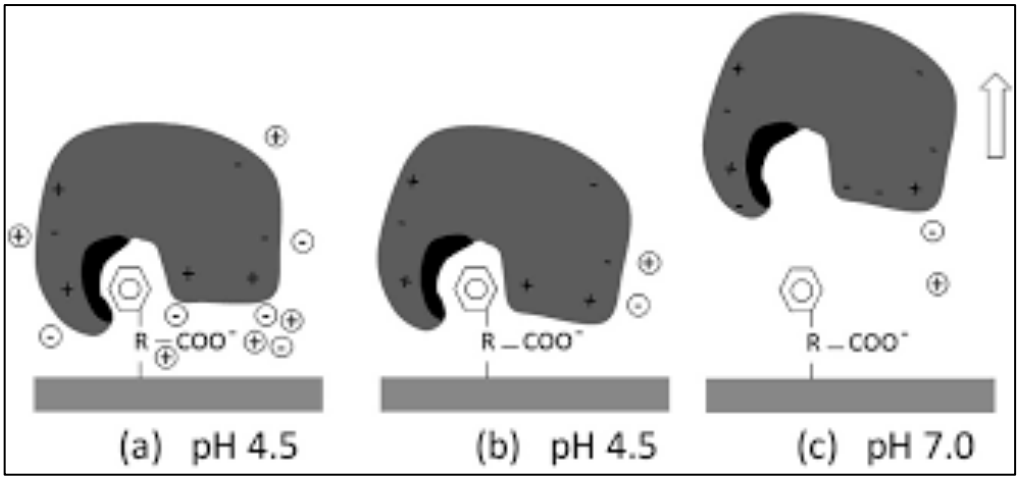


Figure 1: The binding and eluting strategy of target protein to chromatographic medium including a ligand between charge and hydrophobic group. (a) A target protein with a high salt concentration in a mobile phase is bound to the hydrophobic part of the ligand at pH 4.5 (b) The electrostatic part of binding becomes the dominating force upon decreasing the salt concentration at pH 4.5 (c) The increasing the pH up to 7.0, the protein will have same charge as the ligand and repulsion by it and finally eluted (Kallberg et al., 2012).

The electrostatic interactions play an important contribution to the protein adsorption process when an electrostatic engaging interaction exists between adsorbent and protein. Here the mechanism is that adsorption capacity first decreases until a minimum is reached while increasing the salt concentration, and then a further increase in salt concentration results in the increase of adsorption capacity (hydrophobic contribution). Due to the strong repulsion interaction at low pH, the retention factor decreases with decreasing pH value. The relation of salt concentration to retention factor is similar to the change of adsorption capacity under different salt concentrations (Gao, Lin, & Yao, 2008).

Mixed mode (MiMo) ligands for protein chromatography was first used by Yon and co-worker (Simmonds & Yon, 1976; Yon, 1972), who divided hydrophobic interaction ligands into two categories; the hydrocarbon groups and the mixed hydrocarbon and ionic groups. Hydrogen bonding close to the ionic groups in the particular mixed mode ligands could be favorable for protein binding at high salt concentrations (Johansson et al., 2003). According to this finding, the commercial MiMo-resins Capto™ MMC (cationic type) and Capto™ adhere (anionic type) were developed (Figure 2).

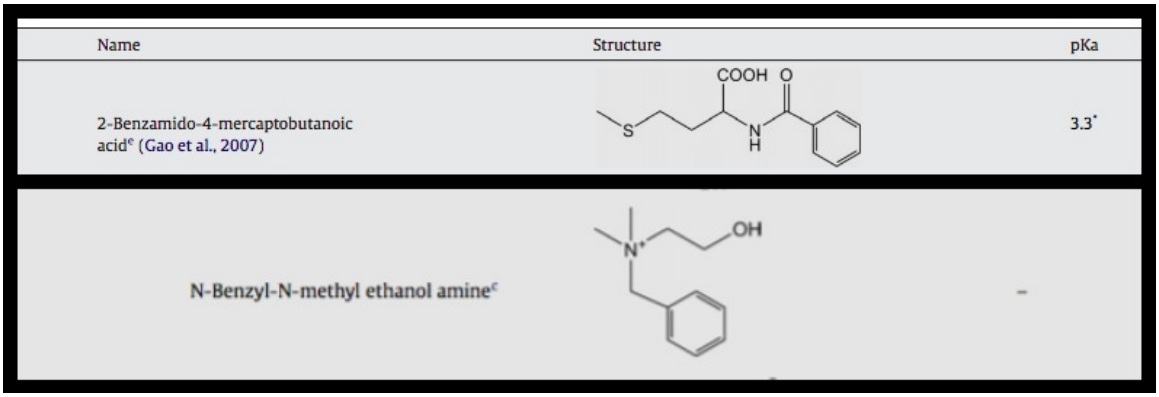


Figure 2: Ligand with negative charge for Capto ™ MMC and Streamline HST I (Gao, Lin, & Yao, 2007) and positive charge ion Capto™ adhere (Zhao et al., 2009).

Capto™ MMC is a weak cation-exchanger with an amide group of hydrogen bonding and phenyl group for hydrophobic interactions, while Capto™ adhere is a strong anion-exchanger with a phenyl group for hydrophobic interactions and a hydroxyl group for hydrogen bonding. In the new multiple interaction ligands, Capto™ MMC absorbents have high potential binding capacities in high salt and is thus termed as "salt tolerant adsorbent". Dissociation of adsorbed proteins from Capto™ MMC adsorbent can be gained by salt gradients as well as by a shift to high pH (Zhao et al., 2009).

2.2 Effects of salts on protein-adsorbent interaction

Whether inorganic or organic, salts play an utmost role in most of the column chromatography methods. They are well known to facilitate binding as well as elution of proteins, or suppress protein–surface or nonspecific protein–protein interactions (Tsumoto, Ejima, Senczuk, Kita, & Arakawa, 2007).

The effects of salts can be classified into two groups, (a) nonspecific and (b) specific effects. The nonspecific salt effects are simply identified due to their ionic properties. At any salt concentrations, salt ions contribute stoichiometric ion binding or charge shielding/electrical double layer on the charged proteins and column surface. At salt concentrations, salts exert specific effects; that is, their effects depend on the type of specific ions. The ions (Na^+, NH_4) influence the different elution steps (Staahlberg, Joensson, & Horvath, 1992).

If any one of the salts is to be used as a source of ions, NaCl could be a good

example. Since binding strength drops as the salt concentration rises, adsorption binding capacity also decreases; which help to prevent protein – protein interactions aggregation occurs (Melander & Horváth, 1977).

That increasing salt concentration causes decrease in the adsorption may be due to the shielding of the electrostatic field, and the adsorbed proteins are displaced by the counter-ions (Figure 3). Then the hydrophobic interactions between protein and adsorbent become more important due to the salting-out effect at higher salt concentration, and then an increasing adsorption capacity can be re-established at higher salt concentration.

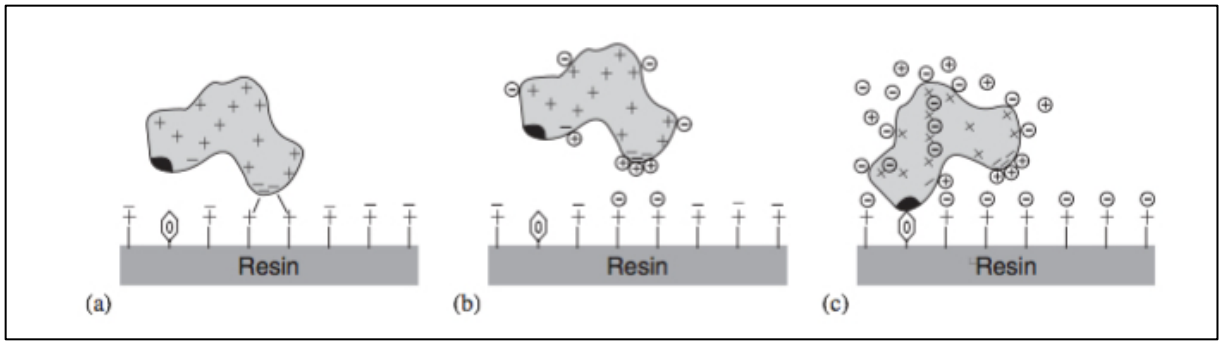


Figure 3: (a) Adsorption of the positively charged protein onto the like charged resin surface is completed in the contact region, where the functional groups are negatively charged; (b) the adsorbed protein is displaced by small ions; (c) adsorption is achieved at higher salt concentration due to the salting-out effect (Gao, Lin, & Yao, 2006).

Inorganic salts, such as NaCl and NH_4Cl are most frequently used because the majority of proteins can be eluted with these salts without adverse effects in purification of proteins (Janson, 2012). For different experiments with organic salts, for example arginine and guanidine; and various charged amino acids (e.g., sodium glutamate, Lysine monohydrocloride) or glycol are used (Huang, J.-X. & Guiochon,1989). Among the organic salts used, special attention has been given to arginine hydrochloride; simply called arginine (Figure 4), which usefulness as a mobile phase additive or modifiers in different column chromatographies is now widely identified (Ejima, Yumioka, Arakawa, & Tsumoto, 2005). Arginine is able to disrupt nonspecific interactions via suppression in ionic and also in hydrophobic interactions. Between column resin and protein, electrostatic interaction will be reduced when NaCl salts applied (Figure 4(A)). Oppositely, hydrophobic interaction will be enhanced by same salts effects (Figure 4(B)).

Organic solvents are opposite (compare A and B). Arginine is effective in suppressing both types of interaction (Figure 4).

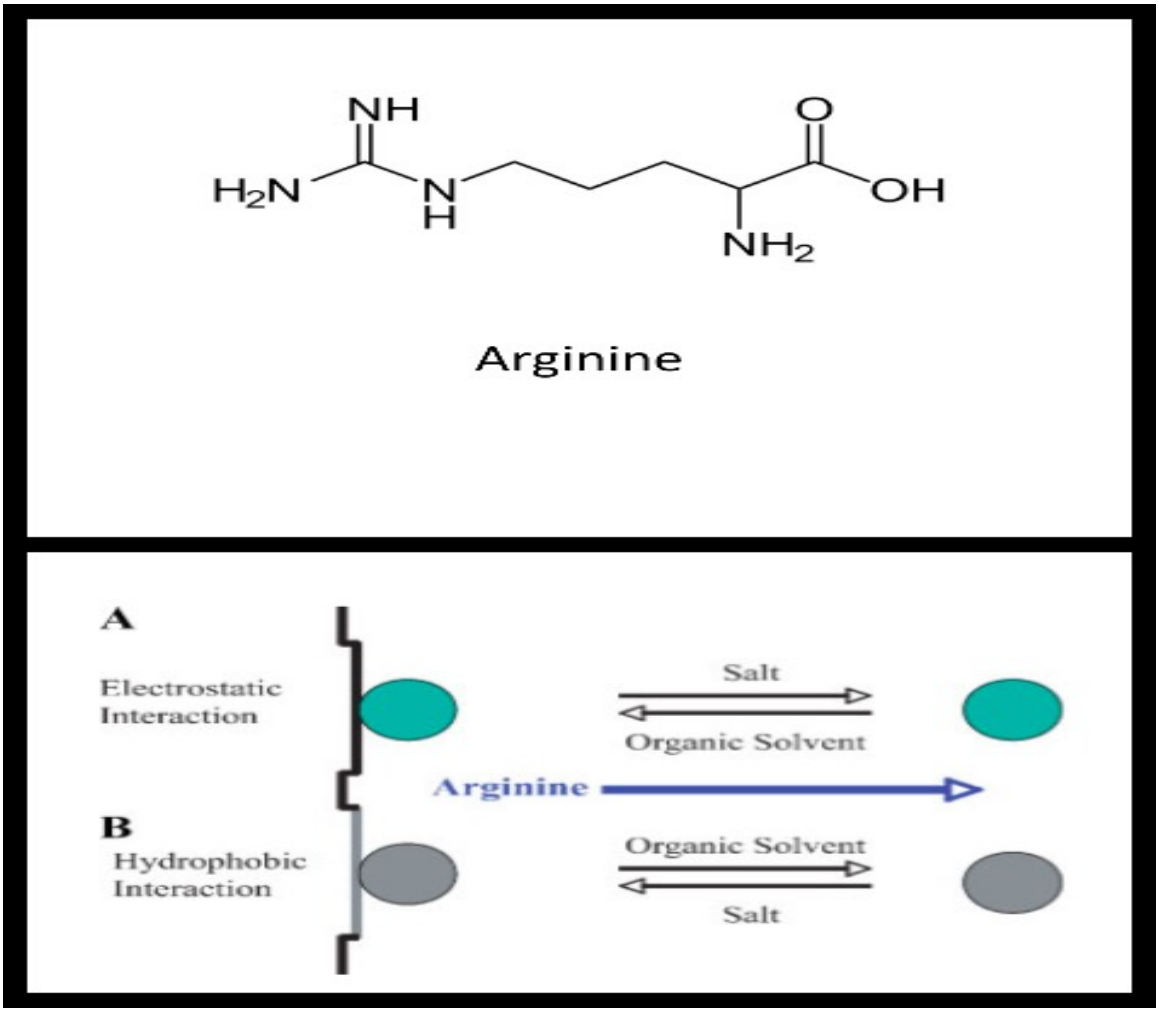


Figure 4: Schematic illustration of the effect of salts on nonspecific interactions of proteins with resin(Tsumoto et al., 2007).

Tsumoto et al. (Tsumoto et al., 2007) have observed that inclusion of arginine in the loading samples increases the recovery and decrease the aggregated contents of the eluted proteins, presumably due to the ability of arginine to suppress aggregation. When the protein was loaded in the absence of argininge salt, the aggregate content in the eluate was low, presumably due to high protein concentration during loading, leading to aggregation (Holstein, Parimal, McCallum, & Cramer, 2012).

Another work by same group also showed that, increases of the arginine concentration in the column elution buffer results in higher resolution and increase the recovery of the proteins by suppressing nonspecific binding. The recovery of eluted protein is close to 100% when using arginine at low sample pH (e.g., 0.5– 0.7 M arginine at pH 4).

2.3 Potato as protein source

The six essential food nutrients for our body are carbohydrates, protein, fat, vitamins, minerals and water. So, protein is one of the important food nutrients among the six dietary food nutrients. Due to the current growth of food demand all over in the world, countries all over the world need to think about the increasing production and sources of most important food nutrients so that people can have the most important nutrients for their body. It is worth to think about some sources of those food nutrients. This thesis paper will focus on protein.

There are three major global crop plants; corn, wheat, potato, which are also sources of dietary proteins. To fulfill the need of protein nowadays consumers search for high quality of protein as well as low price of that source. According to the growth rates reported by the International Food Policy Research Institute (IFPRI) and the International Potato Center (CIP) the total amount of potato production all over the world will be 403.5 million tons in 2020 and the worldwide demand for potatoes for food; processing and animal feed is expected to increase by about 40 percent within this time (Kumar et al., 2012).

Regarding the quality and price criteria, proteins from potato has an advantageous position for its nutritional value. Different nutritional components are mainly proteins, amino acids, sugars, plant phenolics available in potato (Sun et al., 2013). Potato protein is also of interest for the 1-2% of the global population people with food allergies towards egg, gluten, soy, nut and fish since potato protein are less common to cause food allergies (Løkra, Helland, Claussen, Strætkvern, & Egelandsdal, 2008); (Castells, Pascual, Esteban, & Ojeda, 1986)).

The potato tuber contains approximately 2% (w/w) proteins which are roughly subdivided into 50% protease inhibitors, 40% patatins, and 10% high molecular weight proteins such as lectins, phenol oxidases, starch phosphorylase, protein kinase (Schoenbeck et al., 2013).

2.4 Composition of industrial potato fruit juice

Potato fruit juice (PFJ) is an aqueous by-product or side stream of starch manufacturing from tubers. The production rate of PFJ depends on the processing technology, ranging from 0.7 to 7 m^3/ tons (Natu & Mazze, 1991). One example of starch manufacturing from tubers is given Figure 5.

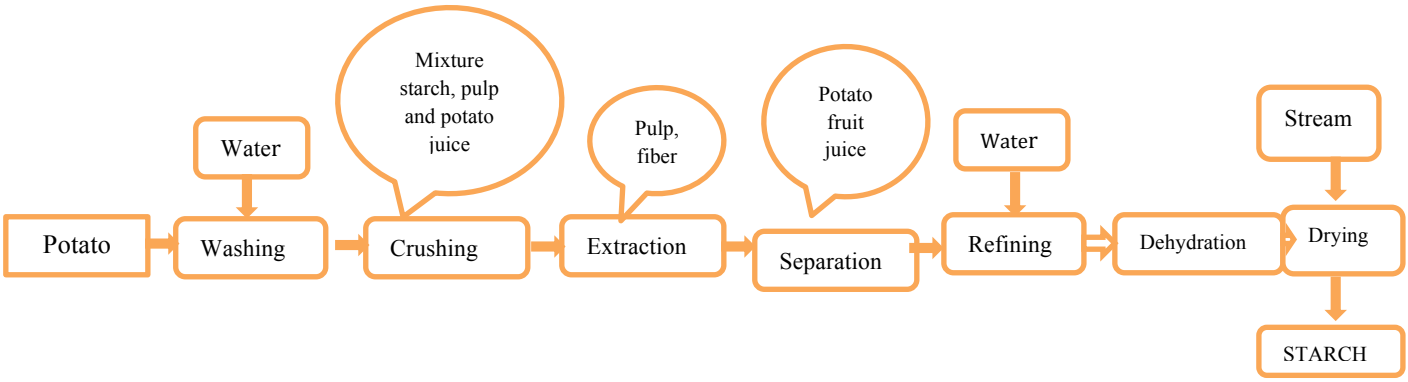


Figure 5: Potato starch manufacturing process at HOFF SA, Brumunddal, Norway. Potato fruit juice is obtained as the aqueous phase after separation by centrifugation of the denser starch material (Adapted from Løkra and Strætkvern, 2009).

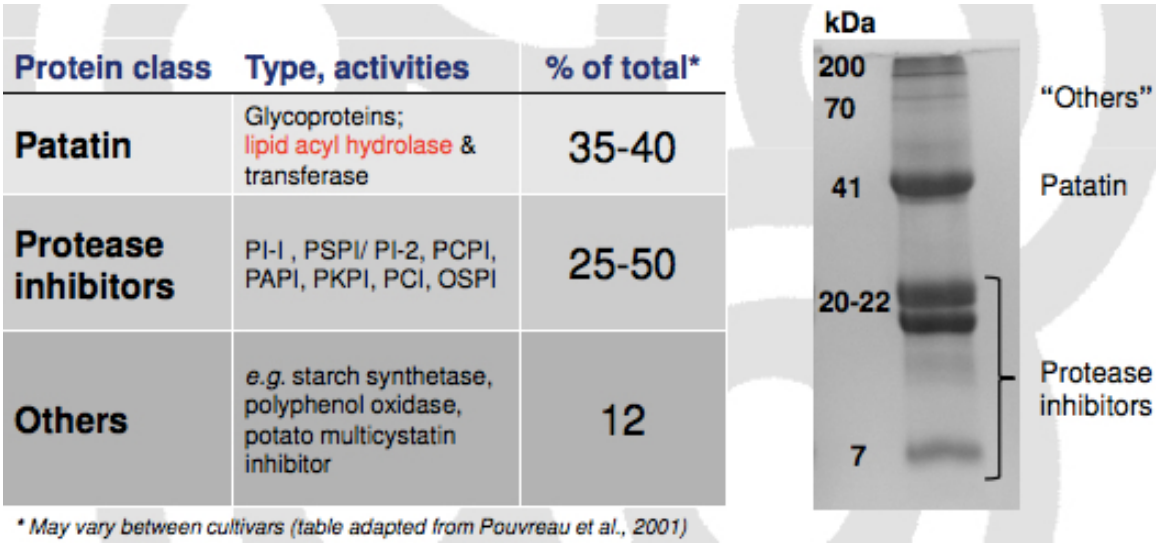
The PFJ is obtained as by-product from starch manufacturing at pH 5.0 to 6.0 and it contains about 5% of dry matter while containing around 22 to 27 % protein (Bártová & Bárta). Average composition of PFJ at dry matter condition is present in Table 1.

Table 1 – Average composition of PFJ (Dry matter) (Adapted from (Koningsveld, 2001)).

Components	Concentration in PFJ (g/l) (Min- Max)	% Of dry matter
Protein (N*6.5)	13.4 (8.5- 22.2)	26.8
Peptides (N*6.25)	2.2 (1.5-3.1)	4.4
Amino acids + amides (N*5.13)	4.8 (3.3 – 7.8)	9.6
Other N-containing compounds	0.9	1.8
Sugars	7.9 (3.0-24.9)	15.8
Lipids	1.1	2.2
Citric acid	5.0 (2.0-12.0)	10.0
Ascorbic acid	0.3 (0.1-0.6)	0.6
Other organic acids (e.g. malic and pyrolidone carboxylic acid)	1.3 (0.7-5.4)	2.6
Chlorogenic acid	0.2 (0.1-0.5)	0.4
Caffeic acid	0.07(0.03-0.3)	0.1
Potassium	5.6 (3.9-7.3)	11.2
Phosphorus	0.5 (0.2-0.9)	1.0
Other components	5.0	10.1

2.5 Main protein groups in potato

There are three classes of protein available in potato which are patatin, 40-42kDa of glycoproteins; potato protein inhibitors (Spelbrink, Gerrits, Mooij, & Giuseppin) 3-23kDa and others (mostly high molecular weight) (Figure 6).



Protein class	Type, activities	% of total*
Patatin	Glycoproteins; lipid acyl hydrolase & transferase	35-40
Protease inhibitors	PI-I , PSPI/ PI-2, PCPI, PAPI, PKPI, PCI, OSPI	25-50
Others	*e.g.* starch synthetase, polyphenol oxidase, potato multicystatin inhibitor	12

** May vary between cultivars (table adapted from Pouvreau et al., 2001)*

Figure 6: Gel electrophoresis analysis of industrial Potato Juice (KO, Strætkvern; personal communication).

Patatin, consists of a family of 40-42 kDa glycoproteins with pI's between pH 4.5 and 5.2((Pots, A. M. 1999);(Straetkvern, Schwarz, Wiesenborn, Zafirakos, & Lihme, 1999)). It is considered a storage protein because of its high accumulation in the tuber (Rexen, 1976). In aqueous solution, it is mainly active on phospholipids, mono acylglycerols. Moreover, it is also active on galactolipids while less active on di-acyl and tri-acyl glycerols(Andrews, Beames, Summers, & Park, 1988). It has been suggested that this LAH-activity may have a role in the plant defense mechanism (Pots, A.M., 1999).

Plant protease inhibitors are found in tubers as well as in seeds (Pouvreau, L. et al. (2001a)). Leguminosae, Solanaceae, and Graminae are the main sources of protease inhibitors specially the serine protease inhibitors (Pouvreau,L., Gruppen, van Koningsveld, van den Broek, & Voragen, 2003). In potatoes, a ubiquitous range of protease inhibitor is expressed. Serine protease inhibitors have inhibitory effects on tumor cell growth (Table 2).

Table 2 – Relative Mass (w/w) Distribution of Protease Inhibitor Families in Potato Juice (Pouvreau, L. et al., 2001b; Pouvreau, L. et al., 2003).

