universität wien

DISSERTATION

Titel der Dissertation

„Investigating the substrate specificity of enteroviral 2A proteases – an attempt to solve high resolution strutures of enzyme-substrate complexes.“

Verfasser

Mag. rer.nat David Neubauer

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 091 490
Dissertationsgebiet lt. Studienblatt:	Molekulare Biologie
Betreuerin / Betreuer:	Ao. Univ.-Prof. Dr. Timothy Skern

To my parents and everybody who supported me during my time as a PhD student by…

… giving scientific input.

… transferring money to my bank account.

… cheering me up.

… providing a nice working atmosphere.

… handing me chocolate.

… listening to me.

… just being a friend.

… giving me the chance to realize my dreams.

I owe you one...

Index

1 ABSTRACT

The family of picornaviruses contains several important human pathogens such as poliovirus (PV), human rhinovirus (HRV) and coxsackievirus (CV) within its genus of enteroviruses. These simple viruses are composed of a naked protein shell with icosahedral symmetry that encapsidates an RNA genome of positive sense. The genome contains only one open reading frame (ORF) from which a single polyprotein is produced. The pathogens use two virally encoded proteases, namely $2A^{pro}$ and $3C^{pro}$, to process the polyprotein in order to yield mature proteins. $2A^{pro}$ initiates the processing cascade by cleaving between the C-terminus of the preceding protein VP1 and its own N-terminus in an intramolecular (*cis*) reaction. $3C^{pro}$ cleaves all other protein junctions on the polyprotein with the exception of the maturation cleavage within VP0. Once freed from the polyprotein, $2A^{pro}$ has been shown to target several cellular substrates.

The molecular structure of HRV2 $2A^{pro}$ has been refined to a resolution of 1.95Å. The coordinates reveal a chymotrypsin like fold. However, the N-terminal β-barrel is severely truncated (compared to chymotrypsin). Furthermore, $2A^{pro}$ uses a cysteine residue and not a serine residue at the active site as its nucleophile and tightly binds a Zn^{2+} ion within its C-terminal domain. These findings and the fact that $2A^{pro}$ shows substrate specificity patterns that are distinct from comparable cellular proteases account for the qualification of $2A^{pro}$ as a potent drug target. Moreover, inactive PV $2A^{pro}$ mutants were shown to exhibit a dominant negative effect on the growth of wt PV. However, our lack in understanding the exact mechanism of self-processing as well as substrate specificity hampers the knowledge driven design of new antiviral compounds. One reason for this gap in our knowledge is the lack of $2A^{pro}$ structures bound to substrate. For this reason, this thesis covers a number of experiments aiming to provide such structures.

In the first set of experiments, we used inactive $2A^{pro}$ constructs with N-terminal extensions. These recombinant proteins were designed with the purpose of trapping the substrate within the substrate binding cleft. Hence, X-ray crystallography should be used to solve the molecular structures of such conformations in order to provide information of the exact mode of substrate binding during self-processing. In order to raise the likelihood of crystallization, we used $2A^{pro}$ from HRV, CV and PV and different lengths of N-terminal extensions. However, despite our intense efforts we were not able to obtain diffraction quality crystals.

In a different approach, we aimed to solve the high resolution structure of $2A^{pro}$ bound to one or both of the two canonical inhibitors chymostatin or elastatinal. To this end, we were able to reproduce crystals of active HRV2 $2A^{pro}$ as described in the literature. Crystals were soaked in inhibitor solutions and X-ray datasets were collected on a synchrotron beam-line. Initial phases were solved by molecular replacement using the apo HRV2 $2A^{pro}$ coordinates. Additional electron density could be traced within the substrate binding cleft suggesting the presence of the inhibitor.

However, the quality of the maps did not allow the placing of the inhibitor molecules, thereby preventing a refinement of the structure.

Die Familie der Picornaviren enthält wichtige humane Pathogene wie Poliovirus (PV), humanes Rhinovirus (HRV) und Coxsackievirus (CV). Sie alle gehören der Gattung der Enteroviren an. Diese simplen Viren bestehen aus einer nackten Proteinhülle mit ikosaedrischer Symmetrie, die ein RNA Genom positiver Polarität beinhaltet. Das Genom besitzt lediglich einen einzigen offenen Leserahmen der ein Polyprotein codiert. Zwei virale Proteasen, $2A^{pro}$ und $3C^{pro}$, sorgen für die Spaltung des Polyproteins in die eigentlichen, einzelnen Proteine. $2A^{pro}$ initiiert dabei die Reihe der Spaltungen durch die Hydrolyse der Peptidbindung zwischen dem C-Terminus des vorhergehenden Proteins VP1 und dem eigenen N-Terminus. Diese Reaktion erfolgt intramolekular (*cis*). $3C^{pro}$ ist für die Summe der restlichen Spaltungen am Polyprotein, mit Ausnahme der Spaltung von VP0, verantwortlich. Außerdem wurde eine Vielzahl von zellulären Substraten für $2A^{pro}$ identifiziert.

Die molekulare Struktur von HRV2 $2A^{pro}$ wurde mithilfe von Röntgenstrukturanalyse mit einer Auflösung von 1.95Å gelöst. Sie zeigt eine dem Chymotrypsin ähnliche Faltung. Die N-terminale Domäne ist jedoch zu einem β-Faltblatt reduziert. Ferner bedient sich $2A^{pro}$ eines Cystein Restes als Nucleophil im aktiven Zentrum und bindet ein Zn^{2+} Ion in seiner C-terminalen Domäne. Die Substratspezifität von $2A^{pro}$ weist außerdem Unterschiede zur Spezifität vergleichbarer zellulärer Proteasen auf. Diese Differenzen machen $2A^{pro}$ zu einem potenten Ziel für die Entwicklung anti-viraler Substanzen. Dieser Ansatz wird durch die kürzlich entdeckten dominant negativen Effekte von inaktivierten $2A^{pro}$ Mutanten auf das Wachstum von wt Virus untermauert. Eine wissensbasierte Entwicklung von Wirkstoffen wird jedoch durch unser schlechtes Verständnis der genauen Spezifität als auch des exakten Mechanismus der intramolekularen Reaktion weitgehend behindert. Als Grund für diese Wissenslücken kann das Nichtvorhandensein der Molekül Strukturen von $2A^{pro}$- Substrat Komplexen genannt werden. Deshalb beschreibt diese Arbeit eine Vielzahl an Experimenten mit dem Ziel, solche Strukturen zur Verfügung zu stellen.

Der erste Teil der Arbeit beschreibt inaktive $2A^{pro}$ Konstrukte mit N-terminalen Verlängerungen. Diese sollten das Beobachten von Substrat in gebundener Form durch Röntgenstrukturanalyse ermöglichen um ein besseres Verständnis für die Erkennung des Substrates bei der intramolekularen Reaktion zu fördern. Um die Wahrscheinlichkeit für die Kristallisation zu verbessern wurden 2A Proteasen von PV, CV und HRV mit unterschiedlich langen N-terminalen Extensionen verwendet. Trotz intensiver Bemühungen konnten jedoch keine qualitativ hochwertigen Kristalle erzeugt werden.

Ein weiterer Aspekt der Arbeit befasst sich deshalb mit Komplexen aus aktiver $2A^{pro}$ mit den kanonischen Inhibitoren Chymostatin und Elastatinal. Dazu wurde HRV2 $2A^{pro}$ kristallisiert

und die Kristalle in Lösungen der Inhibitoren getränkt. Datensätze zur Röntgenstrukturanalyse wurden an einem Synchrotron aufgenommen und das Phasenproblem über Molecular Replacement gelöst. Die resultierenden Elektronendichten wiesen positive Peaks im Bereich des aktiven Zentrums auf, was auf die Anwesenheit der Inhibitor Moleküle schließen lassen könnte. Die Qualität der Elektronendichteverteilung war jedoch zu gering um das Platzieren der Wirkstoff Moleküle zu ermöglichen. Ein sinnvolles Modell konnte aus diesem Grund nicht erstellt werden.

2 INTRODUCTION

2.1 PICORNAVIRUSES

Picornaviruses are a family of small, icosahedral and non-enveloped viruses. Their capsid contains a single stranded RNA genome of positive sense coding for only one open reading frame (ORF). The term picornavirus is composed of the words pico (the Greek word for small) and RNA (from the nature of its genome). The family contains a number of important animal and human pathogens such as foot and mouth disease virus (FMDV), human rhinovirus (HRV), coxsackievirus (CV), hepatitis A virus (HAV) and poliovirus (PV).

Research on picornaviruses has enabled a number of important scientific achievements. Indeed, the first animal pathogen to be identified as a virus was foot and mouse disease virus (Loeffler & Frosch, 1897; 1898). Especially the intense efforts in investigating poliovirus have lead to key discoveries for the entire field of virology. For example, the first evidence for the existence of an RNA dependent RNA polymerase (RdRP) was obtained using poliovirus and mengovirus (Reich *et al.*, 1961; Shatkin, 1962).

According to the website of the international committee on taxonomy of viruses (ICTV) the family contains 12 genera: aphthovirus, avihepatovirus, cardiovirus, enterovirus, erbovirus, hepatovirus, kobuvirus, parechovirus, sapelovirus, senecavirus, teschovirus and tremovirus (ICTV, 2011) (website accessed December 5th, 2012).

2.1.1 Human Rhinovirus

Human rhinoviruses (genus of enterovirus) are the main causative agents of the common cold, an infection of the upper respiratory tract that goes along with sore throat, sneezing and rhinorrhea. Around half of all cases are caused by human rhinoviruses (Makela *et al.*, 1998) with a peaking percentage of up to 90% of all cases in autumn. With symptoms lasting up to 10 days (Arruda *et al.*, 1997), the common cold is one the major causes for short time absence from work and school (Savolainen *et al.*, 2002), causing a major economic burden of around $40 billion per year in the USA alone (Fendrick *et al.*, 2003).

Recently, more and more evidence was presented implying that human rhinoviruses do not exclusively infect the upper but also the lower respiratory tract (Gern *et al.*, 1997; Kaiser *et al.*, 2006; Mosser *et al.*, 2005; Papadopoulos *et al.*, 2000). It has been shown that many HRV

serotypes prefer to replicate at 33°C rather than at 37°C (Papadopoulos *et al.*, 1999; Stott & Heath, 1970). However, this preference seems to be no absolute obstacle for their replication at core temperature (Papadopoulos *et al.*, 1999; Schroth *et al.*, 1999). Furthermore, lower respiratory infections, especially with human rhinoviruses, have been reported to be linked to exacerbations of asthma (Johnston *et al.*, 1995; Kling *et al.*, 2005; Nicholson *et al.*, 1993) and chronic obstructive pulmonary disease (COPD) (Greenberg *et al.*, 2000; Rohde *et al.*, 2003; Seemungal *et al.*, 2001).

To date, neither antiviral drugs nor vaccines against HRV infections are available.

2.1.1.1 Taxonomy

Human rhinoviruses classically have been distinguished into serotypes by their antigenetical and cross-neutralization properties (Hamparian *et al.*, 1987). To date, the 9th report of the international committee on taxonomy of viruses lists a number of 100 HRV serotypes (Knowles *et al.*, 2012). The common classification system divides human rhinoviruses into two genetic groups (A and B) according to their homologies within the VP4/VP2 coding region of the genome (Horsnell *et al.*, 1995; Savolainen *et al.*, 2002). Genetic group A contains 74 whereas group B contains 25 serotypes. HRV87 was found to cluster with enterovirus 68 (EV68) and has therefore been grouped into the species of *Human enterovirus D* (Blomqvist *et al.*, 2002; Oberste *et al.*, 2004; Savolainen *et al.*, 2002). An alternative classification system divides human rhinoviruses into major and minor group according to the receptor molecule used for cell entry. The major group contains 90% of all human rhinovirus serotypes using the intracellular adhesion molecule 1 (ICAM-1) as their receptor (Greve *et al.*, 1989), whereas minor group viruses utilize the low density lipoprotein receptor (LDLR) or a different member of the LDLR family (Hofer *et al.*, 1994) for cell entry (see chapter 2.3.1.1).

Recently, a novel group of rhinoviruses has been detected to be circulating worldwide (Arden *et al.*, 2006; Kaiser *et al.*, 2006; Lamson *et al.*, 2006; Lau *et al.*, 2007; Lee *et al.*, 2007; Renwick *et al.*, 2007). A new species, namely HRV-C, has been proposed for these viruses. As propagation of those viruses in cell culture has been unsuccessful to date, serological characterization and hence serotyping is not possible. Nevertheless, Simmonds *et al.* recently proposed classification into genotypically assigned types according to VP4/VP2 coding sequences (Simmonds *et al.*). To date, the ICTV lists 48 assigned HRV-C types in their last report on taxonomy of viruses (Knowles *et al.*, 2012). The clinical impact of HRV-C is still a matter of debate. Both data

suggesting increased contribution of HRV-C to lower respiratory tract infections and their consequences (Khetsuriani *et al.*, 2008; Lau *et al.*, 2007; Miller *et al.*, 2009; Wisdom *et al.*, 2009) as well as studies with equal contribution of all HRV species to clinical outcome are available (Lau *et al.*, 2007; Piotrowska *et al.*, 2009).

In 2009, Palmenberg *et al.* completed the available list of full length HRV genome sequences to cover every known serotype (Palmenberg *et al.*, 2009). To this end, they sequenced 70 HRV serotypes with incomplete sequences in the repository and 10 additional field isolates. Their phylogenetic analysis, based on the full genome sequences, revealed a sub-clade (HRV-D) within the species of HRV-A. The clade contained the serotypes HRV8, HRV45 and HRV95 (all major group) and exhibited RNA elements and insertions/deletions that phylogenetically clearly isolates the viruses within it from the rest of the HRV-A viruses.

2.1.2 Coxsackievirus

Coxsackievirus was first isolated and described in 1948 in the town of Coxsackie, New York (Dalldorf & Sickles, 1948) where it was found to be connected to cases of children with paralysis. By now, coxsackieviruses have been associated with a wide spectrum of diseases (Centers for disease control and prevention, 2000; 2006) and are regularly detected worldwide (Antona *et al.*, 2007; Centers for disease control and prevention, 2008). Especially the linkage of coxsackie B viruses to inflammatory heart disease (myocarditis) (Baboonian *et al.*, 1997; Esfandiarei & McManus, 2008; Knowlton, 2008; Longson *et al.*, 1969), pancreatitis (Arnesjo *et al.*, 1976; Imrie *et al.*, 1977; Lal *et al.*, 1988; Parenti *et al.*, 1996) and insulin dependent (type 1) diabetes mellitus (T1D) (Hindersson *et al.*, 2005; Nairn *et al.*, 1999; Smith *et al.*, 1998) have been of interest for the scientific community. The virus is typically spread via the fecal-oral route. Neither vaccines nor antiviral compounds against coxsackievirus infections are available.

2.1.2.1 Taxonomy

Historically, coxsackieviruses have been divided into subgroup A and B according to their pathogenicity in suckling mice (Dalldorf *et al.*, 1949). While coxsackieviruses of group A (CVA) caused flaccid paralysis, infection with members of group B (CVB) resulted in spastic paralysis.

According to current taxonomy (Knowles *et al.*, 2012), coxsackieviruses are scattered amongst the species of human enterovirus A, B and C. The report lists a number of 24 serotypes for CVA and 6 serotypes for CVB. However, some serotypes have been misidentified. Indeed, CVA23 equals Echovirus 9 (E9), CVA15 equals CVA11 and CVA18 equals CVA13. A complete set of full length genome sequences is available in GenBank.

2.1.3 Poliovirus

Poliovirus (PV) was discovered by Karl Landsteiner and Erwin Popper over 100 years ago (Landsteiner & Popper, 1909) and is the causative agent of poliomyelitis. However, the disease is thought to be known for a long time. An Egyptian stone tablet from the 14th century BC shows a man with a crippled leg (Figure 2-1) and is thought to be the oldest record of a poliomyelitis case. First studies that suggested that poliovirus replicates in the intestine (Kling *et al.*, 1912a; b) were disregarded till Albert Sabin and Robert Ward demonstrated that poliovirus is predominately found in the nervous system and the intestine leading to the fact that poliovirus is transmitted via the fecal-oral and oral-oral route (Sabin & Ward, 1941). In 1954, John F. Enders, Thomas H. Weller and Frederick C. Robbins received the Nobel prize in physiology/medicine for their experiments leading to the propagation of poliovirus in human tissue (Enders *et al.*, 1949).

Figure 2-1 Egyptian stone tablet from the 18th dynasty (1403-1365 BC). The tablet shows a man with a debilitated leg suggesting that the tablet records a poliovirus victim (Westendorf, 1992).

Most infections with poliovirus are asymptomatic or cause just mild flu-like symptoms. Indeed, the virus crosses into the central nervous system (CNS) in less than 1% of all infected individuals. Upon uptake of the virus, it reaches the gastrointestinal tract where epithelial cells

form a physical barrier for the virus. The virus possibly invades the lymphoid system by being transported by microfold (M) cells located in the follicle-associated epithelium of Peyer's patches (Neutra *et al.*, 1996; Owen & Jones, 1974). Poliovirus is subsequently further detected in tonsils and the pharynx (Bodian, 1956; Wenner *et al.*, 1960) before entering the bloodstream leading to acute viremia. Virus can also be recovered from extraneural tissue of infected chimpanzees and cynomolgus monkeys (Bodian, 1956; Wenner *et al.*, 1960) as well as infected transgenic mice carrying the poliovirus receptor (Ida-Hosonuma *et al.*, 2005) in low amounts.

Two mechanisms on how poliovirus enters the CNS have been described. First of all, poliovirus is believed to be able to cross the blood-brain barrier (BBB) (Yang *et al.*, 1997). However, the exact mechanism how the virus can overcome the BBB remains unclear. Secondly, poliovirus was found to be transported via retrograde axonal transport in neurons (Ohka *et al.*, 2004; Ohka *et al.*, 2009). Hence, virus that has been taken up at the neuromuscular junction can be transported through neurons to invade the CNS (Nathanson & Bodian, 1961; Ren & Racaniello, 1992). Within the CNS, PV strongly replicates within motor neurons in the anterior horn of the spinal cord as well as in neurons of the brain stem and motor cortex (Bodian, 1949; Bodian & Howe, 1940). Cell lysis of the motor neurons caused by the replication of PV leads to the symptoms of flaccid paralysis typically observed with poliomyelitis.

2.1.3.1 Taxonomy

Already in 1931, Burnet F.M. and McNamara J. observed antigenic differences in strains of PV which were by then thought to behave similarly (Burnet & MacNamara, 1931). The three serotypes of PV were finally assigned in the early 1950's (The Committee on Typing of the National Foundation for Infantile Paralysis, 1951). Today, the ICTV lists them among the species of *Human enterovirus C* (Knowles et al., 2012).

2.1.3.2 The PV eradication program

Poliovirus is believed to have been endemic before the 20th century. Rapid urbanization and increased sanitation lead to a loss of population immunity and an increased age of first contact at which individuals were no longer protected by maternal antibodies. Thus, poliomyelitis

developed from a rare into an epidemic disease spreading over the entire world. In contrast to other enteroviruses, the availability of vaccines allowed medicine to fight the virus. In the early 1950's, Jonas Salk started work on the development of an inactivated poliovirus vaccine (IPV) that finally was tested in the biggest field trial so far (Francis, 1955; Salk *et al.*, 1954). In the following six years, the vaccine reduced the number of paralytic poliomyelitis cases by a 25-fold. An oral attenuated live vaccine (OPV), developed in parallel by Albert Sabin, was first tested in 1957 and introduced in 1961, further reducing the number of poliomyelitis cases (Sabin, 1957). The trivalent OPV (tOPV) rapidly developed into the standard vaccine worldwide. Cuba needed only one year of synchronized OPV use to eradicate PV in 1962 (Cruz, 1984). In 1985, Albert Sabin convinced Rotary International to fund PolioPlus, a campaign targeting poliovirus eradication. Soon, the world health organization (WHO) declared the aim to eradicate PV by the year 2000 followed by a worldwide drop of poliomyelitis cases and a rising number of PV free countries (Figure 2-2).

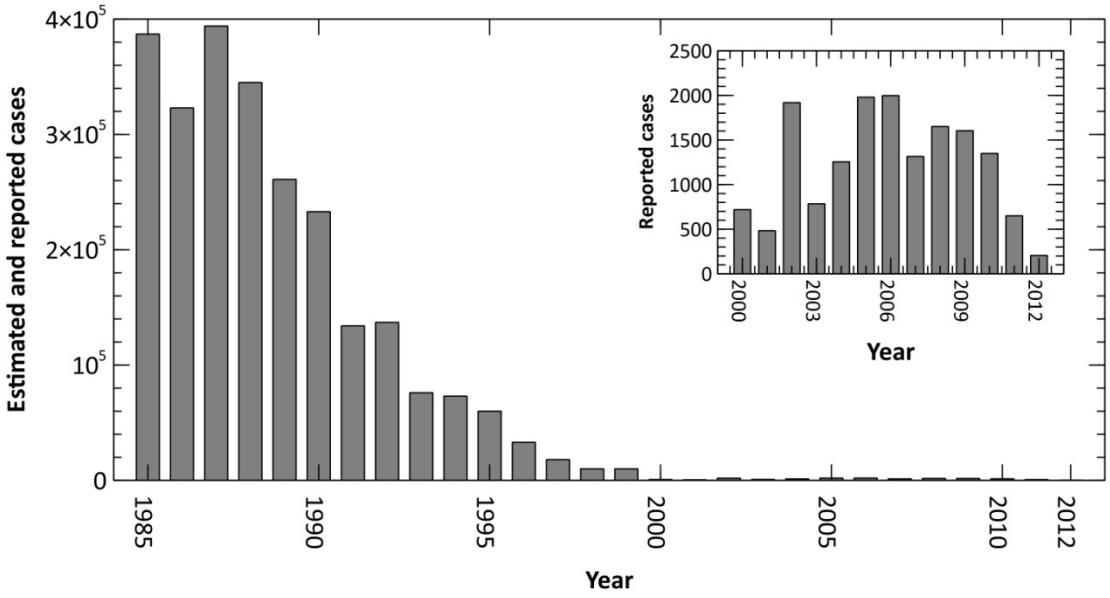


Figure 2-2 Number of estimated and reported paralytic poliomyelitis cases worldwide from 1985-2012 (data from WHO as of Dec 4th, 2012). Numbers from 1985-1999 contain estimates of cases due to non satisfactory disease surveillance (WHO, 2012).

At present, wild PV2 (WPV2) infections seem to have been interrupted worldwide with the last case of a WPV2 infection being reported in India in 1999 (Centers for Disease Control and Prevention (CDC), 1999). Moreover, with India taken off the list of PV endemic countries by the WHO in 2012, only three countries remain that never have stopped wild PV (WPV) circulation.

These are Nigeria, Pakistan and Afghanistan. The case report for 2012 (as of Dec 4th, 2012) lists only four countries with a total of only 205 remaining WPV cases worldwide (Table 2-1) (WHO, 2012). The PV eradication program targets for 2010-2013 state the cessation of all wild poliovirus transmission by the end of 2012, indicated by a period of at least 12 months without a case that is genetically linked to an indigenous virus. The low number of cases for 2012 indicate that humanity is finally on the road to a PV free world.

	WPV1	WPV3	W1W3	Total
Pakistan	53	2	1	56
Afghanistan	33	0	0	33
Nigeria	93	18	0	111
Chad	5	0	0	5
Total	184	20	1	205

Table 2-1 WPV case breakdown by country for the year 2012 (as of Dec 4th, 2012). Reported cases are listed according to country and type: wild PV1 (WPV1), wild PV3 (WPV3) and mixture of type 1 and 3 (W1W3) (WHO, 2012).

2.2 THE VIRION

2.2.1 Capsid

Picornaviral capsids are composed of the four viral proteins VP1, VP2, VP3 and VP4. One copy of each protein is required to form the building block of the capsid, the protomer. Five of which are again combined to form a pentamer (Figure 2-3 A) (see also chapter 2.3.4). In early electron micrographs, viruses appeared roundish with a diameter of around 30nm. Initial X-ray analysis revealed icosahedral symmetry (Finch & Klug, 1959). The first high resolution structures of enteroviral capsids showed that 60 copies of each protein pack together to form a shell with a pseudo T=3 icosahedral symmetry (Hogle *et al.*, 1985; Muckelbauer *et al.*, 1995; Rossmann *et al.*, 1985). Furthermore, these structures revealed that VP1-VP3 form the outer shell, whereas VP4 lines the interior of the capsid. The entire shell is held together by a network of contacts formed by the N-termini of VP1-VP3 in the interior, while the C-termini of those proteins are exposed on the surface.

Enteroviral capsids share some unique features such as the plateaus on the 3-fold and 5-fold axis of symmetry (Figure 2-3 B-E). The star shaped plateau on the 5-fold axis is surrounded by a deep cleft known as the canyon. This depression was found to be the site of receptor binding for major group rhinoviruses, coxsackievirus and poliovirus while minor group viruses bind their receptor closer at the 5-fold axis (see chapter 2.3.1).

Although VP1, VP2 and VP3 share no sequence homology, they all adopt a common wedge shaped jelly-roll fold. This eight-stranded antiparallel β-barrel is composed of two antiparallel β-sheets that are arranged in space to form a wedge like structure (Figure 2-4).

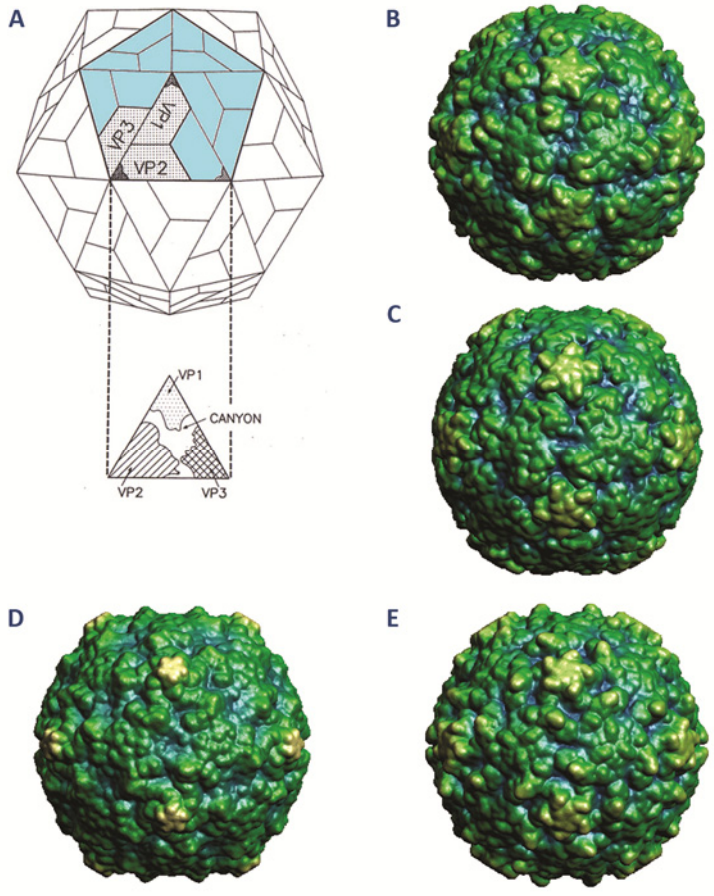


Figure 2-3 Enteroviral capsids and their architecture. (A) Schematic overview of the architecture of enteroviral capsids. One copy of each capsid protein (VP1-VP4) forms a protomer, five of which are combined to form a pentamer (light blue). Adapted from (Fields *et al.*, 2007). **(B-E)** Surface representations of the major group HRV2 (PDB 1fpn) **(B)**, minor group HRV14 (PDB 4hrv) **(C)**, CVB3 (PDB 1cov) **(D)** and PV1 (PDB 1hxs) **(E)**. Capsids are colored according to the distance from the center with central regions in blue, intermediate regions in green and distant regions in yellow. Representations rendered by Viper database (Carrillo-Tripp *et al.*, 2009).

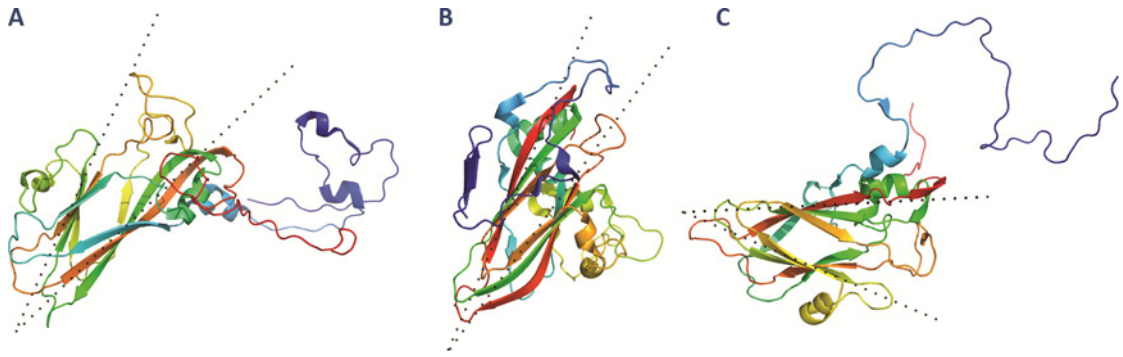


Figure 2-4 Illustrating the wedge shaped jelly roll fold of capsid proteins VP1-VP3. The three capsid proteins VP1 **(A)**, VP2 **(B)** and VP3 **(C)** of the HRV2 capsid (PDB 1fpn) are depicted as cartoon and rainbow colored from N- to C-terminus. The wedge shape of the jelly roll is highlighted by dotted lines. Representations rendered with PyMol (DeLano, 2002).

2.2.2 Genome

The family of picornaviruses contains single stranded RNA of positive sense as its genome. The genome is uncapped at the 5' end and contains a poly(A) tail at its 3' end. The length of picornaviral genomes varies from 7032 bases for *Avian encephalomyelitis virus* (Marvil *et al.*, 1999) to 8828 bases for *Equine rhinitis B virus* (Wutz *et al.*, 1996), excluding the bases of the poly(A) at the 3' end and extensive poly(C) regions in the 5' untranslated region. The length of enteroviral genomes was found to be lower than the median of the family that was calculated to around 7600 bases. Examples of viruses that were of interest for this thesis are listed in Table 2-2. The GC content of enteroviral genomes is typically found to be between 40%-50%.

Virus	Genome length [bases]	GenBank	Reference
HRV2	7102	X02316	(Skern *et al.*, 1985)
HRV14	7212	K02121	(Stanway *et al.*, 1984)
PV1	7440	V01149	(Racaniello & Baltimore, 1981)
CVB4	7395	X05690	(Jenkins *et al.*, 1987)

Table 2-2 Genome length of viruses used within this thesis. Lengths exclude the bases of the poly(A) tail.

The genome contains only one open reading frame that is preceded by a long 5' untranslated region (UTR) and succeeded by a short 3' UTR and the poly(A) tail (Figure 2-5 A). The RNA is uncapped but bound to a small viral protein (VPg) on its 5' end (Flanegan *et al.*, 1977; Lee *et al.*, 1977). VPg is covalently linked to the 5'-most nucleotide of the genome via an O4-

(5'uridylyl)tyrosine bond (Ambros & Baltimore, 1978; Rothberg *et al.*, 1978). The 5' UTR contains several important motives in its RNA secondary structure. Directly succeeding the covalently bound VPg at the 5' end, enteroviruses harbor a cloverleaf structure of about 89 nucleotides (nt) (Andino *et al.*, 1993; Andino *et al.*, 1990) (Figure 2-5 B and C). The cloverleaf is followed by a spacer region of variable length which is rich in cytidines (spacer I, ~10-25nt). The succeeding RNA stretch folds into seven stems known as the internal ribosomal entry segment (IRES) (Figure 2-5 C). Enteroviruses contain an IRES of type I, the stems of which are used to bind proteins as well as the 43S ribosomal subunit and was first described in PV by Pelletier *et al.* (Pelletier & Sonenberg, 1988; 1989). The IRES is followed by a second region of variable length (spacer II, up to ~100 bases) that separates the IRES from the AUG initiation codon. Spacer II folds into a stem in the genomes of HRVs, pairing the AUG initiation codon with an upstream non-coding AUG (Figure 2-5 D).

The single ORF codes for only one polyprotein that will give rise to 11 mature proteins upon full processing. The polyprotein is organized in three regions. P1 contains the four capsid proteins VP1-VP4, whereas P2 and P3 contain so called accessory proteins as well as the proteins required for replication. The historic scheme of protein numbering for picornaviruses (Rueckert & Wimmer, 1984) is used in Figure 2-5 A, but the terms VP1 (1D), VP2 (1B), VP3 (1C) and VP4 (1A) are simultaneously used. The ORF is followed by one or two in-frame stop codons and a short 3' UTR. 3' UTR RNA forms a single unbranched stem (e.g. in HRV-A (Palmenberg *et al.*, 2009) (Figure 2-5 E)) or multiple, mostly unbranched stems with possible kissing interactions (e.g. HEV-B (Zoll *et al.*, 2009) (Figure 2-5 F)).

In 1998, the efforts to establish a replicon system for HRV14 resulted in the discovery of an additional RNA element required for replication within the ORF (McKnight & Lemon, 1996). The *cis* acting replicative element (CRE) was defined (McKnight & Lemon, 1998) and subsequently found to be present in other picornaviruses as well. However, the position of these elements was found to differ dramatically. The CRE was found to be located within the 2C coding region of HEV-A, HEV-B, HEV-C and HEV-D (Goodfellow *et al.*, 2000; Paul *et al.*, 2000), within the 2A coding region of HRV-A (Gerber *et al.*, 2001), within the VP1 coding region of HRV-B (McKnight & Lemon, 1998) as well as the VP2 coding region of HRV-C (Cordey *et al.*, 2008). The CREs of enteroviruses consist of small stems (14nt) (Figure 2-5 G) and were found to function as a template for VPg uridylylation (Paul *et al.*, 2003a; Paul *et al.*, 2000; Paul *et al.*, 2003b).

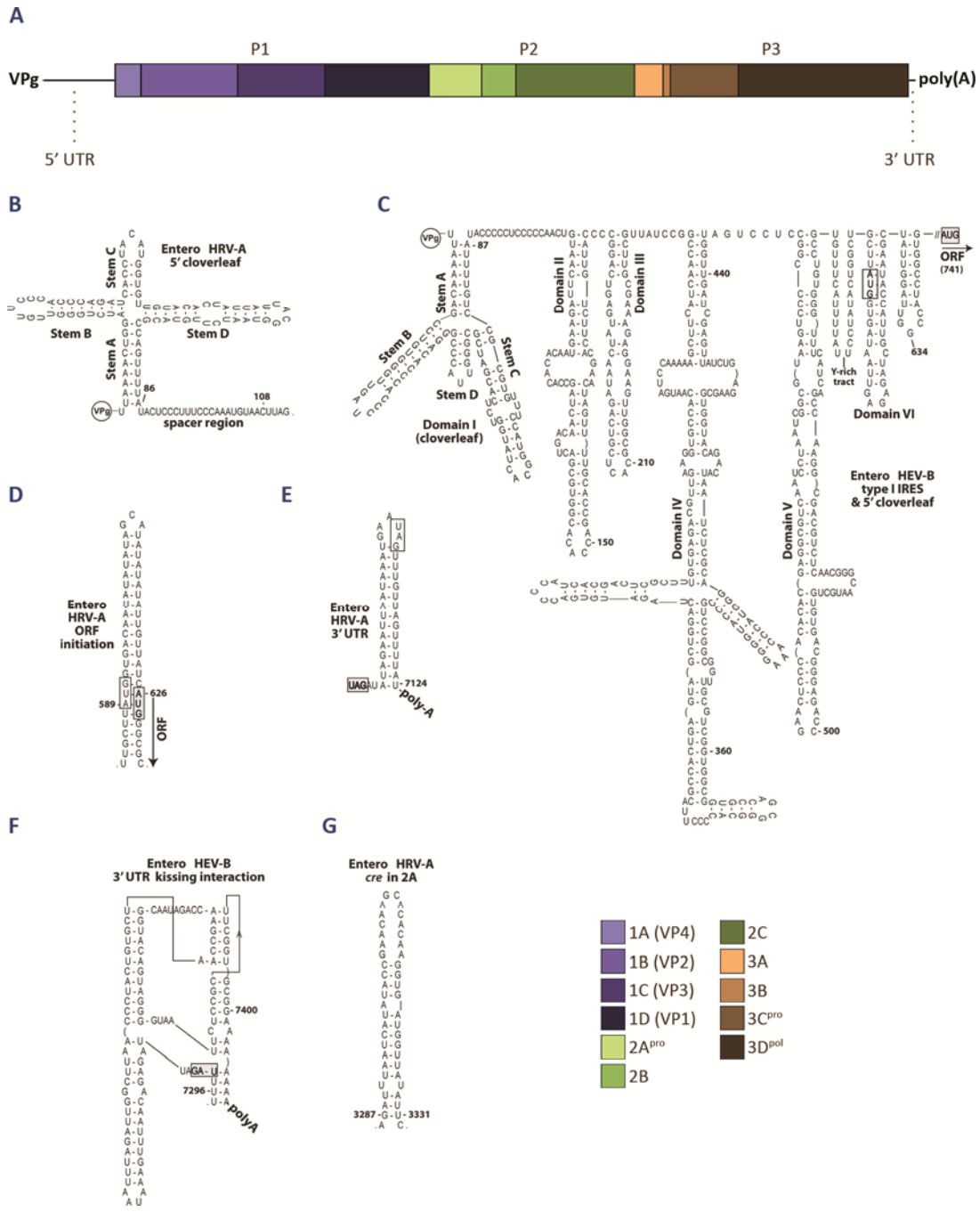


Figure 2-5 Overview of the enteroviral genome and RNA secondary structural elements within it. **(A)** Schematic, organizational view (Rueckert & Wimmer, 1984) of HRV2 viral RNA (Skern *et al.*, 1985). Untranslated regions and coding regions for each viral protein are drawn to scale. The three regions of the ORF are color coded in shades of blue (P1), green (P2) and brown (P3). Each color represents one of the 11 encoded proteins. **(B-G)** Representations of secondary structural RNA elements in enteroviral genomes (Palmenberg *et al.*, 2010).

Enteroviral RNA alone is infective when transfected into susceptible cells. Moreover, cell free systems allow translation of proteins from viral RNA, viral RNA replication and assembly of infectious virus (Barton & Flanegan, 1993; Molla *et al.*, 1991; Todd *et al.*, 1997; van Ooij *et al.*,

2006). In 2002, the lab of Eckard Wimmer showed that it was possible to build the genome of PV without using a nucleic acid template by combining synthetic oligonucleotides to a full length, genomic cDNA. Fully infectious RNA could be transcribed *in vitro* and used to grow PV (Cello *et al.*, 2002).

2.3 LIFE CYCLE

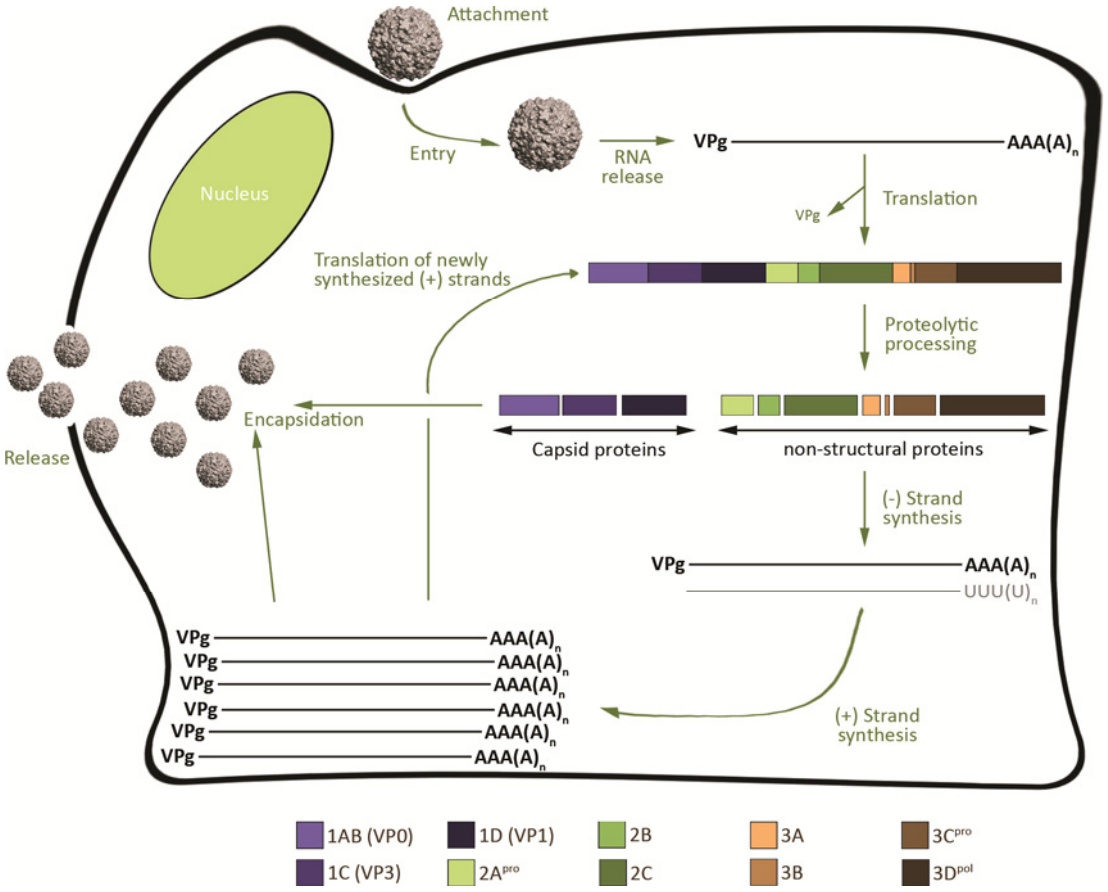


Figure 2-6 Schematic representation of the enteroviral life cycle. The virus uses receptor mediated endocytosis to attach to and enter susceptible cells (chapter 2.3.1). RNA is released into the cytosol, the location of replication. After VPg has been removed from the viral genome, the single open reading frame is translated into a polyprotein that is subsequently processed by viral proteases (chapter 2.3.2). Replication complexes used for (-) strand intermediate and (+) strand synthesis are found attached to virus induced membrane vesicles (not drawn for simplicity). (-) Strand intermediates are used as templates for the amplification of new (+) strands (chapter 2.3.3), which are either translated for further amplification of viral proteins or encapsidated to form new provirions (chapter 2.3.4). A maturation cleavage, that separates the capsid proteins 1A (VP4) and 1B (VP2) turns the particles into fully infectious virions that are released by cell lysis. Adapted from Zoll *et al.* (Zoll *et al.*, 2009).

