Rotavirus VP7 bacterial expression, purification and characterisation

by

Nsovo Tonny Mathebula

Submitted in accordance with the requirements for the degree

Master of Life Science

in the

SCHOOL OF AGRICULTURE AND LIFE SCIENCES

COLLEGE OF AGRICULTURE AND ENVIRONMENTAL SCIENCES: DEPARTMENT OF LIFE AND CONSUMER SCIENCE

at the

UNIVERSITY OF SOUTH AFRICA

FLORIDA CAMPUS

SUPERVISOR

Prof. S. Gildenhuys

Co-SUPERVISOR

Dr. N. Parbhoo

December 2021

Dedications

I dedicate this work to my late grandmother **Mrs Hlamalani Mjaji Machovani** and late sister **Ms Ntshuxekani 'Masesi' Mathebula**. I would have loved to see you, see your grandson and brother, respectively, being the first in the family to start and complete his MSc and moving a step forward to a doctorate in honour of my late grandfather **Mr. Magezi Samuel Mathebula.** To my mother **Mrs Rirhandzu Patricia Mathebula**, it migh have taken long but it is done now, I know you have seen my frustrations and possibly depression because of it, however, I hope I made you proud. To my siblings, **Mr Ripfumelo 'Fana' and Ms Vutlhari Memory Mathebula**, let us continue to make our family more especially our dear manana, mhani, maOlady na sirha kokwani xinuna proud. My love for you **Mr Ripfumelo 'Fana', Ms Vutlhari Memory Mathebula, Mrs Rirhandzu Patricia Mathebula** cannot be measured on earth and heaven. To my late loved ones, **Ms Ntshuxekani 'Masesi' Mathebula, Mr. Magezi Samuel Mathebula** and **Mrs Hlamalani Mjaji Machovani** I hope I've made you proud wherever you are. Remember blood makes us related but Loyalty makes us family.

Declaration

I _**MATHEBULA NSOVO TONNY**_ hereby declare that the dissertation/thesis, which I hereby submit for the degree of **MASTERS IN SCIENCE** at the University of South Africa, is my own work and has not previously been submitted by me for a degree at this or any other institution.

I declare that the dissertation /thesis does not contain any written work presented by other persons whether written, pictures, graphs or data or any other information without acknowledging the source.

I declare that where words from a written source have been used the words have been paraphrased and referenced and where exact words from a source have been used the words have been placed inside quotation marks and referenced.

I declare that I have not copied and pasted any information from the Internet, without specifically acknowledging the source and have inserted appropriate references to these sources in the reference section of the dissertation or thesis.

I declare that during my study I adhered to the Research Ethics Policy of the University of South Africa, received ethics approval for the duration of my study prior to the commencement of data gathering, and have not acted outside the approval conditions.

I declare that the content of my dissertation/thesis has been submitted through an electronic plagiarism detection program before the final submission for examination.

Student signature: _____  ____ Date: **December 2021**_

Acknowledgements

Firstly I would like to thank my **Father God (YHWH)** for giving me the opportunity to further my studies and then giving me the strength and determination to complete it even though it felt and seemed impossible.

I would like to express my sincere gratitude to the following individuals and organisations for their assistance and support towards the completion of this study

Special thanks to **Professor Samantha Gildenhuys** in giving me this opportunity to work in her laboratory and guiding me throughout my study even though I would stress her at times.

Professor **Samantha Gildenhuys** and **Doctor Nishal Parbhoo** for their assistance throughout my studies, by helping me understand the research, teaching me critical thinking and letting me voice my opinions and ideas.

Special thanks to **Samuel G. Kgokolo** and **Bonnie Russell** for their assistance, their shared knowledge, ideas and having fun in the lab.

Professor S.R. Magano for his counselling, guidance and encouragements during tough times, it was never easy and I will never forget it. Stay blessed.

My family and friends for their support and patience though they failed to understand what I do. Thanks to my uncle **Mr Isaac Machovani** for abstract translation to xitsonga

NRF and UNISA for their financial assistance.

Quotes

“Even the young ones can give you advice and knowledge you can’t get from elders, so listen” – **Vutlhari Memory Mathebula**

“Someone one said to me, ‘research is RE… SEARCH, so that means you search and search until you find it’” – **Thabelo Roseline Radzilane**

“I understand your pain, don’t give up yet, it’s just the last lap ma dawg”—**Samuel Given Kgokolo**

“I’m short but if I can’t reach what’s up there, I’ll ask for help. So asking for help doesn’t make you useless but helps you defeat some challenges you might have” – **Nicky Verhoog**

“Whatever the height of imagination and dreams your mind can reach, then the depth of your courage and capabilities can manifest” – **Seanaphoka Tsapi**

“Don’t be scared to share with your superior your thoughts and ideas, those thoughts might help solve the biggest problems the world has” – **Stranger in the walks of life**

“I hated every minute of training, but I said, don’t quit. Suffer now and live the rest of your life as a champion” – **Muhammad Ali**

“Let us sacrifice our today so our children can have a better tomorrow” – **Abdul Kalam**

“Repetition is the father of learning” – **Dwayne “Lil Wayne” Carter**

“The fear of the Lord is the beginning of knowledge, but fools despise wisdom and instruction” – **Proverb 1:7**

“If it is not going your way, surely you took a wrong turn” – **Unknown**

“Life is not fair why play fair”**-- Given Muthuvhuli**

“You don’t understand and probably won’t, because I don’t know how to make you understand and know what I know and how I feel about everything”— **Nsovo T. ‘Hlayisani’ Mathebula**

“Therapy doesn’t help, it makes things worse. I’ve asked those question 10 times before they asked, still no answers to remedy emotions”--- **Nsovo T. ‘Hlayisani’ Mathebula**

“Success and money is the motivation”- **Unknown**

Abstract

Reoviridae is a family of icosahedral viruses such as rotavirus, a virus that infects both humans and animals, resulting in dehydration and diarrhoea. Rotavirus infections cause high mortality rates in infants in low and middle income countries. The virus contains several different structural proteins, with VP7 acting as an outer capsid layer protein. This study aimed to overexpress, purify and characterise rotavirus VP7. VP7 was expressed predominantly in the insoluble fraction of both BL21 (DE3) and C43 (DE3) pLysS bacterial cells regardless of temperature and induction time conditions. Purifying VP7 using ion exchange and affinity (His-Tag) chromatography via immobilised metal affinity chromatography (IMAC) from the soluble fraction was not successful, but purification of solubilised protein from the insoluble fraction resulted in 100% pure VP7 with a single affinity (His-Tag) chromatography step. A yield of 0.013 mg VP7 was obtained from a litre of bacterial culture. Purified VP7 showed secondary structural characteristics of the protein. Intrinsic tryptophan fluorescence indicated a tertiary structure of a folded protein, which, on heating, undergoes conformational rearrangements at temperatures ranging from 20 to 90°C but never completely unfolds.

Key words: Affinity (His-Tag) chromatography, Bacterial expression, Ion-exchange chromatography, Secondary structure CD spectrum, Fluorescence spectrum, Rotavirus, VP7.

Abstract (XiTsonga)

Reoviridae i matsambu wa ndyangu wun'we wa swi tsongwatsongwani w Icosahedral viruses i xitsongwatsongwani xa tlawa wun'we xikan'we na lexi vange i Rotavirus, i swi tsongwatsongwani kumne ti virus leti khomaku kumbe ku tlulela eka vanhu na swiharhi, leswi heteleku switisa ku omeriwa ka mikolo, torha na chuluko. Ntluleto wa Rotavirus swi vanga mafu yale henhla eka vatsongwani kumbe eka vana vatincece eka matiku lawaha hluvukaku na lawa ya hanyaka evuswetini. Xi tsongwatsongwani kumbe virus leyi, yi kumeka yiri na nhlayo yale henhla ya swiga swo hambana hambana swati protein. Yin'wani yatona kunga leyi vuriwaka VP7 leyi tirhaku tani hi xikhamba xa protein. Ntsariso dyondo lowu wuna xikongomelo xo komba ntikelo wa kona, ku basisiwa naku hambana hambana ka Rotavirus VP7. VP7 yi kombisile ntikelo wa yona na matimba yo thlandlukela ehenhla ku tlula swiga swa BL21 (DE3) na C43 (DE3) pLysS bacterial cells handle ka kuhambana ka timheho tama hiselo ya tshamiseko wa mikarhi. Sawutiso kumbe ku basisiwa ka VP7 hiku tirhisa cincaniso wa swihlawulekiso swa ion na ntshembo (His Tag) chromatography kusuka eka

swiga hikwaswo leswi nga tirhisiwa naku kutsandeka is swona hi thlelo ra wona, kambe ntirho wo hluta ku mbhasiso nawo nhlantswa awu humelelangi, kambe mbhasiso wa swihlawulekiso wale hehla wa ti protein ku suka eka swiphemu hikwaswo swa swiga swantirho leswi nga hluleka swi tise mbuyelo wo basa hi xiga xa 100% VP7 hi katsakanyo wo twisiseka wa goza ra (His Tag) chromatography. Ku humelela kale hansi ka 0.013 mg hi litara ya ndhavuko wa switsongwatsongwani kumbe leswi vuriwaka bacterial culture ku kumekile naku humela eri valeni. Mbhasiso katsakanyo hika mbirhi wa VP7 wu konbiso mbuyelo wa xivumbeko xaku hambana hambana kati protein. Nkombiso wa mbuyelo wale henhla wa Intrinsic tryptophan fluorescence wu kombise songano wati protein leti laha ku hisaku ti kombisa mbuyelo wa hlanganiso wa timheho ta mahiselo yo suka eka 20 kuya fika eka 90°C kambe swinga vonaki hiku hetiseka kuya emahlweni.

Abstract (isiZulu)

I-Reoviridae wumndeni wamagciwane e-icosahedral afana ne-rotavirus, igciwane elithelela abantu nezilwane okuholela ekubanjweni yisisu. Ukutheleleka nge-Rotavirus kubangela amazinga aphezulu okufa ezinganeni emazweni asathuthuka. Leli gciwane liqukethe amaprotheni ahlelekile ahlukene, i-VP7 isebenza njengeprotheyini engqimba yangaphandle ye-capsid. Lolu cwaningo kuhloswe ngalo ukusebenzisa i-overexpress, ukuhlanza kanye nokufaka i-rotavirus VP7. I-VP7 ivezwe kakhulu engxenyeni engafakwanga ye-BL21 (DE3) ne-C43 (DE3) pLysS amaseli wegciwane ngaphandle kwezimo zokushisa nezikhathi zokungeniswa. Ukuhlanza i-VP7 kusetshenziswa ukushintshaniswa kwe-ion nobudlelwano (i-HisTag) i-chromatography evela engxenyeni encibilikayo akuphumelelanga kepha ukuhlanjululwa kwamaprotheni ancibilikisiwe aqhamuka engxenyeni engancibiliki kuholele ku-100% VP7 emsulwa ngesinyathelo esisodwa sokusondelana (HisTag) chromatography. Kutholwe isivuno esiphansi se-0.013 mg ilitha ngalinye lesiko lamagciwane. I-VP7 ehlanziwe ikhombise ukwakheka kwesibili kweprotheni. I-tryptophan fluorescence yangaphakathi ikhombise ukwakheka okuphakeme kweprotheyini eligoqiwe okuthi lapho kufudunyezwa kube nokuhlelwa kabusha okuhambisanayo emazingeni okushisa asukela ku-20 kuye ku-90 ° C kepha kungaze kwenzeke ngokuphelele.

Table of contents

Abbreviations

ATP	Adenosine triphosphate
DLP	double layer particle
DTT	dithiothreitol
ER	endoplasmic reticulum
HisTag	poly-hexa-histidine (histidine residue tag)
IL	inner layer
IPTG	isopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
ML	middle layer
NSP	non-structural protein
OL	Outer layer
pI	Isoelectric point
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
TEMED	N, N, N', N'-tetramethylethylene-1,2-diamine
TLP	Triple layer particle
Tris	Tris (hydroxymethyl) aminomethane (Trizma base)
VP	viral protein

The IUBMB-IUPAC abbreviations for amino acids single and triple codes were used

1. Introduction

1.1 Protein expression systems

Protein research requires scientists to be able to overexpress recombinant proteins (Chen, 2012) because the protein quantity expressed by cells is not sufficient for detection in research techniques and processes to continue with the aims and objectives of research studies. Depending on the research objective and the nature of the protein, recombinant proteins are expressed in different expression systems. There are a number of expression systems to choose from, such as the mammalian, plant, insect and bacterial-based expression systems.

Mammalian cells such as the Chinese hamster ovary cells and the baby hamster kidney cells have been the most predominant choices for mammalian cell expression of recombinant proteins (Andersen *et al.,* 2002). The mammalian cells are transfected with a plasmid engineered with the recombinant protein coding gene to over-express the target protein using the cells' machinery (Fan and Steitz, 1998). The use and selection of specific mammalian cells that results in specific environmental and additive control to maintain specific cell metabolism, growth and death rate based on its metabolic capacity (Andersen *et al.,* 2002) play a great deal in recombinant expression. The mammalian cell system has been found to have useful control strategies, such as the control of DNA transcription and translation to polypeptides (Wurm *et al.,* 2016), protein phosphorylation (Corti and Simons, 2016), tumour suppressor protein, and cytostasis product expression (Andersen *et al.,* 2002). Though they have useful control strategies, the mammalian cells have their own set of limits and drawbacks such as the low volumetric productivity, high costs and risks of viral, bacterial and cell contamination (Wurm *et al., 2016)*.

Plants and insects are the other expression systems used in over-expressing recombinant proteins. Plant (Varraelmann *et al.,* 2000; Gleba *et al.,* 2007) and insects cells (van Hulten *et al.,* 2000) are infected with virions containing the genetic makeup of the protein of interest, like in most systems (van Hulten *et al.,* 2000), the virion then utilises the cells' machinery to overexpress the protein of interest. The insects' cells are mostly used for target genetic sequence expression and screening (Andersen *et al.,* 2002) the target gene sequence with the presence of specific post-translational modifications such as glycosylation which are a major challenge for the generation of many recombinant proteins. Most plant expression systems have an advantage of performing proteins post-translational modifications (PTMs) which plays an important role in maintaining the recombinant protein's antigenicity (Rigano and Walmsley, 2005). Plant expression systems also have

the advantage of the possibility of oral administrations of recombinant proteins from transgenic plants with a rigid cell wall that can be protective of the recombinant antigens from the stomach acidic environment (Rigano and Walmsley, 2005); however with oral administration, the recombinant proteins can be adsorbed into the human system (blood stream) through human metabolism. The plant system does have the limitation of proteolysis (Andersen *et al., 2002)* that can affect the expression of recombinant proteins.

A cost effective expression system is the bacterial system (also referred to as the microbial expression system) (Andersen *et al.,* 2002) as it can result in high protein expression and yield (Terpe, 2006) depending on the protein properties. Bacterial cells transformed with plasmid DNA containing the target gene to be expressed, use its metabolic machinery to favour the overexpression of the non-toxic recombinant protein of interest. However, bacterial systems have limitations depending on the nature of the protein such as the risk of overexpression of the desired protein as an inclusion body, rendering it insoluble and no post translational modification. Nevertheless, the bacterial expression system can be used advantageously to overexpress a high yield of protein and the recombinant insoluble protein may be recovered to sufficient purity when solubilised.

During expression, the target gene is transcribed by the cells to mRNA, followed by mRNA translation to the target protein. It should be noted that the plasmid DNA has different promoter regions that are induced by different inducers to over-express the protein of interest (Andersen *et al.*, 2002). The rhaBAD, nar, araE, lac operon promoters are induced by rhamnose, microacrobic, arabinose (Andersen *et al.,* 2002) and IPTG, respectively. There are different bacterial strains, BL21 (DE3) bacterial strain being the most widely used host cell. Cell growth conditions can also be varied to effectively overexpress the target proteins, hence, it is advisable to optimise and determine the appropriate conditions to overexpress the protein of choice (Ahn *et al.,* 2007). The selection of bacterial strain to be transformed with the plasmid that is composed of the target protein coding DNA, is crucial in protein expression as it affects the yield and solubility state of the overexpressed recombinant protein. The soluble or insoluble (inclusion bodies) overexpressed protein (Andersen *et al.*, 2002) has an effect on the aim and objective of the research question; therefore, precautionary steps are taken to best obtain soluble protein. Soluble proteins are proteins that can be dissolved in solution/buffer and reach homogeneity with the buffer solution, whereas, insoluble protein are proteins that precipitate in solution or buffer (Tsumoto, 2004) after cell lysis and centrifugation.

1.2 Protein solubility

Certain target proteins are toxic to certain bacterial cell lines (Nallamsetty *et al.,* 2005), which results in the overexpression of target protein forming aggregates and compartmentalising into inclusion bodies. Inclusion bodies are transient reservoirs (Carrió and Villaverde, 2002) formed from intracellular aggregated polypeptides, observed during the overexpression of the recombinant target gene (Carrió and Villaverde, 2001) in the bacterial cells cytoplasm. Inclusion body formation and protein aggregation is a major obstacle in obtaining high yield expression of soluble protein (Nallamsetty *et al.,* 2005; Maggioni *et al.,* 2014; Rosano and Ceccarelli, 2014) to further determine and characterise protein structure. There is no universal strategy or protocol used to prevent the formation of inclusion bodies due to differences in amino acid composition and structural characteristics of the target protein, environmental conditions of the host cell and folding mechanism; therefore, recombinant protein overexpression strategies must be optimised for a particular polypeptide (Carrió and Villaverde, 2002). Protein precipitation and inclusion body formation is most commonly observed during the expression of the recombinant target genes (Carrió and Villaverde, 2002), especially in the absence of chaperones during high yield protein overexpression. Chaperones are proteins in bacterial cells that help protein assembly, covalent folding, unfolding and disassembly of other proteins and macromolecules (Mirazimi and Svensson, 2000). Inclusion body formation can be avoided by optimising expression conditions, the use of chaperones (Carrió and Villaverde, 2002) and fusing the target protein with fusion proteins (Young *et al.,* 2012). There is a wide range of chaperones (Carrió and Villaverde, 2001, 2003) and fusion proteins to choose from; however, not all are universal, which is therefore essential to search for one that will be most effective for your protein of interest (Vargas-Cortez *et al.,* 2016).

In most cases, glutathione-S-transferase (~ 26 kDa) (Young *et al.,* 2012) and maltose binding protein (~ 44 kDa) are fusion proteins found to enhance protein overexpression and solubility (Nallamsetty *et al.,* 2005). It should be noted that large size recombinant proteins (usually more than 90 kDa) (Young *et al.,* 2012), can result in reducing the efficiency of bacterial machinery and therefore, low yield protein overexpression. Hence, the reason target proteins above 35 kDa are not generally fused with proteins such as maltose binding protein (Lebendiker, 2014). Myelin basic protein (18.5 kDa) and cation efflux system protein (12.2 kDa) are fusion proteins that showed reduced formation of insoluble proteins as they transport the fused target proteins to the periplasm (Vargas-Cortez *et al., 2016)*. The use of small size fusion proteins that will equal to a final protein size of less than 90 kDa when fused with the target protein and the use of fusion proteins that will transport the target proteins to the periplasm are of the best choice to enhance protein overexpression and solubility.

The periplasm contains fewer proteins and hence less proteolysis takes place, which makes it an advantageous place to over-express recombinant proteins and also an appropriate place to form disulphide bonds (Vargas-Cortez *et al.,* 2016).

1.3 Protein purification

Host cells that overexpress target proteins also produce their own biological proteins; hence the protein of interest needs to be purified. There are a number of purification techniques available such as liquid chromatography, precipitation and membrane separation (Hilbrig and Freitag, 2003). Each purification technique has its advantages and limitations; however, liquid chromatography is the most widely used in protein research. Liquid chromatography has an advantage of being very economical, can be easily operated and undergo easy and effective troubleshooting, however, its limitations include having a large solvent consumption and low binding capacity due to low surface area (Hilbrig and Freitag, 2003) and therefore low throughput (Xu *et al.,* 2011). The membrane separation technique, however, has high economic cost (Hilbrig and Freitag, 2003) due to the need of membrane pore size selection and coating stability that influence separation (Xu *et al.,* 2011).

A wide variety of liquid chromatography purification techniques are available such as size exclusion (Irvine, 1997; Hong *et al.,* 2012), affinity, ion exchange (Xu *et al.,* 2011; Zaveckas *et al.,* 2015), favours the importance of designing the protein purification stages based on the known characteristics of the target protein (Zhao *et al.,* 2007; Zaveckas *et al.,* 2015). The knowledge and understanding of the target proteins' physiochemical properties such as molecular weight, isoelectric point (pI) and hydrophobicity (Zhao *et al.,* 2007) is crucial to select and undergo the type and steps of purification to remove critical protein contaminants (Zhao *et al.,* 2007) and undergo successful purification. Size exclusion chromatography separates proteins based on their molecular weight; whereas ion-exchange chromatography separates protein based on their ionic strength and isoelectric point. Affinity chromatography excludes all proteins apart from the target protein based on the presence of an engineered tag (Zaveckas *et al.,* 2015) which helps identify and select the target protein from the range of proteins expressed in the cell. Another is enzyme-based chromatography based on the substrate or protein and antibody (Hilbrig and Freitag, 2003). Ion-exchange is mostly conducted before size-exclusion purification when the proteins' properties are known (Zhao *et al.,* 2007).

1.3.1 Ion-exchange chromatography

The differing amino acid compositions of proteins results in differing isoelectric points for each protein thus allowing for exploitation of this property to separate by varying ionic strength and pH of buffers. Ion exchange chromatography is a liquid purification technique which separates proteins and other molecules based on their surface charge (Rea *et al.,* 2011), ionic strengths and charge difference (Xu *et al.,* 2011). Based on the above mentioned properties, the proteins can either bind the stationary phase (charged matrix that are either cation or anion exchangers) or repel in which case they would flow through (FT) the column (Rea *et al.,* 2011) during protein mixture/sample application. A cation exchanger (e.g. carboxymethylethyl (CM)) is a negatively charged stationery phase that binds and exchange positively charged ions (cations). An anion exchanger (e.g. diethylaminoethanol (DEAE)) is positively charged and binds and exchanges negatively charged ions (anions). Ion exchange protein purification is performed by separation and elution of proteins using ionic strength (salt concentration) and/or buffers pH. The salt concentration and the buffer pH needs to be optimised for purification (Rea *et al.,* 2011). Buffers salt concentration and/or pH get exchanged to purify the protein of interest. Generating an ionic strength gradient helps to determine the specific or closely related conditions (salt concentration) to purify the protein of interest (Rea *et al., 2011)*.

Optimising the pH of buffers to purify the protein of interest (Zhao *et al.,* 2007), requires the consideration of change in temperature as an increase in temperature decreases buffer viscosity and may increase the mobility of ions as well as increase the number of ions in the buffer affecting pH (Barron *et al.,* 2006), which can affect the elution resolution (Kesler, 1953). Systems such as the polyelectrolyte coacervation (an electrostatically driven liquid-liquid phase separation, resulting from association of oppositely charged macro-ions) uses the proteins' pI to purify proteins (Xu *et al.,* 2011) by eluting the soluble complexes when they reach a neutral charge at the buffers pH close to the proteins pI value (Xu *et al.,* 2011).

1.3.2 Affinity chromatography

Affinity chromatography is a purification technique that utilises an affinity ligand immobilised in a column to bind specific proteins. The bound proteins are then eluted by conditions not supporting the non-covalent bonding interaction of the proteins with the ligand (Hilbrig and Freitag, 2003). Conditions that can affect the proteins-ligand interaction can be pH and ligand binding competing agents (Hilbrig and Freitag, 2003) forcing proteins elution. Affinity tags are used to bind target proteins to an immobilised affinity ligand for purification.

Affinity tags were initially designed from naturally occurring epitopes assumed to be the only recognition site for antibody-protein interaction that result in successful protein purification (Hilbrig and Freitag, 2003), however with time, engineered tags have been optimised and used for different host cell expression conditions and purification purposes (Young *et al.,* 2012). A poly-hexa-histidine (HisTag) is found to be an effective tag for high yield protein purification (Xu *et al.,* 2011) due to its small size and robust nature and its ability to easily bind to the immobilised metal affinity column (IMAC) (Nallamsetty *et al.,* 2005) such as nickel-nitrilotriacetate (Ni^{2+}-NTA), Zinc (Zn^{2+}) (Carter *et al.,* 2006) or cobalt (Co_{2+}) metal (Xu *et al.,* 2011) separating the protein of interest from proteins without HisTags. Other fusion proteins are used as tags such as cation efflux system protein (Vargas-Cortez *et al.,* 2016), Halotag (Young *et al.,* 2012) and maltose binding protein (Nallamsetty *et al.,* 2005). The fusion proteins can also be fused with a HisTag linked protein for effective purification and increasing soluble protein overexpression, solubility and purification when using amylose or nickel affinity chromatography (Waugh, 2005). The tags can then be cleaved from the target protein by using a protease, e.g. Tobacco etch virus (TEV) protease (35 kDa) or Factor Xa after purification. Recombinant viral proteins are examples of proteins that are tagged to purify from bacterial proteins using affinity chromatography.

1.4 Viral proteins

A virus is a microscopic infective agent, which consists of a nucleic acid molecule within a protein coat (López-García, 2012). There are vast viral families and species, which infect different types of cells such as microbial, plant and mammalian cells. Viruses infect cells to replicate their nucleic acid molecule and proteins to form a virion (a complete infectious virus form) (Hulo eta al., 2010). Viruses are obligate parasites (Bhowmick *et al.,* 2014) as they are unable to replicate or sustain themselves without the host organisms' metabolic system (López-García, 2012). Therefore, the virus relies on viral manipulation of the host cells machinery and proteins to replicate (Bhowmick *et al.,* 2014). The viral nucleic acid can be ssDNA, dsDNA, ssRNA, dsRNA to name a few which are covered and protected by the virus capsid/protein coat (Hulo *et al.,* 2010). The DNA viruses use the nucleus as their primary site for replication whereas RNA viruses use the cells cytoplasm, however a viroplasm can be formed (Carreño-Torres *et al.,* 2010) in the cytoplasm by the virus (Cheung *et al.,* 2010).

Reoviridae is a virus family which consists of 15 distinct genera such as aquareovirus, dinovernavirus, fijivirus, orthoreovirus, phytoreovirus and rotavirus (Attoui *et al.,* 2006) to name a few that have double stranded (ds) RNA with 10 to 12 genome segments (Desselberger, 2014; Marthaler *et al.,* 2014). The viral genome segments consist of viral proteins genetic information, which is synthesised

using the host system to form virion protein layers (Pesavento *et al.,* 2006). Reoviridae viruses consist of an icosahedral T=2 (inner capsid) and T=13 (outer capsid) symmetry and are regarded to not have an envelope (in which some can obtain during morphogenesis) (Attoui *et al.,* 2006; Pesavento *et al.,* 2006). The reoviridae proteins include the virus capsid layer, regulatory, none structural (NSP) and/or envelope proteins. The viral capsid and glycoprotein play an important role in viral structure, viral DNA protection from host cells protease and viral penetration inside the host cells for replication. However, the viral multi-layered proteins can be considered crucial in virus classifications and vaccine development using neutralising antibodies (Dennehy, 2008). With such crucial information being obtained from viral proteins, the proteins are therefore overexpressed, isolated and studied for use by the science community to benefit the world community such as understanding corona virus to combat the pandemic(Belouzard *et al.,* 2012; Pereira de Albuquerque *et al.,* 2020)). There are ~5000 identified viral proteins which their genetic information are recorded and saved at Protein Data Bank (PDB) (Worldwide PDB protein data bank) (www.rcsb.org). This helps scientists worldwide to identify and link the genetic code or amino acids sequences for every protein sequence verification, or information of the target protein.

1.5 Rotavirus

Rotavirus is a virus from the reoviridae family (Desselberger, 2014) which infects both humans and animals (McDonald and Patton, 2008) and therefore is classified as a zoonotic infectious agent. The name rotavirus comes from the virion structure (rota, means “wheel” in Latin). Rotavirus infection occurs from being in contact with stool of an animal or human individual infected with rotavirus, lack of clean hygiene and poor sanitation (Chia *et al.,* 2018). Rotavirus infects the intestinal enterocytes (Hu *et al.,* 2012) of the jejunum in the small intestine, an environment rich in proteases (Pesavento *et al.,* 2006) affecting disaccharide metabolism and hindering water and sugar absorption (Kobayashi *et al.,* 2007). During infection, the virus arrests the cell’s cycle at growth phase and utilises the energy and machinery to replicate (Bhowmick *et al.,* 2014). The ion transport protein sodium-hydrogen exchanger 3 (NHE3) activity is inhibited, resulting in decreased sodium ion absorption causing osmotic gradient (Ramani *et al.,* 2016) and cell death resulting in dehydrating diarrhoea and malnutrition (Pesavento *et al.,* 2006) in children ≤ 5 years old (Pesavento *et al.,* 2006; Desselberger, 2014), resulting in over ~450 000 child mortalities annually (Dormitzer *et al.*, 2001; Aoki *et al.,* 2011). During the cooler and drier months in countries affected by rotavirus infection, there is an increase in hospital visits (~24 million) and hospitalisation of children (~2.4 million) (Dennehy, 2008) that may lead to compromised health services in developing countries with limited healthcare resources.

Rotavirus infection symptoms include watery diarrhoea, vomiting, fever and dehydration (Ramani *et al.,* 2016).

Rotavirus has been classified into seven (Hoshino and Kapikian, 2000) and possibly eight groups (groups A-H) also called species (Desselberger, 2014) using the VP6 expressed by the virus (McDonald and Patton 2008; Trask *et al.,* 2013) and VP6 gene sequencing (Marthaler *et al.,* 2014) as qualifiers/distinguishers. Human infection is caused by strains A-C (Hoshino and Kapikian, 2000) with A being the predominant strain resulting in child mortality (Parashar *et al.,* 2006) and hence strain A is the most investigated and linked to vaccine development (Hoshino and Kapikian, 2000). Most monoclonal antibodies are directed to VP4 and VP7 because these proteins are located on the outer layer of rotavirus and carry epitopes that evoke and bind neutralising antibodies (Wang *et al.,* 1996; Hoshino and Kapikian, 2000; Dormitzer *et al.,* 2001). The rotavirus neutralisation specificity (serotype) is then used to classify and determine rotaviruses into two serotype groups (Hoshino and Kapikian, 2000; Desselberger, 2014).

VP4 is subjected to proteolysis and VP7 is a glycoprotein (Pesavento *et al.,* 2006) which is used in classification of the virus into two serotype groups (serotype P and serotype G respectively) (Hoshino and Kapikian, 2000). Serotype P characterisation is difficult using the virus neutralisation method; therefore a genotype sequence analysis method is used to determine and characterise the P serotype (Dennehy, 2008).

The dual-nomenclature system which uses VP4 and VP7 is used to classify rotavirus into G and P genotype (Hoshino and Kapikian, 2000; Ianiro *et al.,* 2016; Ramani *et al.,* 2016), however the P-type uses a double nomenclature which uses the serotype (such as P1A) and genotype (such as P[8]) to give a P1A[8] classification (Desselberger, 2014). There are 14 rotavirus G serotypes and 11 rotavirus P serotypes, with 11 of G serotype and 7 of P serotypes detected in humans by neutralising assay with neutralising antibodies (hyperminnue artisera) (Hoshino and Kapikian, 2000), from the 11 G serotype, four G-types (G1-4) identified to be the strains that predominantly infect humans (Kobayashi *et al.,* 2007).

1.6 Rotavirus replication cycle

During human and animal infection by rotavirus, the virus undergoes specific steps to replicate its genome to multiply and survive. The outer layer (OL) protein, VP4, plays an essential role in viral infection by adhering to the host epithelial cell receptors initiating viral entry by endocytosis (Zhou *et al.,* 1994; Dormitzer *et al.,* 2001; Hu *et al.,* 2012). Prior to viral adhesion to the cell, proteolysis

splits VP4 into VP5 and VP8 which remain non-covalently attached during infection mediated by VP8 attachment to the cells glycans (Ramani *et al.,* 2016). VP5 a lipophilic domain, gets exposed and forms post penetration umbrella information which is also favoured by trypsin to convey full infectivity (Desselberger, 2014). The cytoplasmic membrane of the host forms an endosome around the virus during cell entry, which is degraded as VP4 and VP7 dissociate (Trask *et al.,* 2013), releasing the double layered particle (DLP) virus into the cells cytoplasm (Trask *et al.,* 2013). Within the cytoplasm, the viral genome begins replicating and producing viral particles (Figure 1.1). As the viral particles assemble, VP6 interacts with non-structural trans-membrane glycoprotein (NSP4) to form DLPs in the viroplasm. The DLPs then move to the endoplasmic reticulum where they acquire a third layer of VP7 and VP4 forming a triple layered particle (TLP) (Schuck *et al.,* 2001).

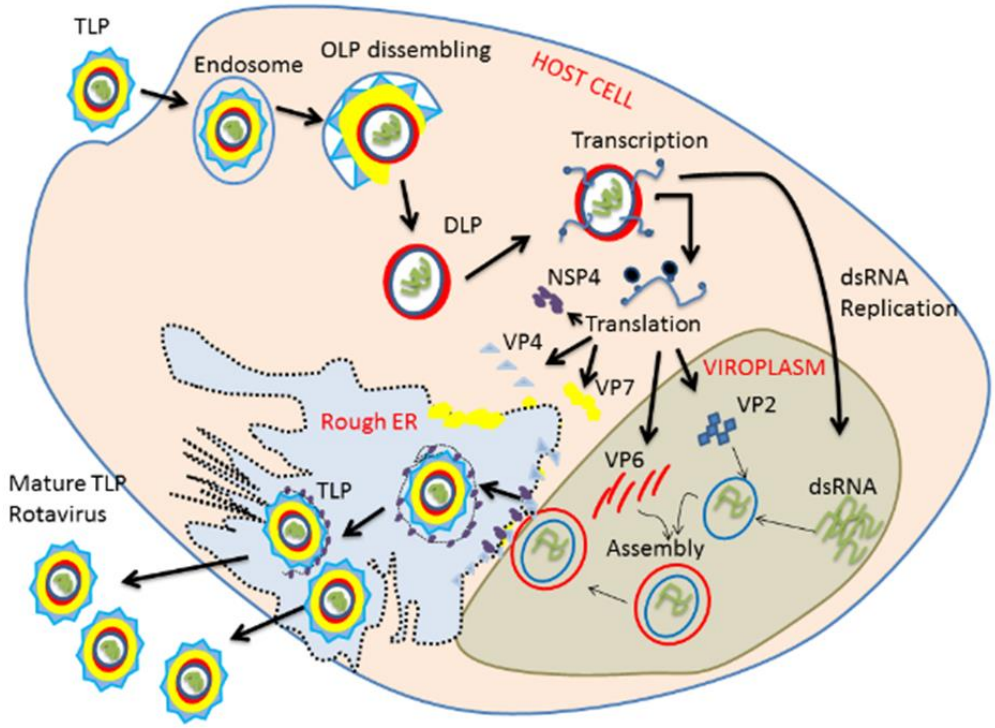


Figure 1.1: Rotavirus replication cycle
Rotavirus replication cycle illustration indicating the epithelial cells with formed endosome (TLP covered blue), the double layered particle (DLP) into the host cell cytoplasm. The DLP begins producing viral particles (VP7 - yellow, VP4 - blue, VP6 - red, VP2 - navy blue etc.) and replicating its genome (dsRNA in green) to assemble and form DLPs in a viroplasm. The assembled virus DLPs move to the endoplasmic reticulum where they acquire their outer layer forming triple layered particles (TLP) which are released to infect adjacent epithelial cells. Author-generated image using MS PowerPoint (2013).

During the formation of the outer protein layer, the correctly folded VP7 and NSP4 interact with raft proteins. In the integral endoplasmic reticulum membrane (Trask *et al.,* 2013), the folded VP7 displaces a lipid layer to form a final functional TLP released to infect other cells (Figure 1.1).

1.7 Roles of rotavirus proteins

Rotavirus consists of three layers in which the inner layer (IL) is made up of VP2, the middle layer (ML) made of VP6 and the outer layer (OL) made up of VP7 and VP4 (Figure 1.2) (Zhou *et al.,* 1994). The viral proteins are translated from the virus 11-segmented double stranded RNA genome (dsRNA) (Hsu *et al.,* 1997; Angel *et al.,* 2007; Amimo *et al.,* 2014) located in the core of the virus (Figure 1.2). Each genome segment is translated to produce a specific protein, i.e. VP2 forming the IL which encloses the viral genome and VP1 & 3 (Hoshino and Kapikian, 2000) are translated from the 2nd genome segment. The spike protein (VP4) which adheres to the host cells receptors during infection is translated from the 4th segment. VP6 forming the ML is transcribed from the 6th segment, and VP7 which forms part of the OL is translated from either the 7th, 8th or/and 9th segment depending on the viral strain (Day *et al.,* 2002).

The non-structural proteins (NSP1-6) are also transcribed from the 11 segmented genome, however some rotavirus strains do not encode NSP6 (Hu *et al.,* 2013). The NSP's are responsible for rotavirus genome replication, however, the role of NSP6 remains unclear (Hu *et al.,* 2012), and NSP1 is involved in a counteractive activity against host antiviral response by interacting with IRF3. However, in some animal models it has a pathogenesis role, such as antagonising the type I interferon response for viral pathogen increase (Pesavento *et al.,* 2006). NSP2, transcribed from the 8th segment of the genome plays a role in viroplasm formation and is found to interact with VP2 (Hu *et al.,* 2013). NSP3 is transcribed from the 7th segment and suppresses host protein synthesis and facilitates virus mRNA translation. NSP4 is transcribed from the 10th segment and plays a role in the formation of the DLP and hence is synthesised as an endoplasmic reticulum (ER) transmembrane glycoprotein (Hu *et al.,* 2012). NSP5 transcribed from the 11th segment plays a role in the formation and regulation of a viroplasm in partnership with NSP2 (Martin *et al.,* 2011). Therefore, if NSP2 and NSP5 are blocked, viral replication won't occur as the viroplasm is required for double layered particle formation (Desselberger, 2014).

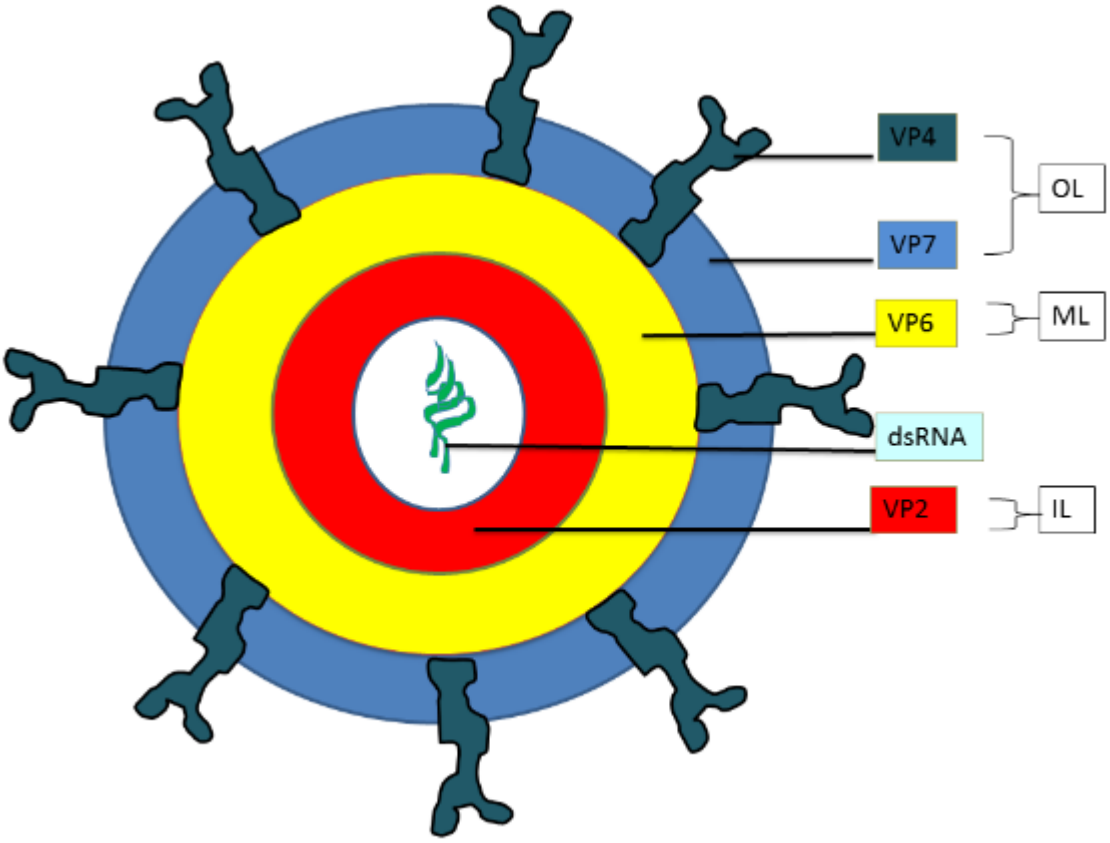


Figure 1.2: Spherical rotavirus triple layer particle (TLP)
Rotavirus is a spherical triple layer particle with its outer layer (OL) consisting of spike VP4 (navy blue) and VP7 (blue), the middle layer (ML) VP6 (yellow) and an inner layer (IL) VP2 (red). All enclosing a double stranded RNA (dsRNA) genome in green. Author-generated image using MS PowerPoint (2013).

1.7.1 Rotavirus VP7

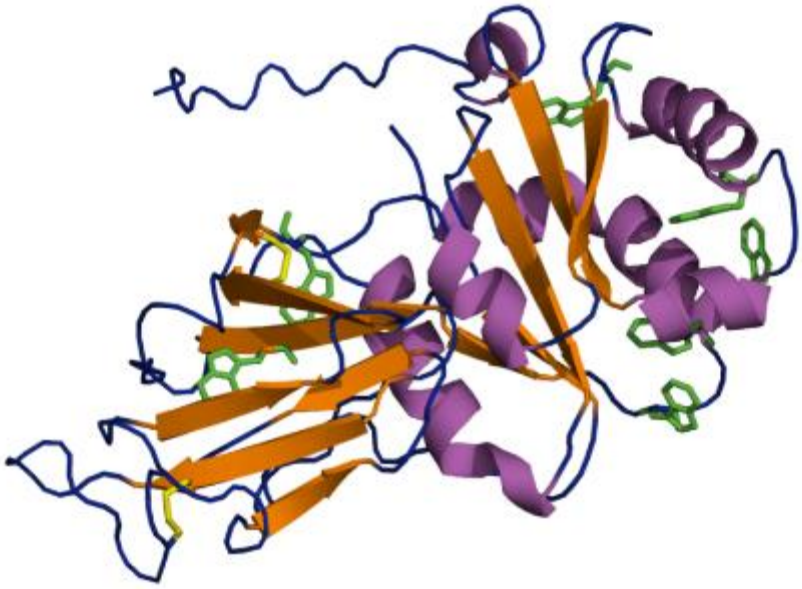
The rotavirus VP7 subunit is a 37 kDa size protein made up of 326 amino acids (Kobayashi *et al.,* 2007). The VP7 monomeric subunit structure consists of 12 beta-strands and 6 alpha-helices (Figure 1.3a). Three VP7 subunits form a trimer stabilized by six calcium ions (Figure 1.3b) (Trask and Dormitzer, 2006) and interacts with other VP7 trimers to form an icosahedral shape. The rotavirus icosahedral OL consists of 780 VP7 monomers which results in 260 VP7 trimers, forming a smooth shell, locking the second outer protein, VP4, in place (Trask et al. 2010).

In previous studies, VP7 antigenicity and gene sequence were used to identify and compare rotavirus strains discovered in different areas (Khetawat *et al.,* 2001). VP7 antigenicity is applied in rotavirus vaccination, in which a vaccine (RotaTeq), a live/attenuated oral rotavirus from the most prevalent G1-4 types (Kobayashi *et al.,* 2007) is administered to children to elicit an immune response against rotavirus. VP7 vaccination studies such as mice immunization from the consumption of VP7 expressing potatoes (Wu *et al.,* 2003) were conducted to test and induce immunogenicity against rotavirus (Tissera *et al.,* 2017). Plants engineered to express VP7 can be ingested to simulate the mucosal system to secrete IgA neutralising antibody activity instead of IgG against rotavirus (Wu *et al.,* 2003).

1.7.2 VP7 expression and purification

In protein research, different types of recombinant protein expression and purification challenges are faced such as the type of expression systems, purification techniques and identifying conditions that favour the target proteins biological structure and activity. In previous studies, for instance, VP7 has been expressed in insect (Trask and Dormitzer, 2006), mammalian cells (Ruiz *et al.,* 2000), plant (Wu *et al.,* 2003) and recently in bacterial cells (Deng *et al.,* 2015; Wu *et al.,* 2015), to study its structure and antigenicity. However, its expression in insect and mammalian cells results in high economical cost to overexpress the protein, and that overexpressed in bacterial cells results in insoluble proteins. The above mentioned expression methods besides the bacterial system require high economical cost to express target proteins. Therefore, a system which results in an economic isolation of VP7 protein will be of great contribution to rotavirus protein research. Previous studies of VP7 purification conducted after expression resulted in low purity (less than 30%) from affinity purification and higher purity (over 90 %) using size exclusion chromatography (Mellado *et al.,* 2008) and monoclonal affinity purification (Dormitzer *et al.,* 2001) indicating VP7 is likely to be purified using different chromatographic techniques. To obtain more information such as the structural characteristics of the purified proteins, additional characterization techniques are conducted.

a



b

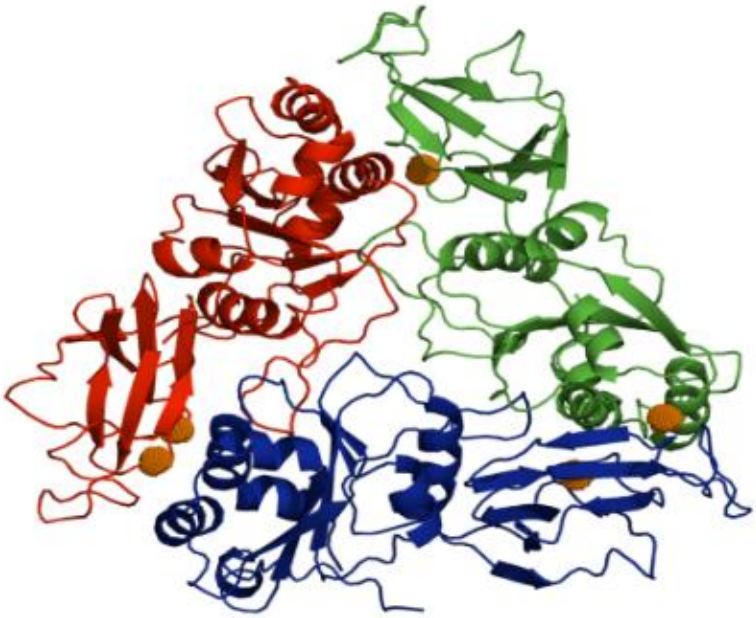


Figure 1.3: Ribbon representation of VP7 monomer and trimer structure
(a) Monomeric unit of VP7 (37 kDa) containing 12 beta strands (orange), 6 alpha helices (magenta) and 7 tryptophan residues (green). The monomer contains 2 disulphide bonds (yellow) which helps in protein folding to form VP7 native structure. (b) VP7 trimeric structure (subunits denoted in blue, red and green). The calcium ion (orange spheres) located at the interface between each monomer unit which stabilises the trimer. Image generated using PyMol software using the engineered VP7 sequence (PDB Code: 3FMG (Aoki *et al.,* 2009)).

1.8 Optical characterisation of proteins

1.8.1 Absorbance spectroscopy

Absorbance spectroscopy is a technique used in chemistry to determine the presence of a particular substance in a sample (Liu, 2009). The technique measures the absorbance of light photons by proteins or chromophores in solution (Liu, 2009) at specific wavelength. Aromatic side chains of amino acids that make up proteins absorb light at 280 nm and the absorbance is used to determine protein concentration at this wavelength by using the Beer-Lamberts law (Lakowicz, 2006). The absorbance reading is recorded and its concentration is determined by using equation 1:

$$A = \varepsilon_{280} c \ell \quad (1)$$

Where A is the absorbance at 280 nm wavelength, Ɛ represents the protein's molar extinction coefficient at 280 nm, c is the protein concentration in M and ℓ is the path length in centimetres (cm).

1.8.2 Intrinsic tryptophan fluorescence spectroscopy

Fluorescence spectra are used to determine proteins tertiary conformational transitions (protein folding and unfolding) by monitoring the fluorescence emission of the protein's chromophores (aromatic amino acids but predominantly contributed by tryptophan) when excited by UV light at a wavelength of 280 nm. However tryptophan can be selectively excited at 295 nm (Vivian & Callis, 2001; Chenal *et al.,* 2005; Tang *et al.,* 2009) when located in the proteins interior site or in the absence of tyrosine residues which emit light that excite tryptophan residues when excited at 280 nm. When proteins unfold/denature, a change in emission fluorescence is observed (Tang *et al.,* 2009).