Superdex fraction (separation by size in HIC)	Group	Proportion in PJ **(%)**
I (48.0%)	patatin	37.5
	PI-1	4.5
II (48.0%)	PI-2	22.3
	PCPI family	11.9
	PAPI family	5.9
	PKPI family	3.6
	OSPI	1.5
III (2.0%)	PCI	0.9
Recovery		
100%	98.0%	88.1%

2.6 PI's group in potato

A characteristic of protease inhibitors is that are they are small, cysteine-rich and heat-resistant proteins of 3-23 kDa. The large number of cysteine residues present in protease inhibitors results in the formation of large number of disulphide bonds, which are necessary to uphold the original peptide conformation of the inhibitor upon hydrolysis by a protease (Jongsma, 1995). The potato protease inhibitors can be classified into seven families, mainly according to the type of proteases inhibited (Pouvreau, L. et al., 2001b).

Potato inhibitor I (PI-1) is a serine protease inhibitor composed of five 7.7 -7.9 kDa iso-inhibitor protomers (Ralet & Guéguen, 2000). There are eight different forms of PI-1 found, which have pI's value between 5.1 and pH 7.8 (Pouvreau, L. et al., 2001b).

Potato Inhibitor II (PI-2) is a dimeric serine protease inhibitor composed of two types 10.2 kDa subunits those are linked by a disulfide bonds. The isoforms of PI-2 have pI's in the range from 5.5 to 6.9 and molecular weight around 20.5 kDa (Pouvreau, L. et al., 2001b).

Potato Cvsteine Protease Inhibitors (PCPI) has at least eight different subunits inhibitors present in PFJ and they differ in molecular weight, ranging from 20.1 to 22.8

kDa. They have pI's ranging from 5.8 to >9 (L. Pouvreau et al., 2001a). They constitute about 12 % of total protein (Rowan, Brzin, Buttle, & Barrett, 1990) they also constitute 16 % and 10 % of the total chymotrypsin and trypsin inhibiting activity present in PFJ, respectively (Pouvreau, L. et al., 2001a).

Potato Aspartyl Protease Inhibitors (PAPI) consists of 6 different inhibitors. Their molecular weights are in the range 19.9 - 22.0 kDa, while their pI's are in the range 6.2 - 8.7 and also comprise 9 % and 2 % of the total chymotrypsin and trypsin inhibiting activity present in PFJ, respectively (Pouvreau, L. et al., 2001b).

Potato Kunitz Protease Inhibitors (PKPI) has only two members, both having a molecular weight of around 20.2 kDa (Walsh & Twitchell, 1991). These subunits have pI's of 8.0 and > 9.0 respectively, and constitute around 4 % of total protein in PFJ. They represent 2 % and 3 % of the total chymotrypsin and trypsin inhibiting activity present in PFJ, respectively (Pouvreau, L. et al., 2001a).

Recent research work on Indian potato varieties, *Solanum tuberosum* and a heterodimer of Kunitz-type protease inhibitor was identified, which called **PotHg**, showing lectin like activity. It was confirmed by SDS and native PAGE analysis that the purified protein was a Kunitz-type serine protease inhibitor having two chains (15 kDa and 5 kDa) and that the protein was glycosylated (Shah, Patel, Pappachan, Prabha, & Singh, 2015).

Potato Carboxypeptidase Inhibitor (PCI) present in only one form in PFJ. PCPI has a molecular weight of 4.3 kDa and represents about 1 % of total protein in PFJ.

Other Serine Protease Inhibitors (OSPI) are represented by two members in PFJ and they represent 1.5 % of total protein. Their molecular weights are 21.0(Valueva, Revina, Kladnitskaya, Mosolov, & Mentele, 1999) and 21.8 kDa (Suh, Peterson, Stiekema, & Hannapel, 1990) and their pI's are 7.5 and 8.8, respectively. They comprise 2 % and 3 % of the total chymotrypsin and trypsin inhibiting activity present in PFJ. Likewise, in potato tubers, PI-1, PI-2, and PCI, and other protease inhibitors have been recognized (Table 3).

Table 3 – Composition of Protease Inhibition Activity groups in PFJ after potato juice fermentation separated by various chromatography steps. Adapted from (L. Pouvreau et al., 2001a).

Name	Chromatography fractions	MW	pI	Subunits
PI-1	I	7.683-7.873	5.1-6.3	5
	IN	7.683-7.873	7.2,7.8	5
	IIC2, IIC3	7.683-7.873	5.1-6.3	5
PI-2	IIA2, IIB1	20.279	6.5	2
	IIB2	20.023	6.0	2
	IIC2, IID3	20.273	6.1	2
	IIC4	20.674	5.8	2
	IID2	20.676	5.5	2
	IIE2	20.315	5.9	2
	IINA2, IINA3	20.265	6.9	2
PIG	IIC4	19.987	6.2	1
NID	IINB2	20.039	8.4	1
PDI	IIND2, IINE2	22.025	8.6	1
PI-8	IINC1	19.878	8.7	1
PI-13	IINC2	20.141	7.5	1
PAPI-8.15	IINA3	19.883	8.2	1
PCPI-23 kDa	IIA1	22.755	6.7	1
PCPI-6.6	IIC1	22.769	6.6	1
PCPI-5.9	IID1	22.674	5.8	1
PCPI-7.1	IINA1	22.773	7.1	1
PCPI-8.0	IINB2	20.096	8.0	1
PCPI-8.6	IIND1, IINE1	20.127	8.6	1
PCPI-9.4	IINF1, IING1	20.134	>9.0	1
PCPI-8.3	IINF2, IING2, IIIG3	20.433	8.3	1
PKPI-9.0	IINH1	20.247	>9.0	1
PKPI-8.0	IINH2, IING2	20.194	8.0	1
HLE inh.	IINE2	21.025	8.8	2
22 kDa inh.	IINA2, IINA3	21.804	7.5	1
PCI	IIIB	4.274	nd[c]	1

SU, subunit; PI-1, potato inhibitor I; PI-2, potato inhibitor II; PAPI, potato aspartate protease inhibitor; PCPI, potato cysteine; protease inhibitor; PKPI, potato Kunitz-type protease inhibitor; PCI, potato carboxypeptidase inhibitor; HLE, human leukocyte elastase; CarboA, Carboxy peptidase A; nd, not determined.

2.8 Bioinformatics for protein surface analysis

Different bioinformatics tools can do protein surface analysis. Currently approximately 68 bioinformatics enrichment tools are available for the data analysis (Huang, D. W., Sherman, & Lempicki, 2009).

The sequence similarity of the protein could be done using the BLAST program available from National Centre for Biotechnology Information (NCBI) (Altschul, Gish, Miller, Myers, & Lipman, 1990). Sequences normally used in various analyses can be downloaded from Uniprot database (http://www.Uniprot.org/) and also available in Protein database (PDB) using the keyword of different types of protease inhibitor (O'Leary et al., 2015).

The Protein Data Bank (PDB) is the worldwide source of 3D structures of proteins and nucleic acids. The PDB has large of data (> 100,000 structures) and related citations provide a well-organized test set for developing and understanding data citation and access metrics (Huang, Y.-H., Rose, & Hsu, 2015). Some examples of these protease inhibitors already established in PDB database are represented in the following Figure 7.

Pymol and Phyre2 are most common tools for protein surface analysis. Though PyMOL can be quite effective for certain visualization tasks, it has a lot of "rough edges" and will frustrate you at times with its limitations and complexity. Pymol normally used to for viewing hydrogen bonding of ligand with protein (Figure 8).

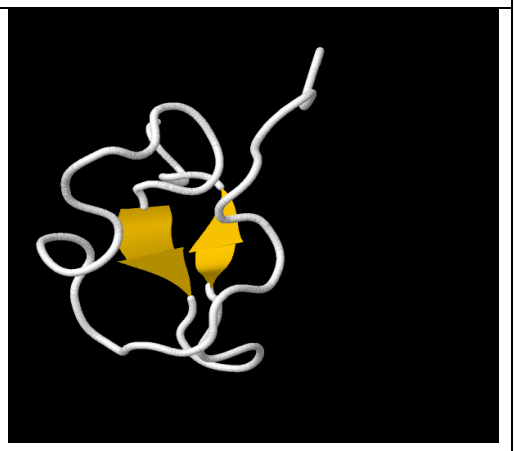
Potato cystatin	Kunitz type chymotrypsin inhibitors	Potato carboxypeptidase inhibitor (PCI)
		

Figure 7: Structures of different types of potato protease inhibitors established in the Protein Data Bases (PDB).

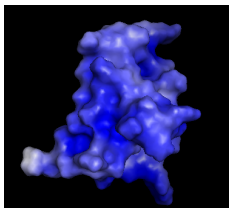
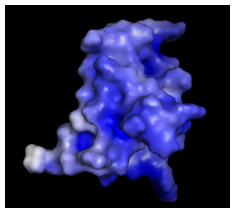
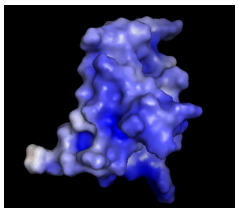
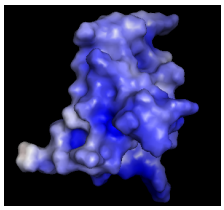
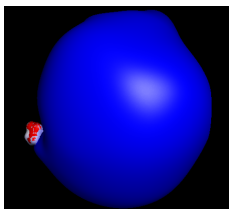
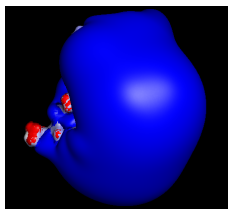
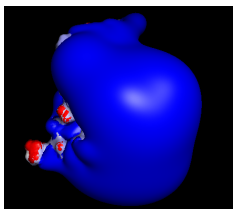
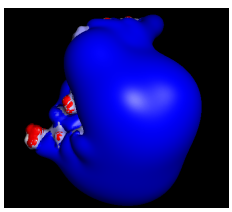
Electrostatic potential map displayed on the surface of chymotrypsin inhibitor				
pH	3.5	4.0	4.5	5.0
Net charge	+10.8	+9.94	+9.01	+8.48
Electrostatic equipotential surfaces for the chymotrypsin inhibitor				

Figure 8: Electrostatic potential map on the surface of protease inhibitors (Kunitz type chymotrypsin inhibitor, 1QH2.pdb) as a function of pH (Chernakova. E, master, 2011).

With the help of phyre2 bioinformatics tool, where one can submit data sequence and predict the tertiary structure of a protein sequence uses the principles and techniques of homology modeling.

2.9 DoE in protein production and purification

Design of experiment (DoE) is a technique for drafting experiments and surveying the information observed. The technique provides several simultaneous experimental factors with a minimum number of experiments, to gain sufficient information at a time. Supported on the obtained data, a bio-statistical model is built for the studied process (e.g, a protein purification scientific experiment or an individual chromatography step). Current statistically software is used to innovate the experimental designs, to obtain an individual model, and also to imagining the result information. DoE approach can improve the experimental conditions in a protein research lab likes, protein separation, study of protein stability, optimization of process, or robustness (i.e correctness) testing (GE Healthcare, 2014).

Optimization of protein purification is a multifactorial activity, the factors together that can play important roles are namely; pH vs ionic strength, flow rate with sample concentration, conductivity vs sample load, temperature vs pH. DoE is a proper technique to identify significant factors and find an optimum for the process (HEALTH, 2014).

- Fundamental research, many parameters of unknown impact
- Development of new products and processes
- Improvement of existing products and processes
- Optimization of quality/performance/cost/time
- Screening to quickly identify the most important factors
- Robustness testing of scientific results, processes, and products

2.10 Objectives of the study

From the recent publication on potato proteins recovery it become clear that the isolation of a high quality as well as functional protein from industrial PFJ is difficult. It requires further investigation on the behavior of the PFJ components under different conditions. However, the advances in the isolation and separation of the major PPI from PFJ are opening possibilities to establish new value added products.

Adsorption chromatography is a promising separation technique to capture PP from a waste effluent of starch manufacture using EBA chromatography which was first demonstrated by Strætkvern and co-workers (Strætkvern, Løkra, Olander, & Lihme, 2005). Research on PFJ also investigated different mixed mode resins to show the possibility of isolate PPI in separated fractions due to manipulation of binding and elution conditions. Can the behavior of the protease inhibitors of the mixed mode column be explained by their surface determinants? Can the behavior of multiple interaction point be identified by surface topology? And can be explained behavior of mixed mode column?

In Multimode chromatography by using Capto MMC column there is a possibility to isolate PIs in separated fractions by manipulation of binding and elution conditions. Large-scale separation can be effectively accomplished by using DoE to screen optimum binding conditions of PIs.

Assay methods for identification and characterization of PIs included 10% Tricine SDS-PAGE and differences between activity and size of fractions performed by TIA (Trypsin inhibitor activity). The main aim of the work was to develop a chromatographic capture process for the total protein fraction on MiMo resin.The objectives of the work included study on:

- Isolation and fractionation proteinase inhibitors from CAP
- Capture PIs with Capto Hi-Trap MMC resin in mixed mode chromatography
- Investigating effect of pH on the protein binding
- Screening for optimum binding conditions of PIs with different levels of sample loading, and conductivity

- Testing salts on elution profile for binding effect
- Bioinformatics analysis for trying to predict surface model

3. MATERIALS AND METHODS

3.1 Materials

3.1.1 Chemicals

Buffers and salts (ammonium phosphate, sodium chloride, argenin, guanidine etc.) for chromatography experiments were supplied by general lab suppliers. Fine chemicals like trypsin, azocasein, acrylamide, glycerol, bovin serum albumin etc. used for different analytical methods were purchased from suppliers (BioRad, California 94547 USA and from Sigma Aldrich, Spruce St. St. Louis).

3.1.2 Potato fruit juice: collection and pretreatment

Fresh potato fruit juice (PFJ) was collected from the commercial potato starch plant of HOFF SA in Brumunddal, from one production date during the fall 2014 starch campaign. PFJ sampled from the process had pH 5.7. In the lab, the juice was added 36% (W/V) citric acid in the proportion of 1:0.004 to prevent the enzymatic activation of phenolic compounds by polyphenol oxidase. After that, it was frozen in blocks (4-5 L) and stored at – 20 ^{o}C.

After thawing to 15 ^{o}C of a sample, PFJ was adjusted down to pH 4.3 with 25% sulphuric acid (approximately1.5-2.0 ml per 500 ml), which raised the conductivity of PFJ from ca 8-9 to ca 10-12 mS/cm (varied in different experiments). On a rocking table, the acidified PFJ was stirred for 30-40 minutes at 4^{o}C t to precipitate most of the patatin protein. Then, the sample was centrifuged at 3,500 x g for 15 minutes at 10℃ (Allegra 25R, BECKMAN Coulter, rotor code 5-5.1). The supernatant was adjusted further to pH 5.5, 6.5 and 7.5 by adding 5M NaOH (approximately 0.5 to 2.0 ml per 200 ml). The supernatant at different pH conditions was identified as clarified acid-treated PFJ (CAP) and retained at -20℃ again until further use. A typical workflow for obtaining clarified PFJ from crude sample is given in Figure 9.

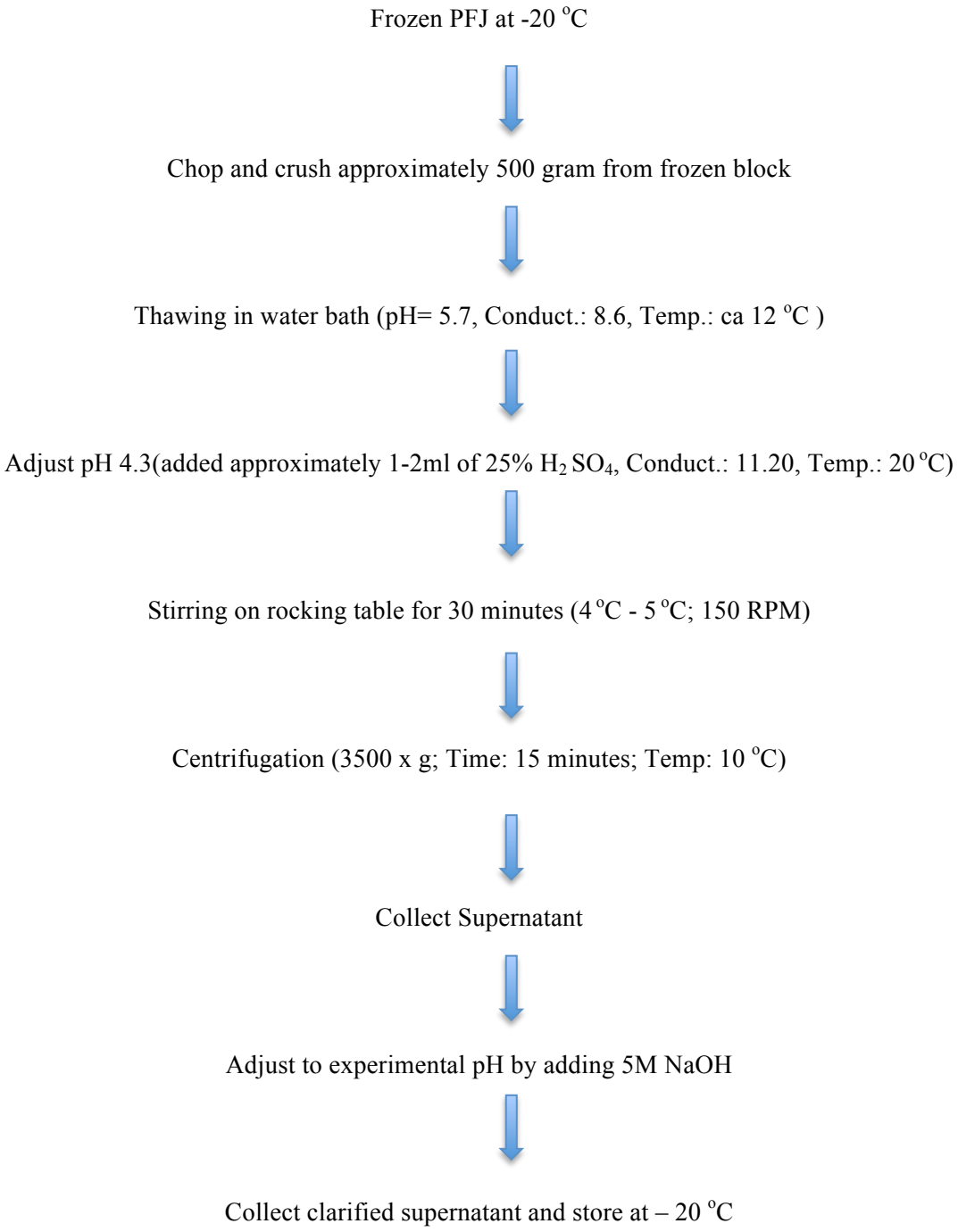


Figure 9: Preparation procedure of CAP from crude PFJ.

3.1.3 Chromatography Columns

The Capto Mixed Mode Cation (MMC) resin (Capto™ MMC – Hitrap™ 1 ml; Capto™ MMC – Hitrap™ 5 ml; Capto™ MMC – HiScreen™ 5 ml) used in all lab-scale experiment was supplied by GE Healthcare (Uppsala, Sweden). This MMC resin provides multiple binding sites with the protein surfaces (Figure 10). The ligand interacts as a cation exchanger while hydrophobic interaction is obtained through hydrogen bonding groups closer to protein surface. The characteristics of the MMC adsorbent are represented in Table 4.

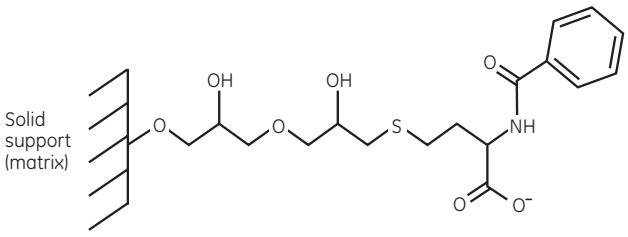


Figure 10: Capto MMC ligand structure (GE Healthcare -Data file11-0035-45 AB).

Table 4 - Some Characteristics of Capto MMC adsorbent (GE Healthcare -Data file11-0035-45 AB).

Characteristics	Value
Matrix Properties	High cross-linked Agarose
Functional group	Multimodal weak cation exchanger
Total ionic capacity	0.07-0.09 mmol H^+/ml medium
Particle size	75 µm
Type of exchanger	High salt tolerant
pH stability	2-12(14)

The sample was injected to the column either through a 5 ml capillary loop. To the column outlet as connected on-line detector and probes for UV 280 nm, conductivity and pH. Scale up experiment was carried out with 50 ml superloop loading 23.5ml sample to a 5 ml column.

3.1.4 PD-10 Desalting Columns

For the convenient sample clean up protein and other macromolecules normally use PD-10 Desalting columns. It generally used in experiments for desalting the PFJ. PD-10 desalting Column contained Sephadex G-25 medium, which allowed rapid group separation of low molecular weight substances to high molecular substances. For desalting of the sample prior to loading, two types of columns were used; one of them is the gravity drip PD-10 desalting column and the Hi-Trap Desalting column (5.0 ml). In each desalting run, 2.5 ml of sample was injected and eluted with 3.5 ml of loading buffer for the next chromatography step.

For desalting of the sample prior to loading, two types of columns were used; one of them is the gravity drip PD-10 desalting column and the Hi-Trap Desalting column (5.0 ml). In each desalting run, 2.5 ml of sample was injected to the Hi-Trap Desalting (5.0 ml) columns.

3.2 Methods

3.2.1 Experimental work flow

Schematic illustration of experimental work is represented in Figure 11 by flowchart, which includes the development and evaluation of PP isolation process carried out by Mixed Mode Chromatography process.