2.3.1 Attachment and entry

2.3.1.1 Receptors

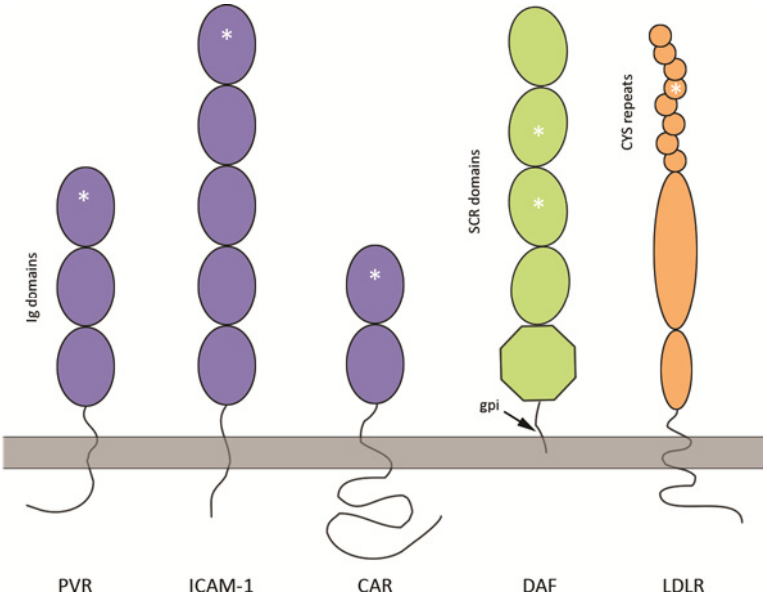


Figure 2-7 Schematic illustration of receptors used by PV, HRV and CVB to enter and infect susceptible cells. All three serotypes of PV use the poliovirus receptor (PVR). Human rhinoviruses are split into two groups (major and minor) according to their receptor usage. Major group HRVs bind to ICAM-1, whereas minor group HRVs utilize members of the low density lipoprotein receptor family (LDLR). Finally, all six serotypes of CVB use the coxsackie and adenovirus receptor (CAR), while only a subset of CVB isolates can bind to the complement decay-accelerating factor (DAF). The white asterisk marks domains involved in binding the virus.

Human rhinoviruses are classified into the major and minor group according to receptor usage of either intracellular adhesion molecule I (ICAM-I) or members of the low density lipoprotein receptor (LDLR) family respectively (see chapter 2.1.1.1). ICAM-1 was identified as the receptor for major group rhinoviruses in the late 1980's (Greve *et al.*, 1989; Staunton *et al.*, 1989; Tomassini *et al.*, 1989). The extracellular, N-terminal part of ICAM-1 consists of five Ig-like domains (D1-D5) and is highly glycosylated (Figure 2-7), whereas the intracellular domain of this transmembrane protein is rather short (Staunton *et al.*, 1988). Its natural function is to bind the ligands leukocyte function-associated antigen-1 (LFA-1) and macrophage-1 antigen (Mac-1) at sites of infection. The N-terminal domain D1 was found to bind within the canyon, a 12Å deep depression surrounding the star shaped plateau at the 5-fold axes of symmetry (Kolatkar *et al.*, 1999; Olson *et al.*, 1993) (Figure 2-8 A). Some major group viruses were also found to alternatively

bind to heparan sulfate for the infection of cells lacking ICAM-1 (Khan *et al.*, 2007; Vlasak *et al.*, 2005).

The LDL receptor was found to consist of eight repeats of 40 amino acids (aa) that are rich in disulfide bonds at its N-terminus. The rest of the extracellular domain contains a mouse epidermal growth factor-like domain and a serine and threonine rich region that is followed by the transmembrane domain (Yamamoto *et al.*, 1984) (Figure 2-7). Its natural function is to bind and take up lipids that are transported within apolipoprotein complexes B and E. After Mischak *et al.* had provided first data on a possible receptor (Mischak *et al.*, 1988), Hofer *et al.* indentified LDLR as the receptor for minor group rhinoviruses (Hofer *et al.*, 1994). Cryo-EM as well as X-Ray data of VLDLR bound to HRV2 has shown that in contrast to ICAM-1, the minor group receptor binds to the star shaped plateau near the five-fold axis of symmetry (Hewat *et al.*, 2000; Verdaguer *et al.*, 2004) (Figure 2-8 B).

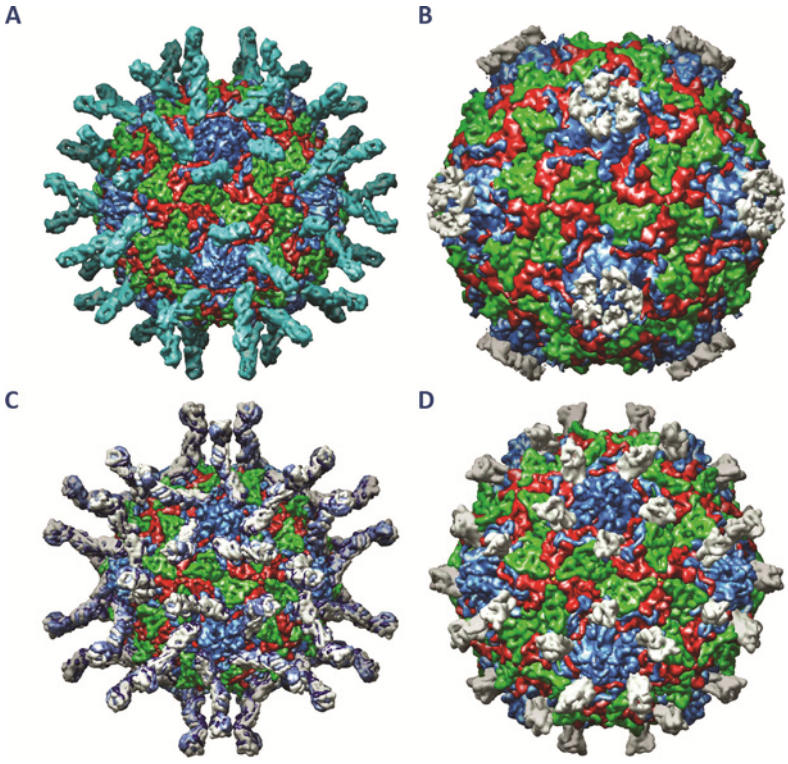


Figure 2-8 Representations of HRV, PV and CVB bound to their receptor molecules. (A) Cryo-EM structure of the major group HRV16 in complex with domain 1 and 2 of its natural receptor ICAM-1 (PDB 1d3e). **(B)** X-Ray structure of the minor group HRV2 in complex with calcium containing modules of the very low density lipoprotein receptor (PDB 1v9u). **(C)** Cryo-EM structure of PV in complex with the PV receptor (PVR) (PDB 1nn8). **(D)** Cryo-EM structure of CVB3 bound to its receptor CAR (PDB 1jew). Capsid proteins are colored according to the historic scheme of dark blue (VP1), green (VP2) and red (VP3). Receptors and receptor fragments are colored in grey and light blue respectively. All representations were rendered by Viper database (Carrillo-Tripp *et al.*, 2009).

A monoclonal antibody raised against an epitope on HeLa cells that was able to block infection by all three PV serotypes gave first hints on the receptor used by PV (Nobis *et al.*, 1985). Subsequently, the receptor was identified as CD155 (now called PVR) (Koike *et al.*, 1990; Mendelsohn *et al.*, 1989), a member of the immunoglobolin superfamily. The extracellular domain contains three Ig-like domains (Figure 2-7) attached to the cell membrane by a transmembrane peptide and a short intracellular domain. The natural function of the receptor was linked to interactions of the cell with the extracellular matrix (Lange *et al.*, 2001) and other cells (Bottino *et al.*, 2003; Seth *et al.*, 2007) as well as to cell migration (Fujito *et al.*, 2005; Sloan *et al.*, 2005). However, its function is poorly understood in detail. Several cryo-EM studies revealed that PVR binds with its N-terminal Ig-like domain (D1) into the canyon of the PV capsid (Belnap *et al.*, 2000; He *et al.*, 2000; Xing *et al.*, 2000) (Figure 2-8 C). It was suggested that the receptor binds the virus in a two step manner. Upon loose and reversible attachment of the receptor, conformational changes in the capsid allow deeper penetration of the D1 domain into the canyon (Tsang *et al.*, 2001; Zhang *et al.*, 2008). The expression of PVR in transgenic mice allows infection of mice with PV, thus establishing an animal model (other than Old World primates) for PV research (Ren *et al.*, 1990).

CVBs have been found to use at least three different cellular receptors. The main one being the coxsackie- and adenovirus receptor (CAR) utilized by all six serotypes of CVB (Bergelson *et al.*, 1997; Carson *et al.*, 1997; Martino *et al.*, 2000; Tomko *et al.*, 1997). Like ICAM-1 and PVR, CAR is a member of the immunoglobulin superfamily with two Ig-like domains on its N-terminal, extracellular part (Figure 2-7). Its natural function was linked to cell-cell adhesion (Honda *et al.*, 2000), especially in tight junctions of polarized epithelial cells (Cohen *et al.*, 2001). CAR attaches to CVB in a similar way as ICAM-1 to HRV and PVR to PV. The N-terminal Ig-like domain (D1) penetrates the canyon of the capsid (He *et al.*, 2001) as determined by cryo-EM (Figure 2-8 D).

A subset of CVBs utilizes an additional cellular receptor to bind susceptible cells. Complement decay-accelerating factor (DAF) is a transmembrane protein with four homologous short consensus repeat (SCR) domains and an additional highly glycosylated domain anchored to the cell membrane by glycosyl phosphatidylinositol (gpi) (Figure 2-7). DAF was shown to protect cells from lysis by the complement system as its natural function (Lublin & Atkinson, 1989). Unlike CAR, DAF binds along the two fold axis of symmetry via its SCR2 and SCR3 domains (Hafenstein *et al.*, 2007). Finally, heparan sulfate was shown to act as a receptor for a single variant of CVB3 in CAR deficient cells (Zautner *et al.*, 2006; Zautner *et al.*, 2003).

2.3.1.2 Entry and uncoating

Enteroviral capsids undergo a series of conformational changes upon uptake into the cell and release of RNA. It has been shown that binding of PVR to PV, ICAM-1 to major group HRVs and CAR to CVB is sufficient to trigger the externalization of VP4 as well as the N-terminus of VP1 to form the 135S particle (De Sena & Mandel, 1977; Fenwick & Cooper, 1962; Hoover-Litty & Greve, 1993; Milstone *et al.*, 2005). In contrast, minor group HRVs are not destabilized upon binding to their receptor and are dependent on the low pH in late endosomes for the formation of 135S particles (Neubauer *et al.*, 1987; Prchla *et al.*, 1994). The 135S particle shows altered antigenicity and is susceptible to proteolytic degradation (Fricks & Hogle, 1990). Further changes include externalization of the RNA leading to the formation of the empty 80S particle. The exact mode of RNA release remains a matter of debate. However, the externalized peptides of VP1 and VP4 have been shown to insert into membranes (Danthi *et al.*, 2003; Fricks & Hogle, 1990; Tuthill *et al.*, 2006) and the N-terminus of VP1 was predicted to form an amphipathic helix which was recently modeled as a stable α-helix by molecular dynamics studies (Roberts *et al.*, 2012). These findings led to suggestions that RNA is released through a pore in order to reach the cytoplasm. Furthermore, a study of Prchla *et al.* demonstrated the selective release of small, fluorescently labeled dextrans from endosomes upon infection with HRV2 (Prchla *et al.*, 1995). In contrast, endosomes seem to be disrupted in HRV14 infected cells leading to the release of a broad range of large dextrans as well as altered virus to the cytoplasm (Schober *et al.*, 1998). Finally, recent studies using HRV2 suggest that RNA is released from the capsid with its 3' end first (Shushan Harutyunyan, personal communication).

The view on cellular pathways used by enteroviruses to enter a cell has strongly diversified over the last decade. The major group HRV14 was shown to use clathrin-, caveolin- and flotillin-independent endocytosis to enter rhabdomyosarcoma cells that overexpress ICAM-1 (Khan *et al.*), whereas internalization may occur via clathrin-dependent endocytosis in HeLa cells (DeTulleo & Kirchhausen, 1998; Grunert *et al.*, 1997). Moreover, also the minor group HRV2 was suggested to use clathrin-dependent endocytosis (Snyers *et al.*, 2003). The exact route of PV entry is unknown. However, it has been shown that inhibitors targeting clathrin-mediated endocytosis, caveoli and macropinocytosis do not affect PV entry in HeLa cells (DeTulleo & Kirchhausen, 1998; Perez & Carrasco, 1993). In contrast, PV was found to utilize dynamin-dependent caveolar endocytosis in human brain microvascular endothelial cells (Coyne *et al.*, 2007a). In addition, Coyne and Bergelson shed light on a possible multi step entry pathway of CVB3 (Coyne *et al.*,

2007b). CVB3 first binds the receptor DAF on the apical surface of polarized epithelial cells followed by migration of the virus to tight junctions where it binds CAR. After conversion to 135S particles, virus is taken up via caveosomes. However, Patel *et al.* observed sensitivity of viral uptake to inhibitors of macropinocytosis, suggesting a combination of caveolin-mediated endocytosis and macropinocytosis (Patel *et al.*, 2009).

2.3.2 Decoding of the genetic information and polyprotein processing

2.3.2.1 Enteroviral translation initiation and host cell shut-off

Unlike most cellular messenger RNAs, the picornaviral RNA genome does not contain the m^7Gppp residue on its 5' end known as the cap structure (Nomoto *et al.*, 1976). The small viral protein 3B (VPg) is covalently linked to the 5' end of all picornaviral genomes instead (see chapter 2.2.2). However, VPg is not required for translation initiation and is removed before binding of the RNA to the translational machinery (Ambros & Baltimore, 1980). Recently, the cellular DNA repair enzyme 5'-tyrosyl-DNA phosphodiesterase-2 (TDP2) has been shown to be responsible for this cleavage (VPg unlinkase) (Virgen-Slane *et al.*).

In the absence of a 5' cap, which commonly promotes attachment of mRNAs to the initiation factor eIF4E (Sonenberg *et al.*, 1978), picornaviruses have evolved a different mechanism to initiate translation of their genome. As described in chapter 2.2.2, the 5' UTR of picornaviral genomes contain highly structured RNA stretches including the internal ribosomal entry site (IRES). These structures were first shown to act as the sites of internal ribosome binding in 1988 (Jang & Wimmer, 1990; Pelletier & Sonenberg, 1988). Picornaviral IRES structures are divided into the groups I-IV according to structural similarities and the set of cellular translation factors required. The type I IRES found in enterovirus requires a set of canonical eukaryotic initiation factors (eIF) to build up the 43S pre-initiation complex. This set consists of the 40S ribosomal subunit, eIF2-GTP-met-$tRNA_i$, eIF4G, eIF4A, eIF3 and ATP (Pestova *et al.*, 1996a; Pestova *et al.*, 1996b). eIF4G plays a key role in cap dependent as well as IRES mediated translation initiation (see Figure 2-9). In the first case, it bridges the cap binding protein eIF4E via its N-terminal domain with the eukaryotic initiation factors 4A (eIF4A) and 3 (eIF3) that bind to the C-terminal domain of eIF4G (Lamphear *et al.*, 1995; Mader *et al.*, 1995). eIF3 in turn binds the 40S ribosomal subunit. In the latter case, the C-terminal domain of eIF4G was shown to directly

bind to domain V of type I IRES structures (de Breyne *et al.*, 2009), thereby promoting the formation of the pre-initiation complex.

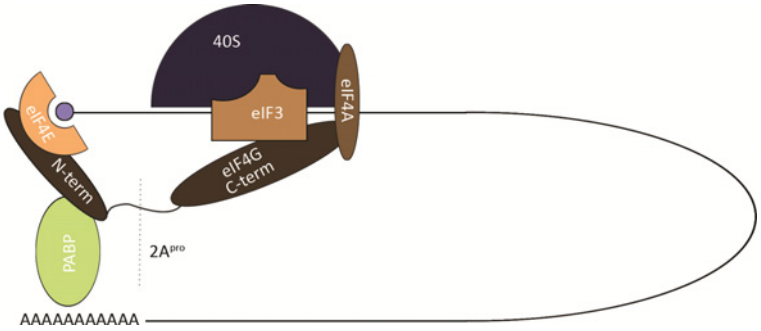


Figure 2-9 Simplified, schematic view of cap dependent translation initiation. The initiation factor eIF4G acts as a scaffolding protein, bridging the cap binding protein (eIF4E) with initiation factors that recruit the 40S ribosomal subunit to build the 43S pre-initiation complex. During enteroviral infection cleavage of eIF4G by the viral protease $2A^{pro}$ leads to the loss of the scaffold and the so called host cell shut-off.

Understanding of the complex process of translation initiation was further complicated by the discovery of polypyrimidine tract-binding protein (PTB) (Jang & Wimmer, 1990; Luz & Beck, 1991). PTB was the first in a long line of non-canonical factors found to modulate translation initiation and termed IRES *trans*-acting factors (ITAF). Additionally, initiation of translation is likely dependent on the melting of RNA secondary structures in vicinity of the start codon (Svitkin *et al.*, 2001).

The host cell shut-off is a phenomenon first described in PV infected HeLa cells. It is triggered by the cleavage of the initiation factor eIF4G (Etchison *et al.*, 1982) and consequently loss of 4G's scaffolding function. Hence, infected cells are no longer able to initiate translation of its own capped mRNAs. In contrast, viral RNA is translated even better in the presence of the C-terminal cleavage products of eIF4G (Borman *et al.*, 1997). For this reason, the virus hijacks the cellular translation machinery giving it advantage over the host cell. After a period of debate, $2A^{pro}$ from enteroviruses was shown to directly cleave eIF4G (Lamphear *et al.*, 1993; Liebig *et al.*, 1993; Sommergruber *et al.*, 1994a). In 1998, Gradi *et al.* discovered a homolog of eIF4G (now eIF4GI and eIF4GII) (Gradi *et al.*, 1998a) and demonstrated that both forms have to be cleaved to achieve complete host cell shut-off (Gradi *et al.*, 1998b). The cleavage sites on the two homologs have been mapped to $LSTR^{681}$*GPPR (eIF4GI) for HRV2 as well as CVB4 $2A^{pro}$ (Lamphear *et al.*, 1993) and $LLNV^{699}$*GSRR (eIF4GII) for HRV2 $2A^{pro}$ (Gradi *et al.*, 2003). More recently, the role of poly-A

binding protein (PABP) cleavage in host cell shut-off has been described (Joachims *et al.*, 1999; Kerekatte *et al.*, 1999).

2.3.2.2 Polyprotein processing

As described in chapter 2.2.2 picornaviral genomes contain only one large open reading frame. Thus, translation of such a genome results in the expression of one big polyprotein. However, as processing is initiated co-translationally, full length polyprotein cannot be detected in infected cells.

The processing cascade is initiated by two intramolecular (*cis*) cleavages performed by $2A^{pro}$ (Toyoda *et al.*, 1986) and $3C^{pro}$ (Gorbalenya *et al.*, 1981; Gorbalenya *et al.*, 1979; Svitkin *et al.*, 1979) (see Figure 2-10). $2A^{pro}$ cleaves at its own N-terminus, thereby releasing the P1 region containing the capsid proteins from the rest of the polyprotein. $3C^{pro}$ processes the site between 2C and 3A. However, an alternative rapid primary cleavage event by $3C^{pro}$ between 2A and 2B has been suggested (Lawson & Semler, 1992).

In a cascade of further proteolytic cleavages, the polyprotein is broken down to its "mature" proteins. The workhorse of this cascade is $3C^{pro}$ that processes all of the remaining sites with exception of the 1A/1B junction (Hanecak *et al.*, 1982; Palmenberg *et al.*, 1979). However, the historical term "mature" proteins is misleading. In fact, immature processing intermediates were shown to play essential roles in the viral life cycle. These include the precursor $3CD^{pro}$ that in fact is the active species that processes the P1 capsid protein precursor (Jore *et al.*, 1988; Ypma-Wong *et al.*, 1988). PV $3C^{pro}$, the most specific of all enteroviral 3C proteases, exclusively cleaves between Gln and Gly. However, not all Gln-Gly pairs in the polyprotein are processed. An alternative processing pattern leads to the formation of 3C' and 3D' by proteolysis of a Tyr-Gly bond performed by $2A^{pro}$ (Toyoda *et al.*, 1986). Nevertheless, viral growth is unaffected by a knock out of this cleavage site in cultured, infected cells (Lee & Wimmer, 1988), raising the question of the biological relevance of this processing event.

During the last steps of virion morphogenesis, the 1AB precursor (VP0) is cleaved into the two capsid proteins 1A (VP4) and 1B (VP2) (maturation cleavage) (Jacobson & Baltimore, 1968). The exact mechanism of this maturation step is still a matter of debate. An initial hypothesis in which the conserved Ser10 of VP2 would act as the nucleophile for autocatalytic cleavage (Arnold *et al.*, 1987) was rejected after a Ser10Ala mutation did not affect the maturation step (Harber *et al.*, 1991). Subsequently, a conserved histidine residue was suggested to act together with the

encapsidated RNA to perform the maturation cleavage (Chen *et al.*, 1989; Curry *et al.*, 1997; Hindiyeh *et al.*, 1999).

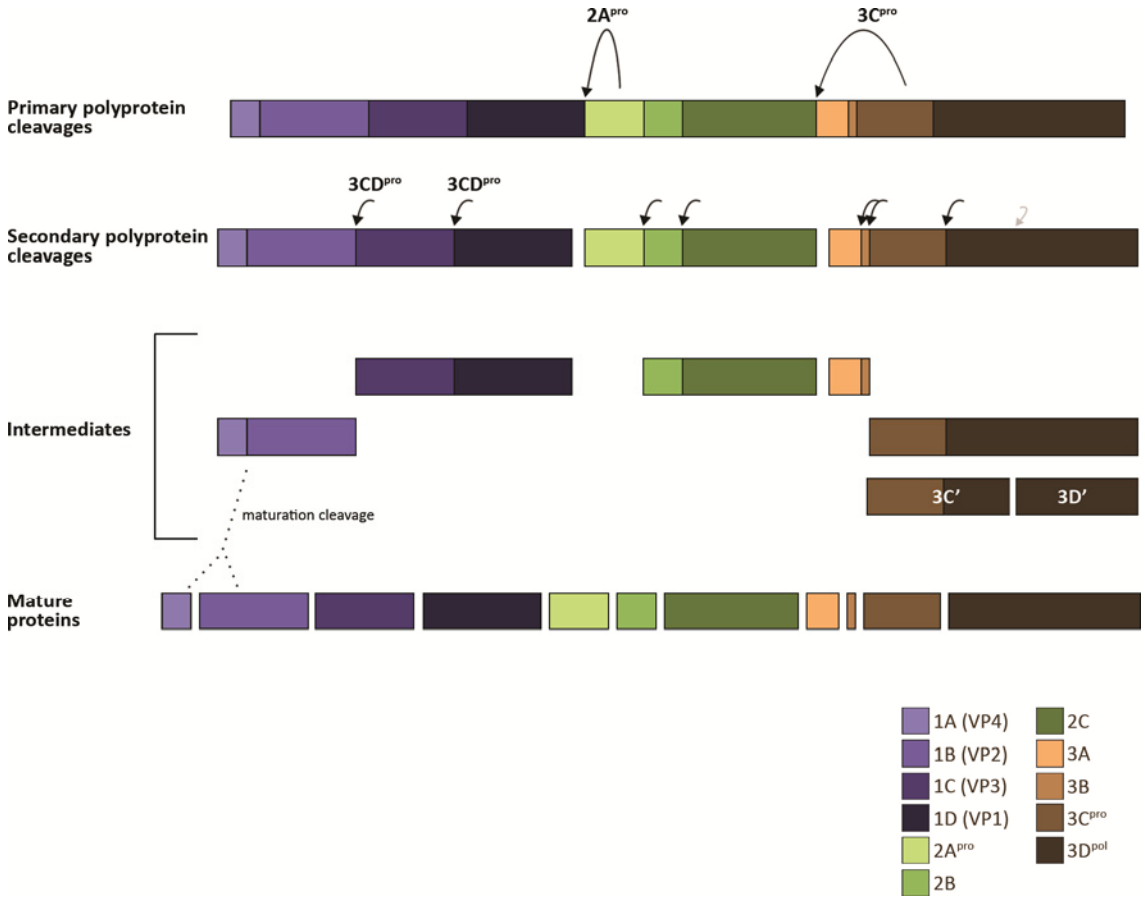


Figure 2-10 Schematic overview of enteroviral polyprotein processing. Proteins are depicted as box (Rueckert & Wimmer, 1984) with their length drawn to scale according to the amino acid sequence of the HRV2 polyprotein (Skern *et al.*, 1985). Large arrows in the primary polyprotein cleavage cascade indicate self-processing reactions by either $2A^{pro}$ or $3C^{pro}$. Black arrows in the cascade of secondary cleavages represent proteolysis in *trans* by $3C^{pro}$ or $3CD^{pro}$. The grey arrow indicates an alternative cleavage event within $3D^{pol}$ by $2A^{pro}$ in *trans* to yield 3C' and 3D'. Finally, the VP0 (1AB) precursor is cleaved by a yet unknown mechanism to yield mature viral capsids. Adapted from (Martínez-Salas & Ryan, 2010).

2.3.3 Replication of the genome

Picornavirus virions do not contain the viral proteins required for replication. Hence, translation of the viral genomes has to precede the replication process in order to produce those proteins. However, this requirement for translation in *cis* persists throughout the infection. For this reason, each copy of RNA has to be translated before being replicated (Collis *et al.*, 1992; Novak & Kirkegaard, 1994). Moreover, it was noted that actively translated RNAs cannot serve as

templates for replication (Barton *et al.*, 1999; Gamarnik & Andino, 1998). As a consequence, a switch from translation to replication is required. This switch is triggered by the accumulation of the viral protein 3CD which binds to stem loop I of the 5' UTR when a certain concentration of the protein is reached. Binding of 3CD induces a change in preference of the cellular protein poly(rC)-binding protein (PCBP2) to bind to stem loop I rather than to stem loop IV of the 5' UTR (Andino *et al.*, 1993; Gamarnik & Andino, 1997; 1998; Parsley *et al.*, 1997). The resulting change in occupancy of the stem loops together with the fact that PCBP2 is cleaved by $3CD^{pro}$ and $3C^{pro}$ later in infection (Perera *et al.*, 2007) is thought to trigger the switch from translation to replication.

Infections with picornaviruses result in extensive reorganization of the membranous system of the infected cell as first detected in 1958 (Kallman *et al.*, 1958). Further studies demonstrated the formation of vesicles, 50 to 400nm in size (Amako & Dales, 1967; Dales *et al.*, 1965; Dales & Franklin, 1962; Skinner *et al.*, 1968). Furthermore, viral replication was associated with these newly formed vesicles (Caliguiri & Tamm, 1969). Electron microscopic immunocytochemistry studies and the presence of membrane-binding domains suggest a role of the viral proteins 2B, 2C and 3A in tethering the RNA to the cytoplasmic face of the vesicles (Bienz *et al.*, 1990). The viral protein 3B (VPg) acts as a protein primer for the synthesis of negative as well as positive RNA strands (Paul *et al.*, 1998). Moreover, an RNA structure with a conserved AAAC sequence was first discovered to serve as the template for VPg uridylylation and termed *cis* replicative element (CRE) (McKnight & Lemon, 1998) (chapter 2.2.2).

The actual replication complex is a highly complex structure involving viral proteins and precursors of mature proteins, cellular proteins as well as RNA secondary structures. Herold and Andino presented a model that includes circularization of the RNA for the synthesis of (-) strand RNA (see Figure 2-11) (Herold & Andino, 2001). $3D^{pol}$ binding and synthesis of the (-) strand is dependent on the crosstalk between the RNA termini that is mediated by 3CD, PCBP2 and PABP. Negative stranded RNA is then used to template synthesis of multiple (+) stranded RNA molecules (Novak & Kirkegaard, 1991).

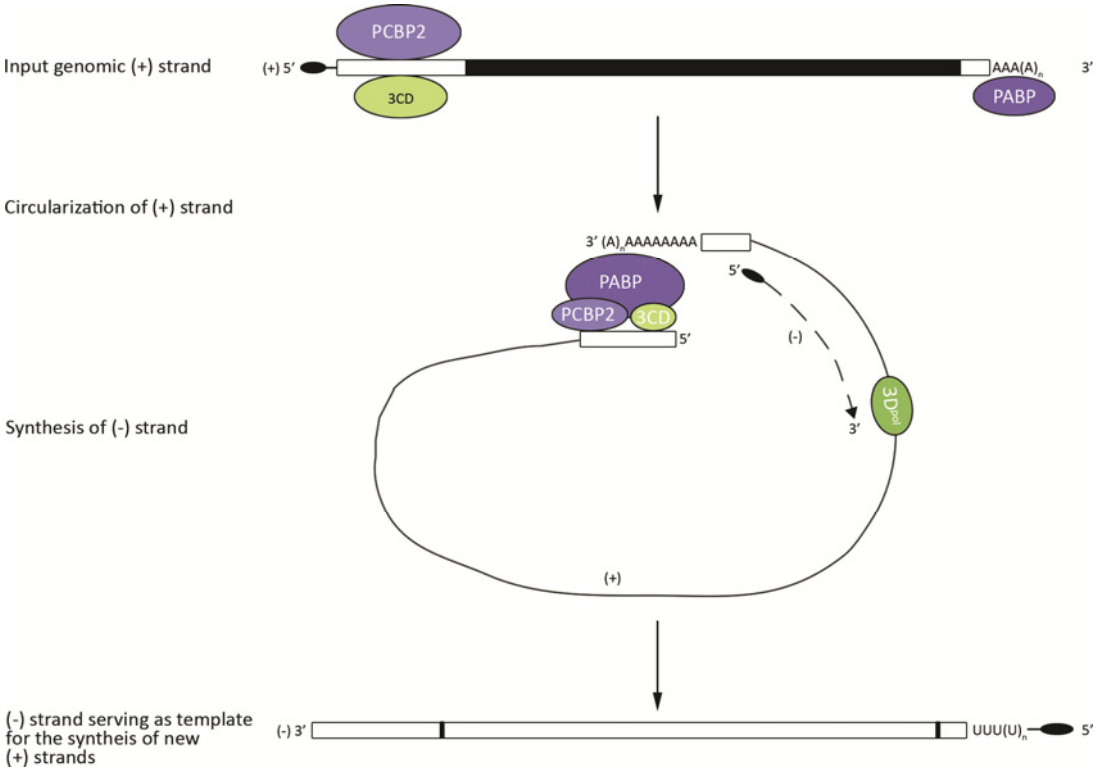


Figure 2-11 Schematic representation of (-) strand synthesis. Circularization of RNA was suggested to allow the long range crosstalk between 5' and 3' end. The ternary complex of 3CD, PCBP2 and stem loop I of the 5' UTR was found to interact with PABP that in turn binds the poly(A) tail at the 3' end of the RNA. Those interactions were found to be a prerequisite for $3D^{pol}$ binding and synthesis of the (-) strand.

2.3.4 Assembly and virus release

As stated in chapter 2.3.2.2 the P1 precursor, containing the four viral capsid proteins, is processed by $3CD^{pro}$ to VP0/VP3/VP1, the so called protomer, sedimenting at 5S (see Figure 2-12). Five of these protomers are able to self-assemble to form the 14S pentamer (Palmenberg, 1982). It has been shown that an empty 80S capsid can be formed by assembling 12 pentamers (Rombaut *et al.*, 1991; Rombaut *et al.*, 1984). However, the exact route of assembling the pentamers is unknown. Subviral particles sedimenting at 45S and 53S have been described (Lee & Colter, 1979; Rombaut *et al.*, 1985; Su & Taylor, 1976). Nevertheless, it is unclear whether these particles are bona fide intermediates of viral assembly. Furthermore, the role of 80S particles and the mechanism of RNA encapsidation is still a matter of debate. Formation of 80S particles seems to be reversible as dissociation to 14S pentamers has been described *in vitro* (Marongiu *et al.*, 1981; Onodera & Phillips, 1987). These observations agree with one possible mechanism for RNA packing. Nugent and Kirkegaard stated that RNA interacts with 14S pentamers before the entire

capsid is built to surround the genome (Nugent & Kirkegaard, 1995). In contrast, RNA could be inserted into fully built 80S particles by a yet unknown energy driven mechanism to build up 150S provirions (Jacobson & Baltimore, 1968). Finally, cleavage of VP0 (see chapter 2.3.2.2) leads to the formation of fully infectious, mature virions sedimenting at 160S. Progeny virus is released from the cell by lysis.

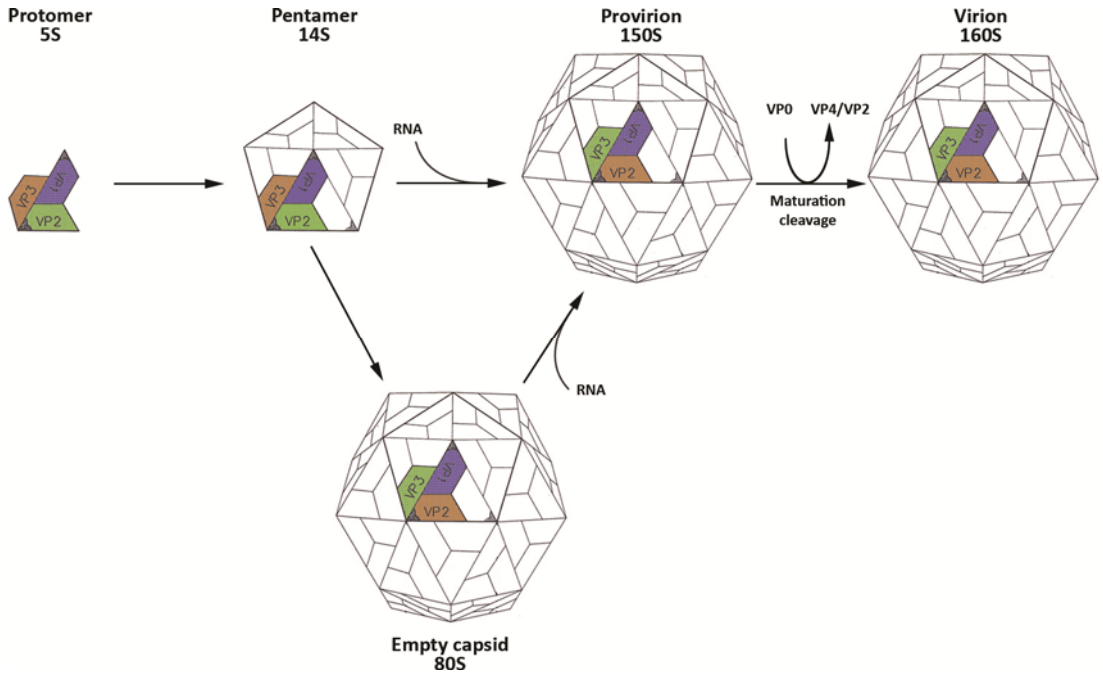


Figure 2-12 Schematic representation of viral assembly. The P1 capsid precursor is cleaved by $3CD^{pro}$ into VP0/VP3/VP1 to form the protomer. Five protomer units are assembled to yield the 14S pentamer. Twelve pentamer units are able to self-assemble into 80S empty capsids. Whether complete 80S capsids are loaded with RNA or 14S pentamers assemble directly with RNA in order to form the 150S provirions is still a matter of debate. Finally, cleavage of VP0 into VP4/VP2 forms the fully infectious 160S virion (maturation cleavage). Adapted from Fields *et al.* (Fields *et al.*, 2007).

2.4 THE 2A PROTEASE

2.4.1 Function

Since the discovery of the 2A protein as a protease in PV by Toyoda *et al.* in 1986 (Toyoda *et al.*, 1986), an appreciable number of proteolytic targets and functions has been assigned to $2A^{pro}$. This section summarizes the most important functions of this multifaceted enzyme.

2.4.1.1 Contribution to polyprotein processing by $2A^{pro}$

As described in chapter 2.3.2, $2A^{pro}$ of PV was first described as a protease by Toyoda *et al.* (Toyoda *et al.*, 1986). $2A^{pro}$ initiates enteroviral polyprotein processing by cleaving between the C-terminus of the preceding capsid protein VP1 and its own N-terminus in an intramolecular manner. This co-translational proteolysis thereby separates the structural from the nonstructural proteins. Furthermore, $2A^{pro}$ is the proteolytic agent responsible for an alternative processing event within the 3CD precursor. The processing of a Tyr-Gly bond leads to the formation of 3C' and 3D' (see chapter 2.3.2.2). Proteolytic activity of HRV and CV 2A proteins was subsequently demonstrated (Liebig *et al.*, 1993; Sommergruber *et al.*, 1989).

2.4.1.2 Cellular targets of $2A^{pro}$

One of the best and longest studied cellular targets of all enteroviral 2A proteases is the host translation factor eIF4G (eukaryotic initiation factor 4G) (see chapter 2.3.2.1) (Alvey *et al.*, 1991; Etchison *et al.*, 1982; Foeger *et al.*, 2003; Glaser & Skern, 2000; Gradi *et al.*, 1998b; Haghighat *et al.*, 1996; Lamphear *et al.*, 1993; Liebig *et al.*, 1993; Sommergruber *et al.*, 1994a; Ventoso *et al.*, 1998). Cleavage of both homologues of this protein is a key step in the so-called host-cell shut-off, a process by which the translation of cellular capped mRNAs is inhibited while IRES driven translation initiation is unaffected or even enhanced. Recently, cleavage of eIF4G by $2A^{pro}$ was also discovered as a prerequisite for the successful formation and the stability of polysomes in PV infected cells (Kempf & Barton, 2008). Moreover, the disruption of eIF4G might not only lead to 5'-cap independent translation initiation but also to independency from eIF2, a factor required for the recruitment of Met-$tRNA_i$ in canonical translation initiation (Redondo *et al.*, 2011).

Furthermore, the degradation of poly-A binding protein (PABP) in cells infected with enteroviruses has been associated with the onset of host-cell shut-off. Joachims *et al.* observed cleavage of PABP by PV $2A^{pro}$ and $3C^{pro}$ as well as CVB3 $2A^{pro}$ (Joachims *et al.*, 1999). In parallel, the ability of CV $2A^{pro}$ to cleave PABP was studied by Kerekatte and coworkers who were also able to map the cleavage site in PABP to $VANTSTQTM^{490}$ * $^{491}GPRPAAAAAA$ (the asterisk indicates the cleaved bond) (Kerekatte *et al.*, 1999).

Besides shutting down host cell translation, enteroviral infection also leads to an altered behavior in the shuttling of cargo between nucleus and cytoplasm. An accumulation of nuclear proteins in the cytoplasm of PV infected cells was first described by Gustin and Sarnow (Gustin & Sarnow, 2001). They demonstrated that nuclear import was blocked by the inability of receptor-cargo complexes to bind to the cytoplasmic face of the nuclear pore complex (NPC) and showed that nuclear pore proteins Nup153 and Nup62 were cleaved in PV infected cells. Soon, EM studies visualized the alterations in the NPCs and $2A^{pro}$ was suggested to be involved in the process (Belov *et al.*, 2004). Moreover, a third target for degradation in PV infected cells was identified as the interferon-induced Nup98 (Park *et al.*, 2008). However, cleavage of Nup98 is much faster than cleavage of Nup153 and Nup62 and relocalization of nuclear proteins was only detected after cleavage of Nup153 and Nup62 was completed. In addition, nuclear export of mRNAs, rRNAs and uridylate-rich snRNAs (U snRNAs) but not tRNAs was blocked upon cleavage of these three Nup's (Castello *et al.*, 2009). Again, $2A^{pro}$ was associated with the cleavage of Nup98 by the addition of purified HRV2 $2A^{pro}$ to cell lysates of uninfected cells. Finally, Nup62 was shown to be cleaved at multiple sites (Park *et al.*, 2010). The use of recombinant Nup62 and HRV2 $2A^{pro}$ allowed to demonstrate direct cleavage of Nup62 by $2A^{pro}$ and the mapping of six *in vitro* cleavage sites.

Besides the translational shut-off described above and in chapter 2.3.2.1, PV infected cells also suffer from inhibition of transcription by RNA polymerase II and III mediated by $3C^{pro}$ (Clark *et al.*, 1991; Clark *et al.*, 1993; Shen *et al.*, 1996; Yalamanchili *et al.*, 1996). However, transient expression of PV $2A^{pro}$ in eukaryotic cells results in a shut-down of both translation and transcription (Davies *et al.*, 1991). Subsequently, the ability of PV $2A^{pro}$ to directly cleave the TATA binding protein was demonstrated but was shown to have no influence on RNA polymerase II transcription *in vitro* (Yalamanchili *et al.*, 1997).

Other cellular targets of 2A proteases contain cytoskeletal proteins. Dystrophin was shown to be cleaved by CV $2A^{pro}$, thereby contributing to dilated cardiomyopathy (Badorff *et al.*, 1999). A second cytoskeletal protein, cytokeratin 8, is directly processed by CVB4 and HRV2 $2A^{pro}$ (Seipelt *et al.*, 2000). Finally, Graham *et al.* demonstrated direct cleavage of the catalytic subunit of the DNA-dependent protein kinase (DNA-PK(CS)) (Graham *et al.*, 2004) and the host factor gemin-3, a protein involved in the assembly of U snRNP complexes, was shown to be cleaved by PV $2A^{pro}$ (Almstead & Sarnow, 2007). However, although a later study confirmed the interference of $2A^{pro}$ with the splicing machinery (Alvarez *et al.*, 2011), gemin-3 remained intact under the conditions of this study. All in all, the role of $2A^{pro}$ in altering the process of pre-mRNA splicing remains to be elucidated.

2.4.1.3 Other functions

$2A^{pro}$ has also been connected to processes independent from its proteolytic activity. Furthermore, targets with unclear direct or indirect cleavage mechanism have been discussed.

$2A^{pro}$ was first shown to be involved in viral genome replication by using mutations that specifically altered its *trans* cleavage activity. When introduced into the genome of PV, these mutants showed defects in genome replication (Yu *et al.*, 1995). Furthermore, a cluster of negatively charged residues (Glu-Glu/Asp-Glu/Asp-Ala-Met-Glu-Gln) that is highly conserved (with the exception of 2A proteases of HRV) is required for replication independently of its proteolytic function (Li *et al.*, 2001). Finally, $2A^{pro}$ activity increases the stability of viral RNA, prolongs translation and stimulates negative strand synthesis by a yet unknown mechanism (Jurgens *et al.*, 2006).

In 2009, Morrison *et al.* presented data that demonstrated a new role of $2A^{pro}$ in interfering with the host immune system. PV, EV70 as well as HRV16 were able to replicate in cells that have been pretreated with interferon alpha (IFN-α), whereas encephalomyocarditis virus (EMCV), a picornavirus lacking 2A protease activity, was not (Morrison & Racaniello, 2009). However, EMCV was able to replicate in IFN-α treated cells when co-infected with PV or when the PV $2A^{pro}$ gene was added to the EMCV genome. These results suggest the interference of $2A^{pro}$ with interferon stimulated genes (ISGs).

The expression of $2A^{pro}$ from PV or EV71 alone was shown to be sufficient to drive cells into an apoptotic program. Hallmarks of apoptosis such as nuclear condensation and fragmentation, DNA breakdown, phosphatidylserine translocation and poly (ADP-ribose) polymerase (PARP) cleavage have been observed in such cells (Calandria *et al.*, 2004; Goldstaub *et al.*, 2000; Kuo *et al.*, 2002). PARP degradation was shown to be induced by $2A^{pro}$ but is inhibited by a caspase-3 inhibitor, suggesting that $2A^{pro}$ does not directly cleave PARP. Whether apoptosis is a consequence of eIF4G cleavage or $2A^{pro}$ induced degradation of yet unknown host factors remains unclear.

2.4.2 Structure

2.4.2.1 The X-ray structure of HRV2 2A^{pro}

Although using cysteine as the nucleophile, 2A protease was proposed to adopt a chymotrypsin-like fold before a high resolution structure could be solved (Bazan & Fletterick, 1988; Gorbalenya *et al.*, 1989; Sommergruber *et al.*, 1997). This proposal was confirmed in 1999 when Petersen *et al.* published the crystal structure of HRV2 2A protease at a resolution of 1.95Å (Petersen *et al.*, 1999) (Figure 2-13 A) and superimposed the coordinates of HRV2 2A^{pro} to the ones of *Streptomyces griseus* protease B (SGPB), a prototype for a minimal chymotrypsin-like fold. The β-strands of HRV2 2A^{pro} were assigned according to their homology to the strands found in SGPB. However, while the β-strands of the C-terminal β-barrels show reasonable superimposition, the N-terminal domain of HRV2 2A^{pro} lacks four of the β-strands (aI, bI1, dI and eI1) found in the N-terminal barrel of SGPB. The structure of HRV2 2A^{pro} can thus be interpreted as a chymotrypsin-like fold with a truncated N-terminal domain. The N-terminal β-barrel is reduced to a four stranded antiparallel β-sheet that packs on the C-terminal β-barrel composed of six antiparallel β-strands. Four short 3_{10}-helices (residues 17-19, 24-26, 62-64 and 133-136) complete the secondary structure elements of the protease. Finally, the N-terminus of HRV2 2A^{pro} adopts the conformation of a type I reverse β-turn that is strongly stabilized by an H-bond formed between the protonated α-amino group of Gly1 and the carboxy group of Asp4 (Figure 2-13 B).