The proteins' fluorescence emission is typically measured between the UV excitation wavelength (280 nm) to 500 nm (Lakowicz, 2006; Tang *et al.,* 2009). Tryptophan has a maximum emission wavelength range between 308-350 nm (Vivian & Callis, 2001) depending on the nature of its local environment. The fluorescence emission maximum between 335 and 345 nm is indicative of tryptophan residues in a hydrophobic environment but as the proteins unfold, the buried (hydrophobic) tryptophan residues are exposed to solvent and a bathochromic shift (red-shift) occurs (Figure 1.4) to wavelength maxima of 350 nm wavelength.

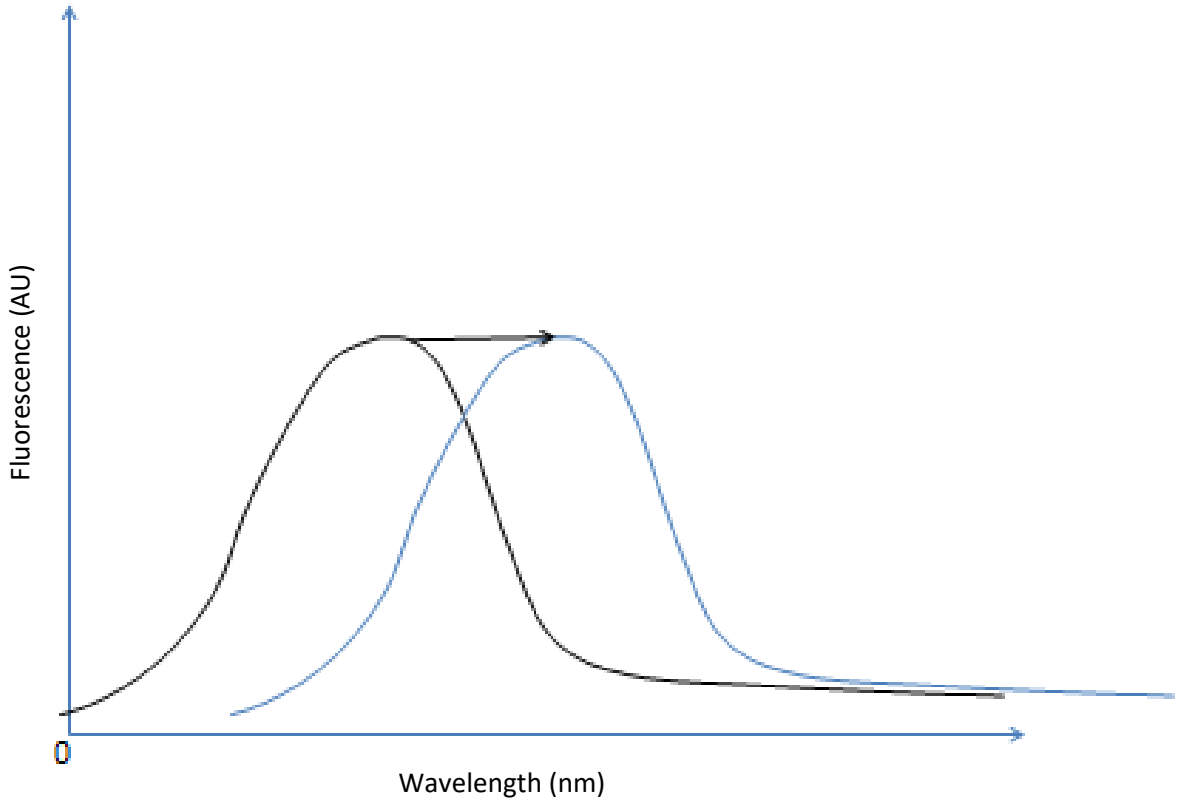


Figure 1.4: Bathochromic shift
Fluorescence spectroscopy of buried tryptophan being exposed to the solvent resulting in a bathochromic shift. Author-generated image using MS PowerPoint (2013).

The bathochromic shift is caused by chromophores absorption of light and the emission of low energy light, which another chromophore and solvent absorbs.

The exposure of the buried (hydrophobic) tryptophan residues into the aqueous solvent (to be hydrophilic) due to the protein complete unfolding, causes a bathochromic shift (Figure 1.4) (Tang *et al., 2009; Zhang et al. 2013)*. Tryptophan consists of an indole ring (benzene fused to a pyrrole ring). When the buried tryptophan is exposed to the solvent, an electron shift from the pyrrole to the benzene ring occurs causing a negative charge at and near pyrrole, a positive charge near the benzene and high energy release which then result in the fluorescence wavelength maxima shift to 350 nm (Vivian and Callis, 2001). Vivian and Callis (2001) study on azurin (with buried tryptophan residues) and glucagon (with solvent exposed tryptophan residues) showed evidence of a buried tryptophan being exposed to the solvent resulting in a bathochromic shift from 320 to 350 nm wavelength maxima.

1.8.3 Far UV circular dichroism spectroscopy

Circular dichroism spectroscopy is a form of light absorption spectroscopy used to measure the right and left circularly polarised light absorbance difference (Chenal *et al.,* 2005) Circular dichroism helps identify the proteins secondary structures (Park *et al.,* 1992; Kelly *et al.,* 2005). The proteins secondary structure is defined by the number and conformation of alpha helices, beta sheets/strands and random coils. Proteins with 100 % alpha helices have a characteristic far-UV CD spectrum where there is a trough at wavelengths 208 nm and at 222 nm (Chenal *et al.,* 2005), and a peak at 190 nm (Park *et al.,* 1992; Kelly *et al.,* 2005) (Figure 1.5).

A protein with 100 % beta strands has a characteristic CD spectrum with a peak at 198 nm and a single trough at 218 nm (Kelly et al. 2005) (Figure 1.5). A protein with a random coil structure shows a CD spectrum with a single trough at 195 nm and a single peak at 212 nm (Figure 1.5). Therefore, based on the characteristics observed in the far UV CD spectrum of given protein, the predominant secondary structure can be estimated.

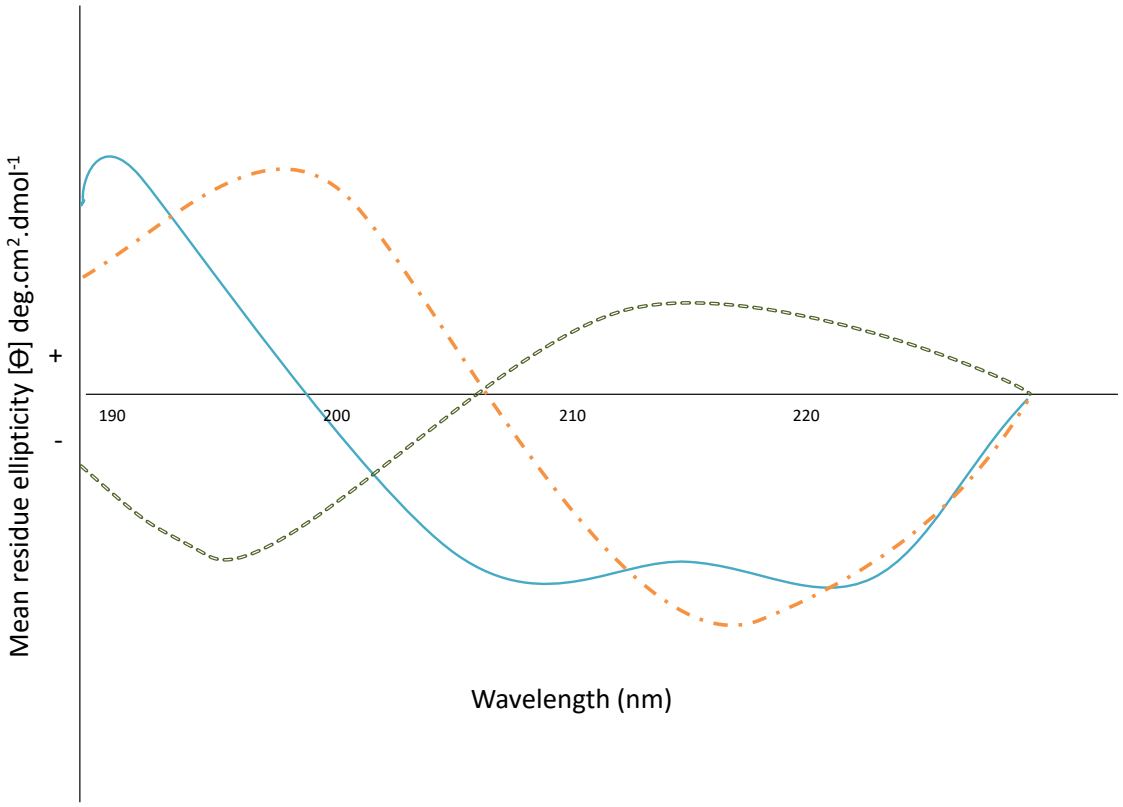


Figure 1.5: Standard Far UV spectra
Secondary structure characteristic of proteins with 100% alpha helices (blue), 100% beta sheets (orange) and 100% random coil (green). Author-generated image using MS PowerPoint (2013).

1.9 Aims and objectives

Rotavirus VP7 has been expressed over the years (Dormitzer *et al.,* 2001; Deng *et al.,* 2015; Wu *et al*., 2015), using different expression systems for viral classification, structural identification and analysis and vaccination research studies. To our knowledge, Deng *et al.* (2015) conducted transformation and over-expression of full-length rotavirus VP7 in TOP 10 bacterial cells and purified using GST affinity tag. Wu *et al*. (2015) overexpressed insoluble bluetongue VP7 in BL23 (DE3) cells and VP7 structure characterisation analysis in the presence of calcium was performed by Dormitzer *et al.* (2001).

This study aimed to obtain purified rotavirus VP7 protein from a bacterial expression system (using BL21 (DE3) and C43 (DE3) pLysS bacterial cells) and characterise the protein using spectroscopic techniques.

The objectives:-

1. To overexpress Rotavirus VP7 protein in bacterial cells, assessing cell lines, inducer concentration, induction temperature and time.
2. Solubilise insoluble over-expressed VP7 protein
3. Purify the VP7 protein using liquid chromatography techniques (ion exchange and affinity chromatography or affinity chromatography alone).
4. Characterise the purified VP7 in terms of secondary and tertiary structure, as well as thermal denaturation using spectroscopic techniques.

2. Materials and methods

2.1 Materials

Luria bertani (LB) agar (LENNOX) and broth media (Miller's LB broth) were purchased (Conda laboratories, Torrejon de Ardoz Madrid, Spain), LB broth, urea crystals, sodium acetate and imidazole were purchased (Merck of EMD Millipore corporation, Massachusetts, USA), Glycine, sodium chloride, sodium phosphate dibasic, Trizma, β-mercaptoethanol, chloramphenicol, Dithiothreitol (DTT) and N,N,N',N'-tetramethylethylene-1,2 diamide (TEMED) were purchased (Sigma, St. Louis, USA). Sodium azide and acrylamide 40% solution of Electran for electrophoresis were purchased (VWR International Prolabo Chemicals, Leicester, England). Isopropylthio-β-galactoside (IPTG), Kanamycin monosulphate were purchased (Melford biolaboratories, Chelsworth, UK). BL21 (DE3) cells and C43 (DE3) pLysS cells were purchased (Bio Labs Incorporation and Lucigen, Middleton, USA). Protein molecular weight markers and 34 mm dry flat width (3.7 m/cm) snakeskin dialysis tubing 10 kDa were purchased (Thermo Scientific, Massachusetts, USA). DEAE-Sepharose resin, columns and GE Healthcare 5 ml HisTrap columns were purchased (GE Healthcare Bio-sciences AB, Danderyd Sweden). The plasmid pET28a + was engineered with the VP7 consensus sequence in Appendix A at BamHI and NdeI restriction site and shipped with confirmation of sequencing insert in plasmid by GenScript (New Jersey, USA). The engineered sequence provided by GenScript was verified on smartblast tool in blast web software (blast.ncbi.nlm.nih.gov), which confirmed 98% similarity with human rotavirus A outer capsid glycoprotein VP7.

2.2 Preparation of competent cells

LB agar and LB media were individually prepared as per the manufactures© description (Conda laboratories, Torrejon de Ardoz Madrid, Spain), autoclaved, allowed to cool down and poured in sterile plates and Erlenmeyer flask, respectively .

C43 (DE3) pLysS and BL21 (DE3) cells were streaked on prepared LB agar and grown over-night at 37 °C in an incubator. A single colony was selected from the LB agar grown cells and placed in 10 ml LB media for an overnight growth with shaking at 225 rpm at 37 °C. 1 ml of the overnight cells growth was transferred into 20 ml of LB media and grown for 2 ½ hours at 37 °C with shaking at 225 rpm. The grown cells were then transferred into a 50 ml micro tube and placed on ice for 20 minutes. The cells were then centrifuged at 4 °C at 3000 × *g* for 10 minutes. The media was discarded and the cells were re-suspended in 30 ml of cold 0.1 M $CaCl_2$ and incubated on ice for 30 minutes. The cells were

then centrifuged for 10 minutes at 3000 × *g* in 4 °C temperature. The supernatant was discarded and the cells were re-suspended by pipetting in 8 ml of 0.1 M $CaCl_2$ with 15% glycerol. Cells were stored in 100 µl aliquots in 1.5 ml micro tubes and stored at -80 °C.

2.3 Cell transformation

LB agar was prepared as indicated by company description, autoclaved, allowed to cool down at room temperature, 100 µg/µl antibiotics (kanamycin) was added and then agar was poured in 10 ml sterile plates for cells transformed wth pET28a+VP7 plasmid. Competent cells were thawed on ice for 30 minutes, 2 µl of pET28a+VP7 DNA plasmid (94 µg/µl) was added into 50 µl of BL21 (DE3) cells, mixed gently in 1.5 ml micro tubes and placed on ice for 30 minutes. The cells were then heat shocked for 60 seconds in a 42 °C heating block and placed back on ice for 2 minutes incubation. 300 µl SOC media was added into the cells micro tubes and incubated in a 37 °C with shaking (225 rpm) for 45 minutes. The cells were then plated on LB agar with kanamycin antibiotic (100 µg/µl) (since pET28a+VP7 plasmid has a kanamycin resistance gene) and incubated at 37 °C incubator to grow overnight from which a single colony was selected. The selected colony was grown in 10 ml LB media overnight in a shaking incubator at 225 rpm at 37 °C. 1 ml of the overnight culture was transferred into 20 ml of fresh LB media containing 100 µg/µl kanamycin and grown for 2 ½ hours at 37 °C in a shaking incubator at 225 rpm. Then 100 µl aliquots of culture with 100 µl 50% glycerol were placed in 1.5 ml micro tubes and stored at -80 °C. C43 (DE3) pLysS cells were also transformed by the same procedure, however, the C43 (DE3) pLysS cells were grown in LB media and agar contains 100 µg/µl kanamycin and 30 µg/µl chloramphenicol antibiotics (C43 (DE3) pLysS contains a pLysS plasmid with chloramphenicol resistant gene and the pET28a+VP7 plasmid contains a kanamycin resistant gene).

2.4 Protein (VP7) overexpression and solubilisation

2.4.1 IPTG induction

The transformed cells with pET28a+ plasmid engineered with rotavirus VP7 gene sequence (pET28a+VP7 plasmid) were grown in 10 ml LB media supplemented with antibiotics (BL21 (DE3) cells in 100 µg/µl kanamycin n LB media; C43 (DE3) pLysS cells in 100 µg/µl kanamycin and 30 µg/µl chloramphenicol in LB media) overnight in a shacking incubator (SI-600 lab companion, Minnesota, USA) set at 225 rpm at 37 °C. A dilution of 1:150 of overnight cell growth to fresh LB media supplemented with antibiotics was conducted for fresh cell growth. The cells optical density at 600 nm ($O.D_{600}$) was measured and at the $O.D_{600}$ 600 of ~0.6 absorbance units, VP7 overexpression was induced with different IPTG (0.2; 0.5 and 1 mM) concentrations. Following induction, cells were

grown for 18 to 20 hours and harvested by centrifugation at 5000 × *g* at 4° C (Beckman Coulter, California, USA) for 20 minutes before resuspension in 20 mM sodium acetate buffer pH 5.7, containing 0.02 % sodium azide. The resuspended cells were lysed by sonication (amplitude of 60 for 30 seconds, in 3 cycles on ice with 1 second interval resting time) using a Q-sonica sonicator (Q-sonica sonicators, New York, USA). The lysate was centrifuged at 16,000 x *g* for 20 minutes, separating the soluble fraction in the supernatant (S) from the insoluble fraction in the pellet (P).

2.4.2 Induction length

The transformed BL21 (DE3) and C43 (DE3) pLysS cells (section 2.3) starter culture growth were conducted overnight in a shaking incubator at 225 rpm (SI-600 lab companion, USA) at 37 °C. A 1:150 dilution of overnight culture into fresh LB media supplemented with antibiotics (100 µg/µl kanamycin for BL21 DE3) cells and 100 µg/µl kanamycin + 30 µg/µl chloramphenicol for C43 (DE3) pLysS cells) was conducted. Cells at O.D 600 of ~ 0.6 absorbance units were induced with optimal (0.5 mM) IPTG concentration to overexpress VP7. Following induction, cells were grown for 4 hours and 20 hours and harvested by centrifugation at 5000 × *g* for 20 minutes at 4° C (Beckman with JA10 rotor, USA). The harvested cells were resuspended in 20 mM sodium acetate buffer pH 5.7, containing 0.02 % (w/v) sodium azide. The resuspended cells were lysed by sonication (as in section 2.4.1) using a Q-Sonica sonicator with a small stainless steel probe. The lysate was centrifuged at 16,000 x *g* for 20 minutes, separating the supernatant (S) from the pellet (P).

2.4.3 Induction temperature variation

Transformation as in section 2.3 was conducted, following cells induction for 20 hours VP7 overexpression at 20 and 37 °C temperatures shaking 225 rpm incubators. The induced cells were harvested by centrifugation at 5000 × *g* at 4° C temperature for 20 minutes (Beckman with JA10 rotor, USA) and resuspended in 20 mM sodium acetate buffer pH 5.7, containing 0.02 % (w/v) sodium azide. The resuspended cells were lysed by sonication (as above) using a Q-sonica sonicator with a small stainless steel probe. The lysate was centrifuged at 16,000 x *g* for 20 minutes, separating the soluble fraction (supernatant) from the insoluble fraction (pellet).

2.4.4 Solubilisation

Bacterial cells containing the overexpressed VP7 following 20 hour growth were harvested by centrifugation for 20 minutes at 5000 × *g* speed at 4° C (Beckman with JA10 rotor, USA) before being resuspended in 20 mM sodium acetate buffer pH 5.7, containing 0.02 % (w/v) sodium azide. The resuspended cells were lysed by sonication (as above in section 2.4.1) using a Q-Sonica sonicators with a small stainless steel probe. The lysate was centrifuged at 16 000 × *g* for 20 minutes, separating the soluble fraction (S) from the insoluble pellet (P). The pellet was resuspended in 20 mM sodium acetate buffer pH 5.7, containing 0.02 % (w/v) sodium azide and lysed two more times by sonication. The lysate was centrifuged at 16,000 x *g* for 20 minutes, separating the soluble fraction from the pellet.

The pellet was resuspended in 100 mM sodium phosphate buffer pH 5.7 with 10 mM Tris and 6 M urea, disrupted by sonication (as above, section 2.4.1) and centrifuged at 16,000 x *g* for 20 minutes to separate the solubilised fraction (Sol S) from the insoluble fraction (Sol P). The method to solubilise VP7 from inclusion bodies was adapted from the QIAGEN purification handbook (The QIAexpressionist, 2003).

2.5 Ion-exchange chromatography techniques

2.5.1 Ion exchange chromatography (salt gradient elution trials)

The pET28a+VP7 plasmid transformed C43 (DE3) pLysS cells were grown, induced with 0.5 mM IPTG for VP7 overexpression at 37°C for 20 hours in a shaking incubator, harvested as above (section 2.4.1), resuspended in 20 mM sodium acetate buffer pH 5.7 and lysed by sonication (as above in section 2.4.1). The lysate was centrifuged at 16,000 x *g* for 20 minutes and separated the supernatant from the pellet. The supernatant was loaded onto a 20 mM sodium acetate buffer pH 5.7 pre-equilibrated DEAE-Sepharose column (GE Healthcare Bio-sciences AB, Sweden) connected to an AKTA Prime plus purification system with pH electrode (GE Healthcare Bio-sciences AB, Sweden) and the flow through (FT) with increasing absorbance was collected. The column was washed with 3 column volumes of 20 mM sodium acetate buffer pH 5.7, the bound proteins elusion was performed by an increasing salt gradient 0.7 M sodium chloride and 4ml eluent fractions were collected. The column was regenerated (R) with 2 M sodium chloride and eluent were collected.

The supernatant obtained after C43 (DE3) pLysS cells growth, inductions, harvesting and sonication as above was loaded onto a 20 mM sodium acetate buffer pH 5.7 pre-equilibrated DEAE-Sepharose

column (GE Healthcare Bio-sciences AB, Sweden) connected to a Bio-Rad NGC purification system (California, USA). The column was washed with 3 column volumes of 20 mM sodium acetate buffer pH 5.7. The bound proteins were eluted by an increasing salt gradient to 0.5 M sodium chloride and column regenerated. The eluted fractions with increasing absorbance were concentrated.

The C43 (DE3) pLysS supernatant with overexpressed VP7 was applied to a DEAE-Sepharose column pre-equilibrated with 20 mM sodium acetate buffer pH 5.7 (GE Healthcare Bio-sciences AB, Sweden) connected to a Bio-Rad NGC purification system and the flow through was collected. The column was washed with 3 column volumes of 20 mM sodium acetate buffer pH 5.7 and the eluted samples (CW) with minimum 20 nm wavelength absorbance were collected. The bound proteins detected by increasing absorbance were eluted and collected in fractions by addition of 20 mM sodium acetate buffer pH 5.7 with 0.1 M sodium chloride for 3 column volumes and then 20 mM sodium acetate buffer pH 5.7 with 0.2 M sodium chloride for 3 column volumes and finally 20 mM sodium acetate buffer pH 5.7 with 0.5 M sodium chloride and the column regenerated. Fractions were collected for each sodium chloride concentration and analysed by SDS PAGE analysis.

2.5.2 Ion exchange chromatography (pH buffer gradient)

The transformed C43 (DE3) pLysS cells were grown, induced with 0.5 mM IPTG, harvested, resuspended in 20 mM sodium acetate buffer pH 5.7 and lysed by sonication. The lysate was centrifuged at 16 000 × *g* for 20 minutes to separate the supernatant from the pellet. The supernatant was applied to a 20 mM sodium acetate buffer pH 5.7 pre-equilibrated DEAE-Sepharose column connected to either an AKTA Prime plus system with pH electrode (GE Healthcare Bio-sciences AB, Sweden), followed by a column wash with 6 column volumes of 20 mM sodium acetate buffer pH 5.7 until the pH detected by the AKTA prime plus system pH meter stabilised. The column was stored at room temperature overnight and washed the next morning with 20 mM sodium acetate buffer pH 5.7 to reach stable pH 5.7 as recorded by the AKTA prime plus system pH meter electrode (GE Healthcare Bio-sciences AB, Sweden). Proteins still bound to the column were then eluted by a decreasing pH gradient to 20 mM sodium acetate buffer pH 4.0 whilst collecting 4 ml sample fractions.

2.5.3 Ion exchange chromatography (pH buffer exchange)

The pH buffer exchange was conducted to observe if VP7 protein was eluted by pH buffer close to its pI value (pI 4.6). The C43 (DE3) pLysS cells transformed with pET28a+VP7 plasmid were grown, induced with 0.5 mM IPTG concentrations at 37°C for 20 hours, harvested and lysed by sonication as

above (section 2.4.1). The supernatant was applied to a DEAE column pre-equilibrated with 20 mM sodium acetate buffer pH 5.7 connected to a Bio-Rad purification system (California, USA), and the bound proteins were eluted with decreasing pH exchange (5.7 to 5.0 to 4.7 to 4.6, 4.5 and finally by 4.0).

The supernatant was applied to a 20 mM sodium acetate buffer pH 5.7 pre-equilibrated DEAE column connected to an AKTA Prime plus system (GE Healthcare Bio-sciences AB, Sweden). The column was washed with 6 column volumes of 20 mM sodium acetate buffer pH 5.7. The bound proteins were eluted by a buffer pH exchange down to 20 mM sodium acetate buffer (pH 5.7 to 5.0 to 4.0) until specific buffer pH reading was detected collecting 4ml sample fractions with increasing absorbance at 280 nm wavelength. The fractions samples eluted by pH 5.7 to 5.0 to 4.0 were concentrated by ultrafiltration for analysis. Ultrafiltration is a membrane system which allows the buffer or solution through membrane filter. The buffer passes through the membrane filter, however excluding proteins or components in the buffer form passing through the membrane resulting in high protein concentration.

2.6 Affinity (HisTag) chromatography

Fractions containing VP7 of ion exchange purification (section 2.5.3) were concentrated and applied into a 20 mM sodium-acetate buffer pH 5.7 with 20 mM imidazole, 20 mM sodium chloride and 1 mM DTT pre-equilibrated GE Healthcare 5 ml HisTrap column, connected to a Bio-Rad NGC purification system. The column was washed with 5 column volumes of 20 mM sodium-acetate buffer pH 5.7 with 20 mM imidazole, 20 mM sodium chloride and 1 mM DTT. The HisTrap bound proteins were eluted by increasing imidazole concentration gradient to 300 mM imidazole.

The solubilised sample (Sol. S) of section 2.4.4 was loaded into a 100 mM sodium-phosphate buffer pH 5.7 with 20 mM imidazole, 20 mM sodium chloride and 6 M urea pre-equilibrated GE Healthcare 5 ml HisTrap column, connected to a Bio-Rad NGC purification system. The column was washed with 5 column volumes of 100 mM sodium phosphate buffer pH 5.7 with 20 mM imidazole, 20 mM sodium chloride and 6 M urea, followed by a 10 column volumes elution of column bound his-tagged protein with increasing imidazole concentration gradient to 300 mM imidazole in a 100 mM sodium phosphate buffer pH 5.7 with 20 mM sodium chloride, and 4 M urea whilst collecting 2 ml fractions. The collected best performing fractions (9 and 10) were concentrated and dialysed as described below (section 2.7) into 50 mM sodium phosphate buffer pH 5.7 with 1 mM DTT.

2.7 Sample denaturant dialysis for spectroscopy analysis

The purified solubilised protein samples (section 2.6) were concentrated and loaded into a buffer soaked dialysis (snakes skin dialysis tubing) tube. The dialysis tube with sample was submerged into 50 mM sodium phosphate buffer pH 5.7 with 4 M urea, 0.02% (w/v) sodium azide and 1 mM DTT for 4 hours while stirring. The buffer was changed to 50 mM sodium phosphate buffer pH 5.7 with 2 M urea, 0.02% (w/v) sodium azide and 1 mM DTT and stirred again for 4 hours. Lastly buffer was changed to 50 mM sodium phosphate buffer pH 5.7 with 0.02% (w/v) sodium azide and 1 mM DTT, and stirred overnight

2.8 SDS-PAGE

2.8.1 Electrophoresis

Sample analysis using SDS-PAGE electrophoresis by Laemmli (1970) was conducted. SDS-PAGE with 4% stacking gel and 12% resolving gels were prepared for each analysis. Samples were diluted (1:1) with prepared SDS sample buffer (0.5 M tris buffer pH 6.8 with 33 % (v/v) glycerol, 16 % (v/v) 20% SDS, 4.1% (v/v) beta-mercaptoethanol and 1.6% (v/v) 1% bromophenol blue) and denatured for 10 minutes at 100 °C. The denatured samples were loaded into the wells of the stacking gel. Protein separation in the resolving gel was conducted using a Bio-Rad PowerPac system from Bio-Rad laboratories (California, USA) for approximately 1 hour at 160 V. The separated protein sizes were determined in relation to the standard protein molecular weight marker from Thermo Scientific (Massachusetts, USA) run under the same separation conditions. The gels were stained with Coomassie brilliant blue stain solution (0.1 % (v/v) Coomassie brilliant blue R-250, 50 % (v/v) methanol and 10 % (v/v) glacial acetic acid) for approximately 1 hour and destained in destain solution (20 % (v/v) methanol, 10 % (v/v) acetic acid and 70 % (v/v) distilled water) overnight. The destained gels were analysed using the Bio-Rad ChemiDoc™ MP image Lab 6.0.1 system from Bio-Rad laboratories (California, USA). The Thermo Fisher unstained protein molecular weight marker with seven native proteins (14.4 to 116 kDa) was used as the size standard in protein SDS-PAGE.

2.8.2 Gel analysis

All gels lanes with sample proteins were analysed by Bio-Rad gel doc system (ChemiDoc™ MP Image Lab 6.0.1 software). The protein size is calculated in relation to the molecular weight marker *(marked STD lane* molecular weight set standards from Thermo Fisher*)* in every gel (SDS-PAGE) using the point to point (semi log) regression method. There is no single equation available for the point-

to-point regression method, the slope (obtained from the log of the molecular weight values) of each segment of the curve between data points was independently calculated. The ChemiDoc™ MP Image Lab 6.0.1 software automatically detected lanes, which were manually adjusted. In each lane, bands were automatically detected at high sensitivity to detect all (including feint) visible bands. The bands band percentage volume (Band %) in comparison to all bands (100%) volumes per lane were determined using the ChemiDoc™ MP Image Lab 6.0.1 software (Bio-Rad user guide, 2017).

2.9 Spectroscopic techniques

2.9.1 Absorbance spectroscopy

Absorbance is used to determine protein concentration at a specific wavelength of light by using the Beer-Lambert law. The absorbance reading was recorded with a Chirascan Plus spectrometer from Applied Photophysics, (Leatherhead, UK). The purified VP7 sample in 50 mM sodium phosphate buffer 7.5 with 1 mM DTT, concentration was determined using the Beer-Lambert law equation ((1) section 1.8.1). The recombinant VP7 extinction coefficient at 280 mM was determined to be 70290 M^{-1} cm^{-1}using the sequence analysing tool (ProtoParam tool) in Expasy web software (web.expasy.org).

2.9.2 Fluorescence spectroscopy

The fluorescence spectra are useful to measure the protein tertiary structure tryptophan residues light emission to determine proteins conformational changes. Tryptophan residues display a maximum fluorescence intensity emission between 308 - 350 nm wavelengths (Vivian & Callis, 2001) when excited at 295 nm l (Tang *et al.,* 2009). Fluorescence spectra were recorded with a Chirascan Plus spectrometer from Applied Photophysics, (Leatherhead, UK). The purified VP7 sample in 50 mM sodium phosphate buffer, pH 7.5 with 1 mM DTT, fluorescence emission intensities spectra were recorded in a 1 cm quartz cuvette at 20 °C with a 1 mm bandwidth. The protein fluorescence spectra were recorded from 295 nm to 500 nm while VP7 tryptophan residues were excited by light at 295 nm wavelength. The averages of 3 fluorescence emission intensities spectra reading were extracted and plotted using Microsoft Excel 2010.

2.9.3 Fluorescence spectroscopy thermal stability

The fluorescence spectrum was recorded with a Chirascan Plus spectrometer from Applied Photophysics, (Leatherhead, UK). The purified 7.40×10^{-7} M VP7 sample in 50 mM sodium phosphate buffer 7.5 with 1 mM DTT fluorescence emission spectra were recorded in a 1 cm quartz

cuvette at increasing temperature from 20 to 90 °C for every 1 °C increase interval controlled by a single cell Peltier temperature controller. The protein fluorescence spectra were recorded from 295 to 500 nm while VP7 tryptophan residues were excited at 295 nm wavelength as the protein solution was heated.

2.9.4 Circular dichroism

Protein structures can be measured and characterised by using spectroscopic techniques such as circular dichroism (CD) (in low resolution in solution) (Brahms *et al.,* 1977). CD refers to the differential absorption of left and right circularly polarized light (Whitmore & Wallace 2008). The CD spectra (far UV spectra) are used to characterise and measure proteins secondary structure (Whitmore & Wallace 2008; Corrêa & Ramos 2009) (Section 1.8.3).

2.9.4.1 Far UV CD

Far UV spectra measured between 190 nm and 250 nm wavelength (Chenal *et al.,* 2005) are used to determine the proteins secondary structure, by checking signals that indicate the presence and/or absence of the proteins alpha helices and beta strands (Mallam and Jackson, 2005). The absorbance spectrum of a protein will be based on secondary structure at wavelengths < 250 nm (Kelly et al. 2005).

The purified $7.40*10^{-7}$ M VP7 sample in 50 mM sodium phosphate buffer 7.5 with 1 mM DTT Far UV was recorded in a 1 cm quartz cuvette at 20 °C, with a data pitch of 0.2 nm and a response time of 0.5 sec. The sample average of 3 spectrum readings at wavelength 222 nm was recorded and subtracted the buffer reading to obtain the protein reading. The data obtained were converted to mean residue ellipticity [Θ] $deg.cm^2.dmol^{-1}$ using the formula:

$$[\Theta] = 100(\text{signal in mdeg})/c.n.l \qquad (2)$$

Where 'Θ' is the mean residue ellipticity in $deg.cm^2.dmol^{-1}$, 'c' the protein concentration in mM, 'n' is the number of polypeptide chain amino acid residues and 'l' the path length in cm

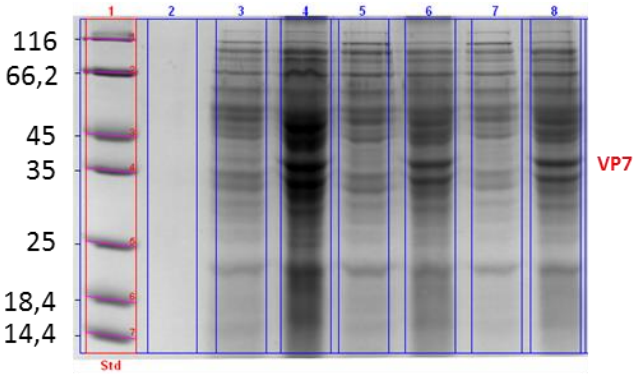
3. Results

3.1. VP7 over expression and solubilisation

Cell lines (BL21 (DE3) and C43 (DE3) pLysS) were transformed with a pET28a+ plasmid engineered by GenScript (New Jersey, USA) to have a rotavirus VP7 protein coding sequence (pET28a+VP7). The plasmid contains a kanamycin resistance gene, which results in the transformed cells able to resist and survive in kanamycin antibiotic present environment. The C43 (DE3) pLysS cells contain a pLysS plasmid, which consists of the chloramphenicol resistance gene, which causes the C43 (DE3) cells to resist and survive the chloramphenicol antibiotic environment. Cells growth in kanamycin antibiotic environment, confirmed the cells successful pET28a+ plasmid transformation.

IPTG is a compound that mimics allolactose and binds a repressor protein inducing a conformational change, causing release of repression (Rosano and Ceccarelli, 2014). This triggers transcription of genes downstream of the *lacI* region in the pET28a+ plasmid upstream of the inserted protein coding gene. IPTG is not metabolised by the cell therefore, the IPTG concentration remains constant during induction. IPTG concentration determines the level of overexpression of the recombinant protein. The transformed cell lines were induced with different IPTG (0.2; 0.5 and 1 mM) concentrations (Figure 3.1a and 3.2a) to determine the optimal inducer concentration required for optimal overexpression of the recombinant protein (VP7). VP7 expression is observed as an approximately 37 kDa size band, by analysis of SDS-PAGE (section 2.8). VP7 expression was observed in cells induced by three different IPTG concentrations in both cell lines (Figure 3.1a and 3.2a). Although cell numbers were not normalized visual inspection shows little difference between BL21 cells induced with 0.2, 0.5 and 1 mM IPTG (Figure 3.1a). The samples were not normalized which could affect the precise reading of the results. Therefore, 0.2 mM IPTG induction results indicate more cells rather than higher expression levels in comparison to cells induced with 0.5 mM IPTG and 1 mM IPTG (Figure 3.1a). Optimal VP7 expression is found to be obtained with 0.5 mM induction when observed in comparison to 0.2 and 1 mM IPTG as the higher concentration will be unnecessarily wasteful. The obtained results indicate that VP7 is predominantly overexpressed in the insoluble form (pellet) in both cell lines (Figure 3.1a and 3.2a) using all three IPTG concentrations. The soluble BL21 (DE3) samples indicate little (faint) to no signs of VP7 expression (Figure 3.1a), however C43 (DE3) pLysS cells indicate traces of soluble VP7 expressed (Figure 3.2). Little difference in expression levels between C43 (DE3) pLysS cells induced with 0.2, 0.5 and 1 mM IPTG are seen (Figure 3.2).

a



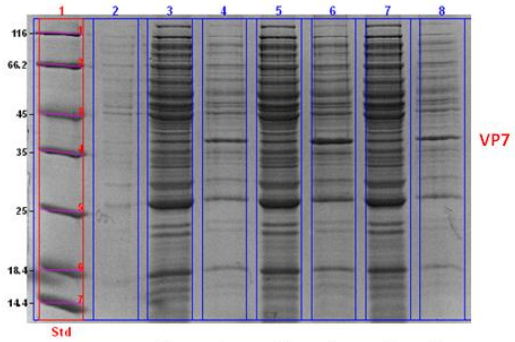
b

Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc™ MP Image Lab 6.0.1 software (kDa)
Lane 3 gel A	37.5
Lane 4 gel A	37.0
Lane 6 gel A	37.3
Lane 8 gel A	37.7

Figure 3.1: BL21 (DE3) IPTG induction

The BL21 (DE3) cells induction samples with different IPTG concentrations (0.2; 0.5 and 1 mM) for VP7 overexpression loaded on SDS PAGE. (a) The supernatant sample (S) indicates soluble protein sample and pellet (P) indicates insoluble protein. (b) The overexpressed bands were quantified and analysed by ChemiDoc™ MP Image Lab 6.0.1 software (b). The labelled bands (VP7) in the SDS-PAGE represent a band of 37 kDa migration (a) of the respective bands (band no) suspected and labelled to be VP7 (b). Table (b) indicates selected bands sizes (kDa) as calculated by ChemiDoc™ MP Image Lab 6.0.1 software.

a



b

S P S P S P
0.2 mM 0.5 mM 1 mM

Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc ™ MP image Lab 6.0.1 software (kDa)
Lane 3 gel A	37.7
Lane 4 gel A	37.8
Lane 5 gel A	37.1
Lane 6 gel A	37.6
Lane 7 gel A	37.3

Figure 3.2: C43 (DE3) pLysS IPTG induction

Samples taken from the induction of expression with different IPTG concentration (0.2; 0.5 and 1 mM) visualised on SDS PAGE. (a) The samples supernatant (S) and pellet (P), are indicated for each IPTG concentration (0.2, 0.5 and 1 mM). (b) The over-expressed bands were quantified and analysed by ChemiDoc™ MP Image Lab 6.0.1 software. The labelled bands in the SDS-PAGE (a) are the respective bands (band no.) suspected and labelled to be VP7. The table indicates calculated protein MW of selected expressed bands with approximate VP7 size by the ChemiDoc™ MP Image Lab 6.0.1 software.

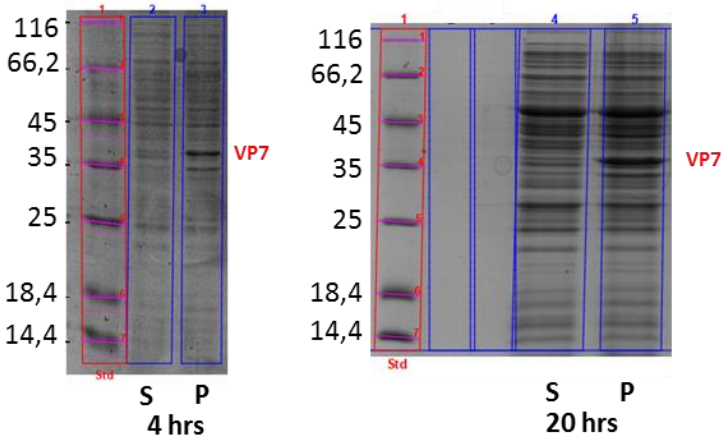
Although 0.5 and 1 mM IPTG seem to produce slightly more VP7 in the pellet samples (Figure 3.2) using the lower (0.5 mM) concentration will therefore produce best level of expression while reducing costs.

These findings suggest VP7 can be expressed in soluble form by C43 (DE3) pLysS cell line compared to BL21 (DE3) cells. Soluble proteins are proteins that are found in homogeneity (the state of being the same or all of the same kind) with solvent (buffer) environment and not forming any aggregates (Voet and Voet, 2004). Insoluble proteins are proteins found to be precipitated aggregates or placed into an inclusion body forming a pellet which is not preferred for spectroscopy studies, because pellets do not represent the native conformation of the protein and insoluble protein give rise to spectral artefacts such as differential light scattering during spectroscopic work (Miles and Wallace, 2016) or analysis when not solubilised.

The time it takes to overexpress protein, can determine the yield of overexpressed protein and can also affect the proteins solubility (Rosano and Ceccarelli, 2014). The overexpression time periods (4 and-20 hours) after IPTG induction were observed to determine the time period resulting in the best overexpression of VP7 in the soluble form. The time period determines the expression time in which the recombinant protein is overexpressed to the highest level by the cells longer expression time can however result in cells death and halting the protein overexpression (Miroux and Walker, 1996). BL21 (DE3) and C43 (DE3) pLysS cells, harbouring the pET28a+VP7 plasmid showed that longer overexpression time periods, resulted in increased cells growth and higher amounts of overexpressed recombinant protein in both cell lines (Figure 3.3 and 3.4). Soluble VP7 (37 kDa) protein were expressed after a growth time of 4 hours and 20 hours overexpression (Figure 3.4). Therefore, shortening the induction growth time period did not increase soluble VP7 overexpression (37 kDa) in both cell lines (BL21 and C43 (DE3) pLysS Cells) (Figure 3.3 and 3.4). The 20 hours induced overexpression time period was selected as it results in high C43 (DE3) pLysS cell growth and therefore more of overexpressed VP7 protein (Figure 3.4).

VP7 overexpression at 20 °C and 37 °C temperature was conducted to determine if temperature impacts the proteins solubility status. The 20 °C environment was expected to lower the cells growth rate, and therefore the rate of protein expression, which favours correct folding of the protein and avoiding formation of misfolded and aggregated protein (Ahmad *et al.,* 2018). The cell growth is relatively similar in both cell lines based on the observed overexpressed bands (Figure 3.5A), however, soluble proteins were expressed in 37 °C temperature, indicating higher temperature favours the formation of more soluble protein.

a

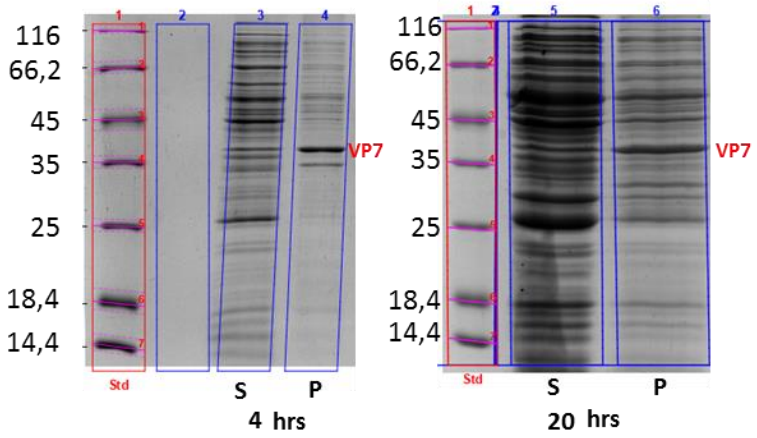


b

Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc™ MP Image Lab 6.0.1 software (kDa)
Lane 2 gel A	37.4
Lane 3 gel A	37.3
Lane 4 gel A	37.0
Lane 5 gel A	37.0

Figure 3.3: VP7 overexpression time period after induction in BL21 (DE3)
The BL21 (DE3) cells induced with 0.5 mM IPTG for VP7 overexpression for 4 and 20 hours at 37°C. The overexpressed VP7 bands (labelled) in SDS PAGE (a) analysed size by ChemiDoc™ MP Image Lab 6.0.1 software (b). The SDS PAGE (a) supernatant (S) indicates soluble protein sample and pellet (P) insoluble protein.

a

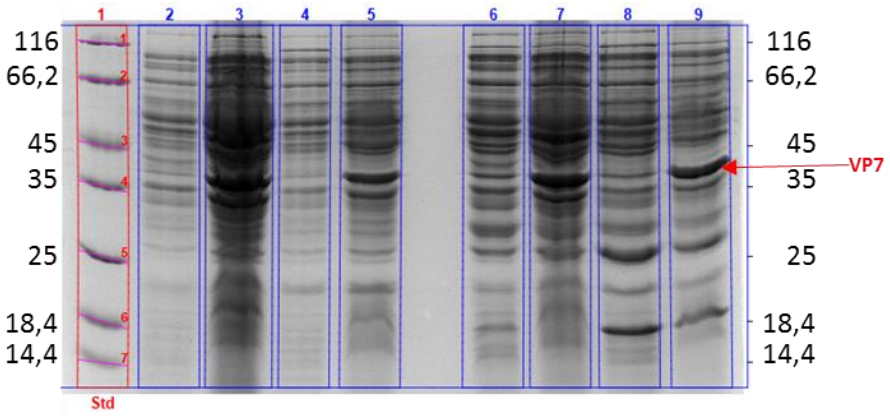


b

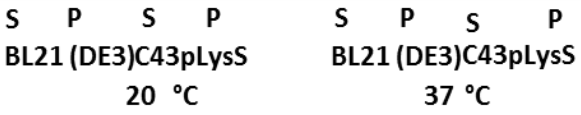
Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc™ MP Image Lab 6.0.1 software (kDa)
Lane 3 gel A	37.4
Lane 4 gel A	37.6
Lane 5 gel A	37.4
Lane 6 gel A	37.2

Figure 3.4: VP7 overexpression time period after induction in C43 (DE3) pLysS
The C43 (DE3) pLysS cells induced with 0.5 mM IPTG for VP7 overexpression for 4 and 20 hours at 37°C. The SDS PAGE (a) lanes loaded with 4 and 20 hours induced expression samples. The bands were analysed using ChemiDoc™ MP Image Lab 6.0.1 software, indicating labelled bands (b) at approximately 37 kDa as target VP7 protein (b).

a



b



Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc™ MP Image Lab 6.0.1 software (kDa)
Lane 2 gel A	37.4
Lane 3 gel A	37.0
Lane 4 gel A	37.1
Lane 5 gel A	37.0
Lane 6 gel A	37.1
Lane 7 gel A	37.0
Lane 8 gel A	37.2
Lane 9 gel A	37.4

Figure 3.5: VP7 overexpression solubility at 20 and 37 °C temperature in C43 (DE3) pLysS and BL21 (DE3)

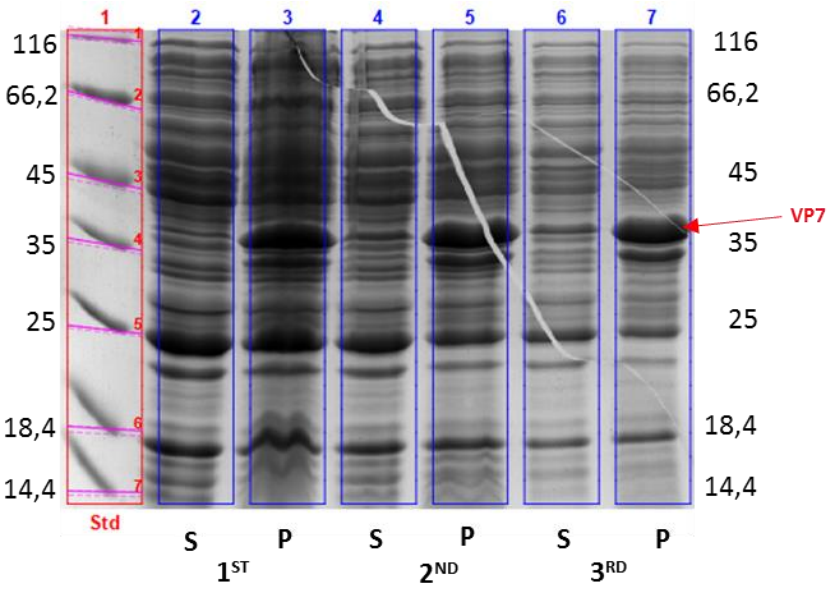
The C43 (DE3) pLysS and BL21 (DE3) cells were induced with 0.5 mM IPTG concentration for VP7 overexpression at 20 and 37 ° C for 20 hours. The soluble (S) and insoluble (P) samples were loaded to an SDS PAGE (a). The bands in SDS PAGE lanes were analysed with ChemiDoc™ MP Image Lab 6.0.1 software for overexpressed VP7 size per lane (b).

Insoluble VP7 is overexpressed in both cell lines at both temperatures (Figure 3.5). Both cell lines indicate more insoluble VP7 overexpressed, over soluble VP7 at both 20 and 37 °C.

The overexpressed VP7 protein is observed as a band with an approximate 37 kDa (between 36.5 and 37.8 kDa) depending on the overexpressed band position/section identified and analysed by ChemiDoc™ MP Image Lab 6.0.1 software in the SDS-PAGE. The results (Figure 3.5b), indicate 37 °C as an ideal temperature over 20 °C to overexpress VP7, with more cells expressed VP7 at 37 °C compared to 20 °C, from the same bacterial cell line. Both cell lines overexpress VP7 at both temperatures (Figure 3.5b) however C43 (DE3) pLysS cell line was selected as it has shown slight expression of soluble VP7 (Figures 3.2; 3.4; 3.5 and 3.6).