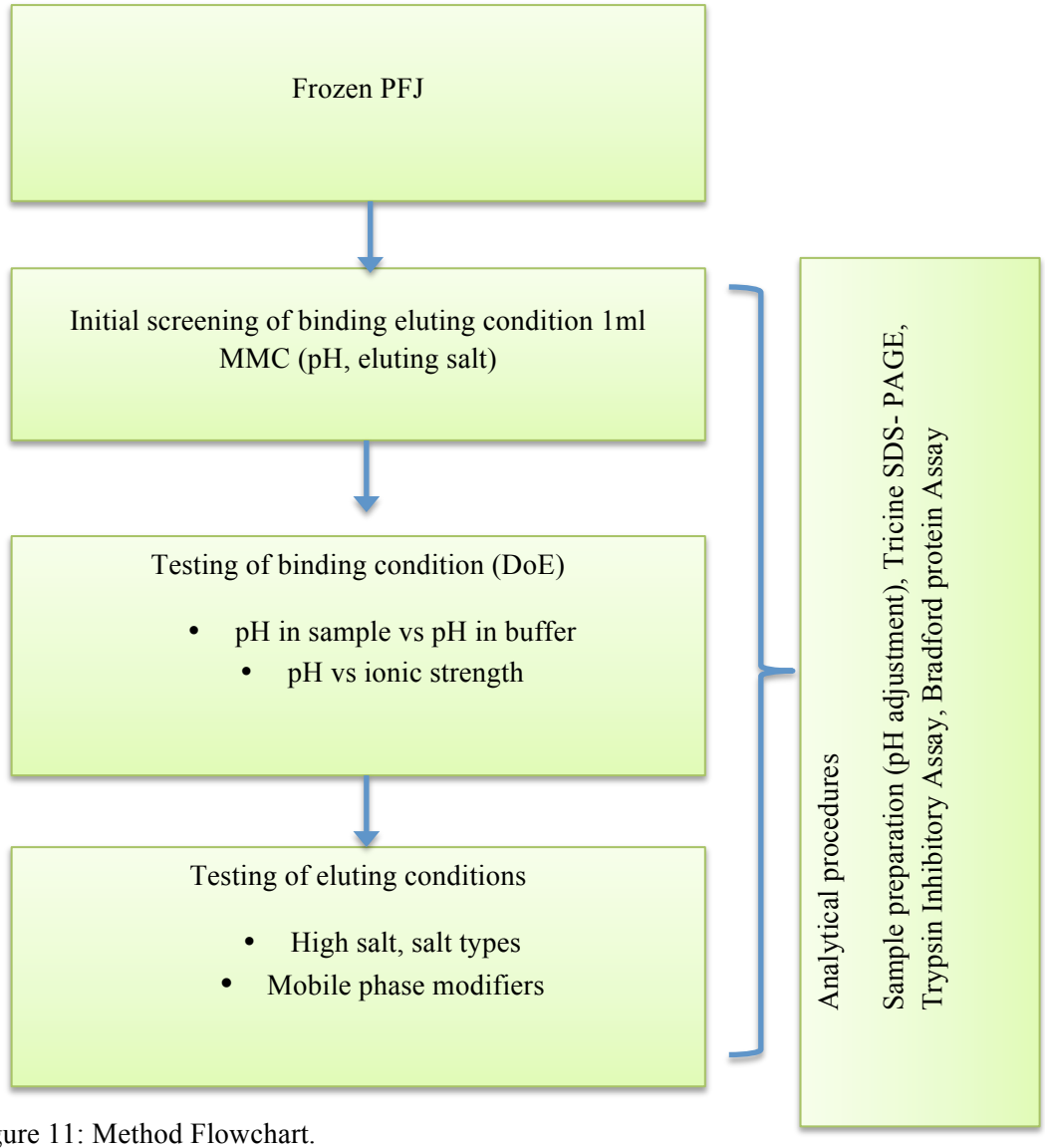


Figure 11: Method Flowchart.

3.2.2 Preparation of buffers

Different column equilibration buffers as well as buffer for the washing step, were based on 10 mM citric acid, adjusted to pH 4.5, 5.5, 6.0 or 7.5, and 20 mM sodium acetate pH 5.5, and was named as buffer A. Primary elution buffer B contained buffer A plus 1.5 M NaCl or NH_4Cl, whereas the secondary elution solution C was 20 mM NaOH

with pH around 11.9. Water used for experiments and for buffers was of ultra-pure grade (Direct-Q, Millipore, Billerica, MA, USA). Glassfiber filter was used for filtering of all buffers prior to chromatography [Schleicher & Schuell, 589^2 (Ø 55 mm)].

3.2.3 Explorative Mixed Mode Chromatography

Automated column chromatography was carried out with the AKTA™ pure chromatography along with Unicorn 6.3 software. A general workflow of the chromatography methods is represented in Figure 12.

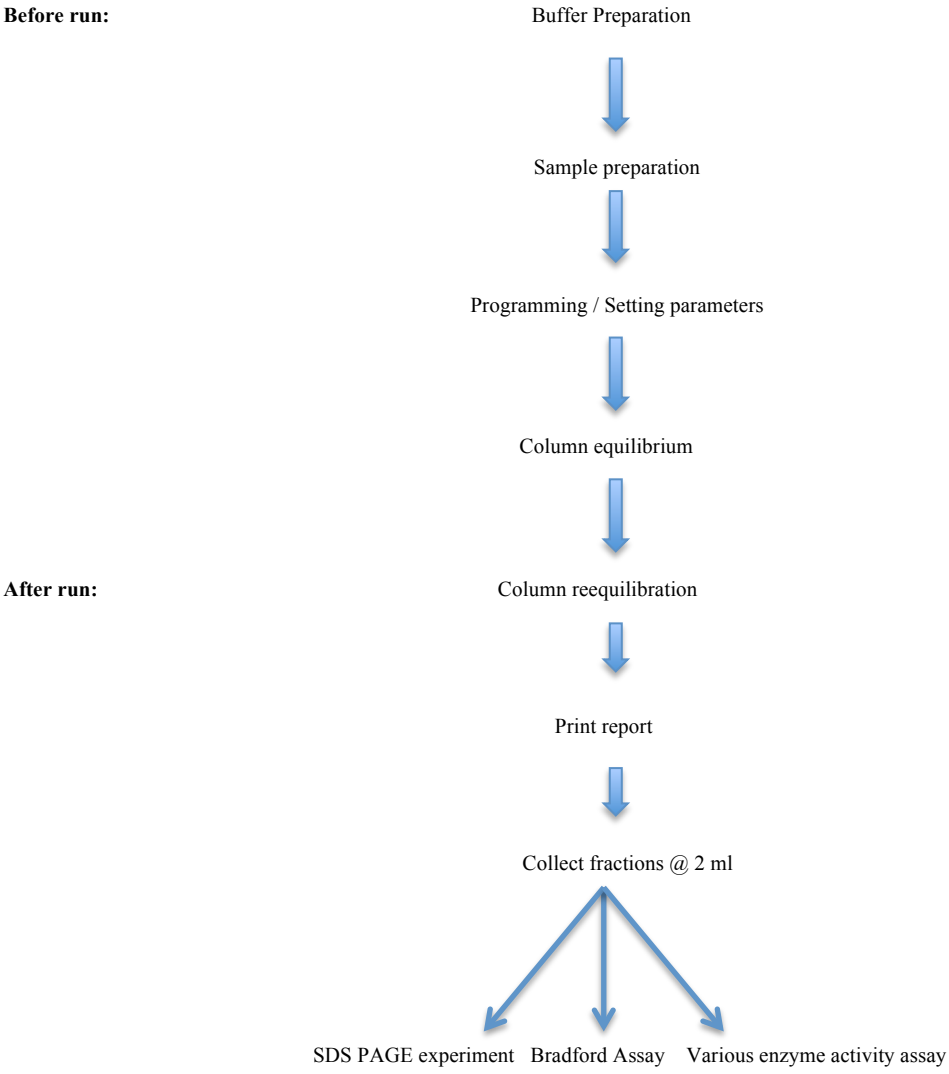


Figure 12: General workflow of the chromatography methods.

The schematic diagram of flow path of chromatography system is shown in Figure 13.

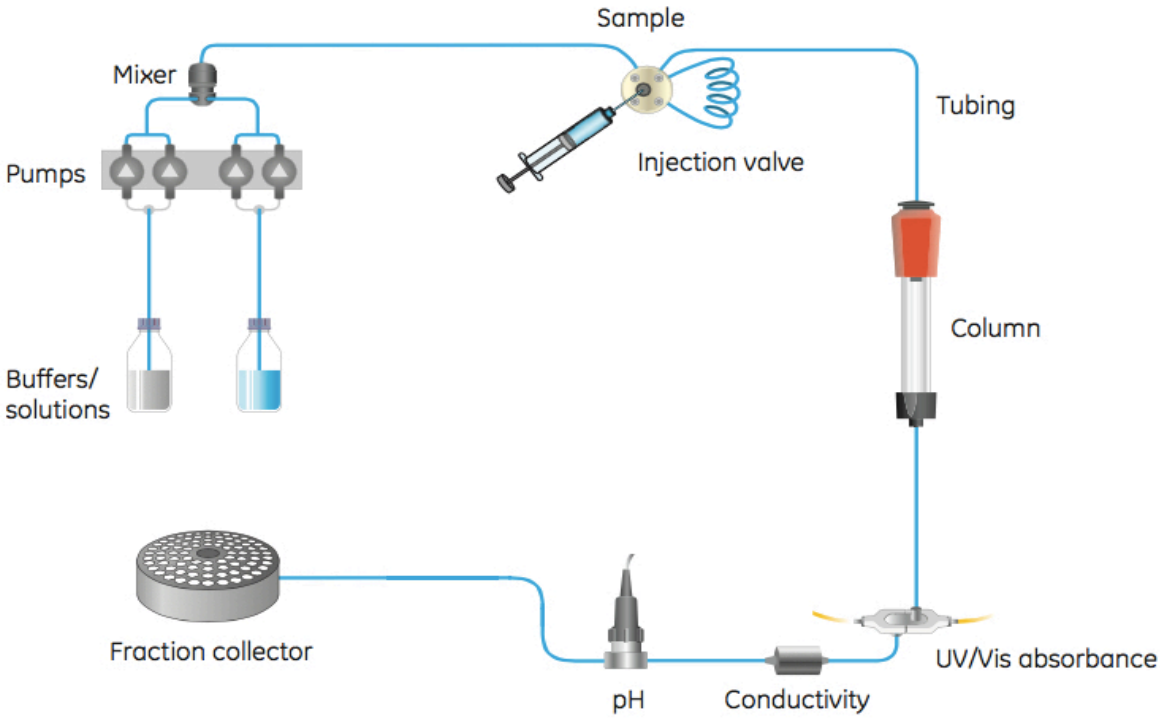


Figure 13: Typical flow path of a chromatography system (ÄKTA Laboratory-scale Chromatography Systems by GE Health care).

In the primary stages of experimentation CAP samples were applied to Hi-Trap MMC 1 and 5 ml columns (Table 5).

Table 5 - Characteristics of Hi-Trap column.

Parameters	Hi-Trap 1ml	Hi-Trap 5ml
Column volume (O'Leary et al.)	1 ml	5 ml
Column dimensions	0.7 × 2.5 cm	1.6 x 2.5 cm
Column hardware pressure limit	5 bar (0.5 MPa)	5 bar (0.5 MPa)

3.2.4 Bradford Coomassie protein concentration assay

Proteins concentration was determined by use of the Bradford Coommasic dye solution (Bradford, 1976). The dye interacts primarily with the sidechains of the aromatic amino acids Tyrosine, Tryptophan, and Phenylalanine in acidic solution to give maximum absorption at 595nm. The assay was using a calibration curve based on dilutions of a standard protein (Bovine Serum Albumin, 1.0 mg/ml) as viewed in Table 6.

Table 6 - Preparation of dilution series of Albumin (BSA) Standards

Tube number	1(x2)	2	3	4	5	6
BSA,µl (1 mg/ml)	0	10	20	30	40	50
Water,µl	50	40	30	20	10	0
Bradford reagent, ml	2.5	2.5	2.5	2.5	2.5	2.5

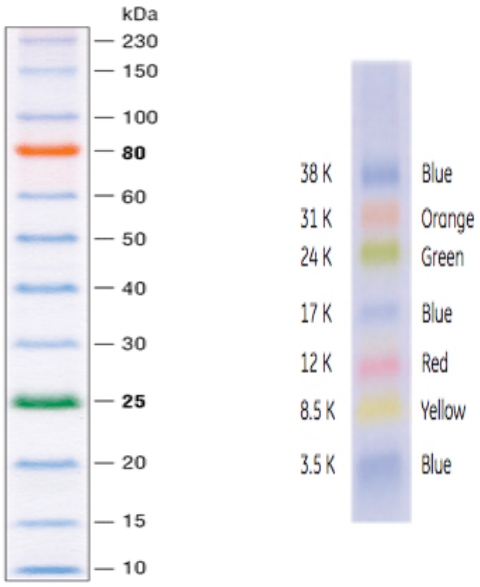
Bradford Coomassie reagent (2.5 ml) was added to each sample. Then, each tube was vortexed and incubated for 5 minutes at room temperature. After developing color of samples, read absorbance in a spectrophotometer (SHIMADZU, UV 1601/VISIBLE) at 595 nm. Unknown samples (crude and clarified PFJ) were measured in replicates and the protein concentration (mg/ml) determined from the calibration curve.

3.2.5 Tricine-SDS-PAGE electrophoresis

The electrophoresis method Tricine SDS-PAGE separates the proteins less than 100 kDa. Gels were composed of staking gel (4%) and separation gel (10% or 16%) (Schägger, 2006). Each sample was mixed with a non-reducing sample buffer at 1:4 ratios and incubated at 37 ^{o}C for 15 minutes and loaded at 10 µl samples per well. Ladder was loaded at 4 µl.

Electrophoresis was performed on ice to avoid excessive heating. Conditions were 35 V, 80 mA, 1 W (Amersham Biosciences electrophoresis Power Supply – EPS 3501 XL) until the samples entered into the stacking gel, then changed to 160V, 160mA, 10W, until the dye front reached the bottom of the gel. After electrophoresis, the gels were soaked for 40 minutes in fixing solution (50 % methanol, 10 %acetic acid, 100 mM ammonium acetate). Incubation with Coomassie® dye (0.025%) was performed for 1 hour 20 minutes until the protein bands became visible, and then the destaining solution

(10% acidic acid) was performed until the bands corresponding with proteinase inhibitors were visible. Two different ladders were used (Figure 14: ColorPlus Prestained Protein ladder Broad Range (10-230kDa) and the ECL™ Rainbow ™ Marker- low range RPN755E (3500-40000 Da)).



(a) (b)

Figure 14: Protein ladders use in Tricine SDS PAGE (a) ColorPlus Prestained Protein ladder Broad Range (b) ECL™ Rainbow ™ Marker- low range RPN755E (New England Biolab ®).

3.2.6 Trypsin Inhibitor Activity

The method developed by Spelbrink et al. (Spelbrink et al., 2011) was used for determination of protease inhibitor activity. A sample containing the protease inhibitor combined with a known amount of a protease enzyme and a chromogenic substrate (azocasein). Following incubation and acidic precipitation the remaining protease activity was detected by the end point absorbance using. To each 125 µl of diluted CAP series and chromatography samples, 25 µl of freshly prepared protease solution was added.. The protease solution was prepared dissolving 0.30 - 0.40 mg Trypsin (Sigma Aldrich 93610), in x µl of 1 mM HCL. As negative control (Tsumoto et al., 2007) 25 µl of water was added to 125 µl of sample and for positive control (PC) 125 µl of water mixed with 25 µl of Trypsin. Each reaction tube was added 225 µl Azocasein (3%) and incubated at 37°C

for 30 minutes. Then the reaction was arrested by adding 150 µl of 15% w/v Trichloroacetic acid (TCA) and pelleting of unhydrolysed substrate was performed by centrifugation at 15,000 x g for 10 minute at 4°C. Finally, 100 µl of the reaction supernatant was transferred to a 96-well microtiter plate, followed by addition of 100 µl 1.5 M NaOH to each well. The spectrophotometer (FLUOstar OPTIMA^{TM}) was calibrated against a Reaction Blank and the absorbance of released color from the substrate measured at 570 nm in replicate samples. The percent trypsin inhibition TI was calculated by the following equation:

$$\%TI = \{[A_{PC} - (A_{sample} - A_{NC})] / (A_{PC})\} \times 100$$

3.2.7 Design of Experiments (DoE) optimization

Design of Experiments (DoE) is a systematic way of changing process inputs and analyzing the resulting process outputs in order to quantify the cause and effect relationship between them as well as the random variability of the process while using a minimum number of runs. It's use in the fundamental research, development of new products, optimization of quality/ performance/cost/time etc. General overview of DoE workflow in UNICORN^{TM} is represented Figure 15.

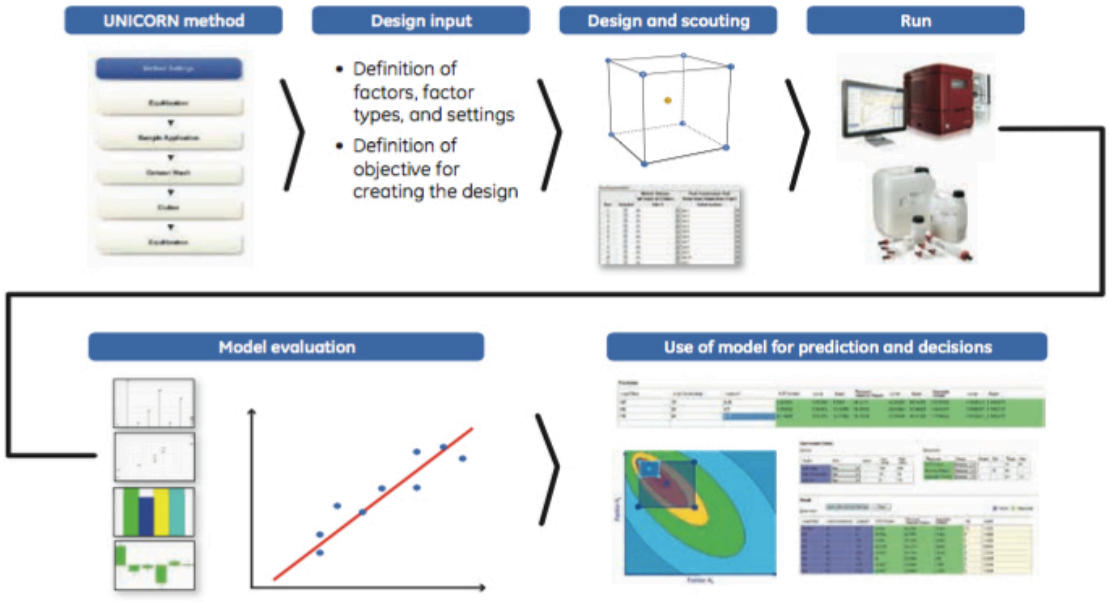


Figure 15: DoE workflow in UNICORN^{TM} (GE Healthcare life science - Design of Experiments in Protein Production and Purification).

Three DoE-processes were designed using the DoE module of the Unicorn 6.3 software (GE Healthcare, Uppsala, Sweden).

Based on a DoE model for scouting, factors were analyzed in three levels

Number of experiments investigation two variables or factors (k) at three levels (low, medium, high): $3^k = 3^2 = 9$

To increase the reliability of the model a middle point in the design space was also included. Thus, a total of nine experiments (3+3+3).

The model under DoE can be written as,

$$\text{Eluted peak area} = \beta_0 + \beta_1 \text{ Column pH} + \beta_2 \text{ Sample pH} + \beta_{12} \text{ Column pH} \times \text{Sample pH} + \epsilon$$

Where β_0 is the intercept and $\beta_1, \beta_2, \beta_{12}$ are called the coefficients. ϵ_i's stands for the error terms in the model (Quinn & Keough, 2002).Moreover, usually there is a hypothesis about the coefficients in design of experiment models which can be formulated as,

$H_0: \beta_i = 0;$ where i=0, 1, 2 and 12 in this model.

Vs

$H_1: \beta_i \neq 0;$

This RSD is a special form of standard deviation which tells how the regular standard deviation is a small or large quantity in comparison to the sample mean and it is represented as percentage. The formula for RSD is,

$$\text{RSD} = \frac{s}{|\bar{y}|} \times 100$$

Where, s=The sample standard deviation

$\bar{y}$=Sample mean

After RSD, there is another statistical measure provided by the experiments is the predictive power Q^2. Q^2 measures the predictive power of the model by using cross validation (a method where the model is fitted by omitting one or some specific number of observations each time)(Han, Kamber, & Pei, 2011) . Thus the PRESS is calculated for the model and after that the following formula has been used for

CalculatingQ^2:

$$R_{cv}{}^2 = Q^2 = 1 - \frac{PRESS}{TSS} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_{i/i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \text{; } Q^2 \leq 1$$

$$\text{Where, PRESS} = \sum_{i=1}^{n}\left(y_i - \hat{y}_{i/i}\right)^2 \text{ and TSS} = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

Here, PRESS is the cross-validated version of R^2 and the term $\hat{y}_{i/i}$ represents the predicted response (eluted peak area) for a particular model where i^{th} element has been omitted. When Q^2 has poor fit for the model and this is because the data have higher variation with fewer observations (Ballabio, Consonni, & Todeschini, 2007). Cross validation is a statistical procedure to calculate the predictive power of a model by omitting one or more observations from the model each time and for this study above the same has been done.

Two sets of DoE were applied: first a CAP sample with pH 4.3, 6.0 and 7.5 were applied to a Hi Trap MMC (1.0 ml) column, for loading the sample in to the instrument. Secondly, a CAP sample with pH 4.3, 6.0 and 7.5 and the conductivity 1.5,10, 20 ms/cm were applied to a Hi Trap MMC (1.0 ml) column. In post-run evaluation the response was measured as the amount of material bound to the column and eluted during gradient elution. The material bound was measured by peak integration and given as Eluted peak area(EPA) (ml x mAU).

3.2.8 Bioinformatics tools for protein surface evaluation

Bioinformatics databases use biological data to predict different structure for authentic evaluation. Different types of bioinformatics tools have been used for the modeling of surface protease inhibitors. Pymol and Phyre 2 software used to build up of

3D moding of 20KDa potato protease and proteinase inhibitors. A general wrok flow on bioinformatics work is presented Figure 16.

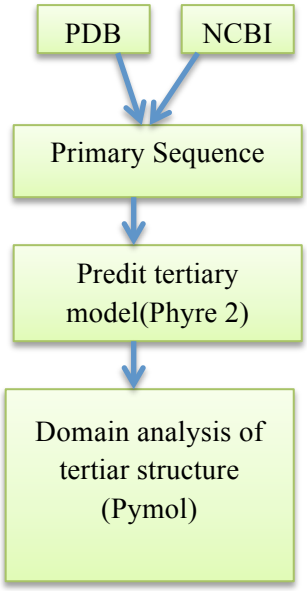


Figure 16: General works flow of bioinfortics work.

4 RESULTS

In the present work the aim was to isolate and purify protease inhibitors from CAP (clarified and acid treated PFJ) samples. In this study the results consist of an experimental part of chromatographic separation of potato protease inhibitors (PPI) and the second part dealing with in silico elucidation of the structure of the same.

4.1 Critical chromatographic parameters for protein binding

To optimize the loading condition of CAP several variables or factors has been used and they are:

- pH of CAP in the range of 4.3 - 7.5
- Conductivity of the CAP in the range of 10- 20 mS cm^{-1}
- Linear flow rate of 150-600 cm h^{-1}
- Loading ratio of CAP to column volume (5-10) x CV

The ranges of the values considered for the experiments were adapted and predicted from the earlier experiments with Mixed Mode Chromatography of PPI (KO Strætkvern, unpublished results).

4.2 Preparing PFJ sample for chromatography

The Capto MMC ligand provides several functional groups such as hydrophobic, electrostatic, hydrogen bonding interactions. Usually, the pH of a solution affects the charge states of the protein and also the resin. Thus, pH has a significant influence on the strength of electrostatic interaction between the protein and the ligand.

In this thesis, the main goal was to separate the protease inhibitors in the PFJ by use of MMC. To reduce the competing interaction from patatin (Chernakova. E, master, 2011), mild acidification was attempted. Patatin becomes less soluble at pH 4.3 or below (Pots, A.M., de Jongh, Gruppen, Hessing, & Voragen, 1998). Reduction of patatin may improve the binding capacity of the PIs on the mixed mode resin (KO Strætkvern, personal communication; unpublished result). The normal pH of untreated PFJ was around 5.3. First, it was acidified to 4.3 with sulphuric acid holding 40 minute at 4^{o} C and was centrifuged for removing precipitated protein. Then, the pH levels of the clarified supernatant were adjusted to 5.5, 6.5, and 7.5. A SDS-PAGE analysis (Figure17) of

untreated PFJ and acidified PFJ's revealed that there was some different between untreated and acidified sample where no specific separation band of PI groups in the untreated PFJ, while several types of PI groups has been found in the CAP sample. The patatin group is going to reduce when different acidification was applied and this was the most crucial part in this experiment.