The two domains are linked by a long interdomain loop that reaches from Thr41 to Thr55. An analogous loop in the second enteroviral protease 3C^{pro} includes the consensus sequence Lys-Phe-Arg-Asp-Ile. This sequence has been connected to RNA binding (Bergmann *et al.*, 1997; Matthews *et al.*, 1994; Mosimann *et al.*, 1997). However, the 2A protease is lacking this consensus sequence. In contrary, the interdomain loop accommodates the residues Cys52 and Cys54 that are involved in the coordination of a tightly bound Zn^{2+}-ion (Figure 2-13 C). The tetrahedral coordination also includes the side-chains of Cys112 and His114. As the distance of the catalytic triad to the Zn^{2+}-ion exceeds 20Å, a role in catalysis seems highly unlikely. The function of the Zn^{2+}-ion was rather suggested to stabilize the overall fold, thereby compensating the truncation of the N-terminal domain (Sommergruber *et al.*, 1994b; Voss *et al.*, 1995).

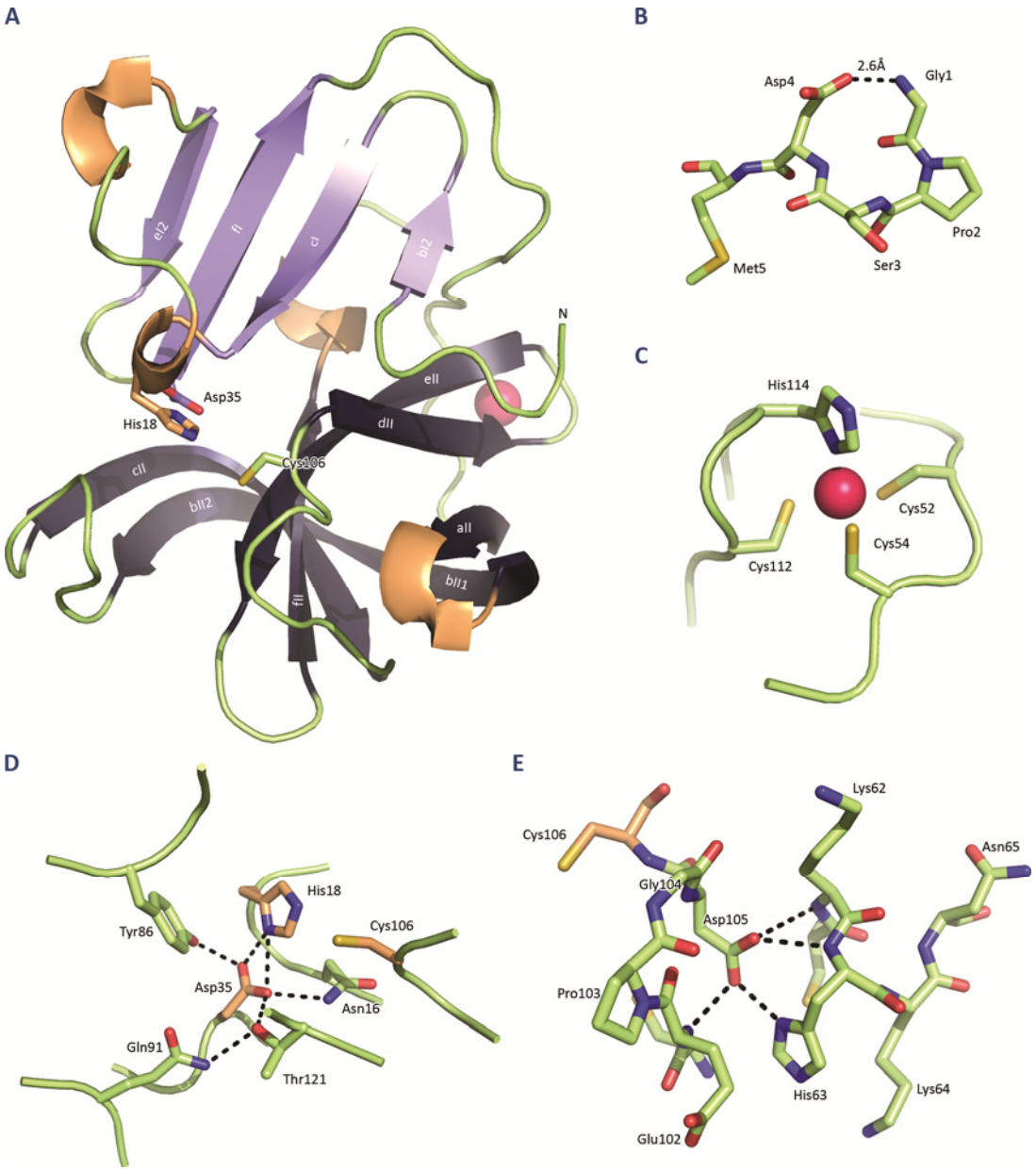


Figure 2-13 X-Ray structure of HRV2 2A^{pro} (PDB 2HRV). Elements are colored according to the following scheme: N (blue), O (red), C (green or rose), S (yellow), Zn (pink sphere). **(A)** Overall structure of HRV2 2A^{pro}. The two domains of chain A are colored in light blue (N-terminal) and dark blue (C-terminal) respectively. The three residues of the active site (His18, Asp35 and Cys106) are shown as sticks. Sheets are depicted as assigned by the dss script included in the molecular graphics program PyMol and labeled according to their homology to *Streptomyces griseus* protease B (Petersen *et al.*, 1999). **(B)** The N-terminus of HRV2 2A^{pro} forms a type I reverse turn stabilized by an H-bond between the protonated α-amino group of Gly1 and the side chain of Asp4. **(C)** The tightly bound Zn^{2+}-ion is tetrahedrally coordinated by Cys52, Cys54, Cys112 and His114. **(D)** The catalytic triad is composed of His18 (general base), Asp35 and Cys106 (nucleophile) and depicted as rose sticks. Black dashed lines indicate the H-bond network that stabilizes the active site. Side chains involved in this network are shown as sticks. **(E)** The H-bond network, stabilizing the oxyanion hole, is shown with dashed lines. All illustrations prepared from chain A of PDB 2HRV and rendered with the molecular graphics program PyMol (DeLano, 2002).

The catalytic triad is composed of His18 (the general base), Asp35 and Cys106 (the nucleophile) and is located at the junction of the two protein domains. His18 (located within the

second 3_{10}-helix) and Asp35 (situated at the beginning of strand fI) belong to the N-terminal domain whereas Cys106 is found between two type II β-turns preceding strand dII of the C-terminal β-barrel (Figure 2-13 A and D). Asp35 was found in the crucial role of positioning His18 in a catalytically active conformation. It builds H-bonds to the main chain amide as well as the $N^{\delta 1}$ of His18. Moreover, it is itself fixed by an H-bond network to Asn16, Tyr86, Gln91 and Thr121 (Figure 2-13 D). The catalytically active His18 and Cys106 were shown to exist as a thiolate-imidazolium ion-pair (Sarkany *et al.*, 2000).

The nucleophilic attack of Cys106 on the carbonyl of the scissile peptide bond causes the formation of a tetrahedral intermediate with a negative charge. This charge is stabilized by the so called oxyanion hole. In HRV2 $2A^{pro}$, two main-chain amides of the type II reverse β-turn preceding Cys106 serve as stabilizing groups. The β-turn itself is held in position by an H-bond network formed by the conserved residue Asp105 (Figure 2-13 E). The side chain carboxy group of Asp105 H-bonds to the main chain amide of Glu102, $N^{\delta 1}$ and the main chain amide of His63 as well as the main chain amide of Lys62, thereby stabilizing its relative position to the loop connecting the aII and bII1 strands.

2.4.2.2 The NMR solution structure of CVB4 $2A^{pro}$

Seven years after the release of the first $2A^{pro}$ high resolution structure, Baxter *et al.* published the NMR solution structure of CVB4 $2A^{pro}$ (Baxter *et al.*, 2006). Unlike the HRV2 $2A^{pro}$ construct described above, CVB4 $2A^{pro}$ was expressed as an inactive protease by mutating the active site cysteine to alanine (C110A). Moreover, the protease was N-terminally elongated by a hexa His-tag and the last eight amino acids of the preceding capsid protein VP1. Using such a construct, Baxter *et al.* aimed to observe a conformational state of the protease with its intramolecular cleavage substrate bound to the active site cleft.

The structure was solved using multidimensional heteronuclear NMR methods. The overall structure was found to be in good agreement with the crystal structure of HRV2 $2A^{pro}$. However, the signal intensities for the His_6-tag, the eight VP1 residues and the $2A^{pro}$ residues Gly1 to Ser7, Phe64 to Ser67 and Pro107 to Gly111 were reduced beyond the limit of detection and were missing in the spectra. The absence of these signals was explained by conformational exchange dynamics of the VP1/N-terminal substrate region with a frequency ranging from 100 to 1000 s^{-1}. The exchange of the substrate would also explain the loss of signals in the Pro107-Gly111

region (containing the nucleophile at position 110) as well as the Phe64-Ser67 region (aII-bII1 loop involved in stabilization of the oxyanion hole).

In conclusion, the NMR structure of CVB4 $2A^{pro}$ did not shed light on the exact binding mode of the substrate to the active site cleft during self-processing. However, the facts that 2A proteases share relatively low K_m values (~10^{-4}M) with k_{cat} constants in the range of $1s^{-1}$ (Liebig *et al.*, 1993; Sommergruber *et al.*, 1992) and the observation that the substrate would visit the active site up to 1000 times per second paint an interesting picture of how 2A proteases avoid product inhibition. Considering these values, the substrate would have to enter the active site 100-1000 times before being processed. Moreover, the low K_m value could either reflect a generally low affinity of the substrate to the binding cleft or the competition by substrate binding sites distinct from the active site.

2.4.3 Substrate specificity

2.4.3.1 PV $2A^{pro}$

Inspection of the VP1-$2A^{pro}$ junctions from different PV1, 2 and 3 isolates reveals a strong sequence conservation around the cleavage site at positions P4 to P7' (Figure 2-14). The only variable amino acid in this area is found at P2'. The aromatic side chains of Phe and Tyr as well as the hydrophobic side chain of Leu are found at this position. In an *in vitro* self-processing assay using mutated cleavage sites, Hellen *et al.* investigated the specificity of the protease in more detail (Hellen *et al.*, 1992). At P5, not only Gly but also Ala was shown to be accepted with only slight decrease in cleavage efficiency. Moreover, the invariable residue Thr at P3 was shown to be exchangeable with Ser. The P2 site was extensively mutated and observed. Mutating the P2 Thr to Ala, Ser, Asn, Gln or Pro did not reduce the cleavage efficiency beyond 83% (compared to wild-type). However, when one of the two bulky and hydrophobic residues Leu or Ile was introduced at P2, the efficiency of self-processing dropped to 50%. An even stronger effect was observed with the bulky and charged side chains of Glu and Arg. Interestingly, a number of amino acids with different physicochemical properties (Leu, Ile, Phe, Arg and Gln) were well accepted when introduced at P1 (cleavage efficiencies above 90%). However, poor cleavage was observed with Pro at P1. Gly at P1' is the only completely invariable residue found in the self-processing sites of all enteroviral 2A proteases. Nevertheless, PV $2A^{pro}$ is capable of processing cleavage sites that

have incorporated amino acids with side chains at P1'. However, cleavage efficiency drops with the size of the side chain with 85% for Ala, 73% for Ser, 32% for Cys, 31% for Leu and 23% for Thr.

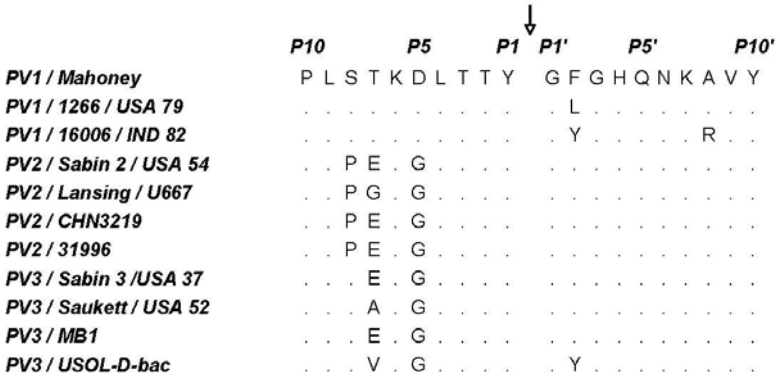


Figure 2-14 Amino acid sequence alignment of the VP1-2A^{pro} junction from different PV isolates. The sequences of the last ten residues of VP1 and the first 10 residues of 2A^{pro} from different PV1, 2 and 3 isolates were aligned using ClustalW (Larkin *et al.*, 2007) and rendered using the program ALINE (Bond & Schuttelkopf, 2009). Dots represent amino acids that are identical to the corresponding position in the reference sequence (PV1/Mahoney). The positions are labeled as P10-P1 and P1'-P10' according to the nomenclature of Berger and Schechter (Schechter & Berger, 1967). The arrow indicates the site of *cis* cleavage by 2A^{pro}.

In agreement with the *cis* cleavage data presented above, a Tyr to Phe mutation at the P1 position of the alternative cleavage site within 3D^{pol} has no influence on *trans* processing at this site to yield 3C' and 3D' (Lee & Wimmer, 1988). However, mutating the P2 Thr to Ala prevented cleavage. In order to further clarify the specificity of PV 2A^{pro} in *trans* Hellen *et al.* tested cleavage of an inactive VP1-2A^{pro} precursor by extracts of infected HeLa cells (Hellen *et al.*, 1992). Comparison of the results in *cis* and *trans* suggests that the specificity of PV 2A^{pro} is more stringent in *trans* as all mutations introduced in the cleavage site led to less effective cleavage compared to the results obtained in *cis*. The cleavage site at which eIF4G is cleaved by PV 2A^{pro} has not yet been mapped exactly. However, a peptide derived from eIF4G is cleaved between Arg and Gly corresponding to the same site utilized by HRV2 and CVB4 2A^{pro} (Ventoso *et al.*, 1998).

2.4.3.2 HRV 2A^{pro}

The specificity of rhinoviral 2A^{pro} has been studied in more detail. The 2A proteases of HRV2 and HRV14 have served as representatives for intensive studies. However, exact understanding is complicated by the differences present in the 2A proteases of genetic group A and B. Figure 2-15 shows the cleavage sites of HRV2 and HRV14 on their own polyprotein as well

as the two cleavages sites of HRV2 2A^{pro} on eIF4GI and eIF4GII (the HRV14 2A^{pro} cleavage sites have not been mapped at present, but are considered identical). Quick inspection of these sequences already reveals substantial variations within the cleavage sites. The sites determining specificity seem to lie within the P4, P2 and P1' positions and are discussed in more detail below.

Early work on substrate specificity concentrated on cleavage assays using synthetic peptides. The minimal length of a peptide that achieves full cleavage efficiency was determined to span P8-P2' (Sommergruber *et al.*, 1992; Wang *et al.*, 1998b). The P4 position shows strong preference for bulky hydrophobic residues such as Ile, Val and Leu in the self-processing reaction in both group A and group B (see Figure 2-16). Leu is also found at P4 of the *trans* substrates eIF4GI and eIF4GII (Figure 2-15). Moreover, all six of the mapped cleavage sites of HRV2 2A^{pro} within the nuclear pore protein Nup62 have hydrophobic side chains (Figure 2-17) (Park *et al.*, 2010). Petersen *et al.* proposed a substrate binding model (Petersen *et al.*, 1999) with the X-ray structure of HRV2 2A^{pro}. By identifying a narrow hydrophobic S4 pocket, the model supplies a reasonable explanation for the preferences at P4. For this reason, peptides with Lys or Asp at P4 could not be cleaved at all by HRV2 2A^{pro}. Moreover, Phe and Ser at P4 strongly decreased cleavage efficiency of those peptides (Sommergruber *et al.*, 1994a).

Assuming a more or less extended β-strand like conformation of the substrate, the P3 side chain would point away from the substrate binding cleft and towards the solvent. A specific recognition of the P3 side chain is therefore unlikely. Indeed, a variety of amino acids can be found at the P3 position in *cis* as well as in *trans* substrates as can be seen in Figure 2-15, Figure 2-16 and Figure 2-17. Nevertheless, peptides with different amino acids at P3 are not equally well accepted by HRV2 2A^{pro}. Whereas Lys and Phe do not drastically lower cleavage efficiency, Ser and Asp at P3 reduce the relative V_{max}/K_m by more than four-fold (Sommergruber *et al.*, 1994a).

The P2 position was extensively examined by the use of synthetic peptide cleavage assays with HRV2 2A^{pro} (Sommergruber *et al.*, 1992). The wild-type Thr was mutated to all other amino acids with the exception of Trp, Cys and Met. All of the mutated peptides were either not cleaved at all or showed severe decrease in cleavage efficiency. Interestingly, the substitution of the wild-type ThrThrAla to LysSerTyr at P3-P1 of the HRV2 2A^{pro} self-processing site did not abolish cleavage in *cis* (Neubauer, 2008). Moreover, although the frequency analysis (Figure 2-16) in the HRV 2A^{pro} self-processing sites reveals a preference for Thr at P2, it is not the only residue found at this position. Asn and Ser are frequently found at P2 in *cis* and in *trans* substrates as well. Indeed, HRV14 2A^{pro} recognizes a peptide resembling its self-processing cleavage site with a Ser to Thr mutation at P2 very badly. Petersen *et al.* suggested a model in which Tyr85 would have to rotate away from its stacked position with His18 in order to provide sufficient space for the P2 Thr

(Petersen *et al.*, 1999) to bind. A hydrogen bond between the P2 Thr and Ser83 should compensate for the loss of stacking interaction. However, though Ser83 is strictly conserved in $2A^{pro}$ of group A HRVs, also Gly and Asn are found in 2A proteases of group B HRVs at this position. Interestingly, three of the twelve serotypes with Gly at the equivalent position of Ser83 in HRV2 $2A^{pro}$ self-process at sites with the bulkier side chains of Gln or Lys at P2.

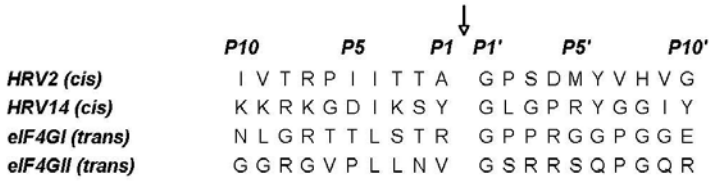


Figure 2-15 Cleavage sites of HRV2 and HRV14 $2A^{pro}$ mapped on *cis* as well as on *trans* substrates. The cleavage sites are aligned according to their P10 to P10' positions. The sequences of the HRV2 and HRV14 self-processing sites as well as the HRV2 cleavage sites on eIF4GI and eIF4GII are given. The arrow indicates the cleavable bond. Rendered with ALINE (Bond & Schuttelkopf, 2009).

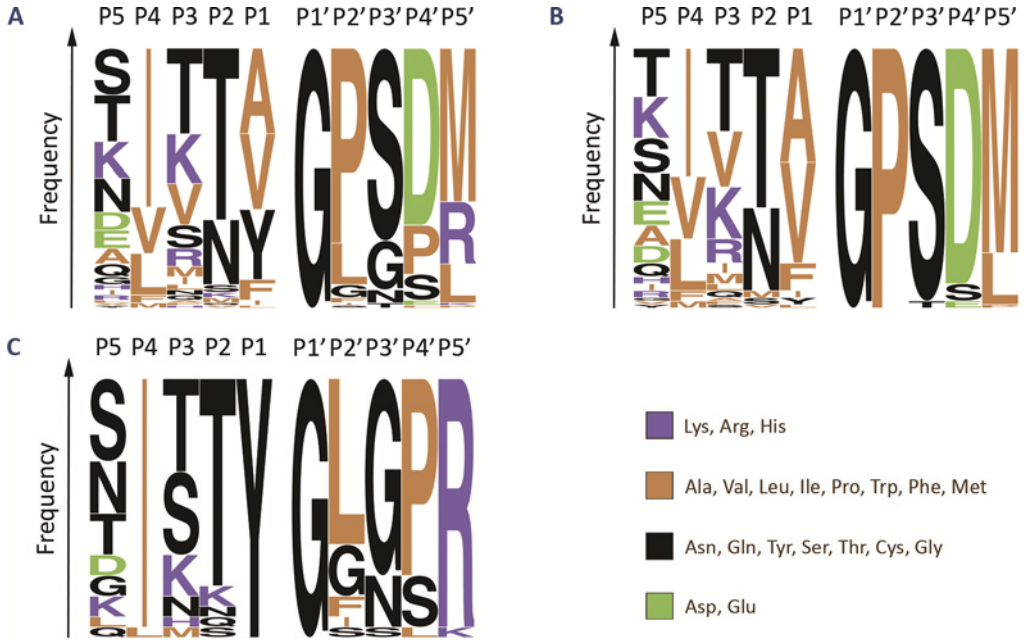


Figure 2-16 Frequencies of amino acids in self-processing sites of HRV $2A^{pro}$. The frequency of a corresponding amino acid at a specific position within a cleavage site is depicted by the size of its one letter code. All illustrations are rendered with WebLogo (Schneider & Stephens, 1990). **(A)** Frequencies in a set of $2A^{pro}$ sequences of all group A and group B HRVs. **(B)** Frequencies in a set of $2A^{pro}$ sequences of all group A HRVs. **(C)** Frequencies in a set of $2A^{pro}$ sequences of all group B HRVs.

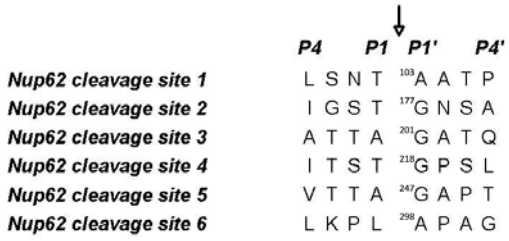


Figure 2-17 P4-P4' positions of all six identified cleavage sites of HRV2 2A^{pro} within the nuclear pore protein Nup62. The arrow indicates the cleavable bond. P1' residues are labeled with their respective amino acid number within the Nup62 protein (Park *et al.*, 2010). Rendered with ALINE (Bond & Schuttelkopf, 2009).

A number of amino acids is found at the P1 position of HRV VP1-2A^{pro} junctions with Ala, Val, Tyr and Phe being the most abundant (Figure 2-16). Interestingly, Tyr is strictly conserved in the self-processing sites of group B viruses. However, HRV2 2A^{pro} was found to cleave peptides with Met or Tyr at P1 even better than peptides with the wild-type Ala at this position. No cleavage of peptides was observed when Pro, Asp or Glu were introduced at P1 and cleavage efficiency dropped drastically with Lys, Gln, Asn or Val at P1 (Sommergruber *et al.*, 1992). The amino acids found at the P1 positions of the six HRV2 2A^{pro} cleavage sites on Nup62 agree well with these experiments (Figure 2-17) (Park *et al.*, 2010). However, the superiority of Tyr and Met at P1 could not be confirmed in *cis* reactions. Although Met and Tyr as well as Phe, Thr and Leu were found to be well accepted at P1 in *cis*, increased cleavage efficiency could not be detected when introduced at P1 of the self-processing site (Skern *et al.*, 1991). On the contrary, Ile and Val at P1 led to a significant reduction of cleavage efficiency. In addition, the P1 position of HRV 2A proteases seems to be a determinant of differences in specificity observed between group A and group B viruses. In a study using a self-processing assay, Sousa *et al.* demonstrated that HRV2 2A^{pro} was able to process the eIF4GI cleavage site (TLSTR * GPPR) in *cis* whereas HRV14 2A^{pro} could not (Sousa *et al.*, 2006). The Arg at P1 was pinpointed as the residue causing this phenotype. A partial rescue of HRV14 2A^{pro} activity could be achieved by mutating Ala104 (strictly conserved within group B viruses) to Cys or Ser, found at the equivalent position of group A viruses.

Gly at P1' is the only strictly conserved amino acid in all enteroviral VP1-2A^{pro} junctions. Consequently, Gly at P1' was defined as a prerequisite for cleavage activity by 2A^{pro}. The introduction of Glu, Lys, Thr and Trp at P1' was demonstrated to abolish self-processing (Skern *et al.*, 1991) and no cleavage of peptides with either Phe, Thr, Asp or Lys at P1' could be detected (Sommergruber *et al.*, 1992). Petersen *et al.* suggested that the side chain of the conserved Leu19 would block the space for amino acids other than Gly at P1' (Petersen *et al.*, 1999). However, the

presence of an Ala at P1' in two of the six Nup62 cleavage sites (Figure 2-17) dismantled the paradigm of the invariant Gly at P1' (Park *et al.*, 2010). Nevertheless, Ala at P1' could not be accepted by HRV2 2A^{pro} in the *cis* reaction. Furthermore, a Leu19Ala mutation could not rescue cleavage activity of this P1' mutant (Neubauer *et al.*, 2013). Hence, HRV2 2A^{pro} displays different specificity at P1' dependent on *cis* or *trans* reactions. The reasons for these differences remain unclear.

The P2' position within the VP1-2A^{pro} junction is predominantly occupied by Pro (group A viruses) as well as Leu (group B viruses) (Figure 2-16). Peptides with Lys, Thr, Phe or Asp are cleaved 2.5-5 fold less effectively than peptides with the wild-type Pro by HRV2 2A^{pro} (Sommergruber *et al.*, 1992). In contrast, the introduction of Pro at P2' in combination with Thr at P2 into a peptide derived from the HRV14 VP1-2A^{pro} junction strongly increases cleavage efficiency of HRV14 2A^{pro} on such a peptide (Wang *et al.*, 1998b).

Recently, we have noticed that the specificity of HRV2 2A^{pro} in *cis* cannot be explained by a site by site analysis of accepted and rejected amino acids but rather must be seen in the light of combinatorial requirements on several sites (Neubauer *et al.*, 2013). HRV2 2A^{pro} was found unable to cleave the HRV14 VP1-2A^{pro} cleavage site (DIKSY*GLGP, underlined residues indicate amino acids that differ from the wild-type HRV2 site) in *cis*. However, introduction of equivalent mutations at just one side of the cleavable bond (IIKSY*GPSD or IITTA*GLGP) led to just minor delay in self-processing. The minimal number of mutations causing a similar phenotype than the DIKSY*GLGP site was found to be a combination of mutations on both sides of the cleavable bond (IIKSY*GLSD).

2.4.3.3 CV 2A^{pro}

The CV 2A proteases show a similar pattern of preferences in their VP1-2A^{pro} self-processing sites as the HRV 2A proteases. Nevertheless, CVB4 2A^{pro} cleaves a peptide resembling the HRV2 VP1-2A^{pro} self-processing site with very poor efficiency (Sommergruber *et al.*, 1994a). The P4 position is predominantly populated by the hydrophobic residues Ile and Leu (Figure 2-18). Thr and Asn are the only amino acids found at the P2 positions of CV 2A^{pro} self-processing sites and the P1' position is exclusively occupied by Gly. Interestingly, six different amino acids are found at the P1 positions of the 21 VP1-2A^{pro} junctions of CVA. In contrast, the P1 position of the six CVB cleavage sites is populated by a conserved Thr. Indeed, CVB4 2A^{pro} suffered from poor performance in cleaving a peptide with Ala or Val at P1 (Sommergruber *et al.*, 1994a). However, a

peptide derived from the HRV2 VP1-2A^{pro} site was cleaved with 12-fold increased efficiency when Ala at P1 was replaced by Tyr.

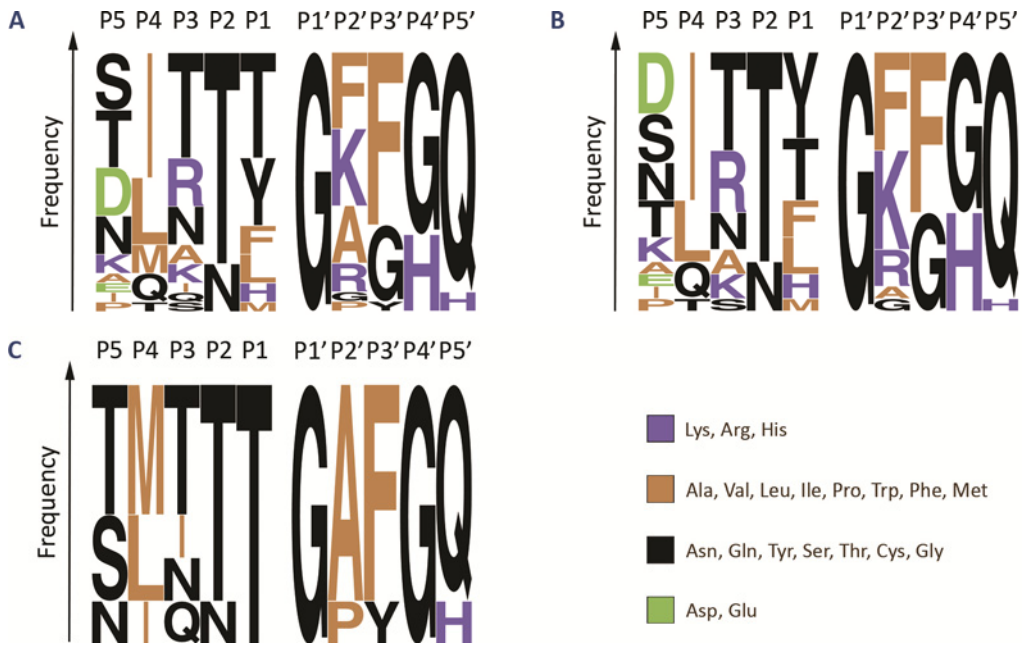


Figure 2-18 Frequencies of amino acids in self-processing sites of CV 2A^{pro}. The frequency of a corresponding amino acid at a specific position within a cleavage site is depicted by the size of its one letter code. All illustrations are rendered with WebLogo (Schneider & Stephens, 1990). **(A)** Frequencies in a set of 2A^{pro} sequences of all CVA and CVB serotypes. **(B)** Frequencies in a set of 2A^{pro} sequences of all CVA serotypes. **(C)** Frequencies in a set of 2A^{pro} sequences of all CVB serotypes.

2.4.4 2A^{pro} – a drug target

Proteases have been of interest for the development of antiviral compounds for a long time. The two enteroviral proteases 2A^{pro} and 3C^{pro} represent good targets for antiviral drugs for two reasons. First of all, inhibition of 2A^{pro} and 3C^{pro} would interfere with polyprotein processing as well as many of the virally induced alterations that enables the virus to hijack the infected cell. Secondly, their usage of cysteine as the nucleophile in a chymotrypsin-like fold, their specificities and the unique Zn^{2+} binding site of 2A^{pro} distinguishes them from cellular proteases. So far, most interest was put into the development of 3C^{pro} inhibitors. Indeed, the compound AG7088, also known as rupintrivir, was already used in clinical trials but was recently stopped from further development. However, in a recent study 2A^{pro} was awarded a great potential as a drug target. Co-transfection of HeLa cells with wild-type PV RNA and PV RNA including mutations that inhibit

$2A^{pro}$ self-processing resulted in the inhibition of both wild-type and mutant virus replication (Crowder & Kirkegaard, 2005). Presumably this phenotype is a consequence of unprocessed VP1-$2A^{pro}$ precursor molecules that interfere with correct capsid formation. These results imply that a $2A^{pro}$ inhibitor could be successful, albeit an efficiency below 100%. Furthermore, the presence of unprocessed VP1-$2A^{pro}$ would theoretically impede the packing of a viral genome that has acquired mutations leading to resistance and therefore the escape of resistant virus. The following section tries to summarize attempts in developing antiviral strategies against the proteolytic activity of enteroviral 2A proteases.

Early studies investigated the effect of canonical inhibitors on the ability of HRV2 $2A^{pro}$ to cleave peptides (Sommergruber *et al.*, 1992). The serine protease inhibitors elastatinal and chymostatin, the cysteine protease inhibitors iodoacetamide and NEM (N-Ethylmaleimide) and the cysteine/serine protease inhibitors antipain and TLCK (tosyl lysyl chloromethyl ketone) all showed IC_{50} values in the 25-90µM range. Similar values for these inhibitors were detected with the use of HRV14 $2A^{pro}$ (Wang *et al.*, 1998b). Furthermore, PV $2A^{pro}$ was shown to be sensitive to elastatinal (IC_{50}=4µm), MPCMK (IC_{50}=20µM), antipain (IC_{50}=23µM), calpain inhibitor 1 (IC_{50}=28µM) and iodoacetamide (IC_{50}=1.2mM) in a *trans* cleavage assay using an inactive VP1-$2A^{pro}$ precursor as a substrate. The presence of 1mM MPMCK in the culture medium also reduced the viral titer obtained from infection of HeLa cell monolayers by 100-1000 fold for PV1-3, HRV2 and 14 as well as CVA21 (Molla *et al.*, 1993). Finally, Wang *et al.* demonstrated the inhibition of HRV2 and HRV14 $2A^{pro}$ by a series of homophtalimides with IC_{50} values in the low to middle µM range (Wang *et al.*, 1998a).

A number of inhibitors based on peptides have been developed and tested against enteroviral $2A^{pro}$. Antiviral activity of the caspase inhibitor zVAD.fmk against HRV was first described in 2004 (Deszcz *et al.*, 2004). Its specificity was drastically increased by the exchange of Asp at P1 by Met (Deszcz *et al.*, 2006). zVAM.fmk inhibits HRV2, HRV14 and HRV16 replication in the µM range but does no longer inhibit caspase 9 activation. Interestingly, the results of the study suggest that the antiviral effect of zVAM.fmk is based on targeting two different functions of $2A^{pro}$. A nanomolar concentration of zVAM.fmk was needed to block cleavage of eIF4GI by HRV2 $2A^{pro}$, whereas a much higher concentration was needed to block eIF4GI cleavage by HRV14 $2A^{pro}$. However, inhibition of intramolecular self-processing was more efficient with HRV14 $2A^{pro}$. Recently, a similar inhibitor was shown to inhibit $2A^{pro}$ of HRV. The peptide LVLQTM was chemically modified to zLVLQTM.fmk and exhibited an IC_{50} value of 0.3µM in peptide cleavage assays of HRV2 $2A^{pro}$ (Falah *et al.*, 2012b). Furthermore, zLVLQTM.fmk was effective in inhibiting eIF4G cleavage in HRV2, HRV14 and EV-6 $2A^{pro}$ expressing A549 cells in the high µM range. A two

log decrease in viral yield (as measured by $TCID_{50}$) was observed when HRV2 and HRV14 infected A549 cells were pretreated with 200µM of zLVLQTM.fmk and the inhibitor also decreased the viral load in the lungs of HRV2 infected mice. Finally, the peptide LVLQTM was shown to act as a pseudosubstrate of EV-71 $2A^{pro}$ with a K_d of 9.6 ± 3.9 µM (as measured by isothermal titration calorimetry (ITC)). In addition, zLVLQTM.fmk was able to inhibit eIF4G cleavage in EV-71 expressing HeLa cells and performed equally well in the inhibition of EV-71 replication in infected HeLa cells (Falah *et al.*, 2012a). Another substrate analogue inhibitor was developed against the 2A protease of CVB3 after its cleavage site within the cellular protein dystrophin has been mapped (Badorff *et al.*, 2000). Again the peptide was C-terminally modified with a fluoromehtyl ketone group (zLSTT.fmk) and exhibited an IC_{50} value of 550nm.

A different approach, using siRNA raised against the 2A genomic region of CVB3 in infected HeLa cells, resulted in a >1000-fold reduction of viral yield in early stages of infection (Merl *et al.*, 2005). Both viral replication as well as expression of viral proteins were shown to be inhibited. However, inhibition could only be maintained as long as siRNA against 2A was efficiently expressed. Administration of siRNA raised against 2A to type I interferon receptor deficient mice significantly increased their life span and even enabled survival of one mouse after infection with a lethal dose of CVB3.

2.5 AIM OF THE WORK

The fight for a PV free world finally seems to come to an end (see chapter 2.1.3.2). Once wild PV circulation has been stopped, the WHO suggests discontinuing of OPV use in order to decrease the risk of vaccine derived polio outbreaks (Aylward *et al.*, 2005). However, such outbreaks cannot be precluded. Furthermore, PV infected, immunodeficient individuals who are chronically shedding virus could put the increasing unprotected portion of the population at risk. A panel assembled by the national research council therefore considered the development of antiviral drugs as a beneficial if not essential tool in the final fight against PV eradication and for the post-eradication era (Couzin, 2006). Though not as devastating as PV, the mild infections caused by HRVs cause enormous economic damages (see chapter 2.1.1). An antiviral drug targeting viral replication rather than drugs targeting the symptoms of the infection could also be of great benefit.

Research on $2A^{pro}$ inhibitors has been limited so far. However, $2A^{pro}$ was identified as an excellent drug target via its *trans* dominant phenotype (Crowder & Kirkegaard, 2005). The example of nowadays excellent possibilities in controlling HIV replication in infected individuals has taught the scientific community that drugs against several different viral targets are essential. Research on $2A^{pro}$ inhibitors in addition to other lines of research should therefore clearly push the efforts towards an antiviral therapy against enteroviral infections.

The lack of a high resolution structure of $2A^{pro}$ bound to substrate in either *cis* or *trans* has hampered our gain in knowledge of how exactly $2A^{pro}$ binds substrates and maintains its specificity. Such information would in fact be beneficial for the development of $2A^{pro}$ inhibitors as a structure guided approach would allow a rationality driven development of new lead compounds as well as the enhancement of already tested inhibitors.

We therefore aimed to obtain new high resolution structures of enteroviral 2A proteases with a focus on PV $2A^{pro}$ and on an X-ray snap-shot of $2A^{pro}$ in the process of *cis* cleavage.

3 Materials and Methods

3.1 PLASMIDS

3.1.1 Expression vectors

The pET System (Novagen) offers the possibility for strong over-expression of recombinant proteins in *E. coli*. Sequences of interest are cloned under the control of the T7 promoter that allows transcription of the target gene by T7 polymerase. Such constructs are transferred into suitable *E. coli* strains, carrying a chromosomal copy of the T7 RNA polymerase gene that is under the control of the *lacUV5* promoter. Induction with IPTG or lactose leads to strong over-expression of the protein of interest.

3.1.1.1 pET-3d (formerly known as pET-8c)

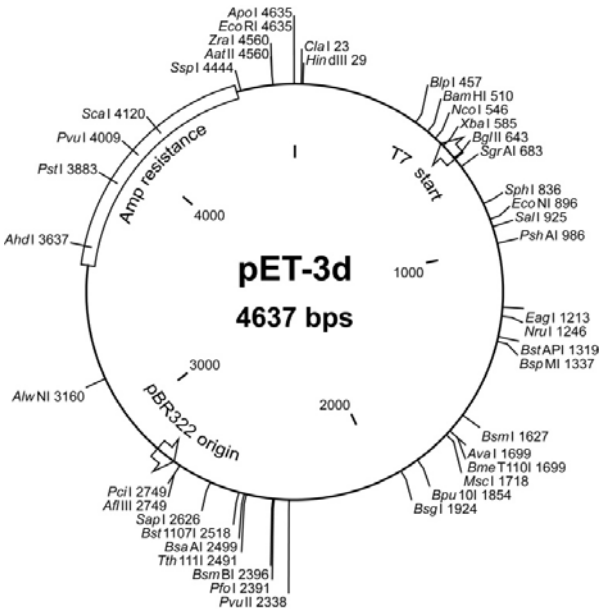


Figure 3-1 Plasmid map of the expression vector pET-3d from Novagen. Single cutter restriction sites are plotted with their respective positions.

pET-3d (Novagen) allows in frame cloning of a fragment between its *Nco*I and *Bam*HI sites to express constructs without tags, or in frame cloning using just the *Bam*HI site to produce protein tagged with the T7-tag® on its N-terminus.

3.1.1.2 Mal-pET

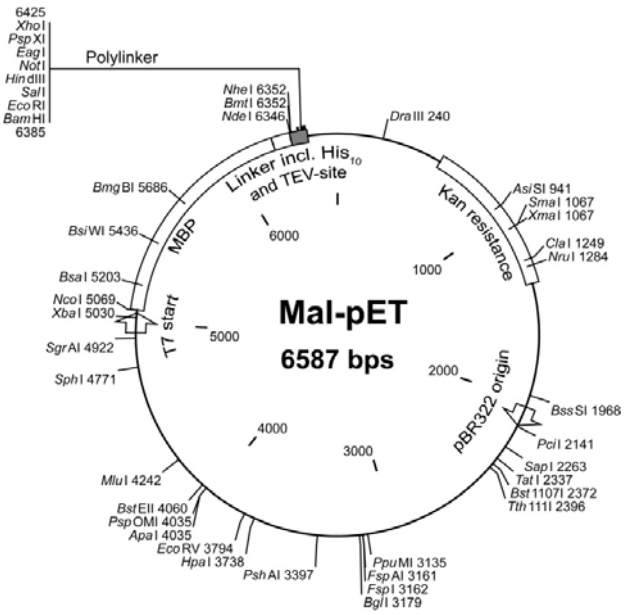


Figure 3-2 Plasmid map of the expression vector Mal-pET (kind gift of Gang Dong). Single cutter restriction sites are plotted with their respective positions.

The plasmid Mal-pET was a kind gift of Gang Dong. It contains the coding sequence of *E. coli* maltose binding protein (MBP) in frame with a 16 amino acids linker, a His_{10}-tag and a recognition site for TEV-protease. cDNA coding for a protein of interest can be cloned between an *NdeI* or *Nhe*I site and one of the sites in the polylinker. Upon induction the protein of interest is expressed with an N-terminal MBP-tag that serves as solubility tag as well as purification tag.

3.1.1.3 pET-MBP 1a

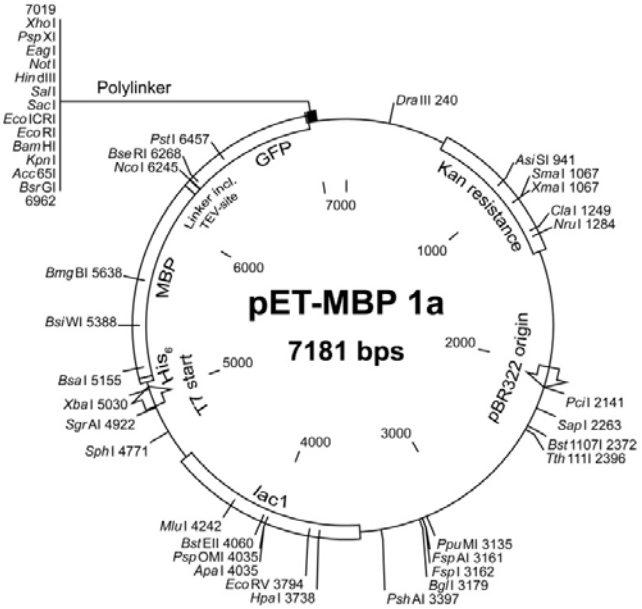


Figure 3-3 Plasmid map of the expression vector pET-MBP 1a. Single cutter restriction sites are plotted with their respective positions.