Sonication degrades the cell membranes, releasing the cells macro and micro particles as lysates (Singh et al. 1990). Recombinant protein can be insoluble (entrapped in the inclusion bodies) and soluble (reach homogeneity with the solvent). Since sonication is used for various purposes such as the extraction of multiple compounds in cells by the act of applying sound energy to agitate particles in a sample (Singh et al. 1990; Rodríguez-Carmona *et al.,* 2010), an attempt to dissociate the inclusion bodies formed and/or aggregated protein by additional sonication (sonication of pellet after centrifugation and resuspension in buffer) was conducted. Additional sonication steps did not result in complete production of soluble VP7 but the pellet was further disrupted, with each round of sonication (Figure 3.6) which suggests that inclusion bodies might be affected by multiple rounds of sonication. With multiple sonication rounds not resulting in complete solubilisation of VP7 (Figure 3.6), large amount of VP7 remains insoluble therefore a solubilisation protocol (unfolding native-like protein structures by using denaturant), adapted from QIAGEN purification handbook ref (The QIA expressionist, 2003). (Section 2.4.4) was tested (Figure 3.7a) to unfold folded protein and dissociate the aggregated polypeptides. Solubilisation with 6 M urea as a denaturant resulted in VP7 being detected in soluble sample (Figure 3.7).

a



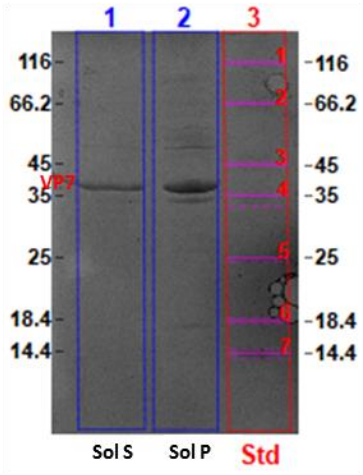
b

Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc™ MP Image Lab 6.0.1 software (kDa)
Lane 2 gel A	37.2
Lane 3 gel A	37.0
Lane 4 gel A	37.3
Lane 5 gel A	37.0
Lane 6 gel A	37.1
Lane 7 gel A	37.5

Figure 3.6: VP7 pellet multiple rounds of sonication

The C43 (DE3) pLysS cells pellet disruption by rounds of sonication. The overexpressed bands analysed on SDS-PAGE (a) and with ChemiDoc™ MP Image Lab 6.0.1 software (b) relative to the standard molecular marker in lane 1 (a). The insoluble pellet (P) and soluble (S) samples after each (1^{st}, 2^{nd} and 3^{rd} round) sonication were loaded in an SDS PAGE and the bands sizes (kDa) (b).

a



b

Lane and gel	Molecular weight of protein marked as VP7 as calculated by ChemiDoc™ MP Image Lab 6.0.1 software (kDa)
Lane 1 gel A	37.3
Lane 2 gel A	37.4

Figure 3.7: VP7 inclusion body solubilisation
The C43 (DE3) pLysS cells inclusion body solubilisation from a pellet using 6 M urea. Samples loaded into an SDS PAGE (a) where Sol P indicates the pellet and Sol S the soluble protein obtained after solubilisation. Lane 1 and 2 (a) were analyzed by ChemiDoc™ MP Image Lab 6.0.1 software for band size (kDa) (b).

3.2 VP7 purification

3.2.1 Ion exchange chromatography (salt elution)

Polypeptides have specific charges based on their amino acids net charges (Pergande and Cologna, 2017). Polypeptides in a buffer or solvent with pH close and equal to its polypeptide charge results in a polypeptide zero net charge which is described as its isoelectric point (pI) (Pergande and Cologna, 2017). Using the sequence analysing tool (ProtoParam tool) in Expasy web software (web.expasy.org), the pI value of 4.67 was generated for target protein (VP7) tagged with histidine residues. VP7 is therefore, expected to be negatively charged and bind DEAE resin in buffer conditions with higher pH than 5 and will elute when the buffers pH reaches pH 4.7 as the protein will reach a net charge zero and therefore eluted from the column.

With the knowledge of the target protein's isoelectric point, ion exchange chromatography was conducted to separate proteins based on their charge due to the environment ionic condition and buffer pH conditions. Negatively charged proteins bind to a DEAE column (which is positively charged) and the bound proteins are then eluted with increasing salt concentration which out-compete the column bound macromolecules and bind negatively charged ions to the column as it becomes more negatively charged causing the bound proteins to be eluted.

The macromolecules and charged proteins with a pI value below the buffers pH (pH 5.7) such as VP7 with pI 4.67 are negatively charged and bind the DEAE column (Pergande and Cologna, 2017). The supernatant sample obtained from the cells lysate with low yield VP7 and a large number of bacterial proteins was loaded onto a DEAE column. Buffer with pH 5.7 favoured the binding of VP7 protein and negatively charged proteins over positively and neutrally charged bacterial proteins present in the sample which are washed off during sample application (flow through (FT)) (Figure 3.8a and b). The binding capacity of the column matrix was compromised as other bacterial protein with similar or higher DEAE resin binding affinity (based on net charge) resulted in the inability of some VP7 to bind the matrix (Figure 3.8). The negatively charged macromolecules that bound the DEAE column at pH 5.7 got eluted from the column by the application of chloride ions (from sodium chloride) loaded to the column to outcompete the bound particles. Increasing the concentration of sodium chloride, results in increased chloride ion interaction with the column, causing the elution of bound proteins. Elution of VP7 occurs at less than 10 % of the gradient, which ends in a final 0.7 M sodium chloride concentration

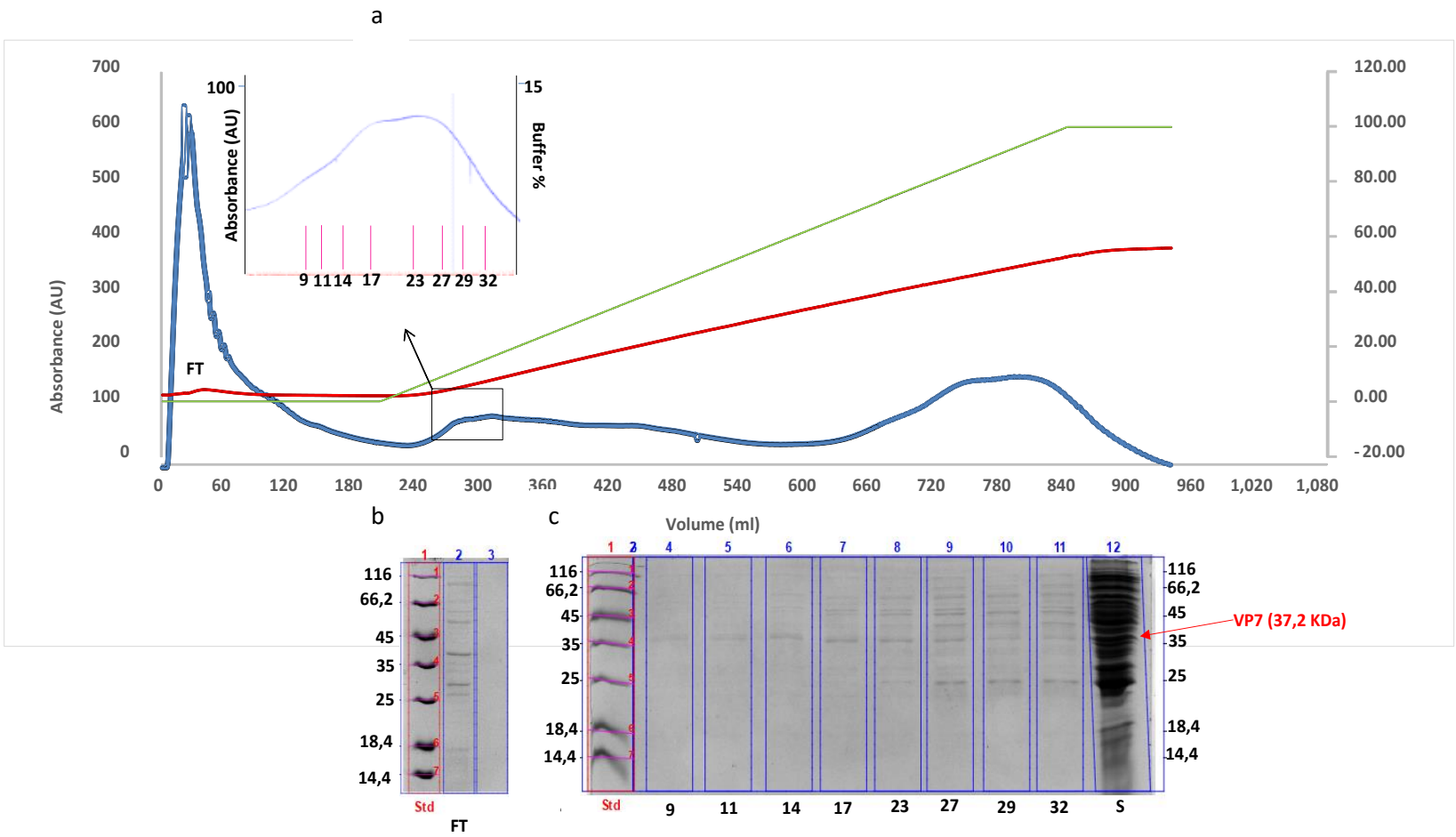


Figure 3.8: Ion-exchange purification by 0-0.7 M sodium chloride gradient

The salt gradient ion exchange chromatography profile (a) indicating the salt gradient (green), conductivity (red) and absorbance (blue). The boxed peak (Inset curve) is the elution peak of fractions (9 to 46) loaded in SDS-PAGE (c). The flow-through (FT) in SDS-PAGE (b), supernatant sample (S) and fractions with increasing absorbance were loaded to an SDS-PAGE (c) and analysed for the presence of VP7.

Therefore, VP7 elutes between approximately 0.2 and 0.3 M sodium chloride as can be seen by the increase in absorbance on the chromatograph (Figure 3.8a) and accompanying bands representing the 37.2 KDa protein as calculated by the ChemiDoc™ MP Image Lab 6.0.1 software within size range for VP7 (Figure 3.8b). The chromatograph profile indicates absorbance increase of two broad peaks, (Figure 3.8a) indicating other proteins eluted with VP7 in the first peak, which is also confirmed by the SDS-PAGE analysis of the fractions (Figure 3.8b). It should be noted that the second higher absorbance peak (Figure 3.8a) does not contain VP7 bands (hence the gel is not included), indicating VP7 binding was outcompeted by more negatively bound bacterial proteins, explaining why some VP7 did not bind the column (Figure 3.8b lane 2).

Ion exchange chromatography was conducted where an increasing salt gradient concentration to 0.5 M sodium chloride was used for elution (Figure 3.9) to increase separation of VP7 from other cellular proteins that eluted along with VP7 during salt gradient elution to 0.7 M sodium chloride (Figure 3.8). The eluted fraction from the first peak (Figure 3.9a) were initially too dilute and no bands were recognised (Figure 3.9b). The dilute fractions were pooled, concentrated and assessed on SDS-PAGE to analyse the presence of contaminants (Figure 3.9c). The second higher absorbance (Figure 3.9a) peak fractions had no VP7 protein bands (Gel not included), further confirming that VP7 binding is outcompeted by more negatively charged bacterial proteins.

The lower final concentration of 0.5 M sodium chloride did not produce conditions that successfully separated VP7 from all the bacterial proteins present in the supernatant sample (Figure 3.9), as VP7 makes up 47.4 % of all the bacterial proteins in the eluted fraction (Appendix 7.1) in lane 5. Different sodium chloride concentrations were used to confirm the concentration (between 0.1; 0.2 and 0.3 M sodium chloride) to elute VP7 from the DEAE column (Figure 3.10a). Supernatant was loaded unto the column with zero sodium chloride in the buffer, and bound proteins were eluted with buffer containing 0.1 M sodium chloride (Figure 3.10c). The eluting buffer was exchanged to buffer with 0.2 M sodium chloride concentration for 3 column volumes at which point stable conductivity indicated the sodium chloride concentration was stable then the elution buffer was changed with buffer containing 0.5 M sodium chloride (Figure 3.10). VP7 was eluted during buffer exchange from 0.1 to 0.2 M sodium chloride containing buffer (Figure 3.10b) but was still not successfully separated from the bacterial proteins. VP7 constituted 45.3% of the eluted fraction in lane 7 indicating that VP7 share the same charge and DEAE binding affinity with the bacterial proteins as they were eluted by the same conditions (Figure 3.10). Therefore, ion exchange purification to purify VP7 protein based on their net charge in buffer pH was conducted

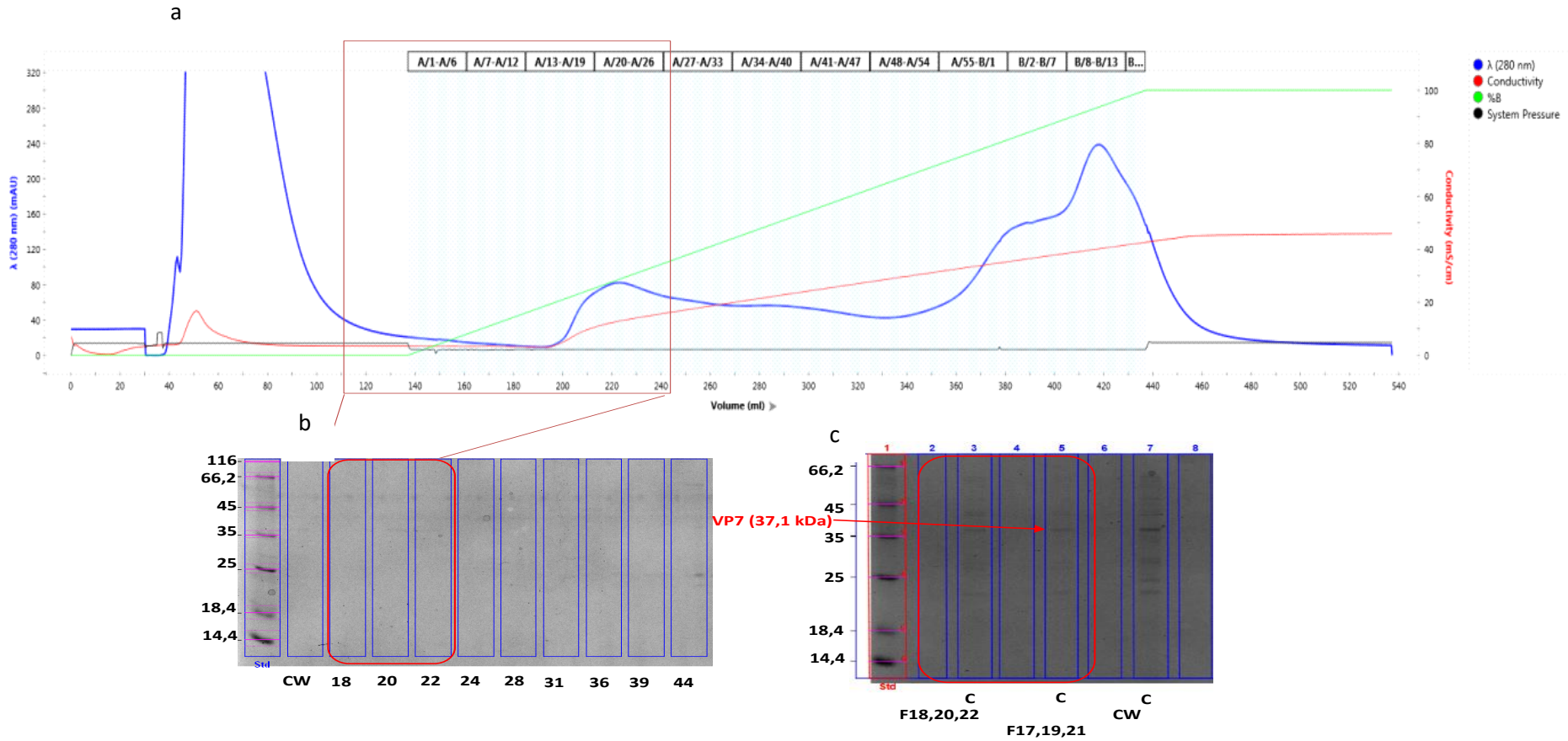
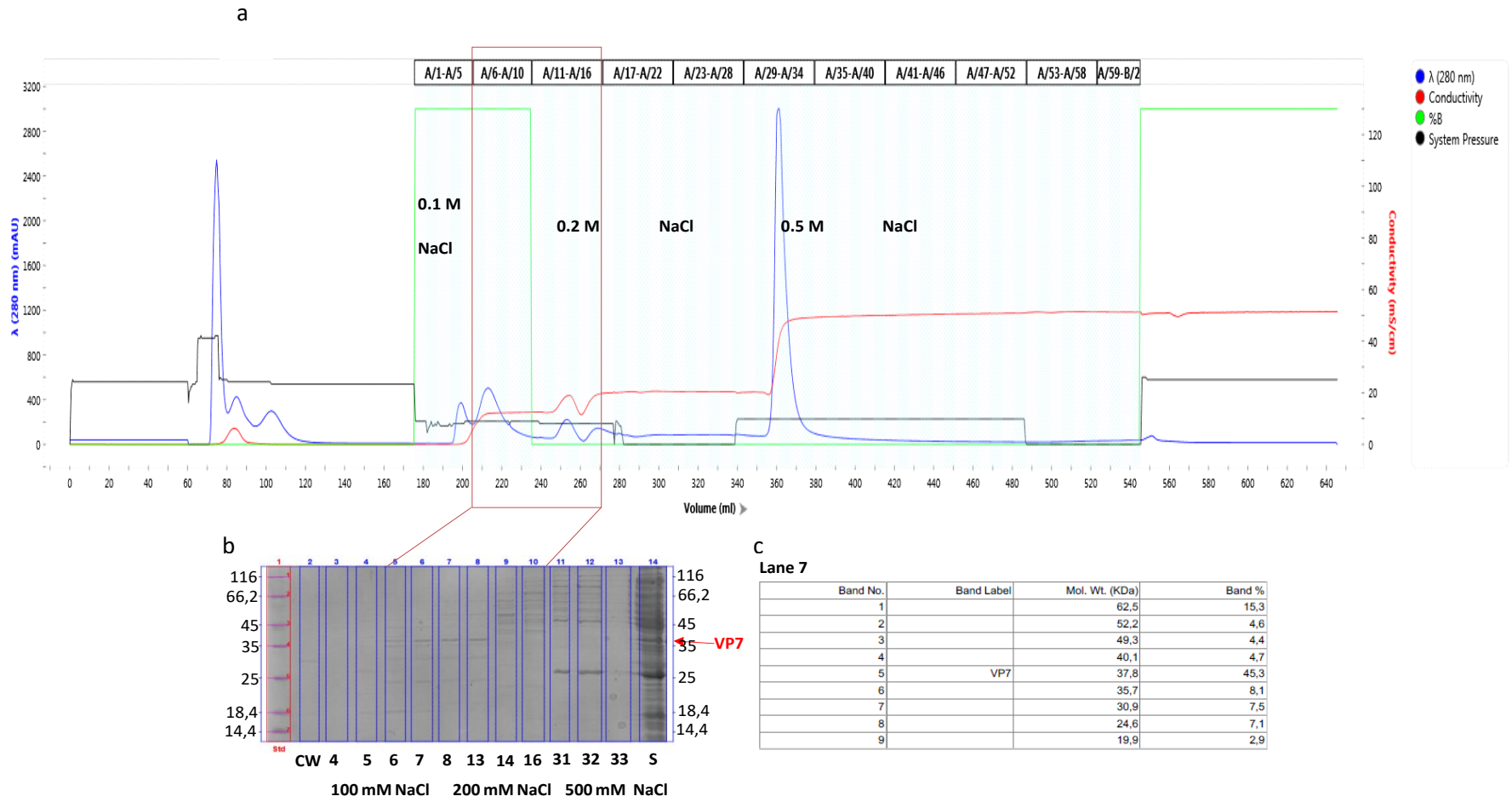


Figure 3.9: VP7 ion-exchange purification elution with a gradient final concentration of 0.5 M sodium chloride

The salt gradient ion exchange chromatography profile (a) indicating the salt gradient (green), conductivity (red) and absorbance (blue). The columns wash (W) and fractions (18 – 44) with increasing absorbance loaded unto SDS-PAGE (b). SDS-PAGE was loaded with the combined and concentrated (c) fractions (F 18, 20 and 22), (F17, 19 and 21), column wash (CW). Band tagged 3 in lane 5 (b) indicate the highlighted bands representing VP7 protein estimate size analysed and calculated by ChemiDoc™ MP Image Lab 6.0.1 software relative to the standard molecular marker. Full SDS-PAGE analysis Appendix 7.1.



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		62,5	15,3
2		52,2	4,6
3		49,3	4,4
4		40,1	4,7
5	VP7	37,8	45,3
6		35,7	8,1
7		30,9	7,5
8		24,6	7,1
9		19,9	2,9

Figure 3.10: VP7 ion-exchange purification by elution with 0.1; 0.2 and 0.5 M sodium chloride buffers.
The salt ion exchange chromatography profile (a) indicating the salt gradient (green), conductivity (red) and absorbance (blue). The column wash (CW), supernatant sample (S) and fractions with increasing absorbance loaded unto SDS-PAGE (b) were analysed for the presence of VP7. Lane 5 was analysed by ChemiDoc™ MP Image Lab 6.0.1 software (c). Full SDS-PAGE analysis in Appendix 7.2

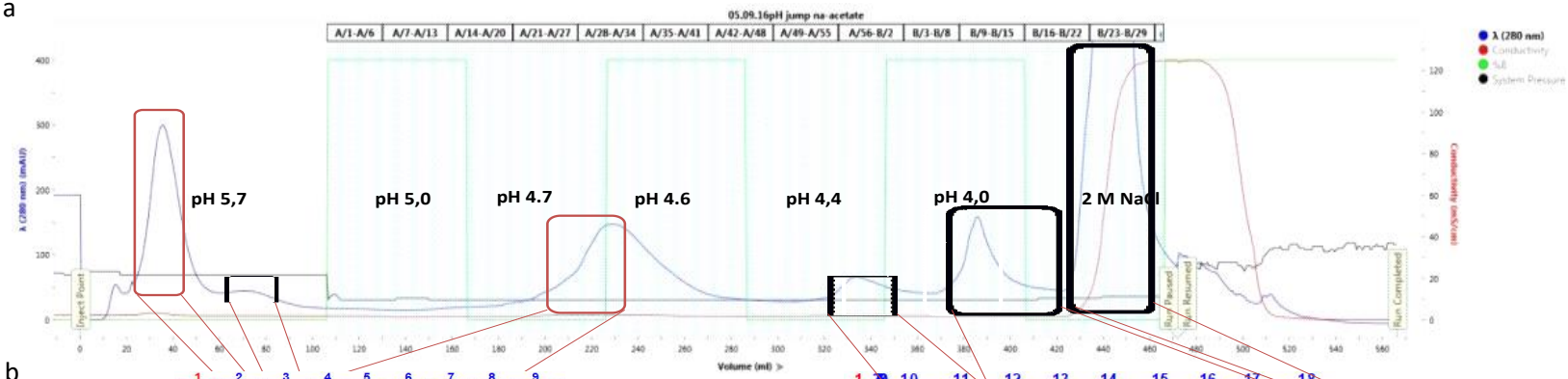
3.2.2 Ion exchange chromatography (pH elution)

With knowledge of VP7 pI value of 4.67, pH exchange was conducted to separate protein based on their charge in relation to buffers pH and not salt concentration. The supernatant was loaded in pre-equilibrated 20 mM sodium acetate buffer pH 5.7 column for purification. The column wash shows no presence of target proteins, indicating that VP7 did bind the column (Figure 3.11). The bound proteins, including VP7 were eluted by different elution buffers with pH values of 5.0; 4.7; 4.6; 4.5 and 4.0 (Figure 3.11 and 3.12). The eluted fractions resulted in 38.8 % and 20.2% VP7 purity eluted by presumed pH 4.6 and 4.4, respectively (Appendix 7.3 lane 8). The elution results indicate VP7 being eluted in different buffers between pH 4.6 and 4.7 and pH 4.4 which could be caused by changing buffers before the set buffer pH was reached (Hirish and Tsonev, 2008; Kroner and Hubbuch, 2013). These results indicate that the pH assumed to be eluting or running through the column was different to the actual pH in the column (pH electrode was not used in these purifications) (Figure 3.11 and 3.12).

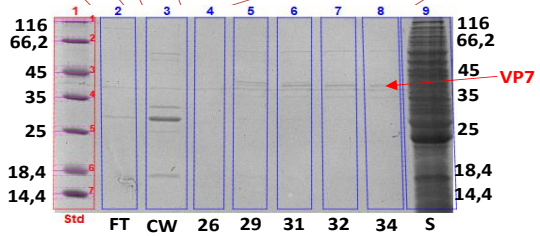
VP7 was eluted with bacterial proteins at different buffers pH thus affecting VP7 purity. Traces of VP7 elution were observed at pH 4.7 (Figure 3.12) with 16.1 % VP7 purity (Appendix 7.4 lane 15), pH 4.7 with 35.8 % (Figure 3.11) and pH 4.0 and 4.4 with 20.6 % and 5.4 % VP7 purity (Figure 3.11 and 3.12, respectively), (Appendices 7.3 lane 11 and 7.4 lane 18, respectively). These results indicate the inconsistency; non-reliable pH buffering that could be caused by the use of low buffer volume (3 column volume) per pH reading and inability to read the buffer pH.

The pH electrode probe on the AKTA prime plus purification system was used to identify and confirm the buffers pH during elution, as pH gradients buffers inherently will resist pH change to a point, a specific stable pH is reached resulting in protein with pI = pH eluted (Sluyterman and Wijdenes, 1977; Shan and Anderson, 2000). Ion exchange purification (Sluyterman and Wijdenes, 1977) was conducted at decreasing pH buffer to identify the pH to elute other non-target proteins and VP7 (Figure 3.13). Supernatant sample with overexpressed VP7 band was loaded at pH 5.7 buffer and the column was washed with the same pH 5.7 buffer. The proteins were then eluted by linear gradient mixing buffer with pH 5.7 and pH 4.0 to decrease buffer pH from pH 5.7 to pH 4.0 (Figure 3.13a). With decreasing buffer pH gradient to pH 4.0, target protein (VP7) was eluted at expected buffer pH range between pH 5.0 and 4.0 (Figure 3.13). The elution results indicate the presence of VP7 (37 kDa) in lane 5 with 44.4% purity (Figure 3.13b) (full SDS-PAGE analysis in Appendices 7.5) indicating more (higher percentage) VP7 eluted, however, not pure.

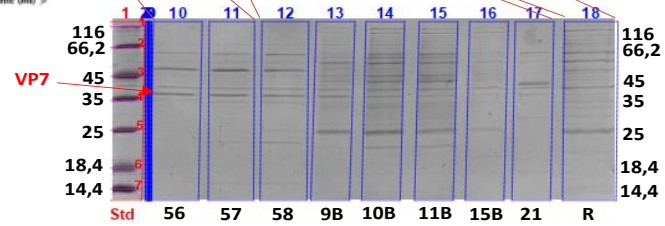
a



b



c



Lane 8

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		58,1	33,5
2	VP7	37,4	35,8
3		34,3	30,7

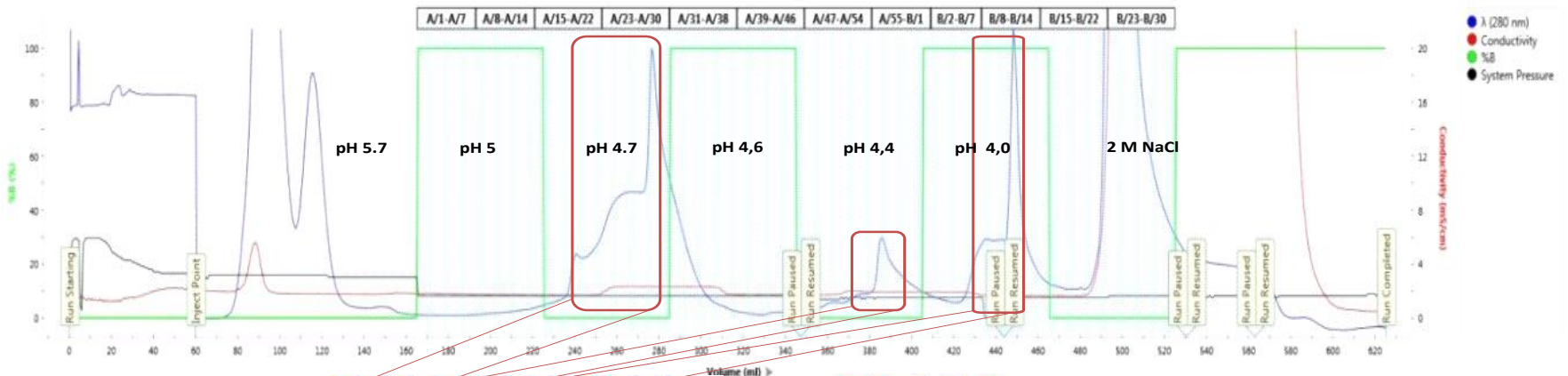
Lane 12

Band No.	Band Label	Mol. Wt. (KDa)	Band %
5		53,9	9,8
6		48,5	27,2
7		45,2	2,3
8		44,0	3,1
9		41,5	1,5
10		40,0	4,7
11	VP7	36,8	20,6

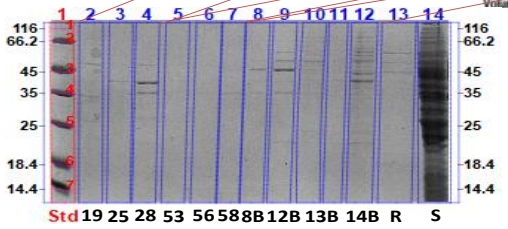
Figure 3.11: VP7 ion-exchange purification by pH exchange

The pH exchange ion exchange chromatography profile (a) indicating the salt gradient (green), conductivity (red) and absorbance (blue). The column wash (CW), supernatant sample (S) and fractions (26 – 34) fractions (56 – 15B and 21) and 2 M sodium chloride regeneration (R) fraction with increasing absorbance loaded in SDS-PAGE (b) and analysed for the presence of VP7. The labelled bands are the bands with approximate 37.4 and 37.2 kDa sizes analysed and calculated by the ChemiDoc™ MP Image Lab 6.0.1 software in relation to the molecular weight marker(lane 1) (c).Full analysis available in Appendix 7.3

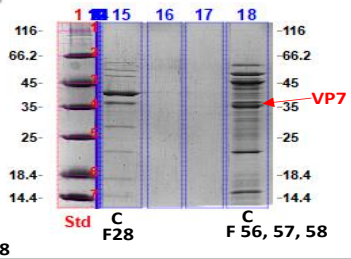
a



b



c



d

Lane 15

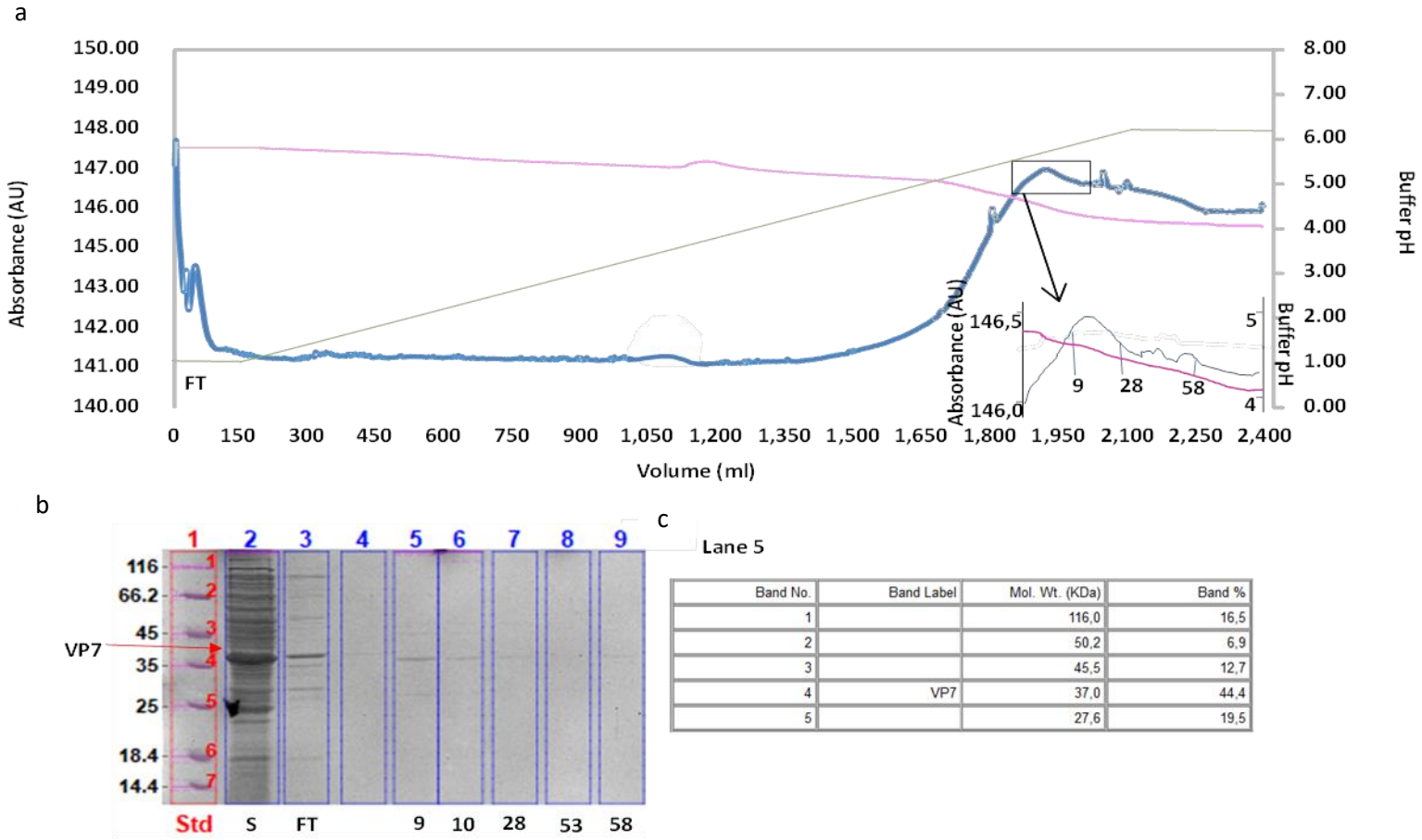
Band No.	Band Label	Mol. Wt. (KDa)	Band %
9		43,2	1,2
10		40,6	37,1
11		39,2	0,6
12	VP7	37,0	16,1

Lane 18

Band No.	Band Label	Mol. Wt. (KDa)	Band %
12		39,9	1,4
13		39,0	1,4
14	VP7	37,1	5,4
15		35,7	5,0

Figure 3.12: Repeat VP7 ion-exchange purification by pH exchange

The pH exchange ion exchange chromatography profile (a). The supernatant (S) and fractions (19 – 24B) loaded onto the SDS-PAGE (b), concentrated fraction 28 (C. F28), combined and concentrated fractions 56 – 58 (C. F56, 57, 58) loaded onto the SDS-PAGE (c) were analysed and calculated by the ChemiDoc™ MP Image Lab 6.0.1 software (d) and the highlighted band show the position of protein size 37 kDa (VP7) (c). Full purification results analysis available in Appendix 7.4.



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	16,5
2		50,2	6,9
3		45,5	12,7
4	VP7	37,0	44,4
5		27,6	19,5

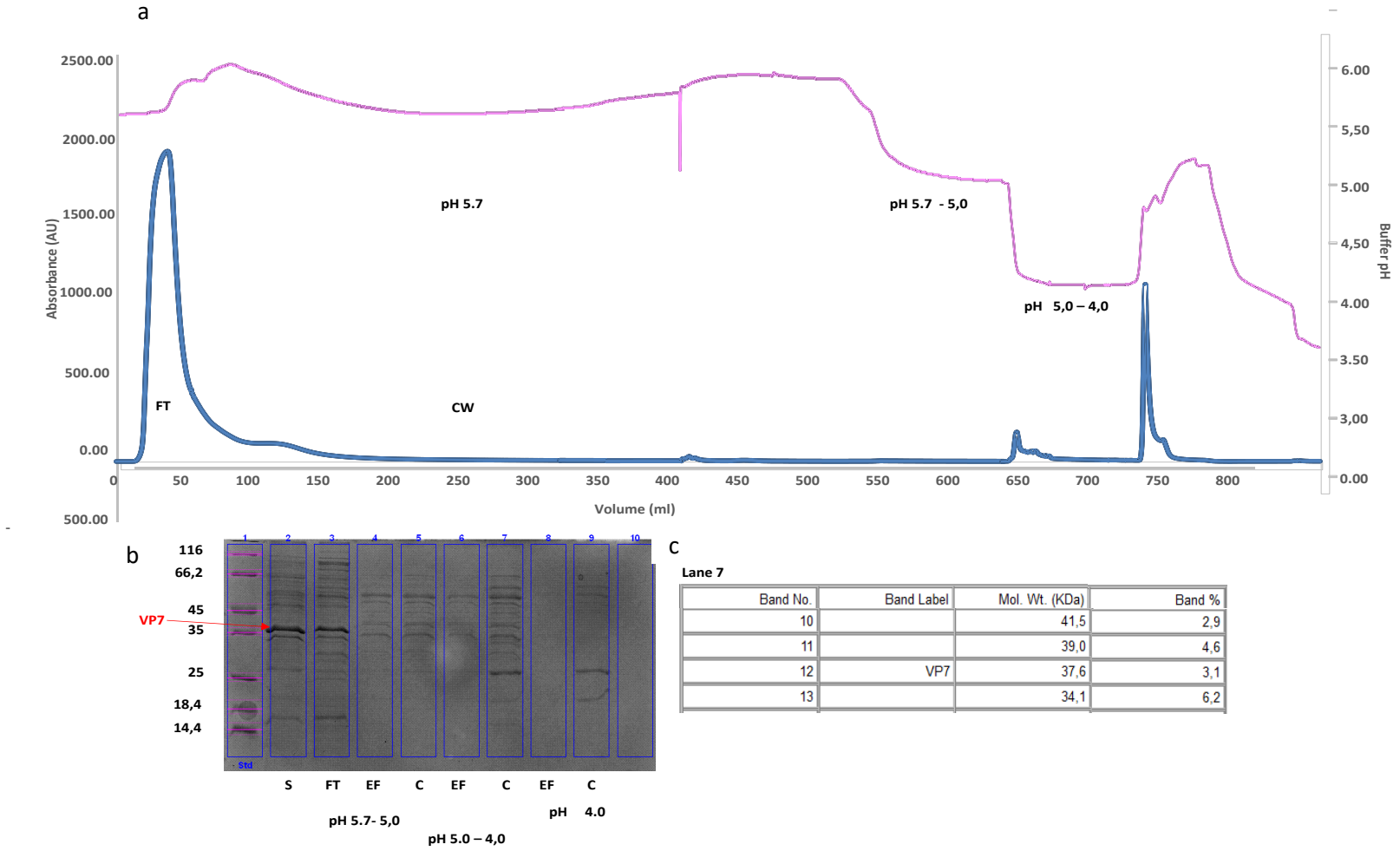
Figure 3.13: VP7 ion-exchange purification by pH gradient

The pH gradient ion exchange chromatography profile (a) indicating, the buffer pH (pink) and the absorbance (blue) and highlighted fractions collections. The column wash (CW), flow through (FT), supernatant sample (S), and fractions with increasing absorbance loaded in SDS-PAGE (b) analysed for the presence of VP7. Fraction 9 band 4 size in lane 5 (b) was analysed and calculated by the ChemiDoc™ MP Image Lab 6.0.1 software to approximate 37 kDa in relation to the molecular weight maker by the Bio-Rad gel dock system (c). Full purification and SDS-PAGE results analysis, available in Appendicx 7.5.

To increase the chances of purification, it was decided to remove other bacterial proteins, the column wash volume would be extended, to ensure stable pH conditions are reached and maintained to wash unbound protein from the resin column. The elution buffer pH range was narrowed down to start from buffer pH 5.7 to 5.0 and from pH 5.0 to 4.0 to purify target protein from bacteria protein based on their pI values (Figure 3.14). During sample application, some VP7 in the sample did not bind the column but flowed through (FT) with the buffer, which could be due to exceeded resin binding capacity (Bolanos-Garcia and Davies, 2006) (Figure 3.14) (Appendix 7.6 lane 2 and 3). The samples eluted (Figure 3.14) at average buffer pH 4.7 contained traces of 3.1 % of VP7 (Appendix 7.6 lane 7), suggesting that some bacterial proteins have the pI value similar to that of VP7 (Figure 3.14) (Bolanos-Garcia and Davies, 2006). It was concluded that VP7 cannot be successfully purified by ion exchange purification alone and therefore an additional chromatography step would be required to purify the protein or an alternative chromatography method such as affinity chromatography could be conducted.

3.2.3 Metal affinity (HisTag) chromatography

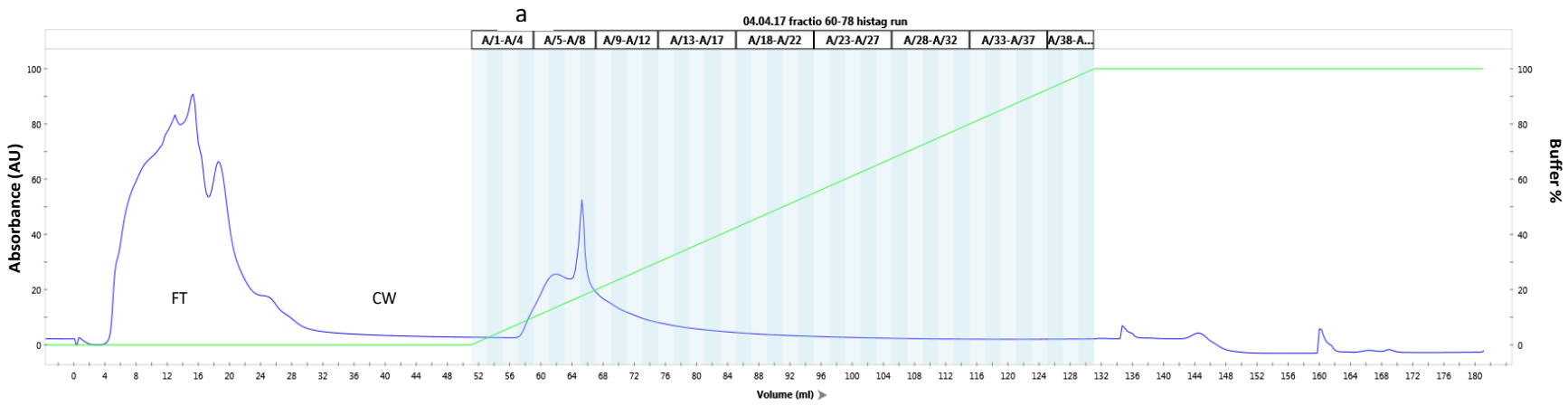
An affinity nickel column purifies protein based on the histidine tag, which will bind to the metal ions of the resin. With VP7 engineered to have a histidine tag attached to the N and C terminus, affinity column to compete with the VP7 histidine residues to bind the nickel column resulting in VP7 elution. The eluted fraction (9 and 10) (Figure 3.13b) (Appendix 7.5 lane 5 and 6, respectively) with VP7 bands sizes from the pH gradient ion exchange purification were combined resulting to 22.3 % of VP7 (Appendix 7.7 lane 2). The combined fractions were then loaded onto a HisTag column (Figure 3.15a) for further purification based on the affinity tag. The chromatograph profile indicates VP7 protein (with histidine residues) binding during sample application and eluting at 25 % of the 300 mM imidazole final concentration gradient (Figure 3.15). Affinity chromatography indicates VP7 purification from most bacterial proteins, which did not bind (FT) the metal ions during sample application in the HisTag column (Figure 3.15b and c lane 3 and 4). The elution of bound protein resulted in partially pure VP7 (69.7% pure) detected (Figure 3.15c lane 6) (Appendix 7.7 lane 6). However, is there a need to conduct ion exchange purification before affinity purification? Affinity purification of the supernatant was conducted to purify the 6.2 % VP7 (Appendix 7.8 lane 2), to identify if VP7 will be purified from the bacterial supernatant (with over 25 detected proteins bands) by affinity purification.

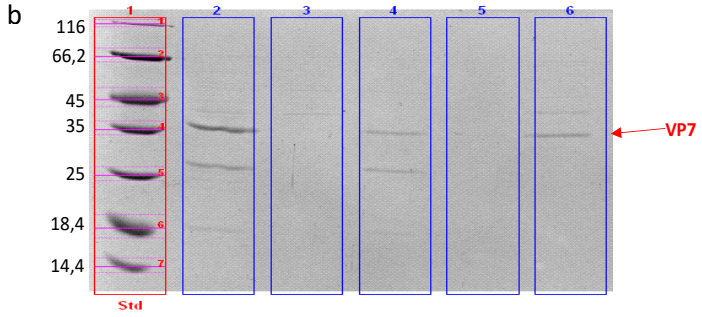


Band No.	Band Label	Mol. Wt. (KDa)	Band %
10		41,5	2,9
11		39,0	4,6
12	VP7	37,6	3,1
13		34,1	6,2

Figure 3.14: VP7 ion-exchange purification by two pH exchange

The pH buffer exchange ion exchange chromatography profile (a) indicating, buffer pH (pink) and the absorbance (Blue). The buffer pH 5.7 – 5.0, pH 5.0 – 4.0 and pH 4.0 elution fractions (EF), concentrated elution fractions (C), flow through (FT), supernatant sample (S), with increasing absorbance loaded in SDS-PAGE (b) were analysed for the presence of VP7. Lane 7 bands sizes were analysed and calculated by the ChemiDoc™ MP Image Lab 6.0.1 software (c) in relation to the standard molecular weight in lane 1 (c). Full SDS-PAGE analysis in Appendix 7.6.





c

	S	CW	FT	6	8

Lane 3

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		66,7	10,7
2		51,3	8,9
3		42,4	15,0
4		24,2	65,5

Lane 4

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		66,6	5,9
2	VP7	36,7	30,8
3		26,7	26,5
4		18,0	36,8

Lane 6

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	8,1
2		49,0	3,7
3		43,4	18,5
4	VP7	36,7	69,7

Figure 3.15: VP7 HisTag purification of ion exchange gradient fractions

HisTag affinity purification profile (a) with elution buffer percentage (green) and the absorbance (blue). SDS-PAGE (b) was loaded with sample (S), flow through (FT), column wash (CW) and fractions 6 and 8 with increase in absorbance. The visible bands (b) were analysed by ChemiDoc™ MP Image Lab 6.0.1 software and calculated the approximate band sizes (c) in kDa in relation to the molecular weight maker (lane 1 of figure b). Full SDS-PAGE lanes in Appendix 7.7.

The bacterial proteins binds the column and gets eluted at different imidazole concentrations (between 30 and 50 % of 300 mM imidazole) (Figure 3.16), resulting in about 20 % pure VP7 from the eluted protein (Appendix 7.8 lane 7 and 8) from a 6.2 % VP7 in the supernatant (Appendix 7.8 lane 2).

These results indicate VP7 HisTag purification directly from the supernatant sample is unsuccessful (Figure 3.16b and c) with less pure target protein (20.3 % VP7) (Figure 3.16) (Appendix 7.8) compared to that from dual purification eluted sample (69.7 % VP7) (Figure 3.15) (Appendix 7.7).