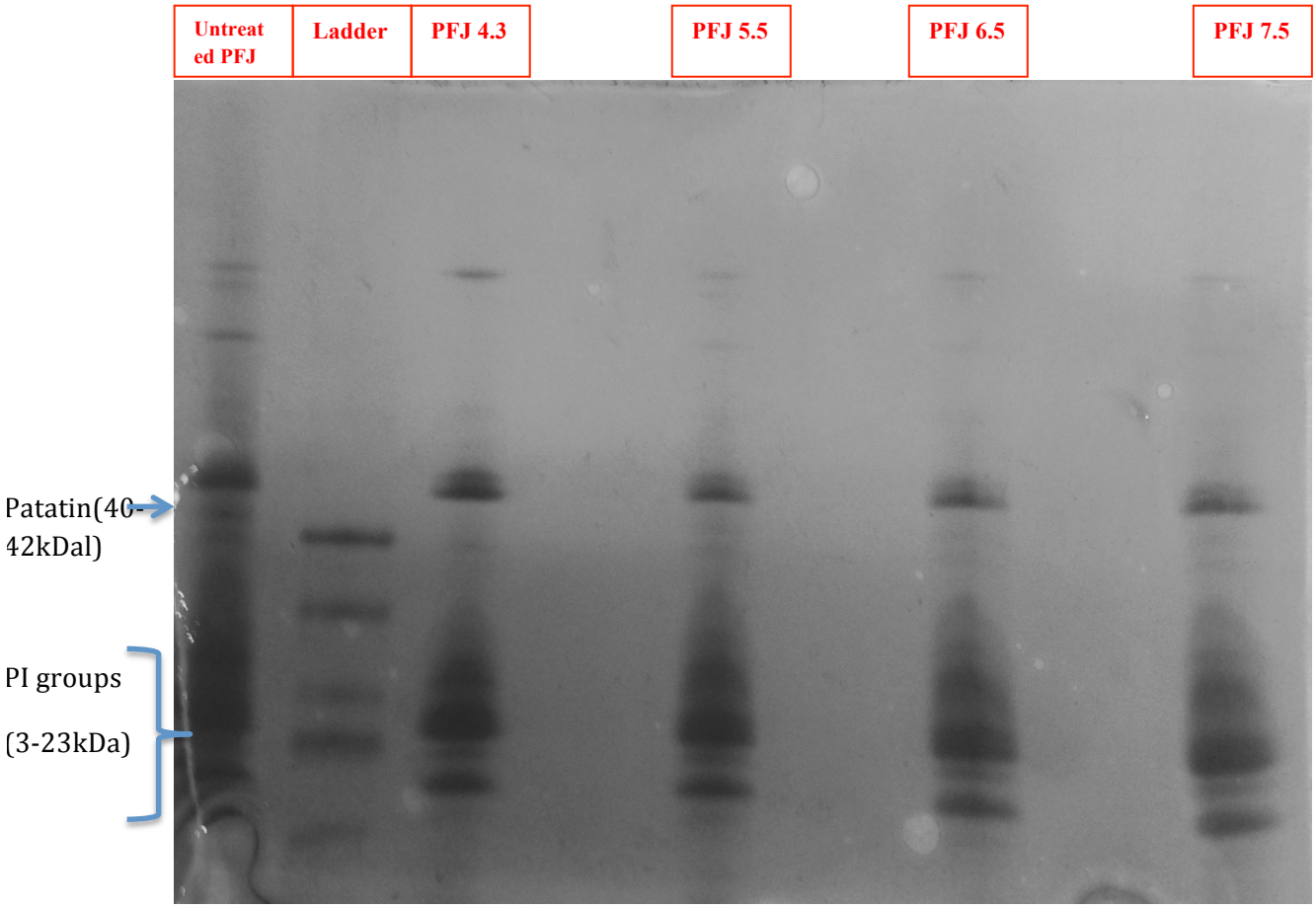


Figure 17: Electrophoretic analysis of untreated PFJ and acidified CAP samples adjusted to different pH.

4.3 Screening of binding pH in Mixed Mode Chromatography

The optimization of CAP pH was carried out by using the 1 ml HiTrap Capto MMC column. At acidic pH 4.3, protein binds strongly to the column and the pH gradient had minor effect on elution. There was significant elution response to the high pH step (7 to 11) but main responses found in the high pH buffer. Added, majority of effects were

visible in the high pH level (Figure 18). So, there was no significant effect of protein in the eluting part as well flow throw fractions did not show any band. On the other hand, the response of the fraction on gel, there was no significant protease inhibitor group appeared for instance: PI 1 or PI 2 groups (7-20kDal). However, there was some patatin group appeared in the eluting section.

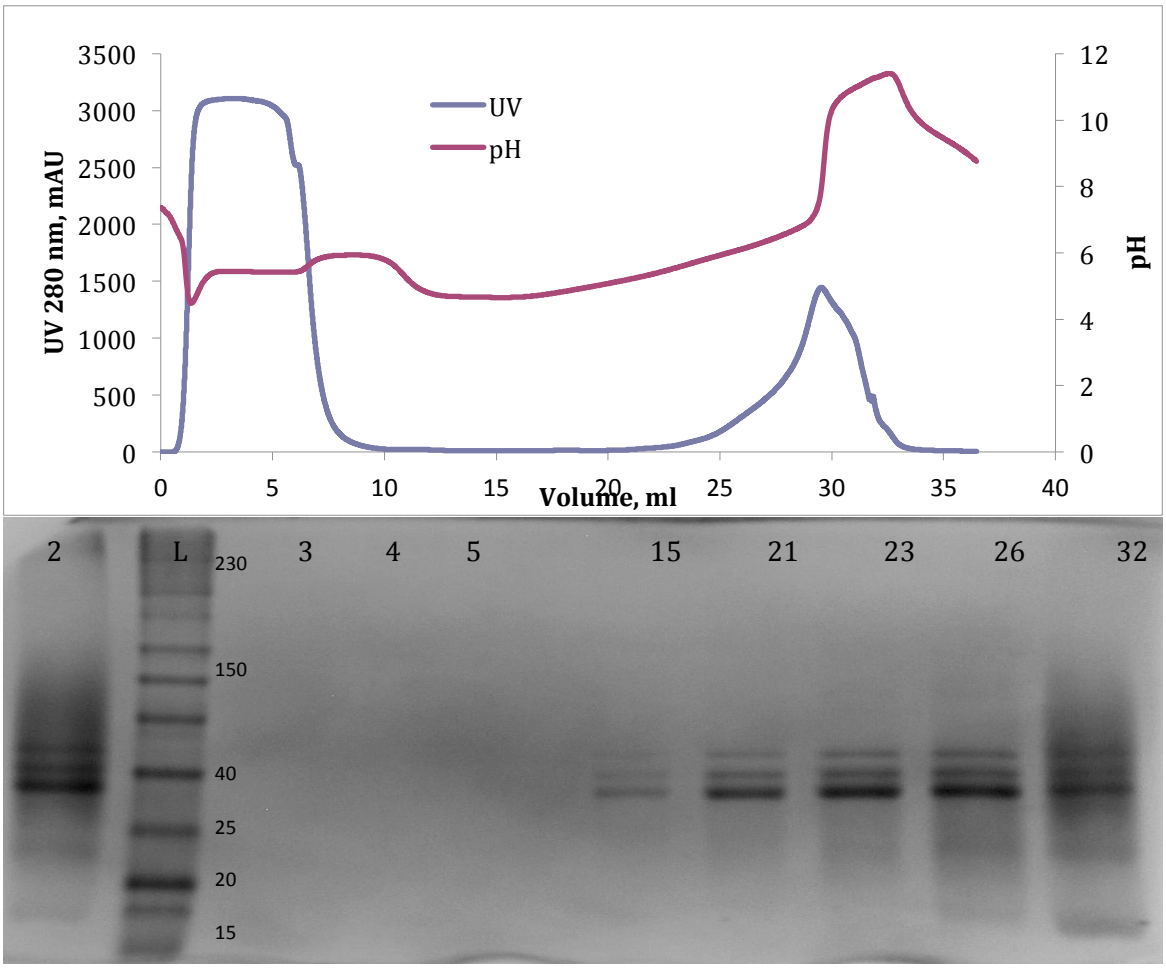


Figure 18: A: Chromatogram (UV 280 nm and outlet pH) from MMC column (1 ml) of CAP pH 4.3 applied at 2xCV. B: Coomassie BB-stained 10% Tricine SDS-PAGE. L; Molecular weight marker, lane 2, injected sample; lanes 3-5, flow throughs; lanes, 15-32, fractions from CAP pH 4.3 elution.

The main information from this scouting experiment revealed that there is less possibility to get protease activity without any salt effect. Therefore, next experiments were carried out with different salts in the elution buffer.

4.4 Eluting condition pH and salts

For the screening of elution conditions of the adsorbed proteins from CAP (pH 4.3) three different conditions (no salt, 1M NaCl and 1.5 M NH_4Cl) were tested using the HiTrap MMC column. The Figure 19 compares the chromatography runs and electrophoresis analysis for each set of eluting conditions.

Noticeably, the flowthrough fractions did not show any protein bands, which signify that, the high UV signal is due to phenolic substances in the sample. The chromatogram analysis of all flow through peaks was different to each other, and for the (pH + 1.5M NH_4Cl) run, there was a markedly sharper peak. However, flowthrough fractions (2-4) performances on the gel have some diversity in those different conditions, indicating sample protein adsorbed efficiently to resin. There was no protein bands appeared except patatin. In the experiment without pH elution only Figure 19 (A), protein peak was visible when high pH was applied and compare to Figure 19 (B) and (C) the corresponding gel gives new information to that experiment, PI elutes at high pH. Under the condition of 0-1.5M NH_4Cl, protein adsorbed more strongly than in others because NH_4^+ has more salting out effect than Na^+. The corresponding gel analysis showed no specific separation of protease inhibitors. Patatin elutes early in the “pH only” and “pH + 1M NaCl” experiments while patatin band is less noticeable in the pH + 1.5 M NH_4Cl experiment.

The difference between the three different conditions of sample were prominent because of the area of bound protein in NaCl was higher than others and also showed better separation band of protease inhibitors group which was the main purpose of this study. Therefore, for approaching the optimal conditions regarding elution buffer, NaCl has been used for the next experiments.

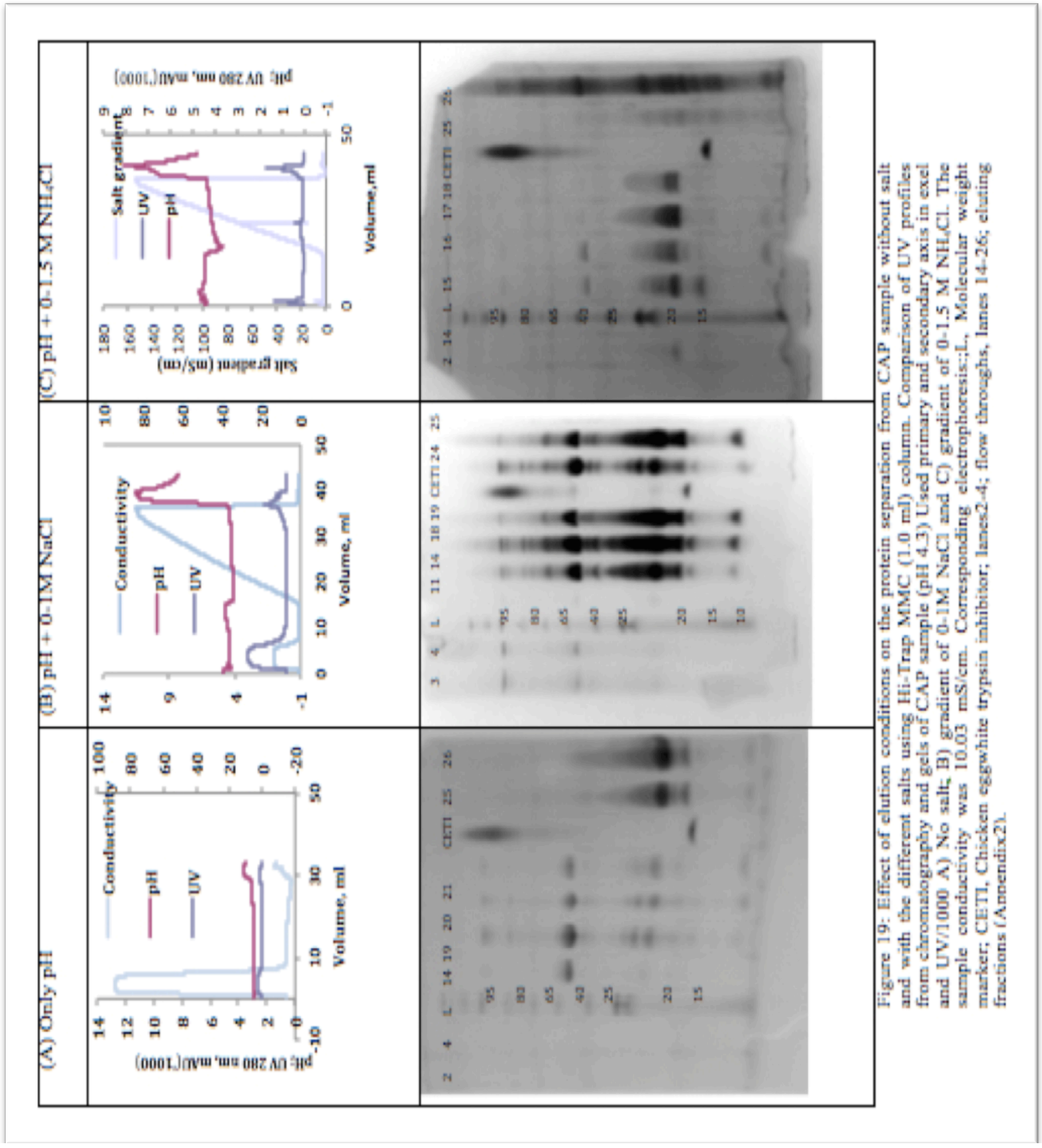
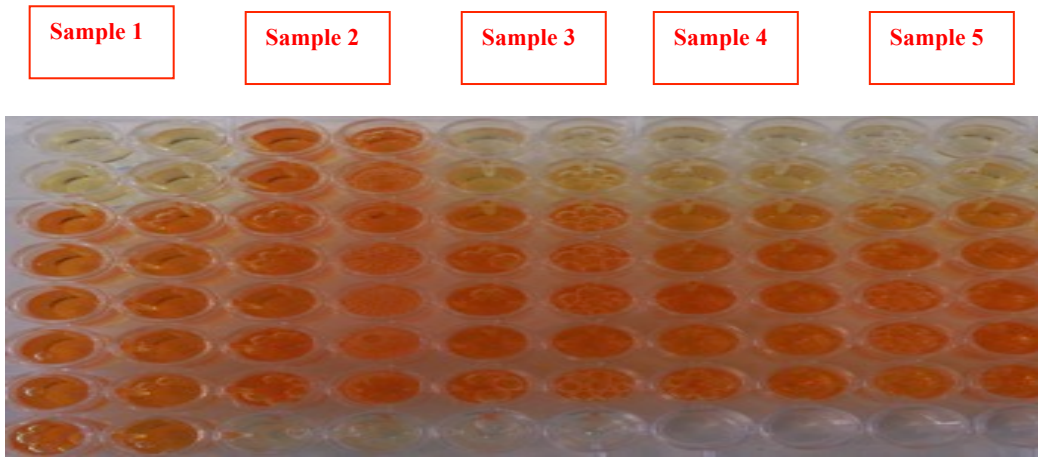
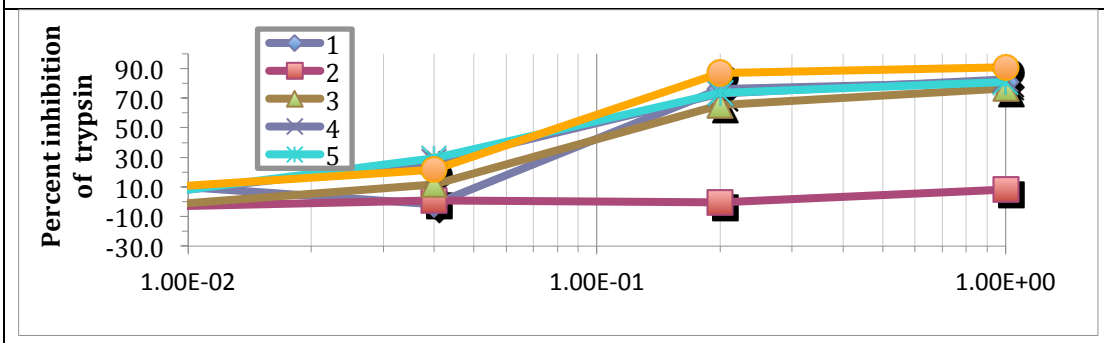


Figure 19: Effect of elution conditions on the protein separation from CAP sample without salt and with the different salts using Hi-Trap MMC (1.0 ml) column. Comparison of UV profiles from chromatography and gels of CAP sample (pH 4.3) Used primary and secondary axis in excel and UV/1000 A) No salt, B) gradient of 0-1M NaCl and C) gradient of 0-1.5 M NH_4Cl. The sample conductivity was 10.03 mS/cm. Corresponding electrophoresis:L, Molecular weight marker; CETI, Chicken eggwhite trypsin inhibitor; lanes2-4; flow throughs, lanes 14-26; eluting fractions (Annendix2).

4.5 Trypsin Inhibitor assay of MMC chromatography

The eluted fractions from CAP pH 4.3 and pH gradient elution pH 4.5-7.5 experiment were analyzed in a trypsin inhibitory assay demonstrating the presence of significant inhibiting activity. TIA experiment has been done from the CAP pH at level 4.3. Here, sample 1 to sample 5 showed more than 50% activity whereas sample 2 had the opposite (Figure 20). Sample no. 5 had the most effect on TIA which activity was 94 (U/ml), as well as best perform on plates. However, flow through 4 had no inhibition activity.





EXPANDED VIEW: Trypsin inhibition curves from PPI samples (1-5), acting on azocasein (450 nm).

Sample no	Dil. at 50%	1/dil.= IC50	Activity (U/ml)
1	0.125	8.0	64
2	>1	N.D	N.D
3	0.135	7.4	59
4	0.09	11.1	89
5	0.085	11.8	94

Figure 20: Trypsin Inhibitory Assay of fractions from CAP pH 4.3 same as figure 19(a) and pH gradient elution pH 4.5-7.5. Sample 1-2; flow throughs 2-4, Sample 3-5; eluting fractions22-24.

4.6 Testing of buffer conditions on protein binding

In order to observe the effect of pH on the protein binding and on peak separation during a pH gradient elution, both column equilibration buffer pH and sample pH were varied. Design of Experiment (DoE) methodology was used to investigate the main and combined effects of the column equilibration pH and the sample pH. Thus, to obtain a response surface screening, a layout of two factors (column pH and sample pH) with 3 levels for each pH was used. The response was defined by the eluted peak areas (EPA).

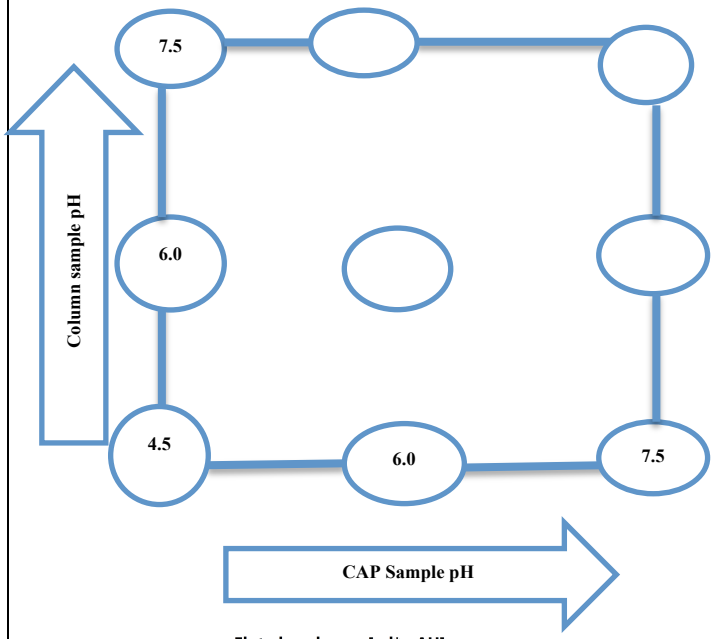
Figure 21 shows the summarized results from the DoE-analysis. The pH levels are 4.5, 6 and 7.5 respectively.

The contour plot (C) shows the eluted peak areas for specific levels of the factors "Column pH" and "Sample pH". Thus, it illustrates the effects of both factors on the eluted peak area. From the plot it is clear that the eluted peak area has higher values with the increase of "column pH". The highest value of EPA (1600 ml*mAU) is obtained for column pH value 7.5. In addition, the contour plot indicates that when the "column pH" is increased by 0.5 unit the eluted peak area increases by about of 200 mAU*ml. The comparable response on TIA in the contour plot (D) also showed that maximum binding of protein in the column pH 7.5.

The Figure 21 (E) reveals that the R^2 value is 0.906. R^2 is a statistical measure which tells how close the data is to the fitted model is. R^2 is usually presented as percentage and 0% means a poor fit whereas, above 90% means a good fit. The percentage of variation explained by the response of the model was 90.6% (R^2=0.906) and this indicates that the model fits well to the data. For the fitted model these values are calculated for the fitted response values and thus the RSD has been calculated. This experiment has a RSD value of 241, which indicates that the standard deviations are higher than the mean values as the formula has standard deviations at the top. More specifically, it can be said that the fitted eluted peak measurements for this experiment have higher amount of variations and the interaction plot gives the same information. The $Q^2 = 31.9\ \%$ shows a poor fit for the model and this is because the data have higher variation with fewer observations (Ballabio et al., 2007).

In addition, this plot for the coefficients or factors shows the confidence intervals as well. From the confidence intervals it can be concluded that whether a factor is significant or not. From this graph it can be said that only “column pH” is significant since the confidence interval does not intersect 0 and it means that the column pH is a significant model term. Sample pH does not significantly affect the protein binding. On the other hand, the scenario is vice versa for “Sample pH” and column buffer pH interaction.