The plasmid pET-MBP 1a is modified from pET-24d. Compared to the plasmid Mal-pET it codes for a His_6-tag, N-terminally fused to MBP, the MBP itself, a shorter 6 amino acid linker plus a TEV protease recognition site. The sequence for the protein of interest is cloned in frame between an *Nco*I site and a restriction site of the polylinker, thereby excising and replacing the GFP gene.

3.2 OLIGONUCLEOTIDES

All oligonucleotides were purchased from either Sigma or Eurofins.

Identifier	5' Sequence 3'
TIM 554	atttaggtgacactatag
TIM 1047	aggccccaaggggttatg
TIM 1085	tccgaaccaggagatgccggtggaattctcagg
TIM 1086	cctgagaattccaccggcatctcctggttcgga
TIM 1704	tctccaggggatgctggtggcatactc
TIM 1705	gagtatgccaccagcatcccctggaga
TIM 1716	agatatatcatgactcaccaccatcaccatcactccaccaaggatctgacc
TIM 1717	agatatatcatgagcaccaaggatctgaacac
TIM 1718	agatatatcatgagggcagtggcgtactac
TIM 1728	acccaactgatcttcagc
TIM 1729	tcagcggcactcgagccag
TIM 1730	gatcctggctcgagtgccgc
TIM 1759	ccgggctagcagcaccaaggatctgaccacatatgg
TIM 1760	atcggatccctattattgttccatggcttcttcttc
TIM 1779	tcgatgaagccctgaaagac
TIM 1788	gcacgcatcctcgcattatc
TIM 1789	gctagcagtctgcgcgtctttc
TIM 1803	gctcattgggcatggattcgc
TIM 1804	gcgaatccatgcccaatgagc
TIM 1805	cgttagccatggcggtgtacactgcagg
TIM 1806	ctggatggatcctattagtaggcatacaagtctc
TIM 1821	gcactagccatggtgaccagaccaattatcactacagc
TIM 1833	acctggggatccctattaatggtgatggtgatggtgtgggccttgttccatggcttcttcttcgtaggc
TIM 1858	gtcgcacatatggagcgcgctagcttgataacc
TIM 1882	tagctcactcatgagcttgataaccacaggcccctatg
TIM 1883	tagctcactcatgagcgccgttacagcggagcgtgcaagc
TIM 1884	tagctcactcatgaagagtgtgaactttgatgttgagg
TIM 1885	gtacatcgggatccttatcactgttccattgcatcatcttccaac
TIM 1899	ctggaccgattcatgagcggggccgtgtatgtgggcaattac
TIM 1900	ctggaccgaggatccttatcacaaccacaacaggtcacggac

Table 3-1 5' to 3' sequences of all oligonucleotides used in this work.

3.3 DNA METHODS

3.3.1 Agarose gel electrophoresis

Agarose gel electrophoresis was used to monitor all DNA work and for purification of DNA fragments. An appropriate amount of agarose (Invitrogen) was melted in 0.5x TAE-buffer and casted in a BioRad mini gel chamber. DNA samples were mixed with 10x DNA loading buffer and *Hind*III digested λ-DNA was used as a molecular weight marker. Gels were run at a constant Voltage of 130V with 0.5x TAE buffer. Subsequently, gels were soaked in ethidiumbromide solution for 15-30 minutes. DNA bands were either visualized on a GelDoc-It^{TM} TS imaging system or on a Herolab UVT-20M (for cutting bands).

0.5x TAE-buffer	20mM Tris base, 5mM Sodiumacetate, 1mM EDTA.
10x loading buffer	1mM EDTA, 0.1% Orange G, 10% Ficoll in 0.5x TAE-buffer.
Ethidiumbromide solution	10^{-4}% Ethidiumbromide in 0.5x TAE-buffer

3.3.2 Restriction reactions

Restriction enzymes of NewEngland BioLabs, Fermentas and Promega were used for standard DNA restriction reactions using appropriate buffers provided by the companies. Reactions were incubated at the stated temperatures on an Eppendorf Thermomixer 5436.

3.3.2.1 Preparative reactions

For the preparation of fragments for the cloning of new constructs up to 10µg of DNA were digested in a 100µl reaction using 0.5-0.7µl of the particular enzyme. Incubation lasted from 2-14 hours till the digest was completed.

3.3.2.2 Analytical reactions

In order to analyze plasmids via restriction digest, up to 1µg of DNA was digested in 20µl reactions using 0.3µl of the particular enzymes. Reactions were incubated for 1-3 hours before the samples were separated via agarose gel electrophoresis (see chapter 3.3.1).

3.3.3 Dephosphorilation of 5' ends

In order to prevent high background caused by relegation of the vector in the cloning procedure, 5' ends of the vectors were dephosphorilated using alkaline phosphatase from calf intestine (CIP, NewEngland BioLabs, 10u/µl). Up to 20µg of DNA were dephosphorilated using NEBuffer 3 and 10 units of CIP in a 50µl reaction. After incubation at 37°C for 60 minutes samples were cleaned using the Wizard® SV Gel and PCR clean-up system (Promega) (see chapter 3.3.6.2).

3.3.4 Kinasing and annealing of oligonucleotides

Preparation of 5' phosphorylated oligonucleotides was conducted using T4 polynucleotide kinase (PNK, NewEngland BioLabs, 10u/µl). The transfer of the γ-phosphate from ATP to the 5'-OH of the oligonucleotides catalyzed by this enzyme, enables the use of complementary oligonucleotides for cassette cloning. A typical 20µl reaction contained ATP (10mM/l), PNK-Buffer, oligonucleotide 1+2 (1µg each) and 10 units T4 PNK. The reactions were carried out using a thermocycler programmed to 37°C/30min, 90°C/30sec, 37°C/5min.

3.3.5 Ligation of DNA fragments

T4 DNA ligase (NewEngland BioLabs, 400u/µl) was used to ligate DNA fragments with compatible 3' or 5' overhangs. Usually 50-100ng of vector DNA was ligated to a roughly 3-fold molar excess of insert DNA. A typical 20µl reaction contained vector and insert together with 0.5µl of T4 DNA ligase and 2µl of 10x ligase buffer. Reactions lacking either ligase or insert served

as negative controls to estimate the background caused by undigested plasmid contaminations in the vector preparation as well as religation of the vector. Ligation reactions were incubated at room temperature for 2-4 hours before they were transformed into a suitable *E. coli* strain.

3.3.5.1 Cassette cloning

Cassette inserts made of complementary oligonucleotides (see 3.3.4) were used to introduce new restriction sites. 50-100ng of vector DNA were ligated to 1µl of phosphorylation mix as described above (see 3.3.5).

3.3.6 Purification of DNA fragments

3.3.6.1 Phenol-chloroform extraction

Phenol-chloroform extraction followed by precipitation with sodium acetate and ethanol was used as standard cleaning procedure. 3µl of 0.5M EDTA (pH8.0) per 100µl DNA solution were added to the sample before vortexing with an equal amount of phenol (containing 0.1% 8-hydroxychinolin and equilibrated with 50mM Tris-HCl pH8.0). Phases were separated by shortly spinning the samples at 13.000 rpm in an Eppendorf centrifuge 5415 and the organic phase was removed. Subsequently the sample was extracted with an equal volume of chloroform. Again the phases were separated, chloroform removed and the DNA in the aqueous solution was precipitated by the addition of 1/10 volume of 3M sodium acetate and 2 volumes of ice-cold, absolute ethanol. Incubation at -80°C for 15 minutes was followed by pelletizing DNA by spinning at 14.000 rpm at 4°C for 15 minutes. The supernatant was removed, the pellet was dried on air and redissolved in an appropriate volume of H_2O MILLI Q.

3.3.6.2 Wizard® SV Gel and PCR clean-up system (Promega)

As an alternative to classical phenol-chloroform extraction the Wizard® SV Gel and PCR clean-up system (Promega) was used as described by the manufacturer. DNA was eluted from the mini anion exchange columns with an appropriate volume of H_2O MILLI Q.

Furthermore, the system was used for the purification of DNA from agarose gel slices. Bands were cut under UV-light and purified as described by the manufacturer.

3.3.7 Polymerase chain reactions (PCR)

3.3.7.1 Standard amplification of DNA fragments

Amplification of DNA fragments was achieved by standard polymerase chain reactions with Pfu polymerase. Pfu polymerase is a high fidelity polymerase with 5'→3' proofreading activity and therefore suitable for the amplification of fragments used for the cloning of new constructs. A typical reaction contained 0.5µl template DNA (~400ng/µl), 1µl of each primer (200ng/µl), 2µl dNTP's (2.5mM, GE Healthcare), 5µl 10x Pfu Buffer (Promega), 1µl Pfu polymerase (Promega, 3u/µl) and 40.5µl H_2O. Amplification was achieved on a Perkin Elmer DNA Thermal Cycler 480 using a temperature program of

95°C	2 min	Initial denaturation
95°C Annealing temperature 72°C	30 sec 1 min 1 min / 500bp	25-30 cycles
72°C	5 min	Final extension
4°C	∞	Soak

The annealing temperature was initially set to 42°C and raised in the case of appearance of side products caused by unspecific priming of the used oligonucleotides. 1µl of the reaction was separated on agarose gel to monitor the amplification before the product was typically cleaned via a spin column of the Wizard® SV Gel and PCR clean-up system (Promega) (see 3.3.6.2).

3.3.7.2 Two step PCR mutagenesis

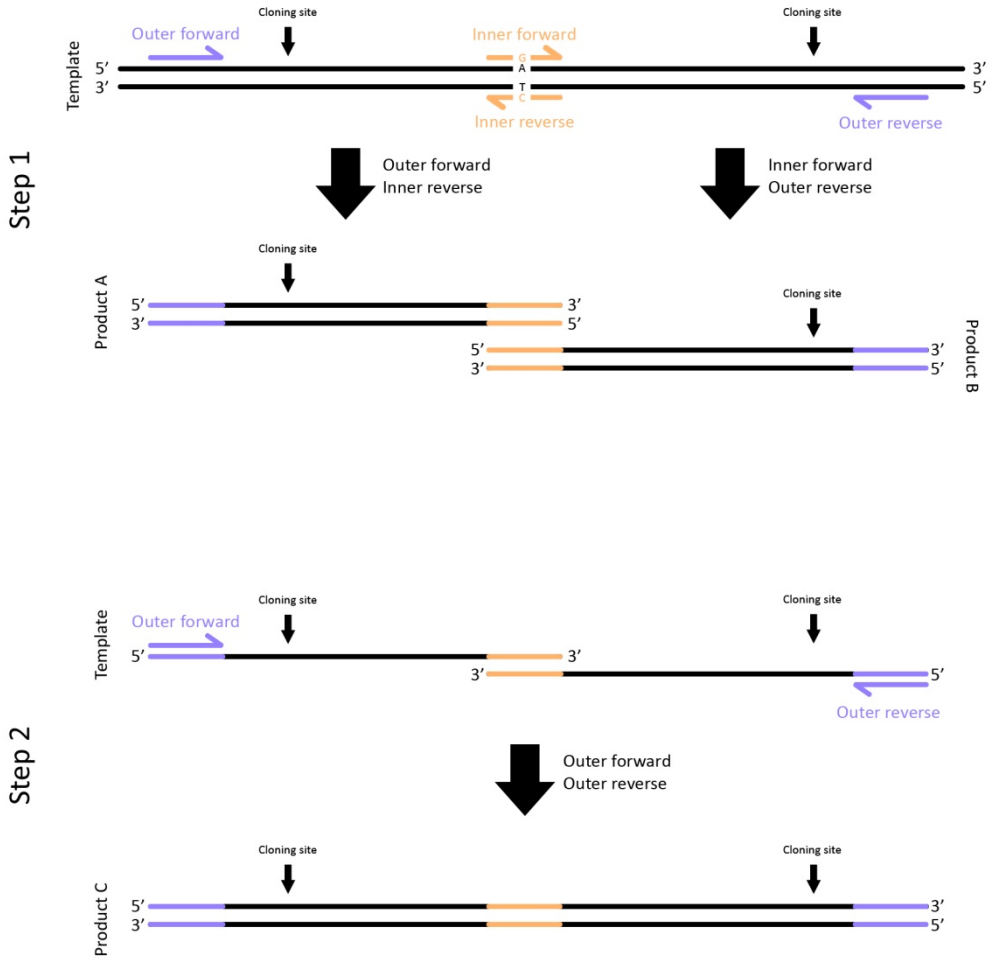


Figure 3-4 Schematic overview of a two step PCR mutagenesis procedure. In a first step, the outer forward and inner reverse primers as well as the inner forward and outer reverse primers are used to amplify two separate, overlapping products that carry the desired mutation. These two products are then used together as a template in a second polymerase chain reaction using the outer forward and outer reverse primers for the amplification of the full length product.

Two step PCR mutagenesis was used for site directed mutagenesis. Mutations were introduced via two PCR reaction steps (Figure 3-4). The method uses a set of four primers. A pair of outer primers, and a pair of inner primers that are complementary and carry a mutation. During the first step two separate reactions using the outer forward and inner reverse primer as well as the inner forward and outer reverse primer amplify two separate, overlapping products (A and B) that carry the desired mutation. The products are cleaned to remove excess primers (see

chapter 3.3.6.2). A mixture of both products is then used as a template in the second step. The overlapping template strands serve as primers for the completion of the full length product (C). Furthermore, the outer forward and outer reverse primers allow amplification of the full length product carrying the mutation.

After digesting the PCR product with suitable restriction enzymes (cloning sites), it can be ligated in a suitable vector. Alternatively, appropriate restriction sites can be introduced via 5' non-binding overlaps of the outer primers.

3.3.8 DNA sequencing

Sequences of all constructs were confirmed via DNA sequencing by LGC genomics.

3.4 WORK WITH BACTERIAL CULTURES

3.4.1 Bacterial strains

3.4.1.1 Cloning strains

E. coli TOP10 cells (Invitrogen) were used as the standard strain in all cloning procedures as well as propagation of plasmids. As an alternative, *E. coli* DH5α (Invitrogen) were used for these applications.

3.4.1.2 Expression strains

The strain *E. coli* BL21 (DE3) (Novagen) contains a chromosomal copy of the T7 RNA polymerase gene under the control of the *lacUV5* promoter. Upon induction with IPTG, T7 RNA polymerase is expressed which itself transcribes the gene of interest under T7 promoter control

that is provided on a pET vector (see chapter 3.1.1) that has been transformed into the cells before.

Leaky expression of target protein can be a problem when expressing proteins that are toxic to the *E. coli* cells. The strains *E. coli* BL21 (DE3) pLysS (Novagen) and *E. coli* BL21 (DE3) pLysE (BIOLINE) offer a solution by carrying an additional plasmid from which a small amount of T7 lysozyme (a natural inhibitor of T7 RNA polymerase) is provided. Thus, basal expression of T7 RNA polymerase before induction does not lead to expression of the target gene before induction. Upon addition of IPTG expression of T7 RNA polymerase overcomes the expression of T7 lysozyme and the target gene is over-expressed. In this respect, *E. coli* BL21 (DE3) pLysE offers even better stringency than its pLysS relative.

Many target genes exhibit codon usage that is suboptimal for *E. coli*. The strain *E. coli* Rosetta2 BL21 (DE3) pLysS (Novagen) supplies tRNAs for seven rare codons (AGA, AGG, AUA, CUA, GGA, CCC and CGG) thereby allowing strong over-expression of recombinant proteins without codon optimization.

The strain *E. coli* SoluBL21 (amsbio) claims better solubility for proteins that could not be found in the soluble fractions of *E. coli* BL21 (DE3). *E. coli* SOLUBL21 is a mutant of the parental *E. coli* BL21 (DE3) obtained by a directed evolutional approach.

3.4.2 Media and antibiotics

Luria Bertani (LB) medium (Roth) was used as the standard medium for growing *E. coli* cells and was composed of 10g/l tryptone, 5g/l yeast extract, 10g/l NaCl, pH 7.0 ± 0.2. All media were autoclaved before being used for growing bacterial cultures.

LB-Gluc medium contains all components of LB medium plus an extra of 2g/l of glucose.

M9 minimal medium was composed of 33.7mM $Na_2HPO_4*2H_2O$, 22mM KH_2PO_4, 8.5mM NaCl, 18.7mM NH_4Cl, 2mM $MgSO_4$, 1% trace elements, 20% glucose, 0.3mM $CaCl_2$, 1µg/ml biotine and 1µg/ml thiamine. A solution of $MgSO_4$, trace elements, glucose, $CaCl_2$, biotine and thiamine was prepared separately and submitted to sterile filtration rather than autoclaving before being added to an autoclaved solution of the former components.

LB agar plates were prepared by adding 1.5% Agar-Agar to LB medium followed by autoclaving at 120°C for 20 minutes. The solution was allowed to cool down before the appropriate antibiotic was added and plates were poured using 10cm petri dishes.

Ampicillin was used at an end concentration of 100µg/L. Stocks were produced by dissolving 100mg/L ampicillin in H_2O MILLI Q followed by sterile filtration and storage at -20°C.

Kanamycin was used at an end concentration of 50µg/L. Stocks were produced by dissolving 50mg/L kanamycin in H_2O MILLI Q followed by sterile filtration and storage at -20°C.

3.4.3 Preparation of competent cells

Chemically competent cells were used to be transformed with plasmids for either propagation of plasmids, cloning of new constructs or expression of recombinant proteins. 6ml of 2xTY medium were inoculated with a strain of *E. coli* cells and grown overnight at 37°C. In the morning the overnight culture was used to start a 200ml culture in 2xTY medium and grown to an OD_{600} of 0.45-0.55 at 37°C/130rpm. Cells were harvested in four 50ml tubes at 1500rpm/4°C/15min. The supernatant was discarded, cells were resuspended in 66ml of ice cold RF1 solution and incubated on ice for 15 minutes. Cells were again pelletized as described above and resuspended in 16ml of ice cold RF2 solution. Subsequently, aliquots of 200µl were frozen in liquid nitrogen and stored at -80°C.

2xTY medium	16g/l Tryptone, 10g/l Yeast extract, 5g/l NaCl. pH adjusted to 7.0 with NaOH.
RF1 solution	100mM KCl, 50mM $MnCl_2$*$4H_2O$, 30mM CH_3CO_2K, 10mM $CaCl_2$+$2H_2O$, 15% v/v glycerol. pH adjusted to 5.8 with HCl.
RF2 solution	10mM MOPS, 10mM KCl, 75mM $CaCl_2$*$2H_2O$, 15% v/v glycerol. pH adjusted to 6.8 with KOH.

3.4.4 Transformation

A 200µl aliquot of an appropriate cloning strain was thawed on ice before 10µl of ligation mix (see chapter 3.3.5) or 0.1µl of a MIDI plasmid prep (see chapter 3.4.6) was added. For expression strains, an aliquot of 5µl vector was added to the thawed cells. Mixtures were incubated on ice for 10 minutes before cells were heat shocked at 42°C for 45 seconds. Immediately, 400µl of preheated (37°C) LB-medium were added and cells were incubated at 37°C for 30 minutes. 150µl of culture was then plated on LB agar plates including the appropriate antibiotic and incubated at 37°C overnight.

3.4.5 DNA Mini preparations

This small scale standard procedure was used to prepare plasmid DNA from *E. coli* cultures in order to screen for positive clones in the cloning procedure of new constructs. Single colonies were picked from LB agar plates and used to inoculate 4ml of LB medium including the appropriate antibiotic. Cultures were grown at 37°/160rpm overnight. In the morning cells were pelletized by spinning 1.5ml of culture in Eppendorf tubes at 6.000rpm for 2 minutes. The supernatant was discarded and cells were collected from another 1.5ml of culture in the same tube. Subsequently, cells were resuspended in 100µl solution I and cell lysis was initiated by the addition of 200µl solution II. Next, 150µl of solution III were added followed by incubation on ice for 5 minutes. Precipitated cell debris was pelletized by spinning at 14.000rpm/4°C/10min. Plasmid containing supernatant was transferred into a fresh tube and DNA was precipitated by the addition of 1ml of ice cold absolute ethanol and spinning at 14.000rpm/4°C/5min. The pellet was resuspended in 200µl TE buffer before 200µl of 5M LiCl were added and insoluble components were removed by spinning at 14.000rpm/4°C/5min. The supernatant was again transferred into a fresh tube and DNA was again precipitated by the addition of 1ml ice cold absolute ethanol followed by a spin at 14.000rpm/4°C/5min. The pellet was dried on air, redissolved in 50µl H_2O MILLI Q and stored at -20°C.

Solution I	50mM Tris-HCl pH 8.0, 10mM EDTA, 100µg/ml RNaseA.
Solution II	20mM NaOH, 1% SDS.
Solution III	2.8M Potassiumacetate pH 5.1.
TE buffer	10mM Tris-HCl pH 8.0, 1mM EDTA.

3.4.6 DNA Midi preparations

The plasmid purification kit NucleoBond® PC 100 (Macherey-Nagel) was used to prepare plasmid DNA of great purity in the 100µg scale. The kit was used following the instructions of the manufacturer.

3.5 PROTEIN METHODS

3.5.1 SDS-PAGE

The discontinuous, tris-glycine buffered gel system of Laemmli was used to monitor protein expression and purification throughout the work (Laemmli, 1970). Proteins are bundled in a 5% stacking gel with pH 6.8, before entering a separating gel with pH 8.8. Separating gels were used with 6%, 12.5%, 17.5% or 20% polyacrylamide depending on the size of the protein of interest (see Table 3-2 for the composition of gels). The BioRad MiniProtean® electrophoresis system was used to cast gels (0.75mm spacers). Samples were mixed with Laemmli sample buffer, heated at 95°C for 5 minutes and separated on the SDS-PAGE using Laemmli running buffer at a constant voltage of 150-200V. BioRad Precision Plus Protein™ Standards as well as PageRuler™ Plus Prestained Protein Ladder from Fermentas were used as molecular weight markers (see Figure 3-5).

	Stacking gel	Separating gel			
	5%	6%	12.5%	17.5%	20%
30% Acrylamide	0.333ml	1.200ml	2.500ml	3.500ml	4.000ml
2.5% Bisacrylamide	0.104ml	0.209ml	0.209ml	0.209ml	0.209ml
4x Stacking/Separating gel buffer	0.500ml	1.500ml	1.500ml	1.500ml	1.500ml
H_2O	1.060ml	3.060ml	1.758ml	0.758ml	0.261ml
10% Ammoniumpersulfate	0.020ml	0.050ml	0.050ml	0.050ml	0.050ml
TEMED	0.002ml	0.005ml	0.005ml	0.005ml	0.005ml

Table 3-2 Composition of SDS-polyacrylamide gels.

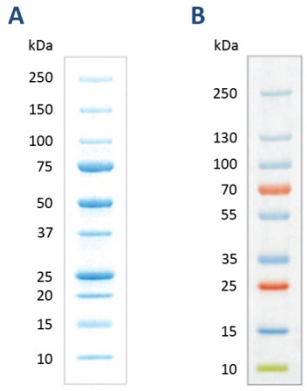


Figure 3-5 Protein molecular weight markers used in SDS-PAGE. (A) BioRad Precision Plus Protein™ Standards all blue separated on a 4-20% Tris-HCl SDS gel. **(B)** Fermentas PageRuler™ Plus Prestained Protein Ladder separated on a 4-20% Tris-glycine SDS gel. Figures taken from manufacturer's product insert.

4x Stacking gel buffer	1.5M Tris-HCl pH 8.8, 0.4% SDS
4x Separating gel buffer	0.5M Tris-HCl pH 6.8, 0.4% SDS

10x Laemmli running buffer	500mM Tris, 3.85M glycine, 1% SDS
2x Laemmli sample buffer	20% v/v glycerol, 10% β-mercaptoethanol, 6% SDS, 125mM Tris-HCl pH 6.8, 0.01% bromphenol blue

3.5.1.1 Coomassie staining of SDS-polyacrylamide gels

Coomassie staining was used as a standard procedure to visualize protein on SDS-PAGE. Gels were soaked in stain for 10-30 minutes and destained with several washes of H_2O. Subsequently, gels were dried on Whatman® paper using a SAVANT slab gel drier.

Coomassie stain	45% Methanol, 10% acetic acid, 0.4% Coomassie brilliant blue R250

3.5.2 Western blotting

3.5.2.1 Transfer of protein onto PVDF membrane

Proteins were transferred from SDS-PAGE to Immobilon®-P Transfer membrane (Millipore) to analyze the samples with specific antibodies. A blotting sandwich in the tank blot system typically consisted of 2 sponges and 3 Whatman® papers soaked in transfer buffer, followed by the polyacrylamide gel and the PVDF membrane that has been activated in 100% methanol topped with 3 more Whatman® papers and 2 sponges (again soaked in transfer buffer). Air bubbles were carefully removed when assembling the sandwich. Transfer was achieved by applying a constant current of 40mA overnight at 4°C.

Transfer buffer	25mM Tris, 192mM glycine, 20% methanol.

3.5.2.2 Blocking the membrane

Unspecific binding of antibodies was prevented by blocking the membrane with Tropix® I-block™ (Applied Biosystems) solution. The membrane was transferred into a 50ml Falcon tube together with 5 ml blocking solution and incubated at room temperature for 30 minutes under

constant rotation. Blocking solution was discarded and the procedure was repeated with another 5ml of blocking buffer.

Blocking solution	0.2% Tween 20, 0.2% Tropix® I-block™ (Applied Biosystems) in PBS.
PBS	1.4mM KH_2PO_4, 2.7mM KCl, 4.3mM Na_2HPO_4, 137mM NaCl.

3.5.2.3 Detection of His-tagged proteins

HisProbe™-HRP (Pierce) was used to detect His-tagged proteins. 1µl of stock solution (4mg/ml) was diluted in 5ml of blocking solution and incubated with the membrane at 4°C for 60 minutes under constant rotation. The solution was discarded and the membrane was washed three times 10 minutes with PBST and one time 10 minutes with PBS. Subsequently, the membrane was soaked for 5 min in 2ml of a 1:1 mixture of the two solutions of the SuperSignal™ West Pico kit (Pierce). Finally, the membrane was wrapped in food foil and exposed to CL-X Posure™ film (Thermo Scientific).

PBST	1.4mM KH_2PO_4, 2.7mM KCl, 4.3mM Na_2HPO_4, 137mM NaCl, 0.1% Tween 20.

3.5.2.4 Detection of eIF4G I

Detection of eIF4G I utilized an antiserum from rabbit that recognizes the N-terminal domain of eIF4G I (α-eIF4G I, kind gift from R. Rhoads, Shreveport, LA). All reactions were performed at room temperature. The blocked membrane was washed with PBS for 10 minutes twice. The primary antibody (α-eIF4G I, 1:8000 in PBST) was incubated with the membrane for 60 minutes under constant rotation followed by washing with PBS for 10 minutes twice. Next, an alkaline phosphatase coupled α-rabbit antibody (Sigma) was used as the secondary antibody (1:5000 in PBST). Incubation for 60 minutes was followed by a final wash of the membrane with PBS (two times 10 minutes). In order to visualize bands the membrane was soaked in 5ml alkaline phosphatase buffer and 25µl NBT solution as well as 25µl BCIP solution were added. The color reaction was conducted until bands were clearly visible before the reaction was finally stopped by soaking the membrane in H_2O.

NBT solution	5% (w/v) nitro blue tetrazolium in 90% dimethylformamide.

BCIP solution	2.5% (w/v) 5-bromo-4-chloro-3-indolyl phosphate disodium salt in dH_2O.
Alkaline phosphatase buffer	5mM $MgCl_2$, 100mM NaCl, 100mM Tris-HCl pH 9.6.

3.6 EXPRESSION AND PURIFICATION OF RECOMBINANT PROTEINS

All purification steps were performed at 4°C unless stated differently.

3.6.1 Expression in *E. coli*

After transforming the *E. coli* expression strain of choice (see chapter 3.4.1.2) with the expression vector coding for the protein of interest, single colonies were picked to inoculate 100ml of LB medium including the appropriate antibiotics. The culture was incubated at 37°C/130rpm overnight and used to inoculate bigger cultures (1.0-1.5 liters per 5L flask) to a starting OD_{600}~0.10.

E. coli cultures grown in M9 minimal medium were initially started by inoculating 4ml of LB medium. This culture was grown at 37°/130rpm overnight and was used to inoculate a 100ml culture of M9 minimal medium including appropriate antibiotics. Again the culture was grown to stationary phase overnight at 37°C/130rpm and then used to inoculate bigger cultures as described above.

Cells were typically grown at 37°C/130rpm till an OD_{600}~0.50 was reached and expression of recombinant protein could be induced by the addition of IPTG and incubation at 110rpm for a suitable time period. If possible, a 20ml aliquot was removed before induction with IPTG and used as an uninduced control, by treating it in an identical way as the induced sample except for IPTG addition. In order to increase the amount of recombinant protein found in the soluble fraction, protein expression was often performed at 16°C. For this purpose, cells were chilled on ice for 20 minutes before addition of IPTG. Alternatively, temperature was reduced to 16°C when an OD_{600}~0.30 was reached. Cells were further grown to an OD_{600}~0.50 and induced by the addition of IPTG.

As 2A protease tightly binds a Zn^{2+}-ion in its C-terminal domain which has been reported to be essential for the structural integrity of the protease, we started to additionally add $ZnCl_2$ (100µM end concentration) at the time of induction in order to prevent Zn^{2+} depletion of the cells.

3.6.2 Cell disruption

Cells were harvested using a Beckmann Coulter J-26 XP centrifuge equipped with a JLA-8.1000 rotor. Cultures were spun 4.000rpm/4°C/20min and the resulting pellet was usually resuspended in 30ml (per liter of culture) of lysis buffer (1ml of lysis buffer was used to resuspend the pellet of the uninduced control if available). Samples that were not immediately processed further were frozen at -80°C and thawed on room temperature when needed. Cells were broken by either sonication or the use of a homogenizer working with a French press principle. Two aliquots of 50µl each (used for SDS-PAGE analysis) were removed before the crude cell extract was cleared by centrifugation in a SS34 rotor (20.000rpm/4°C/30min). One of the aliquots was cleared by spinning in an Eppendorf 5417R centrifuge (14.000rpm/4°C/30min). The supernatant was transferred into a fresh tube and the pellet was redissolved in 50µl lysis buffer. The total, soluble and insoluble fractions were then mixed with 50µl 2x Laemmli sample buffer each and cooked at 95°C for 15 minutes. Finally, sample was separated on SDS-PAGE and stained with Coomassie Blue to assess the degree of over-expression and the ratio of recombinant protein found in the soluble and insoluble fractions.

3.6.2.1 Sonication

Resuspended cells were placed in a Sonoplus RZ3 rosett cell packed in ice and sonicated via a Bandelin Sonoplus HD200 ultrasonic wave generator equipped with a SH213 booster horn and MS73 tip. The series of boosts contained three times 30sec on Cycle40 with 60sec breaks in between. Disruption was finalized by a 15sec continuous boost. The uninduced control (if available) was processed by a series of three boosts of 30sec at Cycle20.

3.6.2.2 Disruption of cells via homogenizer

As an alternative to sonication cells were broken with the Emulsiflex-C3 homogenizer (Avestin). Cells were passed through the outlet valve under a pressure of 1000-1500 bar for three to five times till a homogenous, non viscous sample was obtained. As the device requires a certain

minimal sample volume to be operated, the cells of the uninduced control could not be disrupted with this method.

3.6.3 Fractionated $(NH_4)_2SO_4$ precipitation

This classic purification strategy was used early in the purification process and is based on the diverse solubility of proteins under high ionic strength. Stepwise addition of a saturated aqueous $(NH_4)_2SO_4$ solution at 4°C under stirring was used to salt out proteins in two steps. First, the protein sample was incubated at a lower concentration of $(NH_4)_2SO_4$ at 4°C for at least 12 hours leaving the protein of interest in solution. Precipitated proteins were removed by pelletizing them in a SS34 rotor (20.000rpm/4°C/30min). Subsequently, the supernatant was mixed with further $(NH_4)_2SO_4$ and incubated for at least 4 hours. Typically, the concentration was chosen to precipitate the protein of interest in this second fraction. The precipitated protein was again collected as described above and resuspended in a buffer of choice. Routinely, these samples were dialyzed against the chosen buffer to remove residual $(NH_4)_2SO_4$ that could interfere with subsequent purification steps.

3.6.4 Use of ÄKTA FPLC in protein purification

An ÄKTA FPLC system (GE Healthcare) composed of a P-920 pump unit, a UPC-900 detection unit, several valves and a Frac-950 was used to drive protein purification via anion exchange chromatography, His-trap affinity chromatography, GST-trap affinity chromatography, desalting and size exclusion chromatography. Systems were initially available at room temperature and on 4°C in the later experiments of this work. All solutions as well as samples were filtered through a 0.22µm filter in order to prevent system filters as well as on column filters from clogging.

3.6.5 Anion exchange chromatography

A MonoQ 10/100 GL column (Pharmacia) was used to separate proteins according to their surface charge. The pH of the mobile phase was chosen to allow binding of the protein of interest to the resin. For this purpose the theoretical pI of the recombinant protein was calculated by the online application provided by ExPASy (http://web.expasy.org/compute_pi/pi_tool-ref.html) (Bjellqvist *et al.*, 1994; Bjellqvist *et al.*, 1993) and a pH that was at least 2 units higher than the obtained value was chosen for the mobile phase in order to allow proper binding to the matrix even in the presence of higher salt concentrations in the binding buffer. The column was equilibrated with at least 3 column volumes of binding buffer and sample was typically loaded via a 50ml superloop. Flow-through was collected and unbound sample was washed from the column with binding buffer before bound protein was eluted via a NaCl gradient (up to 1M NaCl) applied by the pumps of the ÄKTA FPLC system. The eluent was collected in tubes of 2ml each. The column was washed with 5 column volumes of buffer containing 1M NaCl to elute tightly bound contaminants and finally washed with H_2O MILLI Q as well as 20% ethanol and stored at 4°C. Aliquots of fractions of interest were separated on SDS-PAGE and stained with Coomassie Blue for analysis.

3.6.6 Affinity chromatography

3.6.6.1 Purification via His-trap

His-Trap HP (Pharmacia) columns were used for high resolution purification of His-tagged proteins. Columns were operated using an ÄKTA FPLC. If stripped, columns were charged by loading 2 column volumes 300mM $NiCl_2$, followed by washing with H_2O to remove excess $NiCl_2$. Against the manufacturer's product information, the use of 5mM DTT or 20mM β-mercaptoethanol in buffers or sample was not possible in our hands. Hence, buffers and sample had to be free of reducing agents for purification with His-trap columns. Columns were equilibrated with 5 column volumes binding buffer (including up to 30mM imidazole). Sample was loaded onto the column with a flow rate of 0.5-1.0ml/min and flow-through was collected. Unbound protein was removed from the column by washing with binding buffer till the UV signal

reproached the baseline (5-10 column volumes). Bound protein was eluted from the column by applying a linear imidazole gradient up to 500mM imidazole. The eluent was typically collected in fractions á 1.0-2.0ml. The column was washed with 10 column volumes H_2O followed by 10 column volumes 20% ethanol and stored at 4°C. Fractions of interest were inspected by SDS-PAGE.

Stripping of the column was achieved by passing 5-10 column volumes of binding buffer including 50mM EDTA over the beads. Subsequently, the matrix was washed with water and stored in 20% ethanol.

3.6.6.2 Purification via amylose resin

Maltose binding protein (MBP) served not only as a solubility tag but also as a purification tag. Typically, 10ml of amylose resin (New England BioLabs) were packed into Econo-Pac gravity flow columns (BioRad). Beads were equilibrated with 5 column volumes of binding buffer before sample was loaded and flow-through was collected. Unbound protein was washed off the column with binding buffer (5 column volumes) and protein was eluted using a step gradient of maltose concentration (binding buffer +10mM maltose). The eluent was collected in fractions of 10ml each which were inspected by SDS-PAGE and Coomassie staining.

Resin was regenerated by washing with H_2O (5 column volumes) followed by stripping with 3 column volumes 0.1% SDS at room temperature (to avoid precipitation of SDS). SDS was removed by extensive washing with H_2O and the resin was stored in 20% ethanol.

3.6.6.3 Purification via GSTrap

A 5ml GSTrap HP (Pharmacia) was used to purify proteins tagged with glutathione S-transferase (GST) on an ÄKTA FPLC system. The column was equilibrated with 10 column volumes of binding buffer before sample was loaded at a flow rate of 1ml/min and flow-through was collected. Unbound protein was removed by washing with 10 column volumes of binding buffer before elution of bound proteins was induced by a step gradient of reduced glutathione (binding buffer +10mM reduced glutathione) at a flow rate of 5ml/min. The eluent was collected in

fractions of 2ml each and screened for protein of interest by SDS-PAGE and Coomassie staining. The column was washed with 10 column volumes H_2O and stored in 20% ethanol at 4°C.

3.6.7 Size exclusion chromatography

3.6.7.1 Purification via Superdex

HiLoad Superdex 75 prep grade 26/60 as well as HiLoad Superdex 200 prep grade 26/60 columns (Pharmacia) were used with an ÄKTA FPLC system for gel filtration as the last step in the purification procedure. Columns were equilibrated with 2 column volumes of running buffer. Sample was concentrated to a maximal volume of 5ml (typically 1-2ml) before being loaded onto the matrix. Isocratic elution was performed with the equilibration buffer at a typical flow rate of 2.0-2.5ml/min and eluent was collected in fractions of 2ml each. Fractions of interest were screened via SDS-PAGE and Coomassie staining. Columns were washed with 2 column volumes of H_2O and stored in 20% ethanol.

3.6.7.2 Desalting and buffer exchange

Fast and efficient desalting and buffer exchange was executed via a HighPrep Desalting 26/10 column (GE Healthcare). The column was equilibrated with the new buffer of choice, concentrated sample (<15ml) was loaded and the column was developed with the new buffer of choice at flow rates up to 30ml/min. The column was washed with H_2O and stored in 20% ethanol.

3.6.8 Other protein methods

3.6.8.1 Dialysis

SnakeSkin® dialysis tubing (Thermo Scientfic) with 10kDa or 3.5kDa cutoff was used for dialyzing protein samples. Typically the sample was dialyzed against three changes of dialysis buffer at 4°C under constant stirring for at least 4 hours per change.

3.6.8.2 Concentrating protein samples

Centriprep® centrifugal filter units (Millipore) with a cutoff of 10kDa were mainly used to concentrate protein samples. Units including protein sample were spun at a maximum of 3.000g at 4°C in a swinging bucket rotor till the protein concentration of choice was reached. Centriprep® centrifugal filter units exhibit an internal hold-up volume of 0.5ml at which no more concentration can be obtained. For further concentration Amicon® Ultra-4 centrifugal filters (Millipore) with a cutoff of 10kDa were used at a maximum of 4.000g at 4°C in a swinging bucket rotor.

3.6.8.3 Assessing protein concentration

Protein concentration was assessed using a NanoDrop ND-1000 spectrophotometer by measuring absorbance at 280nm. Molar extinction coefficients were estimated using the ProtParam web server (provided by ExPASy) (Wilkins *et al.*, 1999). The Server uses the equation by Gill *et al.* to predict the molar extinction coefficient (Gill & von Hippel, 1989). Having detected these values, the concentration was calculated using Lambert Beer's law.

3.6.9 Reductive lysine methylation

Reductive lysine methylation was performed following the protocol of Walter *et al.* (Walter *et al.*, 2006). Protein was re-buffered to 50mM HEPES pH 7.5, 250mM NaCl using a

HighPrep Desalting 26/10 column (see chapter 3.6.7.2). Protein was diluted to 0.5mg/ml and methylation was performed in a closed hood at 4°C. 20µl 1M dimethyl-borane and 40µl 1M formaldehyde per ml of protein solution was added, the reaction vessel was carefully closed with parafilm and incubated under stirring at 4°C for 2 hours. Addition of 1M dimethyl-borane and 40µl 1M formaldehyde per ml of protein solution was repeated and the mixture was again incubated for 2 hours under stirring at 4°C. Finally, another 10µl of 1M dimethyl-borane complex per ml of protein solution were added and the reactions was finalized by incubation overnight. Protein was re-purified via size exclusion chromatography (see chapter 3.6.7.1) using the buffer of choice.

1M dimethyl-borane complex	Freshly dissolve in H_2O MILLI Q directly before usage.

3.6.10 Thermofluor

Thermofluor allows fast screening of buffer conditions for maximal stability of proteins. The method is based on the binding of a fluorophore (Sypro Orange) to the hydrophobic patches of proteins that will be exposed upon stepwise heating. Binding to those hydrophobic patches is leading to a sigmoid increase of fluorescence that is otherwise quenched by the aqueous solution. Upon total unfolding of the protein fluorescence will again decline. The inflection point of the sigmoid rise is defined as the melting point of the protein. Buffer, pH, ionic strength and different additives are mixed with the protein to detect the buffer that stabilizes the protein of interest best.