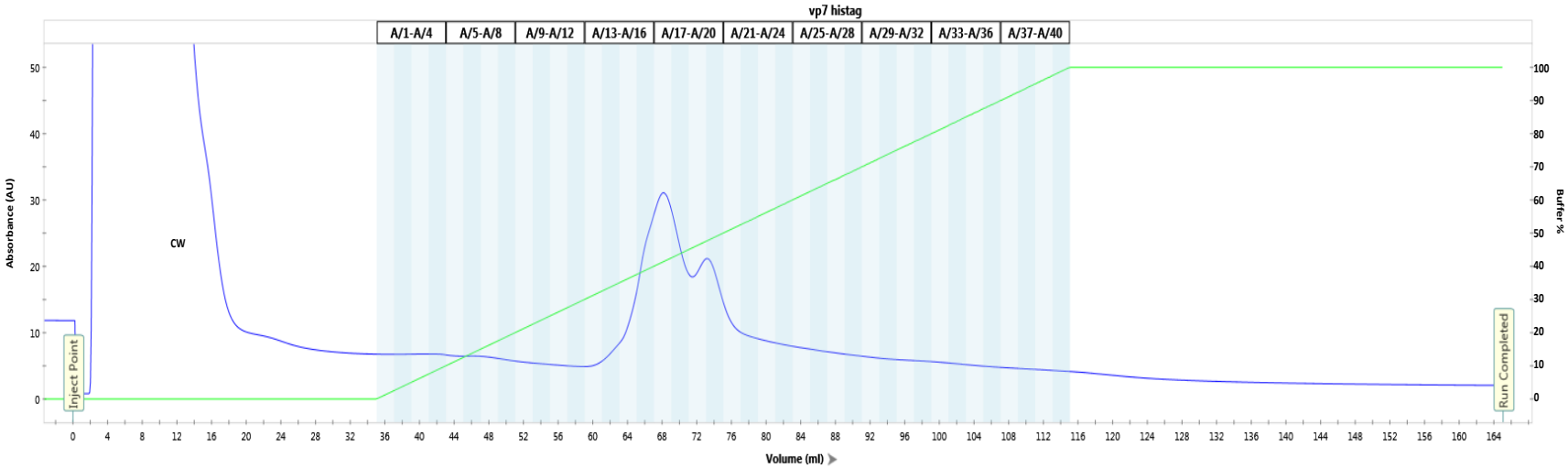
VP7 overexpression resulted in less soluble target protein (VP7) expressed in the supernatant (S) fractions as compared to the insoluble in the pelleted (P) fractions (Figure 3.5). With expression results indicating VP7 to be overexpressed more in insoluble form than soluble (Figure 3.5), solubilisation of VP7 from the insoluble fraction was conducted to obtain more VP7 protein (Figure 3.7) in a soluble state. The solubilised sample (Figure 3.7) was purified by affinity purification (Figure 3.17). VP7 solubilisation (Figure 3.7) resulted in exclusion of bacterial proteins found in the lysate supernatant (Figure 3.4 lane 3 and 5) which favours VP7 purification (Figure 3.17) from solubilised supernatant sample (Sol S) (Figure 3.7) by single purification of affinity chromatography resulting in 100 % pure VP7 (Figure 3.17d) (Appendix 7.9) as compared to VP7 purification by affinity chromatography from the bacterial lysate supernatant resulting in 20 % pure VP7 (Figure 3.16c). The eluted soluble fractions (Figure 3.17b) were concentrated (Figure 3.17c) using stirred ultrafiltration cell (Amicon bio separation, Millipore, Bedford USA), indicating pure VP7 in lane 11 (Figure 3.17d).

3.3 Structural characterisation

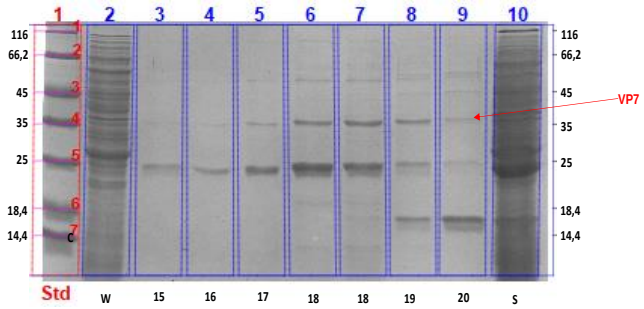
3.3.1 Absorbance and concentration

The purified solubilised fraction with over 95 % pure VP7 absorbance spectra was recorded between 225 and 360 nm wavelengths. The absorbance reading of 0.05 AU was recorded at 280 nm wavelength and the molar extinction coefficient of 70 290 M^{-1} cm^{-1} were used to calculate the proteins concentration of $7.41*10^{-7}$ M (0.013 mg from 1 litre of cell culture) (Figure 3.18) using the Beer-Lamberts law.

a



b



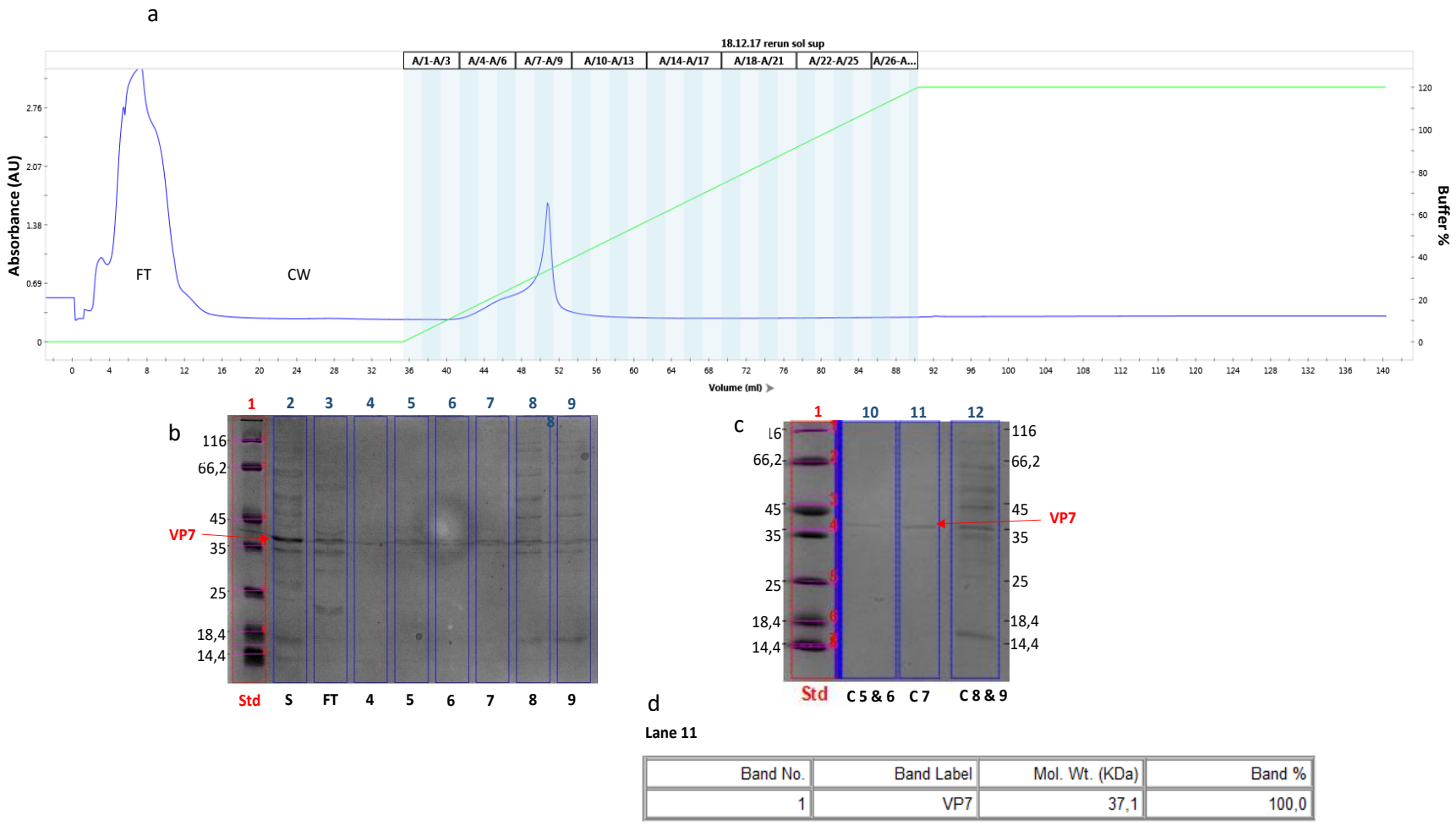
c

Lane 8

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		61,8	6,0
2		54,4	7,9
3		52,4	9,8
4		45,0	6,5
5	VP7	36,7	20,3
6		24,7	11,1
7		23,7	3,7
8		16,9	19,8
9		15,8	15,0

Figure 3.16: VP7 HisTag purification of supernatant

HisTag affinity purification profile (a) with the elution buffer percentage (green) and the absorbance (blue). SDS-PAGE (b) was loaded with supernatant (S), column wash (CW) and fractions 15 to 20 with increase in absorbance. The visible bands in lane 8 (b) was analysed by ChemiDoc™ MP Image Lab 6.0.1 software and calculated the approximate band sizes (c) in kDa in relation to the molecular weight maker (lane 1 (b)). Full information included in Appendices 7.8.



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	VP7	37,1	100,0

Figure 3.17: VP7 HisTag purification of solubilized sample

HisTag affinity purification profile (a) with the conductivity (Red), elution buffer percentage (green) and the absorbance (blue). SDS-PAGE (b) was loaded with solubilized sample (S), flow through (FT), and fractions 4 to 10 increase in absorbance, SDS-PAGE (c) indicates all the concentrated fractions (5 & 6; 7; and 8 & 9) with VP7 protein. Lane 11 band was analysed and calculated by the ChemiDoc™ MP Image Lab 6.0.1 (d) in relation to the molecular weight maker (lane 1 of figure b).

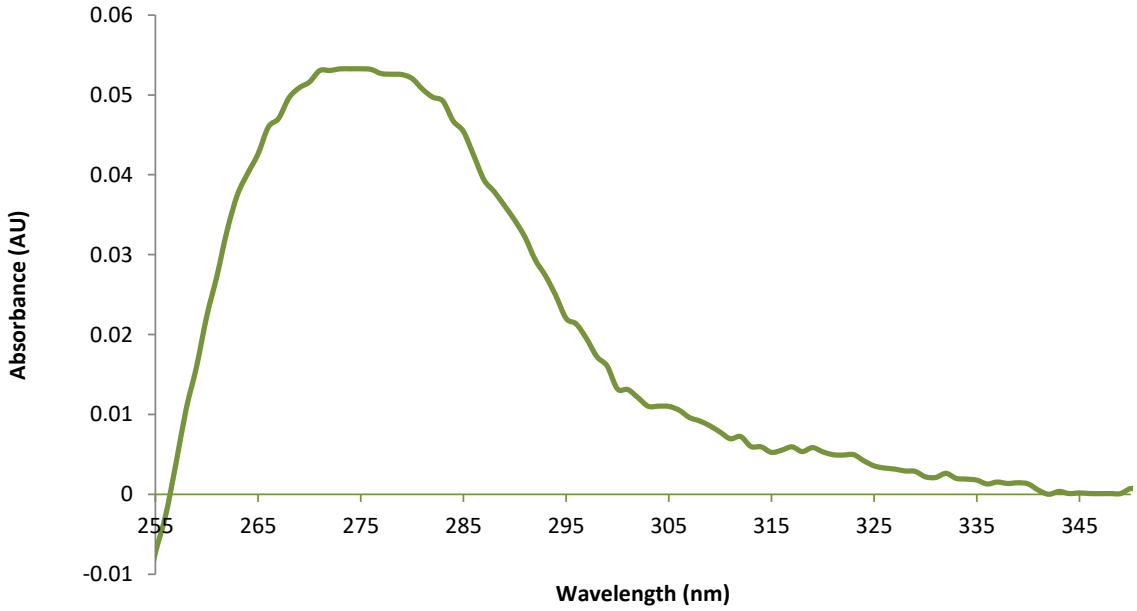


Figure 3.18: VP7 absorbance spectra
Chirascan Plus spectrometer absorbance spectra results of the purified VP7 sample in 50 mM sodium phosphate buffer pH 5.7 with 1 mM DTT at 280 nm wavelength. Spectrum results with buffer subtracted plotted using MS excel by Mathebula NT

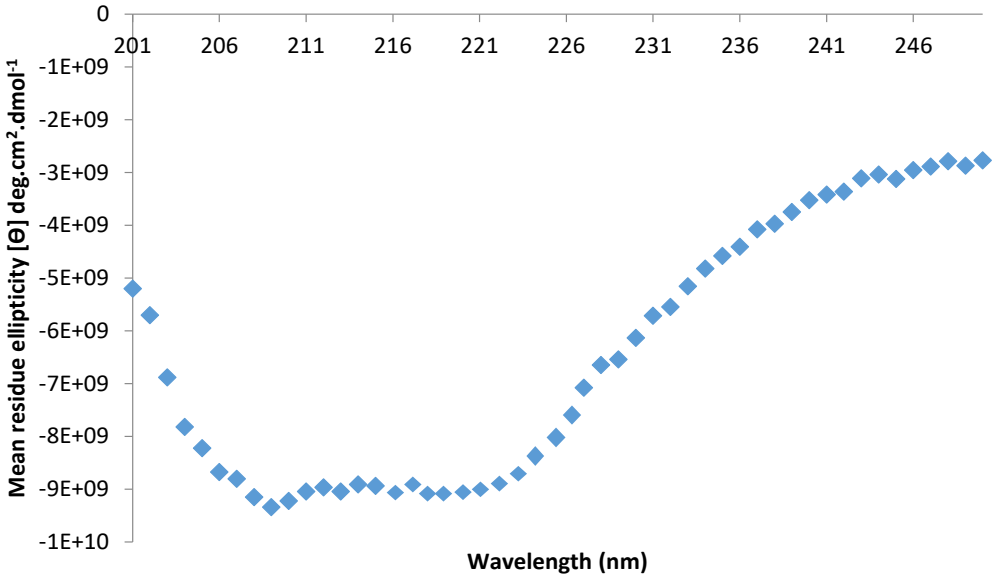
3.3.2 Far UV circular dichroism

VP7 far-UV spectrum was recorded between 250 and 200 nm wavelengths, to characterise its secondary structure. A far-UV CD spectrum for 100 % alpha helices consists of troughs at 208 and 222 nm and a peak at 218 nm, whereas the 100 % beta sheet proteins consists of a trough at 218 nm and an increasing ellipticity thereafter with increasing wavelength (Figure 1.5) (Sreerama and woody, 2004). The far UV CD spectrum for the dilute protein sample could not be measured below 200 nm as the voltage reading was above 600 HV (Figure 3.19b) indicating a high noise to signal ratio. Dithiothreitol (DTT) absorbs light below 220 nm wavelengths (Kelly *et al.,* 2005) which can explain the HV signal observed at 200 nm (Figure 3.19) that resulted in omitting the Far UV CD spectra results below 200 nm.

The recombinant VP7 secondary structure consists of 12 beta strands and 6 alpha helices which amounts to double the beta strands in respective to its alpha helices (figure 1.3a) (Dormiter *et al.,* 2001; Aoki *et al.,* 2009). The VP7 far UV CD spectra reading between 200 and 250 nm shows a trough at 208 nm wavelength, which indicates the presence of alpha helices. A beta sheet trough at 218 nm cancels the alpha helices peak, indicating the predominance of VP7 beta sheets over its alpha helices (Figure 3.19). Lower alpha helices peak at 218 nm cancelled the alpha helices trough by raising ellipticity reading at 222 nm, confirms the expected secondary structure of folded native VP7 protein.

Ellipticity readings at 208 nm are affected and based only on the alpha helices presence, however at 218 and 222 nm the beta sheets and alpha helices affect the ellipticity reading. Therefore, beta sheets being double the amount of the alpha helices (14:7), the beta sheets reduces the alpha helices reading (i.e 2:1 ratio which is similar as 1:0) causing the beta sheet reading at 212 nm and above to be detected over that of the alpha helices (Figure 3.19).

a



b

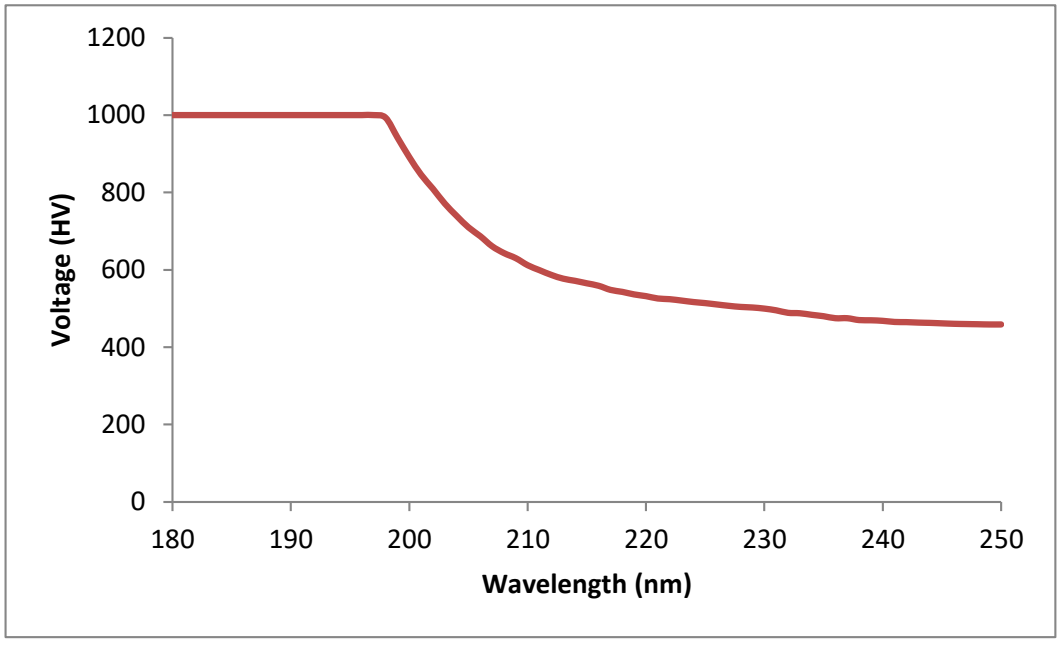


Figure 3.19: VP7 far UV CD and high tension voltage (HV) spectra

Far UV CD spectra (a) and High tension voltage (b) collected from a $7.40*10^{-7}$ M protein in 50 mM sodium phosphate buffer pH 5.7 with 1 mM DTT. Spectrum results plotted using MS excel by Mathebula NT

3.3.3 Intrinsic fluorescence spectroscopy

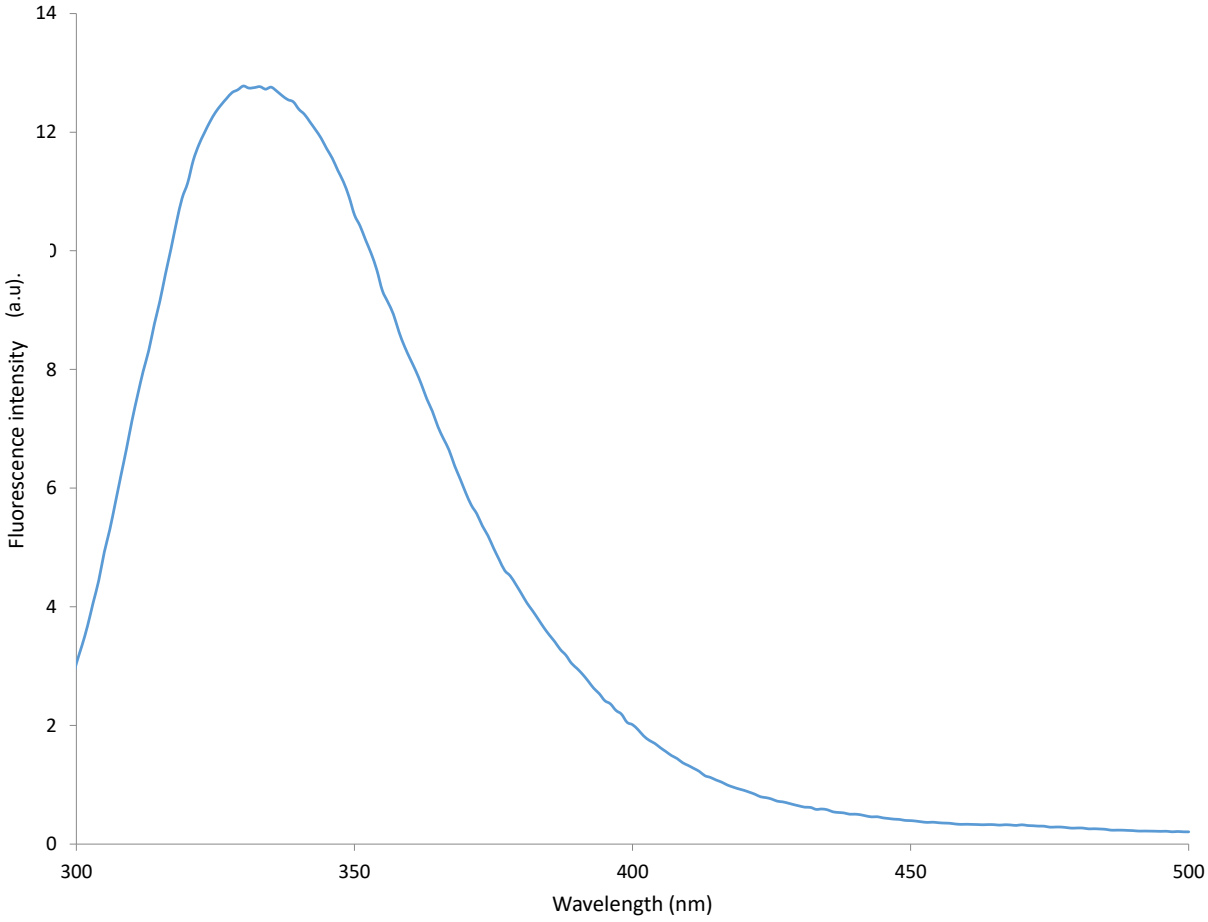
Intrinsic fluorescence emission spectrum of proteins can be used to research proteins tertiary structure and folding (Eftink, 1991). Proteins containing tryptophan residues have a maximum wavelength emission range from 308 to 355 nm (Vivian and Callis, 2001). Exposed tryptophan residues have a maximum emission wavelength range between 345 and 355 nm and the buried tryptophan residues around 308 and 340 nm. When the protein is unfolded a red shift is seen in the emission resulting in a peak between 345 and 355 nm (Eftink, 1991) upon excitation by 295 nm wavelength light. The fluorescence spectra can be recorded even for small amounts of protein (Alston et al. 2004). When proteins conformation changes, the emitted fluorescence energy can also get absorbed by a fluorophore which is excited by low energy light emitted by tryptophan residue, resulting in the decrease of fluorescence intensity (Lakowicz, 2006).

VP7 consists of 7 tryptophan residues buried within the protein structure (Figure 1.3a). Whether VP7 is in monomeric or trimetric form, the majority of tryptophan residues are found in the hydrophobic environment (Figure 1.3) which explains the observed maximum wavelength at 335 nm (Figure 3.20) confirming the hydrophobic position of VP7 tryptophan residues and the native folded tertiary structure of the purified protein.

3.3.4 Thermal denaturation

Proteins tend to rearrange their structural form with changing temperature. As proteins structure rearranges, the proteins chromophores positions change and therefore results in fluorescence emission spectrum differences. Tryptophan excited by light with 295 nm wavelength (high energy) emits low energy light with wavelengths between 308 and 355 nm wavelengths. Therefore, chromophores which absorb light with lower energy (with 300 nm wavelengths and above) will therefore absorb tryptophan emitted light. Certain chromophores might not be able to absorb emitted light because of their position in the proteins structure, however as the protein structure rearranges which can be caused by temperature changes, the chromophores position might change causing it to be able to absorb the emitted light and therefore reduce fluorescence emission intensity, this phenomenon is known as quenching (Rae *et al.,.,* 2003).

Temperature changes can also cause the protein backbone to unfold, exposing the buried (Hydrophobic) tryptophan residues to the solvent. When the hydrophobic tryptophan residues get exposed to the hydrophilic environment, due to unfolding, this can lead to a maximum wavelength shift from high energy between 308 and 340 nm wavelength range to 345 and 355 nm range causing a redshift (Eftink, 1991).



0.

Figure 3.20: VP7 fluorescence spectra

VP7 fluorescence spectra at 335 nm wavelength from $7.40*10^{-7}$ M protein in 50 mM sodium phosphate buffer pH 5.7 with 1 mM DTT. VP7 tryptophan residues excited by 295 nm wavelength. Spectrum results plotted using MS excel by Mathebula NT.

Temperature ramping resulted in VP7 conformational changes, observed with protein fluorescence intensity quenching at 335 nm wavelength (Figure 3.21a and b), indicating VP7 structure conformation was altered by rearrangement when temperature increased, however the temperature did not cause protein unfolding as a redshift was not observed. VP7 conformational change was observed from 24 °C after temperature increase from 20 °C, the conformational change occurred at a constant rate until it reached 48 °C after which minor changes occurred until final temperature 90 °C (Figure 3.21b). The fluorescence intensity at 295 nm wavelength indicates that temperature increase from 20 to 48 °C could results in the dissociation of oligomeric protein structures in solution, however, further temperature increase from 48 °C resulted in aggregate formations which cause 295 nm light scattering (Figure 3.21c). The increase in temperature to 90 °C resulted in VP7 irreversible conformational change, even after it was stored at 20 °C (Figure 3.22). These results indicate that an increase in temperature has an effect on VP7 structural conformational state. A decreasing temperature effect on VP7 structure study needs to be conducted to determine which temperature favours a stable VP7 structure, as we only focused on increasing temperature, however, as the temperature cooled down back to 20 °C, there was no structural changes observed.

a

Fig re 3.21: VP7 fluorescence thermal stability spectra

V conformation changed with increased temperature from 20 °C to 90 °C indicated by a decrease in fl escence intensity at 335 nm wavelength maxima. The fluorescence spectrum (a) of tryptophan residues e ed by 295 nm wavelength at different temperatures (20; 25; 31;36; 40; 46 and 90 °C) and the fluorescence i sity trend at 335 nm wavelength (b) as temperature increases. (c) Different VP7 fluorescence spectra at 295 n avelength in increasing temperature. Spectrum results plotted using MS excel by Mathebula NT.

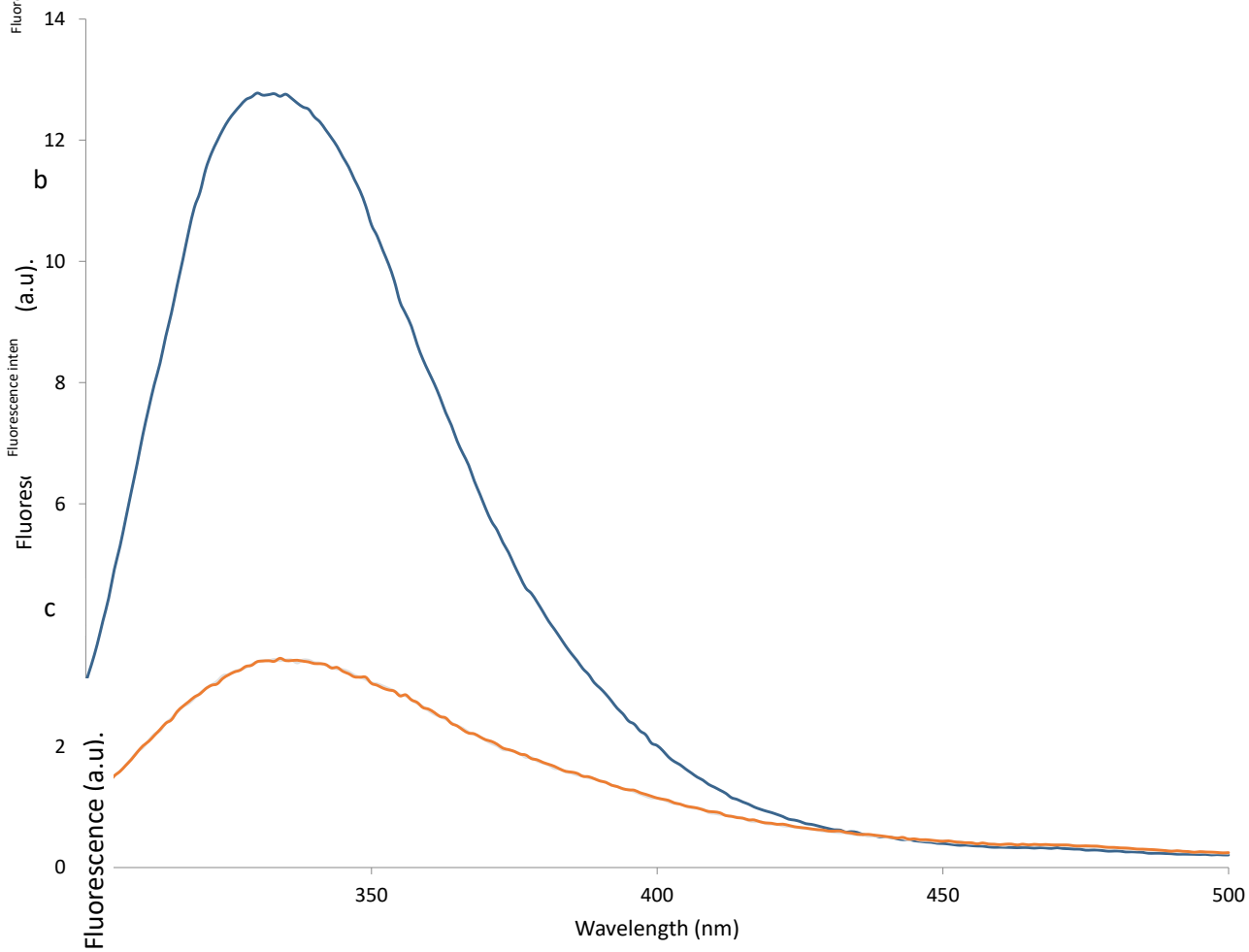


Figure 3.22: VP7 fluorescence int nal ramp

Temperature (°C)

Fluorescence emission spectra re ure of VP7 pre- and post-thermal ramp after sample fluorophore was excited with 295 nm wavelength. Blue curve represents fluorescence intensity before temperature ramp and orange curve represents after temperature ramp of samples at 20 °C Spectrum results plotted using MS excel by Mathebula NT.

4. Discussion

Previous studies indicate that rotavirus VP7 has been predominantly overexpressed using other expression systems besides microbial/bacterial systems such as plant expression system for vaccination purposes and infection of mammalian and insects cells with rotavirus (Rigano and Walmsley, 2005; Mellado *et al.,* 2008), which indicates possible challenges and limited ability to overexpress VP7 using microbial systems.

Rotavirus VP7 has been expressed using other protein expression techniques (Ruiz *et al.,* 2000; Trask and Dormitzer, 2006) which made us question why the bacterial system has not been favoured regardless of its advantages. Deng and colleagues (2015) recently conducted an overexpression of rotavirus VP7 in TOP 10 bacterial cells. Bacterial cell lines (BL21 (DE3) and C43 (DE3) pLysS) were selected based on BL21 (DE3) as the universal host choice for overexpression and C43 (DE3) pLysS ability to tolerate toxic level of recombinant protein. C43 (DE3) pLysS cells (also known as Walker strain) is a bacterial strain derived from the BL21 (DE3) strains with enhanced ability for high expression of heterologous toxic recombinant membrane proteins (Ren *et al.,* 2009; Kopanic, J.L. *et al.,* 2013; Jia, B. and Jeon C.O., 2016) compared to BL21 (DE3). C43 (DE3) strain are said to be superior to BL21 (DE3) strain as a host for protein overexpression because it is a double mutant (a mutant from C41 (DE3) which is a BL21 (DE3) strain mutant) that can overexpress and tolerate membrane and globular proteins in inclusion bodies while limiting and preventing bacterial cell death (Miroux and Walker, 1996).

Bacterial cell lines were transformed with a plasmid (pET28a+) engineered with VP7 gene inserted for VP7 overexpression. C43 (DE3) pLysS showed traces of low soluble VP7 expressed, suggesting that C43 (DE3) pLysS cells can produce soluble protein, which can be caused by its tolerance for membrane proteins (Miroux and Walker, 1996) as a mature VP7 is a membrane associated protein (Dormitzer, *et al.,* 2001). The disturbance of insoluble pellet proteins by sonication has resulted in obtaining small fractions of soluble proteins (Singh et al. 1990; Rodríguez-Carmona *et al.,* 2010) which can result in solubilisation of the particles. The expression results also indicate VP7 expression level to slightly differ between both cells lines, favouring the hypothesis that C43 (DE3) pLysS tolerance of toxic levels of recombinant protein allows higher expression levels of rotavirus VP7 proteins (Sorensen *et al.,* 2003). Rotavirus VP7 contains disulphide bonds, which play a role in protein folding and solubility. Therefore, could the presence of disulphide bond presence (Siurkus and Neubauer, 2011), results in formation of misfolded proteins and inclusion bodies in bacterial cells, hence affecting VP7 solubility.

The insoluble form of expressed target proteins are the results of specific interactions and aggregation between partially folded polypeptides (Kopito, 2000; Villaverde and Carrio, 2003), resulting in cells response to stress by forming inclusion bodies (Bowden *et al.,* 1991; Waelter *et al.,* 2001). Inclusion body formation for recombinant proteins results in insoluble protein production caused by expression conditions such as host strain, inducer

concentration, induction period and environment temperature during induction (Chalmer *et al.,* 1990; Van den Berg *et al.,* 1999)

An IPTG concentration of 0.5 mM showed traces of soluble VP7 expression observed in C43 (DE3) pLysS cells, whereas BL21 (DE3) cells expression resulted in insoluble VP7. The optimal induction period of 20 hours was selected as it resulted in optimal cell growth, resulting in high concentration of VP7 overexpression. In both cell lines, VP7 protein was highly overexpressed in insoluble fractions by 0.5 mM IPTG concentration at 37 °C for 20 hours. The sonication of the insoluble fractions resulted in obtaining very low soluble VP7 as observed with Ortwerth and Sharma (1986). The challenges of not overexpressing more soluble VP7 protein, agrees with Deng and colleagues (2015) statement that in *E.coli* cells, VP7 is unstable and toxic.

The use of urea as a denaturant to solubilise (Njengele *et al.,* 2016; Vargas-Cortez *et al.,* 2016) the insoluble VP7 from the pellet was conducted which resulted in obtaining over 95 % purity of soluble VP7 after affinity purification. The solubilisation of proteins from inclusion bodies resulted in soluble VP7 protein with few bacterial proteins as compared from the supernatant. These results indicate and confirm that solubilisation of insoluble protein can be achieved, and facilitates target protein purification.

The unsuccessful VP7 separation from bacterial protein by salt gradient elution, led to protein purification based on protein charge differences caused by change in buffer pH in respective to the proteins pI values (Ahamed *et al.,* 2007). However, the ion exchange pH change elution resulted in unsuccessful VP7 purifications. The results indicate that there are *Escherichia coli* proteins with pI values between pI 4.8 and 4.6 as mentioned by Bolanos-Garcia and Davies (2006) and Kroner and Hubbuch (2013) which share ionic properties with VP7 hence resulting in an unsuccessful VP7 purification. The unsuccessful purification of VP7 based on its charge could be the reason why VP7 was always not purified by ion exchange chromatography.

The recombinant VP7 expressed in this study contains a HisTag and therefore can be purified by affinity chromatography. Ion exchange chromatography alone did not result in a successful purification; therefore, a dual purification (Alibolandi and Mirzahoseini 2011; Qi *et al.,* 2001) of ion exchange and affinity purification was conducted. The dual purification resulted in 69.7 % pure VP7 suggesting that other bacterial protein with similar charge as VP7, contain histidine residues (Bolanos-Garcia and Davies, 2006) which binds the nickel ions during affinity purification as VP7 does (Figure 3.15 and 3.16) and thus, affect VP7 purity. Soluble VP7 in the supernatant fraction was purified by affinity chromatography; however, the purification was also unsuccessful (Figure 3.16) (Appendix 7.8), which confirms that C43 (DE3) pLysS bacterial cells contains proteins in the soluble fraction with histidine residues that bind the nickel ions during purification (Bolanos-Garcia and Davies, 2006). Dormitzer and colleagues (2001) success in a single specific binding (Affinity) purification resulted in 95% pure rhesus rotavirus

VP7. These results indicate that some bacterial proteins contain histidine residues as mentioned by Bolanos-Garcia and Davies (2006) and do bind the metal ions and thus affect VP7 purification. Some bacterial protein required 150 mM imidazole concentration (Figure 3.16), which is more than what is required for VP7 elution (Figure 3.17), shows that the bacterial proteins containing histidine residues had more binding affinity for the column than VP7 (Figure 3.16). The solubilised VP7 from the inclusion bodies only required affinity purification (Figure3.17) to obtain over 95 % VP7 purity. The results (Figure 3.17) indicate that, the bacterial proteins with an approximate 20 to 25, 45 and 55 to 58 kDa (Figure 3.16) which contain histidine residues (Bolanos-Garcia and Davies, 2006) that bind the nickel ion during affinity chromatography (Figure 3.15 and 3.16) are not in the solubilised sample to hinder VP7 purification (Figure 3.17). This indicates that the formation of inclusion bodies is not just a setback for VP7 purification from bacterial protein, however, an alternative route to obtain pure proteins.

These results also indicate that the microbial expression of VP7 results in less lower VP7 yield (0.013 mg per 1 litre of C43 (DE3) pLysS bacterial cells) as compared to 0.25 to 0.5 mg VP7 per litre of insects' cells obtained by Domitzer (2001). Thus suggesting that VP7 is optimally expressed in insect cells over the two selected bacterial cells in our project indicating that the over expression condition and cell lines we used results in low yield VP7 overexpression when using the microbial systems (Terpe, 2006), at low cost (Andersen *et al.,* 2002) as said and therefore, confirms Deng *et al.* (2015) statement that "VP7 gene sequence has underlying promoters that affect its expression in *E. coli*". All affinity purification showed a consistent elution of VP7 by imidazole concentration below 150 mM (Figure 3.15; 3.16 and 3.17) indicating conditions to purify VP7.

The obtained CD spectrum is similar to the derived CD spectrum of VP7 expressed in insects' cells published by Dormiter *et al.,* (2001). The CD spectrum shows one trough with the lowest ellipticity reading at 210 nm wavelength which gradually increases until 230 nm wavelength. Dormitzer *et al.,* (2001) CD spectra results also indicate and confirms the hypothesis, that the beta sheets reading at 212 nm and above to be detected over that of the alpha helices resulting in the observed trough between 200 and 250 nm wavelength reducing the alpha helices peak at 218 nm due to the beta sheet and alpha helices ratio.

The fluorescence intensity change indicates the change in VP7 tertiary structure as described by Pando *et al.* (2002) that the chromophore that absorbs low energy light rearrange to a position in which they are able to absorb light emitted by tryptophan residues (Lackowics, 1999) during conformational change as compared to its initial detected conformation (Figure 3.21). VP7 conformational changes were observed immediately as the temperature changed which suggests that temperature has an effect (Figure 3.21b). However, conformational changes might be prevented by the presence of calcium (Emslie *et al.,* 1998; Dormitzer et al. 2001) and DTT over

time (Svensson et al. 1994; Mirazima and Svensson, 1998). The results suggest that VP7 conformation changes as mentioned by Deng *et al.* (2015) could have affected VP7 expression and purification.

5. Conclusion

Recombinant protein overexpression in bacterial cells is economically inexpensive and results in high yield protein expression compared to other expression systems. Rotavirus VP7 was optimally expressed in BL21 (DE3) and C43 (DE3) pLysS when induced with 0.5 mM IPTG for 20 hours at 37 °C incubation, however, C43 (DE3) pLysS showed traces of soluble VP7. Soluble VP7 purification in supernatant by dual purification of ion exchange chromatography and affinity chromatography was unsuccessful due to bacterial proteins sharing properties with recombinant VP7. However, the urea solubilised VP7 from inclusion bodies, resulted in over 95% VP7 purification by affinity chromatography alone, suggesting inclusion body formation favour purification of VP7. VP7 overexpression in C43 (DE3) pLysS, inclusion body solubilisation and purification resulted in 0.013 mg rotavirus VP7 from 1 litre of bacterial cell culture. The purified protein has identical secondary structure (Figure 3.19a) as identified by Dormiter *et al.,* (2001). In future, VP7 overexpression in other bacterial cell lines can be conducted and compare the purity difference between VP7 extracted from inclusion bodies and cells supernatant.

6. References

Ahamed, T., Nfor, B.K., Verhaert. T., van Dedem, G.W.K., van der Wielem, L.A.M., Eppink, M.H.M.,van der Sandt, E.J.A.X and Ottens, M., 2007. pH gradient ion-exchange chromatography: An analytical tool for design and optimization of protein separation. Journal of Chromatography A, 1164 (1-2), 181-188

Ahn, O.K., Lim, S.W., Yang, H.J., Li, C., Sugawara, A., Ito, S., Choi, B.S., Kim, Y.S., Kin, J., Yang, C.W., 2007. Induction of PPAR gamma mRNA and protein expression by rosiglitazone in chronic cyclosporine nephropathy in the rat. National Library of Medicine, 48(2):308-316

Ahmad, I., Nawaz, N., Darwesh, N.M., Raman, S., Mustafa, M.Z., Khan, S.B. and Patching, S.G., 2018. Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. Protein Expression and Purification, 144, 12-18

Alibolandi, M. and Mirzahoseini, H., 2011. Purification and refolding of overexpressed human basic fibroblast growth factor in *Escherichia coli*. Biotechnology Research International, 973741, 1-6

Alston, R.W., Urbanikova, L., Sevcik, J., Lasagna, M., Reinhart, G.D., Scholtz, J.M. and Pace, C.N., 2004. Contribution of single tryptohan residue to the fluorescence and stability of ribonuclease Sa. Biophysical Journal, 87 (6), 4036-4047

Amimo, J.O., Vlasova,A.N. and Saif, L.J., 2014. Prevalence and genetic heterogeneity of porcine group C rotaviruses in nursing and weaned piglets in Ohio, USA and identification of a potential new VP4 genotype. Veterinary Microbiology NIH Public Access, 164(0), 27–38.

Amir, M., Dar, A.M., Haque, A., Islam, A., Ahmed, F. and Hassan, I., 2016. Purification and characterization of Ras related protein, Rab5a from Tinospora cordifolia. International Journal of Biological Macromolecules, 82, 471–479.

Andersen, D. C. and Krummen, L., (2002). Recombinant protein expression for therapeutic applications. Current Opinion in Biotechnology, 13(2), 117–23.

Angel, J., Franco, M.A. and Greenberg, H.B., (2007). Rotavirus vaccines: recent developments and future considerations. Nature Reviews. Microbiology, 5(7), 529–539.

Aoki, S.T., Settembre, E.C., Trask, S.D., Greenberg, H.B., Dormitzer, P.R. and Harrison, S.C., 2009. Structure of rotavirus outer-layer protein VP7 bound with a neutralizing Fab.Science, 324, 1444-1447.

Aoki, S.T., Trask, S.D., Coulson, B.S., Greenberg, H.B., Dormitzer, P.R. and Harrison, S.C., 2011. Cross-linking of rotavirus outer capsid protein VP7 by antibodies or disulfides inhibits viral entry. Journal of virology, 85(20), 10509–10517.

Attoui, H., Jaafar, F.M., Belhouchet, M., De Micco, P., De Lamballerie, X. and Brussaard, P.D., 2006. *Micromonas pusilla* reovirus: a new member of the family reoviridae assigned to a novel proposed genus (Mimoreovirus). Journal of General Virology. 87, 1375–1383

Barron, J.J., Ashton, C. and Geary, L., 2006. The effects of temperature on pH measurement. International society of electrochemistry, 1-7

Belouzard, S., Millet, J.K., Licitra, B.N., and Whittaker, G.R., 2012. Mechanism of coronavirus cell entry mediated by the viral spike prtein. Viruses. 4, 1011-1033

Bennion, B.J. & Daggett, V., 2003. The molecular basis for the chemical denaturation of proteins by urea. , PNAS, 100(9), 5142–5147

Bhowmick, R., Banik, G., Chanda, S., Chattopadhyay, S. and Chawla-Sarkar, M., 2014. Rotavirus infection induces G1 to S phase transitionin MA104 cells via Caþ2/Calmodulin pathway. Virology. 454-455, 270–279

Bolanos-Garcia, V.M. and Davies, O.R., 2006. Structural analysis and classification of native proteins from *E.coli* commonly co-purified by immobilised metal affinity chromatography. Biochemica et Biophysiaca Acta. 1760, 1304-1313

Bowden, G.A., Peredes, A.M. and Georgiou, G., 1991. Structure and morphology of protein inclusion bodies in Escherichia coli. Nature, 9(8), 725-730.

Brahms, S., Brahms, J., Spach, G., and Brack, A., Identification of #,3-turns and unordered conformations in polypeptide chains by vacuum ultraviolet circular dichroism. PNAS. 74(8), 3208-3212,.

Carreño-Torres, J.J., Gutiérrez, M., Arias, C.F., López, S. and Isa, P., 2010. Characterization of viroplasm formation during the early stages of rotavirus infection. Virology Journal, 7(350), 1-11

Carrió, M. M. and Villaverde, A. (2001) Protein aggregation as bacterial inclusion bodies is reversible, FEBS Letters, 489(1), 29–33

Carrio´, M.M. and Villaverde A., 2002. Construction and deconstruction of bacterial inclusion bodies. Journal of Biotechnology. 96, 3–12

Carrió, M. M. and Villaverde, A. (2003) Role of molecular chaperones in inclusion body formation, FEBS Letters, 537(1–3), 215–221.

Carter, S., Rimmer, S., Rutkaite,R., Swanson, L., Fairclough, J.P., Sturdy, S. and Webb, M., 2006. Highly Branched Poly (N-isopropylacrylamide) for Use in Protein Purification, 1124–1130.

Chalmer, J., Kim, E., Telford, J., Wong, E., Tacon, W., Shuler, M. and Wilson, D., 1990. Effects of temperature on Escherichia coli overproducing β-lactamase or human epidermal growth factor. Applied and Environmental Microbiology, 56, 104-111

Chen, R., 2012. Bacterial expression systems for recombinant protein production: *E. coli* and beyond. Biotechnology Advances, 30(5), 1102–1107.

Chenal, A., Vernier, G., Savarin, P., Brushmarina, N.A., Geze, A., Guillain, F., Gillet, D. and Forge, V., 2005. Conformational states and thermodynamics of alpha-lactalbumin bound to smembranes: a case study of the effects of pH, calcium, lipid membrane curvature and charge. Journal of molecular biology, 349(4), 890–905.

Cheung, W., Gill, M., Esposito, A., Kaminski, C.F., Courousse, N., Chwetxoff, S., Trugnab, G., Keshavan, N., Lever, A. and Desselgerger, U., 2010. Rotaviruses associate with cellular ipid droplet components to replicate in viroplasms, and compounds disrupting or blocking lipid droplets Inhibit viroplasm formation and viral replication. Journal of Virology, 84, 6782-6798

Chia, G., Ho, J.H., Ng, C., Neo, F. J., Win, M., Gui, L., Leo, Y. and Chow, A., 2018. An unusual outbreak of rotavirus G8P[8] gastroenteritis in adults in an urban community, Singapore, 2016. Journal of Clinical Virology. 105, 57-63

Corti, F. and Simons, M., 2016. Modulation of VEGF receptor 2 signaling by protein phosphatases. Pharmacological research, 115,107-123

Day, R., Bennion, B.J., Ham, S. and Daggett, V., 2002. Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. Journal of Molecular Biology, 322(1), 189–203.

Deng, X., Xiong, G., Cong, W., Liu, Z., QI, C. and Yang, J., 2015. Monoclonal antibodies directed against VP7 protein of human group B rotavirus. Mary Ann Liebert, 33(1), 38-41.

Dennehy, P.H., 2008. Rotavirus vaccines: an overview. American Society for Microbiology. 21, 198–208

Desselberger, U., 2014. Rotaviruses. Virus Research. 190, 75–96

Dormitzer, P. R., Greenberg, H.B. and Harrison, S.C., 2001. Proteolysis of monomeric recombinant rotavirus VP4 yields an yligomeric VP5* core. Journal of virology. 75(16), 7339-7350

Dormitzer, P.R., Greenberg, H.B. and Harrison, S.C., 2001. Purified recombinant rotavirus VP7 forms soluble, calcium-dependent trimers. Virology 277, 420-428

Eftink, M.R., 1991. Fluorescence technique for studying protein structure. Methods of Biochemical Analysis 35, 127-205

Emslie, K.R., Coukell, B., Birch, D. and Williams, K.L., 1996. Calcium influences the stability and conformation of rotavirus SAl 1 glycoprotein VP7 expressed in Dictyostelium discoideum. Journal of Biotechnology, 50, 149-159

Fan, X. C., & Steitz, J. A., 1998. Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs. EMBO Journal, 17(12), 3448–3460.

Gleba, Y., Klimyuk, V. and Marillonnet, S., 2007. Viral vectors for the expression of proteins in plants. Current Opinion in Biotechnology, 18,134–141

Hilbrig, F. and Freitag, R., 2003. Protein purification by affinity precipitation. Journal of Chromatography B, 790, 79–90

Hirsh, A. and Tsonev, L.I., 2008. Application of well-controlled pH gradient at variable isoratic salt concentration to IEX chromatography. American Biotechnology Laboratories, 1-3

Hong, P., Koza, S. and Bouvier, E.S., 2012. Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. Journal of Liquid Chromatography & Related Technologies, 35(20), 2923–2950.

Hoshino, Y. and Kapikian, A.Z., 2000. Rotavirus serotypes: slassification and importance in epidemiology, immunity, and vaccine development. Journal of Health Population Nutrition. 18(1), 5-14

Hsu, G.G., Bellamy, A.R. and Yeager, M., 1997. Projection structure of VP6, the rotavirus inner capsid protein, and comparison with bluetongue VP7. Journal of Molecular Biology, 272(3), 362–368.

Hu, L., Crawford, S.E., Czako, R., Cortes-Penfield, N.W., Smith, D.F., Pendu, J.L., Estes, M.K. and Prasad, B.V., 2012. Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. Nature, 485(7397), 256–259.