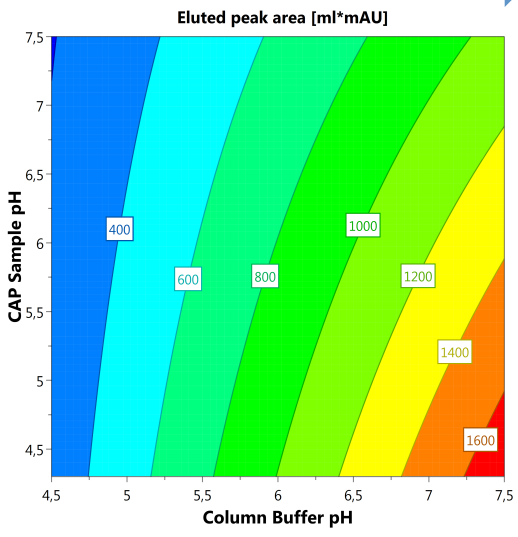
A)



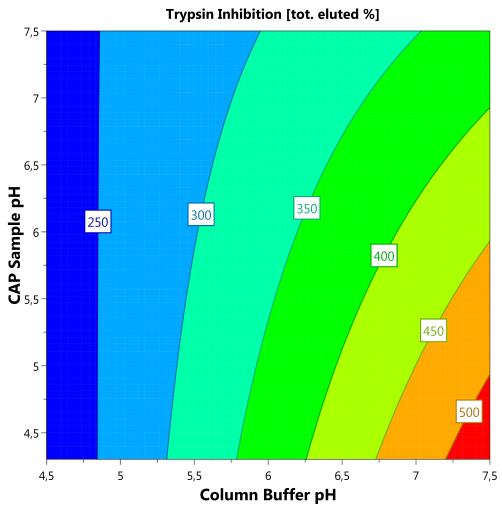
B)

	1	2	3	4	5	6	7
1	Exp No	Exp Name	Run Order	Incl/Excl	Buffer pH	Sample pH	Eluted peak area
2	1	5	1	Incl	4,5	4,3	2724,5
3	2	7	2	Incl	6	4,3	3120,9
4	3	8	3	Incl	7,5	4,3	1563,9
5	4	10	5	Incl	4,5	6	504,3
6	5	11	6	Incl	6	6	1170,5
7	6	9	4	Incl	7,5	6	1553,4
8	7	14	9	Incl	4,5	7,5	503,85
9	8	12	7	Incl	6	7,5	706,7
10	9	13	8	Incl	7,5	7,5	864,3

C)



D)



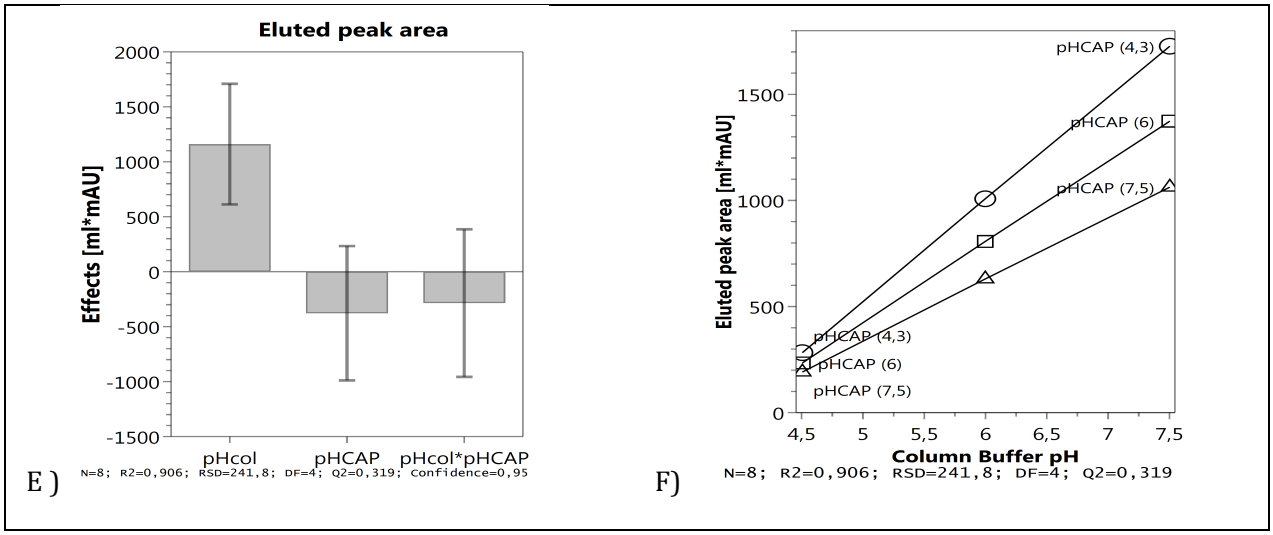


Figure 21: A) 3x3 Design of experiment model (B) Eluted peak areas numerical value from 3x3 model (C) Contour plot of Eluted peak area (ml*mAU) showing the effects of column pH and sample pH (D) Trypsin inhibitory activity of eluted peak area (E) Coefficient plots showing the effects and significance level of the factors (buffer pH and sample pH) (F) Plot of the fitted linear model with two factors or independent variables where the response is the eluted peak area.

The interaction plot (F) shows the effect on EPA for increasing levels of column pH at three values of sample pH. The response values for column pH shows a general positive correlation, but highest effect for acidic sample pH loaded to the neutral column pH. This result gives the idea that; the column pH has more pronounced effect adsorbed sample protein and thereby on eluted peak areas, than the sample pH.

The screening experiment indicated that neutral buffer pH and acidic sample pH may perform better. Then two experiments were performed, in which sample pH was the same (CAP pH 4.3), but where one experiment was with buffer pH 6.0 and another with buffer pH 7.5(Figure 22). In the Figure 22(A) protein elution was not significantly affected during salt gradient elution, only when high salt was applied. The gel analysis shows that some patatin eluted in the gradient and PI eluted in the major peak.

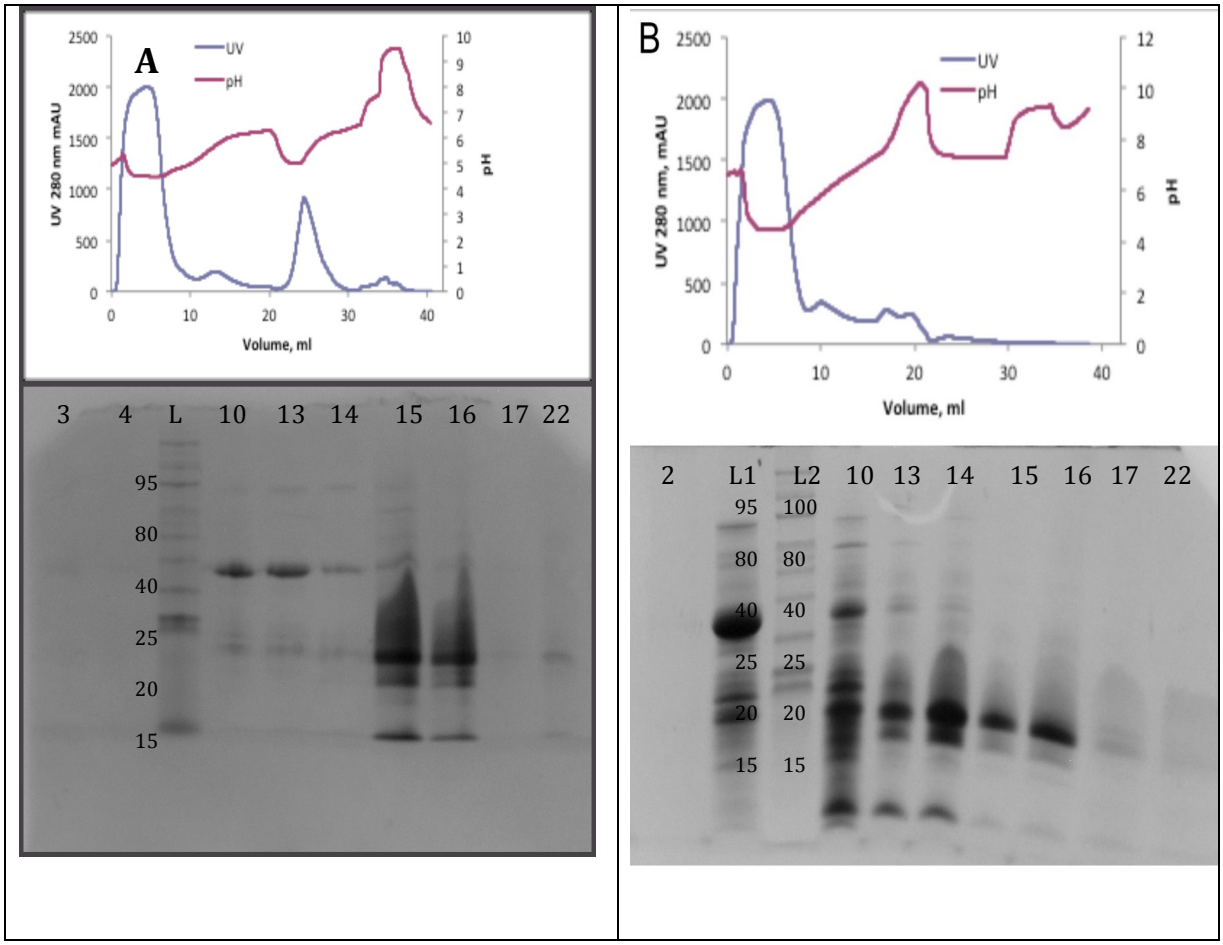


Figure 22: Two different chromatogram with different buffer pH. 10% Tricine SDS-PAGE analysis of protein ihibitors from different pooled fractions analyzed by Coomassie Brilliant blue staining for both gels A) Chromatogram with buffer pH 6.0 B) Chromatogram with buffer pH 7.5. L1-L2; Molecular weight marker, 2-4;Flow throughs, 10-22; eluates from sample pH 4.3.

On the other hand, when column buffer pH 7.5 (Figure 22 B) was applied, protein eluted earlier. Proteins came in the flowthrough peak, including patatin, and adsorbed proteins eluted earlier in the gradient. However, the gradient peaks did not separate into discrete PI bands in the gel analysis. Thus, column buffer conditions at pH 7.5 and pH 6.0 either caused elution of protein earlier or caused more of the patatin to bind. Considering maximum binding of capacity for PPI, CAP pH 4.3 and column pH 5.5-6.0 were therefore chosen for further experiment.

4.7 Effect of binding condition in the elution salt

Consequently, for improved binding of protease inhibitor proteins, column equilibration buffer was set to pH 5.5 and sample pH 4.3. It appears; when sample was applied the column buffer pH adjusted the sample pH. Trypsin activity test showed, the better performance found in the elution condition and high protein significant inhibition completely found in the elution part (Figure23). The TIA activity (% inhibition) in the elution part was found to be more than 80%(Appendix4). Protein in the eluted fractions corresponded with the expected PI groups. High molecular proteins were observed in several lanes and to the group of PI (17-20 kDa) was detected (Figure 23).

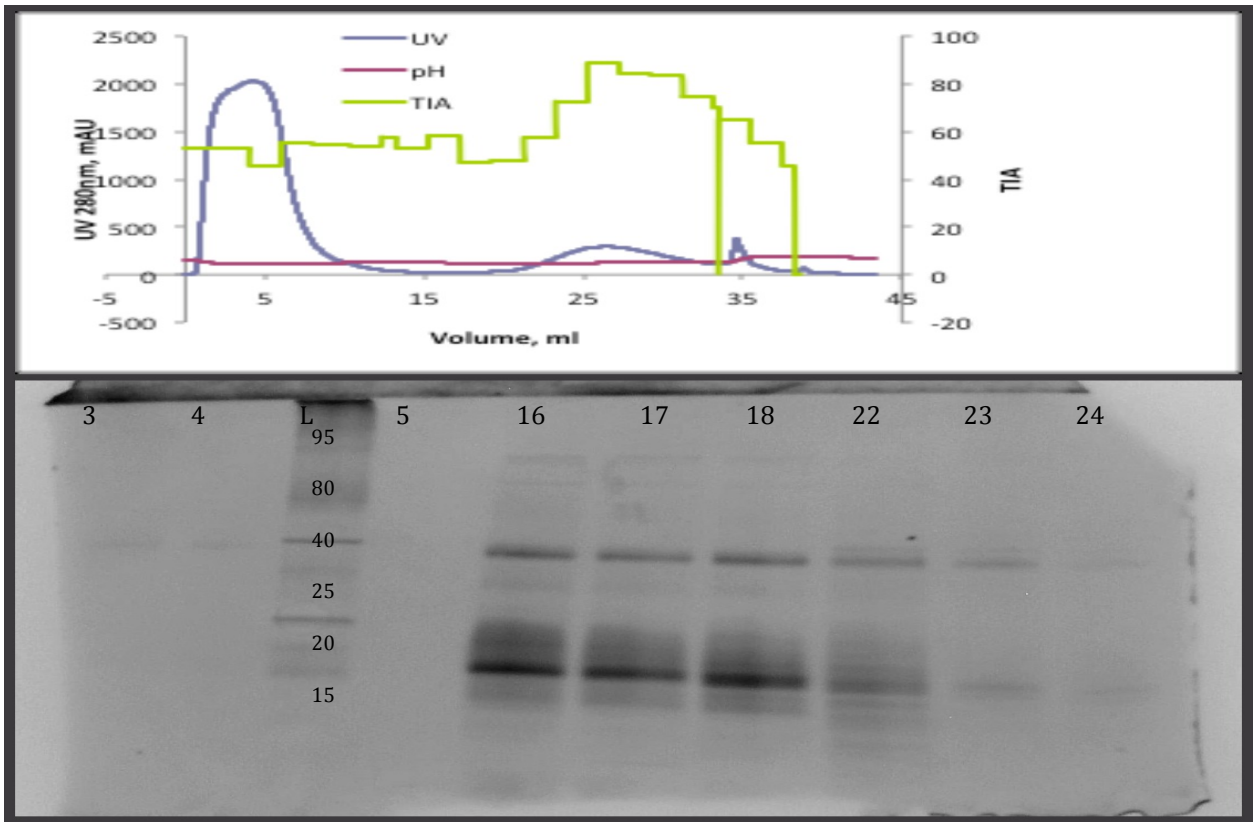


Figure 23: Chromatographic run (UV 280 nm, mAU)and TIA performance of eluting salt and Tricine SDS PAGE(10%) of flow throughs and elution fractions from MiMo pH optimization experiment. 3 - 5;Flow throughs, L; Molecular weight marker, 16- 24; eluates from sample pH 4.3.and buffer pH 5.5

Considering maximum binding of capacity for PPI, CAP pH 4.3 and column pH 5.5 was therefore chosen for further experiments in the study.

4.8 Investigating connectivity versus Sample

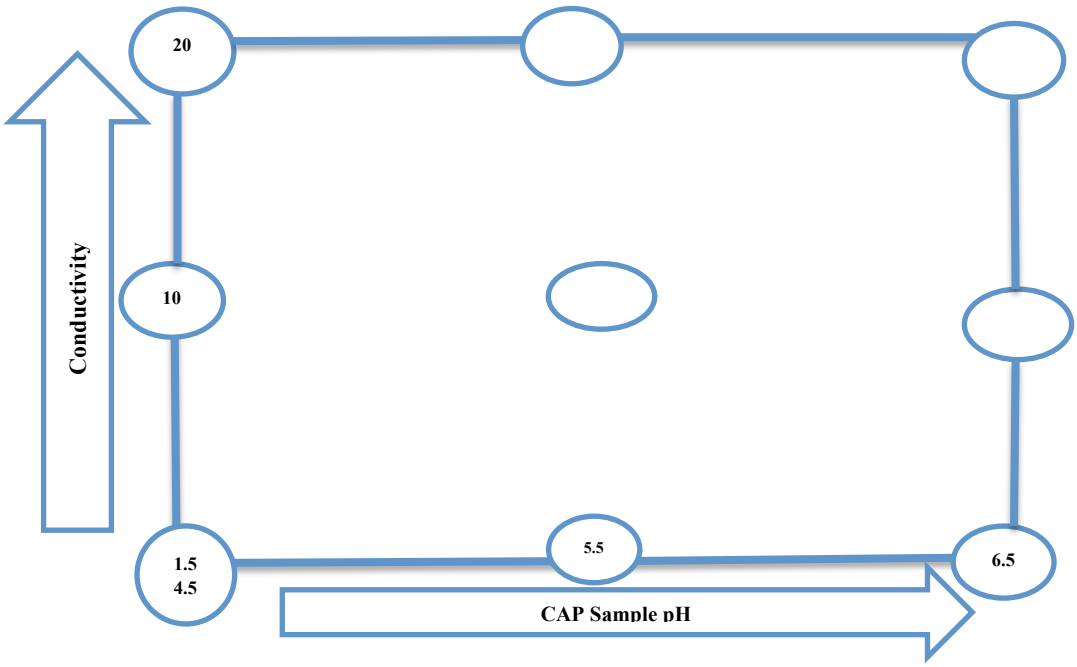
Similar to the experiment in 4.6, another Design of Experiment on conductivity and sample pH was performed to be more confident on the conditions for protein binding. This too, was executed in a 1ml HiTrap Capto MMC Colum. To obtain a response surface screening, a layout of 2 factors (conductivity and CAP sample pH) with 3 levels for each pH has been used. In addition to eluted peak area, trypsin inhibitor activity was also used as response. Thus, a MLR (multilinear regression) model of the following format was fitted to find out the optimal protein binding condition.

Trypsin Inhibitor Activity (i.e. Sum of % inhibition) = $\beta_0 + \beta_1$ Conductivity + $\beta_2 CAP$ Sample pH + β_{12} Conductivity ×Sample pH + ϵ_I

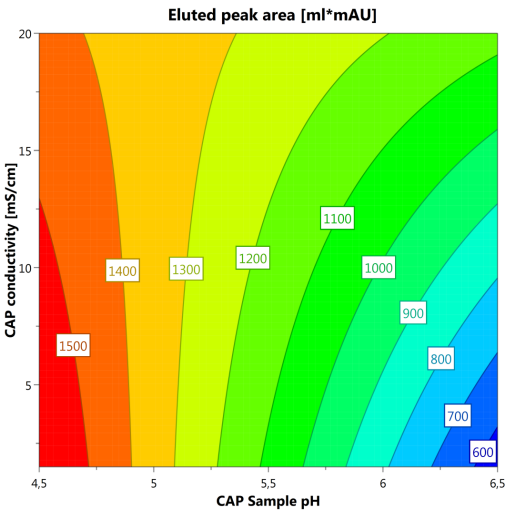
In Figure 24 an overview about this experiment has been represented with the help of design layout, contour plots, co-efficient plot and interaction plot.

The contour plot (B) shows the eluted peak areas for specific levels of the two factors and it is clear that the CAP sample pH has effect on eluted peak area. The highest value (approximately 1500 ml*mAU) is obtained for pH 4.5. When the CAP sample pH was increased by 0.5 unit, the eluted peak area decreased by about of 100 mAU*ml, which indicate that protein dissociates from the ligand when high pH is applied.

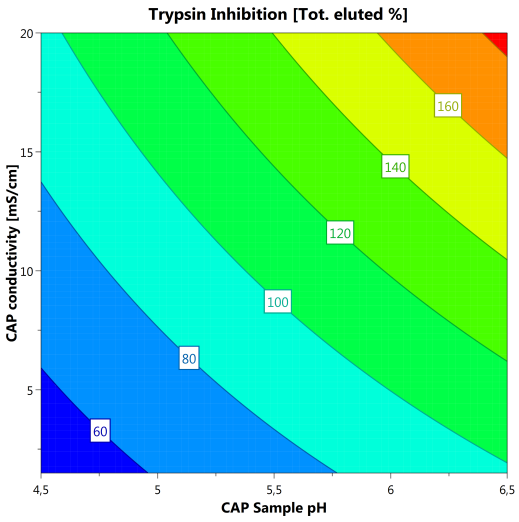
A)



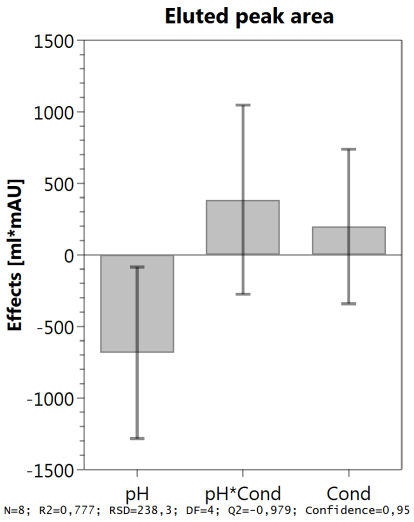
B)


Eluted peak area [ml*mAU]
CAP conductivity [mS/cm]
CAP Sample pH
1500
1400
1300
1200
1100
1000
900
800
700
600

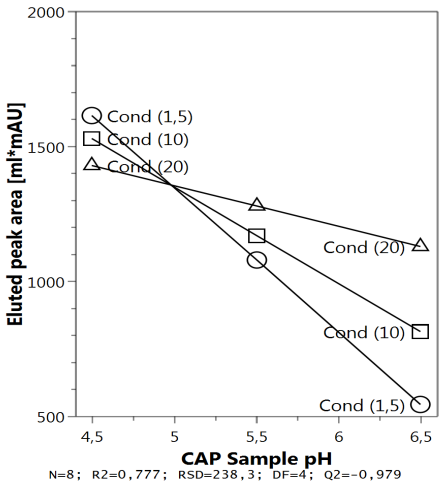
C)


Trypsin Inhibition [Tot. eluted %]
CAP conductivity [mS/cm]
CAP Sample pH
160
140
120
100
80
60

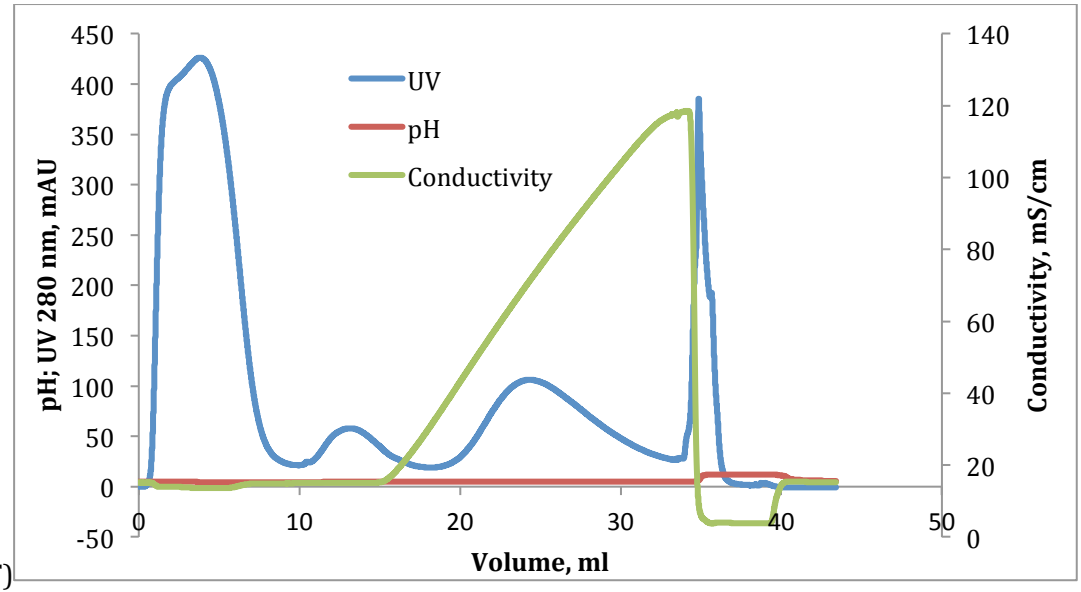
D)


Eluted peak area
Effects [ml*mAU]
pH
pH*Cond
Cond
N=8; R2=0,777; RSD=238,3; DF=4; Q2=-0,979; Confidence=0,95

E)


Eluted peak area [ml*mAU]
Cond (1,5)
Cond (10)
Cond (20)
CAP Sample pH
N=8; R2=0,777; RSD=238,3; DF=4; Q2=-0,979

F)



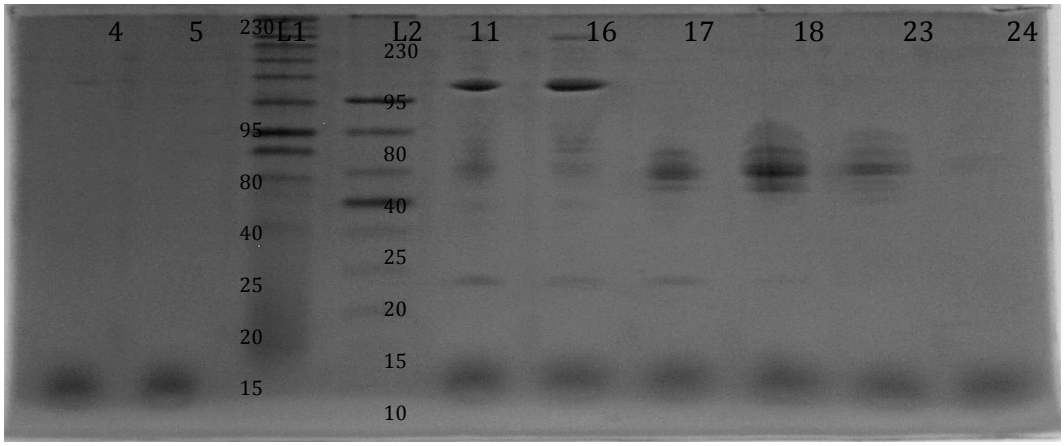


Figure 24: shows the summarized results from the DoE-analysis. (A) 3x3 layout of Design of experiment; (B) Contour plot of Eluted peak area (ml*mAU) showing the effects of conductivity and CAP sample pH (C) Trypsin inhibitory activity of eluted peak area (D) Coefficient plots showing the effects and significance level of the factors (Conductivity and CAP sample pH) (E) Plot of the fitted linear model with two factors or independent variables where the response is the eluted peak area. (F) Chromatogram and gel analysis of the center experiment (sample pH 5.5 and conductivity 10), Frac. No. 4-5,Flow throughs; L1, L2; Two types of molecular weight marker, 11 - 24; eluates from CAP sample pH 4.5 and conductivity 10.