Preliminary experiments without screening of different conditions were conducted to identify the optimal ratio of protein and dye. In a 5µl reaction all combinations of protein concentrations (0.20mg/ml, 0.10mg/ml and 0.05mg/ml) and dye concentrations (100x, 10x and 1x (from a 5000x stock)) were tested. The combination with the strongest signal was chosen and 2x96 different conditions were screened by mixing 4µl of screening solution with 1µl of a mix made of protein and dye. Fluorescence was measured using a qPCR system equipped with a 492nm extinction and 516nm emission filter. The temperature gradient was established in steps of 1°C from 20°C to 95°C. Screened conditions are listed in Table 3-3 and Table 3-4.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Sodium acetate pH 4.5	Sodium citrate pH 4.7	Sodium acetate pH 5.0	Potassium phosphate pH 5.0	Sodium phosphate pH 5.5	Sodium citrate pH 5.5	MES pH 5.8	Potassium phosphate pH 6.0	MES pH 6.2	Sodium phosphate pH 6.5	Sodium cacodylate pH 6.5	MES pH 6.5
B	Sodium acetate pH 4.5 100mM NaCl	Sodium citrate pH 4.7 100mM NaCl	Sodium acetate pH 5.0 100mM NaCl	Potassium phosphate pH 5.0 100mM NaCl	Sodium phosphate pH 5.5 100mM NaCl	Sodium citrate pH 5.5 100mM NaCl	MES pH 5.8 100mM NaCl	Potassium phosphate pH 6.0 100mM NaCl	MES pH 6.2 100mM NaCl	Sodium phosphate pH 6.5 100mM NaCl	Sodium cacodylate pH 6.5 100mM NaCl	MES pH 6.5 100mM NaCl
C	Sodium acetate pH 4.5 500mM NaCl	Sodium citrate pH 4.7 500mM NaCl	Sodium acetate pH 5.0 500mM NaCl	Potassium phosphate pH 5.0 500mM NaCl	Sodium phosphate pH 5.5 500mM NaCl	Sodium citrate pH 5.5 500mM NaCl	MES pH 5.8 500mM NaCl	Potassium phosphate pH 6.0 500mM NaCl	MES pH 6.2 500mM NaCl	Sodium phosphate pH 6.5 500mM NaCl	Sodium cacodylate pH 6.5 500mM NaCl	MES pH 6.5 500mM NaCl
D	Sodium acetate pH 4.5 100mM KCl	Sodium citrate pH 4.7 100mM KCl	Sodium acetate pH 5.0 100mM KCl	Potassium phosphate pH 5.0 100mM KCl	Sodium phosphate pH 5.5 100mM KCl	Sodium citrate pH 5.5 100mM KCl	MES pH 5.8 100mM KCl	Potassium phosphate pH 6.0 100mM KCl	MES pH 6.2 100mM KCl	Sodium phosphate pH 6.5 100mM KCl	Sodium cacodylate pH 6.5 100mM KCl	MES pH 6.5 100mM KCl
E	Sodium acetate pH 4.5 500mM KCl	Sodium citrate pH 4.7 500mM KCl	Sodium acetate pH 5.0 500mM KCl	Potassium phosphate pH 5.0 500mM KCl	Sodium phosphate pH 5.5 500mM KCl	Sodium citrate pH 5.5 500mM KCl	MES pH 5.8 500mM KCl	Potassium phosphate pH 6.0 500mM KCl	MES pH 6.2 500mM KCl	Sodium phosphate pH 6.5 500mM KCl	Sodium cacodylate pH 6.5 500mM KCl	MES pH 6.5 500mM KCl
F	Sodium acetate pH 4.5 100mM $(NH_4)_2SO_4$	Sodium citrate pH 4.7 100mM $(NH_4)_2SO_4$	Sodium acetate pH 5.0 100mM $(NH_4)_2SO_4$	Potassium phosphate pH 5.0 100mM $(NH_4)_2SO_4$	Sodium phosphate pH 5.5 100mM $(NH_4)_2SO_4$	Sodium citrate pH 5.5 100mM $(NH_4)_2SO_4$	MES pH 5.8 100mM $(NH_4)_2SO_4$	Potassium phosphate pH 6.0 100mM $(NH_4)_2SO_4$	MES pH 6.2 100mM $(NH_4)_2SO_4$	Sodium phosphate pH 6.5 100mM $(NH_4)_2SO_4$	Sodium cacodylate pH 6.5 100mM $(NH_4)_2SO_4$	MES pH 6.5 100mM $(NH_4)_2SO_4$
G	Sodium acetate pH 4.5 10% glycerol	Sodium citrate pH 4.7 10% glycerol	Sodium acetate pH 5.0 10% glycerol	Potassium phosphate pH 5.0 10% glycerol	Sodium phosphate pH 5.5 10% glycerol	Sodium citrate pH 5.5 10% glycerol	MES pH 5.8 10% glycerol	Potassium phosphate pH 6.0 10% glycerol	MES pH 6.2 10% glycerol	Sodium phosphate pH 6.5 10% glycerol	Sodium cacodylate pH 6.5 10% glycerol	MES pH 6.5 10% glycerol
H	Sodium acetate pH 4.5 20% glycerol	Sodium citrate pH 4.7 20% glycerol	Sodium acetate pH 5.0 20% glycerol	Potassium phosphate pH 5.0 20% glycerol	Sodium phosphate pH 5.5 20% glycerol	Sodium citrate pH 5.5 20% glycerol	MES pH 5.8 20% glycerol	Potassium phosphate pH 6.0 20% glycerol	MES pH 6.2 20% glycerol	Sodium phosphate pH 6.5 20% glycerol	Sodium cacodylate pH 6.5 20% glycerol	MES pH 6.5 20% glycerol

Table 3-3 **Thermofluor ionic strength screen I.** Composition of the 96 conditions of the ionic strength screen (plate I) used for thermofluor measurements.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Potassium phosphate pH 7.0	HEPES pH 7.0	Ammonium acetate pH 7.3	Sodium phosphate pH 7.5	Tris pH 7.5	Imidazole pH 8.0	HEPES pH 8.0	Tris pH 8.0	Bicine pH 8.0	Tris pH 8.5	Bicine pH 9.0	H_2O
B	Potassium phosphate pH 7.0 100mM NaCl	HEPES pH 7.0 100mM NaCl	Ammonium acetate pH 7.3 100mM NaCl	Sodium phosphate pH 7.5 100mM NaCl	Tris pH 7.5 100mM NaCl	Imidazole pH 8.0 100mM NaCl	HEPES pH 8.0 100mM NaCl	Tris pH 8.0 100mM NaCl	Bicine pH 8.0 100mM NaCl	Tris pH 8.5 100mM NaCl	Bicine pH 9.0 100mM NaCl	H_2O 100mM NaCl
C	Potassium phosphate pH 7.0 500mM NaCl	HEPES pH 7.0 500mM NaCl	Ammonium acetate pH 7.3 500mM NaCl	Sodium phosphate pH 7.5 500mM NaCl	Tris pH 7.5 500mM NaCl	Imidazole pH 8.0 500mM NaCl	HEPES pH 8.0 500mM NaCl	Tris pH 8.0 500mM NaCl	Bicine pH 8.0 500mM NaCl	Tris pH 8.5 500mM NaCl	Bicine pH 9.0 500mM NaCl	H_2O 500mM NaCl
D	Potassium phosphate pH 7.0 100mM KCl	HEPES pH 7.0 100mM KCl	Ammonium acetate pH 7.3 100mM KCl	Sodium phosphate pH 7.5 100mM KCl	Tris pH 7.5 100mM KCl	Imidazole pH 8.0 100mM KCl	HEPES pH 8.0 100mM KCl	Tris pH 8.0 100mM KCl	Bicine pH 8.0 100mM KCl	Tris pH 8.5 100mM KCl	Bicine pH 9.0 100mM KCl	H_2O 100mM KCl
E	Potassium phosphate pH 7.0 500mM KCl	HEPES pH 7.0 500mM KCl	Ammonium acetate pH 7.3 500mM KCl	Sodium phosphate pH 7.5 500mM KCl	Tris pH 7.5 500mM KCl	Imidazole pH 8.0 500mM KCl	HEPES pH 8.0 500mM KCl	Tris pH 8.0 500mM KCl	Bicine pH 8.0 500mM KCl	Tris pH 8.5 500mM KCl	Bicine pH 9.0 500mM KCl	H_2O 500mM KCl
F	Potassium phosphate pH 7.0 100mM $(NH_4)_2SO_4$	HEPES pH 7.0 100mM $(NH_4)_2SO_4$	Ammonium acetate pH 7.3 100mM $(NH_4)_2SO_4$	Sodium phosphate pH 7.5 100mM $(NH_4)_2SO_4$	Tris pH 7.5 100mM $(NH_4)_2SO_4$	Imidazole pH 8.0 100mM $(NH_4)_2SO_4$	HEPES pH 8.0 100mM $(NH_4)_2SO_4$	Tris pH 8.0 100mM $(NH_4)_2SO_4$	Bicine pH 8.0 100mM $(NH_4)_2SO_4$	Tris pH 8.5 100mM $(NH_4)_2SO_4$	Bicine pH 9.0 100mM $(NH_4)_2SO_4$	H_2O 100mM $(NH_4)_2SO_4$
G	Potassium phosphate pH 7.0 10% glycerol	HEPES pH 7.0 10% glycerol	Ammonium acetate pH 7.3 10% glycerol	Sodium phosphate pH 7.5 10% glycerol	Tris pH 7.5 10% glycerol	Imidazole pH 8.0 10% glycerol	HEPES pH 8.0 10% glycerol	Tris pH 8.0 10% glycerol	Bicine pH 8.0 10% glycerol	Tris pH 8.5 10% glycerol	Bicine pH 9.0 10% glycerol	H_2O 10% glycerol
H	Potassium phosphate pH 7.0 20% glycerol	HEPES pH 7.0 20% glycerol	Ammonium acetate pH 7.3 20% glycerol	Sodium phosphate pH 7.5 20% glycerol	Tris pH 7.5 20% glycerol	Imidazole pH 8.0 20% glycerol	HEPES pH 8.0 20% glycerol	Tris pH 8.0 20% glycerol	Bicine pH 8.0 20% glycerol	Tris pH 8.5 20% glycerol	Bicine pH 9.0 20% glycerol	H_2O 20% glycerol

Table 3-4 **Thermofluor ionic strength screen II.** Composition of the 96 conditions of the ionic strength screen (plate II) used for thermofluor measurements.

3.7 CRYSTALLIZATION OF PROTEINS

3.7.1 Automated setup of screens

Large scale screens were set up using a nanoliter crystallization robot (Phoenix). MRC 2-well plates (Swissci) were used to equilibrate two sitting drops of different composition against 50µl of reservoir solution. Drops consisted of 100nl protein and 100nl reservoir solution or 200nl protein and 100nl reservoir solution respectively. Drops were set at 22°C, plates were sealed with CrystalClear tape and incubated at either 4°C or 22°C

3.7.2 Manual setup of drops

The hanging drop vapor diffusion setup was chosen to screen around crystal hit conditions and to grow crystals at established crystallization conditions. When using 48-well plates, 2-3 drops of 2-3µl each were equilibrated against 100µl reservoir solution per well. In contrast, up to 7 drops of 3-5µl each per well were used in 24-well plates and equilibrated against 500µl of reservoir solution. Plates were stored at 4°C or 22°C and monitored either on an automated Rigaku imaging system or manually under the microscope.

3.7.3 Monitoring of crystallization screens

An automated imaging system (Rigaku CrystalTrak) assisted in easy online inspection of drops by imaging drops at time points 0h, 12h, 1d, 2d, 3d, 7d, 12d, 19d and 31d. Subsequently, plates were inspected manually under the microscope once a month over a period of 6-12 month.

3.7.4 Harvesting and soaking of crystals

Crystals were carefully fished from the drops with a loop under the microscope at 4°C and washed in a fresh drop of reservoir solution. Subsequently crystals were transferred into a drop of reservoir solution containing the inhibitor concentration of interest and soaked for an appropriate time. Finally, crystals were transferred into cryo solution (reservoir solution including 21% v/v glycerol) followed by freezing in liquid nitrogen.

3.8 DATA COLLECTION AND PROCESSING

3.8.1 Data collection

All crystals were mounted on an in house X-ray source (Bruker Microstar (Bruker AXS Inc.) rotating anode at λ=1.54Å) and typically exposed for 10 seconds to the beam. Crystals with good diffraction were saved and re-frozen in liquid nitrogen whereas crystals with low or no diffraction were discarded.

For the collection of a full dataset crystals were sent to the BESSY synchrotron in Berlin. Data was collected at beamline BESSY_14.1 (λ=0.918Å) on a MarMosaic 225 detector.

3.8.2 Data processing

Indexing, cell refinement and integration of the data set was calculated using Mosflm (Leslie & Powell, 2007). Data was scaled and intensities were converted to structure factors with SCALA (CCP4, 1994).

3.9 SOLVING THE STRUCTURE

3.9.1 Solving the phase problem by molecular replacement

Molecular replacement was used to solve initial phases to model HRV2 $2A^{pro}$ bound to inhibitors. The programs Phaser (McCoy *et al.*, 2007) implemented in the PHENIX program suite (Adams *et al.*, 2010) or MolRep (Vagin & Teplyakov, 2010) implemented in the CCP4 program suite (Winn *et al.*, 2011) were used for this purpose.

3.9.2 Model building and refinement

Models were built using the molecular graphics system COOT (Emsley & Cowtan, 2004) and refined using programs implemented in PHENIX and CCP4.

3.10 EXPRESSION AND PURIFICATION OF PROTEASES NEEDED FOR THE ELIMINATION OF TAGS

3.10.1 Tobacco Etch Virus (TEV) protease

A bacterial glycerol stock of *E. coli* BL21 (DE3) transformed with an expression construct coding for an MBP-TEV protease fusion protein was obtained as a kind gift from Gang Dong. Mature TEV-protease is yielded upon self-processing of the protease on its own N-terminus. A single colony of cells was used to inoculate a 20ml starter culture of LB-Amp which was grown at 37°C/200rpm overnight. On the following day two liters of LB-Amp were inoculated with the starter culture till an OD_{600}~0.1 was reached. Cells were grown to and OD_{600}~0.5-0.6 at 37°C/150rpm. Expression was induced by the addition of 0.5mM IPTG and incubation was continued at 37°C/150rpm for three to four hours.

Cells were harvested, resuspended in lysis buffer (20mM Tris-HCl pH 8.0, 300mM NaCl, 5% v/v glycerol, 20mM imidazole, 15mM BME) and disrupted via a homogenizer as described in chapter 3.6.2. The supernatant was then loaded onto a 5ml His-Trap column that had been equilibrated with 5 column volumes of lysis buffer. Purification was performed as described in 3.6.6.1 with the exception that a step gradient (200mM imidazole) was used to elute TEV-protease directly into 20ml of lysis buffer including 20% v/v glycerol to prevent TEV-protease from precipitating. Aliquots of 500µl each were stored at -80°C.

3.10.2 HRV14 3C protease (PreScission protease)

Expression constructs for GST-PreScission protease and His_{10}-GST-Prescission protease fusions were obtained as a kind gift from Georg Mlynek. *E. coli* BL21 (DE3) Rosetta2 pLysS cells were transformed with either of the plasmids and plated on LB-Amp plates. Proteins were expressed as described in chapter 3.6 from 1l of LB-Amp using 0.2mM IPTG at 16°C. Harvesting of cells (lysis buffer: 50mM Tris-HCl pH 8.1, 100mM NaCl, 1mM EDTA, 1mM DTT) and cell disruption via a homogenizer was performed as described in chapter 3.6.2 and PreScission protease was purified via affinity chromatography using a GSTrap column (see chapter 3.6.6.3). Finally, protease was stored in 50mM Tris-HCl pH8.1, 150mM NaCl, 10mM EDTA, 1mM DTT and 20% v/v glycerol at -20°C.

4 RESULTS

4.1 SNAPSHOOTING 2A^{PRO} DURING SELF-PROCESSING – AN ATTEMPT

Only two entries on enteroviral 2A^{pro} can be found in the PDB-database. These are the X-ray structure of HRV2 2A^{pro} (PDB 2HRV) (Petersen *et al.*, 1999) and the NMR structure of CVB4 2A^{pro} (PDB 1Z8R) (Baxter *et al.*, 2006). The coordinates of HRV2 2A^{pro} reveal a conformation corresponding to a post-self-processing state. The N-terminus is rotated out of the active site, allowing the protease to cleave substrates in *trans* without self-inhibition (see chapter 2.4.2.1). The data is therefore of limited value in terms of interpreting the substrate specificity of the 2A^{pro} self-processing reaction. In order to overcome this problem, Baxter *et al.* expressed and purified inactive 2A^{pro} with an N-terminal extension of eight amino acids (see chapter 2.4.2.2). The solution structure was solved with a set of NMR experiments. Nevertheless, information on self-processing could not be extracted as no signals could be obtained for the N-terminal residues due to conformational flexibility.

We speculated that the protein variant used in the NMR studies or rather similar constructs could be useful for crystallization as crystal packing and optimization of the length of the N-terminal extension could favor trapping of the substrate in the binding cleft during self-processing. Hence, we designed a number of 2A^{pro} variants from different enteroviruses differing in expression and purification strategy as well as in the length of the N-terminal extension. Table 4-1 gives an overview of all variants described in the following section together with their length, molecular weight and theoretical pI value. The purified proteins were then subjected to crystallization trials. The outcomes of these experiments are given in this section.

Virus	Plasmid	Protein variant alias	Protein variant expressed from the plasmid: Length [aa]	M_w [kDa]	pI (theoretical)	$2A^{pro}$ variant after cleavage by TEV protease or PreScission protease (if applicable): Length [aa]	M_w [kDa]	pI (theoretical)	Appendix
PV1	pET 3d PV1 $VP1_8$-$2A^{pro}$ (C109A)	PV1a	157	17.5	5.79				APPENDIX I
	pET 3d PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A)	PV1b	163	18.3	6.31				APPENDIX II
	Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A)	PV1c	586	64.7	5.77	162	18.0	5.93	APPENDIX III
	Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) short linker	PV1d	551	60.5	5.18				APPENDIX IV
	Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) short linker His_6	PV1e	559	61.5	5.58	551	60.5	5.18	APPENDIX V
HRV2	pET MBP HRV2 $VP1_9$-$2A^{pro}$ (C106A)	HRV2a	535	59.2	5.03	154	14.4	5.54	APPENDIX VI
CVB4	pET 3d CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110a)	CVB4a	166	18.4	5.72				
	pET MBP CVB4 $VP1_5$-$2A^{pro}$ (C110A)	CVB4b	548	60.2	5.20	158	17.3	4.90	APPENDIX VII
	Mal pET CVB4 $VP1_8$-$2A^{pro}$ (C110A)	CVB4c	585	64.4	5.49	161	17.7	5.09	APPENDIX VIII
	pET MBP CVB4 $VP1_{12}$-$2A^{pro}$ (C110A)	CVB4d	556	61.0	5.21	166	18.0	4.94	APPENDIX IX
	pET MBP CVB4 $VP1_{20}$-$2A^{pro}$ (C110A)	CVB4e	563	61.8	5.15	173	18.9	4.83	APPENDIX X

Table 4-1 Overview of the protein variants covered in section 4.1. The names of the plasmids are given along their aliases used in the text in order to improve readability. The length, molecular weight and theoretical pI are given for all proteins expressed from the corresponding plasmids. If applicable, these specifications are also given for the $2A^{pro}$ variants after being freed from their MBP tags by TEV protease (PV1c, HRV2a, CVB4b, CVB4c, CVB4d and CVB4e) or PreScission protease (PV1e). Exact sequence information about all new variants is given in the appendix.

4.1.1 Poliovirus 2A^{pro} constructs

4.1.1.1 PV1a: pET 3d PV1 VP1$_8$-2A^{pro} (C109A)

All purification steps on ÄKTA FPLC were performed at room temperature.

To produce an expression vector for PV1 2A^{pro} an *Xho*I site was introduced into pET 3d CVB4 His$_6$-VP1$_8$-2A^{pro} (C110A) (Baxter *et al.*, 2006) by replacing the *Bam*HI/*Blp*I fragment with an oligonucleotide cassette containing an *Xho*I site (TIM 1729/1730) (Table 3-1). Two step PCR mutagenesis (see chapter 3.3.7.2) was used to amplify a 1558 bp fragment containing the sequence coding for a start codon followed by PV1 VP1$_8$-2A^{pro} (C109A) (TIM 1704, TIM 1705, TIM 1717, TIM 1728) (Table 3-1) from the plasmid pCITE PV1 VP1-2A^{pro} (Andreas Rötzer, unpublished). The fragment was treated with *Bsp*HI and *Xho*I and used to replace an *Nco*I/*Xho*I fragment of the pET 3d CVB4 His$_6$-VP1$_8$-2A^{pro} (C110A) / *Xho*I plasmid (APPENDIX I).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3); **Medium:** 1 liter LB-Amp; **[IPTG]:** 0.1mM; **Expression temperature:** 16°C; **Expression period:** 5 hours; **[ZnCl$_2$]:** 0µM. Cells were collected by centrifugation (see chapter 3.6.2). The pellet was resuspended in 30ml Buffer A/PV1a (50mM Tris-HCl pH 7.5, 50mM NaCl) per liter of culture. The uninduced control was resuspended in 1ml Buffer A/PV1a. Cells were frozen at -80°C overnight and thawed on RT at the next day. Lysis of cells was finalized by sonication (see chapter 3.6.2.1). Aliquots of total, soluble and insoluble fractions were subjected to 17.5% SDS-PAGE and stained with Coomassie Blue.

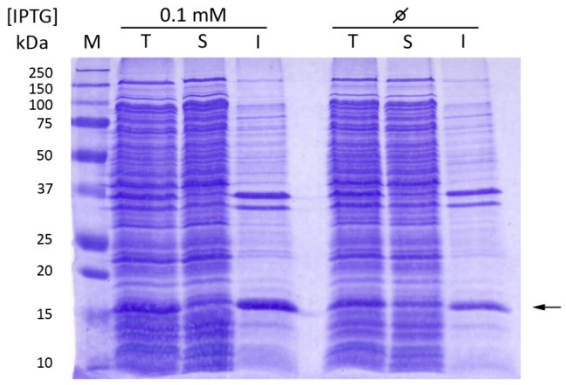


Figure 4-1 Expression of PV1 VP1$_8$-2A^{pro} (C109A) in *E. coli* BL21 (DE 3). 9µL of total (T), soluble (S) and insoluble (I) fractions of induced as well as uninduced samples were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

A strong band corresponding to the size of the construct was detected in the induced as well as the uninduced fraction (marked with an arrow) (Figure 4-1). In the absence of a specific antibody for PV1 $2A^{pro}$, we were unable to verify whether the band contained the recombinant protein. We speculated that the presence of $2A^{pro}$ in the uninduced control could be a result of leaky expression in the BL21 (DE3) cells. A His-tagged version of the construct (see chapter 4.1.1.2) was used to confirm the identity of the bands as $2A^{pro}$.

However, the protein was mainly found in the insoluble fraction, even under the mild over-expression conditions of 0.1mM IPTG and 16°C. Nevertheless, we headed for purification via fractionated ammonium sulfate precipitation (see chapter 3.6.3) using steps of 20% and 45% saturation of salt. Proteins in the pellet of the second step were dissolved in 4ml Buffer B/PV1a (50mM Tris-HCl pH 8.1, 50mM NaCl, 5mM DTT) per liter of bacterial culture. The sample was dialyzed against three washes of Buffer B/PV1a of 500ml each to remove residual ammonium sulfate. Protein aggregates formed during dialysis were collected by centrifugation at 4°C/20.000rpm/30min and redissolved in Buffer B/PV1a. Aliquots of the dialyzed sample, redissolved protein aggregates after dialysis and supernatant of the second precipitation step were then analyzed on 17.5% SDS PAGE (Figure 4-2).

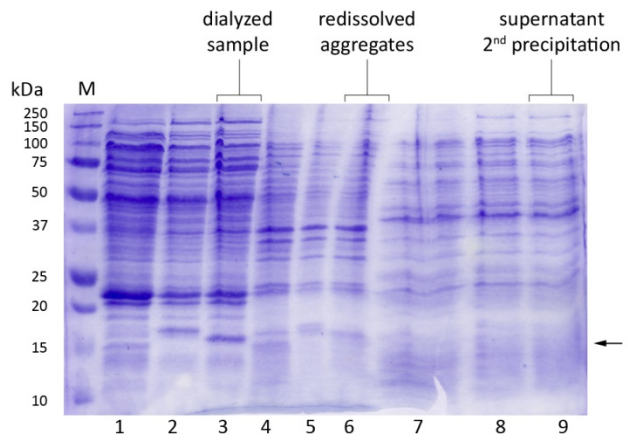


Figure 4-2 Analyzing the fractionated ammonium sulfate precipitation of PV1 $VP1_8$-$2A^{pro}$ (C109A) by 17.5% SDS-PAGE. Lanes 3, 6 and 9 show aliquots of the redissolved and dialyzed sample after the second precipitation step, the redissolved aggregates and the supernatant of the second precipitation step respectively. Protein bands of interest are highlighted with an arrow. Lanes 1, 2, 4, 5, 7 and 8 are from unrelated experiments.

PV1 $VP1_8$-$2A^{pro}$ (C109A) precipitated in the 45% ammonium sulfate fraction (see arrow in Figure 4-2). Hence, the sample was used for further purification by anion exchange chromatography on an ÄKTA FPLC system on room temperature. A MonoQ 10/100 GL was equilibrated with 5 column volumes of Buffer B/PV1a (50mM Tris-HCl pH 8.1, 50mM NaCl, 5mM DTT). The sample was loaded via a superloop and unbound sample was washed off the column using buffer B/PV1a until the UV-signal reapproached the baseline. Subsequently, the column was developed with a NaCl gradient using Buffer C/PV1a (50mM Tris-HCl pH 8.1, 50mM NaCl, 5mM

DTT, 1M NaCl) on pump B of the ÄKTA FPLC system. Eluted protein was detected by tracing absorption of UV-light at 280 nm and fractions of 2ml each were collected (Figure 4-3).

The chromatogram showed only minor peaks in the low salt regions and only one pronounced peak at around 750mM NaCl (retention volume=217ml). Running 9µl of each peak-fraction on a 17.5% SDS-PAGE should confirm whether one of the peaks contained the recombinant protein. However, no protein corresponding to the size of PV1 $VP1_8$-$2A^{pro}$ (C109A) could be detected, neither in the fractions nor in the collected flow-through. In fact, the only peak containing any detectable protein by Coomassie staining was found to be the one at a retention volume of (V_r) 84ml (data not shown). Consequently, no further purification steps could be performed.

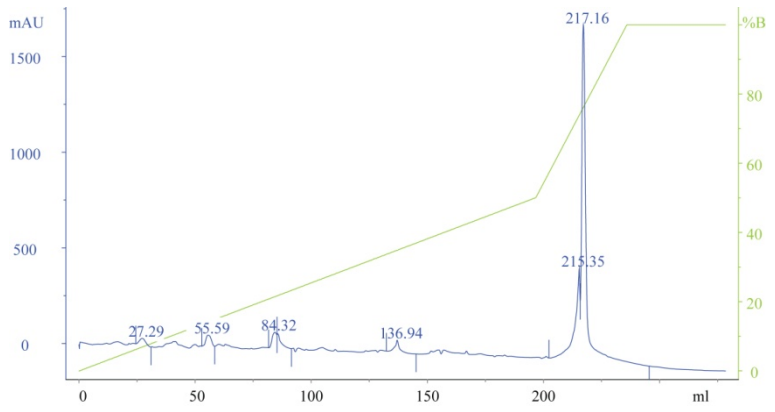


Figure 4-3 Chromatogram of PV1 $VP1_8$-$2A^{pro}$ (C109A) MonoQ run. PV1 $VP1_8$-$2A^{pro}$ (C109A) containing sample was loaded on a MonoQ 10/100 GL using an ÄKTA FPLC system. Bound proteins were eluted by applying a NaCl gradient (green). Peaks in the UV trace (blue) are labeled with the respective retention volumes.

4.1.1.2 PV1b: pET 3d PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A)

All purification steps on ÄKTA FPLC were performed at room temperature.

Cloning of this fragment with a His_6-tag was identical to pET 3d PV1 $VP1_8$-$2A^{pro}$ (see section 4.1.1.1) with exception of the usage of TIM 1716 (instead of TIM 1717) (Table 3-1). The construct consists of an N-terminal hexa-His-tag directly fused to the last eight residues of VP1 and the inactivated $2A^{pro}$ (APPENDIX II).

E. coli BL21 (DE3) were transformed with the corresponding plasmid and recombinant protein was expressed as described in section 4.1.1.1. Cells were processed as in chapter 4.1.1.1

and 9µL of the total, soluble and insoluble fractions were analyzed on a 17.5% SDS-PAGE (Figure 4-4).

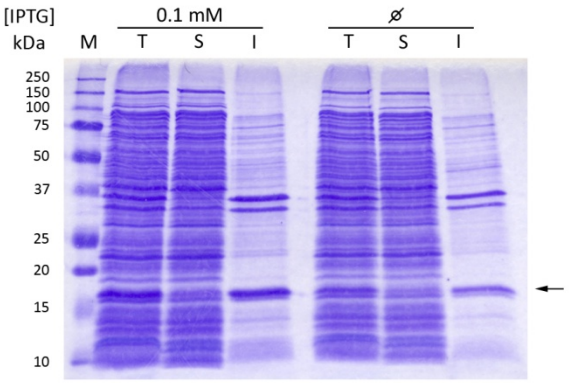


Figure 4-4 Expression of PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A) in *E. coli* BL21 (DE3). 9µL of total (T), soluble (S) and insoluble (I) fractions of induced as well as uninduced samples were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

A strong band running between 15 and 20 kDa was detected in the induced as well as in the uninduced fraction. Again, the protein seemed to be highly insoluble under the expression conditions mentioned above. In order to confirm the identity of the bands, proteins were blotted on PVDF membrane and analyzed via His-probe (Figure 4-5) (see chapter 3.5.2.3). The blot clearly identified the bands as the protein of interest. However, it also confirmed the insolubility of poliovirus $2A^{pro}$.

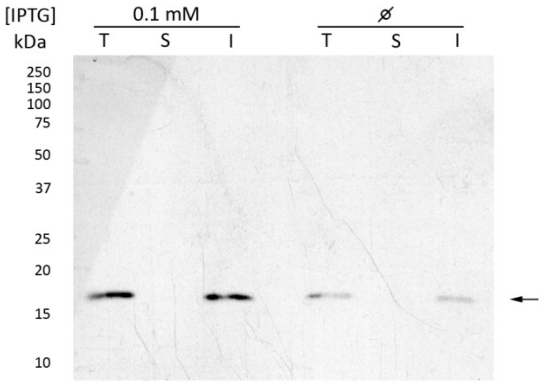


Figure 4-5 α-His blot of *E. coli* BL21 (DE3) crude cell extracts containing PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A). 9µL of total (T), soluble (S) and insoluble (I) fractions of induced as well as uninduced samples were separated on 17.5% SDS-PAGE, blotted on PVDF membrane and analyzed using His-probe.

In order to enhance solubility we aimed to slow down the rate of expression by using M9 minimal medium (see chapter 3.4.2). Furthermore, the effect of different cell lysis methods was to be examined. After expression, cells were divided into two parts and either broken by sonicaton or by a high pressure homogenizer (Figure 4-6) (see chapter 3.6.2).

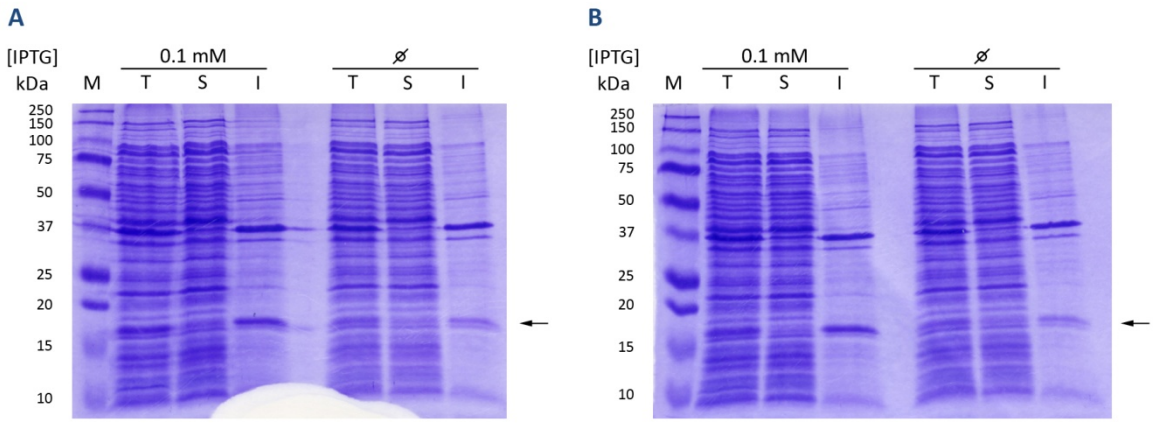


Figure 4-6 Expression of PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A) in *E. coli* BL21 (DE3) grown in M9 minimal medium. 9µL of total (T), soluble (S) and insoluble (I) fractions of induced as well as uninduced samples were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow. **(A)** Crude extract from cells broken by sonication. **(B)** Crude extract from cells broken with a high pressure homogenizer.

No essential differences between the two modes of cell disruption could be detected. Neither the total amount of protein nor the fraction of soluble recombinant protein is dependent on the method. Moreover, the use of M9 minimal medium also did not seem to stimulate the production of soluble protein. The Coomassie Blue stained gels as well as the blots developed with His-probe (Figure 4-7) did not indicate the presence of any soluble recombinant protein.

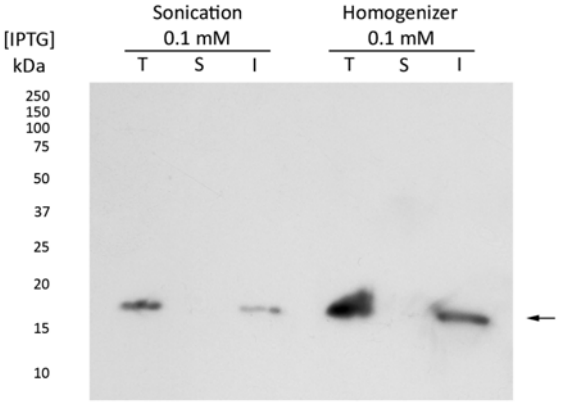


Figure 4-7 α-His blot of *E. coli* BL21 (DE3) crude cell extracts including PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A). 9µL of total (T), soluble (S) and insoluble (I) fractions of cells broken either by sonication or a high pressure homogenizer were separated on 17.5% SDS-PAGE, blotted on PVDF membrane and analyzed using His-probe.

Nevertheless, soluble fractions of both breakup methods were pooled and fractionated ammonium sulfate precipitation was applied as described before. The precipitated protein of the second fraction was dissolved in 4ml Buffer B/PV1b (50mM Tris-HCl pH 8.1, 50mM NaCl) and dialyzed against three washes of 500ml of the same buffer. A 1ml His-Trap column (see chapter 3.6.6.1) was equilibrated with 10ml of Buffer B/PV1b (50mM Tris-HCl pH 8.1, 50mM NaCl). The

sample was loaded via a superloop at a flow rate of 0.5ml/min and the column was developed with a gradient of imidazole (up to 500mM) (Figure 4-8).

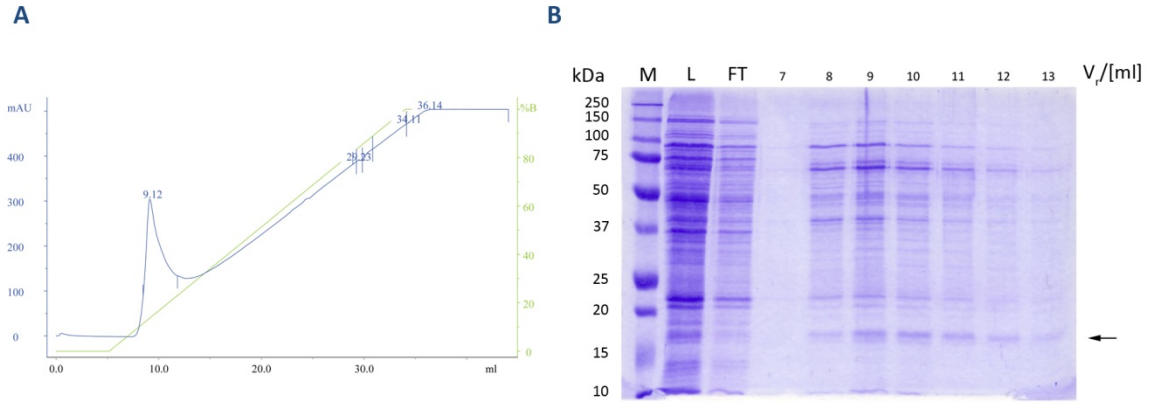


Figure 4-8 His-Trap purification of PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A). (A) PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A) containing sample was loaded on a 1ml nickel charged His-Trap column. Protein (blue UV trace) was eluted from the column with rising concentrations of imidazole in the elution buffer (green curve) and fractions of 1ml each were collected. **(B)** Aliquots (9µL of load (L), flow-through (FT) and elution fractions) of the protein containing peak (V_r=9.12 ml) were separated on a 17.5% SDS-PAGE and stained with Coomassie Blue. Recombinant protein bands are marked with an arrow.

The fractions contained a substantial amount of impurities (Figure 4-8, lanes 8-13). Addition of 10mM imidazole to the binding buffer improved the efficiency of this affinity step in the subsequent purification of protein in a scaled up experiment. Nevertheless, a low amount of PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A) could be detected (see lanes corresponding to V_r=8-13ml in Figure 4-8 B). These fractions were pooled, concentrated to a volume of 1ml, dialyzed against Buffer B/PV1b to remove imidazole and loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/PV1b (50mM Tris-HCl pH 8.1, 50mM NaCl). Isocratic elution was performed with Buffer B/PV1b and fractions of 2ml were collected. However, the amount of loaded protein seemed to be too low as no protein could be recovered from the size exclusion run (data not shown). Even a scale up of the expression to 6 liters of bacterial culture did not allow the recovery of any recombinant protein from the size exclusion column (data not shown).

4.1.1.3 PV1c: Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A)

In order to increase the solubility of PV1 $2A^{pro}$ we aimed to produce a fusion protein using maltose binding protein (MBP) as a solubility tag. The protein of interest can be cleaved from the

tag via TEV protease. The plasmid Mal pET was a kind gift of Gang Dong (Max F. Perutz laboratories) and has been described in chapter 3.1.1.2.

A fragment coding for PV1 $VP1_8$-$2A^{pro}$ (C109A) was amplified from pET 3d PV1 His_6-$VP1_8$-$2A^{pro}$ (C109A) using the primers TIM 1759 and TIM 1760 (Table 3-1). TIM 1759 contained the sequence of an *Nhe*I site, TIM 1760 the sequence of a *Bam*HI site in their 5' nonbinding overhang. Ligation of the *Nhe*I/*Bam*HI digested PCR fragment into a similarly treated Mal pET vector resulted in a plasmid coding for a fusion protein with five additional amino acids between TEV cleavage site and the $VP1_8$-$2A^{pro}$ (C109A) sequence (APPENDIX III). *E. coli* BL21 (DE3) Rosetta2 pLysS were transformed with the plasmid and plated on LB-agar plates containing kanamycin as the selecting antibiotic. Single colonies were used to inoculate LB-medium.

A small scale test expression was performed to test optimal expression conditions. Three batches of 100ml each of LB-Kan medium were inoculated with an overnight culture of *E. coli* BL21 (DE3) Rosetta pLysS cells transformed with Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) and grown to an OD_{600}~0.5 at 37°C. Subsequently, samples were induced via addition of 0.5mM IPTG and expression was performed for either five or twenty hours at 16°C or six hours at 25°C. Cells were harvested by centrifugation, resuspended in Buffer A/PV1c (50mM Tris-HCl pH 7.5, 50mM NaCl) and broken by sonication (chapter 3.6.2.1). 4µl of total, soluble and insoluble fractions were then separated on a 12.5% SDS-PAGE and analyzed by Coomassie staining (Figure 4-9).

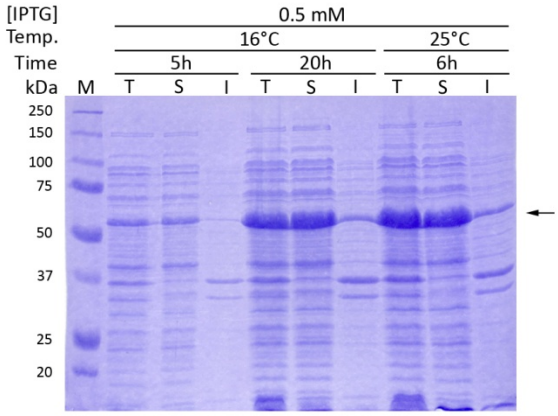


Figure 4-9 Test expression of PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) in *E. coli* BL21 (DE3) Rosetta2 pLysS. 4µL of total (T), soluble (S) and insoluble (I) fractions of induced samples of different expression conditions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

The fusion protein was strongly over-expressed at both 16°C and 25°C and was found nearly exclusively in the soluble fraction. Further test expressions at 37°C yielded insoluble fusion protein (data not shown). Hence, we decided to scale the experiment up to a one liter culture and express at 25°C for six hours with an IPTG concentration of 0.5mM (Figure 4-10).

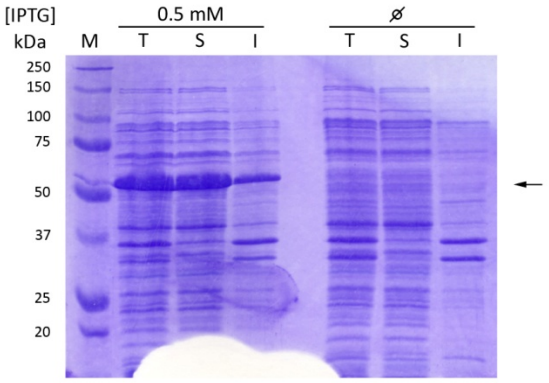


Figure 4-10 Expression of PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) in *E. coli* BL21 (DE3) Rosetta2 pLysS. 3µL of total (T), soluble (S) and insoluble (I) fractions of induced and uninduced samples were separated on 12.5% SDS PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

The strong over-expression yielded an estimated total amount of at least 50mg of recombinant fusion protein per liter of culture, which was predominantly found in the soluble fraction.

Initially, we intended to apply a classical purification protocol using fractionated ammonium sulfate precipitation, ion exchange chromatography followed by polishing via size exclusion chromatography. Thus, the soluble fraction was subjected to protein precipitation with 20%, 40% and 60% saturation of ammonium sulfate. Proteins were precipitated for 12 hours at 4°C at each step followed by collection of precipitated protein via centrifugation. Precipitated protein of each step was then redissolved in 8ml Buffer B/PV1c (50mM Tris-HCl pH 8.5, 50mM NaCl, 5mM DTT, 5% v/v glycerol). All three samples were then dialyzed against Buffer B/PV1c to remove residual ammonium sulfate. Protein that aggregated during dialysis was pelletized and dissolved in 8ml Buffer B/PV1c. In order to investigate which fraction included the fusion protein, 3µl of each fraction were separated on 12.5% SDS-PAGE and stained with Coomassie Blue (Figure 4-11).

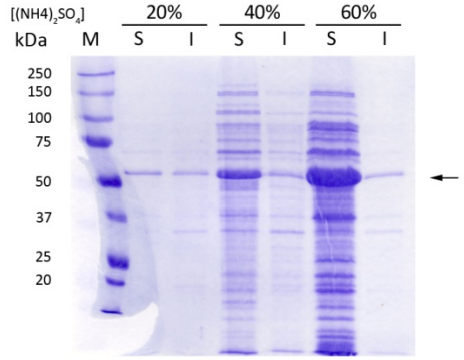


Figure 4-11 Fractionated ammonium sulfate precipitation of PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) containing crude cell extract. 3µl of each fractionation step were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. (S) Supernatant after removal of residual ammonium sulfate via dialysis. (I) Redissolved aggregated protein after removal of residual ammonium sulfate via dialysis. The arrow marks over expressed recombinant protein.

The majority of recombinant protein was found in the fraction that was precipitated by the addition of ammonium sulfate to 60% saturation. Hence, this fraction was submitted to further purification by anion exchange via a MonoQ 10/100 GL column preequilibrated with Buffer B/PV1c (50mM Tris-HCl pH 8.5, 50mM NaCl, 5mM DTT, 5% v/v glycerol). Bound proteins were eluted from the column by the use of a NaCl gradient with a maximum NaCl concentration of 1M and fractions of 1ml each were collected (Figure 4-12).

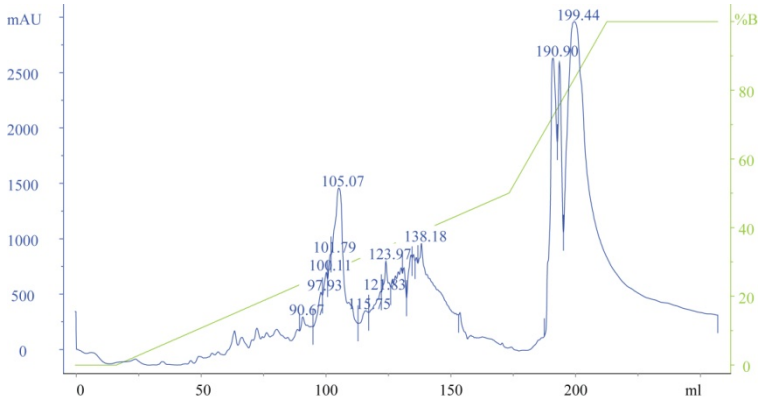


Figure 4-12 Chromatogram of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) MonoQ run. PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing sample was loaded on a MonoQ 10/100 GL using an ÄKTA FPLC system. Bound proteins were eluted by applying a NaCl gradient (green). Peaks in the UV trace (blue) are labeled with the respective retention volumes.

In order to verify the identity of proteins in the different peaks, fractions with V_r=82-166ml and V_r=188-204ml were analyzed on 12.5% SDS-PAGE (Figure 4-13).

Interestingly, recombinant fusion protein was detected over a wide range of the NaCl gradient, peaking first at V_r=105ml followed by a broad distribution from V_r=112ml to V_r=166ml. We speculated that this behavior was either due to the bi-globular nature of the fusion protein or due to unspecific binding caused by aggregation of a large fraction of the fusion protein. For this reason, we split the fractions into three pools (Table 4-2).