Hulo, C., De Castro, E., Masson, P.,Bougueleret, L., Bairoch, A., Xenarios, I. and Mercier, P., 2010. ViralZone: a knowledge resource to understand virus diversity. Nucleic Acids Research. 39, D576–D582

Ianiro, G., Delogu, R. amd Fiore, L., 2016. Genetic variability of VP7, VP4, VP6 and NSP4 genes of commonhuman G1P[8] rotavirus strains circulating in Italy between 2010 and 2014. Virus Research. 220, 117–128

Irvine, G. B. 1997. Size-exclusion high-performance liquid chromatography a review of peptides : a review. Analytica Chimica Acta, 352, 387–397.

Jia, B and Jeon, C. O., 2016. High-throughout recombinant protein expression in Escherichia coli: current status and future perspectives. Open Biology. 6(160196), 1-17

Kelly, S.M., Jess, T.J. and Price, N.C., 2005. How to study proteins by circular dichroism. Biochimica et Biophysica Acta - Proteins and Proteomics, 1751(2), 119–139.

Kesler, R. B. 1953. Rapid Quantitative Anion-Exchange Chromatography. Analytical Chemistry, 74(2090), 1416–1422.

Khetawat, D., Chosh, T., Bhattacharya, M.K., Bhattacharya, S.K. and Chakrabarti, S., 2001. Molecular characterization of the VP7 gene of Rotavirus isolated from a clinical sample of Calcutta, India. Virus Research, 74(1–2), 53–58.

Kroner, F. and Hubbuch, J., 2013.systematic generation of buffer systems for pH gradient ion exchange chromatography and their application. Journal of Chromatography A, 1285, 78-87

Kobayashi, N., Ishino, M., Wang, Y.H., Chawla-Sarkar, M., Krishnan, T. and Naik, T.N., 2007. Diversity of G-type and P-type of human and animal rotaviruses and its genetic background. Communicating Current Research and Educational Topics and Trends in Applied Microbiology, 847–858.

Kopanic, J.L., Al-Mugotir, M., Zach, S., Das, S., Grosley, R and Sorgen, P.L., 2013. An *Escharichia coli* strain for expression of the connexin45 carboxyl terminus attached to the transmembrane doamin. Frontiers in Pharmacology, 4(106), 1-8

Kopito, R.R., 2000. Aggresomes, inclusion bodies and protein aggregation. Cell Biology,10, 524-530

Laemmli U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259), 680–685

Lakowicz J. R., 2006. Principles of Fluorescence Spectroscopy 3rd edition, Springer, 954

Lebendiker, M., and Danieli, T., 2014. Production of prone-to-aggregate proteins. FEBS Press, 588(2), 236-246

Liu, P.F., Avramova, L.V. and Park, C., 2009. Revisiting absorbance at 230 nm as a protein unfolding probe. Analytical Biochemistry, 389(2), 165-170

López-García, P., 2012. The place of viruses in biology in light of the metabolismversus-replication-first debate. Stazione Zoologica Anton Dohrn. 34, 391-406

Maggioni, A., Hadley, B., Von Itzstein, M. and Tiralongo, J., 2014. Expression, solubilisation, and purification of a functional CMP-sialic acid transporter in Pichia pastoris, Protein Expression and Purification. Elsevier Inc., 101, 165–171.

Mallam, A.L. & Jackson, S.E., 2005. Folding studies on a knotted protein. Journal of Molecular Biology, 346(5), 1409–1421.

Marthaler, D., Suzuki, T., Rossow, K., Culhane, M., Coliins, J., Goyal, S., Tsunemitsu, H., Ciarlet, M. and Matthijinssens, J., 2014. VP6 genetic diversity, reassortment, intragenic recombination and classification of rotavirus B in American and Japanese pigs. Veterinary Microbiology. 172, 359–366

Martin, D., Ouldali, M.M. and Poncet, D., 2011. Structural organisation of the rotavirus nonstructural protein NSP5. Journal of Molecular Biology. 413(1), 209–221.

McDonald, S.M. & Patton, J.T., 2008. Molecular characterization of a subgroup specificity associated with the rotavirus inner capsid protein VP2. Journal of virology, 82(6), 2752–2764.

Mellado, M. M., Peixoto, C., Cruz, P.E., Carrondo, M. T. and Alves, P.M., 2008. Purification of recombinant rotavirus VP7 glycoprotein for the study of in vitro rotavirus-like particles assembly. Journal of Chromatography B. 874, 89–94

Mena, J. A., Ramirez, O.T. and Palomares, L.A., 2005. Quantification of rotavirus particles by gel permeation chromatography. Journal of Chromatography B, 824 (1-2), 267-276

Miles, A. J. and Wallace, B.A., 2016. Circular dichrism spectroscopy of membrane proteins. Royal society of chemistry.45,4859-4872

Mirazimi, A. and Svensson, L. 1998. Carbohydrates facilitate correct disulfide bond formation and folding of rotavirus VP7.American Society fot Microbiology. 3887-3892

Mirazimi, A. and Svensson, L. 2000. ATP is required for correct folding and disulfide bond formation of rotavirus VP7. Journal of Virology, 74(17), 8048–8052.

Miroux, B. and Walker, J.E., 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of Molecular Biology, 260, 289-298

Nallamsetty, S., Austin, B.P., Penrose, K.J. and Waugh, D.S., 2005. Gateway vectors for the production of combinatorially-tagged His6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*. Protein Science : A Publication of the Protein Society, 14(12), 2964–2971.

Njengele, Z., Kleynhans, R., Sayed, Y. and Mosebi, S., 2016. Expression, purification and characterization of a full-length recombinant HIV-1 *Vpu* from inclusion bodies. Protein Expression and Purification. 128, 109-114

Orthwerth, B.J. and Sharma, K.K., 1986. Solubilization of the lens water-insoluble fraction by sonication. Experimental Eye Research, 43 (6), 955-963

Pando, V., Isa, P., Arias, C.F. and Lopez, S., 2002. Influence of calcium on the early steps of rotavirus infection. Virology, 295, 190–200

Parashar, U. D., Gibson, C.J., Bresee, J.S. and Blass, R.I., 2006. Rotavirus and severe childhood diarrhea, 12(2), 304–306.

Park, K., Perczel, A., and Fasman, G.D., 1992. Differentiation between transmembrane helices and peripheral helices by the deconvolution of circular dichroism spectra of membrane proteins. Protein Science, 1032-1049

Pereira de Albuquerque, L., Leite de Siqueira Patriota, L., Gonzatto, V., Pontual, E.V.,Paiva, P.M.G., and Napoleao,T.H,. 2020 Coronavirus spikes (S) protein: A brief review on structure-function relationship, host receptors and role in cell function. Advance in Research, 21(9): 116-124

Pergande, M.R. and Cologna, S.M., 2017. Isoelectric point separations of peptides and proteins. Proteomes. 5(4), 1-14

Pesavento, J.B., 2006. Rotavirus proteins: structure and assembly. Springer-Verlag Berlin Heidelberg 309, 189–219

QIAexpressionist, 2003. A handbook for high-level expresson and purification of 6xHistagged protein. QIAGEN, 1024473, 80.

Rae, M., Fedorov, A., and Berberan-Santos, M. N., 2003. Fluorescence quenching with exponential distance dependence: Application to the external heavy-atom effect, Journal of Chemical Physics. 119, 2223–2231

Ramani, S., Hu, L., Prasad, B.V. and Estes, M.K., 2016. Diversity in rotavirus–host glycan interactions:A “sweet” spectrum. Cellular and Molecular Gastroenterology and Hepatology. 2 (3), 263-273

Rcsb.org/structure/3FMG

Ren, H., Yu, D., Ge, B., Cook, B., Xu, Z. and Zhang, S., 2009.High-level production, solubilization and purification of synthetic human GPCR chemokine receptors CCr5,CCr3,CXCR4 and CX3CR1. PLoS One, 4(2),1-15

Rigano, M.M. and Walmsley, A.M. 2005. “Expression systems and developments in plant-made vaccines,” (February), 271–277.

Rodríguez-Carmona, E., Cano-Garrido, O., Seras-Franzoso, J., Villaverde, A. and Garcia-Fruitos, E., 2010. Isolation of cell-free bacterial inclusion bodies. Microbial Cell Factories. 9 (71), 1-9

Rosano, G.L. and Ceccarelli, E.A., 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. Frontiers in Microbiology, 5(172), 1-17

Ruiz, M.C., Cohen, J., and .Michelangeli, F., 2000. Role of Ca2+in the replication and pathogenesis of rotavirus and other viral infections. Cell Calcium. 28(3), 137-149

Schuck, P., Taraporewala, Z., McPhie, P. and Patton, J.T., 2001. Rotavirus nonstructural protein NSP2 self-assembles into octamers that undergo ligand-induced conformational changes. The Journal of biological chemistry, 276(13), 9679–9687.

Shan, L. and Anderson, D.J., 2000. Effect of buffer concentration on gradient chromatofocusing performance separating proteins on a high perforemance DEAE column. Journal of Chromatography A, 909, 191-205

Singh, N.K., Donovan, G.R., Batey, I.L and MacRitchie, F., 1990. Use of sonication and size exclusion high-performance liquid chromatography in the study of wheat fluor protein. I. Dissolation of total proteins in th absence of reducingagents. American Association of Cereal Chemists, 67(2), 150-161

Siurkus, J. and Neubauer, P., 2011. Reducing conditions are the key for efficient production of active ribonuclease inhibitor in Escherichia coli. Microbial Cell Factories, 10(31), 1-15

Sluyterman, L.A.AE. and Wijdenes, J., 1977. Chromatofocusing: Isoelectric focusing on ion-exchange columns ii. Experimental verification. Joournal of Chromatography, 150, 31-44

Sorensen, H.P., Sperling-Peterson, H.U. and Mortensen, K.K., 2003. Production of recombinant thermostable proteins expressed in *Escherichia col*:. Completion of protein sysntheis is the bottleneck. Journal of Chromatography B, 786 (1-2), 207-214

Sreerama, N. and Woody, R.W., 2004. Computation and analysis of protein circular dichroism spectra. Method in Enzymology. 14(383), 318-351

Svensson, L., Dormitzer, P.R., Von Bonsdorff, C., Maunula, L. and Greenberg, H.B. 1994. Intracellular manipulation of disulfide bond formation in rotavirus proteins during assembly. Journal of Virology, 5204-5215

Tang, B., Wang, Y., Zhang, D. and Zhang, H., 2009. Fluorescence properties and conformational stability of the hemocyanin from Chinese mitten crab *Eriocheir japonica sinensis* (Decapoda, Grapsidae). Journal of Molecular Structure, 920(1–3), 454–458.

Terpe, K., 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems, Applied Microbiology and Biotechnology, 72, 211-222

Tissera, M.S., Cowley, D., Bogdanovic-Sakran, N., Hutton, M.L., Lyras, D., Kirkwod, C.D., and Buttery, J.P., 2017. Options for improving effectiveness of rotavirus vaccines in developing countries. Human Vaccines & Immunotherapeutics, 13(4): 921–927.

Trask, S.D. and Dormitzer, P.R., 2006. Assembly of highly infectious rotavirus particles recoated with recombinant outer capsid proteins. Journal of virology, 80(22), 11293–11304.

Trask, S.D., Kim, I.S., Harrison, S.C., and Dormitzer. P.R., 2010 A Rotavirus spike protein conformational intermediate binds lipid bilayers. Journal of Vitology, 84(4): 1764–1770.

Trask, S.D., McDonald, S.M. and Patton, J.T., 2013. Structural Insights into the Coupling of Virion Assembly and Rotavirus Replication, Nature Reviews Microbiology NIH Public Access. , 10(3), 165–177

Trask, S.D., Ogden, KM. and Patton, J.T., 2013. Interactions among capsid proteins orchestrate rotavirus particle functions. Current Opinion in Virology, 2(4), 373–379.

Tsumoto, K., Umetsu, M., Kumangai I., Ejima, D., Phio, J.S. and Arakawa, T., 2004. Role of arginine in protein refolding, solubilization, and purification. Biotechnol. Prog. 20, 1301-1308

Vargas-Cortez, T., Morones-Ramirez, J.R. and Baideras-Renteria, I., 2016. Expression and purification of recombinant proteins in *Escherichia coli* tagged with a small metal-binding protein from Nitrosomonas. Protein Expression and Purification. 118, 49-54

Van den Berg, B., Ellis, R and Dobson, C., 1999. Effects of macromolecula crowding on protein folding and aggregation. EMBO journals, 18 (24), 6927-6933

Van Hulten, M.C.W., Westenberg, M., Goodall, S.D. and Vlak,J.M., 2000. Identification of two major virion protein genes of white spot syndrome virus of shrimp. Virology, 266, 227-236

Vicente, T., Roldao, A., Peixoto, C., Carrondo, M.J.T and Alves, P.M., 2011. Large-scale production and purification of VLP-based vaccines. Journal of Intertebrate Pathology, 107, s42-s48

Villaverde, A. and Carrio, M. M., 2003. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. Biotechnology Letters, 25(17), 1385-1395

Vivian, J.T. and Callis, P.R., 2001. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. Biophysical Journal, 80(5), 2093–2109.

Voet, D.and Voet, J.G., 2004. Biochemistry 3rd edition, Wiley 9, 276-302

Qi, Y., Yan, Z. Huang, J., 2001. Chromatography on DEAE ion-exchange and protein G affinity columns in tandem for the separation and purification of proteins. Journal of Biochemical Biophysical Methods. 49, 263–273

Waelter, S., Boeddrich, A., Lurz R., Scherzinger, E., Lueder, G., Lehrach, H. and Wanker, E.E., 2001. Accunulation of mutant Huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation. Molecular Biology of the Cell, 12, 1393-1407

Wang, L., Huang, J., Phelps, A., Firth, S., Holmes, I.H. and Reeves, P.R., 1996. Periplasmic expression of part of the major rotavirus capsid protein VP7containing all the three antigenic regions in *Escherichia coli*. Gene. 177, 155-162

Waugh, D.S., 2005. Making the most of affinity tags. TRENDS in Biotechnology. 23(6), 317-320

Whitmore, L. and Wallace, B. A., 2008. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. Biopolymers, 89(5), 392–400.

Wu, X., Liu, Q., HE, J., Zang, M., Wang, H., Li, Y. and Tang, L., 2015. Preparation and characterization of a Monoclonal antibody against the core protein VP7 of the 25th serotype of bluetongue virus. Mary Ann Liebert, 34(2), 116-120

Wu, Y. Z., Li, J., Mou, Z., Fei, L., Ni, B., Geng, M., Jia, Z.C. and Tang, Y., 2003. Oral immunization with rotavirus VP7 expressed in transgenic potatoes induced high titers of mucosal neutralizing IgA. Virology, 313(2), 337–342.

Wurm, D. J. Veiter, L., Ulonska, S., Eggenreich, B., Herwig, C. and Spadiul, O., (2016). The *E. coli* pET expression system revisited—mechanistic correlation between glucose and lactose uptake. Applied Microbiology and Biotechnology, 100(20), 8721–8729.

Xu, F., Geiger, J.H., Baker, G.L. and Bruening, M.L., 2011. Polymer brush-modified magnetic nanoparticles for his-tagged protein purification. Langmuir, 27(6), 3106–3112.

Xu, Y., Mazzawi, M., Chen, K., Sun, L. and Paul, L.D., 2011. Protein purification by polyelectrolyte coacervation: Influence of protein charge anisotropy on selectivity. Biomacromolecules, 12(5), 1512–1522.

Young, C. L., Brittan, Z.T. and Robinson, A.S., 2012. Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications. Biotechnology Journal, 7(5), 620–634.

Zaveckas, M., Snipaitis, S., Pesliakas, H., Nainys, J. and Gedviaite, A., 2015. Purification of recombinant virus-like particles of porcine circovirus type 2 capsid protein using ion-exchange monolith chromatography. Journal of Chromatography B. 991, 21–28

Zhang, C., Gao, C., Mu, J.,Qiu, Z. and Li, L., 2013. Spectroscopic studies on unfolding processes of apo-neuroglobin Induced by guanidine hydrochloride and urea. BioMed Research International, 349542, 1-7

Zhao, X.B., Wang, L. and Liu, D.H., 2007. Effect of several factors on peracetic acid pretreatment of sugarcane bagasse for enzymatic hydrolysis. Journal of Chemical Technology & Biotechnology, 82, 1115–1121.

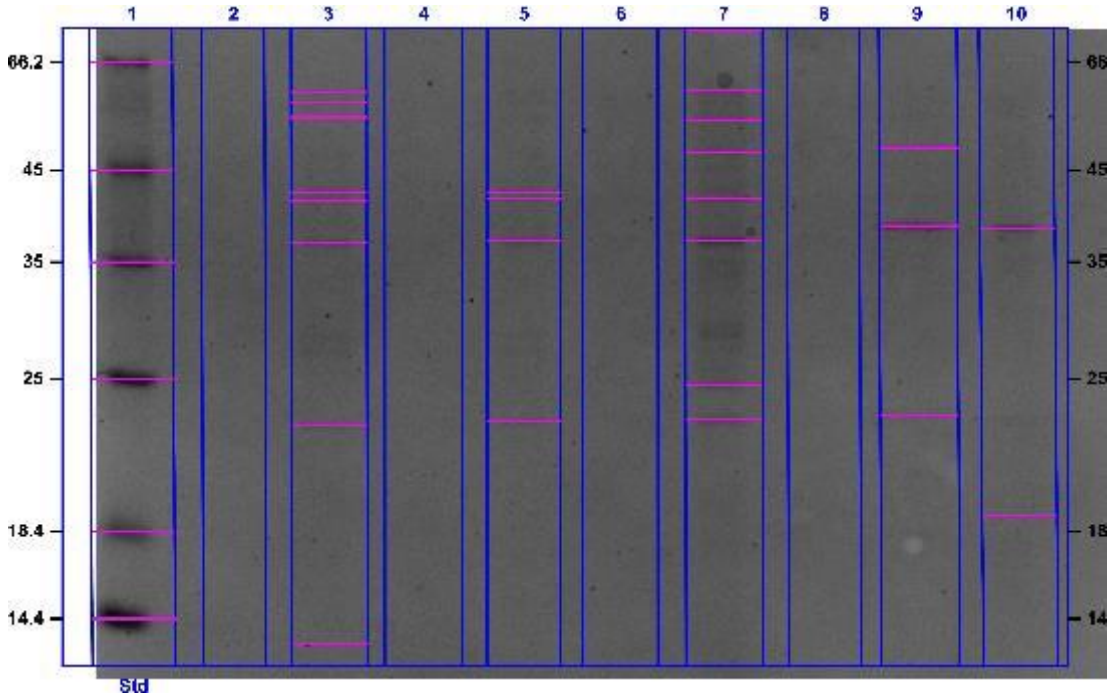
Zhou, Y.J., Burns, J.W., Morita, Y., Tanaku, T. and Estes, M.K., 1994. Localization of rotavirus VP4 neutralization epitopes involved in antibody-induced conformational changes of virus structure. Journal of virology, 68(6), pp.3955–64.

7. Appendices

Appendix 7: The pET28a+ vector used to insert the VP7 protein with the His-Tag for transformation

(MYGIEYTTILTFLISIILLNYILKSITRIMDYIIYRFLLIIVILSPFVRAQNYGINLPITGSMDTAYANSTQEETFLTSTLCLYYPTEASTQIND
GEWKDTLSQLFLTKGWPTGSVYFKEYSDIASFSVDPQLYCDYNLVLMKYDQTLELDMSELADLILNEWLCNPMDITLYYYQQTDES
NKWISMGSSCTVKVCPLNTQTLGIGCQTTNVDTFEEVATSEKLVITDVVDGVNHKINVTTTTCTIRNCKKLGPRENVAVIQVGGSN
VLDITADPTTNPQTERMMRVNWKKWWQVFYTVVDYINQIVQVMSKRSRSLNSAAFYYRV)

Appendix 7.1: Analysis results for figure 3.9: VP7 ion-exchange purification elution with a gradient final concentration of 0.5 M sodium chloride



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.061 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2016/04/13 9:38:38 AM
User Name	User
Image Area (mm)	X: 60.1 Y: 38.1
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	12251 - 25240

Analysis Settings

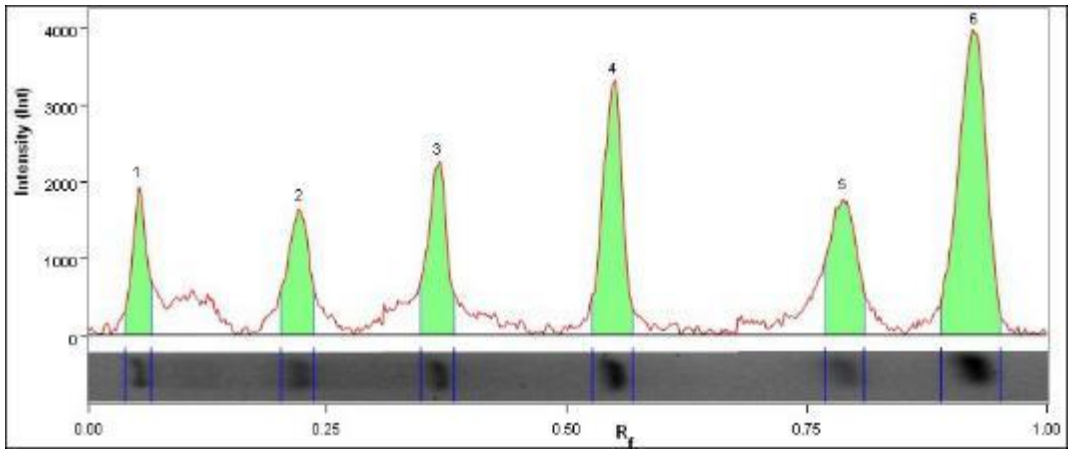
Detection	Lane detection: Automatically detected lanes with manual adjustments Band detection: Automatically detected bands with sensitivity: High
	Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: Variable
Mol. Weight Analysis	Standard: Copy of BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	18 870 522	151 389 466	25 826 179	551 985 524	526 159 345
2	N/A	N/A	18 432 360	445 371 372	426 939 012
3	2 088 060	81 661 840	7 813 520	546 005 200	538 191 680
4	N/A	N/A	13 066 680	552 956 976	539 890 296
5	2 235 114	39 113 235	13 041 756	530 626 383	517 584 627
6	N/A	N/A	8 008 384	543 849 216	535 840 832
7	14 496 702	85 169 040	73 069 326	569 262 210	496 192 884
8	N/A	N/A	9 280 594	514 020 424	504 739 830
9	3 393 684	48 290 250	18 897 417	551 981 510	533 084 093
10	1 635 072	27 527 232	15 633 728	521 489 216	505 855 488

Lane and Band Analysis

Lane 1 - Copy of BIOTECH STD

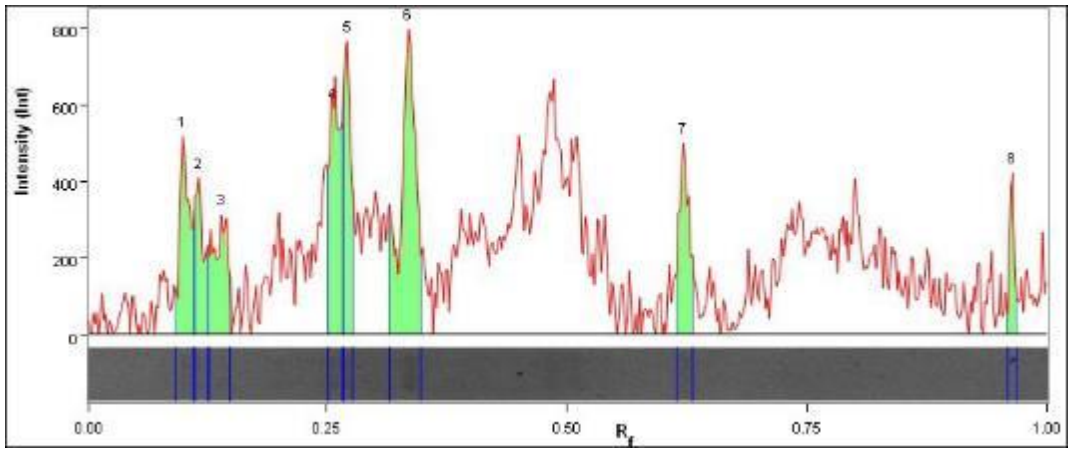


Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	B	66,2	8,2
2	C	45,0	10,5
3	D	35,0	12,9
4	E	25,0	19,8
5	F	18,4	14,2

6	G	14,4	34,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.85 mm
Regression Equation	A single equation is not available for this method

Lane 3

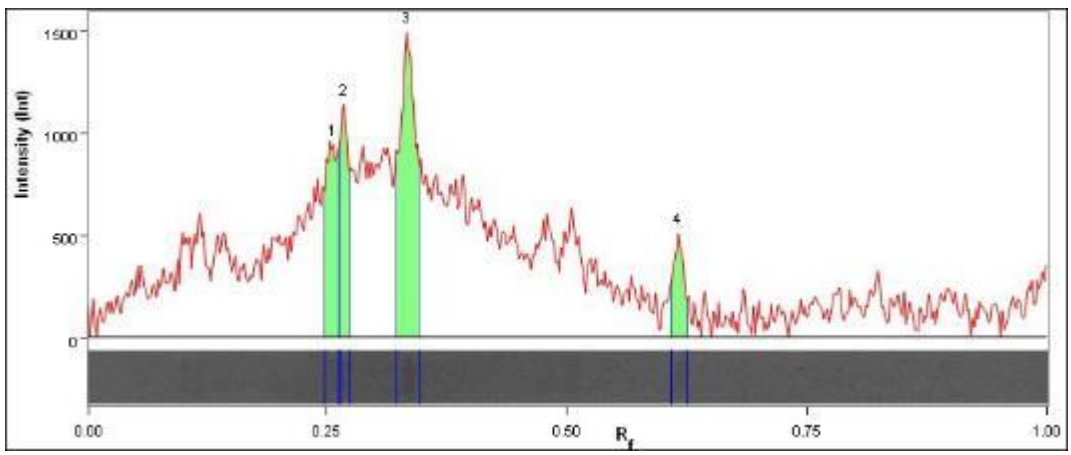


Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		59,6	11,7
2		57,4	7,9
3		54,4	10,1
4		42,4	16,7
5		41,3	12,2
6	VP7	36,9	25,7
7		22,8	11,0
8		14,4	4,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.44 mm
Regression Equation	A single equation is not available for this method

Regression Equation	A single equation is not available for this method

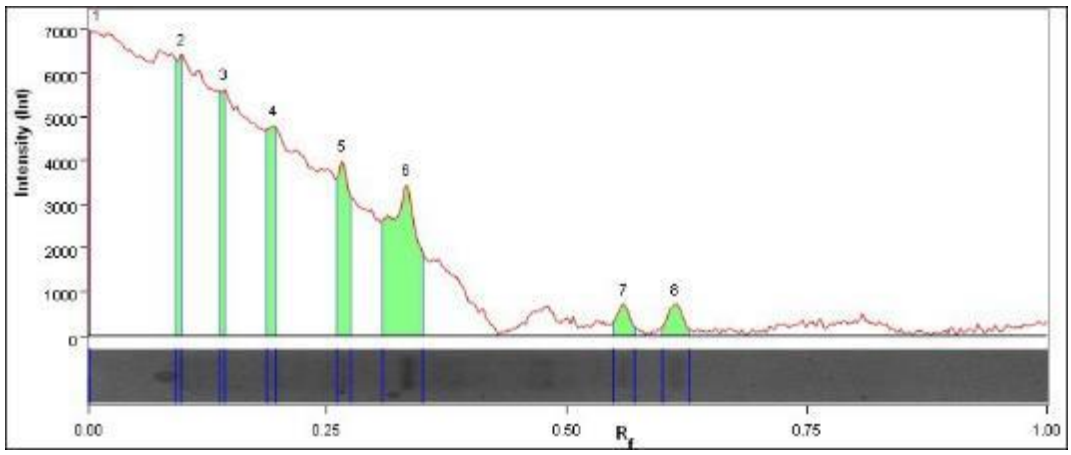
Lane 5



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		42,4	24,0
2		41,6	18,1
3	VP7	37,1	47,4
4		22,9	10,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.30 mm
Regression Equation	A single equation is not available for this method

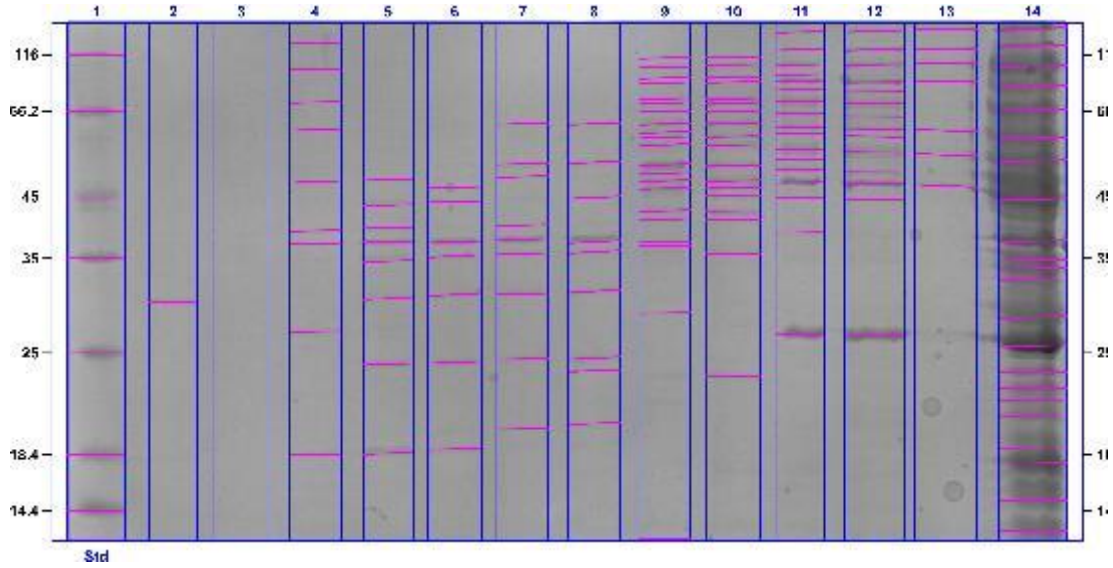
Lane 7



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		66,2	6,8
2		59,9	12,3
3		53,9	10,8
4		47,9	16,1
5		41,7	15,8
6	VP7	37,1	31,5
7		24,7	3,0
8		23,0	3,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.50 mm
Regression Equation	A single equation is not available for this method

Appendix 7.2: Analysis results for figure 3.10: VP7 ion-exchange purification by elution with 0.1; 0.2 and 0.5 M sodium chloride buffers



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.075 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2016/05/13 5:14:50 PM
User Name	User
Image Area (mm)	X: 81.6 Y: 42.4
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	3873 - 30261

Analysis Settings

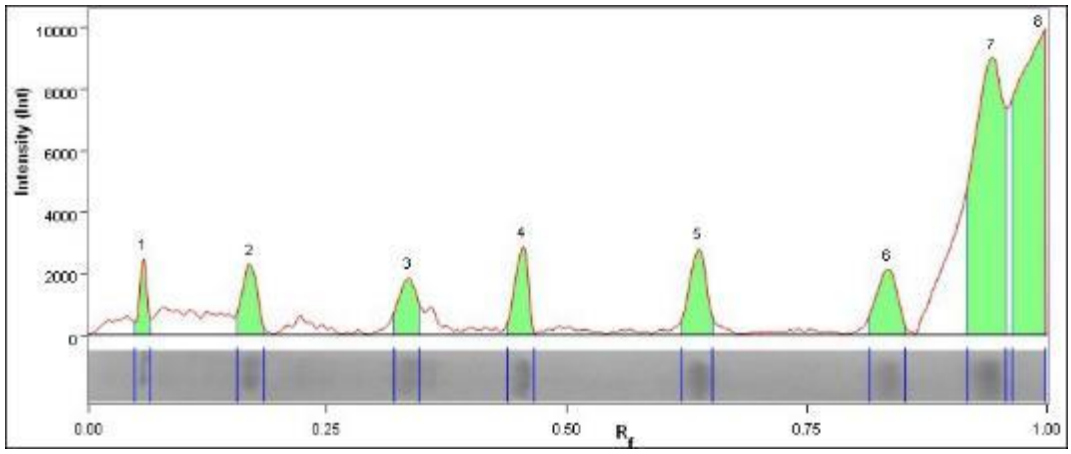
Detection	Lane detection: Manually created lanes Band detection: Automatically detected bands with sensitivity: High Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10
	Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	42 617 330	139 408 150	57 956 990	506 867 060	448 910 070
2	545 896	11 144 816	45 359 886	447 135 514	401 775 628
3	N/A	N/A	11 825 385	508 421 485	496 596 100
4	1 859 280	35 295 149	28 937 851	479 798 428	450 860 577
5	1 513 620	30 213 480	29 643 480	483 325 260	453 681 780
6	3 640 768	78 769 792	31 243 200	516 055 872	484 812 672
7	2 770 966	52 356 768	9 564 306	491 131 884	481 567 578
8	2 274 552	38 668 959	19 410 993	497 356 335	477 945 342
9	12 831 086	170 236 686	29 795 526	503 998 868	474 203 342
10	9 488 952	157 124 220	24 938 694	540 032 658	515 093 964
11	15 319 598	154 137 146	25 293 220	469 447 650	444 154 430
12	17 880 143	182 059 336	27 053 769	590 161 088	563 107 319
13	2 597 326	58 355 808	16 144 062	591 290 784	575 146 722
14	77 110 866	552 580 218	139 544 127	927 830 619	788 286 492

Lane And Band Analysis

Lane 1 - BIOTECH STD



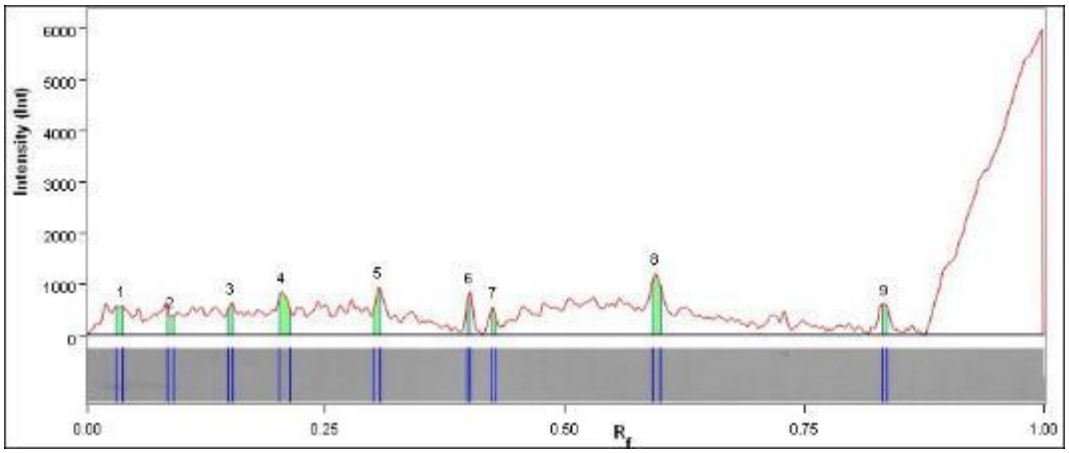
Band No.	Band Label	Mol. Wt. (KDa)	Volume (Int)	Band %
1	A	116,0	9 324 560	2,6
2	B	66,2	16 560 460	4,7
3	C	45,0	15 327 060	4,3
4	D	35,0	16 080 120	5,1
5	E	25,0	19 081 860	6,4
6	F	18,4	21 276 920	6,1
7	G	14,4	24 365 390	35,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.78 mm
Regression Equation	A single equation is not available for this method

Regression Equation	A single equation is not available for this method

Regression Equation	A single equation is not available for this method

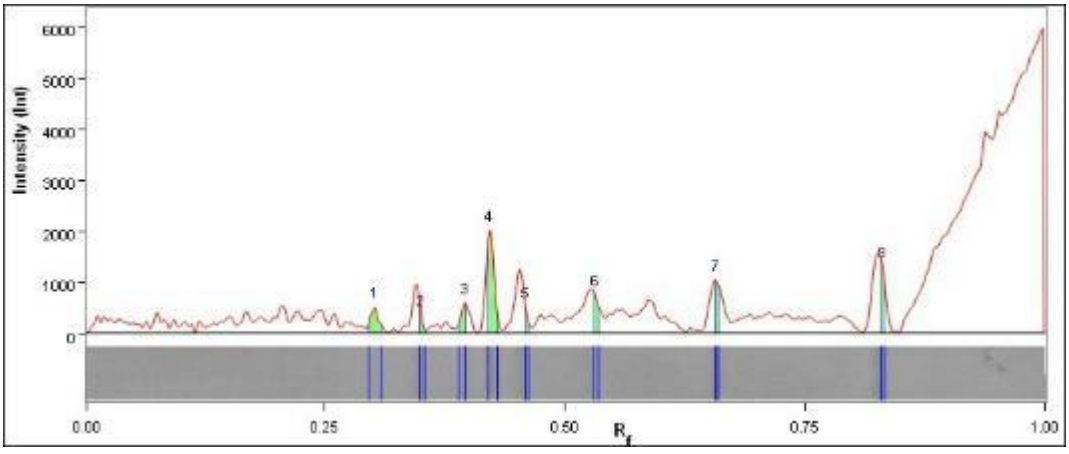
Lane 4



Band No.	Band Label	Mol. Wt. (KDa)	Band %	
1		116,0	9,5	
2		100,0	5,1	
3		72,5	9,7	
4		60,9	17,3	
5		48,2	15,6	
6		39,2	8,1	
7	VP7	37,2	5,1	
8		27,0	22,0	
9		18,4	3 132 838	7,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.16 mm
Regression Equation	A single equation is not available for this method

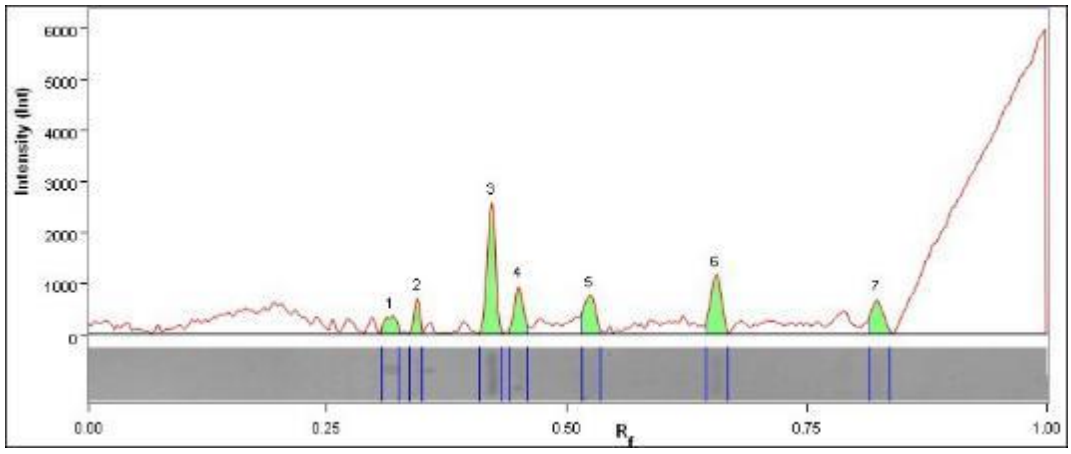
Lane 5



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		48,5	12,7
2		43,5	2,4
3		39,6	11,4
4	VP7	37,5	33,4
5		34,7	4,3
6		30,3	10,1
7		24,2	12,5
8		18,5	13,2

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.09 mm
Regression Equation	A single equation is not available for this method

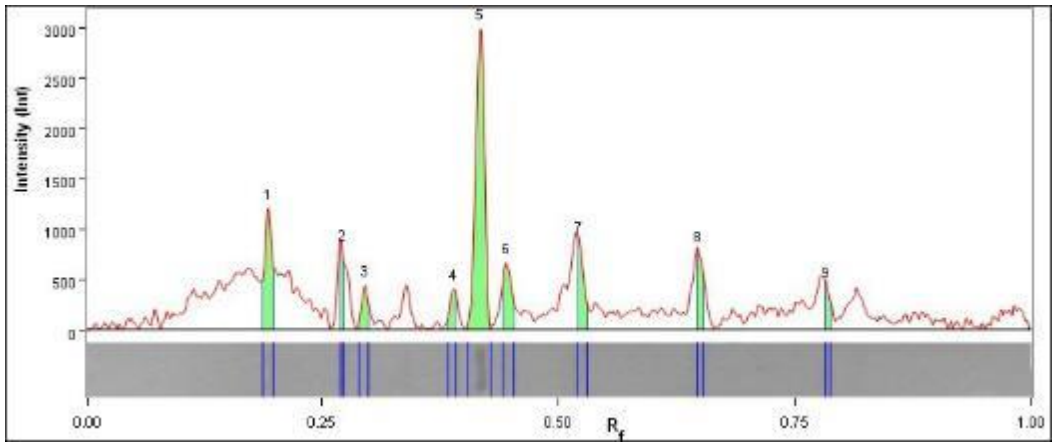
Lane 6



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		46,9	6,4
2		44,1	5,9
3	VP7	37,5	31,8
4		35,4	12,3
5		30,8	13,0
6		24,3	18,8
7		18,7	11,7

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.37 mm
Regression Equation	A single equation is not available for this method

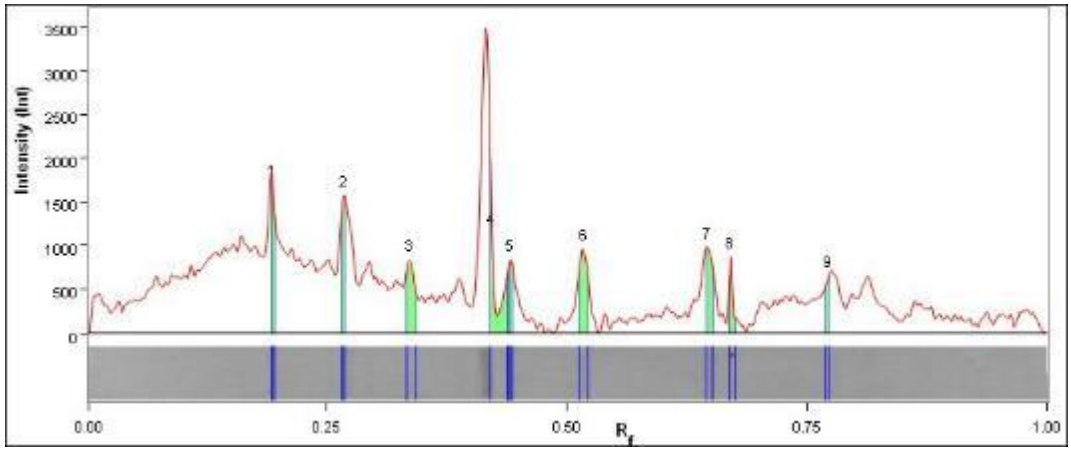
Lane 7



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		62,5	15,3
2		52,2	4,6
3		49,3	4,4
4		40,1	4,7
5	VP7	37,8	45,3
6		35,7	8,1
7		30,9	7,5
8		24,6	7,1
9		19,9	2,9

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.23 mm
Regression Equation	A single equation is not available for this method

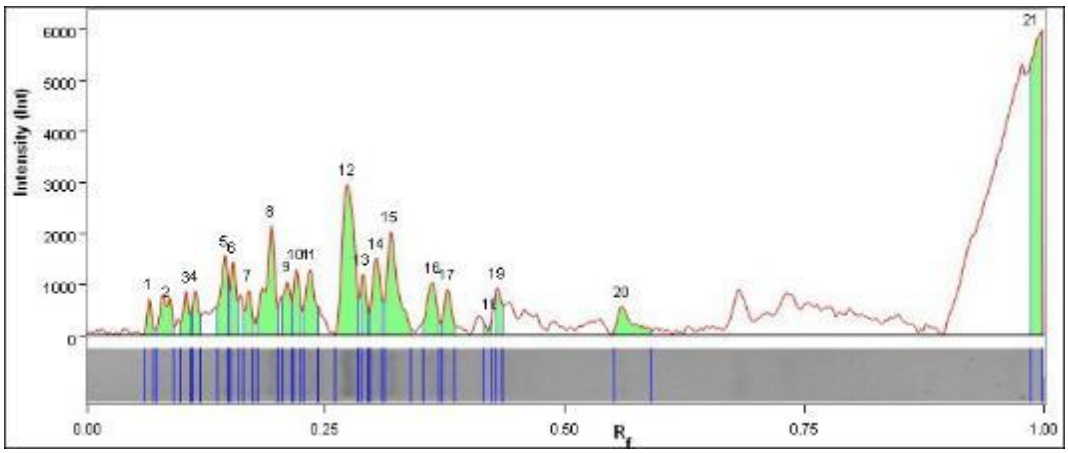
Lane 8



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		62,5	12,2
2		52,6	13,4
3		44,8	12,2
4	VP7	37,5	17,0
5		36,0	6,8
6		31,2	15,9
7		24,6	12,2
8		23,7	4,8
9		20,2	5,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.30 mm
Regression Equation	A single equation is not available for this method

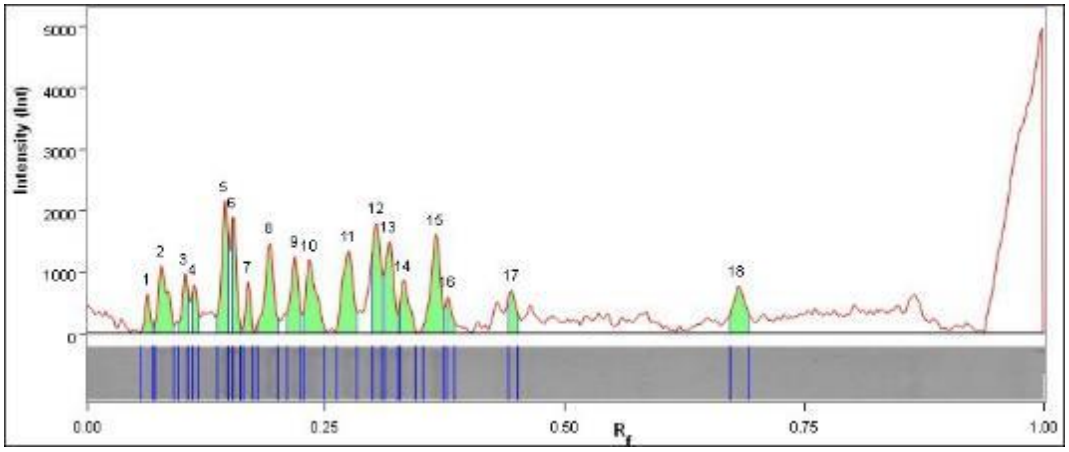
Lane 9



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		112,2	1,2
2		102,5	3,1
3		92,1	2,0
4		87,6	1,9
5		74,9	4,4
6		71,9	3,2
7		66,2	2,2
8		62,5	7,8
9		60,2	3,0
10		58,9	3,0
11		56,9	4,4
12		52,0	13,4
13		50,0	2,3
14		48,4	4,5
15		46,7	8,5
16		42,5	3,7
17		41,1	2,6
18	VP7	37,4	0,7
19		36,8	1,6
20		28,8	3,3
21		14,4	23,2

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.23 mm
Regression Equation	A single equation is not available for this method

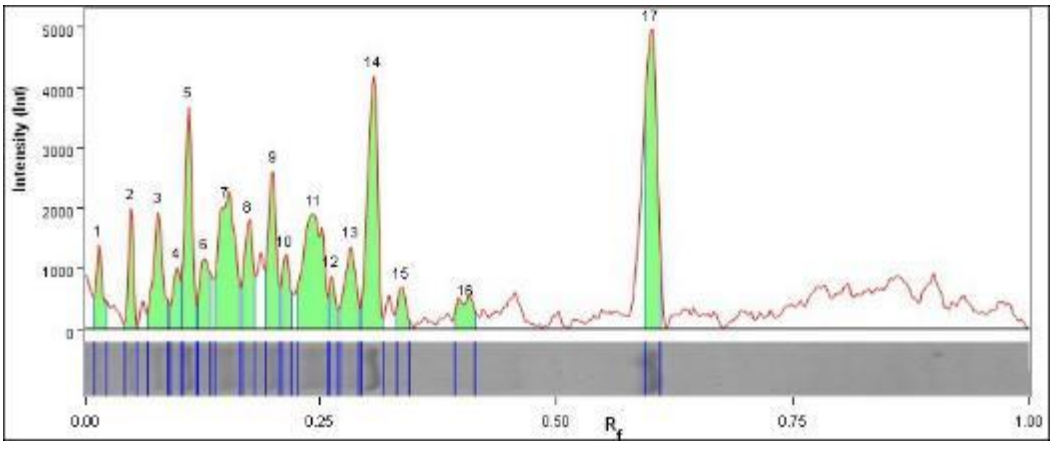
Lane 10



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		113,2	2,0
2		105,1	6,4
3		92,8	3,3
4		88,4	2,5
5		74,9	8,4
6		71,9	5,1
7		66,2	2,4
8		62,8	7,9
9		59,1	6,1
10		56,9	7,6
11		51,8	8,9
12		48,4	7,5
13		46,9	7,7
14		45,3	4,4
15		42,2	9,3
16		41,1	2,0
17		35,7	2,7
18		23,4	5,7

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.50 mm
Regression Equation	A single equation is not available for this method