The contour plot (C) shows TIA response surface, which revealed that the maximum response found on the high sample pH 7.5 as well as found high response at conductivity 20.

Coefficient plot shows the confidence intervals and from those intervals it can be concluded that the sample pH is a significant negative factor while conductivity and the conductivity-pH interaction do not significantly affect the protein binding. Moreover, the results are also displayed with R^2, RSD and Q^2. R^2 is a statistical measure, which tells how close the data are to the fitted model. R^2 is usually presented as percentage and 0% means a poor fit whereas, above 90% means a good fit. The percentage of variation explained by the response of the model was 77.7% (R^2=0.777) and this indicates that the model does not fit well to the data. Then, the RSD (Relative standard deviation) for the fitted MLR model has been reported as 238.3. This RSD is a special form of standard deviation which tells how the regular standard deviation is a small or large quantity in compared to the sample mean and it is represented as percentage. Since the RSD seems a large value here, it can be said that the standard deviations might be larger than the mean values for this data. Another essential term in the results is the predictive power (Q^2 = -.979) which has another name called cross validated R^2.

A negative Q^2 means that the PRESS (sum of square differences between the experimental response y_i and the response predicted by the regression model) is greater than TSS (total variance or variation that a regression model can explain), this leads to the conclusion that the predicted model does not perform well compared to the cross validated models.

The interaction plot provides the same statistical information as the coefficient plot. However, from the fitted lines there is some information about protein binding condition regarding three different pH conductivity levels. When the pH level has been increased from 4.5 to 5.5 and then to 6.5 the lines go far away from each other, which mean there is a huge difference between the eluted peak areas. On the contrary, for same pH level with different conductivities (1.5, 10 and 20) there is not too much difference between the response values eluted peak area. The same conclusion has been found in the previous plot of coefficient where the pH value shows significant effects on the response values, whereas, the conductivities do not.

4.9 Testing salts on elution profile

Salt gradient elution (0-1.5 M NaCl) with mobile phase modifiers was performed in the presence of either arginine or guanidine. Figure 25 shows the elution profiles for the protein on Capto MMC in the absence and presence of 0.025M arginine and of 0.025M guanidine in the mobile phase. In the absence of modifier exhibited did not show any protein, while applying specific modifiers with same conditions proteins bind significantly to the column. It should be noted that the pH increased during the gradient elution. This is important observation since the previous DoE showed the effect of pH on retentionArginine salt is preferable in this experiment because protease inhibitors were found in eluted peaks as demonstrated on 10% tricine SDS gels. On the other hand, protein in general exhibited weaker retention in the presence of 0.025M guanindine. While PI exhibited a decrease in retention with guanidine in the multimodal system (Figure 25), the decrease was much less pronounced than that observed with arginine at the same concentrations. It is not clear why this particular family of proteins exhibited such a different behavior. It may be due to the propensity of these proteins to undergo minor structural changes in the presence of low concentrations of guanidine (Holstein, Parimal, McCallum, & Cramer, 2012). Another experiment with 20% of glycol in the elution buffer was conducted, but unfortunately no significant result was found in that experiment. According to obtained results arginine salt was therefore chosen for further experiments in the study.

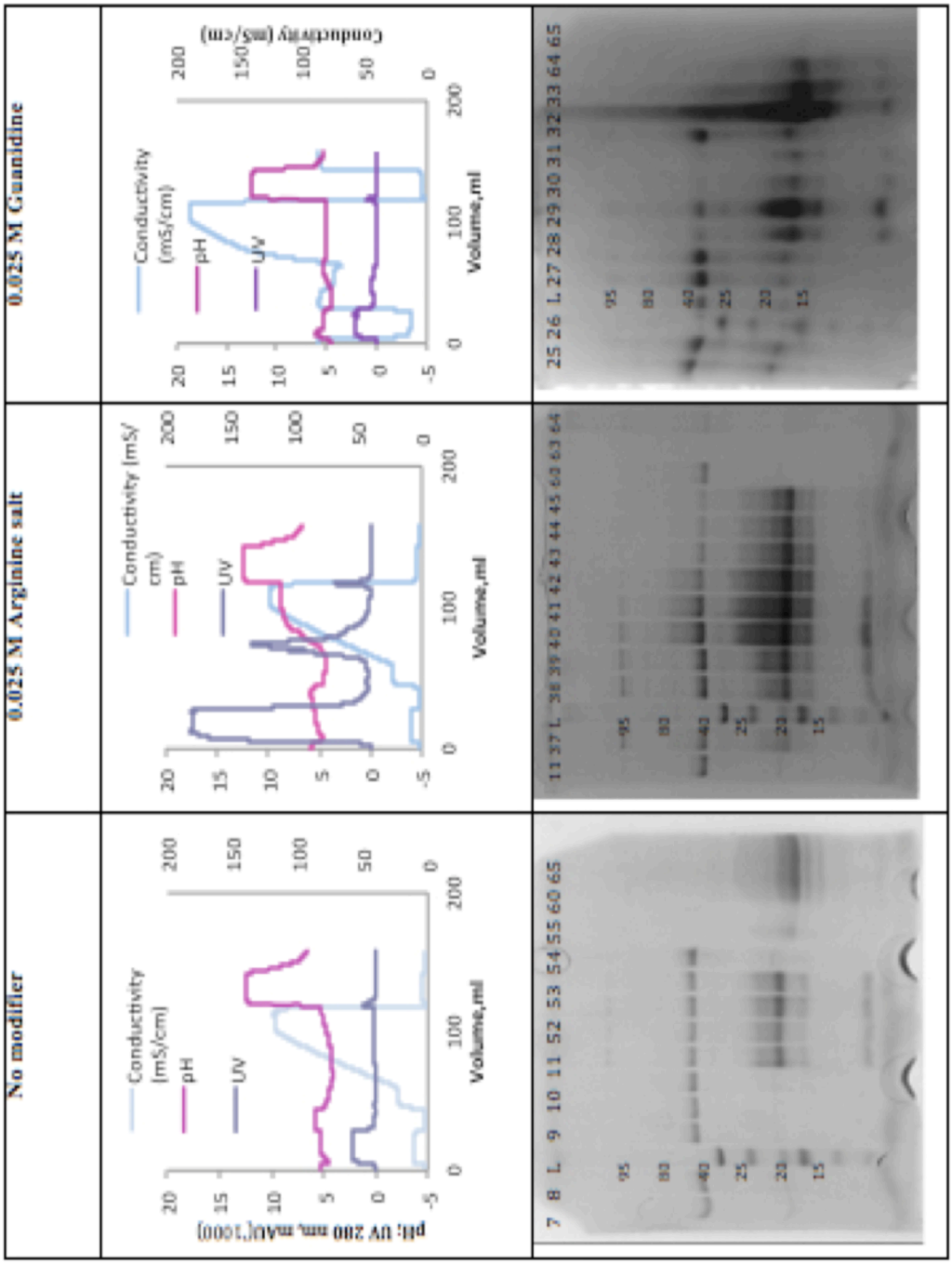


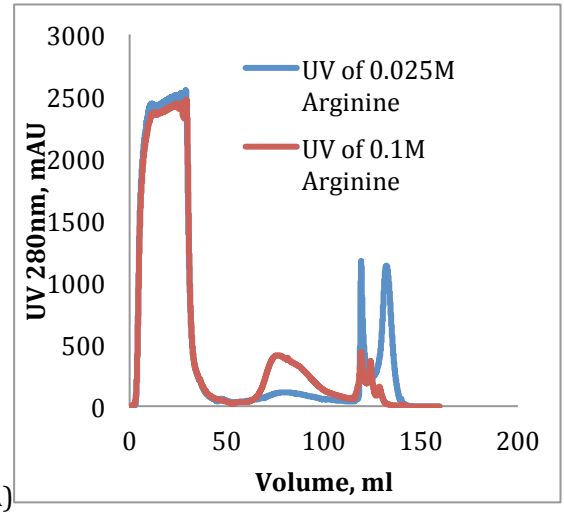
Figure 25: Effect of mobile phase modifier on the protein separation from CAP sample using Hi-Trap MMC 1.0 ml column. Comparison of UV profiles from chromatography and gels of CAP sample (pH 4.3) UV/1000 in exel A) No modifier, salt gradient only, followed by high pH elution step; B) same as in A but with 0.025M arginine in elution buffer and C) same as A but with 0.025 M guanidine. The sample conductivity was 10.03 mS/cm. Corresponding electrophoresis: L, Molecular weight marker; lanes7-25; flow throughs, lanes 26-65; eluting fractions.

4.10 Effects of Arginine concentration

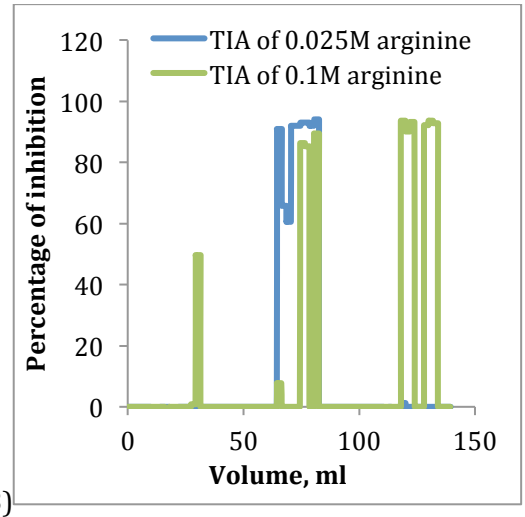
To be more confident on the effect of arginine (guanidino pKa 12.5) the experiments were carried out with different concentration and also compared with another basic amino acid, namely lysine (ε-NH_2 pKa 13.5). Addition of 0.025 or 0.1 M arginine into the elution buffer showed a greater recovery of aggregated proteins. When there was no arginine, a low recovery of aggregated materials caused underestimate of aggregation protein versus in its presence a much larger peak in the presence of 0.025-0.1 M arginine. When the concentration of arginine was increased from 0.025 to 0.1M, the eluted peaks come earlier and protease inhibitors was also found in the in the gradient section of the elution (Figure 26 a). The protease inhibitors exhibited a decrease in retention in the presence of 0.1 M arginine. According to the previous work with arginine on protein researchers also suggested, higher arginine concentrations to induce protein recovery (Tsumoto et al., 2007). Trypsin Inhibitory assay also showed that, 0.1M arginine effect on protein more than the effect of 0.025M arginine. The protease activity was more than 80% in the gradient elution section in the presence of 0.1M arginine while the low concentration arginine (0.025M) had lesser influence on elution (Figure 26 b).

The experiment also aimed to compare the effect of protein elution with lysine. The performance on chromatogram as well as SDS gels result showed lysine and arginine salt effects are almost same except better separation of PI. However, arginine has better separation of protease inhibitors II (20 kDa) on 10% tricine gel performance.

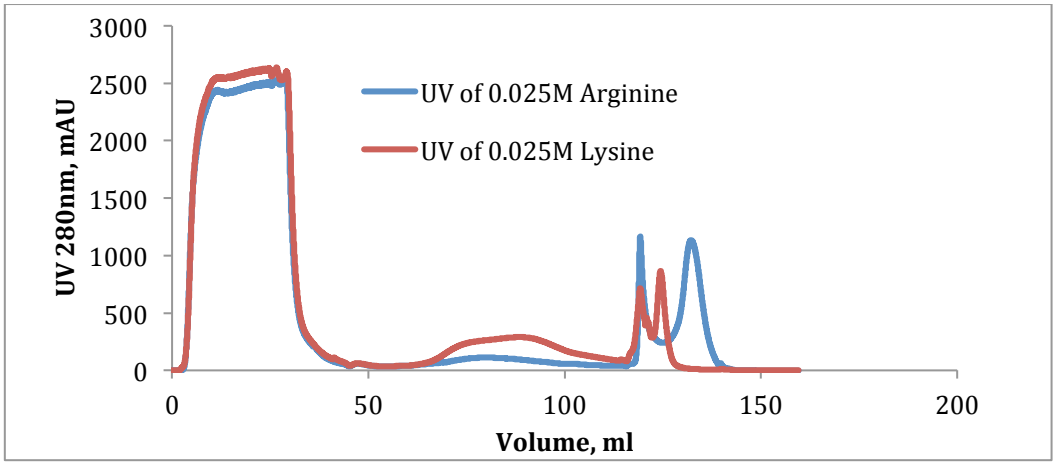
A)



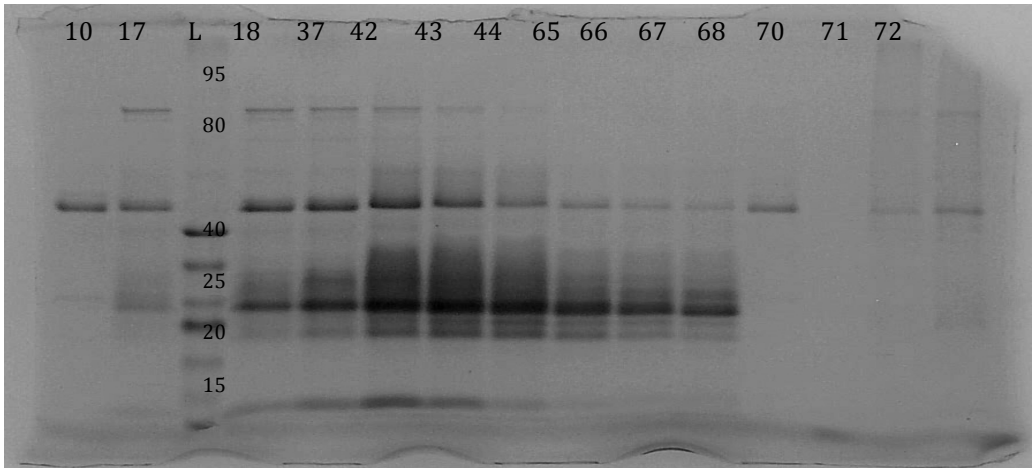
B)



C)



D)



E)

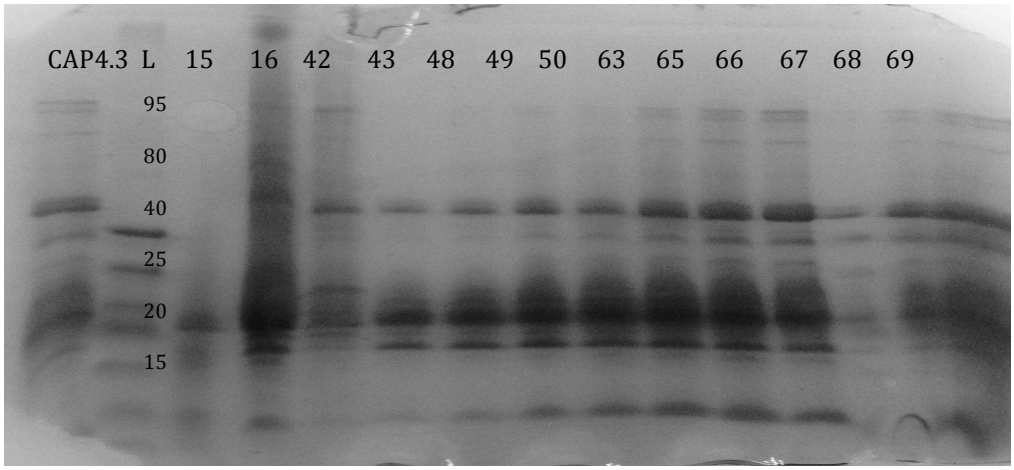


Figure 26 : The effects of different arginine salt concentration. A) comparison between 0.025M arginine and 0.1M arginine on Mixed mode chromatography B) TIA performance of different concentration of arginine C) Comparsion chromtogram of Arginie versus Lysine D & E) 10% Tricine SDS PAGE analysis of Arginine and lysine. Corresponding electrophoresis: CAP4.3, Undiluted sample at pH 4.3; L, Molecular weight marker; lanes10-18; flow throughs, lanes 37-72; eluting fractions.

4.11 Bioinformatics Analysis

Tertiary structure of some protease inhibitors are already available (Birk, 2003) in the protein database where they are functionally bound with the chromatography ligands. A search of the UniProt database provided 6 non-redundant sequences from *S. tuberosum*, which are listed as Kunitz-type protease inhibitors (aspartic protease inhibitors, serine protease inhibitor), potato cystatin and potato carboxypeptidase inhibitor (PCI).

This study attempted to predict the structures of various protease and proteinase inhibitors by the Phyre 2 software. By definition proteases are any kind of enzymes, including the endopeptidases and exopeptidases that catalyze the hydrolytic breakdown of proteins into peptides or amino acids. A proteinase is a protease that begins the hydrolytic breakdown of proteins, usually by endolyttic splitting into polypeptide chains. It is also a synonym of endopeptidases (e.g. pepsin, trypsin, or papain) (McDonald, Schwabe, & Owers, 1987).

The information about these types of protease inhibitors and proteinase inhibitors were found in the Phyre2 (Table 7, figure 27). Confidence level (%) indicates the confidence or accuracy in the prediction. Tertiary structure and coverage prediction is on average 78-80% accurate (i.e. 78-80% of the residues are predicted to be in their correct state). However, this accuracy is only reached if there are a substantial number of diverse sequence homologues detectable in the sequence database. If target sequence has very few homologues (something that can be checked by looking at the PSI-Blast results) then accuracy falls to approximately 65% (Kelley, Mezulis, Yates, Wass, & Sternberg, 2015).

Table 7 - Predicted tertiary model and their characteristics data on Phyre2 database of different types of protease and proteinase inhibitors from *S. tuberosum*. Red being high confidence and blue refers to low confidence.

PDB header: hydrolase inhibitor **Chain:** C **PDB Molecule:** kunitz-type proteinase inhibitor p1h5; **PDBTitle:** crystal structure of potato serine protease inhibitor.			
Confidence:	**100.0%**	**Coverage:**	**81%**
178 residues (81% of target sequence) have been modelled with 100.0% confidence by the single highest scoring template (Figure 27 A).			
Fold: Plant proteinase inhibitors **Superfamily:** Plant proteinase inhibitors **Family:** Plant proteinase inhibitors			
Confidence and coverage			
Confidence:	**100.0%**	**Coverage:**	**68%**
104 residues (68% of target sequence) have been modelled with 100.0% confidence by the single highest scoring template (Figure 27 B)			
PDB header: hydrolase inhibitor **Chain:** C **PDB Molecule:** kunitz-type proteinase inhibitor p1h5; **PDBTitle:** crystal structure of potato serine protease inhibitor.			
Confidence and coverage			
Confidence:	**100.0%**	**Coverage:**	**79%**
172 residues (79% of target sequence) have been modelled with 100.0% confidence by the single highest scoring template (Figure 27 C)			

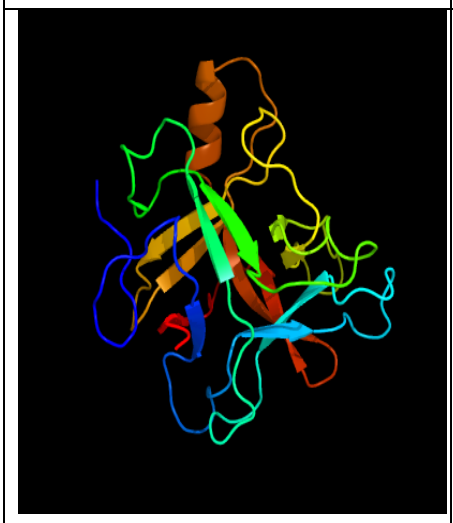
(A) Serine protease inhibitor1 (>sp|P58514)	(B) Proteinase inhibitor type-2 (>sp|Q41488)	(C) Aspartic protease inhibitor 10 (>sp|Q03197)
		

Figure 27: Predicted tertiary structure of different types of protease and proteinase inhibitors from potato (*S. tuberosum)*. All Images are coloured by rainbow N → C terminus.

In this study tried to model surface some of protease and proteinase inhibitors. Three types of inhibitors (Serine protease inhibitor 1(>sp|P58514), Proteinase inhibitor type-2 P303.51 (>sp|Q41488), Aspartic protease inhibitor 10(>sp|Q03197)) shows 100% confidence of this sequence and coverage are 81%, 68% and 79% respectively. By this result, it is clear that there is a higher chance of building up 3D model of Serine protease inhibitor 1 and this will help to predict to surface analysis of protease inhibitors. Other types of inhibitors also have possibility to surface modeling. However, due to the limitation of available sources and time the establishment of full tertiary model of this predicted structure has not been finished.

5. DISCUSSION

Three main groups of proteins found in PFJ; (a) The patatin protein; (b) the PIs; and (c) other proteins with higher molecular mass. The PIs include about 25-50 percent of total protein in PFJ (Pouvreau et al., 2001). The PIs consists of a heterogeneous group of proteins, which have approximately 3-23 kDa. Moreover, PIs perform a significant role in plant defense system against parasites and pests (Grosse-Holz & Hoorn, 2016).

For extracting PIs from PFJ, the phenolic component should be prevented from undergoing enzymatic oxidation. Due to this reason, citric acid was added to inhibit polyphenoloxidase by complexing its metal cofactor. Also for obtaining relatively purer sample of PI, sulfuric acid was added to the PFJ to promote the precipitation of patatin proteins.