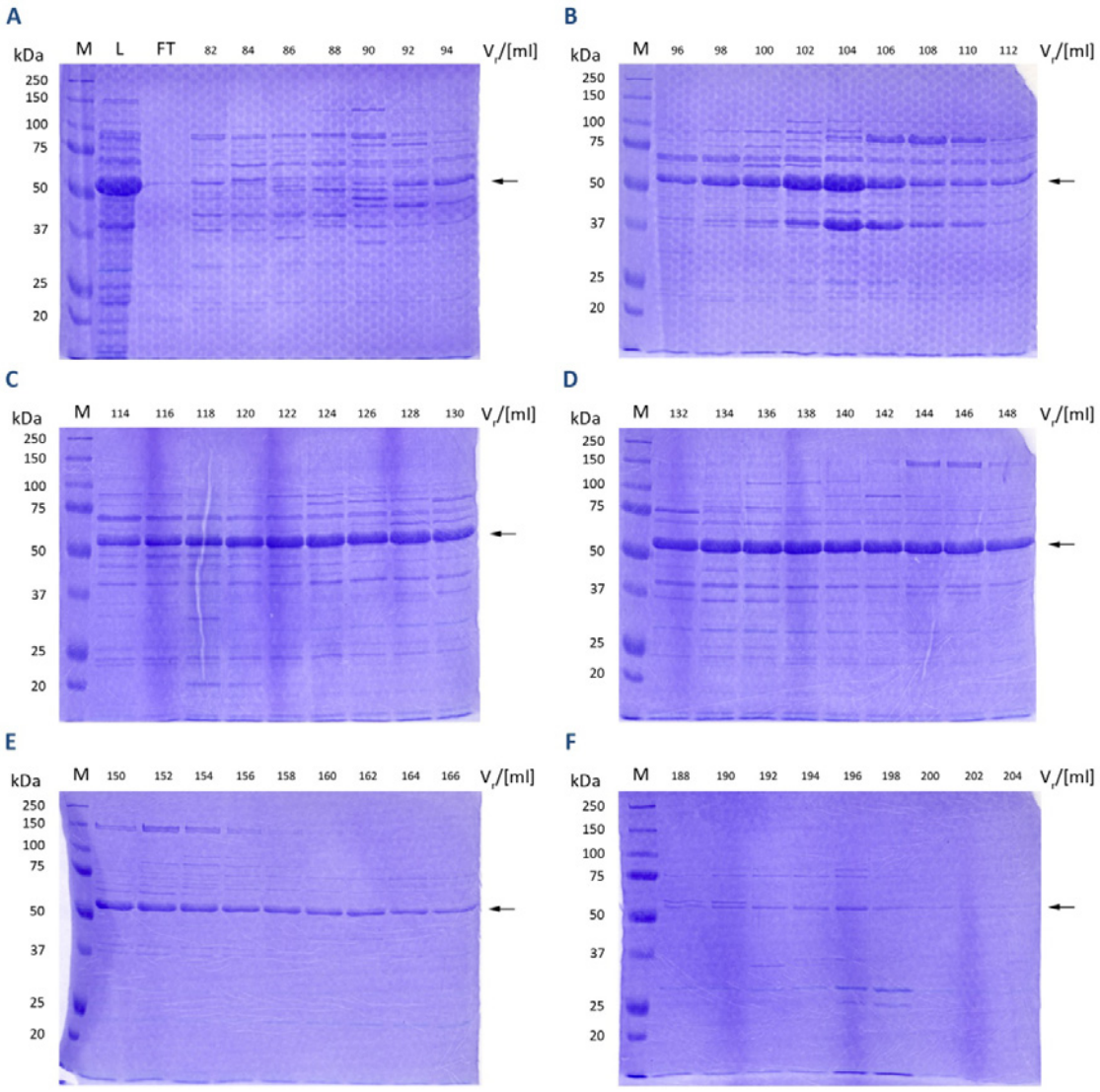


Figure 4-13 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from MonoQ purification by 12.5% SDS-PAGE. 3µl of load (L), 9µl of flow-through (FT) and 5µl of fractions were loaded and separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

Pool Nr.	V_r (first fraction) / [ml]	V_r (last fraction) / [ml]
1	94	112
2	112	130
3	130	166

Table 4-2 Distribution of the three pools composed of fractions from MonoQ purification of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A).

Each pool was then concentrated to 4ml and run on a HiLoad Superdex 75 prep grade 26/60 that was preequilibrated and developed with Buffer B/PV1c (50mM Tris-HCl pH 8.5, 50mM NaCl, 5mM DTT, 5% v/v glycerol). Fractions of 2ml each were collected.

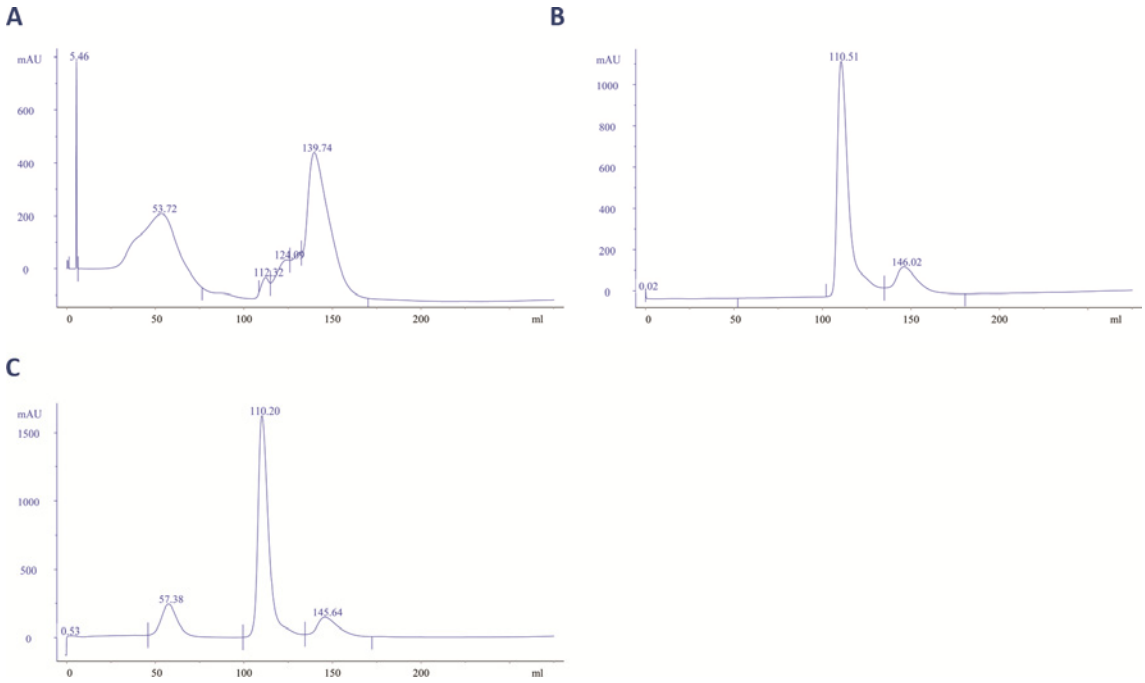


Figure 4-14 Size exclusion profiles of pools 1-3 from PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) loaded onto a HiLoad Superdex 75 prep grade 26/60. PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing sample was loaded on a HiLoad Superdex 75 prep grade 26/60 using an ÄKTA FPLC system. The column was developed with Buffer B/PV1c. Peaks in the UV trace (blue) are labeled with the respective retention volumes. **(A)** Pool 1. **(B)** Pool 2. **(C)** Pool 3.

The size exclusion profile of pool 1 exhibited a strong peak at a V_r~140ml (Figure 4-14 A). The profiles of pool 2 and 3 displayed a similar (but much smaller) peak at V_r~145ml. Investigation of the protein fractions on 12.5% SDS-PAGE demonstrated that these peaks contained mainly recombinant fusion protein with major impurities, as shown in Figure 4-15 (representing pool 1) and Figure 4-16 (representing pool 2). However, the majority of protein in the pools 2 and 3 eluted in a peak with V_r~110ml corresponding with the void volume of the column. SDS-PAGE confirmed that these fractions were mainly composed of recombinant protein. Thus, this fraction of recombinant protein formed dimers, oligomers or soluble aggregates too big to be retained on the column (Figure 4-14 B and C).

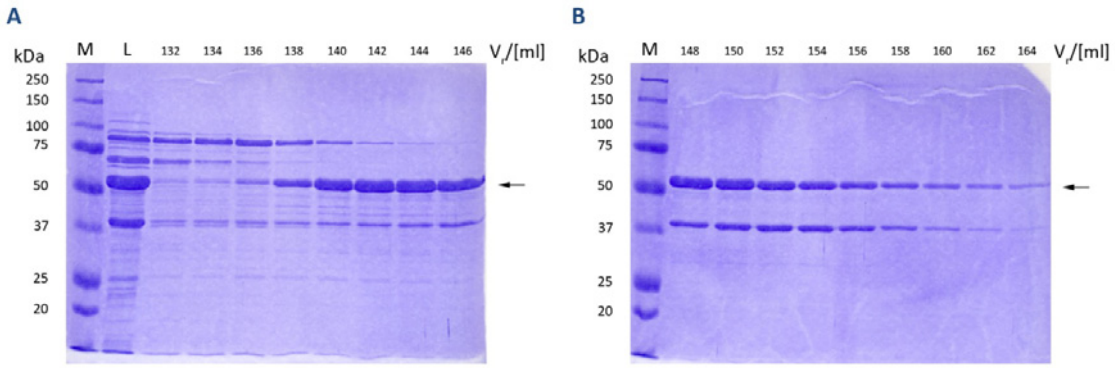


Figure 4-15 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from size exclusion purification by 12.5% SDS-PAGE. Pool 1. 3µl of load (L) and 6µl of fractions were loaded and separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

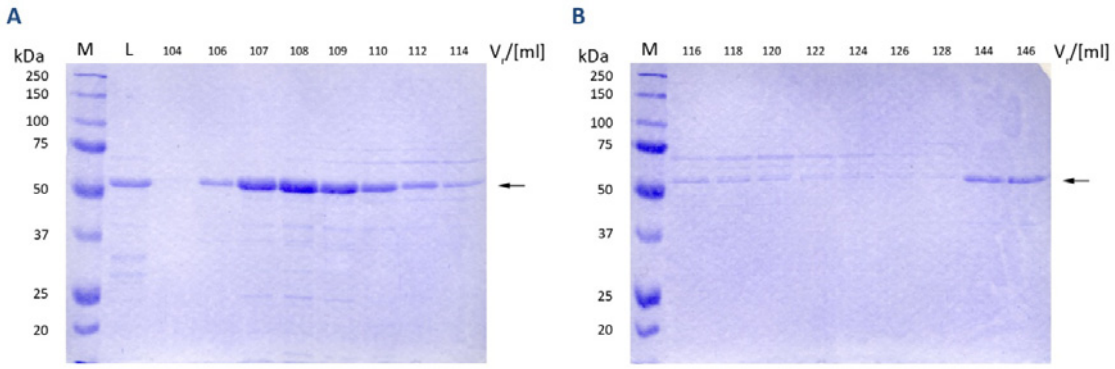


Figure 4-16 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from size exclusion purification by 12.5% SDS-PAGE. Pool 2. 3µl of load (L) and 5µl of fractions were loaded and separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

As the aggregated proteins of pool 2 and pool 3 were of higher purity, we attempted to use higher ionic strength in the buffer to dissolve the aggregates. We therefore re-concentrated fractions with V_r=106-114ml of the initial size exclusion runs of pool 2 and pool 3 and ran further gel filtrations using Buffer B2/PV1c (50mM Tris-HCl pH 8.5, 300mM NaCl, 5mM DTT, 5% v/v glycerol) for re-concentrated protein from pool 2 (Figure 4-17) and Buffer B3/PV1c (50mM Tris-HCl pH 8.5, 300mM $(NH_4)_2SO_4$, 5mM DTT, 5% v/v glycerol) for re-concentrated protein from pool 3 (Figure 4-18).

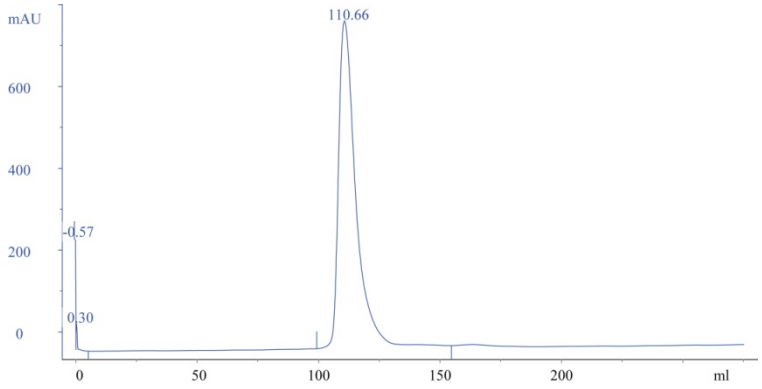


Figure 4-17 Chromatogram of PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) size exclusion run (pool 2) using buffer B2/PV1c. PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) containing sample was loaded on a HiLoad Superdex 75 prep grade 26/60 using an ÄKTA FPLC system. The column was developed with Buffer B2/PV1c. Peaks in the UV trace (blue) are labeled with the respective retention volumes.

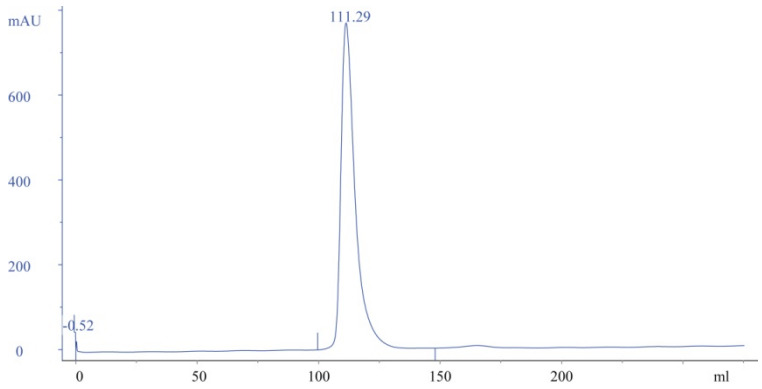


Figure 4-18 Chromatogram of PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) size exclusion run (pool 3) using buffer B3/PV1c. PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A) containing sample was loaded on a HiLoad Superdex 75 prep grade 26/60 using an ÄKTA FPLC system. The column was developed with Buffer B3/PV1c. Peaks in the UV trace (blue) are labeled with the respective retention volumes.

Neither higher concentrations of NaCl nor addition of $(NH_4)_2SO_4$ could reverse protein aggregation. Both size exclusion profiles show only extensive peaks corresponding to the void volume. The protein samples as well as the classical purification strategy via ammonium sulfate precipitation, anion exchange and size exclusion chromatography were therefore abandoned.

The potentially non-aggregated samples of pool 1 were used to test further purification by affinity chromatography via amylose resin. Fractions with a V_r=136-164ml (see Figure 4-15) were loaded on a gravity column containing 9ml of amylose resin (NewEngland BioLabs) that was equilibrated with Buffer B/PV1c (50mM Tris-HCl pH 8.5, 50mM NaCl, 5mM DTT, 5% v/v glycerol). Unbound proteins were removed by washing with 5 column volumes Buffer B/PV1c and bound fusion protein was eluted via Buffer B/PV1c including 10mM maltose. The eluate was fractionated

into 2ml and 9µl of flow-through and wash fractions as well as 5µl of elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue (Figure 4-19).

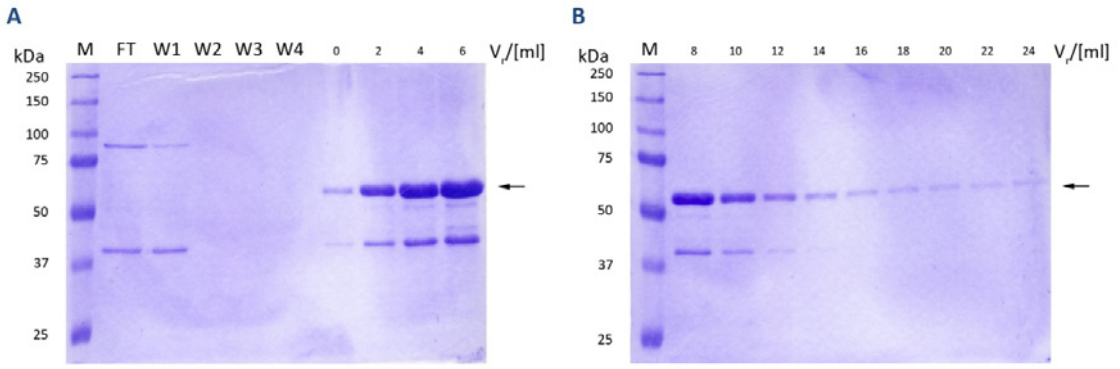


Figure 4-19 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from affinity chromatography via amylose resin by 12.5% SDS-PAGE. 9µl of flow-through (FT) and wash fractions (W) as well as 5µl of elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Recombinant fusion protein bands are marked with arrows.

The eluted protein sample still contained a major impurity running above the 37kDa band of the molecular weight marker. As it co-purified with the MBP fusion protein and corresponded to the size of MBP (42.9 kDa), we suspected the band to be intrinsic maltose binding protein. Hence, a new expression in LB-medium containing extra glucose (2g/l) (LB-Gluc) was set up in order to prevent the bacteria from fermenting maltose.

Fusion protein was expressed in LB-Kan-Gluc in BL21 (DE3) Rosetta2 pLysS cells at 25°C and an IPTG concentration of 0.3mM for six hours. Cells were harvested by centrifugation and resuspended in 30ml of Buffer C/PV1c (Tris-HCl ph 8.5, 200mM NaCl). Cells were broken by sonication and 3µl aliquots of total, soluble and insoluble fraction were separated and analyzed on 12.5% SDS-PAGE (Figure 4-20).

A satisfactory fraction of the fusion protein was found in the soluble fraction. As the classical purification strategy with ammonium sulfate fractionation, anion exchange and gel filtration was not successful, we aimed to purify the protein via affinity chromatography (amylose resin) and gel filtration. The supernatant of the crude cell extract was loaded onto 14ml of amylose resin that had been equilibrated with Buffer D/PV1c (50mM Tris-HCl pH 8.5, 200mM NaCl, 5mM DTT, 5% v/v glycerol). Unbound protein was removed by extensive washing with 10 column volumes of Buffer D/PV1c. Subsequently, proteins were eluted with Buffer D/PV1c including 10mM maltose. The eluate was divided into fractions of 9ml each and samples were analyzed on 12.5% SDS-PAGE (Figure 4-21).

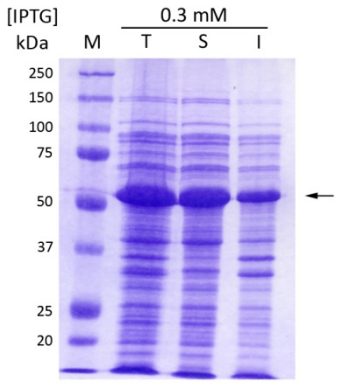


Figure 4-20 Expression of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) in *E. coli* BL21 (DE3) Rosetta2 pLysS using LB-Kan-Gluc. 3µL of total (T), soluble (S) and insoluble (I) fractions were separated on 12.5% SDS PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

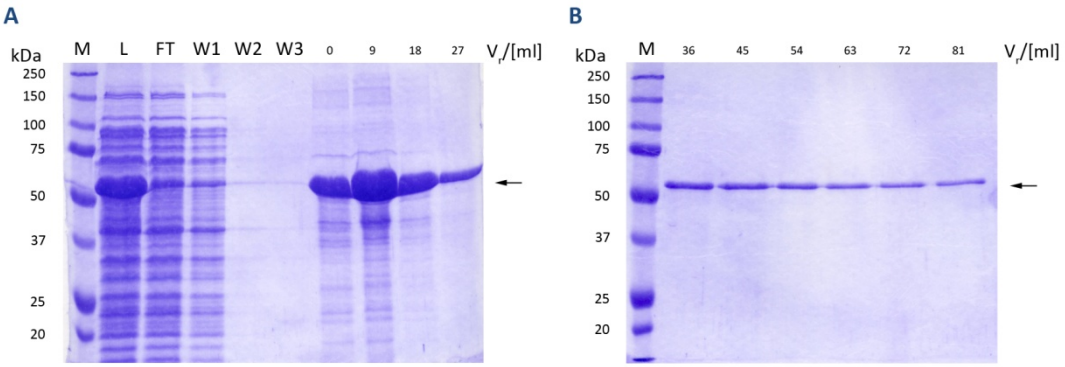


Figure 4-21 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from affinity chromatography via amylose resin by 12.5% SDS-PAGE. 9µl of load (L), flow-through (FT) and wash fractions (W) as well as 9µl of elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Recombinant fusion protein is marked with arrows.

The first three elution fractions (corresponding to retention volumes up to 27ml) seemed to be less pure than the following fractions. Consequently, we decided to split the samples to build pool 1 containing fractions one to three and pool 2 containing the remaining fractions. Each of the two pools was concentrated to around 6ml and loaded onto a HiLoad Superdex 200 prep grade 26/60 that had been equilibrated with Buffer D/PV1c (50mM Tris-HCl pH 8.5, 200mM NaCl, 5mM DTT, 5% v/v glycerol) (Figure 4-22).

Aliquots of the peaking fractions were analyzed on 12.5% SDS-PAGE (Figure 4-23 and Figure 4-24). The size exclusion profiles of both pools showed three distinct peaks. The peak with V_r of about 117ml corresponded to the void volume peak and contained again soluble aggregates of the fusion protein. Nevertheless, the second peak in the chromatograms contained non aggregated fusion protein in satisfying amounts. However, in pool 1 this peak showed a clear shoulder towards higher retention volume. Analysis of these fractions showed that they

contained a contaminant (possibly intrinsic MBP) (Figure 4-23). Peak 3 of both pools did not contain any detectable protein.

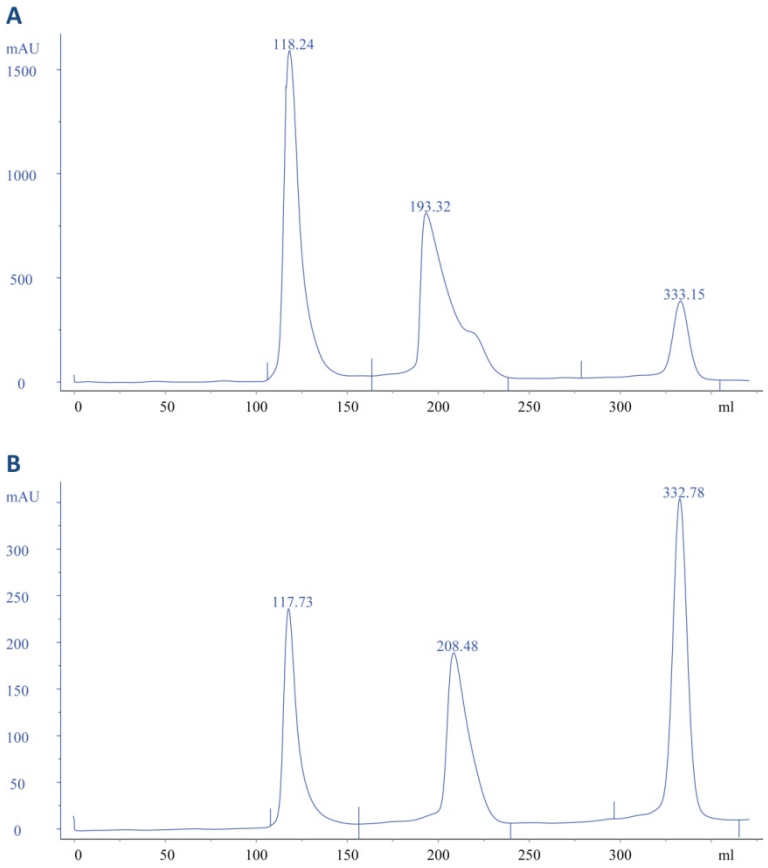


Figure 4-22 Chromatogram of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) size exclusion runs (pool 1 (A), pool 2 (B)). PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing sample was loaded on a HiLoad Superdex 200 prep grade 26/60 using an ÄKTA FPLC system. The column was developed with Buffer D/PV1c (50mM Tris-HCl pH 8.5, 200mM NaCl, 5mM DTT, 5% v/v glycerol). Peaks in the UV trace (blue) are labeled with the respective retention volumes.

In order to further investigate the aggregation behavior of the fusion protein, fractions with a V_r of 187-211ml of pool 1 were again pooled, concentrated to a volume of 5ml and reloaded to a HiLoad Superdex 75 prep grade 26/60 that was equilibrated with Buffer E/PV1c (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% Glycerol). The column was developed with Buffer E/PV1c and fractions of 2ml were collected (Figure 4-25). Analysis of the fractions around the peak with V_r~152ml on SDS-PAGE confirmed its identity as 2A^{pro} (data not shown). The void volume peak (V_r=105ml) has almost disappeared and was only present as a minor shoulder to the main peak. Hence, we considered the main peak to be stable and not in an equilibrium with an aggregated state.

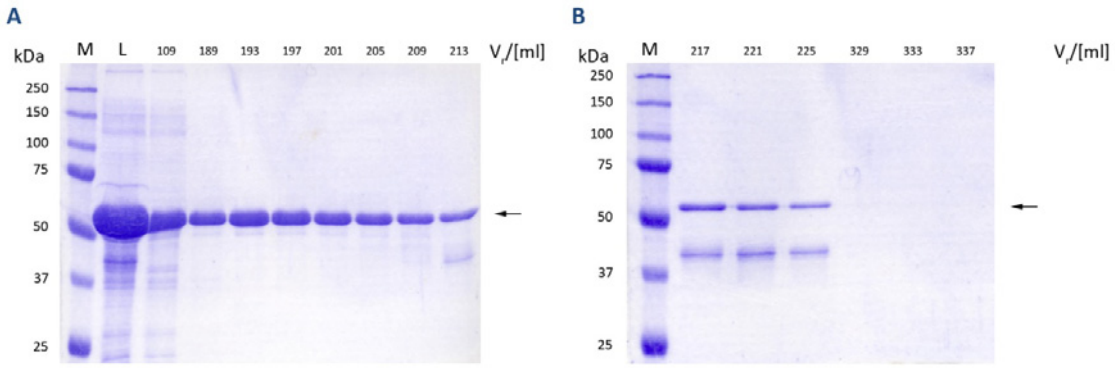


Figure 4-23 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from size exclusion purification by 12.5% SDS-PAGE. Pool 1. 5µl of load (L) and 5µl of fractions were loaded and separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

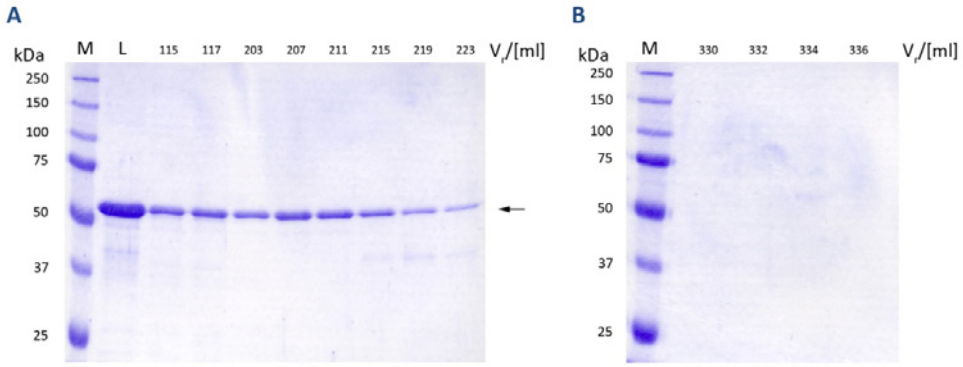


Figure 4-24 Analysis of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) containing fractions from size exclusion purification by 12.5% SDS-PAGE. Pool 2. 5µl of load (L) and 5µl of fractions were loaded and separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

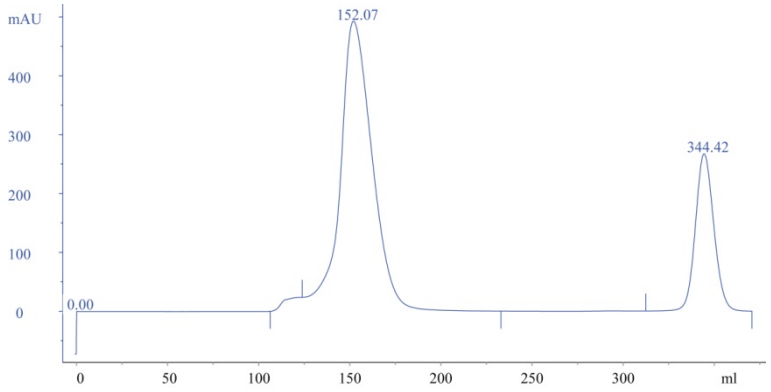


Figure 4-25 Chromatogram of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) size exclusion. Investigation of aggregation behavior of re-concentrated protein. PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) was loaded on a HiLoad Superdex 75 prep grade 26/60 using an ÄKTA FPLC system. The column was developed with Buffer E/PV1c (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). Peaks in the UV trace (blue) are labeled with the respective retention volumes.

The fractions of the main peak were again pooled and concentrated to 47mg/ml (725µM) as calculated by absorbance at 280nm on a NanoDrop-1000 device (chapter 3.6.8.3). With regard to subsequent crystallization trials, addition of maltose to a concentration of 1mM was thought to enhance the conformational stability of maltose binding protein. Figure 4-26 shows the final sample on a Coomassie stained 12.5% SDS-PAGE.

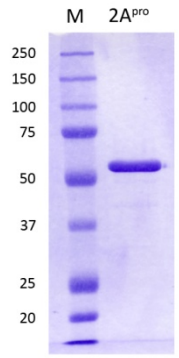


Figure 4-26 Final, purified sample of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A). 0.1µl of a protein solution containing 47mg/ml of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A) in Buffer E/PV1c (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% glycerol) and 1mM maltose were separated on 12.5% SDS-PAGE and stained with Coomassie Blue.

An aliquot of the sample was diluted to 12 mg/ml in Buffer E/PV1c (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol) and samples of these two concentrations were used for initial crystal screens using the commercial screens "Crystal Screen 1+2" from Hampton Research. A nanoliter crystallization robot (Phoenix) was used to set up 96well screens with sitting drops (see chapter 3.7.1). Indeed, crystals were detected in the wells of condition G7 (100mM HEPES pH 7.5, 20% Jeffamine M-600) (Figure 4-27) of both starting concentrations of protein. At 22°C, crystals were already detected at the first day after setup whereas it took seven days at 4°C to grow crystals of detectable size. Moreover, crystals at 22°C grew bigger and reached a maximum dimension of 0.6-0.7 mm. Unfortunately, crystals started to disintegrate at the time of the next inspection. Crystals lost both size and sharp edges and finally could not be detected any more before further experiments could be conducted.

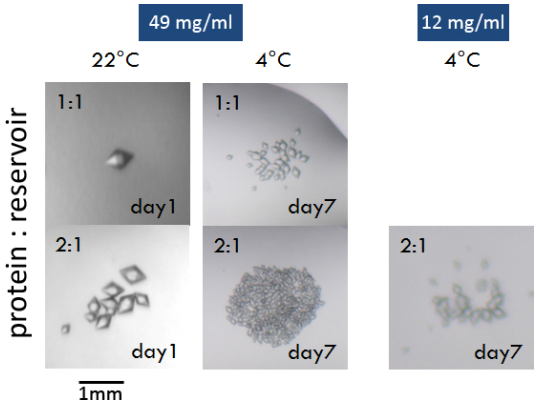


Figure 4-27 Crystal hits in initial crystal screens of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A). Screens were set up as sitting drops with 50µl reservoir solution and drops of different size. Drop 1 contained 0.1µl of protein solution as well as precipitating agent. Drop 2 contained 0.2µl of protein solution and 0.1µl of precipitating agent. Crystals were detected in condition G7 (100mM HEPES pH 7.5, 20% Jeffamine M-600).

A repeat of the screen with the same batch of protein and identical conditions did not yield any crystals. For this reason, the chemical space around the original condition was screened by varying the pH as well as the concentration of Jeffamine M-600 (Table 4-3).

	1	2	3	4	5	6	7
A	pH 6.5 14% J. M-600	pH 6.5 16% J. M-600	pH 6.5 18% J. M-600	pH 6.5 20% J. M-600	pH 6.5 22% J. M-600	pH 6.5 24% J. M-600	pH 6.5 26% J. M-600
B	pH 7.0 14% J. M-600	pH 7.0 16% J. M-600	pH 7.0 18% J. M-600	pH 7.0 20% J. M-600	pH 7.0 22% J. M-600	pH 7.0 24% J. M-600	pH 7.0 26% J. M-600
C	pH 7.5 14% J. M-600	pH 7.5 16% J. M-600	pH 7.5 18% J. M-600	pH 7.5 20% J. M-600	pH 7.5 22% J. M-600	pH 7.5 24% J. M-600	pH 7.5 26% J. M-600
D	pH 8.0 14% J. M-600	pH 8.0 16% J. M-600	pH 8.0 18% J. M-600	pH 8.0 20% J. M-600	pH 8.0 22% J. M-600	pH 8.0 24% J. M-600	pH 8.0 26% J. M-600
E	pH 8.5 14% J. M-600	pH 8.5 16% J. M-600	pH 8.5 18% J. M-600	pH 8.5 20% J. M-600	pH 8.5 22% J. M-600	pH 8.5 24% J. M-600	pH 8.5 26% J. M-600

Table 4-3 Optimization screen for initial crystal hits of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A). pH as well as the concentration of Jeffamine M-600 were varied around the original hit conditions (blue). The concentration of HEPES buffer was maintained at 100mM.

A hanging drop system with a reservoir volume of 100µl and drop sizes of 2µl (1:1 protein:reservoir solution) as well as 3µl (2:1 protein:reservoir solution) using 48-well crystallization plates was used to perform the crystallization screens at 22°C. However, drops remained clear and no crystals could be obtained (data not shown). In a final attempt to reproduce the crystals, a fresh batch of protein was produced and used in the optimization screen. Fusion protein was expressed under the conditions described before (see page 97), cells were broken via sonication and recombinant protein was purified via amylose resin (with buffer C/PV1c) and gel filtration on a HiLoad Superdex 200 prep grade 26/60 (with buffer E/PV1c). Fractions of satisfactory purity where then pooled and concentrated to 31mg/ml. Samples of the different purification steps were separated on 12.5% SDS PAGE and stained with Coomassie Blue to monitor the success of purification (Figure 4-28). Maltose was added to the final sample to a final concentration of 0.5mM and used for further crystallization screens using the setup described (Table 4-3) at 4°C as well as 22°C. Drops were regularly monitored. However, drops remained clear and no crystals could be detected.

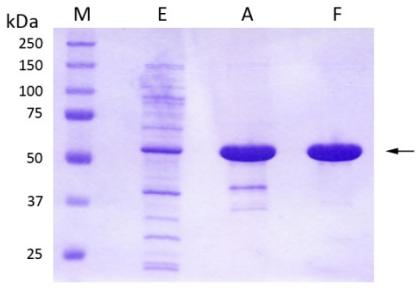


Figure 4-28 Monitoring the progress of purification of PV1 MBP-His$_{10}$-VP1$_8$-2A^{pro} (C109A). 3µl of cleared cell extract (E), 3µl of pooled fractions after affinity chromatography via amylose resin (A) and 1µl of final, concentrated sample after gel filtration were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Fusion protein is marked with an arrow.

For this reason, the initial hit conditions were dropped and additional screens were performed. Fresh protein batches were produced as described above and used to set up a number of commercial crystallization screens using a nanoliter crystallization robot (Phoenix) (see chapter 3.7.1) (APPENDIX XIV). Drops were regularly monitored. However, no crystals could be observed.

The 37 amino acids long stretch between MBP and PV1 $VP1_8$-$2A^{pro}$ (C109A) made up of linker, His_{10} and TEV recognition site is most likely unfavorable for crystallization due to the introduced internal flexibility. Hence, we aimed to make use of the TEV recognition site to cleave our protein of interest from the MBP tag after affinity chromatography. In a preliminary test, we used 40µg of purified fusion protein and 10 units of AcTEV™ (Invitrogen) and incubated them in 50mM Tris-HCl pH 8.0, 0.5mM EDTA and 1mM DTT at 4°C overnight. An identical mix without TEV protease served as negative control. On the next morning, aggregated protein was pelleted by centrifugation and resuspended in cleavage buffer. Subsequently, 9µl of total, soluble and insoluble fractions as well as 1µl of untreated fusion protein were separated on 17.5% SDS-PAGE and stained with Coomassie Blue (Figure 4-29). The experiment clearly showed that cleavage by TEV protease was efficient. However, whereas the sample lacking TEV protease remained stable and soluble, about one third of processed PV1 $VP1_8$-$2A^{pro}$ (C109A) was found to be precipitated after the reaction. Nevertheless, approximately two third of product still seemed to be soluble, providing a promising result.

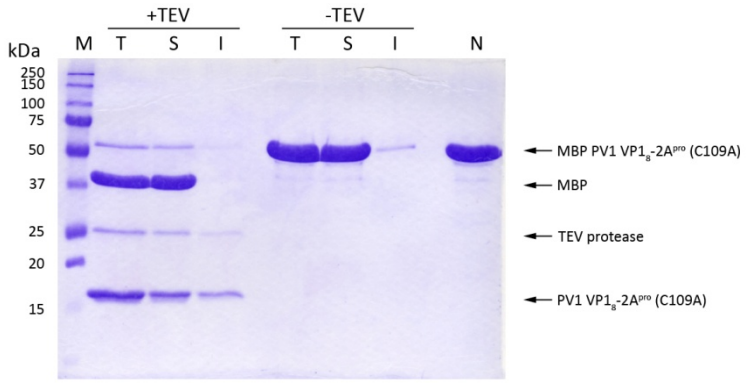


Figure 4-29 TEV protease cleavage tests on PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A). 40µg of fusion protein were incubated in the presence of 10U AcTEV™ (Invitrogen) at 4°C overnight. The reaction was buffered by TEV cleavage buffer (50mM Tris-HCl pH 8.0, 0.5mM EDTA and 1mM DTT). An identical mixture without protease served as negative control. Aliquots of total (T), soluble (S) and insoluble (I) fractions as well as an aliquot of non incubated (N) fusion protein were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. The identities of the four different species on the gel are given on the side.

We therefore expressed and purified TEV-protease in larger quantities (see chapter 3.10.1). MBP PV1 $VP1_8$-$2A^{pro}$ (C109A) was expressed as described above (see page 97). After cell lysis, the crude cell extract was cleared by centrifugation and the supernatant was loaded onto amylose resin, pre-equilibrated with buffer C/PV1c (50mM Tris-HCl pH 8.5, 200mM NaCl, 5mM DTT, 5% v/v glycerol). Extensive washing with buffer C/PV1c should remove all non bound impurities. Subsequently the resin was equilibrated with TEV cleavage buffer (50mM Tris-HCl pH 8.0, 0.5mM EDTA, 1mM DTT, 5% v/v glycerol, 100mM NaCl) and soaked with ~2mg TEV protease that had been diluted in TEV cleavage buffer. The resin was incubated at 4°C overnight to allow efficient cleavage of the fusion protein. Subsequently, processed protein was eluted from the column with the help of buffer C/PV1c, followed by elution with buffer C/PV1c including 10mM maltose to remove the still bound MBP and uncleaved fusion protein (Figure 4-30).

Fractions with V_r of 0-77ml contained processed PV1 $VP1_8$-$2A^{pro}$ (C109A). The presence of fusion protein as well as MBP in these fractions indicated a slightly leaking resin. Moreover, the fractions received by elution with 10mM maltose contained a substantial amount of uncleaved fusion protein, illustrating incomplete processing.

Fractions according to V_r=0ml to V_r=77ml were pooled and concentrated for further purification via size exclusion chromatography. However, heavy precipitation occurred during removal of the first third of buffer, reflecting the insolubility of PV1 $2A^{pro}$ lacking a solubility tag, an observation that had been made previously. Precipitate was removed by centrifugation and the supernatant was loaded onto a HiLoad Superdex 75 prep grade 26/60. As expected, PV1 $VP1_8$-$2A^{pro}$ (C109A) could not be detected in the eluate any more (data not shown). For this reason the strategy was abandoned.

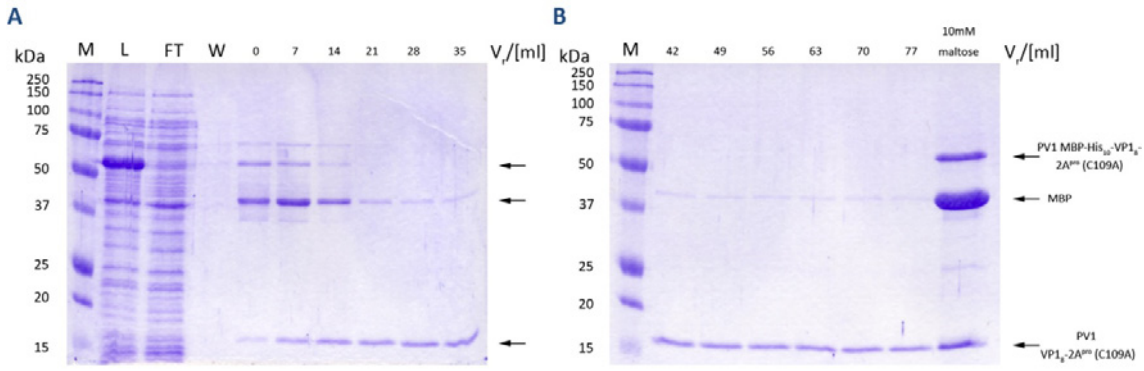


Figure 4-30 Purification via amylose resin and TEV cleavage of PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A). Fusion protein was bound to amylose resin and processed at the TEV recognition site on the column. Aliquots of load (L), flow-through (FT) and wash (W) as well as aliquots of elution fractions (with or without maltose) were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. The three main species of the sample (PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A), MBP, PV1 $VP1_8$-$2A^{pro}$ (C109A)) are marked with arrows.

4.1.1.4 PV1d: Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) short linker

In order to increase the likelihood of crystallization while maintaining the MBP fusion tag to keep PV1 $2A^{pro}$ in solution, we designed a construct in which the long linking region between MBP and PV1 $VP1_8$-$2A^{pro}$ (C109A) was truncated to the minimal feasible length.

The PCR primers TIM 1788 (binding upstream of a *BsiW*I site within the MBP coding region) and TIM 1789 (binding at the C-terminal end of the MBP coding region and containing an *Nhe*I site in its non binding 5' overhang) (Table 3-1) were used to amplify a 1164bp fragment from Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) (see chapter 4.1.1.3). The fragment was digested with *BsiW*I and *Nhe*I and ligated into *BsiW*I/*Nhe*I digested Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A); this eliminated the region coding for the linker, the His_{10}-tag and the TEV cleavage site. The new construct (Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) short linker) consisted of the maltose binding protein fused to $VP1_8$-$2A^{pro}$ (C109A) by a stretch of two amino acids (alanine and serine) (APPENDIX IV).

One possible problem of such a variant is that $2A^{pro}$ could clash with MBP when folding back on its own N-terminus if the linker between MBP and $2A^{pro}$ would be too short. This would preclude monitoring of self-processing. In order to rule out such a possibility, we also created an active $2A^{pro}$ version of this construct to assay the general activity of $2A^{pro}$ in such a context (see chapter 4.2.1.2).

Expression of the recombinant proteins followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3); **Medium:** 2 liter LB-Kan; **[IPTG]:** 0.5mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 0µM. Cells were harvested by centrifugation (chapter 3.6.2) and resuspended in 60ml of Buffer A/PV1d (50mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol, 5mM DTT). Subsequently, bacteria was broken by sonication (chapter 3.6.2.1) and 4µl of the total, soluble and insoluble fractions were separated on 12.5% (inactive construct) or 17.5% (active construct) SDS-PAGE (Figure 4-31).