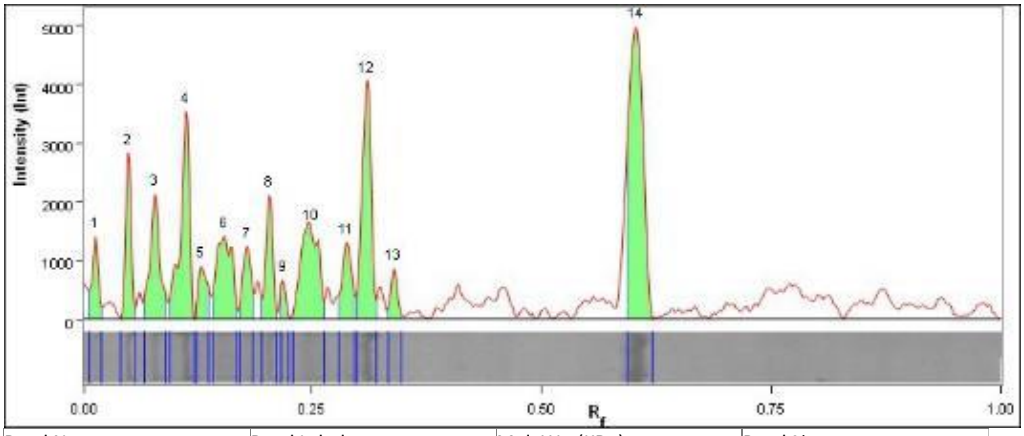
Lane 11



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	2,9
2		116,0	3,0
3		105,1	5,5
4		95,2	2,4
5		89,1	8,0
6		82,0	3,5
7		73,1	11,0
8		65,5	4,7
9		61,6	6,6
10		59,8	2,7
11		55,6	11,5
12		53,4	1,5
13		50,8	4,5
14		48,0	12,7
15		44,8	1,7
16		38,9	2,4
17		26,8	15,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	3.96 mm
Regression Equation	A single equation is not available for this method

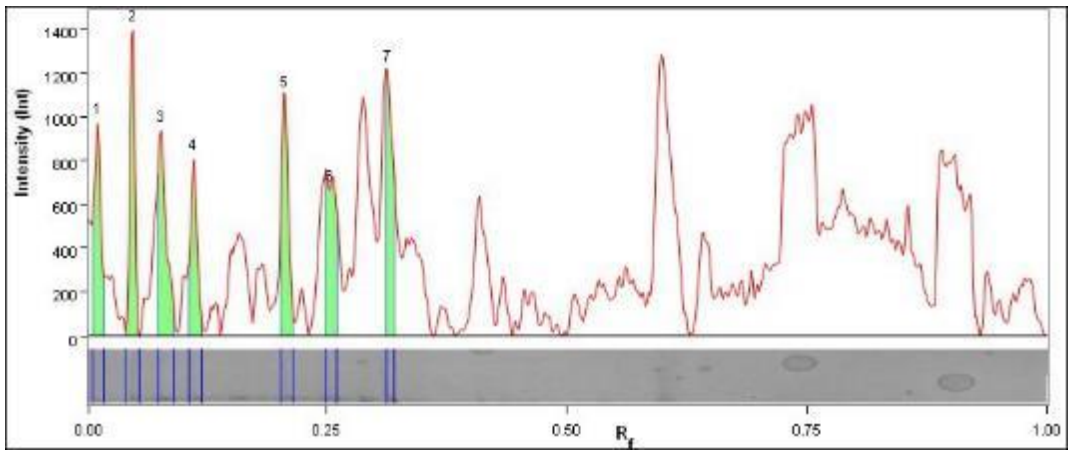
Lane 12



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	3,1
2		116,0	5,5
3		105,1	7,1
4		88,4	11,1
5		81,4	2,7
6		71,3	7,8
7		64,7	3,7
8		61,1	5,2
9		59,1	0,9
10		55,0	10,2
11		50,2	4,3
12		47,6	13,2
13		44,5	2,0
14		26,6	23,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	4.85 mm
Regression Equation	A single equation is not available for this method

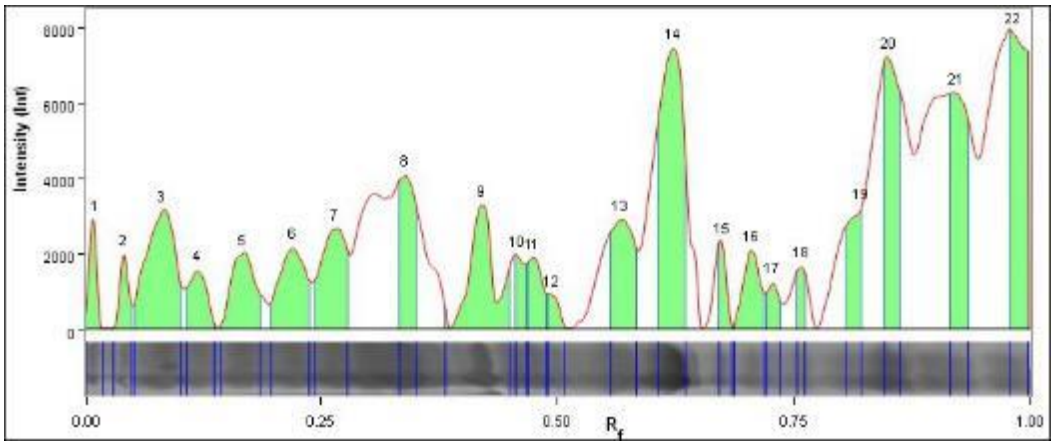
Lane 13



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	14,4
2		116,0	14,7
3		105,9	14,5
4		89,1	11,3
5		60,7	14,8
6		54,4	14,9
7		47,3	15,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	5.05 mm
Regression Equation	A single equation is not available for this method

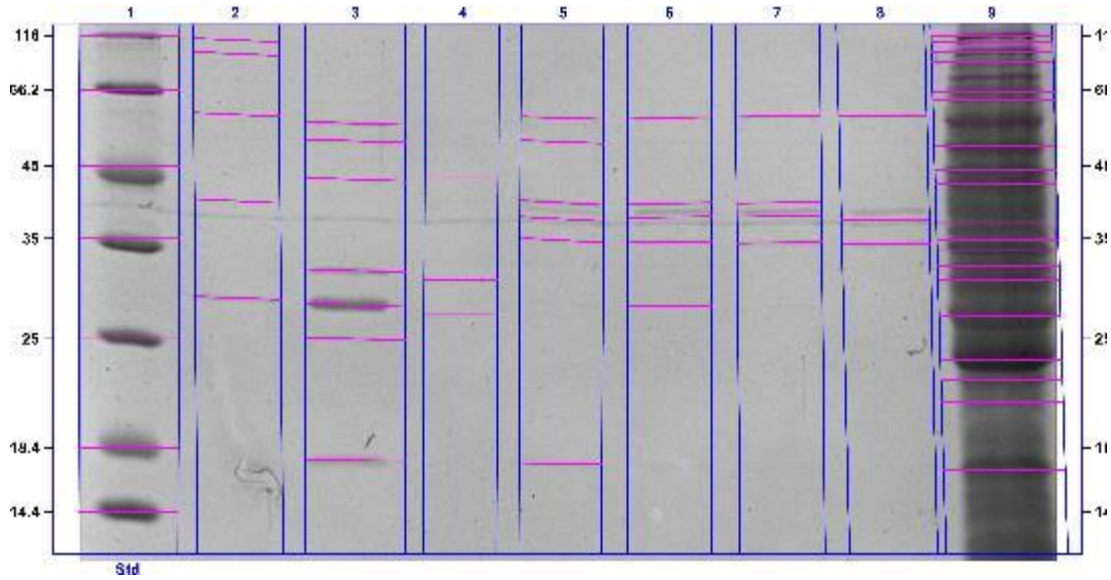
Lane 14



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	1,9
2		116,0	1,5
3		103,3	7,7
4		85,5	2,3
5		67,3	4,0
6		58,9	4,7
7		53,0	5,5
8		44,5	5,1
9	VP7	37,6	6,9
10		34,8	1,4
11		33,8	2,3
12		32,5	0,9
13		28,4	5,1
14		25,7	13,8
15		23,6	1,2
16		22,5	2,9
17		21,7	1,1
18		20,7	1,0
19		18,8	3,3
20		17,7	8,4
21		15,1	8,8
22		14,4	10,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	5.53 mm
Regression Equation	A single equation is not available for this method

Appendix 7.3: Analysis results for figure 3.11: VP7 ion-exchange purification by pH exchange



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.060 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2016/09/06 4:44:15 PM
User Name	User
Image Area (mm)	X: 83.2 Y: 45.9
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	5123 - 36007

Analysis Settings

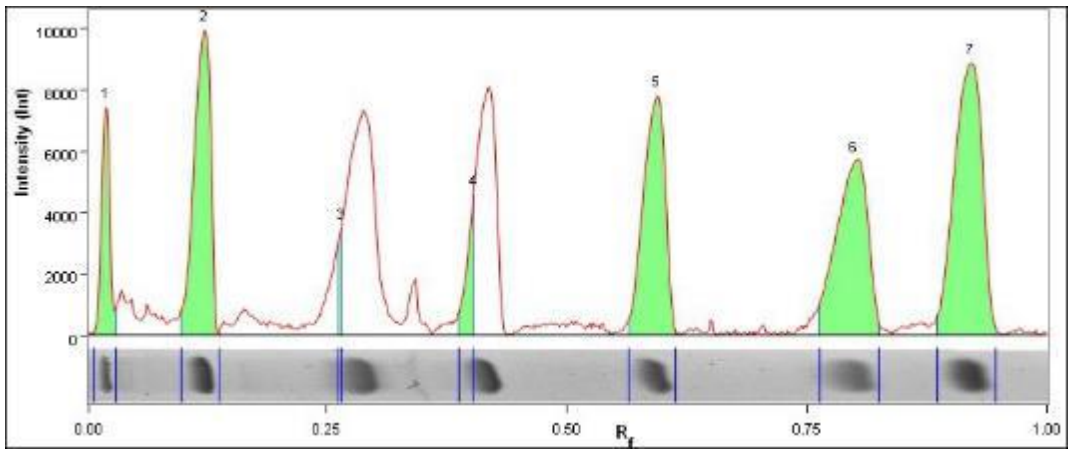
Detection	Lane detection: Automatically detected lanes with manual adjustments Band detection: Automatically detected bands with sensitivity: High Manually adjusted bands
	Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	99 224 000	289 586 625	158 143 000	888 736 625	730 593 625
2	3 786 623	50 217 347	29 800 998	706 056 192	676 255 194
3	28 655 375	171 802 250	47 584 250	857 194 000	809 609 750
4	928 188	24 032 792	12 536 196	630 311 504	617 775 308
5	4 838 080	74 580 272	18 934 448	713 494 808	694 560 360
6	4 748 625	82 793 655	30 795 345	709 510 620	678 715 275
7	2 824 476	39 999 418	23 318 092	702 593 652	679 275 560
8	1 155 220	13 125 200	23 237 830	769 426 460	746 188 630
9	65 525 680	797 886 152	140 948 536	2 216 360 640	2 075 412 104

Lane And Band Analysis

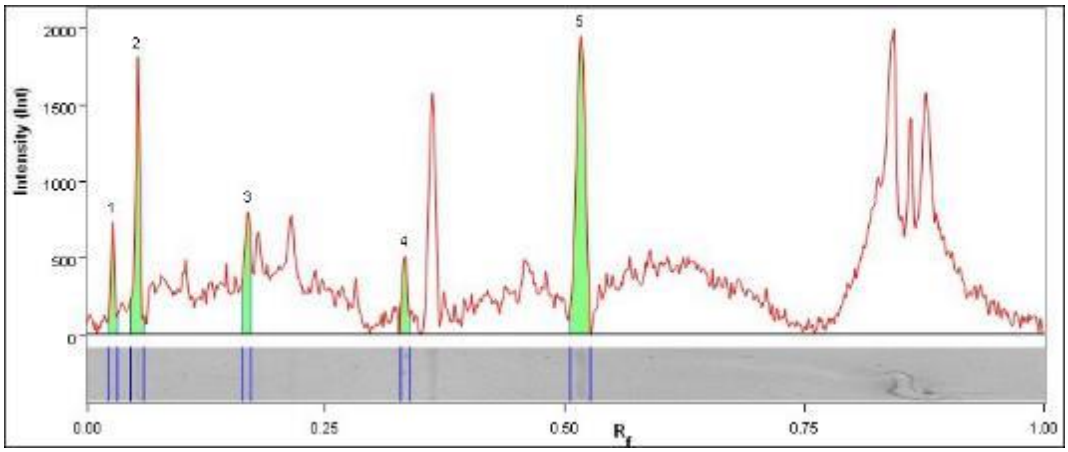
Lane 1 - BIOTECH STD



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	7,1
2	B	66,2	18,9
3	C	45,0	1,5
4	D	35,0	4,3
5	E	25,0	19,4
6	F	18,4	20,5
7	G	14,4	28,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.53 mm
Regression Equation	A single equation is not available for this method

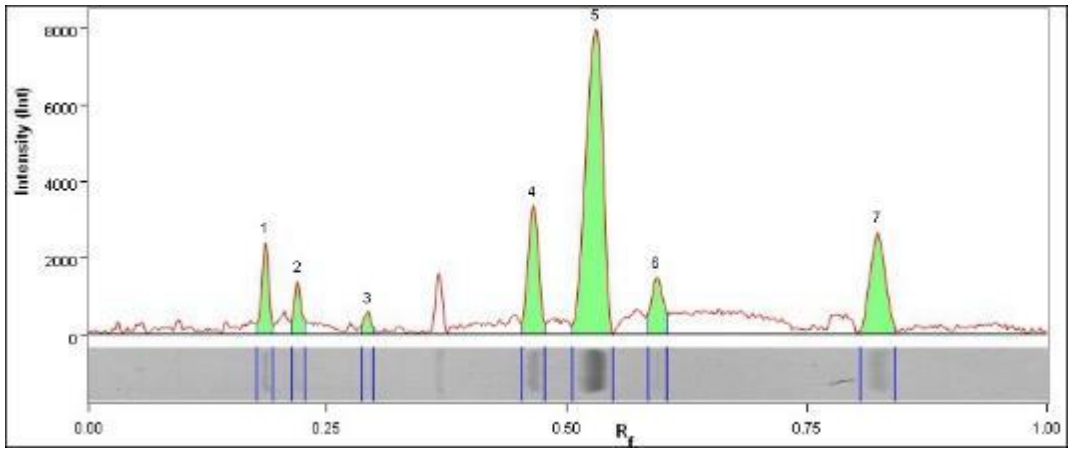
Lane 2



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		110,4	7,7
2		96,0	22,2
3		58,3	14,0
4		39,8	7,6
5		28,7	48,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.30 mm
Regression Equation	A single equation is not available for this method

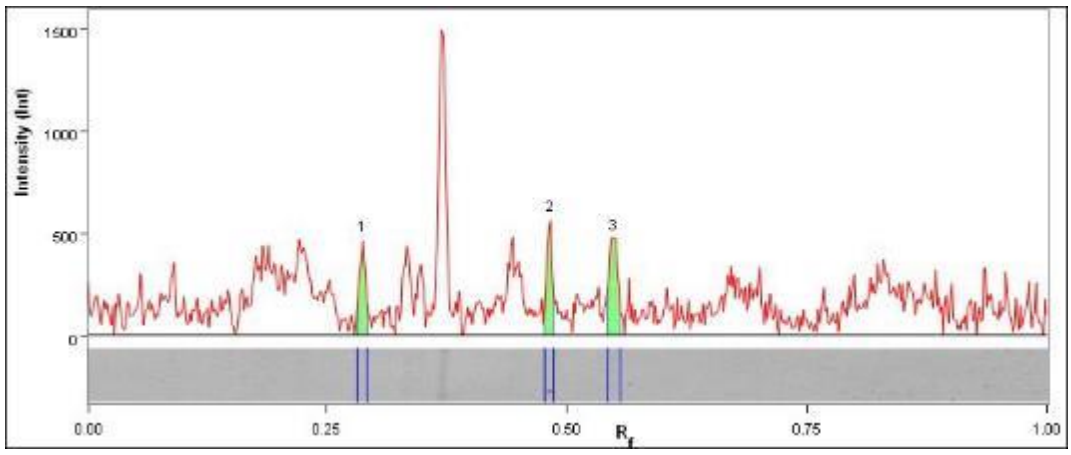
Lane 3



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		55,8	6,4
2		51,0	3,8
3		42,9	1,6
4		31,4	13,9
5		28,0	52,1
6		25,0	6,7
7		17,5	15,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.53 mm
Regression Equation	A single equation is not available for this method

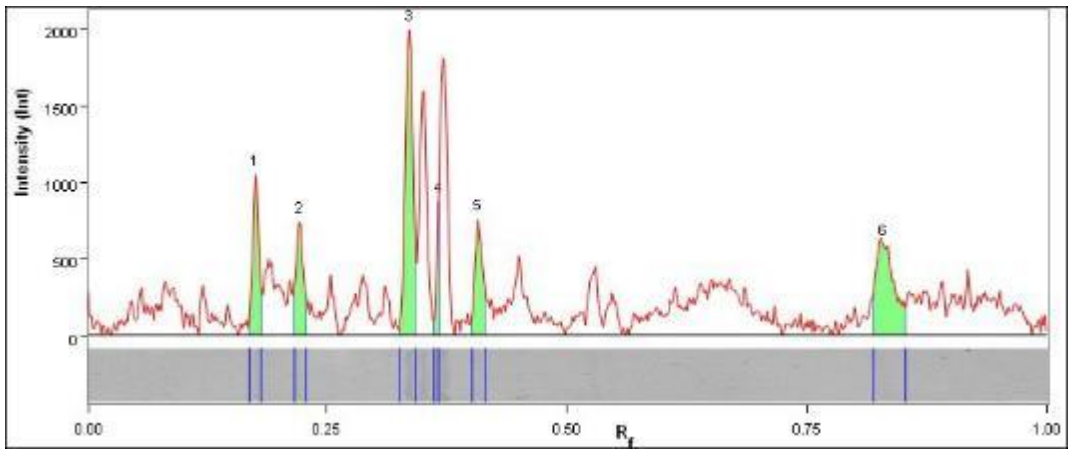
Lane 4



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		43,3	28,1
2		30,4	29,8
3		27,1	42,1

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.28 mm
Regression Equation	A single equation is not available for this method

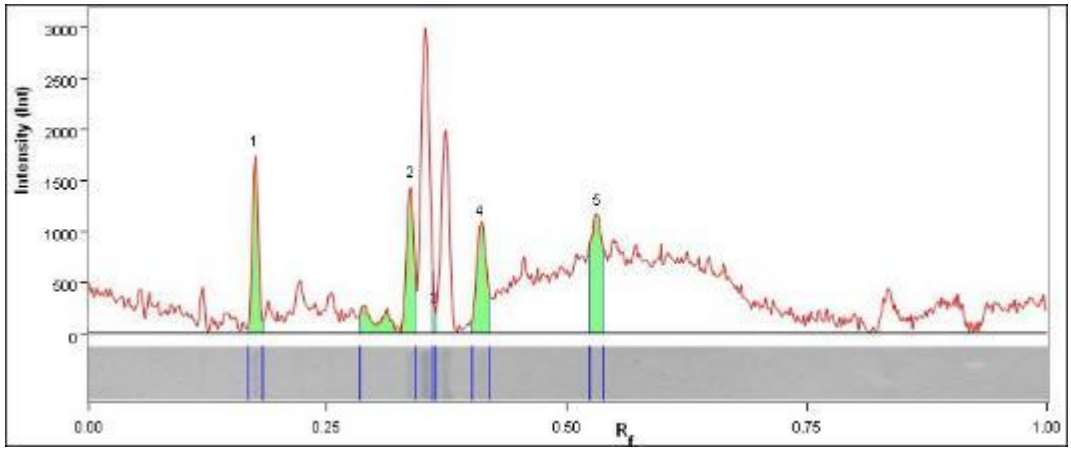
Lane 5



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		57,4	13,5
2		50,8	11,2
3		39,6	32,0
4	VP7	37,5	7,8
5		34,8	11,8
6		17,3	23,7

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.10 mm
Regression Equation	A single equation is not available for this method

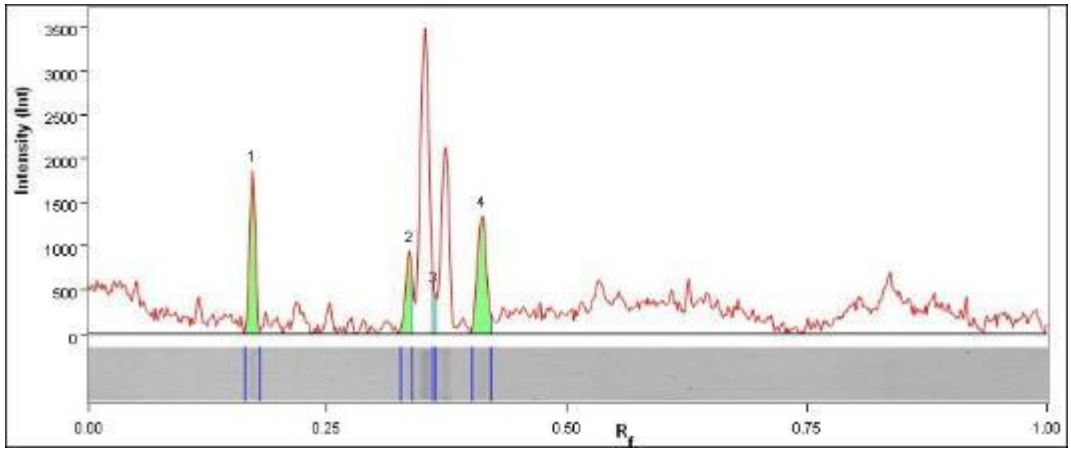
Lane 6



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		57,4	20,7
2		39,5	30,7
3	VP7	37,7	1,9
4		34,6	21,4
5		27,9	25,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.17 mm
Regression Equation	A single equation is not available for this method

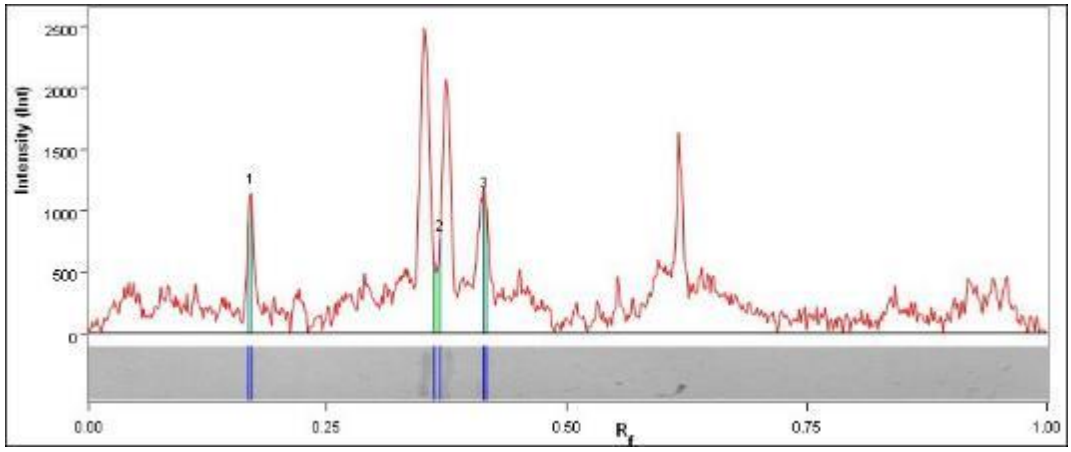
Lane 7



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		57,9	35,1
2		39,6	19,4
3	VP7	37,8	5,1
4		34,5	40,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

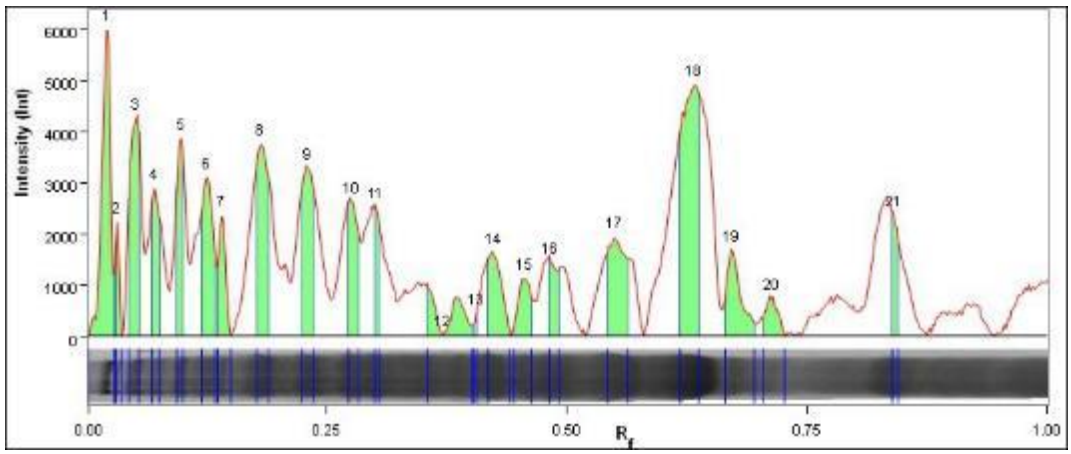
Lane 8



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		58,1	33,5
2	VP7	37,4	35,8
3		34,3	30,7

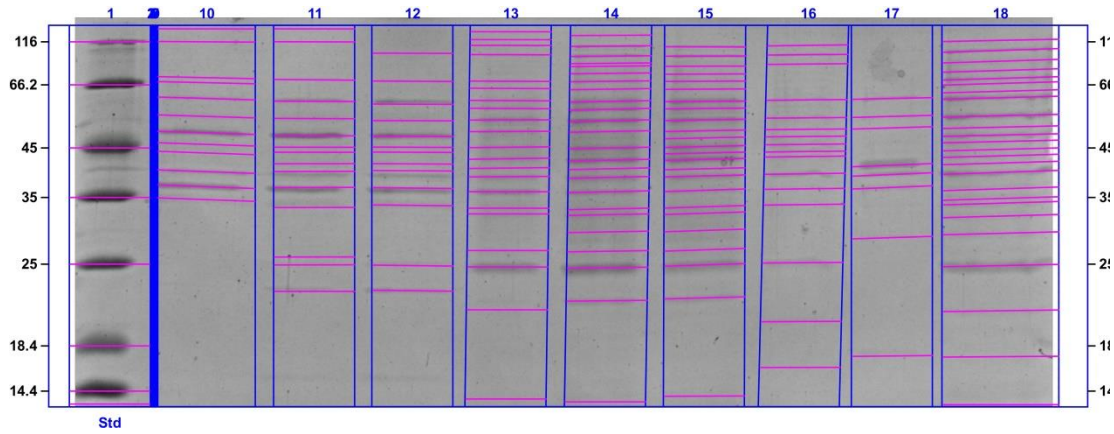
Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.51 mm
Regression Equation	A single equation is not available for this method

Lane 9



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		115,0	10,1
2		108,6	1,5
3		97,6	7,5
4		87,6	3,7
5		75,5	4,3
6		65,7	5,7
7		63,0	3,2
8		56,5	7,5
9		49,6	6,8
10		44,2	4,1
11		42,2	2,2
12	VP7	37,2	3,3
13		34,9	0,3
14		33,8	3,8
15		31,9	2,6
16		30,4	2,6
17		27,0	6,2
18		23,6	15,7
19		22,2	4,5
20		20,9	1,8
21		16,9	2,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	10.37 mm
Regression Equation	A single equation is not available for this method



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.059 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2016/09/06 3:49:01 PM
User Name	User
Image Area (mm)	X: 93.0 Y: 36.9
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	145 - 31214

Analysis Settings

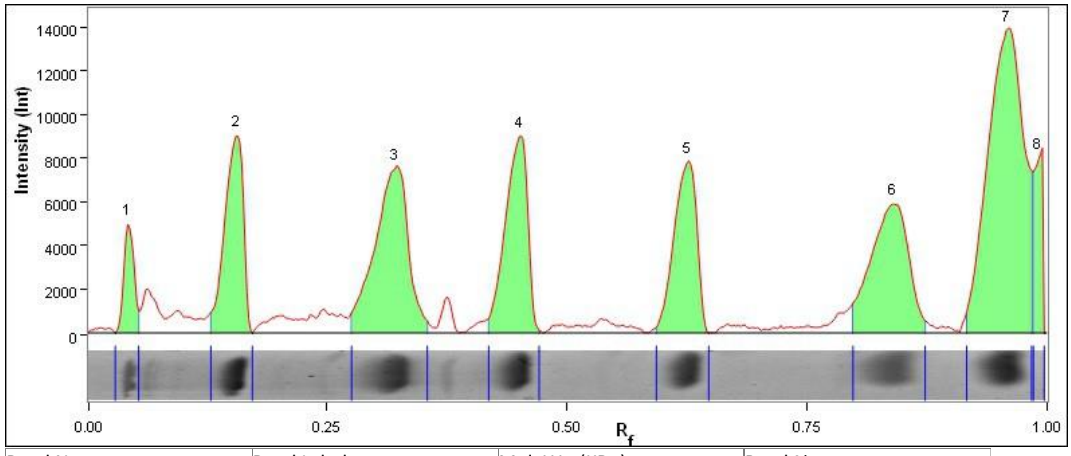
Detection	Lane detection: Manually created lanes Band detection: Bands detected with different sensitivity per lane Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first
	Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	135 155 684	376 203 668	153 664 406	762 329 075	608 664 669
2	N/A	N/A	1 306 984	6 407 104	5 100 120
3	N/A	N/A	740 355	6 408 248	5 667 893
4	N/A	N/A	667 862	6 445 441	5 777 579
5	N/A	N/A	667 866	6 445 445	5 777 579
6	N/A	N/A	813 062	6 486 748	5 673 686
7	N/A	N/A	3 833 808	19 474 359	15 640 551
8	N/A	N/A	796 483	6 487 618	5 691 135
9	N/A	N/A	919 859	6 507 036	5 587 177
10	24 898 744	303 923 280	106 109 240	885 120 096	779 010 856
11	15 838 532	189 938 649	135 106 755	747 637 719	612 530 964
12	9 339 224	131 826 704	88 949 193	759 494 357	670 545 164
13	46 731 037	420 512 324	105 237 126	780 877 121	675 639 995
14	64 574 302	536 492 360	78 089 102	819 301 528	741 212 426
15	108 611 419	661 763 708	113 967 619	797 421 790	683 454 171
16	88 649 743	362 443 432	135 825 774	718 073 190	582 247 416
17	6 243 137	74 147 210	84 324 103	707 204 059	622 879 956
18	43 734 856	509 876 714	72 223 833	972 811 873	900 588 040

Lane And Band Analysis

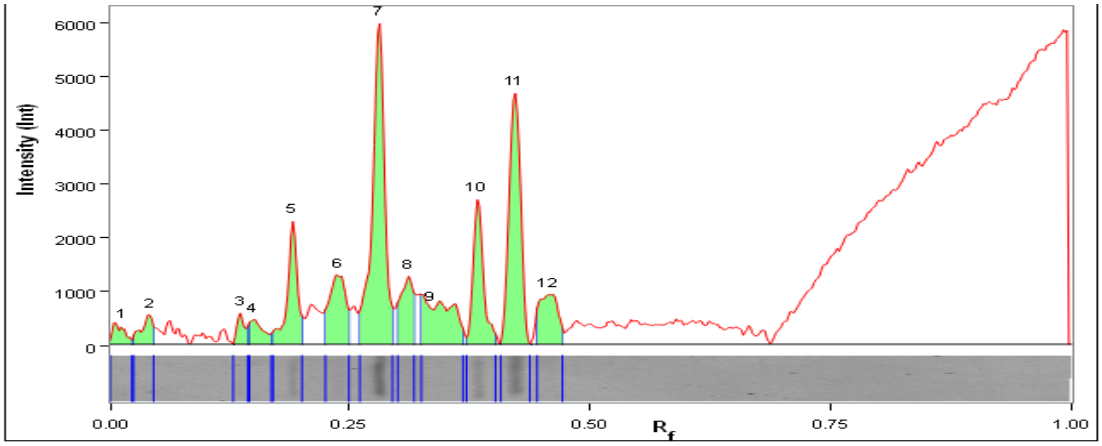
Lane 1 - BIOTECH STD



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	3,1
2	B	66,2	10,2
3	C	45,0	15,4
4	D	35,0	12,2
5	E	25,0	10,8
6	F	18,4	13,5
7	G	14,4	30,4
8		N/A	4,4

Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

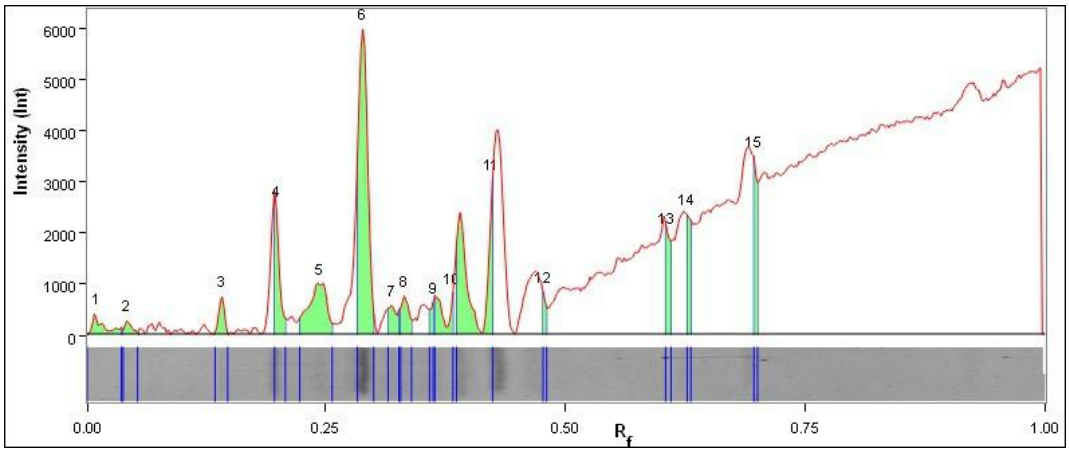
Lane 10



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	1,7
2		116,0	2,3
3		72,7	1,6
4		68,1	2,3
5		60,9	8,2
6		54,6	7,7
7		49,3	26,1
8		46,0	5,7
9		43,9	9,2
10		39,9	9,9
11	VP7	37,0	19,3
12		34,6	6,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	9.28 mm
Regression Equation	A single equation is not available for this method

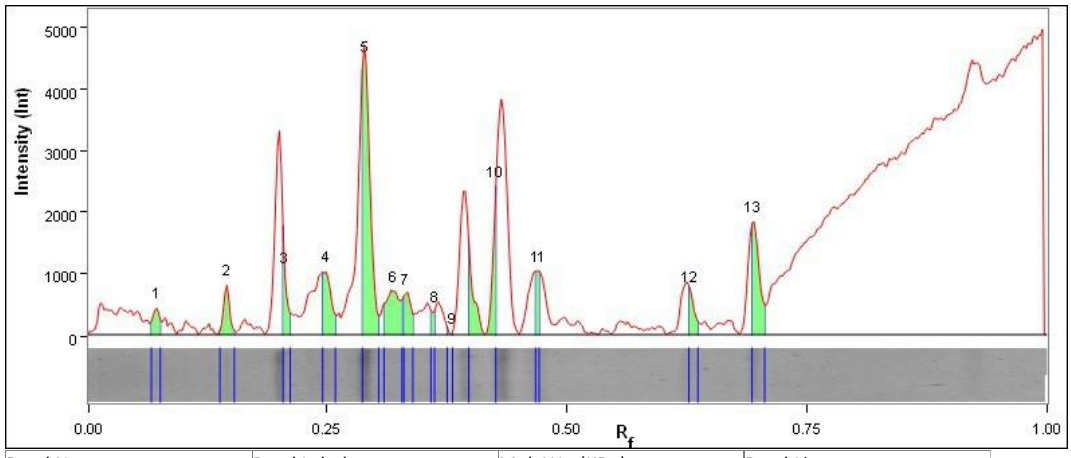
Lane 11



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	2,2
2		116,0	1,0
3		70,7	2,2
4		59,8	5,4
5		53,9	9,8
6		48,5	27,2
7		45,2	2,3
8		44,0	3,1
9		41,5	1,5
10		40,0	4,7
11	VP7	36,8	20,6
12		33,3	1,7
13		25,9	4,7
14		24,9	5,7
15		22,6	7,9

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

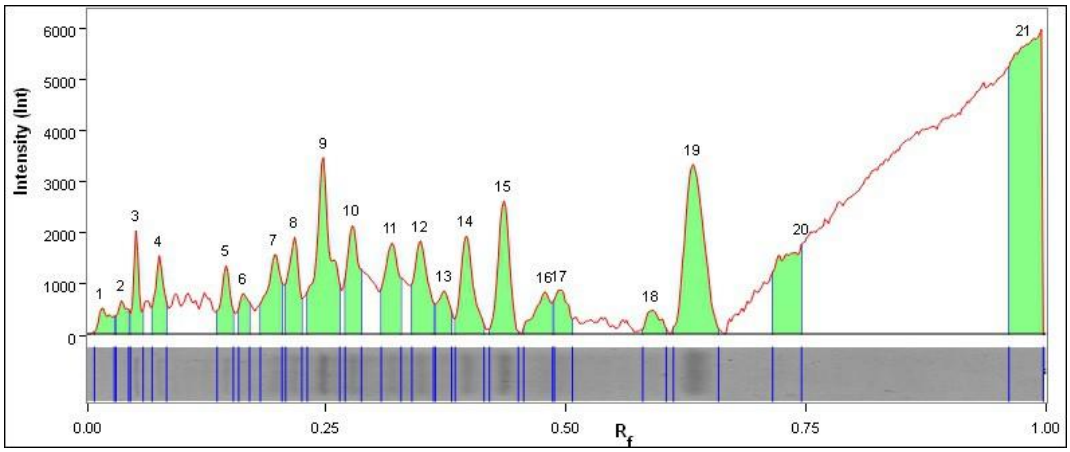
Lane 12



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		99,9	2,8
2		69,4	4,6
3		58,8	3,5
4		53,2	7,1
5		48,3	28,8
6		45,2	9,7
7		44,0	4,8
8		41,5	1,8
9		40,1	0,3
10	VP7	36,7	17,9
11		33,6	4,4
12		24,9	3,1
13		22,6	11,1

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

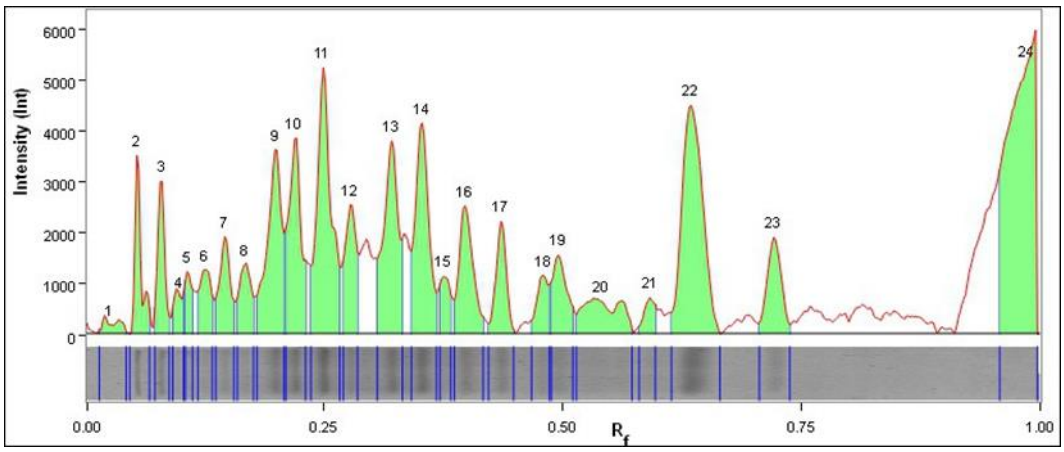
Lane 13



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	1,0
2		116,0	1,1
3		110,7	2,0
4		98,0	2,3
5		69,4	2,1
6		64,8	1,2
7		60,1	3,6
8		57,3	3,4
9		53,4	8,6
10		49,8	4,2
11		45,2	4,3
12		42,6	4,2
13		40,6	1,8
14		38,9	4,2
15		36,0	5,0
16		33,2	2,3
17		32,2	1,9
18		26,8	1,1
19		24,7	10,8
20		21,1	6,8
21		14,4	28,1

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

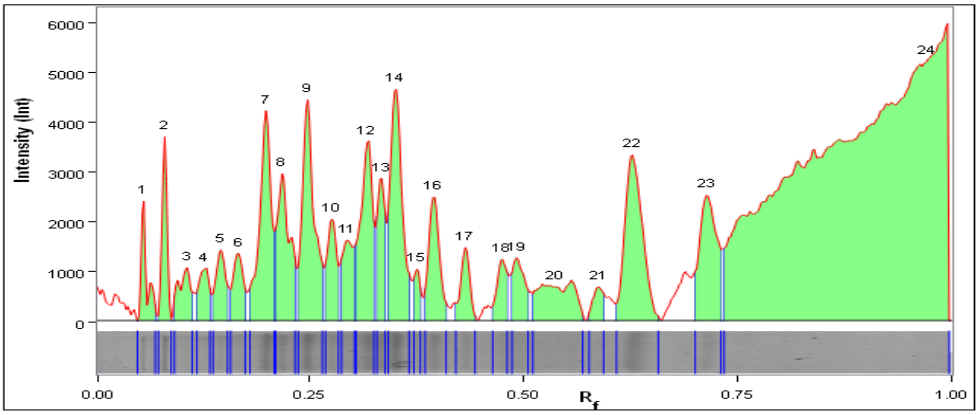
Lane 14



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	0,6
2		109,7	2,6
3		96,2	2,6
4		87,6	1,0
5		84,4	1,0
6		76,9	1,7
7		69,4	2,4
8		64,5	2,0
9		59,8	6,2
10		57,0	5,7
11		53,2	8,7
12		49,8	3,3
13		45,0	6,8
14		42,3	6,8
15		40,5	1,3
16		38,8	4,2
17		36,0	2,7
18		33,0	1,6
19		32,1	2,6
20		29,4	3,1
21		26,7	0,9
22		24,7	11,1
23		21,8	3,2
24		14,4	18,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

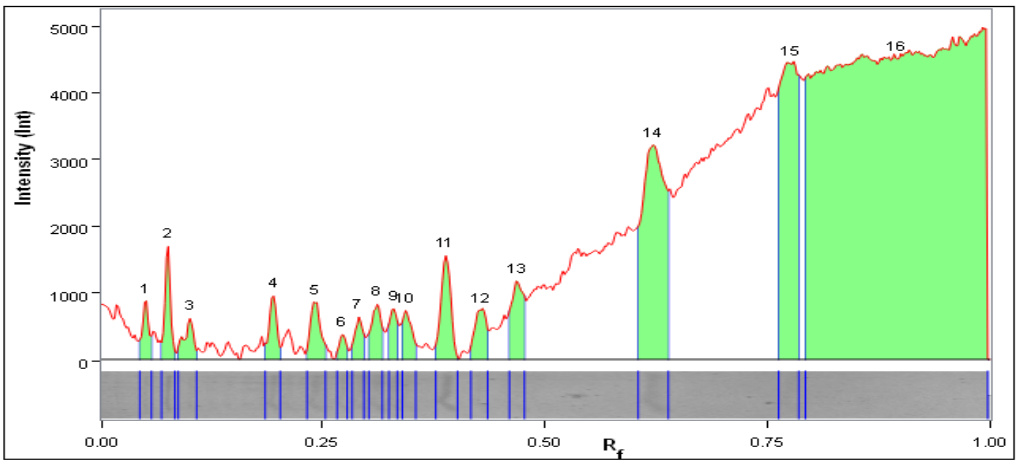
Lane 15



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		108,7	1,1
2		96,2	1,4
3		84,4	1,0
4		76,2	0,8
5		69,4	1,1
6		64,5	1,1
7		59,8	3,8
8		57,3	2,9
9		53,4	4,2
10		50,0	1,5
11		47,9	1,4
12		45,4	3,3
13		43,9	1,6
14		42,5	4,3
15		40,5	0,5
16		39,0	2,1
17		36,3	1,1
18		33,4	0,9
19		32,3	1,2
20		29,6	2,2
21		26,9	0,5
22		24,9	4,6
23		22,0	3,4
24		14,4	53,9

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

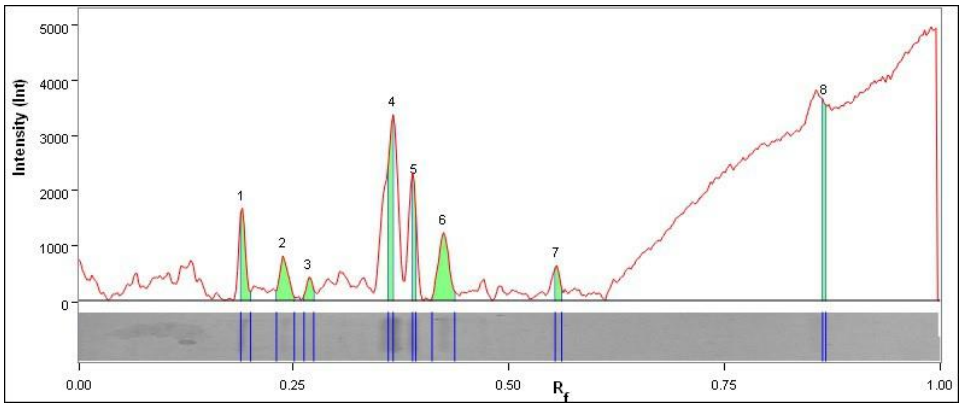
Lane 16



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		110,7	0,6
2		98,0	1,0
3		86,8	0,6
4		60,4	0,8
5		54,1	0,9
6		50,4	0,3
7		48,3	0,5
8		46,0	0,8
9		44,2	0,6
10		43,1	0,7
11		39,5	1,6
12		36,6	0,9
13		33,8	1,5
14		25,2	7,7
15		20,2	8,3
16		16,4	73,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

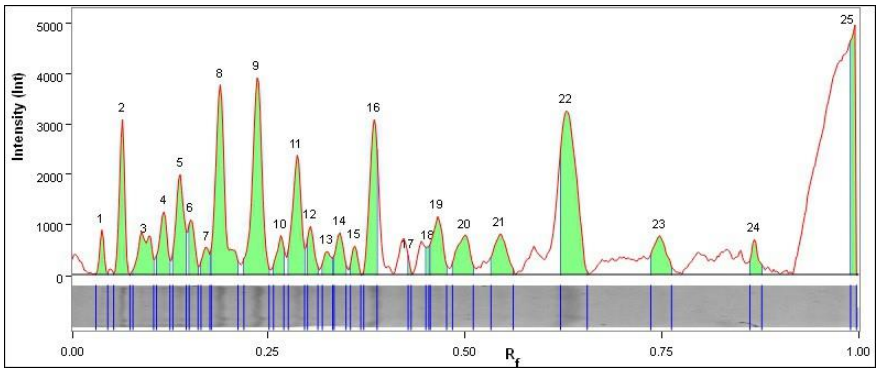
Lane 17



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		60,9	9,2
2		54,6	10,0
3		50,9	3,8
4		41,2	24,8
5		39,3	9,3
6	VP7	36,8	18,5
7		28,6	3,4
8		17,4	21,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.71 mm
Regression Equation	A single equation is not available for this method