The pH affects the surface charge of PIs and ligand of the resin. Therefore, pH has a salient effect on the strength of electrostatic interaction between a protein and pIs of proteins. The isoelectric point range for potato PIs are pH 5-10 (L. Pouvreau et al., 2001b). The PIs are stable at low pH and have stable functional properties (L. Pouvreau et al., 2001b).The high positively charged PIs have enough interaction with the weakly charged MMC ligand at low pH 5.5 , therefore the major fractions of PIs binding to MMC appear to be a fraction of PIs with relatively low PIs.

5.1 Mixed mode chromatography is an excellent way for isolation and separation of several protein

In terms of protein isolation, mixed mode chromatography is an outstanding technique, where MiMo adsorbent displays both ion-exchange and hydrophobic interaction in the chromatographic processes. The aim of this thesis project was to isolate and separate a series of protease inhibitors from potato fruit juice (PFJ) by using multimodal chromatography. Thus, the separation has been mainly carried out using Hi-Trap MMC column. Although the Capto MMC ligand is attached to agarose, agarose matrices are considered to be relatively inert to binding of most proteins (Woo, Parimal, Brown, Heden, & Cramer, 2015).

The multimodal functionality gives a different selectivity compared to traditional ion exchangers and also provides the possibility of operating in different regions with respect to pH and conductivity. The hydrophilic and electrostatic features of molecules were considered as the effective force for optimal binding condition of target proteins with the cationic MMC resin. Therefore, a positively charged and hydrophilic surface around the protein is necessary for the optimal binding capacity (Janson, 2012).

According to the chemical characteristics of MiMo ligand, the interaction and electric charge between protein and ligand maybe considered as a combination of:

i. Hydrogen bonds (8-21 kJ mol^{-1}) between amide groups and hydrogen donors;
ii. Electrostatic bond carboxylate group and positively charged patches over the molecule (42 kJ mol^{-1}) and;
iii. Hydrophobic bond between phenyl (or/and) thiophilic group and hydrophobic patch of protein (4-8 kJ mol^{-1}) (Lehninger, Nelson, & Cox).

With the binding of the PI groups, the electrostatic and hydrophobic interactions were to reinforce the interaction, because of negatively ion charged and hydrophobic patches presence at optimal binding's conditions. The use of several types of interactions at the same time with relatively higher ligand density can achieve higher protein binding at a relatively higher salt concentration. The patatin was loosely bound in some of the experiments.

In mixed mode type of ligand, normally the elution process is provided simply by the electrostatic charge repulsion between charged groups of protein and the ionic group of the ligand (i.e. the carboxylic acid group). To induce this electrostatic charge repulsion, the pH was increased up to 12. The high pH step would help electrostatic repulsion to overcome the weaker hydrophobic binding interactions between aromatic/aliphatic ligand groups and the patches on the protein. However, the study of Chang et al. (Chang, Chou, Liu, & Tasi, 2007) showed that the adsorption protein ligand although with an elution buffer of pH 12 was more easily facilitated with 1M NaCl. From the MiMo experiment conducted in this work; it was observed that PI as well as patatin cannot be eluted completely with 3-5 of eluent (20 mM NaOH with pH around 12). So,

experiments carried out with 20 mM NaOH with pH around 12 at washing step in the chromatography. It also informed that Proteins can be more strongly retained on the multimodal resin and a different elution pattern was obtained as compared to anion exchange.

At the preliminary phases of this project, the first chromatography was performed with elution of 5 CV of pH 12. Protein bond strongly to the column and pH step elution did not significantly affect protein desorption. However, the main elution response was found in the high salt step. Thus, a higher concentration of salt or conductivity was necessary to elute the target proteins from the resin.

For the confirmation of binding pH several experiments were done. The pH of buffer with salts has been confirmed as the most important parameters for affecting the binding capacity for the PI adsorption on Capto MMC resin. Besides, the change of salt effects on buffer altered the charge distribution of PI significantly and possibly the conformation structure as well. The binding strength increased dramatically at same sample pH while applying salts with the elution buffer under an investigated range. Comparison between three different gradient elution showed that with the same sample pH with two different salts (1M NaCl; 1.5M NH_4Cl) thereby increasing the binding capacity and the better separation of peaks of PIs samples (Figure 19-a,b, c). It was considered that the large flow through peak in the chromatography consisted of sample proteins as well; therefore, the decision was made to carry out next experiment with salts to obtain higher amount of the target protein.

5.2 Effects of operating parameters on binding capacity

On the basis of this prediction the adsorption behavior might by caused by following reasons:

- It is known that numerous PI are presented in PFJ and their isoelectric point range from 5 to 10. The pH of a solution affects the charge states of both the protein and the resin. Thus, pH has a significant influence on the strength of electrostatic interaction between a protein and the ligand (Koningsveld, 2001).

- At adsorption pH > 10, the carboxyl groups of the ligand dissociate progressively, resulting in an increase of the negative charge of ligand. The negatively charged PI will be less retained by electrostatic interaction on the adsorbent as adsorption pH will increase (Chernakova, E., master, 2011).

Testing of binding condition plays a crucial role in protein separation process. The pH of CAP and of column equilibration buffer was confirmed as the most important parameters affecting the binding capacity of the potato protein on the Capto MMC resin. Variataion of pH both in CAP and buffer altered the binding capacity for PI. An optimal pH range was confirmed by the experimental results and this optimal range is 4.5 – 7.5. This condition was then applied both in buffer and CAP sample. After that, to find out the exact parameters for binding condition, a design of experiment (DoE) was performed. The binding capacity increases dramatically at buffer pH level 4.5 to 7.5 in the investigated range. The assessment of this optimal pH resulted in the most favorable compromise of total positive charge on PI and negative charge of ligand. It appeared that the electrostatic interaction contributes significantly in the binding process while the hydrogen bond and hydrophobic interaction contribute to the total binding process. The coefficient plot also discloses that there was a significant effect of buffer pH and design of experiment for this model is valid because statistical analysis R^2 has optimum score for an ideal experiment.

When high buffer pH (pH 7.5) and low sample pH (pH4.3) has been applied the PI (a) came earlier in flow throw areas; (b) there exists more patatin and (c) this patatin interrupts PI. For ignoring patatin and early PIs exitances, high buffer pH should be avoided. Concluding, while the optimum pH completely favorable for the binding of the PIs showed acceptable binding elution behavior on Capto MMC at adsorption CAP pH of 4.3 and buffer pH 5.5(Figure 22).

5.3 CAP adsorption vs. conductivity

It was observed that the increase of conductivity of CAP had no effect the overall protein binding as evaluated from eluted peak area. However, from the contour plot (figure 24) it is clear that the eluted peak area has effect on the low CAP pH at 4.5. The results were presented in different types of plots. The sample load, along with sample

conductivity was significant factors on of the PIs yields. Interestingly, major eluting fractions, PI were more completely bound at the low CAP pH, which means that some interaction (more likely, electrostatic) was contributed significantly to the binding process.

The surface properties protease inhibitors have an important influence on the binding behavior of the protein. The protein is surrounded by an electrical double layer so it can be considered as a colloidal particle. The potato PIs thus, has positive surface charge at optimal pH 4.5. The negative charge at the protein surface affects the ion distribution in the nearby region forming an electrical double layer in the region of the particle liquid surface. In other words, the thickness of the double layer depends on the concentration of the ion in the solution, the higher the conductivity the more compressed is the double layer (Everett, 1988). This statement can be suitable for the investigation of high salt concentration, which most likely was encountered in the screening experiment of conductivities 1.5- 20 mS/cm. When the CAP pH level was increased from 4.5 to 5.5 and then to 6.5 the linear lines go far away from each other, which mean there is a significant difference between the eluted peak areas at pH 6.5. In contrast, for the same pH (at 4.5) level with different conductivities (1.5, 10 and 20) there is not too much difference between the response values of eluted peak area. It is more revealed from the contour plots that high conductivity does not effect on CAP pH. And this process, actually, cause the increase of protein solubility due too electrostatic repulsion between CAP proteins and negatively charge ligands.

Based on mentioned DoE, sample conductivity is not the only factor, while need to investigate buffer load as well for optimizing the amount of PIs. Investigations of behavior of potato proteins at higher salt concentrations are required.

5.4 Effect of mobile phase modifiers on elution profile. Arginine as critical selectable modifier

The PFJ proteins were selected to sample a wide range of chemical and physical properties, with proteins possessing various degrees of hydrophobicity, charge, and size. Upon initial inspection, it is clear that all of the proteins in CAP exhibited all exhibited strong binding to the multimodal resin. Chromatographic retention time was determined

for proteins binding in Hi Screen Capto MMC chromatographic system.

A variety of mobile phase modifiers (arginine, guanidine, lysine, glycol) were employed in chromatographic experiments to examine the effects of these modifiers on protein-binding behavior. According to the Holstein al and co-workers (Holstein et al., 2012) it was reported that arginine and guanidine were found to have dramatic effects on protein binding, yielding changes in selectivity with both ion exchange and multimodal chromatographic systems.

Figure 25 shows the elution salt concentrations for the protein library on a multimodal cation exchanger (Capto MMC) in the absence and presence of 0.025M arginine and guanidine in the mobile phase. As it can be seen in the figure, protein exhibited a decrease in concentration in the absence of modifier, when the presence of modifiers had vice versa effect. The PI in CAP exhibited more binding when arginine was applied than for the same amount of guanidine. According to obtained results and gel performance arginine salt has been therefore chosen for further experiments in the study because arginine has more ability to capture the protein surface due to charge repulsion effect as well as hydrogen destabilized bonding effect to aggregation.

Addition of 0.025 – 0.1 M arginine into the column elution buffer showed a greater recovery of aggregated proteins. When applying four times higher concentration of arginine salt, it was found higher recovery of PI measured as eluted peak area. It is also compared with lysine salt and showed that there was no significant different between arginine and lysine (Figure 26). Arginine and lysine are the most basic amino acids with a pK of ~12.5 and 10.5. Arginine and lysine forms electrostatic complexes with buffer (salt bridges), destabilized surface structure. It also interacts with acidic groups being bound to a protein, such as the phosphate groups of nucleic acids.

Arginine and lysine also mostly exposed at surface solubilize proteins and interact with nucleic acids negatively charged phosphate backbone. The long side chains of lysine and arginine are highly flexible and due to its amine group a great solubilizer of protein inhibitors.

PI recovery is obtained because of charge repulsion effect and hydrogen destabilized bonding effect the protein binding with the help of different modifiers.

5.5 Bioinformatics analysis to predict protein surface characterization

In the continuation of this study on potato protease inhibitors, the present study tried to predict the protein surface characterization of a protein termed Kunitz-type protease inhibitors (aspartic protease inhibitors, serine protease inhibitor) from S. tuberosum. Some examples of protease inhibitors, which already were established in the PDB database butother types of PIs from potato (*S. tuberosum*) were not yet established. This thesis paper aimed to predict some of those proteins employing Phyre2 online server. From this study were revealed the predicted structure of three types of inhibitors (Serine protease inhibitor 1(>sp|P58514), Proteinase inhibitor type-2 P303.51 (>sp|Q41488), Aspartic protease inhibitor 10(>sp|Q03197)), which showed 100% confidence of this sequence and coverage are 81%, 68% and 79% respectively. By this result, it is clear that there is a higher chance of building up 3D model of Serine protease inhibitor 1 and this will help to predict to surface analysis of protease inhibitors and other types of inhibitors has possibility to build up surface modeling. According to (Woo et al., 2015) Capto MMC ligand could be recognizing different hydrophobic surface features based on the type of side chain (aromatic or aliphatic) exposed on the protein surface. Based on this hypothesis, the proteins in the library could be classified according to the hydrophobic clusters present on their surfaces. These results indicate that it may be possible to employ mobile phase modifiers under appropriate conditions to enhance the selectivity in systems (Holstein et al., 2012).

For further work, it is suggested to repeat the experiments in Pymol and molecular docking software, which may be achieved more precise outcomes.

5.6 Suggestion for further work

It was revealed during this study that, the Capto MMC resin has a narrow adsorption range for the total protein fraction. It was also predicted and observed that patatin could not be completely removed from the protease inhibitors by the acidification. Indeed, patatin is the sample protein fraction which could be easier adsorbed to the ligand. The PFJ should be more acidified to remove patatin group. If patatin group is

removed, capacity for binding protease inhibitors will increase. That is why, the further investigation of Capto MMC at other binding conditions are suggested for the protein isolation from PFJ. Along with this, the study on new elution condition is also suggested.

- This study only carried out 1-5ml columns. For the further work it is suggested to evaluated process at higher flow rates as well as using larger columns (50,125 ml) to investigate large scale production.
- It would be interesting to investigate chemical interaction of the protein with Capto MMC resin. In particular, it will help in understanding the behavior of the arginine salt with the protein surface.
- Other MiMo ligands, are suggested to investigate, since broad range of resins is offered on the market today, some of them more inexpensive than MMC.
- Regarding the surface modeling tertiary model of some protease inhibitors can be developed with different bioinformatics tools Pymol, Conserved domain database (CDD), Structural homologues searching tool DALI server AutoDock tool (ADT) etc.

- In consideration to this study model, if it can efficiently isolate protease inhibitors from this available protein source, then it might be applied on an industrial scale to meet market demand for inhibitors.

6. CONCLUSIONS

The main objectives have been addressed during the different steps of the study, and the following conclusions are evident:

- Potato proteins, including the major fraction of proteinase inhibitors can be isolated and fractionated them into groups acoording to charge and hydrophobicity from CAP.
- PIs can be captured by Mixed Mode chromatography with HiTrap Capto MMC column and observation from gel analysis is also clearly supporting evidence.
- The result shows, because of isoelectric point and charge of ligand, in combination with hydrophobic interaction pH have significant effect on the protein binding from acid treated PFJ (CAP). In comparison with CAP samples pH 4.3- 6.5, the maximum protein is bound from CAP sample precipitated at pH 4.3.
- The three different conditions of sample were prominent because of the bound protein represented by eluted peak area was higher in NaCl-gradients. Therefore, in the elution buffer, NaCl was used for each experiment.
- Optimizing binding conditions through screening by DoE attained the maximum yield of PIs. Maximum protein as observed by EPA is bound at column buffer pH 5.5 and CAP sample pH 4.50. The result of the experiment indicated that the highest amount of protein could be obtained from CAP pH 4.5 in combination with the buffer pH 5.5 in chromatography.
- Screening of conductivity versus sample CAP pH showed that the maximum of the total protein fractions is affected by both, showing interaction, although not significant. Maximum protein bound at 4.5 and conductivity was 11.5 mS/cm. These values conformed to the ranges predicted theoretically because carboxyl groups of the ligand dissociate progressively when increasing pH, resulting in an increase of the negative charge of ligand.
- Testing of mobile phase modifiers effects on eluting condition confirmed that same as high pKa value of lysine, arginine has the ability to capture the protein

surface due to charge repulsion effect and hydrogen destabilized bonding effect leading to aggregation.

- By knowledge of data sequence and Phyre 2 it was possible to obtain tertiary structures of three PPIs not yet solved. But did the structures – both the ones obtained and the published structures bring any closer to the surface topology required to explain the behavior of the PPIs in mixed mode chromatography.
- The PPI a useful model system to understand interaction between mixed mode ligands and protein surface, further it suggested to analyze the chemical interaction of the protein.

7. REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology, 215*(3), 403-410.

Andrews, D. L., Beames, B., Summers, M. D., & Park, W. D. (1988). Characterization of the lipid acyl hydrolase activity of the major potato (Solanum tuberosum) tuber protein, patatin, by cloning and abundant expression in a baculovirus vector. *Biochem. J, 252*, 199-206.

Ballabio, D., Consonni, V., & Todeschini, R. (2007). Classification of multiway analytical data based on MOLMAP approach. *analytica chimica acta, 605*(2), 134-146.

Bártová, V., & Bárta, J. Effect of heat treatment on re-solubility of potato proteins isolated from industrial potato fruit juice.

Birk, Y. (2003). *Plant protease inhibitors: significance in nutrition, plant protection, cancer prevention and genetic engineering*: Springer Science & Business Media.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry, 72*(1), 248-254.

Burton, S. C., & Harding, D. R. K. (1998). Hydrophobic charge induction chromatography: salt independent protein adsorption and facile elution with aqueous buffers. *Journal of Chromatography a, 814*(1), 71-81.

Burton, S. C., & Harding, D. R. K. (2001). Salt-independent adsorption chromatography: new broad-spectrum affinity methods for protein capture. *Journal of biochemical and biophysical methods, 49*(1), 275-287.

Castells, M. C., Pascual, C., Esteban, M. M., & Ojeda, J. A. (1986). Allergy to white potato. *Journal of Allergy and Clinical Immunology, 78*(6), 1110-1114.

Chang, Y.-K., Chou, S.-Y., Liu, J.-L., & Tasi, J.-C. (2007). Characterization of BSA adsorption on mixed mode adsorbent: I. Equilibrium study in a well-agitated contactor. *Biochemical Engineering Journal, 35*(1), 56-65.

Chernakova, E. (2011). Chromatographic isolation of native potato proteins on a mixed mode resin: Optimazation, scale up, *in silico* simulation and economic evaluation of an industrial process, Master thesis, Hedmark university college, Campus hamar. 91-92.

Ejima, D., Yumioka, R., Arakawa, T., & Tsumoto, K. (2005). Arginine as an effective additive in gel permeation chromatography. *Journal of Chromatography a, 1094*(1), 49-55.

Everett, D. H. (1988). *Basic principles of colloid science*: Royal Society of Chemistry.

Gao, D., Lin, D.-Q., & Yao, S.-J. (2006). Protein adsorption kinetics of mixed-mode adsorbent with benzylamine as functional ligand. *Chemical engineering science, 61*(22), 7260-7268.

Gao, D., Lin, D.-Q., & Yao, S.-J. (2007). Mechanistic analysis on the effects of salt concentration and pH on protein adsorption onto a mixed-mode adsorbent with cation ligand. *Journal of Chromatography B, 859*(1), 16-23.

Gao, D., Lin, D.-Q., & Yao, S.-J. (2008). Patch controlled protein adsorption in mixed-mode chromatography with benzylamine as functional ligand. *Biochemical Engineering Journal, 38*(3), 355-361.

Grosse-Holz, F. M., & Hoorn, R. A. L. (2016). Juggling jobs: roles and mechanisms of multifunctional protease inhibitors in plants. *New Phytologist*.

Han, J., Kamber, M., & Pei, J. (2011). *Data mining: concepts and techniques: concepts and techniques*: Elsevier.

HEALTH, G. *GEHealthcare. (2014). Design of Experiments in Protein Production and Purification Hands book from GE Healthcare.*

Holstein, M. A., Parimal, S., McCallum, S. A., & Cramer, S. M. (2012). Mobile phase modifier effects in multimodal cation exchange chromatography. *Biotechnology and bioengineering, 109*(1), 176-186.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research, 37*(1), 1-13.

Huang, J.-X., & Guiochon, G. (1989). Applications of preparative high-performance liquid chromatography to the separation and purification of peptides and proteins. *Journal of Chromatography B: Biomedical Sciences and Applications, 492*, 431-467.

Huang, Y.-H., Rose, P. W., & Hsu, C.-N. (2015). Citing a Data Repository: A Case Study of the Protein Data Bank. *PloS one, 10*(8), e0136631.

Janson, J.-C. (2012). *Protein purification: principles, high resolution methods, and applications* (Vol. 151): John Wiley & Sons.

Johansson, B.-L., Belew, M., Eriksson, S., Glad, G., Lind, O., Maloisel, J.-L., & Norrman, N. (2003). Preparation and characterization of prototypes for multi-modal separation media aimed for capture of negatively charged biomolecules at high salt conditions. *Journal of Chromatography a, 1016*(1), 21-33.

Jongsma, M. A. (1995). The resistance of insects to plant proteinase inhibitors.

Kallberg, K., Johansson, H. O., & Bulow, L. (2012). Multimodal chromatography: an efficient tool in downstream processing of proteins. *Biotechnology journal, 7*(12), 1485-1495.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols, 10*(6), 845-858.

Koningsveld, v. G. (2001). Physico-chemical and functional properties of potato proteins. *Wageningen University: Wageningen, The Netherlands*.

Kumar, M., Baishaya, L. K., Ghosh, D. C., Gupta, V. K., Dubey, S. K., Das, A., & Patel, D. P. (2012). Productivity and soil health of potato (Solanum tuberosum L.) field as influenced by organic manures, inorganic fertilizers and biofertilizers under high altitudes of eastern Himalayas. *Journal of Agricultural Science, 4*(5), p223.

Lehninger, A. L., Nelson, D. L., & Cox, M. M. Principles of Biochemistry, (1993). *Worth, New York*.

Løkra, S., Helland, M. H., Claussen, I. C., Strætkvern, K. O., & Egelandsdal, B. (2008). Chemical characterization and functional properties of a potato protein concentrate prepared by large-scale expanded bed adsorption chromatography. *LWT-Food science and Technology, 41*(6), 1089-1099.

McDonald, J. K., Schwabe, C., & Owers, N. O. (1987). Peptidases in connective tissue degradation and remodelling in reproductive and invasive tissues *Cells, Membranes, and Disease, Including Renal* (pp. 335-349): Springer.

Melander, W., & Horváth, C. (1977). Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: an interpretation of the lyotropic series. *Archives of biochemistry and biophysics, 183*(1), 200-215.

Natu, R. B., & Mazze, G. (1991). Waste utilization. Potato: Production, Processing, and Products.(Salunke, DK, Kadam, SS, & Jadhav, SJ, Eds.): Boca Raton: CRC Press. 0.

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., . . . Ako-Adjei, D. (2015). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic acids research*, gkv1189.

Pots, A. M. (1999). Physico-chemical properties and thermal aggregation of patatin, the major potato tuber protein.

Pots, A. M., de Jongh, H. H. J., Gruppen, H., Hessing, M., & Voragen, A. G. J. (1998). The pH dependence of the structural stability of patatin. *Journal of agricultural and food chemistry, 46*(7), 2546-2553.

Pouvreau, L., Gruppen, H., Piersma, S. R., van den Broek, L. A., van Koningsveld, G. A., & Voragen, A. G. (2001a). Relative abundance and inhibitory distribution of protease inhibitors in potato juice from cv. Elkana. *J Agric Food Chem, 49*(6), 2864-2874.

Pouvreau, L., Gruppen, H., Piersma, S. R., Van den Broek, L. A. M., Van Koningsveld, G. A., & Voragen, A. G. J. (2001b). Relative abundance and inhibitory distribution of protease inhibitors in potato juice from cv. Elkana. *Journal of agricultural and food chemistry, 49*(6), 2864-2874.

Pouvreau, L., Gruppen, H., van Koningsveld, G. A., van den Broek, L. A. M., & Voragen, A. G. J. (2003). The most abundant protease inhibitor in potato tuber (cv. Elkana) is a serine protease inhibitor from the Kunitz family. *Journal of agricultural and food chemistry, 51*(17), 5001-5005.

Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*: Cambridge University Press.

Ralet, M.-C., & Guéguen, J. (2000). Fractionation of potato proteins: solubility, thermal coagulation and emulsifying properties. *LWT-Food science and Technology, 33*(5), 380-387.

Rexen, B. (1976). Studies of protein of potatoes. *Potato Research, 19*(2), 189-202.