Both constructs yielded a satisfactory amount of overexpressed protein. However, recombinant protein could also be detected in the uninduced control, indicating a leaky expression in the *E. coli* BL21 (DE3) cells. The solubility of PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker was satisfactory with about one third to one half of the fusion protein found in the soluble fraction (Figure 4-31 A). The equivalent active construct showed that $2A^{pro}$ was generally active in such a construct. The fusion protein migrating at around 50 kDa in the SDS-PAGE could not be detected. Instead, additional bands could be observed at about 37 kDa and 15 kDa, representing

processed MBP-VP1$_8$ and 2A^{pro}. However, this experiment again illustrated the insolubility of PV1 2A^{pro} as it could only be detected in the insoluble fraction (Figure 4-31 B).

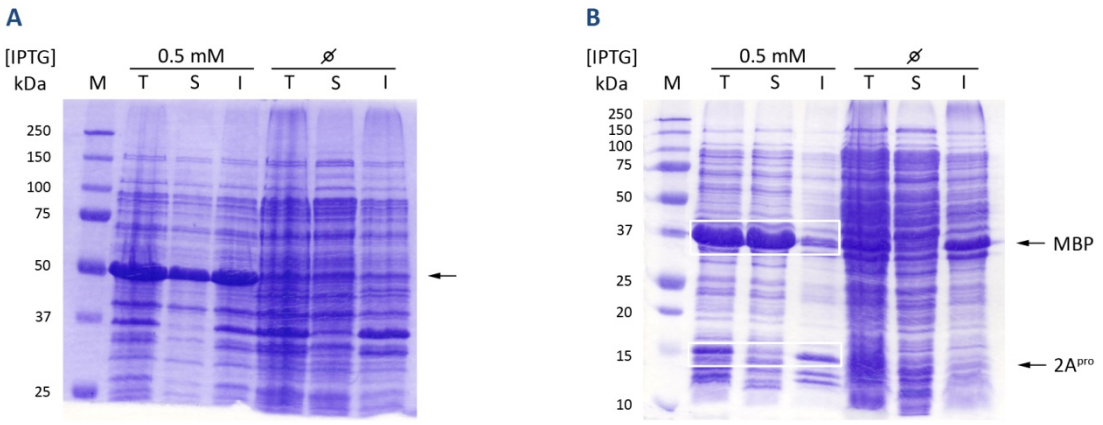


Figure 4-31 Expression of inactive and active PV1 MBP-VP1$_8$-2A^{pro} short linker constructs in *E. coli* BL21 (DE3). (A) 4µL of total (T), soluble (S) and insoluble (I) fractions of induced and uninduced samples of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow. **(B)** 4µL of total (T), soluble (S) and insoluble (I) fractions of induced and uninduced samples of PV1 MBP-VP1$_8$-2A^{pro} short linker were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression and processing of recombinant protein are marked with an arrow (MBP running at approximately 37kDa, VP1$_8$-2A^{pro} running at approximately 15kDa).

The soluble fraction of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker was loaded on 10ml of amylose resin preequilibrated with 10 column volumes of Buffer A/PV1d (50mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol, 5mM DTT). Unbound sample was washed off with 10 column volumes of Buffer A/PV1d and the fusion protein was eluted using Buffer A/PV1d including 10mM maltose. Eluted protein was collected in fractions of 10ml each and analyzed on 12.5% SDS-PAGE as were fractions of load, flow-through and wash (Figure 4-32).

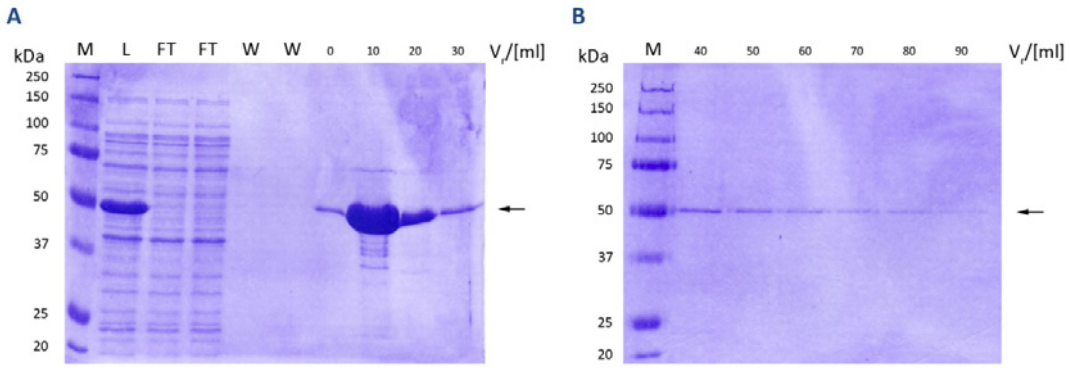


Figure 4-32 Analysis of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker containing fractions from affinity chromatography via amylose resin by 12.5% SDS-PAGE. 5µl of load (L), flow-through (FT) and wash fractions (W) as well as 5µl of elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Recombinant fusion protein is marked with arrows.

Analysis of the fractions eluted from the amylose resin showed signs of what was first considered to be degradation (Figure 4-32 A). Distinct bands of smaller molecular weight than the fusion protein were clearly visible in the fraction with V_r=10-20ml. We then examined how the sample would behave on a size exclusion column. Hence, we pooled fractions containing recombinant protein, concentrated the pooled sample to 2.5ml and loaded it onto a HiLoad Superdex 75 prep grade 26/60 preequilibrated with Buffer B/PV1d (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). The chromatogram and subsequent analysis of the fractions (2ml each) on 12.5% SDS-PAGE are shown in Figure 4-33.

The resolution power of the HiLoad Superdex 75 prep grade 26/60 was not high enough to separate full length recombinant protein from the impurities of smaller molecular weight. However, the fact that the protein seemed to be relatively stable and showed no further degradation let us speculate whether the impurities were the result of premature translation stops rather than degradation of the protein. Furthermore, we speculated that the massive over-expression of the zinc binding 2A^{pro} could lead to Zn^{2+} depletion within the cells, causing problems in the proper folding of 2A^{pro}. Addition of 100µM $ZnCl_2$ to the growth medium should therefore be tested.

In order to rule out growth defects of *E. coli* due to possible toxicity of $ZnCl_2$, *E. coli* BL21 (DE3) cells were grown in LB medium containing 0µM, 10µM, 50µM, 100µM or 500µM $ZnCl_2$. Bacterial growth was monitored by regular measurement of OD_{600} (Figure 4-34). The $ZnCl_2$ concentrations tested did not appear to have an effect on bacterial growth as even bacteria in 500µM $ZnCl_2$ grew equally well as in 0µM $ZnCl_2$. Consequently, $ZnCl_2$ was added to an end concentration of 100µM at the time of induction in the following expressions.

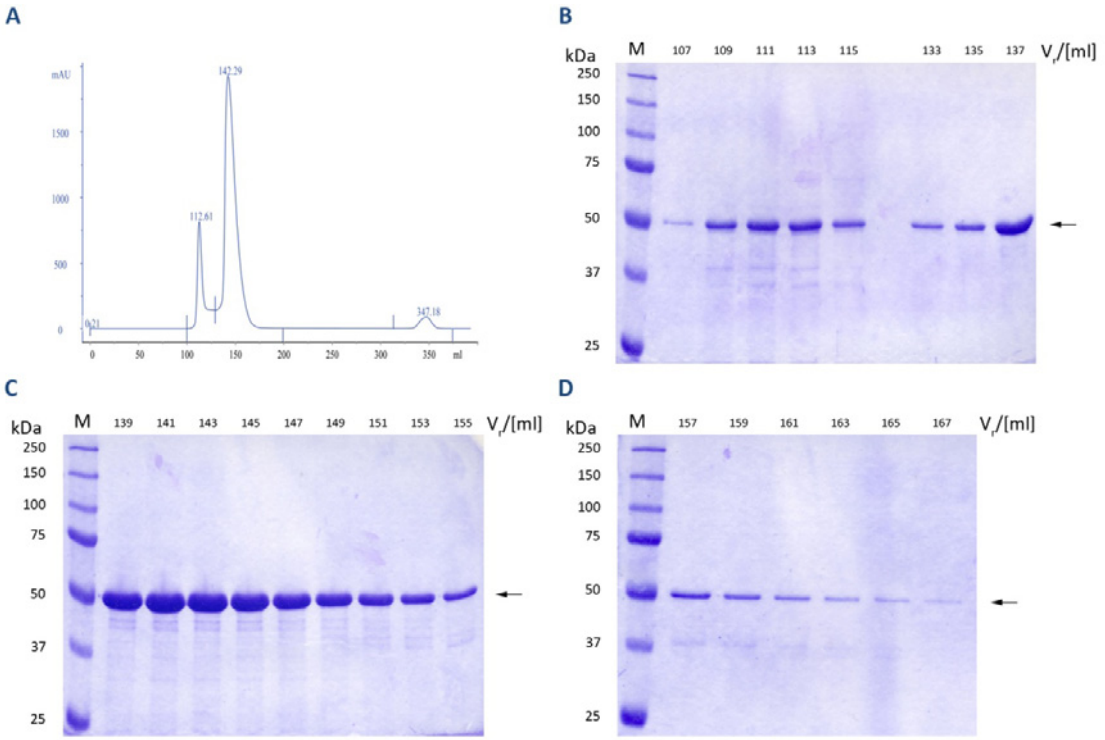


Figure 4-33 Size exclusion profile and analysis of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker containing fractions by 12.5% SDS-PAGE. (A) Chromatogram including retention volumes of the peaks of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker separated on a HiLoad Superdex 75 prep grade 26/60 developed with Buffer B/PV1d (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). **(B)** 3µl of fractions were loaded and separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

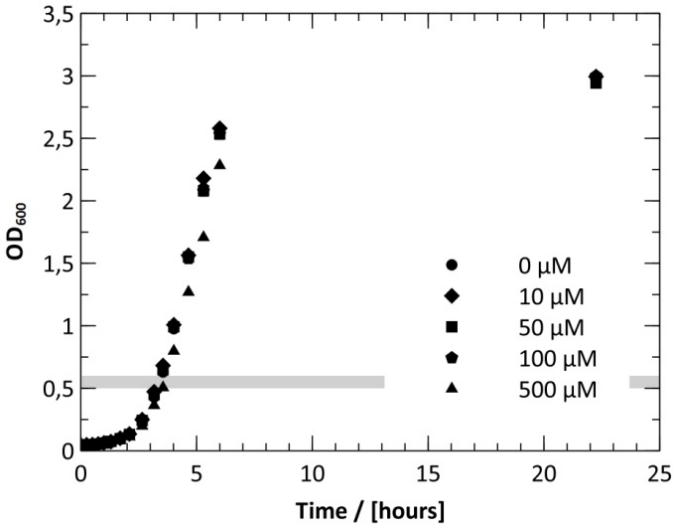


Figure 4-34 Monitoring bacterial growth rate in LB-medium including different concentrations of $ZnCl_2$. 100ml of LB-medium including 0µM, 10µM, 50µM, 100µM or 500µM $ZnCl_2$ were inoculated with 1ml of overnight culture of *E. coli* BL21 (DE3) in LB-medium without $ZnCl_2$. Cells were incubated at 37°C/200rpm and bacterial growth was monitored by regular measurement of optical density at 600nm. The grey bar highlights the area of OD_{600}=0.5-0.6, the typical range used throughout the entire work for induction of expression by the addition of IPTG.

Furthermore, we analyzed codon usage of our VP1$_8$-2A^{pro} gene to check for codons rarely used by *E. coli*. The software RaCC (rare codon calculator) provided by the NIH MBI laboratory for structural genomics and proteomics (http://nihserver.mbi.ucla.edu/RACC/) was used to screen for six codons rarely used by *E. coli*. These are AGG, AGA and CGA coding for arginine; CTA coding for leucine; ATA coding for isoleucine and CCC coding for proline (Figure 4-35). The software reported eight rare arginine codons (all of them being AGA or AGG, the two rarest arginine codons) and two rare isoleucine codons. Given the very strong over-expression of our construct, we suspected that the presence of the rare codons could be the reason for premature translation stops. Hence, we used a codon optimized *E. coli* strain in the subsequent expressions (*E. coli* BL21 (DE3) Rosetta2 pLysS (see chapter 3.4.1.2)).

Figure 4-35 Analyzing the PV1 VP1$_8$-2A^{pro} (C109A) cDNA for rare codon usage. The cDNA sequence coding for PV1 VP1$_8$-2A^{pro} (C109A) was fed into the rare codon calculator (RaCC) of the NIH MBI laboratory for structural genomics and proteomics (http://nihserver.mbi.ucla.edu/RACC/). Codons rarely used by *E. coli* are highlighted in color. Red: rare arginine codons (AGA, AGG). Blue: rare isoleucine codon (ATA).

A fresh expression of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker was performed as described above with the following modifications. The *E. coli* strain BL21 (DE3) Rosetta2 pLysS was used to overcome possible problems in expression caused by rare codon usage and 100µM $ZnCl_2$ were added to the bacterial culture at the time of induction by IPTG. Recombinant protein was purified as described above (data not shown) and PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker containing fractions from the gel filtration were pooled and concentrated. The sample was analyzed by separating different amounts (0.5µl, 1.0µl and 1.5µl) on a 12.5% SDS-PAGE (Figure 4-36 A). Despite the usage of a codon optimized *E. coli* strain as well as the use of $ZnCl_2$ in the expression, the impurities co-purifying with PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker could not be avoided. In a final attempt to remove the impurities we reapplied the sample onto a HiLoad

Superdex 200 prep grade 26/60 column and developed it with Buffer B/PV1d (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). However, the change in pore size of the gel filtration material was not sufficient to obtain a pure sample (Figure 4-36 B) for crystallization

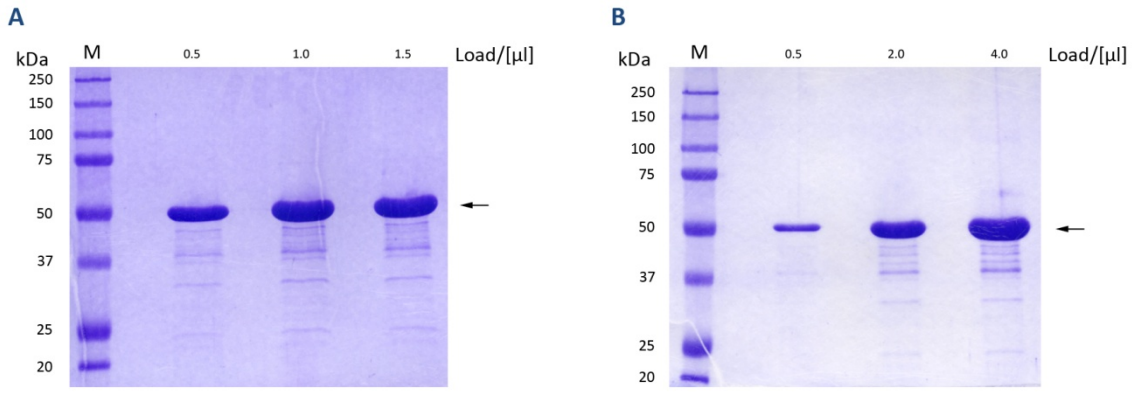


Figure 4-36 Final, purified sample of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker. Different amounts of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker were run on 12.5% SDS-PAGE and stained with Coomassie Blue. **(A)** Protein after purification by amylose resin and gel filtration on a HiLoad Superdex 75 prep grade 26/60. **(B)** Protein after additional run on a HiLoad Superdex 200 prep grade 26/60.

4.1.1.5 PV1e: Mal pET PV1 VP1$_8$-2A^{pro} (C109A) short linker His$_6$

As we suspected that the impurities of our PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker samples originate of premature translation, we intended to include a C-terminal hexa-histidine tag which can be used for affinity chromatography to pull out full length recombinant protein. Additionally, we included a glycine-proline stretch between the C-terminus of 2A^{pro} and the His$_6$-tag which together with the C-terminus of 2A^{pro} should act as a PreScission protease recognition site. PreScission protease is the 3C protease of HRV14, naturally processing the 2A^{pro}-2B site in the viral polyprotein. We therefore rationalized that we could obtain a native C-terminus of 2A^{pro} by removing the His$_6$-tag with PreScission protease.

PCR primers TIM1779 (fwd, binding at the end of the MBP coding sequence) and TIM1833 (rev, binding the end of 2A^{pro} and coding for a GlyProHis$_6$StopStop stretch as well as a *Bam*HI restriction site in its 5' non binding overhang) (Table 3-1) were used to amplify a 547bp fragment from Mal pET PV1 VP1$_8$-2A^{pro} (C109A) short linker. The fragment was digested with *Pst*I and *Bam*HI and ligated into a *Pst*I/*Bam*HI digested Mal pET VP1$_8$-2A^{pro} (C109A) short linker. The resulting fusion protein differs from PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker by the addition of a GlyProHis$_6$ stretch at the C-terminus of 2A^{pro} (APPENDIX V).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3) Rosetta2 pLysS; **Medium:** 2 liter LB-Gluc-Kan; **[IPTG]:** 0.1mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 100µM. The next morning, cells were harvested by centrifuging the cultures at 4000rpm/4°C/20min. Subsequently, cells were resuspended in 60ml Buffer A/PV1e (50mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol, 5mM DTT) and broken by sonication (see chapter 3.6.2.1). 2.5µl aliquots of total, soluble and insoluble fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue (Figure 4-37).

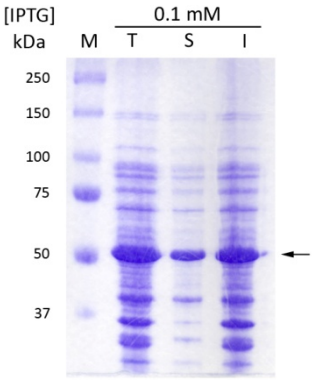


Figure 4-37 Expression of PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker-His_6 constructs in *E. coli* BL21 (DE3) Rosetta2 pLysS. 2.5µL of total (T), soluble (S) and insoluble (I) fractions of the induced samples of PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker-His_6 were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

The recombinant protein was over-expressed well. The generally large amount of protein in the insoluble fraction was taken to indicate insufficient lysis of the cells rather than insolubility of these proteins. Nevertheless, the soluble fraction was purified over 10ml of amylose resin that has been preequilibrated with Buffer A/PV1e (50mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol, 5mM DTT). Bound protein was eluted with Buffer A/PV1e including 10mM maltose and fractions of 10ml each were collected. 2µl of the fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue to monitor the progress of purification (Figure 4-38).

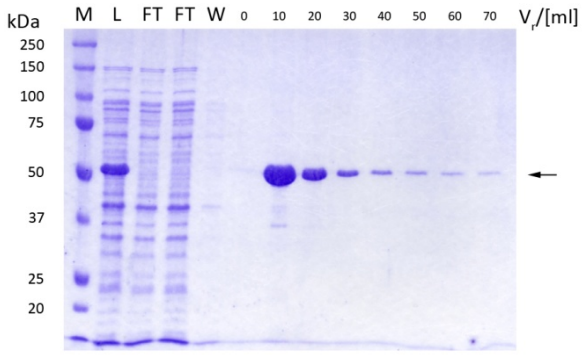


Figure 4-38 Analysis of PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker-His_6 containing fractions from affinity chromatography via amylose resin by 12.5% SDS-PAGE. 2µl of load (L), flow-through (FT) and wash fractions (W) as well as 2µl of elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Recombinant fusion protein is marked with arrows.

The protein purified over amylose resin still contained the familiar impurities. For this reason, we pooled the fractions with V_r=10-40ml, dialyzed against three changes of Buffer B/PV1e (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol), and loaded the sample onto a Ni^{2+} charged 5ml HiTrap™ Chelating HP that had been equilibrated with Buffer B/PV1e. Unbound sample was washed off with several column volumes of Buffer B/PV1e. Subsequently, 5mg of PreScission protease (see chapter 3.10.2) were diluted in 7ml of Buffer BPV1e and loaded onto the column. The column was incubated at 4°C overnight to allow sufficient cleavage by PreScission protease at the C-terminus of $2A^{pro}$. Fusion protein that had been cleaved from its C-terminal His_6-tag was washed off the column using Buffer B/PV1e. Finally, the column was developed with buffer B/PV1e including 500mM imidazole to elute protein still bound to the beads. All fractions were analyzed on 12.5% SDS-PAGE (Figure 4-39).

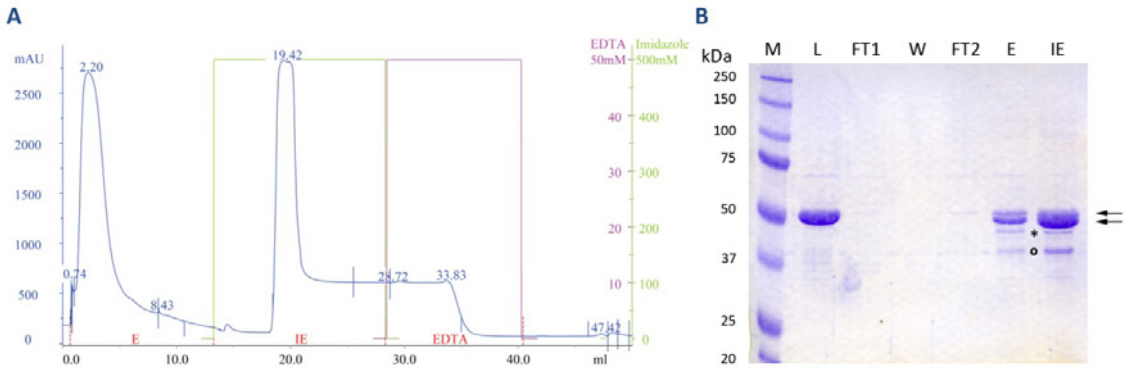


Figure 4-39 HiTrap™ Chelating HP capture and on column cleavage with PreScission protease of PV1 MBP-VP1$_8$-2A^{pro} (C109A) short linker-His$_6$. Protein was loaded on a 5ml HiTrap™ Chelating HP column that had been equilibrated with Buffer B/PV1e (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). Unbound sample was washed off with Buffer B/PV1e before 5mg of PreScission protease (in 7ml Buffer B/PV1e) were loaded onto the column and incubated at 4°C overnight. Cleaved material was eluted from the column by washing with Buffer B/PV1e followed by elution with Buffer B/PV1e including 500mM imidazole and stripping of the column with Buffer B/PV1e including 50mM EDTA. **(A)** Chromatogram of elution (E), imidazole elution (IE) and stripping (EDTA). **(B)** 2µl of load (L), flow-through of loading (FT1), wash (W), flow-through at PreScission protease loading (FT2), elution (E) and imidazole elution (IE) were separated on a 12.5% SDS-PAGE and stained with Coomassie Blue. Cleaved and uncleaved recombinant proteins are marked with arrows. PreScission protease bands are indicated with a circle. A band possibly arising from aberrant cleavage is highlighted with an asterisk.

The SDS-PAGE clearly shows that cleavage of the His-tag was incomplete. Moreover, cleaved as well as uncleaved recombinant protein co-purified in the fraction eluted with Buffer B/PV1e as well as in the fraction eluted with Buffer B/PV1e including 500mM imidazole. Furthermore, a weak band migrating slightly faster than the cleaved recombinant protein (highlighted with an asterisk in Figure 4-39 B) indicated a low rate of unspecific cleavage, resulting in an aberrant cleavage product. For all these reasons, we decided to express a fresh batch of

protein which should be purified via amylose resin, HiTrap™ Chelating HP and gel filtration to examine whether crystals of this construct could be obtained without removing the C-terminal His_6-tag.

Expression and purification via amylose resin was performed as described above. After dialysis against Buffer C/PV1e (buffer A/PV1e lacking DTT), the sample was loaded onto a 5ml HiTrap™ Chelating HP column followed by extensive washing with Buffer C/PV1e containing 10mM imidazole. Recombinant protein was eluted with a step gradient using Buffer C/PV1e including 500mM imidazole (data not shown). Protein containing fractions were pooled and concentrated to a volume of 3ml, and loaded onto a HiLoad Superdex 200 prep grade 26/60 that had been equilibrated with Buffer B/PV1e (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). The column was developed with Buffer B/PV1e and protein containing fractions were pooled and concentrated. Figure 4-40 shows the chromatogram of the size exclusion chromatography as well as final protein in two different concentrations separated on a 12.5% SDS-PAGE. The protein eluted at a V_r of 187ml, with the peak tailing towards higher retention volumes. The pooled and concentrated samples are of high purity and no bands of lower molecular weight could be detected in the Coomassie stained SDS-PAGE.

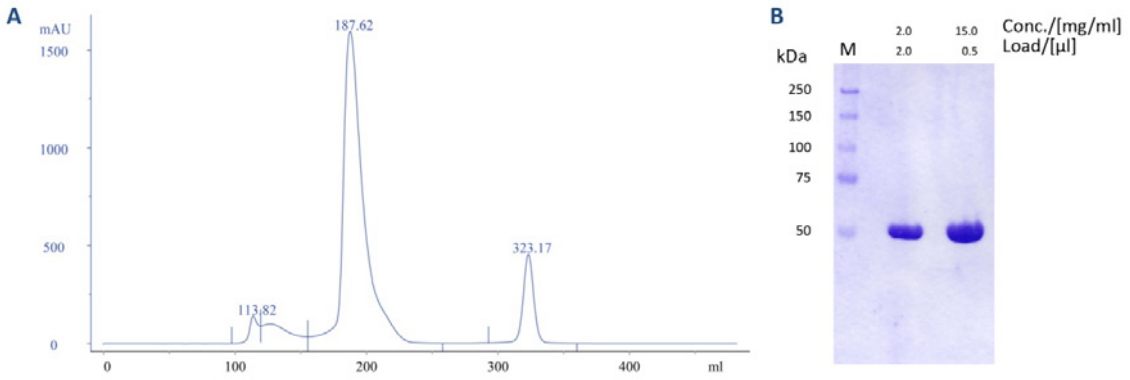


Figure 4-40 Size exclusion profile and examination of the final protein sample for its purity by SDS-PAGE. (A) Size exclusion profile of PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker-His_6 separated on a HiLoad Superdex 200 prep grade 26/60 developed with Buffer B/PV1e (10mM Tris-HCl pH 8.5, 200mM NaCl, 5% v/v glycerol). **(B)** 2µl of final sample at 2mg/ml and 0.5µl of final sample at 15mg/ml were separated on 12,5% SDS-PAGE and stained with Coomassie Blue.

Both sample concentrations were used to perform crystal screening using the commercial screens given in APPENDIX XIV. A nanoliter crystallization robot (Phoenix) was used to set up 96well screens with sitting drops (see chapter 3.7.1). Despite the large set of screens, no crystal hits could be obtained.

4.1.1.6 Summary

A lot of effort was put into the expression, purification and crystallization of PV 2A^{pro}. The initial problems with protein solubility were tackled by using MBP as a solubility tag. Fusion proteins exhibited strongly increased solubility. However, protein crystals from MBP-2A^{pro} fusion proteins could not be obtained or were not reproducible. On the contrary, 2A^{pro} that was cleaved from its solubility tag was insoluble just like the protein variants expressed without tag before. For these reasons we decided to stop experiments on PV 2A^{pro}.

4.1.2 Human Rhinovirus 2 2A^{pro} constructs

As our attempts to crystallize a PV 2A^{pro} variant failed, we wanted to examine 2A proteases from different enteroviruses. The HRV2 2A^{pro} was the first one to be crystallized and its structure was solved 1.95Å (see chapter 2.4.2.1). In doing so, active HRV2 2A^{pro} behaved well in terms of solubility and crystallization. Hence, we rationalized that an inactive HRV2 2A^{pro} variant with N-terminal extension might perform better than the comparable variants from PV 2A^{pro}.

4.1.2.1 HRV2a: pET MBP HRV2 VP1$_9$-2A^{pro} (C106A)

An *Nco*I/*Bam*HI digested PCR fragment amplified from the plasmid pCITE HRV2 VP1-2A^{pro} (C106A) (Glaser et al., 2003) with the oligonucleotides TIM 1821 and TIM 554 (Table 3-1) was ligated into the *Nco*I/*Bam*HI cut backbone of the expression plasmid pET MBP 1a (see chapter 3.1.1.3). This plasmid differs from the Mal pET vector by the presence of a His$_6$-tag at the N-terminus of MBP instead of a His$_{10}$-tag between MBP and TEV cleavage site. Regardless, VP1$_9$-2A^{pro} can be cleaved from the MBP tag by the use of TEV protease leaving a glycine-alanine stretch fused to the N-terminus of HRV2 VP1$_9$-2A^{pro} (APPENDIX VI).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3) Rosetta2 pLysS; **Medium:** 1 liter LB-Gluc-Kan; **[IPTG]:** 0.2mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[ZnCl$_2$]:** 100µM. The bacterial pellet was resuspended in 60ml Buffer A/HRV2a (50mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol, 5mM DTT) and cells were broken by sonication (see chapter

3.6.2.1). 3µl of the total, soluble and insoluble fractions were separated on 12.5% SDS-PAGE and analyzed by Coomassie staining (Figure 4-41).

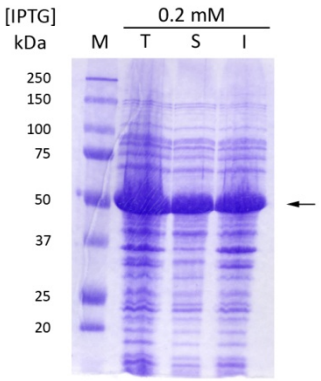


Figure 4-41 Expression of HRV2 MBP-VP1$_9$-2A^{pro} (C106A) in *E. coli* BL21 (DE3) Rosetta2 pLysS. 3µL of total (T), soluble (S) and insoluble (I) fractions of the induced samples of HRV2 MBP-VP1$_9$-2A^{pro} (C106A) were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

The over-expression and solubility of the HRV2 construct was sufficient. Hence, we continued with purification via amylose resin. The supernatant of the crude cell extract was loaded onto a 12ml bed of amylose resin that had been equilibrated with Buffer A/HRV2a. Unbound sample was removed by extensive washing with Buffer A/HRV2a before the column was equilibrated with TEV-protease cleavage buffer (50mM Tris-HCl pH 8.0, 1mM EDTA, 5mM DTT) for on-column TEV-protease cleavage. 3mg of TEV-protease in cleavage buffer were then loaded onto the column followed by incubation at 4°C overnight. The next morning, cleaved material was washed from the column using Buffer A/HRV2a (50mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol, 5mM DTT) followed by a wash with Buffer A/HRV2a including 10mM maltose in order to remove MBP and uncleaved fusion protein. Aliquots of the fractions were separated on 20% SDS-PAGE and stained with Coomassie Blue to monitor the purification (Figure 4-42).

Figure 4-42 shows that the efficiency of on column TEV cleavage was insufficient as roughly two third of the protein was not cleaved. Furthermore, uncleaved fusion protein as well as MBP co-eluted with HRV2 VP1$_9$-2A^{pro} when washing with Buffer A/HRV2a. The reason for this behavior is unclear. Unfortunately, these observations made on-column cleavage unattractive. Nevertheless, we concentrated the first elution fraction corresponding to V_r=0-15ml to about 4ml for further purification via size exclusion. Already in the early stage of the concentration process, precipitation of the protein could be detected. Hence, precipitated protein was removed by centrifugation and only the supernatant was loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/HRV2a (10mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol). Figure 4-43 shows the size exclusion profile as well as 20% SDS-PAGE analysis of the corresponding protein fractions.

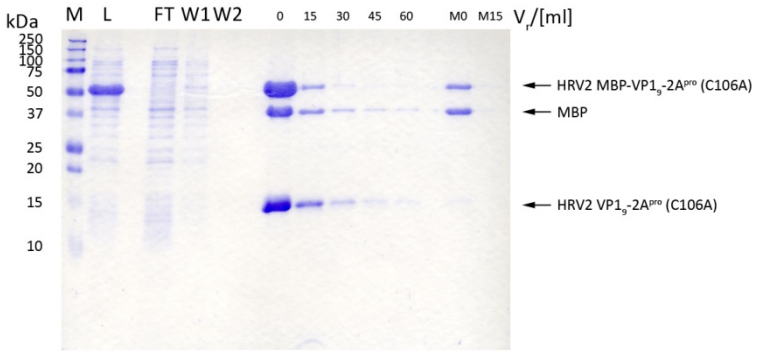


Figure 4-42 Monitoring HRV2 MBP-VP1$_9$-2A^{pro} (C106A) purification by affinity chromatography via amylose resin and cleavage of the fusion protein by TEV-protease. HRV2 MBP-VP1$_9$-2A^{pro} (C106A) containing supernatant from crude cell extract was loaded onto amylose resin that had been equilibrated with Buffer A/HRV2a (50mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol, 5mM DTT). After washing with Buffer A/HRV2a, the column was equilibrated with TEV-protease cleavage Buffer (50mM Tris-HCl pH 8.0, 1mM EDTA, 5mM DTT), 3mg of TEV-protease were added and incubated at 4°C overnight. On the following morning, cleaved protein was eluted with Buffer A/HRV2a followed by elution of uncleaved material by Buffer A/HRV2a including 10mM maltose. 1µl of load (L), 2µl of flow-through (FT), 2µl of wash (W1, W2), 4µl of elution fractions with buffer A/HRV2a (see retention volumes) and 1µl of maltose elution fractions (see "M" retention volumes) were separated on 20% SDS-PAGE and stained with Coomassie Blue. The identity of the relevant bands is indicated with arrows.

The size exclusion profile exhibited three peaks at V_r~110, 130 and 162ml. The corresponding lanes on the SDS-PAGE showed that separation of uncleaved fusion protein from MBP was rather successful. However, the presence of a band at around 15kDa over the whole elution range of proteins indicated that HRV2 VP1$_9$-2A^{pro} (C106A) is in a polydisperse, oligomeric state and tended to aggregate. As possible reasons for aggregation, we identified the lack of salt and glycerol in the TEV-cleavage buffer. For this reason, we decided to modify this step. A fresh batch of fusion protein was expressed as described above. Supernatant of the crude cell extract was then loaded onto amylose resin that had been equilibrated with Buffer A/HRV2a (50mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol, 5mM DTT). After removing unbound sample by extensive washing with Buffer A/HRV2a, fusion protein was eluted from the beads with Buffer A/HRV2a including 10mM maltose. EDTA was added to the eluate to a final concentration of 1mM and 3mg of TEV-protease were added. TEV-protease cleavage was performed in solution at 4°C overnight. Cleavage efficiency was monitored by separating 4µl of TEV-protease cleaved material as well as 4µl of uncleaved sample on a 20% SDS-PAGE (Figure 4-44). Nearly 100% of fusion protein was processed by TEV-protease to yield two fragments with sizes corresponding to MBP and HRV2 VP1$_8$-2A^{pro} (C106A). Due to this increase in TEV-protease cleavage efficiency, future corresponding experiments were always carried out in solution rather than on the column.

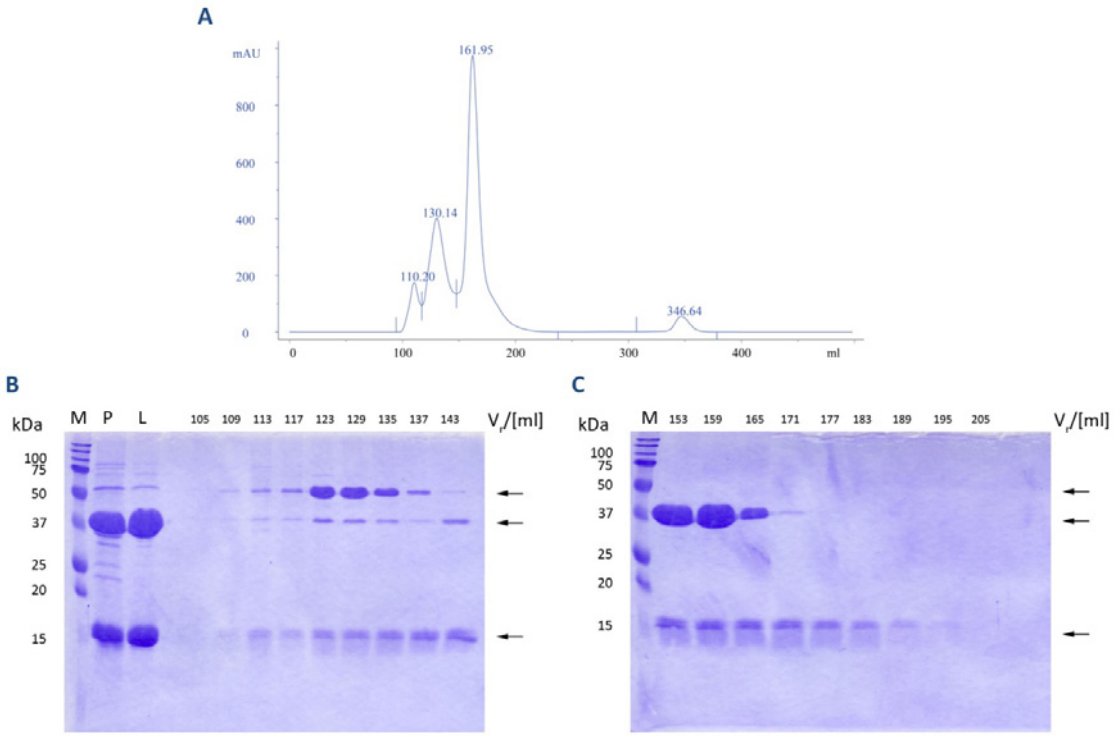


Figure 4-43 Monitoring size exclusion chromatography of TEV-protease cleaved HRV2 MBP-VP1$_9$-2A^{pro} (C106A). Protein sample was loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/HRV2a (10mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol) and the column was developed with the same buffer. **(A)** Size exclusion profile of the eluted proteins. **(B+C)** Separation of 4µl of precipitated, pelletized and resuspended protein (P), 1µl of load (L) and 4µl of corresponding fractions on 20% SDS-PAGE. Proteins were stained with Coomassie Blue and protein bands of interest were labeled with arrows.

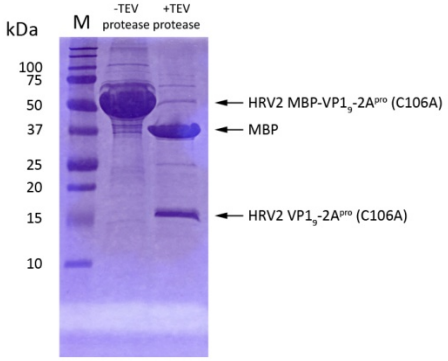


Figure 4-44 Cleavage of HRV2 MBP-VP1$_9$-2A^{pro} (C106A) with TEV-protease in solution. EDTA was added to the fusion protein to a final concentration of 1mM and 3mg of TEV protease were added. Cleavage was performed at 4°C overnight. 4µl of cleaved sample as well as of a negative control without TEV-protease were separated on 20% SDS-PAGE and stained with Coomassie Blue. Protein bands of interest are labeled.

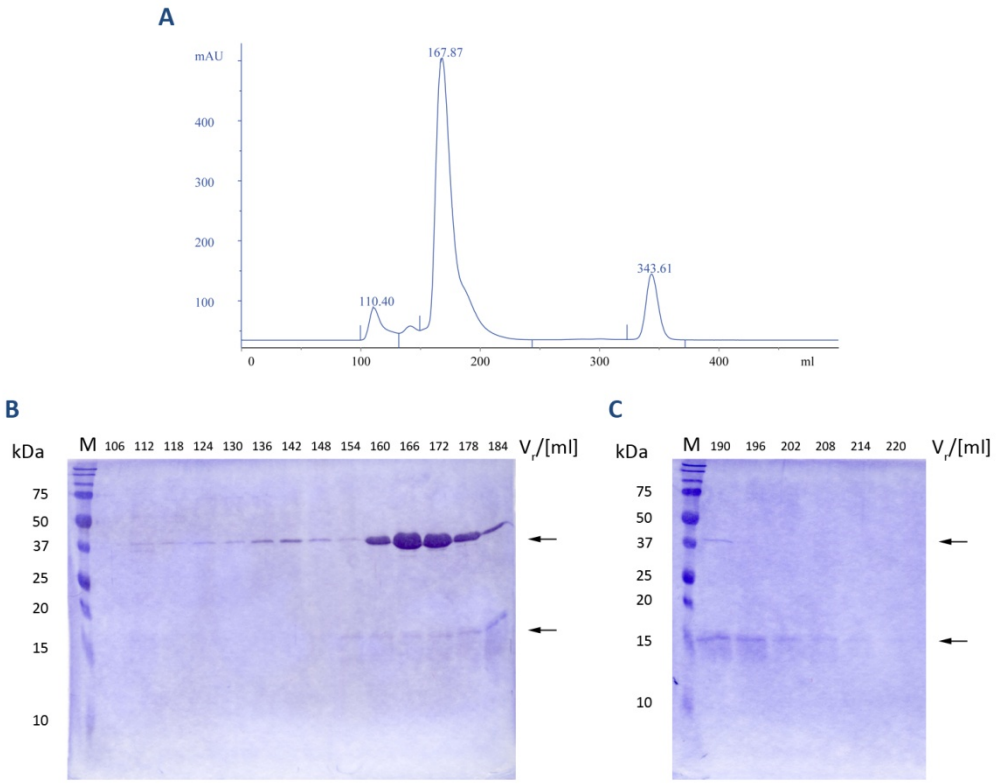


Figure 4-45 Monitoring size exclusion chromatography of TEV-protease cleaved HRV2 MBP-VP1$_9$-2A^{pro} (C106A) Pool 1. Protein sample was loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/HRV2a (10mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol) and the column was developed with the same buffer. **(A)** Size exclusion profile of the eluted proteins. **(B+C)** Separating 5µl of corresponding fractions on 20% SDS-PAGE. Proteins were stained with Coomassie Blue and protein bands of interest were labeled with arrows.

Given the observed tendency of HRV2 VP1$_9$-2A^{pro} (C106A) to aggregate, we split the sample into two pools. Pool 1 was used to run size exclusion chromatography on a HiLoad Superdex 75 prep grade 26/60 that has been equilibrated with Buffer B (10mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol) without prior concentration of the sample or removal of MBP (Figure 4-45). Pool 2 was extensively dialyzed against Buffer B to remove maltose, passed over amylose resin to remove MBP, concentrated to 3ml and finally purified over a HiLoad Superdex 75 prep grade 26/60 that has been equilibrated with Buffer B (Figure 4-46).

Both experiments confirmed the tendency of HRV2 VP1$_9$-2A^{pro} (C106A) to aggregate. Again the protein was detected over the whole range of retention volumes in which proteins were eluted from the column. However, Liebig *et al.* have demonstrated that active wt 2A^{pro} from HRV2 (without N-terminal extensions) does not show this behavior (Liebig *et al.*, 1993), eluting from a HiLoad Superdex 75 prep grade 26/60 column with an apparent molecular mass of 30kDa (compare also chapter 4.3.2.1). As the buffer conditions of Liebig *et al.* were very close to the ones used in our studies, we speculated that the N-terminal extension was the reason for the

aggregation of HRV2 $VP1_9$-$2A^{pro}$ (C106A), precluding further crystallization trials. Consequently, experiments on $2A^{pro}$ variants from HRV2 were discontinued.