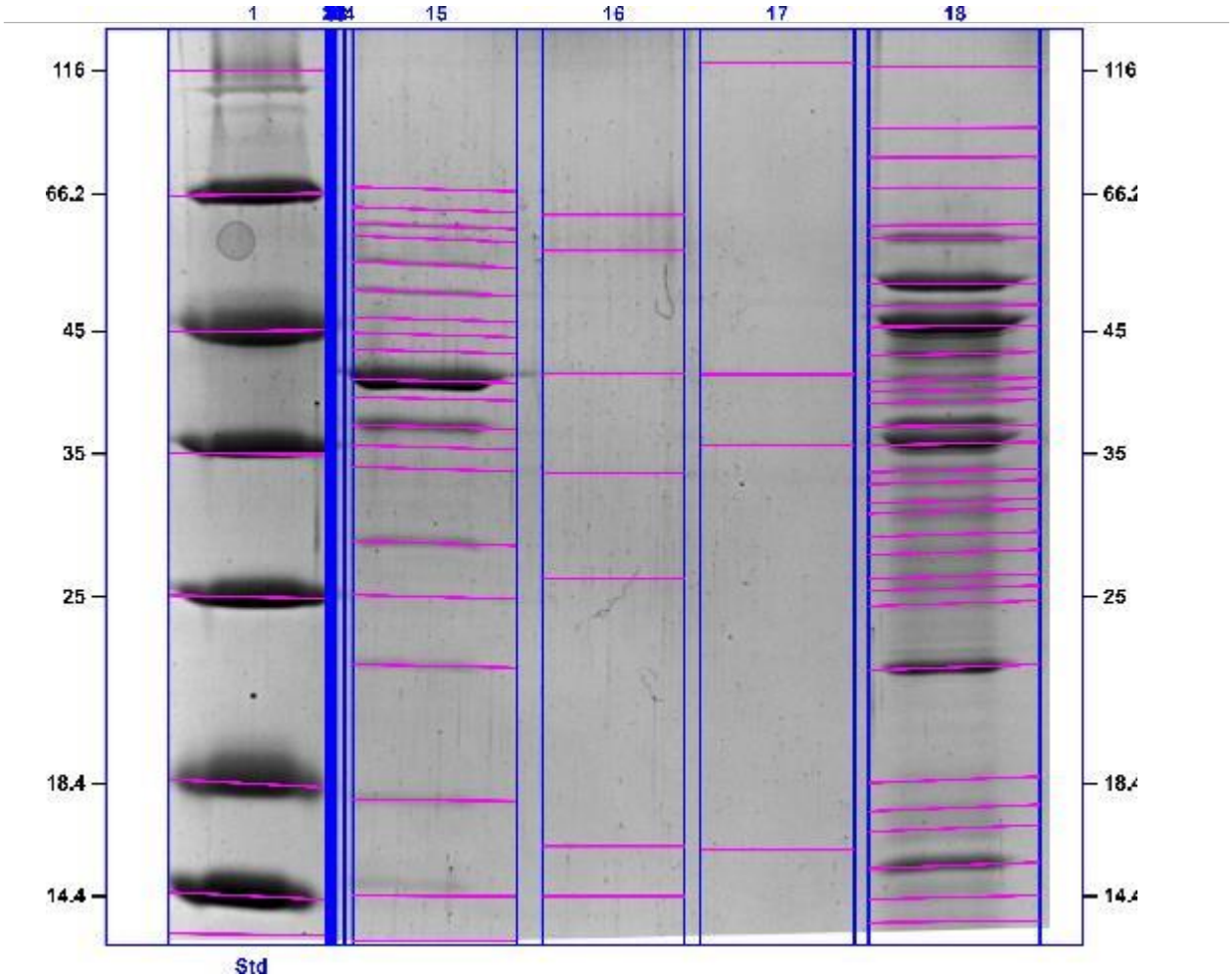
Lane 18



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	1,2
2		103,7	4,9
3		90,1	3,1
4		79,8	2,9
5		72,0	4,9
6		67,4	2,0
7		63,6	1,2
8		61,2	10,7
9		54,8	11,4
10		51,1	1,6
11		48,7	6,3
12		46,8	1,9
13		44,7	1,1
14		43,2	1,9
15		41,7	1,0
16		39,8	7,6
17		36,6	0,2
18		34,8	0,6
19		34,0	3,5
20		31,9	3,1
21		29,2	3,2
22		24,9	13,8
23		21,0	3,0
24		17,4	1,5
25		14,4	7,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	11.12 mm
Regression Equation	A single equation is not available for this method

Appendix 7.4: Analysis results for figure 3.12: Repeat VP7 ion-exchange purification by pH exchange



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.075 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2016/09/26 4:42:03 PM
User Name	User
Image Area (mm)	X: 47.1 Y: 48.7
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	0 - 39698

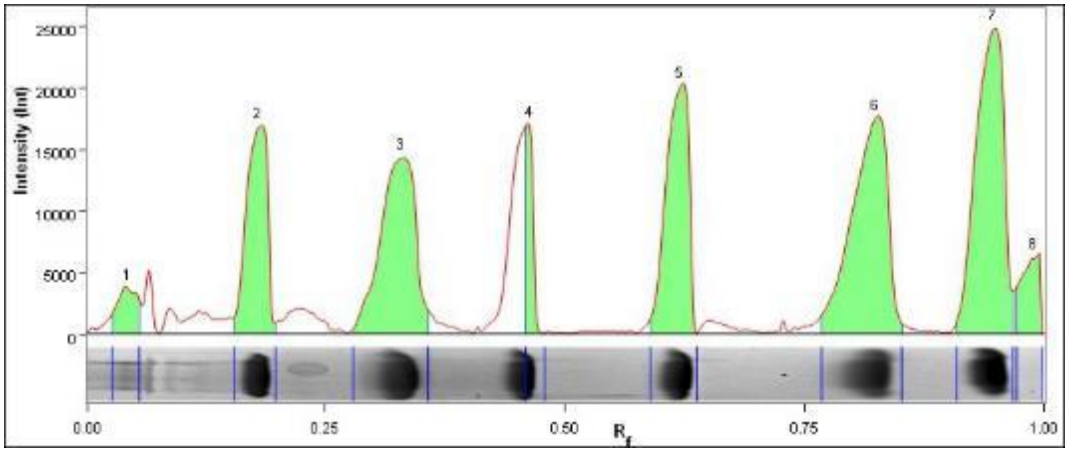
Analysis Settings

Detection	Lane detection: Manually created lanes Band detection: Bands detected with different sensitivity per lane Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	342 522 606	655 020 075	410 734 581	1 267 139 265	856 404 684
2	N/A	N/A	2 489 766	8 783 462	6 293 696
3	N/A	N/A	1 794 895	8 650 937	6 856 042
4	N/A	N/A	1 794 895	8 650 937	6 856 042
5	N/A	N/A	1 794 895	8 650 937	6 856 042
6	N/A	N/A	1 794 895	8 650 937	6 856 042
7	N/A	N/A	1 794 895	8 650 937	6 856 042
8	N/A	N/A	3 368 604	17 283 870	13 915 266
9	N/A	N/A	1 621 935	8 652 569	7 030 634
10	N/A	N/A	1 463 918	8 532 907	7 068 989
11	N/A	N/A	1 463 916	8 532 905	7 068 989
12	N/A	N/A	1 540 821	8 458 105	6 917 284
13	N/A	N/A	2 709 108	16 816 800	14 107 692
14	N/A	N/A	1 277 537	8 129 196	6 851 659
15	104 290 048	477 443 328	176 725 376	1 121 761 408	945 036 032
16	84 088 494	276 900 156	175 032 015	863 258 766	688 226 751
17	268 316 411	615 444 841	280 000 897	945 496 299	665 495 402
18	262 603 150	1 031 343 328	294 387 012	1 393 029 428	1 098 642 416

Lane And Band Analysis

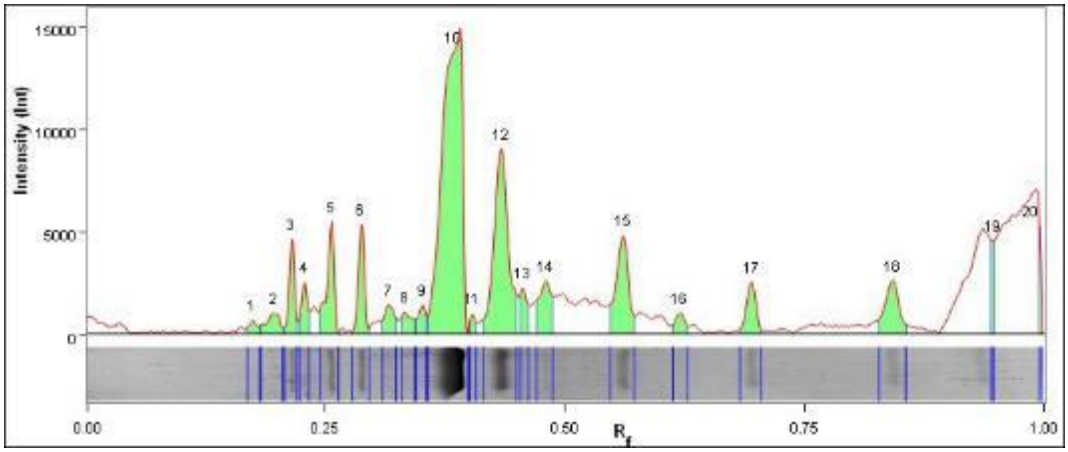


Lane 1 - BIOTECH STD

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	2,5
2	B	66,2	12,2
3	C	45,0	17,1
4	D	35,0	3,6
5	E	25,0	15,7
6	F	18,4	21,4
7	G	14,4	23,6
8		N/A	3,9

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.80 mm
Regression Equation	A single equation is not available for this method

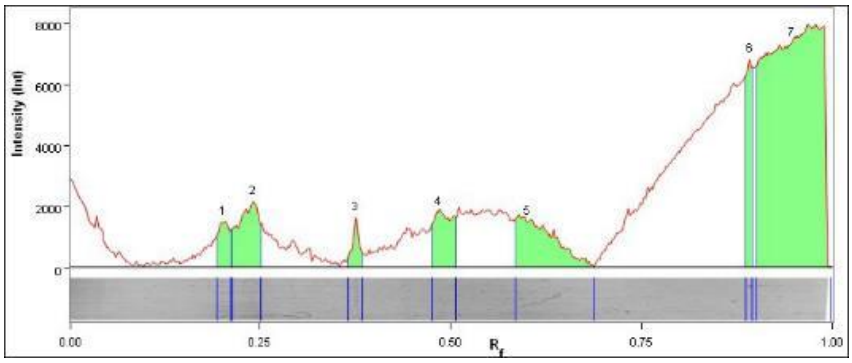
Lane 15



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		67,7	0,6
2		63,4	1,7
3		60,5	3,2
4		58,3	2,0
5		54,3	5,0
6		50,1	4,1
7		46,5	1,7
8		44,6	1,3
9		43,2	1,2
10		40,6	37,1
11		39,2	0,6
12	VP7	37,0	16,1
13		35,5	1,9
14		33,8	3,9
15		28,4	7,5
16		25,0	1,2
17		22,3	3,0
18		17,7	4,6
19		14,4	2,0
20		14,4	1,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.74 mm
Regression Equation	A single equation is not available for this method

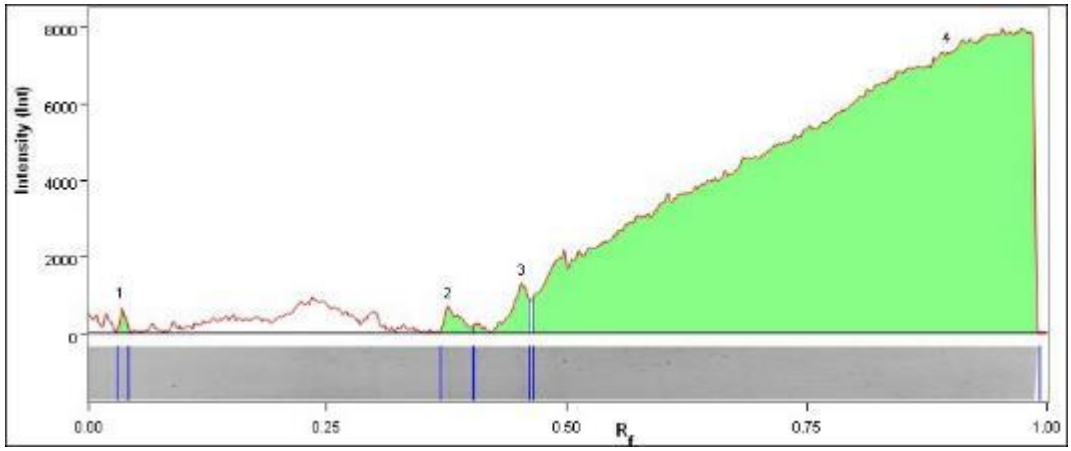
Lane 16



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		62,5	2,6
2		56,5	6,5
3		41,3	1,6
4		33,5	5,3
5		26,1	9,4
6		16,1	6,4
7		14,4	68,1

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.58 mm
Regression Equation	A single equation is not available for this method

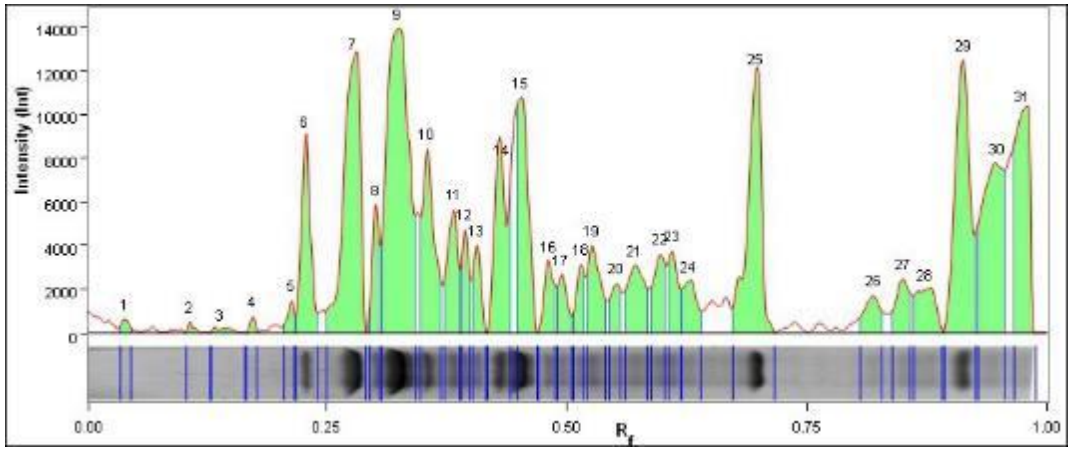
Lane 17



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	0,2
2		41,2	0,5
3		35,6	1,0
4		15,9	98,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.26 mm
Regression Equation	A single equation is not available for this method

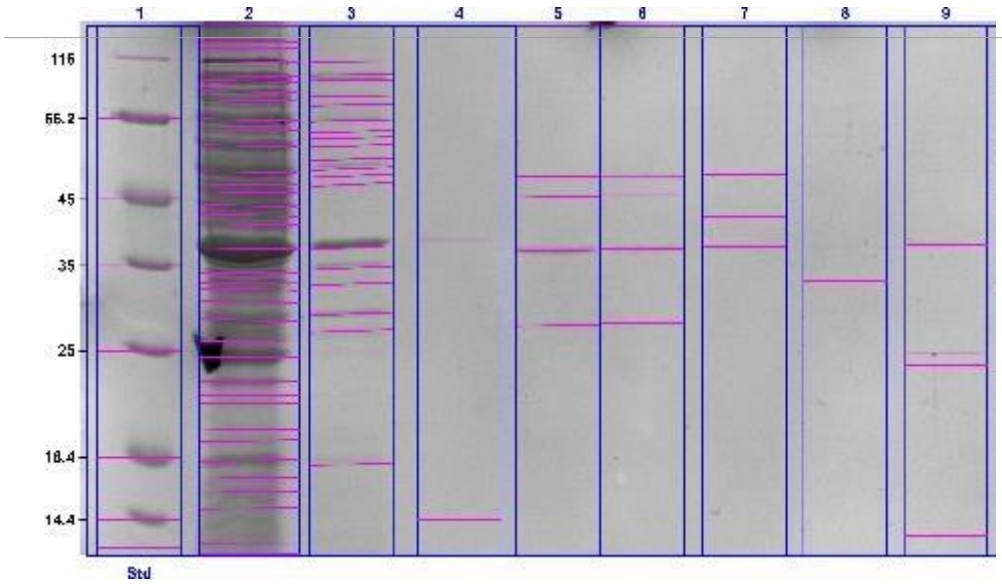
Lane 18



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	0,2
2		89,4	0,1
3		78,3	0,2
4		68,1	0,2
5		60,7	0,4
6		58,6	3,8
7		51,4	9,7
8		48,4	1,7
9		45,7	14,1
10		42,9	4,6
11		40,7	2,7
12		39,9	1,4
13		39,0	1,4
14	VP7	37,1	5,4
15		35,7	5,0
16		33,7	1,4
17		32,8	1,1
18		31,3	1,1
19		30,6	2,3
20		29,0	0,9
21		27,8	2,3
22		26,2	1,7
23		25,5	1,6
24		24,7	1,5
25		22,3	7,7
26		18,5	1,1
27		17,5	1,4
28		16,7	1,9
29		15,4	8,5
30		14,4	7,7
31		14,4	6,7

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	9.15 mm
Regression Equation	A single equation is not available for this method

Appendix 7.5: Analysis results for figure 3.13: VP7 ion-exchange purification by pH gradient



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.080 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2017/03/23 4:31:04 PM
User Name	User
Image Area (mm)	X: 90.5 Y: 46.2
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	0 - 44663

Analysis Settings

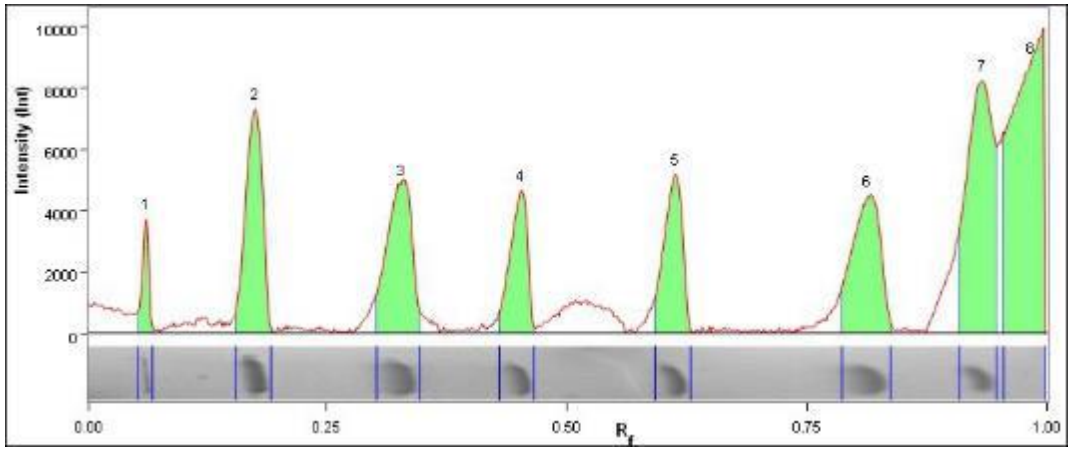
Detection	Lane detection: Manually created lanes Band detection: Automatically detected bands with sensitivity: High Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10
	Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	102 644 146	310 316 590	125 453 226	872 581 718	747 128 492
2	223 709 992	1 439 384 091	298 445 301	1 825 222 664	1 526 777 363
3	17 514 274	205 500 504	74 940 622	937 990 396	863 049 774
4	216 179 050	511 809 552	221 487 742	856 080 062	634 592 320
5	13 269 186	87 251 250	102 343 742	834 209 718	731 865 976
6	13 173 150	90 938 566	60 890 428	776 190 194	715 299 766
7	3 236 180	35 932 622	135 160 176	763 646 578	628 486 402
8	1 296 910	7 330 642	135 382 564	725 213 204	589 830 640
9	90 607 104	264 059 568	102 653 792	681 682 714	579 028 922
10	63 111 559	383 712 980	108 402 049	922 134 087	813 732 038

Lane and Band Analysis

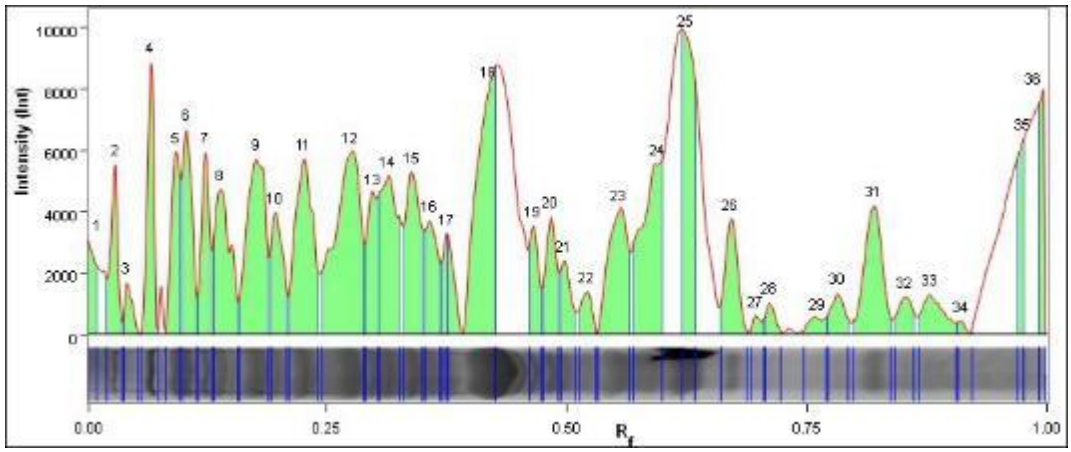
Lane 1 - BIOTECH STD



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	2,3
2	B	66,2	11,5
3	C	45,0	11,4
4	D	35,0	7,4
5	E	25,0	8,8
6	F	18,4	11,7
7	G	14,4	20,0
8		N/A	26,9

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

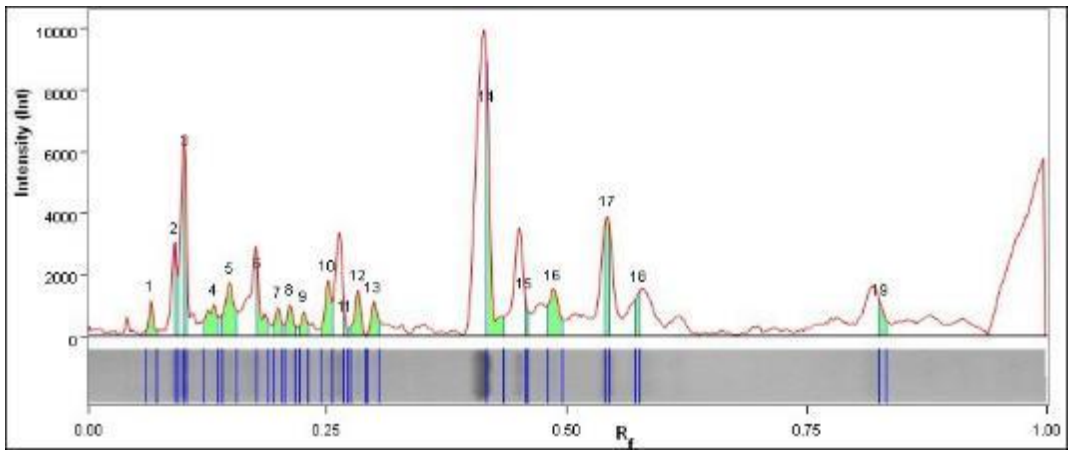
Lane 2



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	1,2
2		116,0	2,1
3		116,0	0,6
4		113,5	3,0
5		99,3	2,6
6		94,3	3,3
7		85,7	2,5
8		79,0	3,8
9		65,9	5,6
10		62,8	2,3
11		58,2	5,2
12		51,4	7,9
13		48,4	2,6
14		46,6	4,4
15		43,9	3,9
16		42,4	2,3
17		41,0	0,9
18	VP7	37,4	9,4
19		34,0	1,6
20		32,8	2,0
21		31,9	1,2
22		30,3	0,8
23		28,2	4,1
24		26,0	5,9
25		24,5	5,9
26		22,9	2,5
27		22,0	0,3
28		21,5	0,5
29		19,9	0,5
30		19,3	0,9
31		18,2	3,9
32		17,0	0,9
33		16,1	1,4
34		15,1	0,2
35		14,4	2,0
36		14,4	2,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.67 mm
Regression Equation	A single equation is not available for this method

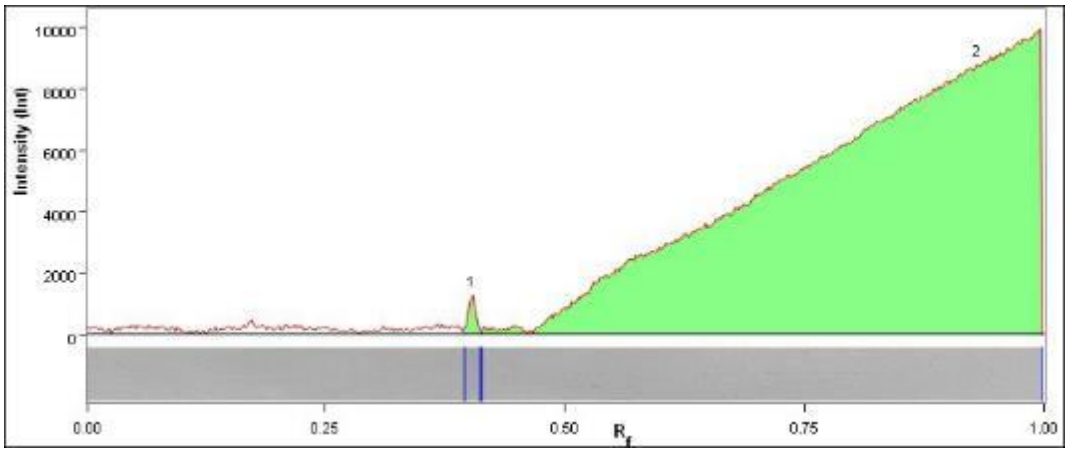
Lane 3



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		113,5	2,9
2		100,1	5,0
3		95,0	7,0
4		82,0	5,6
5		75,6	8,7
6		65,7	4,4
7		62,3	3,0
8		60,5	3,5
9		58,2	2,5
10		54,6	6,5
11		52,2	0,7
12		50,4	5,1
13		48,4	4,0
14	VP7	37,5	14,9
15		34,7	2,0
16		32,6	8,3
17		28,9	9,9
18		27,0	3,8
19		17,9	2,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

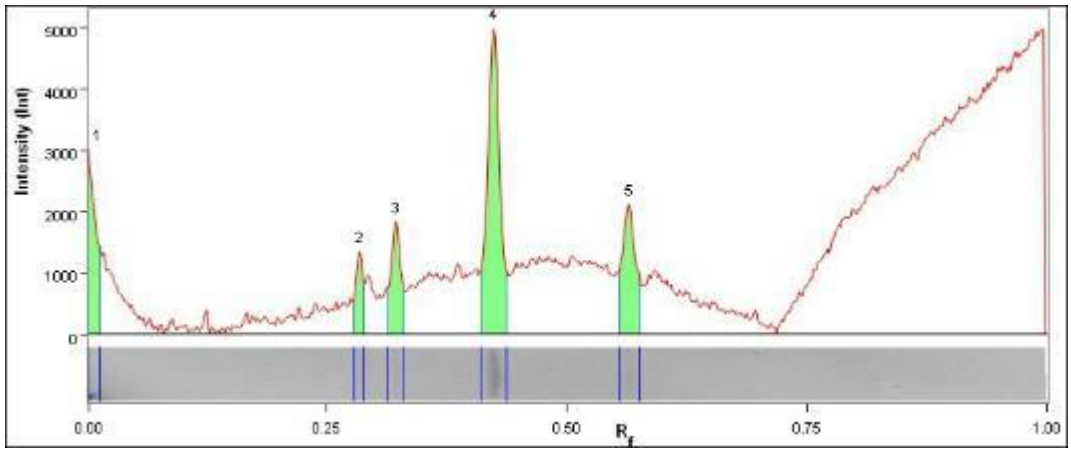
Lane 4



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		38,6	0,4
2		14,4	99,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

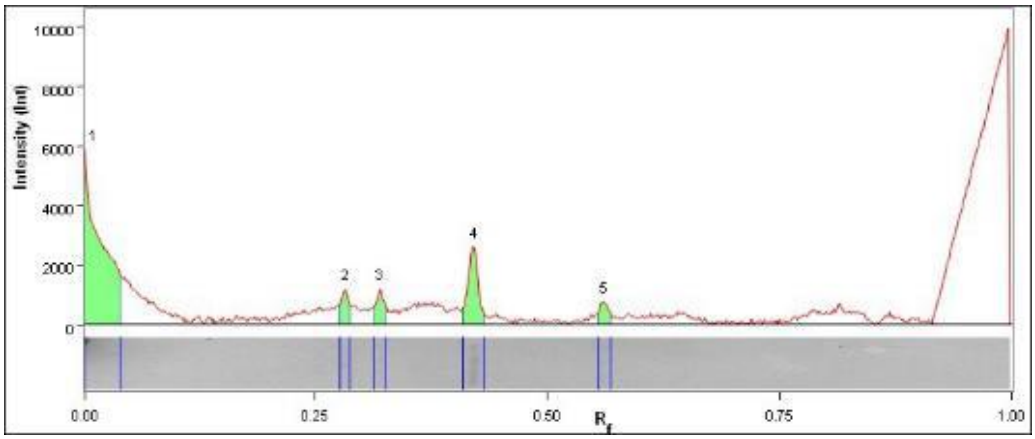
Lane 5



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	16,5
2		50,2	6,9
3		45,5	12,7
4	VP7	37,0	44,4
5		27,6	19,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

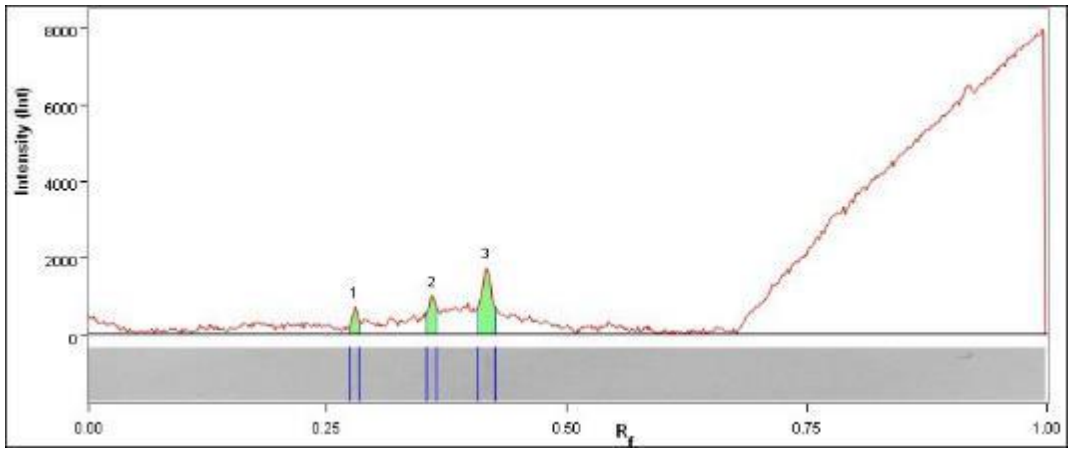
Lane 6



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	64,8
2		50,4	5,9
3		45,9	6,2
4	VP7	37,3	18,4
5		27,9	4,7

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

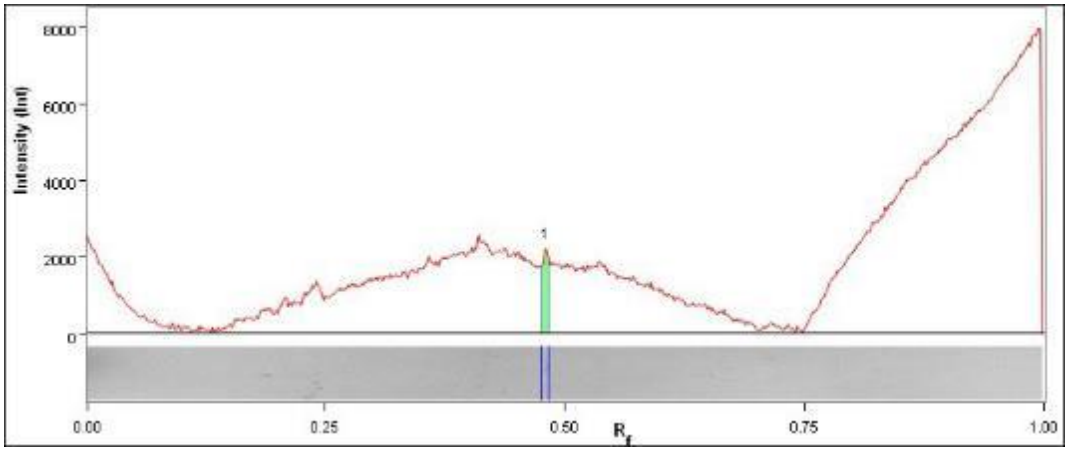
Lane 7



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		50,8	14,2
2		42,1	25,2
3	VP7	37,6	60,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

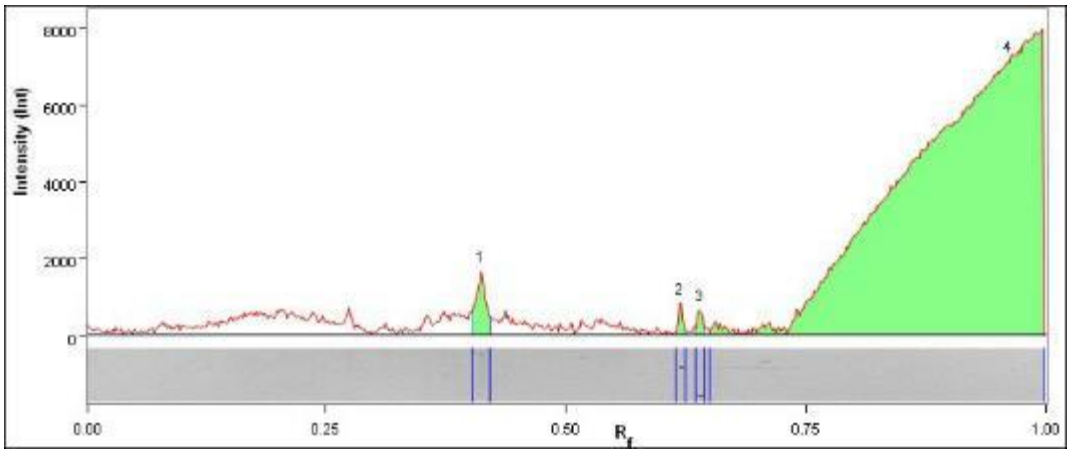
Lane 8



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		33,0	100,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

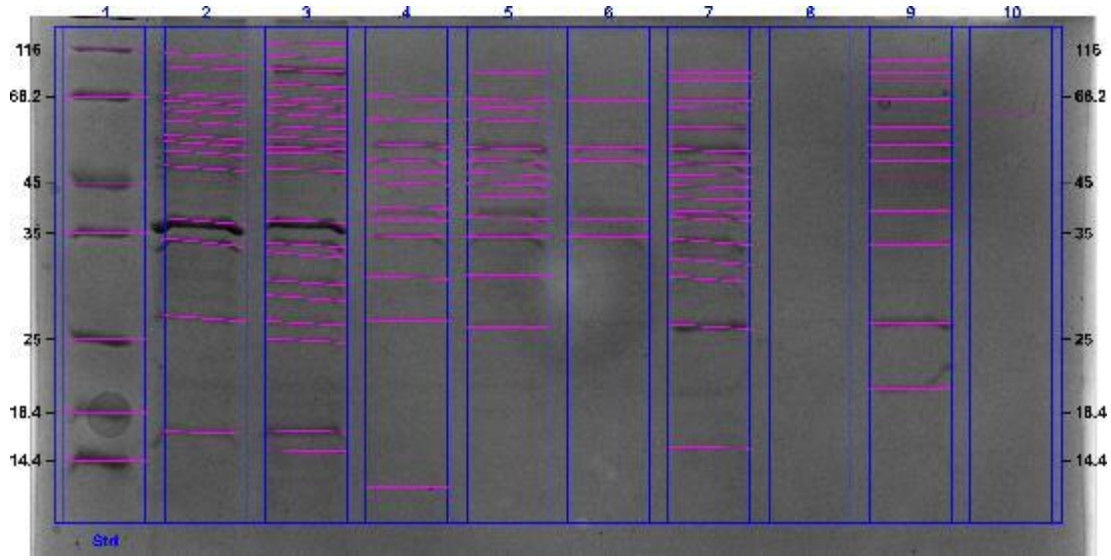
Lane 9



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		37,9	12,6
2		24,8	10,4
3		24,0	10,4
4		14,4	66,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.23 mm
Regression Equation	A single equation is not available for this method

Appendix 7.6: Analysis results for figure 3.14: VP7 ion-exchange purification by two pH exchange



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.060 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2017/12/07 5:40:17 PM
User Name	User
Image Area (mm)	X: 89.0 Y: 45.4
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	0 - 35995

Analysis Settings

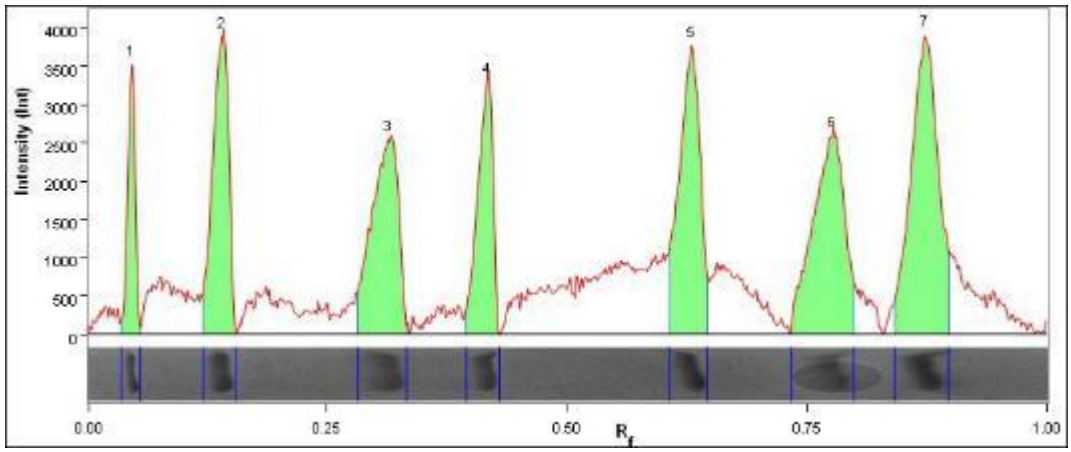
Detection	Lane detection: Manually created lanes Band detection: Automatically detected bands with sensitivity: High Manually adjusted bands
	Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: 6.82 mm
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	35 218 600	305 530 900	54 647 400	920 445 700	865 798 300
2	21 845 600	296 016 100	44 579 600	1 002 715 900	958 136 300
3	32 627 500	441 670 900	49 295 600	1 007 926 500	958 630 900
4	4 779 300	135 327 400	22 265 200	947 181 600	924 916 400
5	5 030 500	151 963 700	19 191 000	954 188 700	934 997 700
6	4 801 700	88 373 700	30 613 300	949 598 400	918 985 100
7	13 163 500	279 246 500	27 713 500	983 450 200	955 736 700
8	N/A	N/A	17 013 900	962 712 000	945 698 100
9	14 418 600	137 617 300	75 110 100	982 475 600	907 365 500
10	15 109 200	201 372 900	49 629 900	932 215 200	882 585 300

Lane And Band Analysis

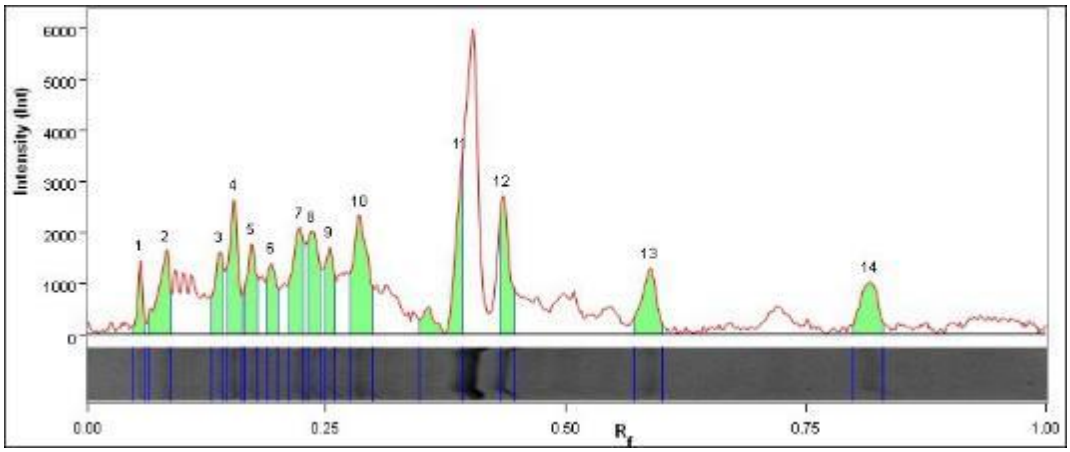
Lane 1 - BIOTECH STD



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	5,5
2	B	66,2	13,3
3	C	45,0	14,5
4	D	35,0	10,7
5	E	25,0	16,5
6	F	18,4	17,5
7	G	14,4	22,1

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

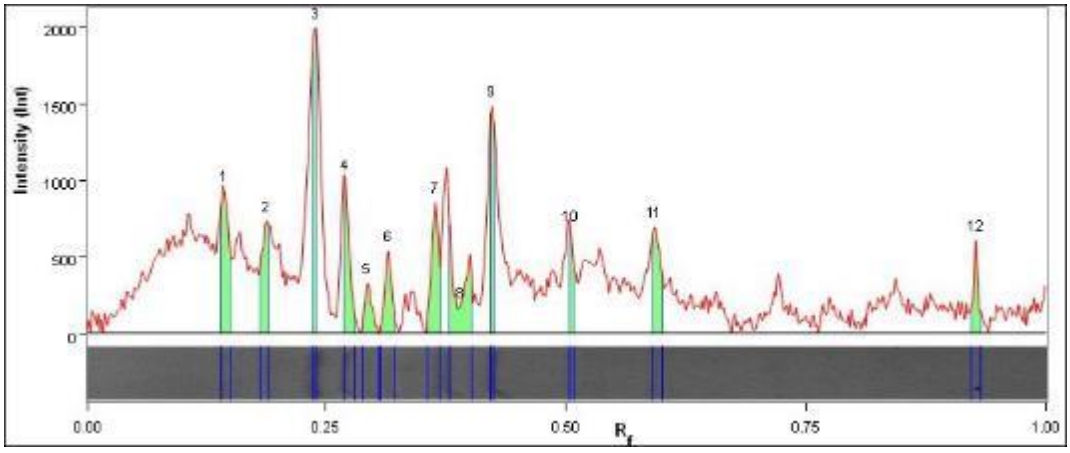
Lane 2



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		109,3	2,6
2		93,4	6,8
3		66,9	4,8
4		64,3	8,5
5		61,7	5,9
6		58,9	4,6
7		55,1	8,0
8		53,5	7,9
9		51,4	5,0
10		47,9	12,2
11	VP7	37,2	12,3
12		34,0	7,9
13		26,7	6,8
14		16,6	6,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

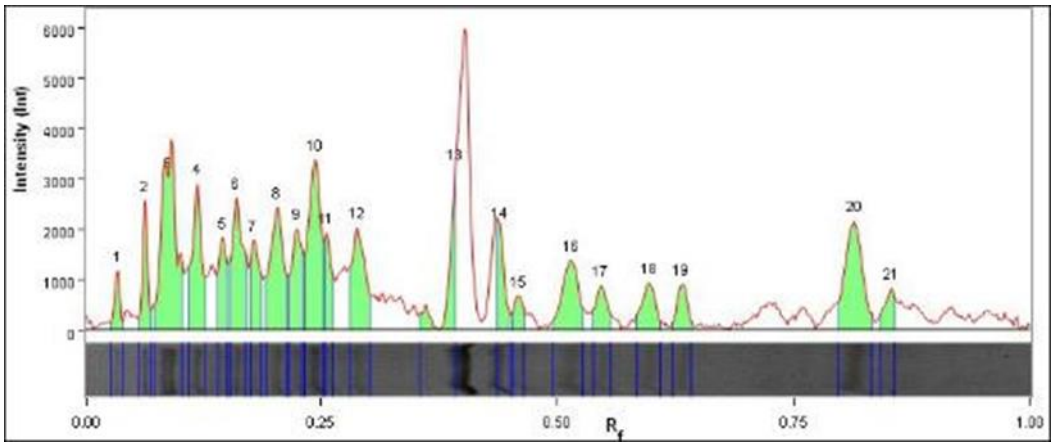
Lane 3



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		65,7	10,7
2		59,5	10,3
3		53,2	13,8
4		49,5	8,0
5		47,0	3,9
6		44,8	6,2
7		39,8	10,8
8	VP7	37,3	10,5
9		34,6	9,5
10		30,4	3,7
11		26,5	7,7
12		14,4	4,8

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

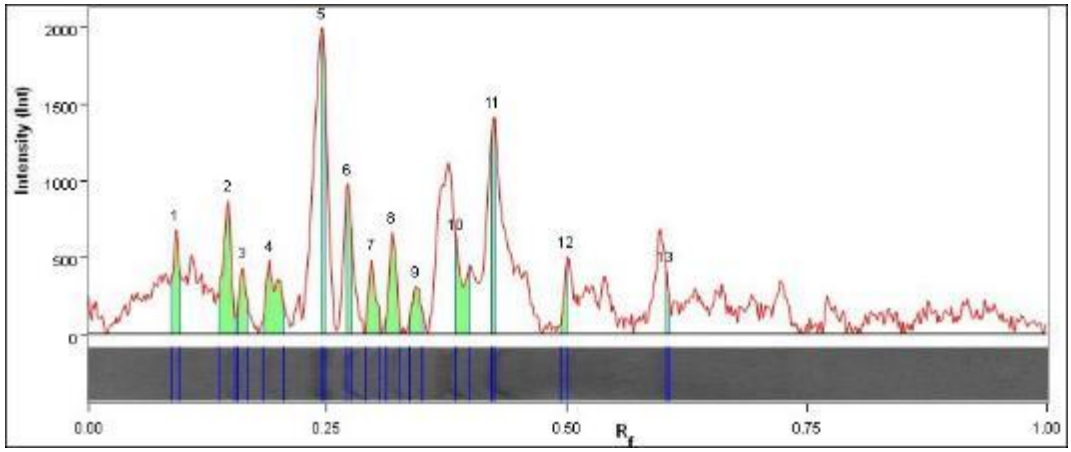
Lane 4



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	1,5
2		105,1	3,0
3		90,7	12,5
4		75,2	6,5
5		65,5	3,3
6		63,3	6,4
7		60,8	3,6
8		57,6	7,2
9		54,9	4,4
10		52,5	9,8
11		51,2	2,9
12		47,6	6,1
13	VP7	37,1	5,7
14		33,8	2,9
15		32,7	1,3
16		30,0	5,1
17		28,5	2,2
18		26,3	2,8
19		24,9	2,1
20		16,7	8,6
21		15,1	1,8

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

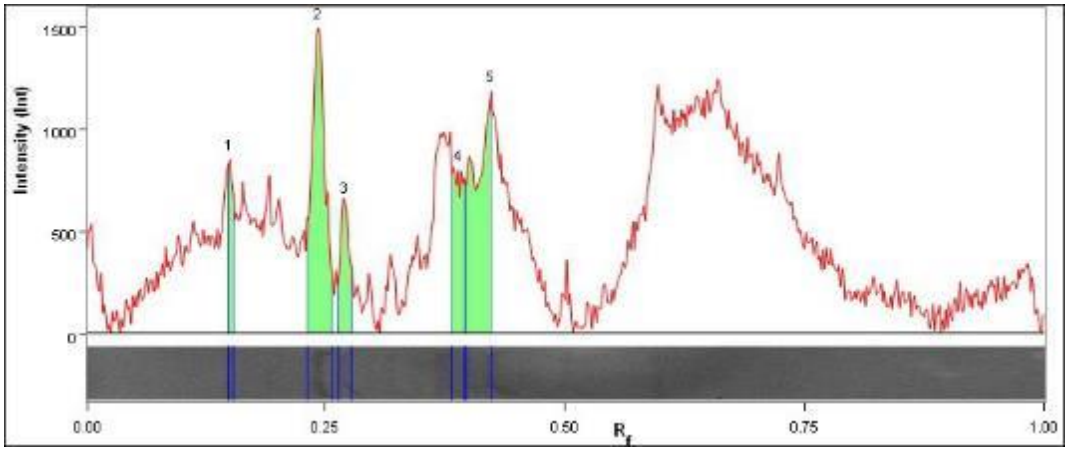
Lane 5



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		88,1	7,1
2		65,2	12,7
3		63,1	5,1
4		59,3	10,1
5		52,3	12,5
6		49,3	7,1
7		46,7	6,1
8		44,5	8,4
9		41,9	4,8
10	VP7	37,8	9,7
11		34,6	9,7
12		30,7	4,4
13		26,1	2,2

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

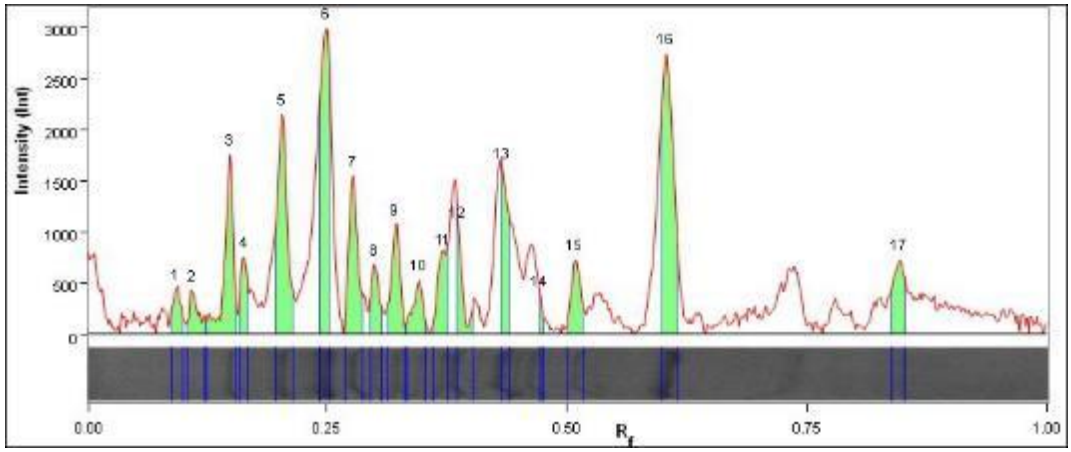
Lane 6



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		64,8	5,2
2		52,7	34,8
3		49,5	9,8
4	VP7	37,5	13,9
5		34,6	36,2