Rowan, A. D., Brzin, J., Buttle, D. J., & Barrett, A. J. (1990). Inhibition of cysteine proteinases by a protein inhibitor from potato. *FEBS letters, 269*(2), 328-330.

Sasaki, I., Gotoh, H., Yamamoto, R., Hasegawa, H., Yamashita, J., & Horio, T. (1979). Hydrophobic-Ionic Chromatography Its Application to Purification of Porcine Pancreas Enzymes. *Journal of biochemistry, 86*(5), 1537-1548.

Sasaki, I., Gotoh, H., Yamamoto, R., Tanaka, H., Takami, K.-i., Yamashita, K., . . . Horio, T. (1982). Hydrophobic-ionic chromatography: Its application to microbial glucose oxidase, hyaluronidase, cholesterol oxidase, and cholesterol esterase. *Journal of biochemistry, 91*(5), 1555-1561.

Schägger, H. (2006). Tricine–SDS-PAGE. *Nature Protocols, 1*, 16-22.

Schoenbeck, I., Graf, A. M., Leuthold, M., Pastor, A., Beutel, S., & Scheper, T. (2013). Purification of high value proteins from particle containing potato fruit juice via direct capture membrane adsorption chromatography. *Journal of biotechnology, 168*(4), 693-700.

Shah, K. R., Patel, D. K., Pappachan, A., Prabha, C. R., & Singh, D. D. (2015). Characterization of a Kunitz-type serine protease inhibitor from Solanum tuberosum having lectin activity. *International Journal of Biological Macromolecules*.

Simmonds, R. J., & Yon, R. J. (1976). Protein chromatography on adsorbents with hydrophobic and ionic groups. Chromatography of dodecyl sulphate-solubilized proteins of the human erythrocyte membrane on N-(3-carboxypropionyl) aminodecyl-sepharose. *Biochem. J, 157*, 153-159.

Spelbrink, R. E. J., Gerrits, P. J., Mooij, C., & Giuseppin, M. L. F. (2011). Quantitative Determination of Trypsin Inhibitory Activity in Complex Matrices. *Open Food Science Journal, 5*, 42-46.

Staahlberg, J., Joensson, B., & Horvath, C. (1992). Combined effect of coulombic and van der Waals interactions in the chromatography of proteins. *Analytical chemistry, 64*(24), 3118-3124.

Strætkvern, K. O., Lokra, S., Olander, M. A., & Lihme, A. (2005). *Food-grade protein from industrial potato starch effluent recovered by an expanded bed adsorption process.* Paper presented at the Journal of biotechnology.

Straetkvern, K. O., Schwarz, J. G., Wiesenborn, D. P., Zafirakos, E., & Lihme, A. (1999). Expanded bed adsorption for recovery of patatin from crude potato juice. *Bioseparation, 7*(6), 335-347.

Suh, S.-G., Peterson, J. E., Stiekema, W. J., & Hannapel, D. J. (1990). Purification and characterization of the 22-kilodalton potato tuber proteins. *Plant physiology, 94*(1), 40-45.

Sun, Y., Liu, L., Jiang, L.-Z., Zhang, G.-F., Li, G.-M., & Wu, N. (2013). Preparation, identification, structure, and in vitro anti-obesity effects of protease inhibitors isolated from potato fruit juice. *European Food Research and Technology, 237*(2), 149-157.

Tsumoto, K., Ejima, D., Senczuk, A. M., Kita, Y., & Arakawa, T. (2007). Effects of salts on protein–surface interactions: applications for column chromatography. *Journal of pharmaceutical sciences, 96*(7), 1677-1690.

Valueva, T. A., Revina, T. A., Kladnitskaya, G. V., Mosolov, V. V., & Mentele, P. (1999). Primary structure of a 21-kD protein from potato tubers. *Biochemistry c/c of biokhimiia, 64*(11), 1258-1265.

Walsh, T. A., & Twitchell, W. P. (1991). Two Kunitz-type proteinase inhibitors from potato tubers. *Plant physiology, 97*(1), 15-18.

Woo, J., Parimal, S., Brown, M. R., Heden, R., & Cramer, S. M. (2015). The effect of geometrical presentation of multimodal cation-exchange ligands on selective recognition of hydrophobic regions on protein surfaces. *Journal of Chromatography a, 1412*, 33-42.

Yon, R. J. (1972). Chromatography of lipophilic proteins on adsorbents containing mixed hydrophobic and ionic groups. *Biochemical Journal, 126*(3), 765.

Zhao, G., Dong, X.-Y., & Sun, Y. (2009). Ligands for mixed-mode protein chromatography: principles, characteristics and design. *Journal of biotechnology, 144*(1), 3-11.

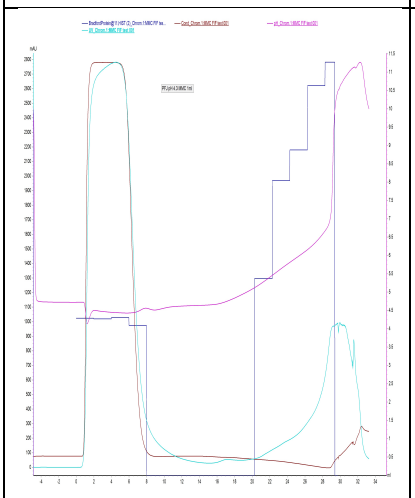
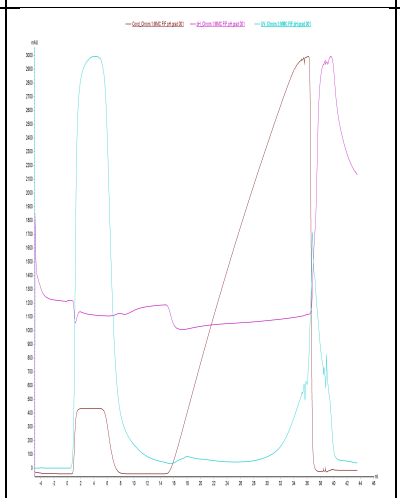
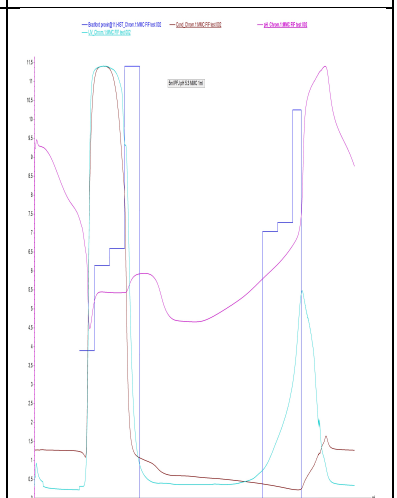
Zhao, G., & Sun, Y. (2007). Displacement chromatography of proteins on hydrophobic charge induction adsorbent column. *Journal of Chromatography a, 1165*(1), 109-115.

8. LIST OF APPENDIXES

Appendix1 - Mixing table for Tricine SDS-PAGE gel.

		Stacking gel (4 %)	**10 % gel**	**16 % gel**
Acrylamide	(ml)	0,5	3	5
Gel buffer (3X)	(ml)	1,5	5	5
Glycerol, 85 %	(g)	-	1,75	1,75
Add water to final volume:	(ml)	6 (=4 ml)	15 (=5.25)	15 (=5.25)
Polymerize by adding:				
APS (10 %)	(µl)	45	75	50
TEMED	(µl)	4.5	7.5	5

Appendix2 - Eluting condition pH and salts.

(A) Only pH	(B) pH + 0-1M NaCl	(C) pH + 0-1.5 M NH_4Cl
		

Appendix3 - Factor levels using two factors salt gradient with sample pH for each DoE, TIA response, ANOVA table and chromatography experiments.

Experiment no	Experiment name	Run order	Buffer pH	Sample pH	Eluted peak area
1	5	1	4.5	4.3	222.3
2	7	2	6.0	4.3	2953.8
3	8	3	7.5	4.3	1563.9
4	10	5	4.5	6.0	188.5
5	11	6	6.0	6.0	1154.2
6	9	4	7.5	6.0	1553.4
7	14	9	4.5	7.5	108.9
8	12	7	6.0	7.5	654
9	13	8	7.5	7.5	864.3

Blank corrected raw data (abs450)

	1	2	3	4	5	6	7	8	9	10	11	12
A	2.974	2.742	2.632	2.16	2.858	2.374	2.276	0.635	0.34			
B	2.694	2.591	2.293	2.56	2.119	0.383	0.356	2.746	1.811			
C	2.828	2.34	0.623	0.412	0.372	0.328	0.219	1.584	1.452			
D	1.045	0.711	0.586	0.592	0.871	0.422	0.503	0.532	3.636			
E	2.566	0.627	0.738	2.751	2.937	2.959	2.625	2.485	2.566			
F	3.016	0.556	0.534	1.489	0.347	0.53	0.77	1.96	1.769			
G	0.416	0.607	0.738	0.804	0.528	0.486	0.711	1.821	2.533			
H	3.224	3.025	0.043	0.073								
avg		3.1245		0.058								

% inhibition = [(PC - (Smpl -NC) / PC] * 100

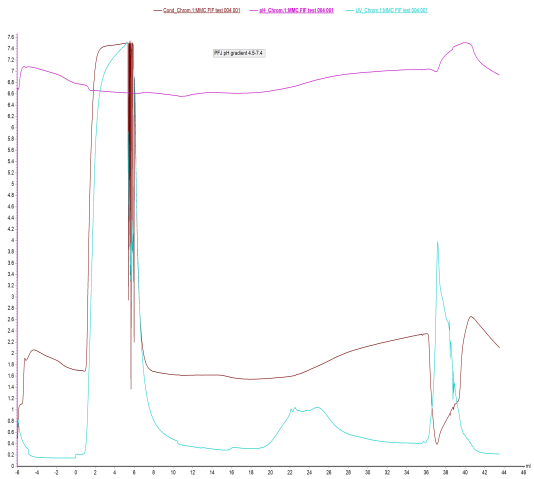
	1	2	3	4	5	6	7	8	9	sum of %inhibition	Buffer pH	CAP pH
A	6.7	14.1	17.6	32.7	10.4	25.9	29.0	81.5	91.0	**237.8**	4.5	4.3
B	15.6	18.9	28.5	19.9	34.0	89.6	90.5	14.0	43.9	**320.4**	6.0	4.3
C	11.3	27.0	81.9	88.7	90.0	91.4	94.8	51.2	55.4	**553.3**	7.5	4.3
D	68.4	79.1	83.1	82.9	74.0	88.4	85.8	84.8	0.0	**332.9**	7.5	6.0
E	19.7	81.8	78.2	13.8	7.9	7.2	17.8	22.3	19.7	**88.7**	4.5	6.0
F	5.3	84.1	84.8	54.2	90.8	84.9	77.2	39.1	45.2	**391.4**	6.0	6.0
G	88.5	82.4	78.2	76.1	85.0	86.3	79.1	43.6	20.8	**390.8**	6.0	7.5

DOE 1: sum of elute data

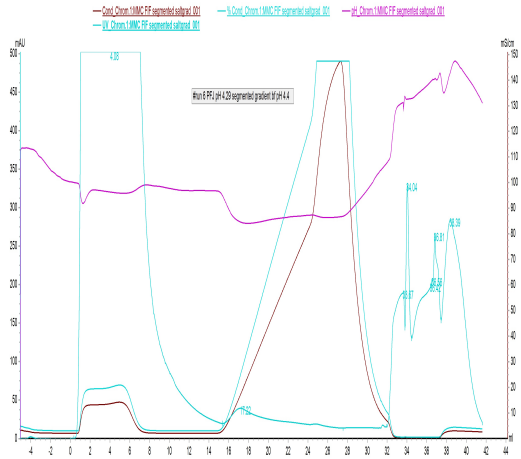
A= 8.483 for PFJ pH 4.3 bfr ppH 4.5; B= 12.268 for PFJ pH 4.3 bfr pH 6.0; C= 4.99 for PFJ pH 4.3 bfr pH 7.5; D= 5.964 for PFJ pH 6.0 bfr pH 7.5; E = 16.323 for PFJ pH 6.0 bfr pH 4.5; F = 6.865 for PFJ pH 6.0 bfr pH 6.0; G = 6.883 for PFJ pH 7.5 bfr pH 6.0

Eluted peak area	DF	SS	MS (variance)	F	p	SD
Total	8	7,46255e+006	932819			
Constant	1	4,97622e+006	4,97622e+006			
Total corrected	7	2,48633e+006	355190			595,978
Regression	3	2,25241e+006	750803	12,8387	**0,016**	866,489
Residual	4	233918	58479,6			241,826
Lack of Fit	4	--	--	--	--	--
(Model error)						
Pure error	0	--	--			--
(Replicate error)						
	N = 8	Q2 = 0,742		Cond. no. = 1,168		
	DF = 4	R2 = 0,906		RSD = 241,8		
	Comp. = 2	R2 adj. = 0,835				

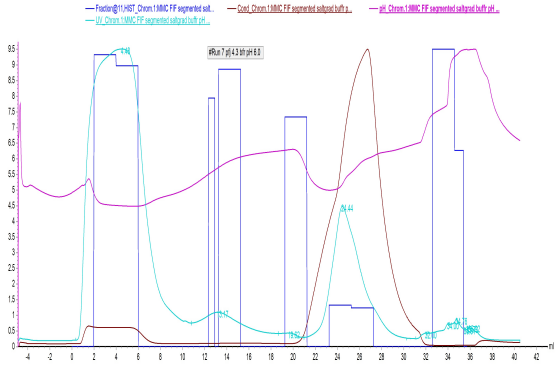
Trypsin Inhibition	DF	SS	MS (variance)	F	p	SD
Total	6	786986	131164			
Constant	1	663005	663005			
Total corrected	5	123981	24796,2			157,468
Regression	3	80766,2	26922,1	1,24597	**0,474**	164,079
Residual	2	43214,7	21607,4			146,994
Lack of Fit	2	--	--	--	--	--
(Model error)						
Pure error	0	--	--			--
(Replicate error)						
	N = 6	Q2 = 0,431		Cond. no. = 1,168		
	DF = 2	R2 = 0,651		RSD = 147		
	Comp. = 2	R2 adj. = 0,129				



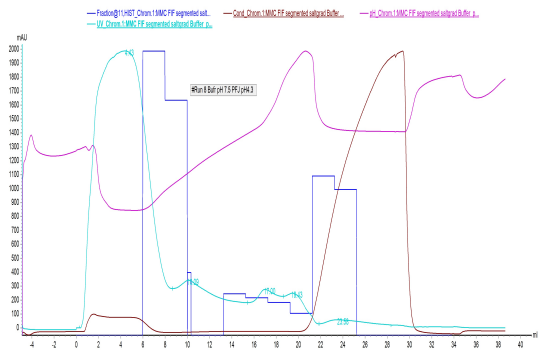
a) CAP pH 4.5 and buffer pH 7.5



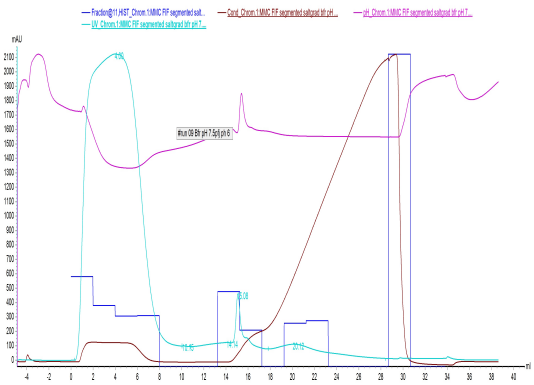
b) CAP pH 4.5 and buffer pH 4.5



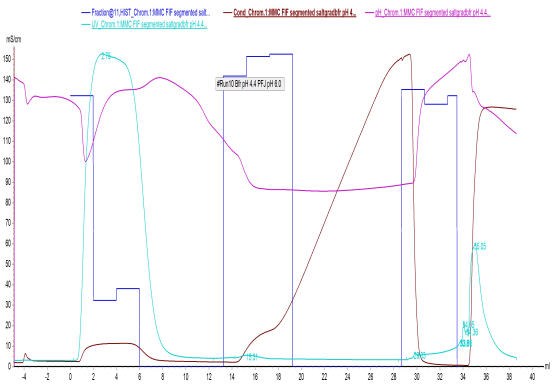
c) CAP pH 4.5 and buffer pH 6.0



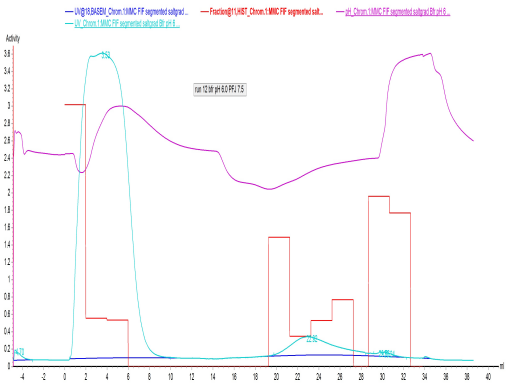
d) CAP pH 6.0 and buffer pH 6.0



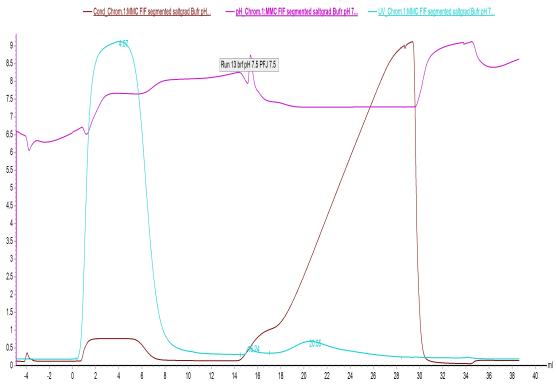
e) CAP pH 6.0 and buffer pH 7.5



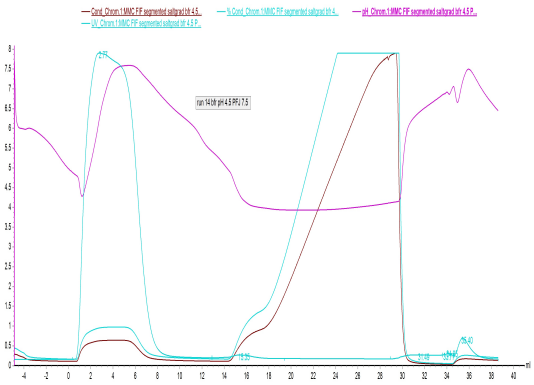
f) CAP pH 6.0 and buffer pH 4.5



g) CAP pH 7.5 and buffer pH 6.0



h) CAP pH 7.5 and buffer pH 7.5



i) CAP pH 7.5 and buffer pH 4.5

Appendix4- Effect of binding condition in the elution salt (TIA response).

1.115	1.166	1.110	1.225	1.098	1.057	1.160	1.145	1.133
1.146	0.575	0.479	0.461	0.406	0.400	0.598	0.485	0.526
0.944	0.958	1.063	0.923	0.935	0.920	0.969	0.889	0.878
1.030	0.893	0.677	0.450	0.510	0.529	0.644	0.717	0.787
1.374	1.487	0.242	0.330	0.278	0.391			
1.430	=PC	0.286	=NC	0.334	=Blank			
1.096		0						

nhibition = [(PC - (Smpl -NC) / PC] * 100

42.0	38.5	42.4	34.4	43.2	46.1	38.9	39.9	40.8
39.9	79.8	86.5	87.8	91.6	92.0	78.2	86.1	83.2
54.0	53.0	45.7	55.5	54.6	55.7	52.2	57.8	58.6
48.0	57.5	72.7	88.5	84.4	83.0	75.0	69.9	65.0

2 i.e. subtracted "blank" from all sample readings and using corrected PC and NC

38.4	33.8	38.8	28.4	39.9	43.7	34.3	35.6	36.7
35.5	87.7	96.4	98.1	103.1	103.6	85.5	95.9	92.1
54.0	52.7	43.1	55.9	54.8	56.2	51.7	59.0	60.0
46.2	58.6	78.3	99.0	93.6	91.9	81.4	74.7	68.3

Appendix5- Factor levels using two factors desalted sample gradient with connectivity for each DoE. ANOVA Table and TIA response.

Eluted peak area	**DF**	**SS**	**MS (variance)**	**F**	**p**	**SD**
Total	8	1,1293e+007	1,41163e+006			
Constant	1	1,0274e+007	1,0274e+007			
Total corrected	7	1,01896e+006	145566			381,531
Regression	3	791869	263956	4,64929	**0,086**	513,767
Residual	4	227094	56773,5			238,272
Lack of Fit	--	--	--	--	**--**	--
(Model error)						
Pure error	--	--	--			--
(Replicate error)						
	N = 8	Q2 = -0,979		Cond. no. = 1,442		
	DF = 4	R2 = 0,777		RSD = 238,3		
		R2 adj. = 0,610				
Trypsin Inhibition	**DF**	**SS**	**MS (variance)**	**F**	**p**	**SD**

Total	5	66560,5	13312,1			
Constant	1	56519,7	56519,7			
Total corrected	4	10040,8	2510,2			50,1019
Regression	3	9646,6	3215,53	8,1574	**0,251**	56,7057
Residual	1	394,186	394,186			19,8541
Lack of Fit (Model error)	--	--	--	--	--	--
Pure error (Replicate error)	--	--	--			--
	N = 5	Q2 = -2,257	Cond. no. = 1,119			
	DF = 1	R2 = 0,961	RSD = 19,85			
		R2 adj. = 0,843				

Blank corrected raw data (abs450)

1	2	3	4	5	6	7	8	9	10
1.242	1.199	1.261	1.121	0.704	1.205				
1.007	1.05	0.876	1.043	0.925	0.931	0.885	0.909	1.207	
0.506	1.3	1.24	1.578	1.17	0.904	0.984	1.214	0.894	1.204
1.094	1.135	1.281	1.594	1.798	1.748	1.222	0.608	0.036	
0.892	2.05	1.311	1.198	1.095	0.66	0.696	1.041	0.422	0.993
1.456	1.361	-0.168	-0.127						

DOE : sum of elute data

A = 4.291 bfr pH 4.5 Connectivity 1.5; B = 4.857 bfr pH 6.5 Connectivity 1.5; C = 9.188 bfr pH 5.5 Connectivity 10; D = 8.287 bfr pH 4.5 Connectivity 20; E = 4.907 bfr pH 6.5 Connectivity 20

Appendix6 - Different protease and proteinase inhibitors data sequences.

Potato cystatin

>2W9P:A|PDBID|CHAIN|SEQUENCEGIVNVPNPNNTKFQELARFAIQDYNKKQNA
HLEFVENLNVKEQVVAGIMYYITLAATDDAGKKKIYKAKIWVKEWEDFKKVVE
FKLV

Kunitz type chymotrypsin inhibitors

>3TC2:A|PDBID|CHAIN|SEQUENCELPSDATPVLDVTGKELDPRLSYRIISTFWGA
LGGDVYLGKSPNSDAPCANGVFRNSDVGPSGTPVRFIGSSSHFGQGIFEDELLNIQ
FAISTSKMCVSYTIWKVGDYDASLGTMLLETGGTIGQADSSWFKIVKSSQFGYNL
LYCPVTTSSDDQFCLKV GVVHQNGKRRLALVKDNPLDVSFKQVQ

Potato carboxypeptidase inhibitor (PCI)

>1H20:A|PDBID|CHAIN|SEQUENCEEQHADPICNKPCKTHDDCSGAWFCQACWN
SARTCGPYVG