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

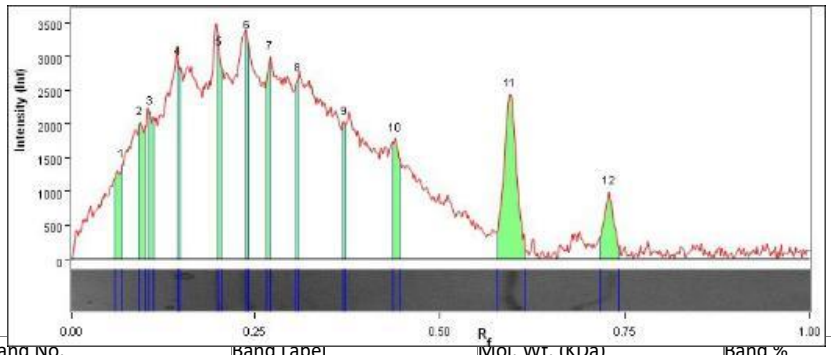
Lane 7



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		88,1	2,0
2		79,8	2,2
3		65,0	8,6
4		62,9	2,8
5		57,6	11,2
6		51,9	14,8
7		48,8	7,9
8		46,3	3,0
9		44,1	5,4
10		41,5	2,9
11		39,0	4,6
12	VP7	37,6	3,1
13		34,1	6,2
14		32,1	0,6
15		30,3	3,9
16		26,1	16,4
17		15,4	4,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

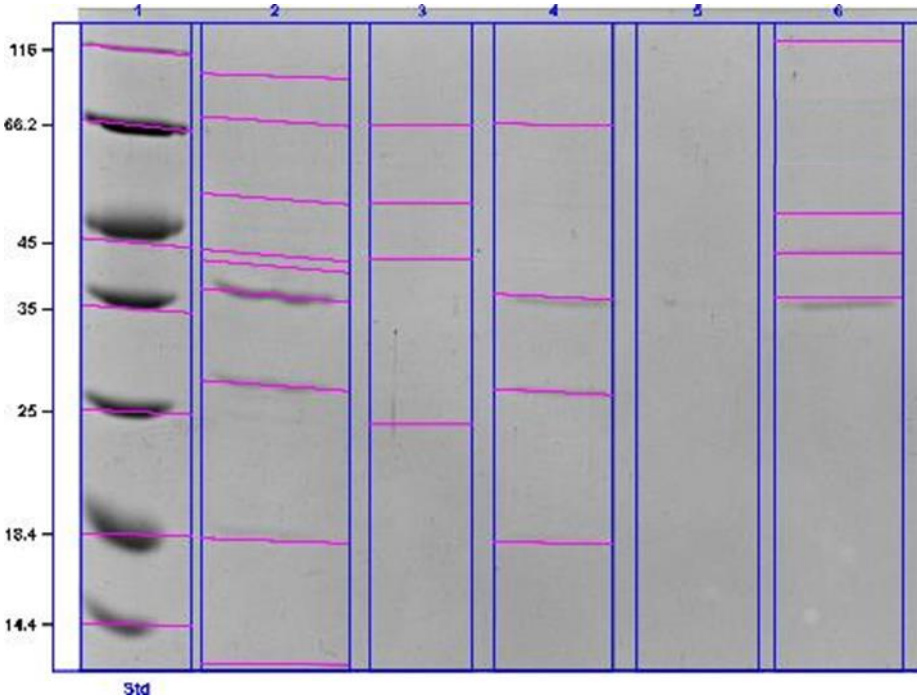
Lane 9



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		101,5	5,8
2		87,2	7,3
3		80,6	7,8
4		65,5	6,4
5		57,8	6,5
6		53,1	7,1
7		49,5	10,5
8		45,4	6,0
9		39,1	6,0
10		33,8	7,3
11		26,4	22,4
12		20,3	6,9

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.82 mm
Regression Equation	A single equation is not available for this method

Appendix 7.7: Analysis results for figure 3.15 VP7 HisTag purification of ion exchange gradient fraction



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.089 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2017/04/05 4:45:38 PM
User Name	User
Image Area (mm)	X: 63.5 Y: 42.9
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	6410 - 36076

Analysis Settings

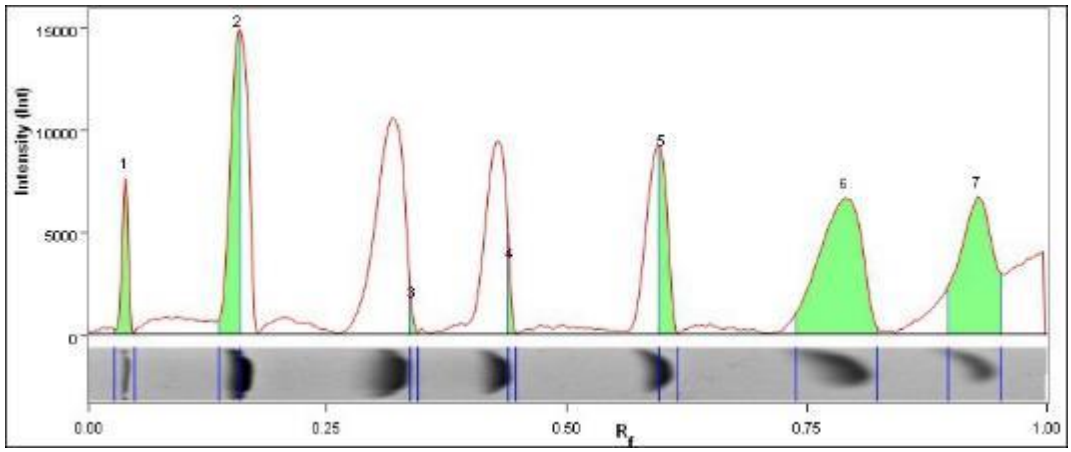
Detection	Lane detection: Manually created lanes Band detection:
	Automatically detected bands with sensitivity: High Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	78 933 540	217 034 790	182 505 015	812 490 525	629 985 510
2	39 973 641	397 179 516	66 246 735	982 948 365	916 701 630
3	4 535 832	94 646 440	117 126 366	683 044 908	565 918 542
4	3 652 446	57 625 746	118 099 668	799 542 762	681 443 094
5	N/A	N/A	142 475 800	827 911 220	685 435 420
6	3 744 180	85 460 878	109 143 152	860 419 030	751 275 878
7	297 153 675	734 885 025	297 153 675	734 885 025	437 731 350

Lane and Band Analysis

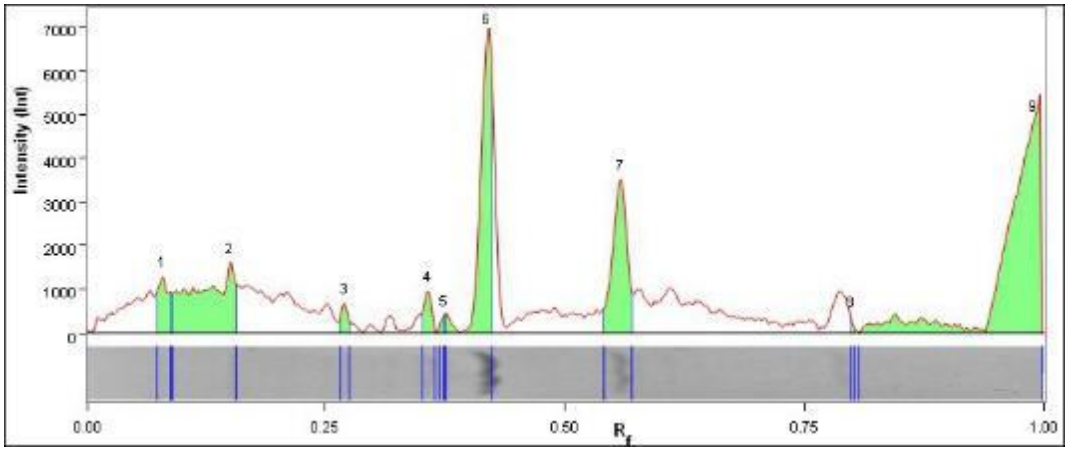
Lane 1 - BIOTECH STD



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	6,1
2	B	66,2	19,9
3	C	45,0	0,5
4	D	35,0	1,0
5	E	25,0	8,9
6	F	18,4	36,0
7	G	14,4	27,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.17 mm
Regression Equation	A single equation is not available for this method

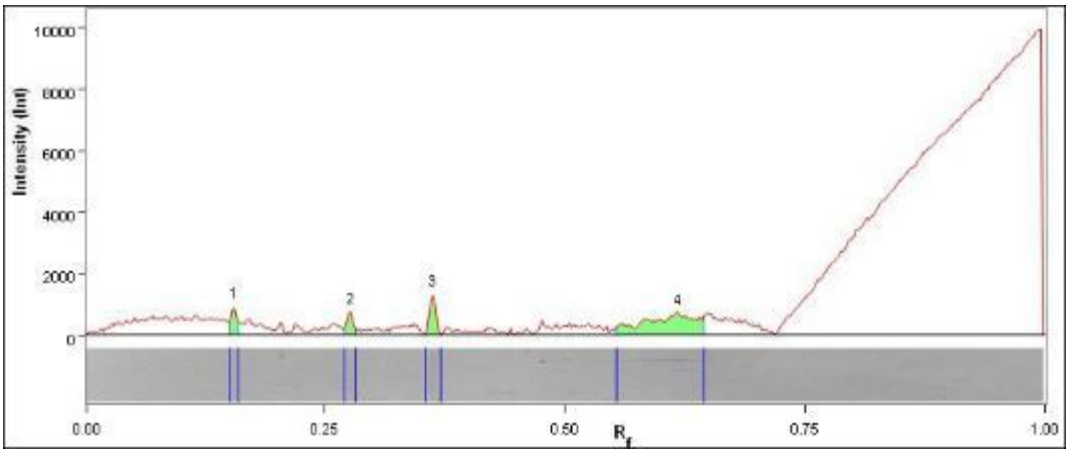
Lane 2



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		95,5	3,7
2		68,3	15,2
3		52,0	1,2
4		42,9	2,0
5		41,2	0,6
6	VP7	36,9	22,3
7		27,2	13,4
8		18,1	0,2
9		14,4	41,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	9.62 mm
Regression Equation	A single equation is not available for this method

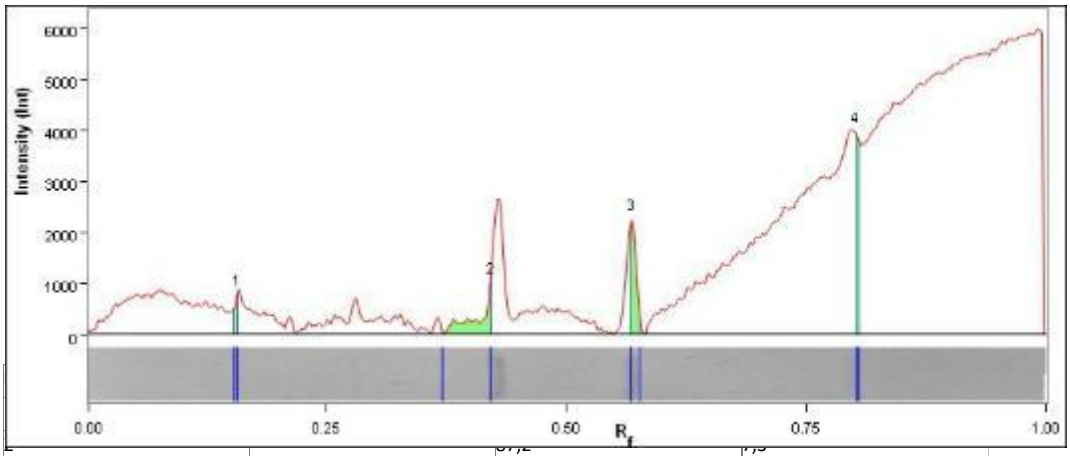
Lane 3



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		66,7	10,7
2		51,3	8,9
3		42,4	15,0
4		24,2	65,5

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	6.69 mm
Regression Equation	A single equation is not available for this method

Lane 4



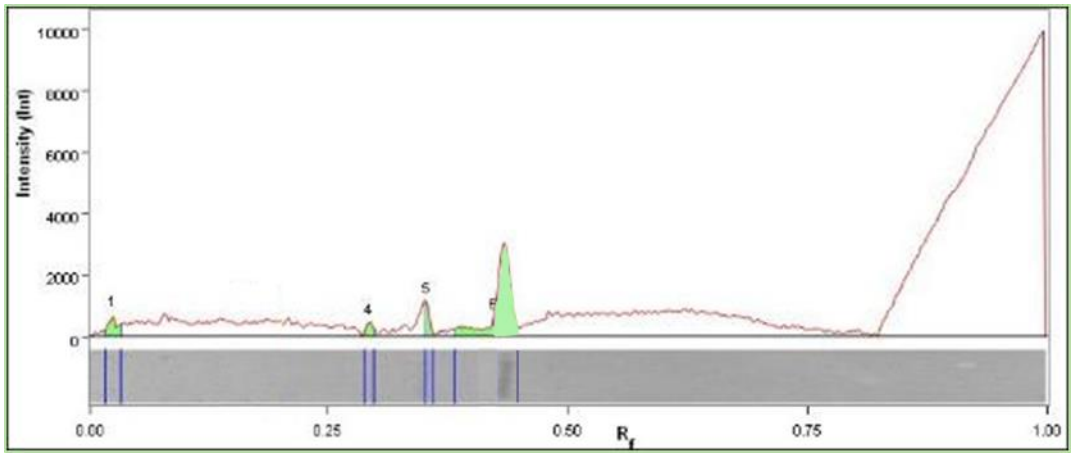
3		80,6	7,8
4		65,5	6,4
5		57,8	6,5
6		53,1	7,1
7		49,5	10,5
8		45,4	6,0
9		39,1	6,0
10		33,8	7,3
11		26,4	22,4
12		20,3	6,9

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		66,6	5,9
2	VP7	36,7	30,8
3		26,7	26,5
4		18,0	36,8

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.78 mm
Regression Equation	A single equation is not available for this method

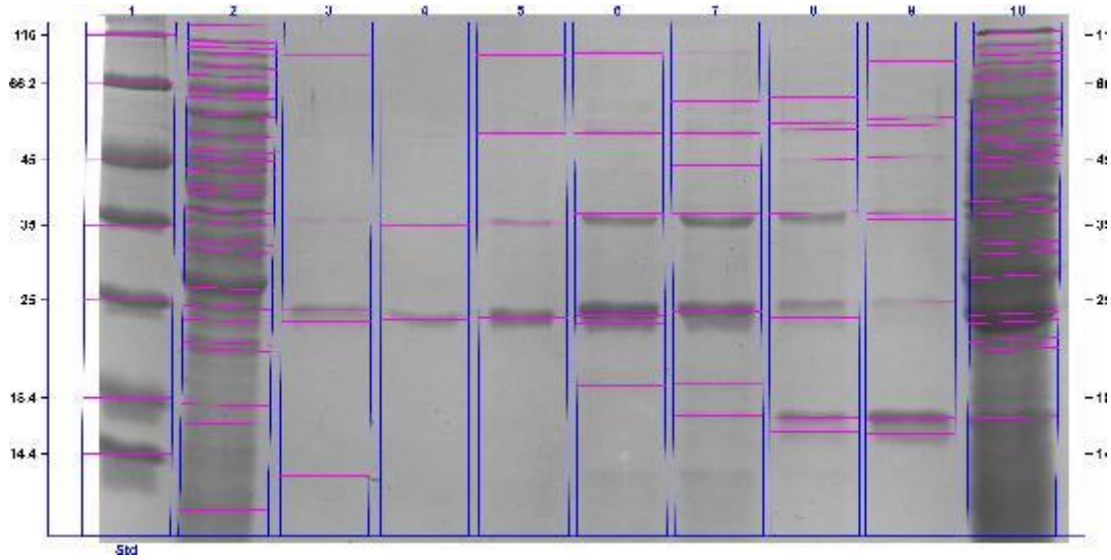
Lane Width	7.85 mm
Regression Equation	A single equation is not available for this method

Lane 6



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	8.1
2		49,0	3.7
3		43,4	18.5
4	VP7	36,6	69,7

Appendix 7.8 Analysis results for figure 3.16: VP7 HisTag purification of supernatant



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.059 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2016/08/18 5:23:36 PM
User Name	User
Image Area (mm)	X: 92.2 Y: 50.0
Pixel Size (µm)	X: 68.2 Y: 68.2
Data Range (Int)	0 - 26482

Analysis Settings

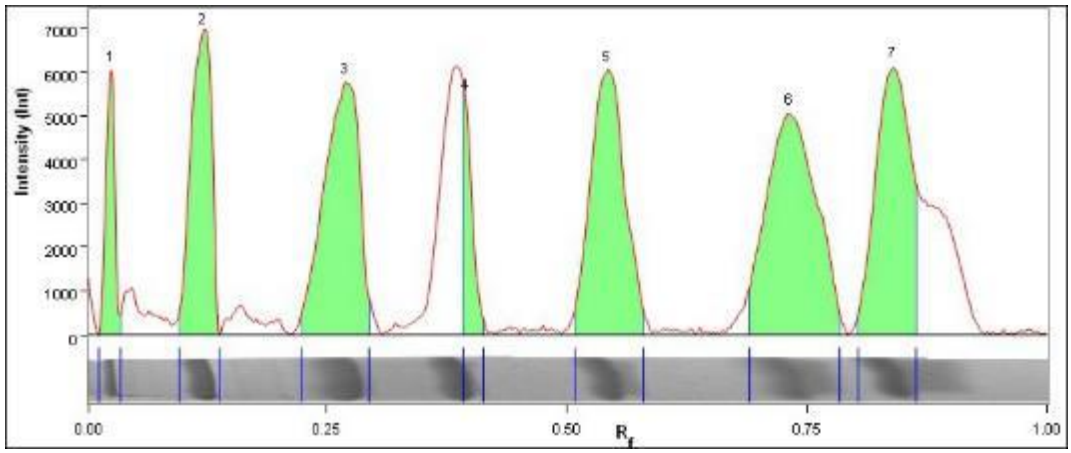
Detection	Lane detection: Automatically detected lanes with manual adjustments Band detection: Automatically detected bands with sensitivity: High Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10
	Lane width: 8.46 mm
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	121 403 812	437 470 512	158 364 616	965 599 284	807 234 668
2	68 632 760	829 750 960	112 228 184	1 446 755 244	1 334 527 060
3	11 535 348	116 526 272	32 187 300	1 024 886 412	992 699 112
4	7 547 384	70 147 420	22 197 736	978 936 848	956 739 112
5	19 885 632	116 580 088	47 528 580	1 005 799 960	958 271 380
6	15 489 460	161 820 124	82 202 576	1 040 017 884	957 815 308
7	22 670 424	282 076 068	74 933 696	1 044 593 360	969 659 664
8	13 599 080	163 033 588	54 889 840	960 351 728	905 461 888
9	18 201 340	148 748 168	59 540 088	984 436 620	924 896 532
10	53 080 804	538 258 828	124 220 596	1 666 265 128	1 542 044 532

Lane and Band Analysis

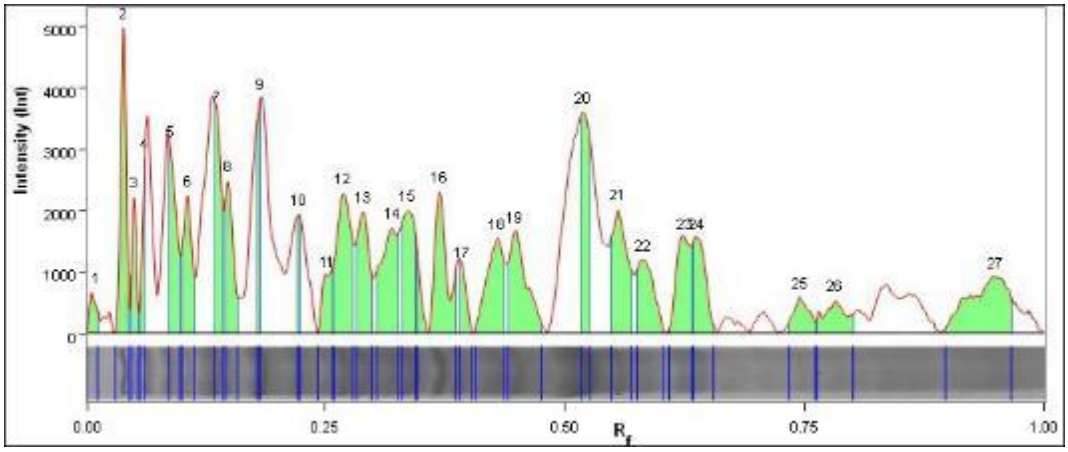
Lane 1 - BIOTECH STD



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	4,8
2	B	66,2	13,1
3	C	45,0	18,3
4	D	35,0	4,0
5	E	25,0	18,6
6	F	18,4	22,6
7	G	14,4	18,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

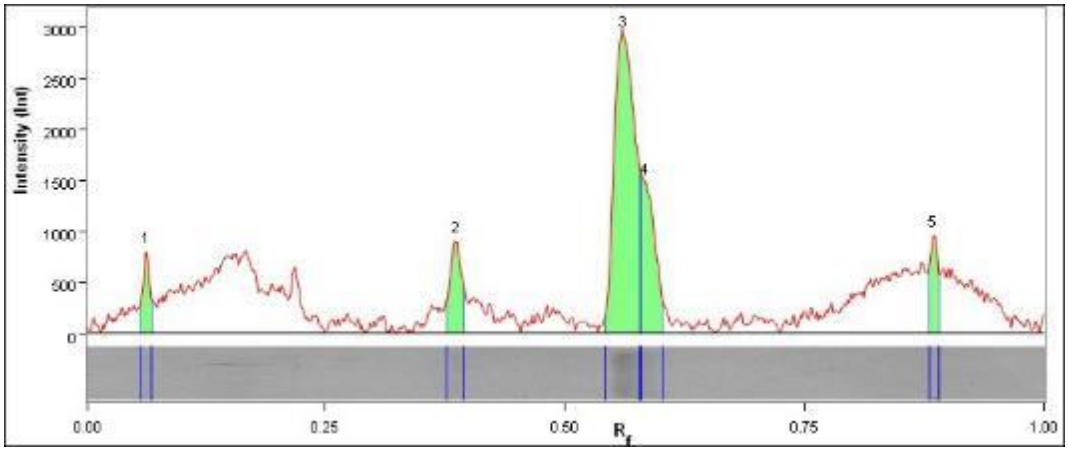
Lane 2



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	0,8
2		106,8	5,2
3		100,0	1,8
4		94,2	2,2
5		80,0	3,7
6		71,9	3,5
7		63,4	3,9
8		61,5	3,7
9		56,4	2,4
10		50,7	1,2
11		46,8	1,9
12		45,0	5,9
13		43,2	4,1
14		40,6	5,2
15		39,4	4,3
16	VP7	36,8	6,2
17		35,1	1,1
18		32,3	4,7
19		31,0	5,2
20		26,3	5,8
21		24,4	5,4
22		23,4	3,6
23		21,8	4,3
24		21,4	3,5
25		17,8	1,6
26		16,4	2,1
27		14,4	6,6

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

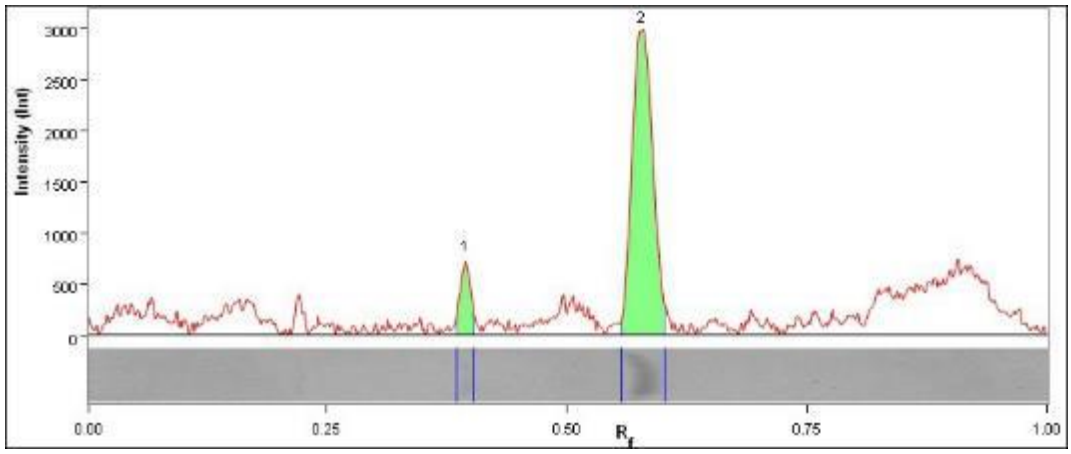
Lane 3



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		92,8	5,2
2	VP7	36,8	9,8
3		24,2	57,9
4		23,4	20,0
5		14,4	7,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

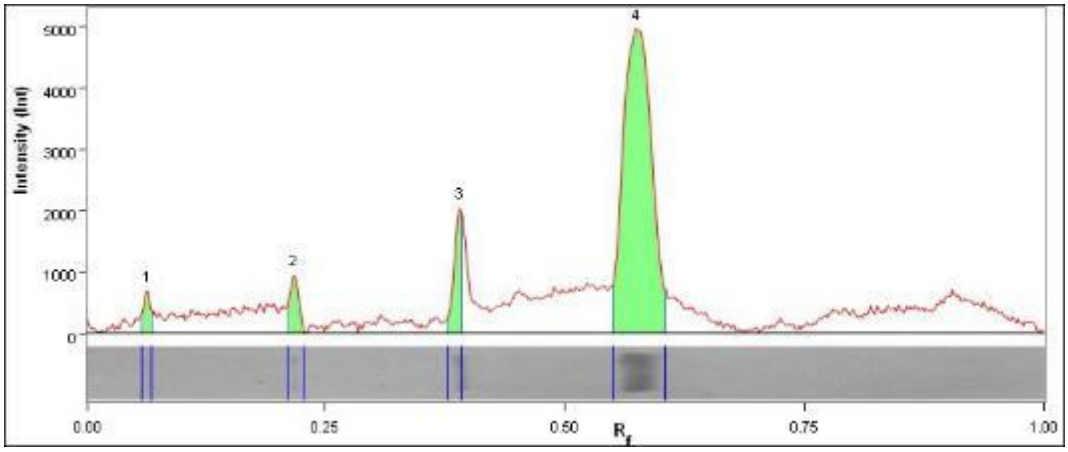
Lane 4



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	VP7	37,0	10,8
2		23,6	89,2

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

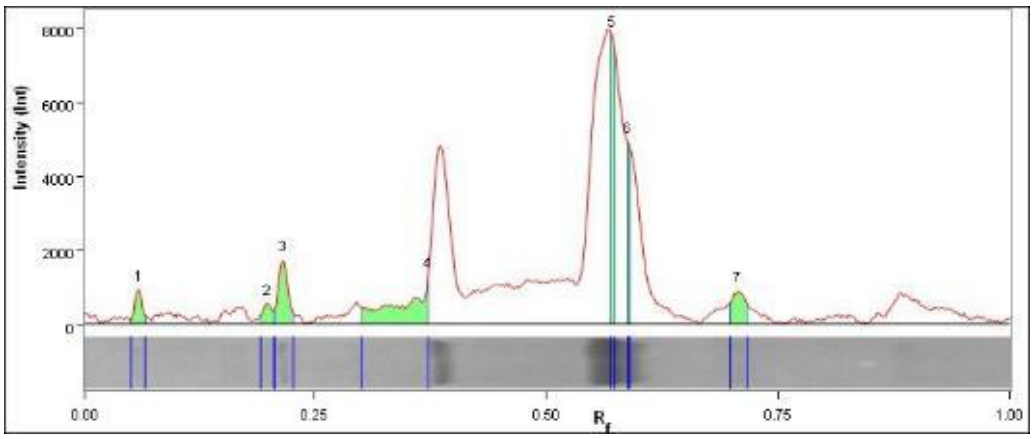
Lane 5



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		92,1	2,7
2		51,5	4,9
3	VP7	37,3	10,0
4		23,7	82,4

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

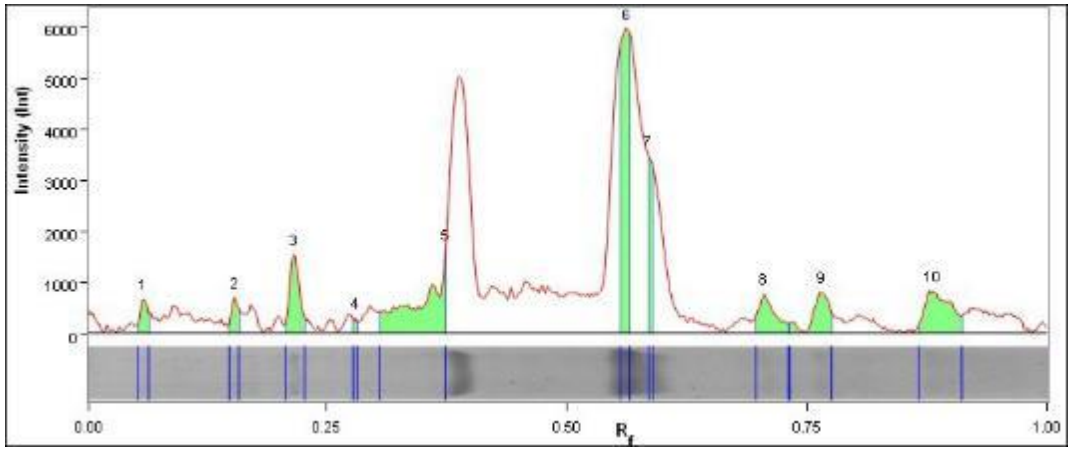
Lane 6



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		94,4	5,4
2		54,0	4,0
3		51,7	13,3
4	VP7	36,9	24,8
5		23,8	26,4
6		23,2	16,9
7		19,2	9,1

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

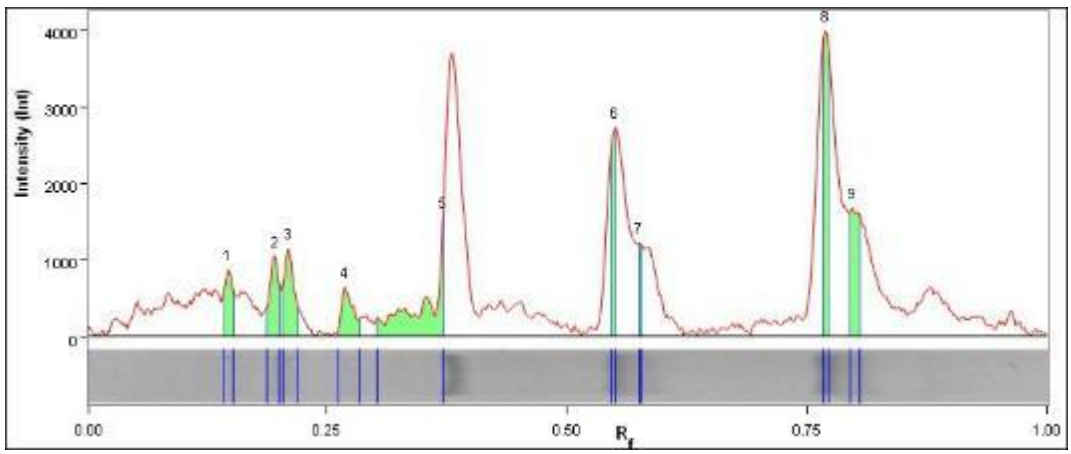
Lane 7



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		95,2	2,7
2		60,7	2,5
3		51,7	8,2
4		44,0	0,6
5	VP7	36,9	20,3
6		24,2	29,8
7		23,3	10,6
8		19,2	6,6
9		17,0	7,1
10		14,4	11,7

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

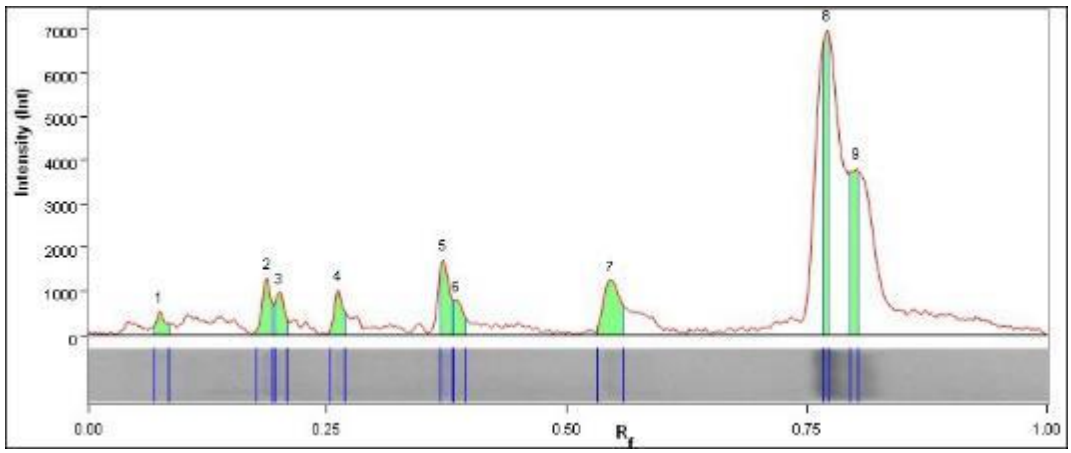
Lane 8



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		61,8	6,0
2		54,4	7,9
3		52,4	9,8
4		45,0	6,5
5	VP7	36,7	20,3
6		24,7	11,1
7		23,7	3,7
8		16,9	19,8
9		15,8	15,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

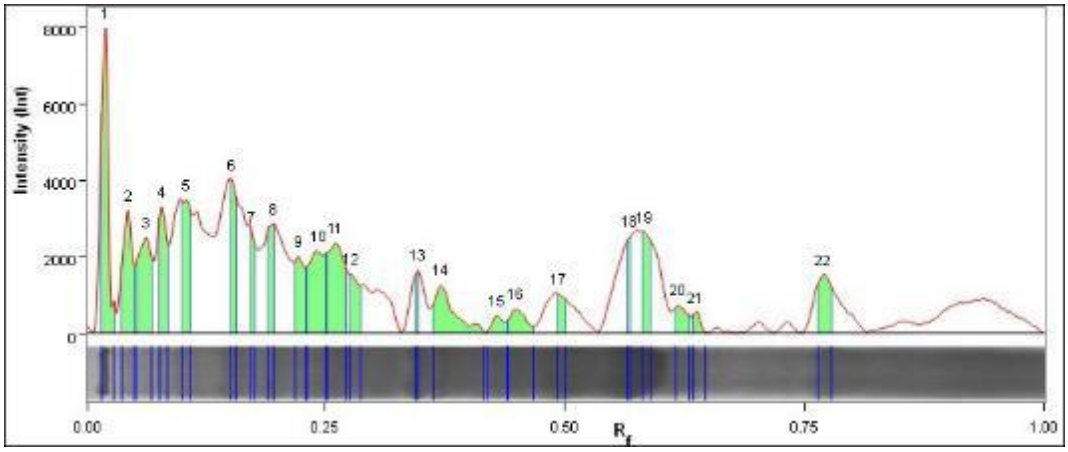
Lane 9



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		86,2	3,2
2		55,6	7,4
3		53,8	5,8
4		45,8	5,9
5	VP7	36,7	9,1
6		35,7	5,0
7		24,8	13,6
8		16,8	26,7
9		15,7	23,3

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm
Regression Equation	A single equation is not available for this method

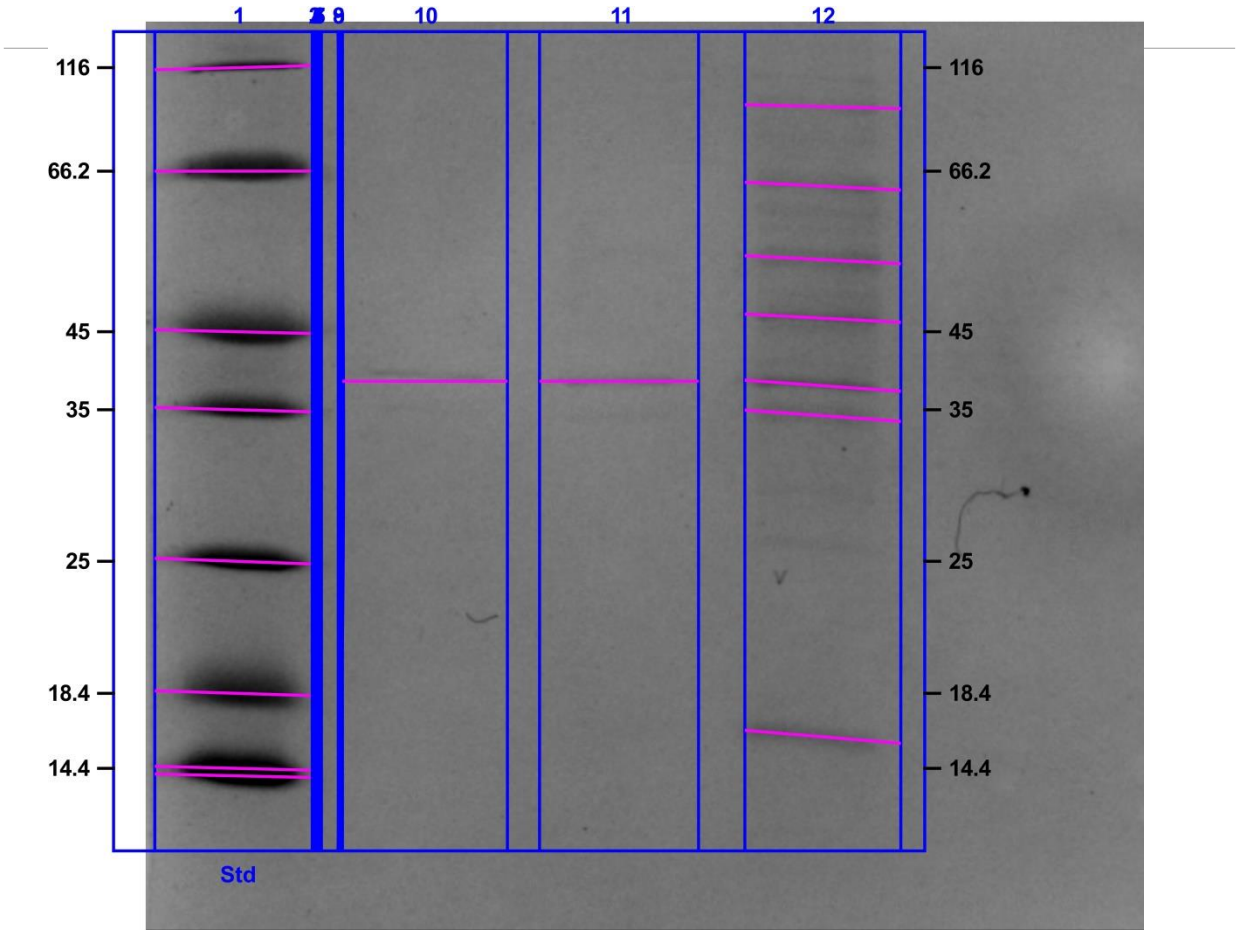
Lane 10



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		116,0	10,7
2		103,3	6,6
3		92,8	7,6
4		84,1	5,0
5		72,5	5,8
6		60,9	5,3
7		57,6	2,1
8		54,5	4,0
9		50,7	4,7
10		48,0	8,6
11		45,8	9,1
12		44,3	3,9
13		38,5	1,4
14	VP7	36,9	5,6
15		32,3	1,3
16		30,9	2,4
17		27,9	1,8
18		24,0	2,1
19		23,4	5,0
20		22,1	1,9
21		21,4	0,9
22		16,9	4,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	8.46 mm

Appendix 7.9 Analysis results for figure 3.17: VP7 HisTag purification of solubilized sample



Acquisition Information

Imager	ChemiDoc™ MP
Exposure Time (sec)	0.052 (Auto - Intense Bands)
Flat Field	Applied (White)
Serial Number	731BR02313
Software Version	5.0
Application	Coomassie Blue
Excitation Source	White Trans illumination
Emission Filter	Standard Filter

Image Information

Acquisition Date	2017/12/20 3:03:13 PM
User Name	User
Image Area (mm)	X: 52.0 Y: 43.5
Pixel Size (µm)	X: 71.8 Y: 71.8
Data Range (Int)	10403 - 31177

Analysis Settings

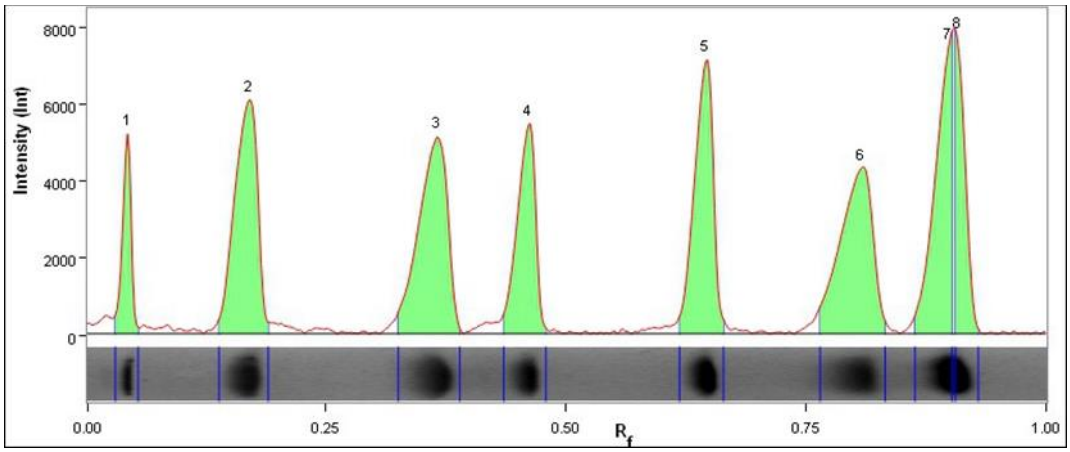
Detection	Lane detection: Manually created lanes Band detection: Bands detected with different sensitivity per lane Manually adjusted bands Lane Background Subtraction: Lane background subtracted with disk size: 10 Lane width: Variable
Mol. Weight Analysis	Standard: BIOTECH STD Standard lanes: first Regression method: Point to Point (semi-log)

Lane Statistics

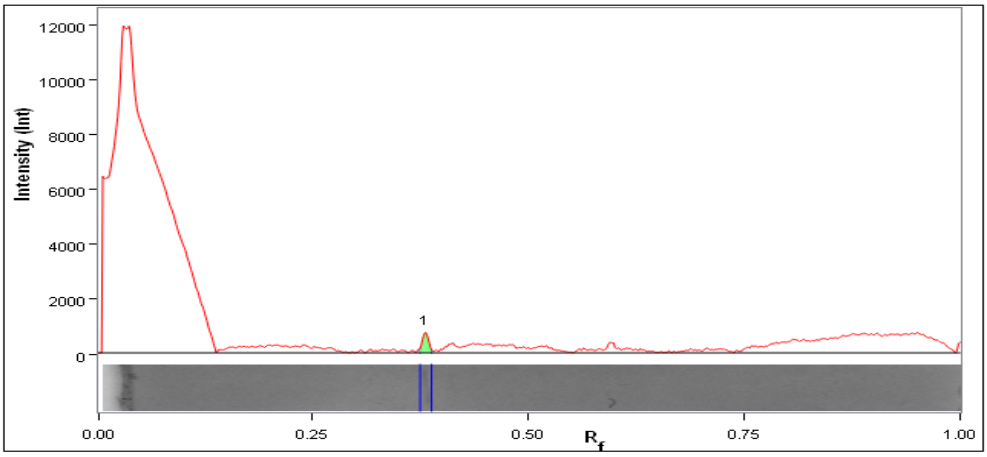
Lane No.	Adj. Total Band Vol. (Int)	Total Band Vol. (Int)	Adj. Total Lane Vol. (Int)	Total Lane Vol. (Int)	Bkgd. Vol. (Int)
1	83 450 290	414 779 750	89 713 470	965 728 280	876 014 810
2	N/A	N/A	355 597	8 139 084	7 783 487
3	N/A	N/A	702 322	16 278 710	15 576 388
4	N/A	N/A	324 145	8 101 623	7 777 478
5	N/A	N/A	327 043	8 089 541	7 762 498
6	N/A	N/A	334 855	8 082 065	7 747 210
7	N/A	N/A	890 442	24 229 188	23 338 746
8	N/A	N/A	248 238	7 926 581	7 678 343
9	N/A	N/A	272 376	7 947 963	7 675 587
10	58 190	4 538 270	14 230 370	853 627 060	839 396 690
11	923 472	21 045 240	17 543 954	818 184 320	800 640 366
12	8 764 392	174 976 880	16 932 760	860 570 256	843 637 496

Lane and Band Analysis
Lane 1 - BIOTECH STD

Band No.	Band Label	Mol. Wt. (KDa)	Band %
1	A	116,0	4,9



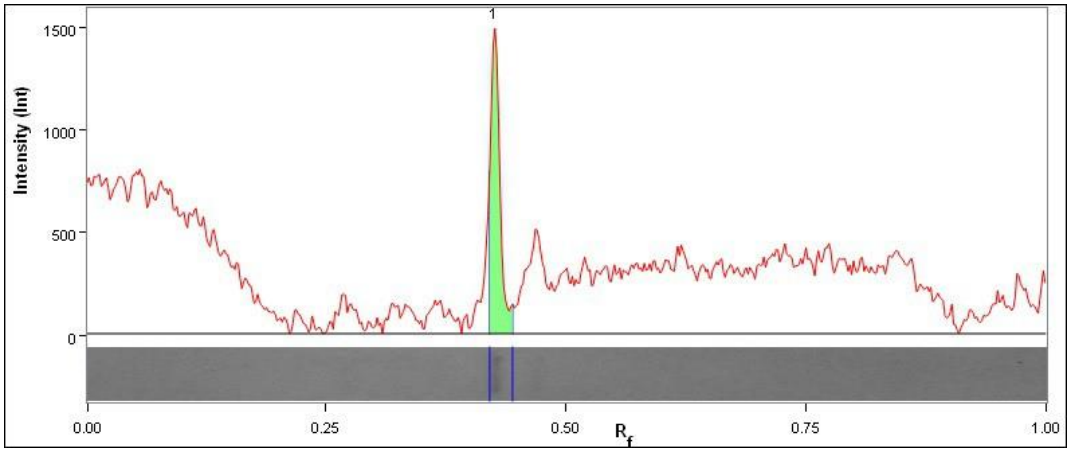
Lane 10



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		37,4	100,0
2	B	66,2	15,3
3	C	45,0	16,1
4	D	35,0	10,9
5	E	25,0	14,7
6	F	18,4	15,4
7	G	14,4	14,0

Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.90 mm
Regression Equation	A single equation is not available for this method

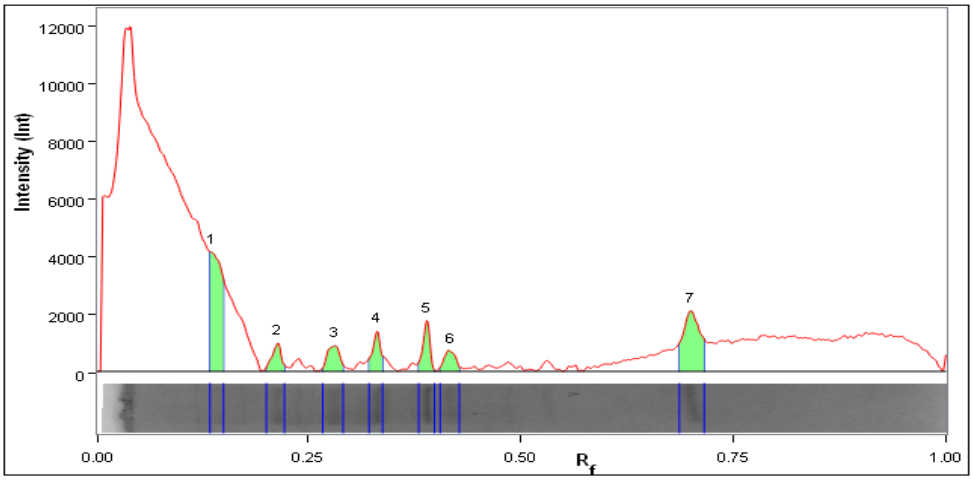
Lane 11



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		37,4	100,0

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.61 mm
Regression Equation	A single equation is not available for this method

Lane 12



Band No.	Band Label	Mol. Wt. (KDa)	Band %
1		93,9	7,6
2		63,9	12,9
3		53,5	16,1
4		46,5	15,7
5		37,8	7,4
6		34,5	12,2
7		16,0	28,2

Band Detection	Automatically detected bands with sensitivity: High
Lane Background	Lane background subtracted with disk size: 10
Lane Width	7.47 mm
Regression Equation	A single equation is not available for this method