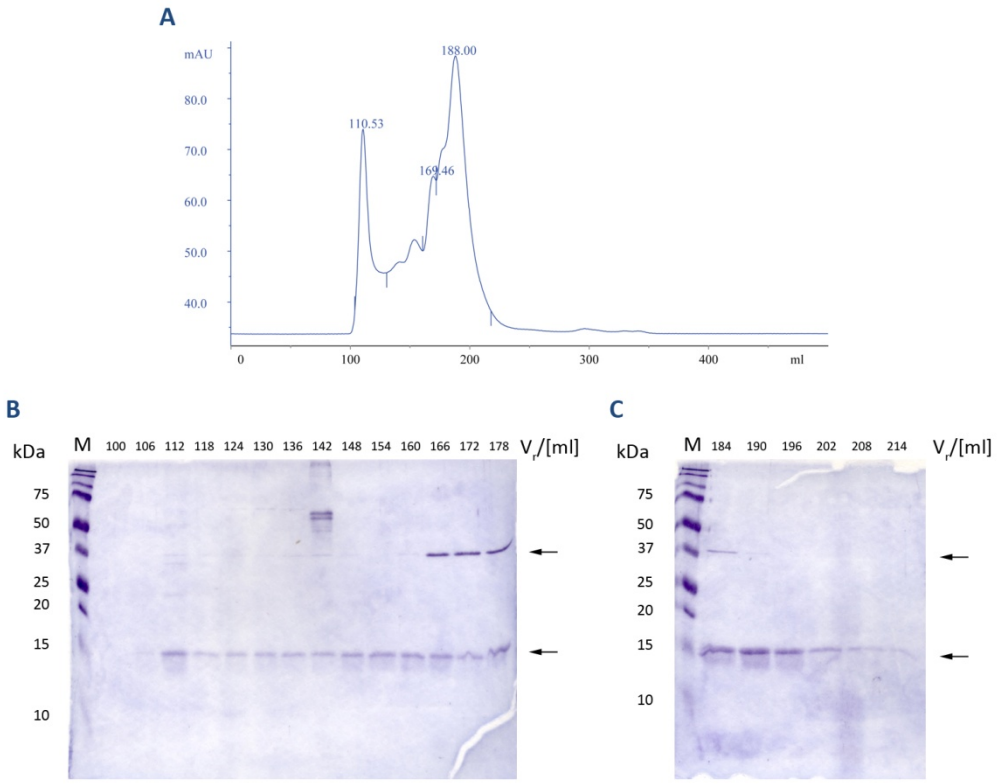


Figure 4-46 Monitoring size exclusion chromatography of TEV-protease cleaved HRV2 MBP-$VP1_9$-$2A^{pro}$ (C106A) Pool 2. Protein sample was loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/HRV2a (10mM Tris-HCl pH 8.1, 200mM NaCl, 5% v/v glycerol) and the column was developed with the same buffer. **(A)** Size exclusion profile of the eluted proteins. **(B+C)** Separating 8µl of corresponding fractions on 20% SDS-PAGE. Proteins were stained with Coomassie Blue and protein bands of interest were labeled with arrows.

4.1.3 Coxsackievirus B4 $2A^{pro}$ constructs

The CVB4 $2A^{pro}$ (including N-terminal extension) has successfully been expressed and purified. However, the conducted NMR experiments did not yield the desired information on self-processing (see chapter 2.4.2.2). For this reason, we aimed to make use of the satisfying solubility and expression levels of this protein variant in order to obtain samples suitable for crystallization.

4.1.3.1 CVB4a: pET 3d CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A)

All purification steps on ÄKTA FPLC were performed at room temperature.

The plasmid was a kind gift of Nicky Baxter (University of Sheffield) and has been described (Baxter *et al.*, 2006). The construct codes for a His_6-tag (with an additional alanine residue at the N-terminus) fused to the eight last residues of CVB4 VP1 and $2A^{pro}$ with the active site cysteine mutated to alanine (C110A).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3); **Medium:** 2 liter LB-Amp; **[IPTG]:** 1.0mM; **Expression temperature:** 25°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 0µM. The bacterial pellet was harvested by centrifugation, resuspended in 60ml Buffer A/CVB4a (50mM Tris-HCl pH 7.5, 50mM NaCl) and frozen at -80°C. After thawing, breakage of cells was completed by sonication and the crude cell extract was cleared by centrifugation. 9µl of total, soluble and insoluble fractions of induced as well as uninduced samples was separated on 17.5% SDS-PAGE and stained with Coomassie Blue (Figure 4-47).

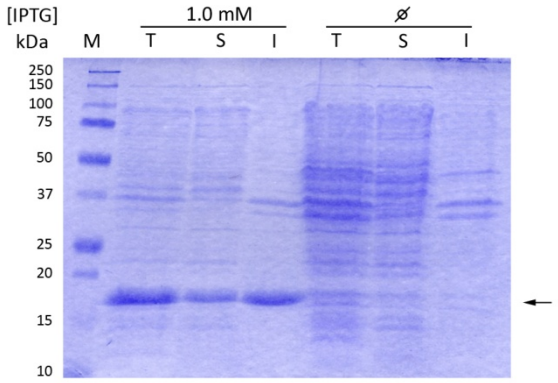


Figure 4-47 Expression of CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) in *E. coli* BL21 (DE3). 9µL of total (T), soluble (S) and insoluble (I) fractions of the induced samples as well as of the uninduced control were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

A satisfactory amount of recombinant protein was found in the supernatant which was subsequently subjected to fractionated ammonium sulfate precipitation using steps of 20% and 40% saturation of salt (see chapter 3.6.3). The pellet of the second step was redissolved in 16ml of Buffer B/CVB4a (50mM Tris-HCl pH 8.1, 50mM NaCl, 5mM DTT) and dialyzed against Buffer B/CVB4a to remove residual ammonium sulfate. Recombinant protein was further purified via ion exchange chromatography using a MonoQ 10/100 GL column. The sample was loaded onto the column that had been equilibrated with Buffer B/CVB4a and eluted with a gradient of rising NaCl

concentrations (up to 1M) in Buffer B/CVB4a. Figure 4-48 shows the chromatogram as well as 17.5% SDS-PAGE of the corresponding fractions.

CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) eluted at retention volumes of around 30-45ml corresponding to a NaCl concentration of 350-400mM. Fractions containing recombinant protein were pooled and concentrated to a volume of around 2ml in order to be loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/CVB4a (50mM Tris-HCl pH 8.1, 50mM NaCl, 5mM DTT). The column was developed with Buffer B/CVB4a and the collected fractions were inspected by loading 9µl aliquots on 17.5% SDS-PAGE (Figure 4-49 A and B). CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) eluted with a retention volume of around 201ml in a sharp, symmetric peak (Figure 4-49 C). Corresponding fractions were pooled, concentrated to around 12mg/ml and 0.5µl of sample were loaded onto a 17.5% SDS-PAGE and stained with Coomassie Blue to verify the purity (Figure 4-49 D).

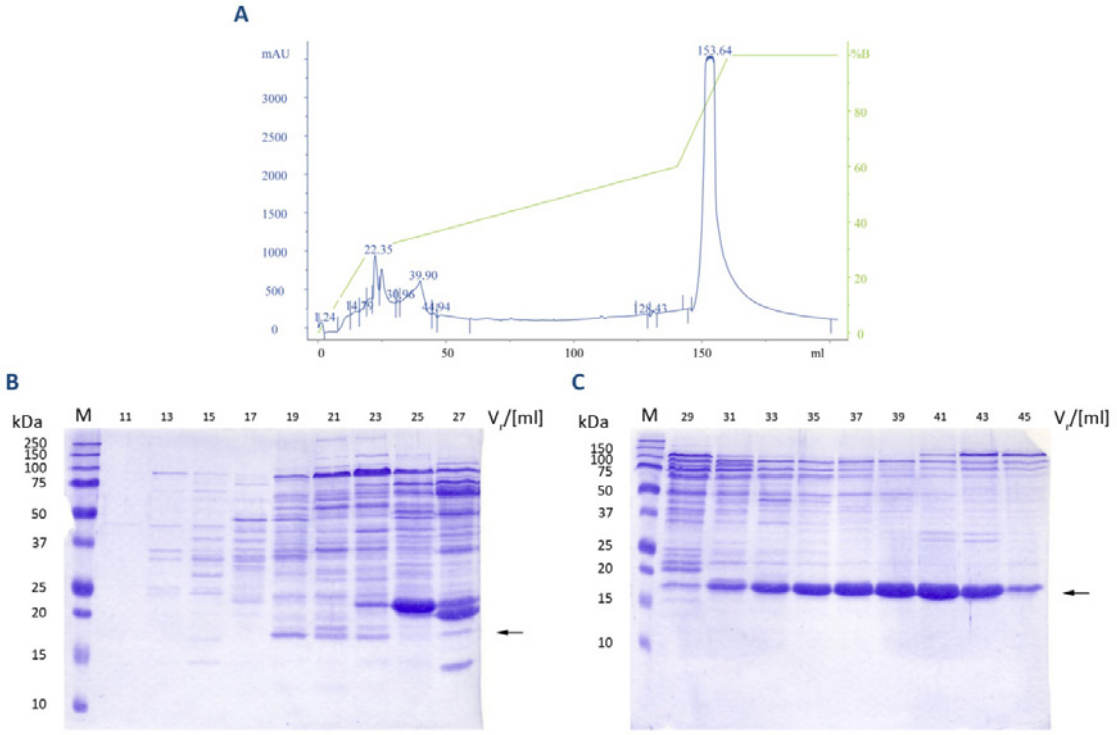


Figure 4-48 Chromatogram of CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) purification and analysis of CVB4 His6-$VP1_8$-$2A^{pro}$ (C110A) containing fractions by 17.5% SDS-PAGE. (A) UV-trace (blue) of proteins bound to a MonoQ 10/100 GL and eluted with a NaCl gradient (green). **(B+C)** 9µl of fractions were loaded and separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Arrows indicate the position of recombinant protein.

The expression and purification protocol was robust and reproducible. However, we could not observe the susceptibility to cold denaturation of CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) described by

Baxter *et al.* (Baxter *et al.*, 2006). Before heading to crystallization trials, we aimed to check the stability of CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A). The melting points of the protein in the chosen buffer as well as in a set of 192 different salt and buffer conditions (Table 3-3 and Table 3-4) were detected by thermofluor measurements and compared to each other. Figure 4-50 A shows the melting curve of CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) in Buffer B/CVB4a (red curve) as well as the quartile of melting curves with the highest values found within the 192 conditions. The melting point in Buffer B/CVB4a was measured to be 56°C. Figure 4-50 B gives an overview of the distribution of the measured melting points (red line indicates melting point in Buffer B/CVB4a). The melting point in Buffer B/CVB4a was within the upper quartile of detected melting points. Inspection of the different buffer compositions revealed that the conditions with the highest melting temperatures were found at pH 7.5-8.5, held a tendency towards 50-100mM NaCl and sometimes contained up to 10% v/v glycerol. For these reasons, we considered the buffer in use (50mM Tris-HCl pH 8.1, 50mM NaCl, 5mM DTT) as appropriate in terms of protein stability.

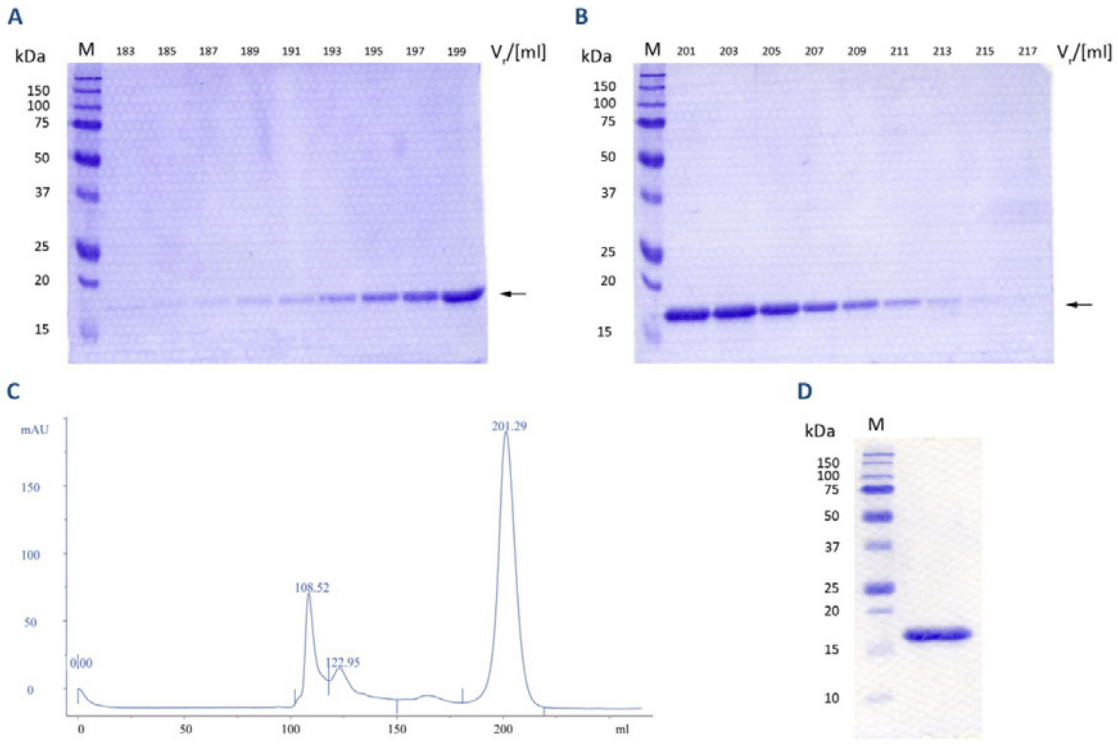


Figure 4-49 Final steps in CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) purification. (A+B) Analysis of protein fractions from size exclusion chromatography using a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/CVB4a. 9µl of fractions were separated on 17.5% SDS-PAGE and stained with Coomassie Blue. Protein bands arising from recombinant protein are labeled with arrows. **(C)** Size exclusion profile showing the UV-trace of eluted proteins. **(D)** Protein quality check of final purified CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) on 17.5% SDS-PAGE.

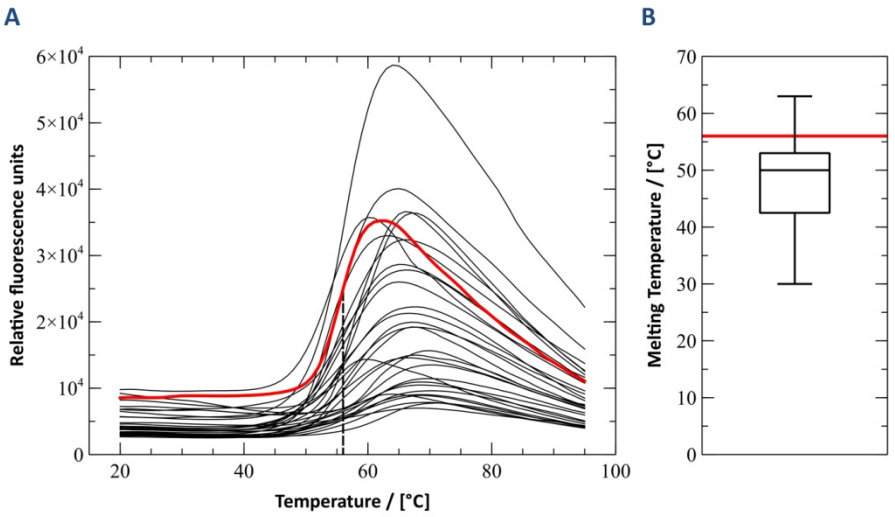


Figure 4-50 Reviewing CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) stability in buffer B/CVB4a and other conditions by thermofluor measurements. (A) Melting curves of CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) in Buffer B/CVB4a as well as in the 192 different buffer conditions were collected. The quartile of curves with the highest melting temperatures is depicted in black whereas the melting curve for CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) in Buffer B/CVB4a is given in red. **(B)** The distribution of melting points of all 192 conditions is depicted as a box-plot with the whiskers representing the minimum and maximum of the data. The melting point in Buffer B/CVB4a is depicted as red line.

Crystal screens were set up using the commercial screens given in APPENDIX XIV (see chapter 3.7.1). Unfortunately, no crystals could be detected during the regular observation of the drops.

4.1.4 Coxsackie B4 $2A^{pro}$ constructs with varying VP1 extension length

As the 2A protease with 8 amino acids of VP1 from CVB4 appeared to be the best construct tested so far in terms of solubility and stability, we decided to perform further experiments with this protease. Hence, different lengths of the VP1 extension were to be tested. Furthermore, the N-terminal His_6-tag was to be eliminated. As MBP fusions previously proved to have very satisfactory expression levels, we decided to again fuse CVB4 2A proteases with different VP1 extensions to the *E. coli* maltose binding protein. Furthermore, the TEV cleavage site would allow us to remove MBP during the purification process. Four constructs were created using either the pET MBP 1a or the Mal pET vector. We chose 5, 8, 12 and 20 amino acids as the length of extensions. However, it should be mentioned that all of the constructs held additional three or four amino acids (remnants of the TEV-cleavage site as well as the cloning site) at their N-

terminus. The variation of the length in VP1 extension should hopefully facilitate crystallization of the protein.

4.1.4.1 CVB4b: pET MBP CVB4 $VP1_5$-$2A^{pro}$ (C110A)

A PCR product coding for CVB4 $VP1_5$-$2A^{pro}$ (C110A) was amplified from pBS CVB4 VP1-$2A^{pro}$ (a kind gift from Andreas Rötzer, Max F. Perutz laboratories). Primer TIM 1882 (that anneals to cDNA coding for the C-terminus of VP1 and contains a *Bsp*HI restriction site in its 5' non-binding part) and TIM 1885 (that binds to cDNA coding for the C-terminus of $2A^{pro}$ and contains two stop codons as well as a *Bam*HI restrictions site in its 5' non binding part) (Table 3-1) were used to amplify the product in a two step PCR mutagenesis (see chapter 3.3.7.2) using DNA oligos TIM 1085 and TIM 1086 to introduce the C110A mutation. The PCR product was treated with *Bsp*HI and *Bam*HI and ligated into an *Nco*I and *Bam*HI processed pET MBP vector. The fusion protein expressed from this construct consists of an N-terminal His_6-tag followed by MBP, a six amino acid long linker, the TEV-protease recognition site, an alanine-methionine stretch and finally the $VP1_5$-$2A^{pro}$ (C110A) from CVB4 (APPENDIX VII).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3) Rosetta2 pLysS; **Medium:** 2 liter LB-Gluc-Kan; **[IPTG]:** 0.1mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 100µM. Cells were collected by centrifugation, resuspended in 60ml Buffer A/CVB4b (50mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol) and frozen at -80°C. Due to our experience with insufficient cell disruption by sonication, lysozyme (from chicken egg white, Sigma) was added to the solution to a final concentration of 1mg/ml and incubated for 60 minutes on ice to aid the breakup of the *E. coli* cells. Cell disruption was finalized by sonication and insoluble cell debris was removed by centrifugation. 2.5µl of total, insoluble and soluble fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue to examine the level of over-expression (Figure 4-51).

The protein was expressed at excellent levels and was predominantly found in the soluble fraction. Hence, the supernatant was loaded onto a 10ml amylose resin column that had been equilibrated with Buffer A/CVB4b (50mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol). Unbound sample was washed off the column with Buffer A/CVB4b before bound recombinant protein was eluted with the help of Buffer A/CVB4b including 10mM maltose. Aliquots of this purification step were analyzed on 12.5% SDS-PAGE (Figure 4-52).

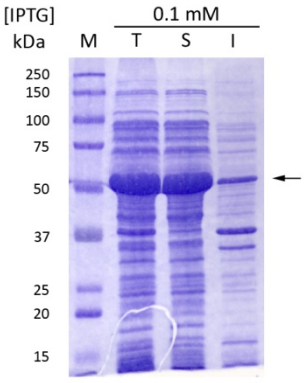


Figure 4-51 Expression of CVB4 MBP-VP1$_5$-2A^{pro} (C110A) in *E. coli* BL21 (DE3) Rosetta2 pLysS. 2.5µL of total (T), soluble (S) and insoluble (I) fractions of the induced samples of CVB4 MBP-VP1$_5$-2A^{pro} (C110A) were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands that potentially arose from the over-expression of recombinant protein are marked with an arrow.

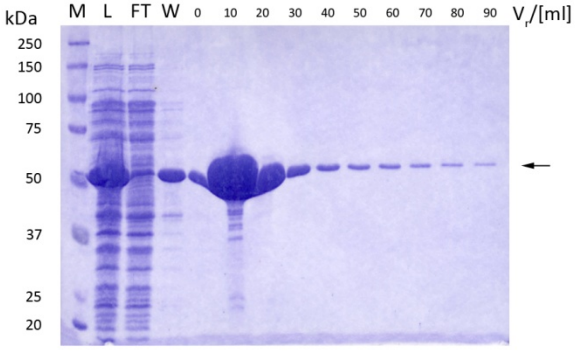


Figure 4-52 CVB4 MBP-VP1$_5$-2A^{pro} (C110A). Monitoring the purification process by amylose resin affinity chromatography on 12.5% SDS-PAGE. 2µl of load (L), flow-through (FT), wash (W) and elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. The arrow marks the bands originating from recombinant protein.

Fractions corresponding to retention volumes 0-40ml were pooled, EDTA was added to a final concentration of 1mM and 3mg of TEV-protease were added and incubated at 4°C overnight. On the following morning, the solution appeared cloudy. The precipitated protein was removed by centrifugation and resuspended in Buffer A/CVB4b. 2µl of uncleaved fusion protein, supernatant after cleavage and resuspended protein were separated on 17.5% SDS-PAGE (data not shown). Cleavage efficiency was satisfactory as nearly all fusion protein appeared to be cleaved. However, the sample still contained a substantial amount of a protein running at ~14kDa. This band, most likely being lysozyme, was also detected in the lane of precipitated protein. In order to remove the remaining lysozyme from the sample the supernatant was loaded onto a MonoQ 10/100 GL that had been equilibrated with Buffer A/CVB4b (50mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol). Chicken egg white lysozyme has an unusually high pI of 11.35 (Wetter & Deutsch, 1951). Consequently, it should not bind to an anion exchanger at pH 8.1 and should be found in the flow-through.

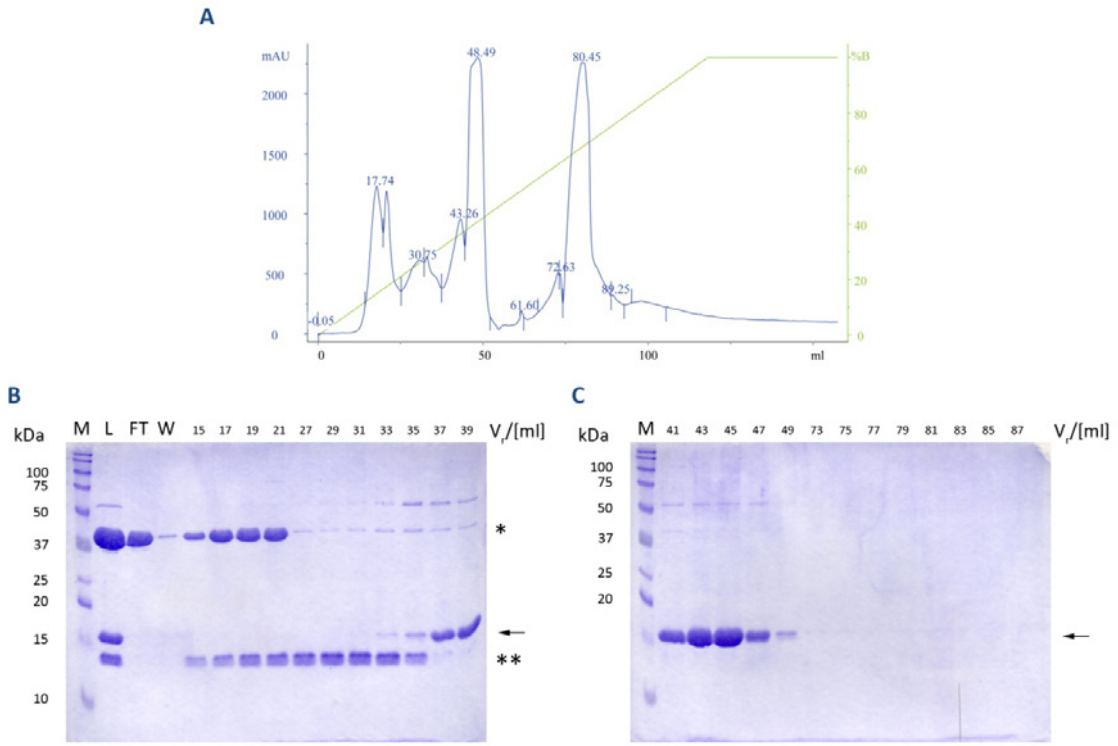


Figure 4-53 Anion exchange chromatography of CVB4 MBP-VP1$_5$-2A^{pro} (C110A) using a MonoQ 10/100 GL column. (A) Profile of CVB4 MBP-VP1$_5$-2A^{pro} (C110A) eluted from the column with the help of a NaCl gradient reaching from 50-1000mM NaCl. **(B+C)** Monitoring the purification: 3µl of load (L), flow-through (FT), wash (W) and the elution fractions were separated on 20% SDS-PAGE and stained with Coomassie Blue. The migration levels of MBP (asterisk), CVB4 VP1$_5$-2A^{pro} (C110A) (arrow) and lysozyme (double asterisk) are marked.

Figure 4-53 shows the elution profile as well as the corresponding SDS-PAGE of the samples. Interestingly lysozyme was not found in the flow-through but in the fractions eluting at 160-380mM NaCl (V_r=15-35ml). This unspecific binding behavior of lysozyme was most probably caused by aggregation due to the reducing conditions of the buffer. Nevertheless, CVB4 VP1$_5$-2A^{pro} (C110A) was successfully separated from MBP and lysozyme. Fractions corresponding to V_r=37-49ml were pooled, concentrated and loaded onto a HiLoad Superdex 75 prep grade 26/60 that had been equilibrated with Buffer B/CVB4b (10mM Tris-HCl pH 8.1, 50mM NaCl, 5mM DTT, 5% v/v glycerol). The protein eluted as a sharp and symmetric peak with a retention volume of 204ml (Figure 4-54 A). Corresponding fractions were pooled and dialyzed against Buffer C/CVB4b (10mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 1mM TCEP). Subsequently, protein was concentrated to ~13mg/ml and the purity was inspected by 20% SDS-PAGE (Figure 4-54 B). No impurities could be detected on the Coomassie stained gel.

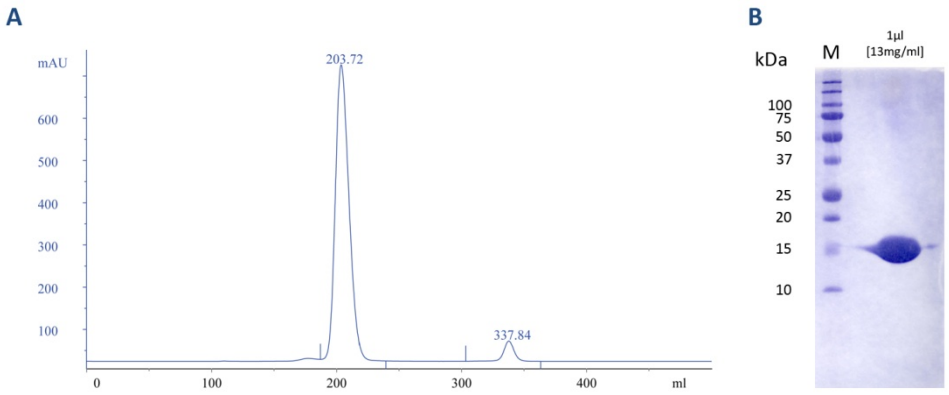


Figure 4-54 Final steps in CVB4 VP1$_5$-2A^{pro} (C110A) purification. (A) Size exclusion profile of the final purification step using a HiLoad Superdex 75 prep grade 26/60 column. CVB4 VP1$_5$-2A^{pro} (C110A) elutes at 204ml. **(B)** Aliquot of final CVB4 VP1$_5$-2A^{pro} (C110A) sample in Buffer C/CVB4b (10mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 1mM TCEP) separated on 20% SDS-PAGE and stained with Coomassie Blue.

A nanoliter crystallization robot (Phoenix) was used to set up crystallization screens in 96-well MRC-2 well plates (see chapter 3.7.1). Initially, we only used the Crystal Screen I+II (Hampton research) to set one plate on 4°C and 22°C each. After six days, the number of drops with precipitate were analyzed and used to readjust the protein concentration for further screens. As the number of drops with protein precipitate seemed to be too high (about 70%), the concentration of CVB4 VP1$_5$-2A^{pro} (C110A) was reduced to 10mg/ml by dilution with Buffer C/CVB4b. Subsequently, the commercial screens found in APPENDIX XIV were used to set up more screens at 4°C and 22°C. However, no crystals could be detected in these drops either.

4.1.4.2 CVB4c: Mal pET CVB4 VP1$_8$-2A^{pro} (C110A)

A PCR fragment amplified with the help of TIM 1858 (containing an *Nde*I site in its 5' non-binding overhang) and TIM 1047 (Table 3-1) from the plasmid pET CVB4 VP1$_8$-2A^{pro} (C110A) was treated with *Nde*I and *Bam*HI and ligated into an *Nde*I/*Bam*HI digested Mal pET vector. A protein containing MBP followed by a linker, His$_{10}$-tag and a TEV protease recognition site fused to CVB4 VP1$_8$-2A^{pro} (C110A) is produced from this plasmid. A CVB4 VP1$_8$-2A^{pro} (C110A) fragment containing three additional residues (Gly-His-Met) on its N-terminus can be isolated upon cleavage with TEV protease (APPENDIX VIII).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3) Rosetta2 pLysS; **Medium:** 1

liter LB-Gluc-Kan; **[IPTG]:** 0.1mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 100µM. The next morning, cells were harvested by centrifugation, resuspended in 30ml Buffer A/CVB4c (50mM Tris-HCl pH 8.5, 50mM NaCl, 5% v/v glycerol, 5mM DTT) and frozen at -80°C.

In order to purify the recombinant protein, cells were thawed and cell disruption was finalized by incubation with lysozyme (1mg/ml end-concentration) for 60min on 4°C and sonication. 3µl of total, soluble and insoluble fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue in order to monitor the expression of recombinant protein (Figure 4-55 A). Purification was performed as described for CVB4 MBP-$VP1_5$-$2A^{pro}$ (C110A) (see chapter 4.1.4.1). The protein eluted in a sharp and symmetrical peak with a retention volume of 210ml from a HiLoad Superdex 75 prep grade 26/60 column (Figure 4-55 B). Fractions were pooled, dialyzed against Buffer C/CVB4c (50mM Tris-HCl pH 8.5, 50mM NaCl, 5% v/v glycerol, 1mM TCEP) and concentrated to around 15mg/ml. Separation of a 1µl aliquot on 20% SDS-PAGE confirmed the purity of the sample (Figure 4-55 C).

A nanoliter crystallization Robot (Phoenix) was used to set drops for crystal screening (see chapter 3.7.1). The commercial screens listed in APPENDIX XIV were utilized with protein solutions at 15mg/ml as well as 10mg/ml at 4°C and 22°C respectively. Drops were constantly monitored. However, no hits could be detected in the time period of observation.

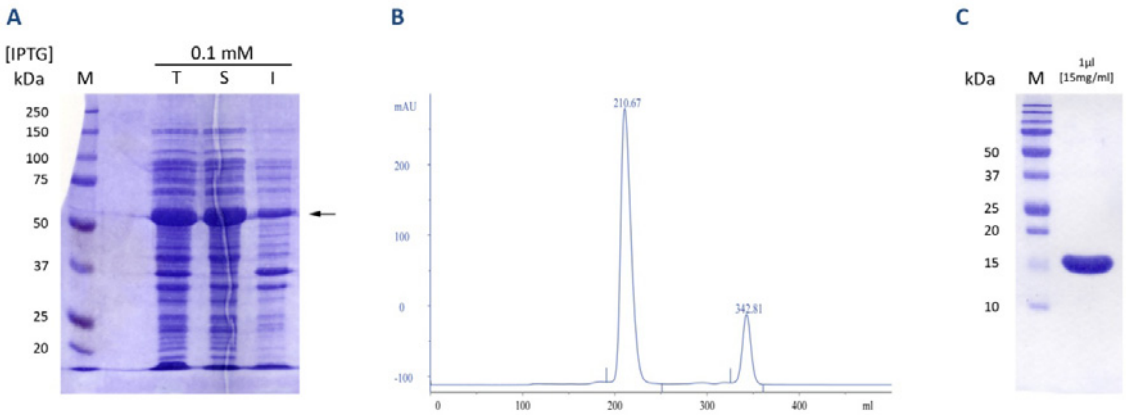


Figure 4-55 Expression of CVB4 MBP-$VP1_8$-$2A^{pro}$ (C110A) and purification of CVB4 $VP1_8$-$2A^{pro}$ (C110A) after TEV-protease cleavage of the fusion protein. (A) CVB4 MBP-$VP1_8$-$2A^{pro}$ (C110A): 3µl of total (T), soluble (S) and insoluble (I) fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands arising from the over-expression of recombinant protein are marked with an arrow. **(B)** Chromatogram of the last purification step of CVB4 $VP1_8$-$2A^{pro}$ (C110A) by size exclusion chromatography. The protein elutes in a sharp and symmetrical peak with a retention volume of 210ml. **(C)** Final protein sample of CVB4 $VP1_8$-$2A^{pro}$ (C110A) concentrated to 15mg/ml. 1µl of sample was run on a 20% SDS-PAGE and stained with Coomassie Blue to assess protein quality and purity.

4.1.4.3 CVB4d: pET MBP CVB4 $VP1_{12}$-$2A^{pro}$ (C110A)

The plasmid was constructed by an identical strategy than pET MBP CVB4 $VP1_5$-$2A^{pro}$ (C110A) using the forward primer TIM1883 (Table 3-1) instead of TIM1882 (see chapter 4.1.4.1). A recombinant protein consisting of MBP with N-terminal His_6-tag fused to a linker followed by a TEV-protease recognition site and CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) is expressed from this plasmid upon induction with IPTG. The MBP tag can be cleaved from the target protein CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) by the help of TEV-protease, leaving a Gly-Ala-Met-Ser tetra peptide fused N-terminally to CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) (APPENDIX IX).

Expression of the recombinant protein followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3) Rosetta2 pLysS; **Medium:** 1 liter LB-Gluc-Kan; **[IPTG]:** 0.1mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 100µM. Cells were harvested by centrifugation, resuspended in 30ml Buffer A/CVB4d (50mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 5mM DTT) and frozen at -80°C. After thawing, cell disruption was finalized by passing the suspension through a homogenizer for four times (see chapter 3.6.2.2). Expression levels were examined by separating 3µl of total, soluble and insoluble fractions on 12.5% SDS-PAGE and Coomassie staining (Figure 4-56 A). Cell disruption via the homogenizer was more efficient compared to sonication. Thus, the homogenizer was routinely used in all further experiments.

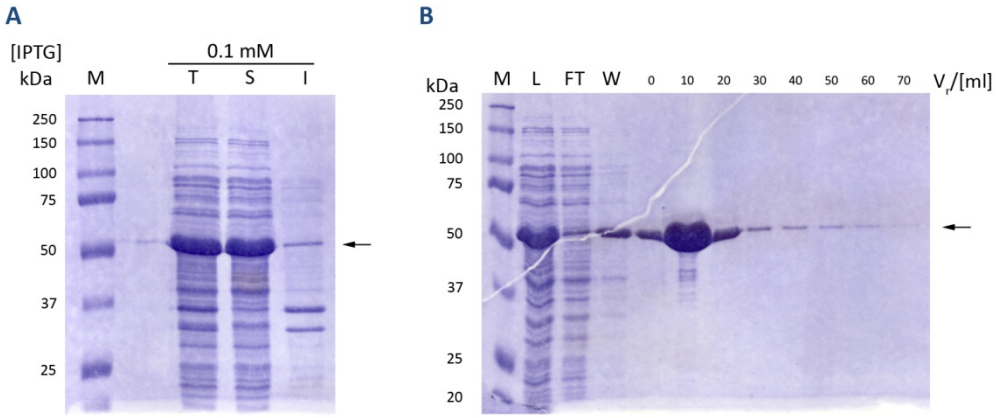


Figure 4-56 Expression and purification of CVB4 MBP-$VP1_{12}$-$2A^{pro}$ (C110A). (A) 3µl of total (T), soluble (S) and insoluble (I) fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. **(B)** Purification of CVB4 MBP-$VP1_{12}$-$2A^{pro}$ (C110A) by affinity chromatography using amylose resin. 2µl of load (L), flow-through (FT), wash (W) and the corresponding elution fractions were separated on 12.5% SDS-PAGE and stained with Coomassie Blue. Bands arising from the presence of fusion protein are marked with an arrow.

As the usage of the homogenizer made the addition of lysozyme for efficient cell disruption unnecessary, a two step purification using amylose resin and size exclusion chromatography could be applied. The supernatant of the crude cell extract was loaded onto a 10ml amylose resin that had been equilibrated with Buffer A/CVB4d (50mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 5mM DTT). Unbound sample was removed by extensive washing with Buffer A/CVB4d and fusion protein was eluted by applying Buffer A/CVB4d including 10mM maltose to the column. Fractions of 10ml each were collected and screened by separating 2µl aliquots on 12.5% SDS-PAGE (Figure 4-56 B). Elution fractions corresponding to a V_r=0-40ml were pooled and EDTA was added to a final concentration of 1mM. 3mg of TEV-protease were added and the sample was incubated at 4°C overnight. Cleavage efficiency was examined by analyzing aliquots of cleaved and uncleaved sample on 20% SDS-PAGE (data not shown). Subsequently the sample was concentrated to a final volume of 5ml and loaded onto a HiLoad Superdex 75 prep grade 26/60 column that had been equilibrated with Buffer B/CVB4d (10mM Tris-HCl pH 8.1, 50mM NaCl, 5% glycerol, 5mM DTT). The size-exclusion profile showed two main peaks corresponding to MBP (V_r=167ml) and CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) (V_r=197ml) with satisfactory separation of the two peaks (Figure 4-57 A). CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) containing fractions were pooled and concentrated to a final concentration of about 10mg/ml. Subsequently the sample was dialyzed against Buffer C/CVB4d (10mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 1mM TCEP). 2µl aliquots of the fractions were loaded onto 20% SDS-PAGE and stained with Coomassie Blue in order to verify the identity and purity of the protein sample (Figure 4-57 B).

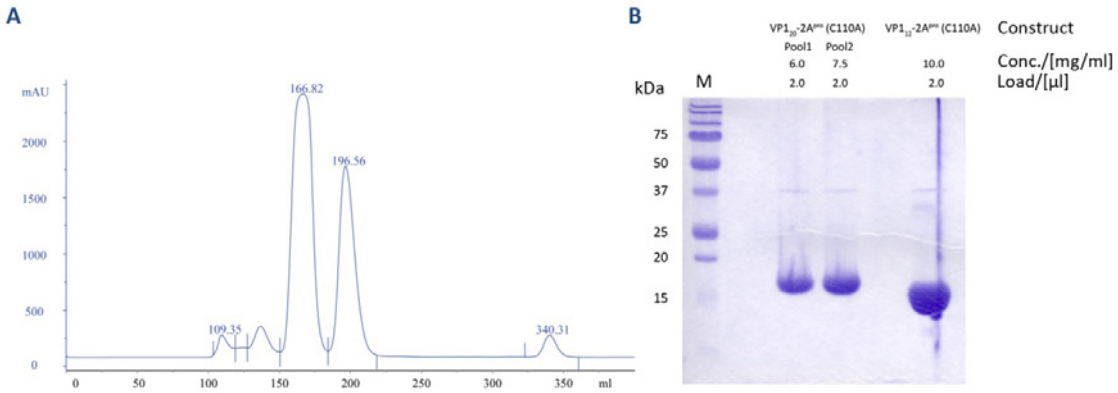


Figure 4-57 Final steps in the purification of CVB4 $VP1_{12}$-$2A^{pro}$ (C110A). (A) Size exclusion profile of MBP and CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) from a HiLoad Superdex 75 prep grade 26/60 column. CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) elutes with a retention volume of about 197ml. **(B)** Inspection of the quality and purity of the final CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) and CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) samples (see chapter 4.1.4.4). 2µl aliquots were separated on 20% SDS-PAGE and stained with Coomassie Blue.

A nanoliter crystallization robot (Phoenix) was used to set up crystallization screens in MRC-2 well plates using the commercial screens listed in APPENDIX XIV (see chapter 3.7.1). Although drops were inspected regularly no crystal hits could be detected.

4.1.4.4 CVB4e: pET MBP CVB4 $VP1_{20}$-$2A^{pro}$ (C110A)

The strategy for the construction of the plasmid was identical to that of pET MBP CVB4 $VP1_5$-$2A^{pro}$ (C110A) except for the use of TIM 1884 (Table 3-1) instead of TIM 1882 (see chapter 4.1.4.1). A fusion protein consisting of N-terminal His_6-tagged MBP fused to CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) via a linker and a TEV-protease recognition site is expressed upon induction with IPTG. A Gly-Ala-Met stretch is found at the N-terminus of the CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) after TEV-protease processing as a remainder of cloning (APPENDIX X).

For the expression and purification of this protein we followed the established protocol for CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) (see chapter 4.1.4.3). CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) eluted from the HiLoad Superdex 75 prep grade 26/60 column in a sharp, symmetrical peak with a retention volume of 192ml (Figure 4-58 A). Analysis of the fractions from gel filtration on 20% SDS-PAGE revealed that MBP and CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) could not be separated completely (Figure 4-58 B). We therefore decided to split the CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) fractions to produce two pools. Pool 1 contained fractions with V_r=185-196ml whereas pool 2 contained fractions with V_r=196-212ml. The samples were concentrated to 6mg/ml (pool 1) and 7.5mg/ml (pool 2) and dialyzed against Buffer C/CVB4e (10mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 1mM TCEP). 2µl of the final samples were separated on 20% SDS-PAGE to estimate quality and purity of the recombinant protein (Figure 4-57 B).

Crystallization screens of both pools were set up identical to the way as described for CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) using the same screens (see chapter 4.1.4.3). However, as the constructs investigated previously, the screens did not give rise to any crystals.

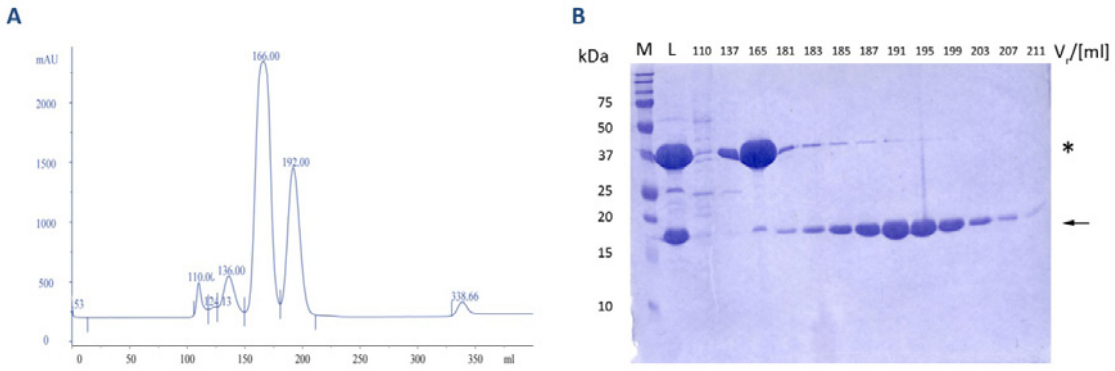


Figure 4-58 Final steps in purification of CVB4 $VP1_{20}$-$2A^{pro}$ (C110A). (A) Size exclusion profile of MBP and CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) from a HiLoad Superdex 75 prep grade 26/60 column. MBP elutes with a retention volume of 166ml whereas CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) elutes at 192ml. **(B)** Analysis of fractions from the final size exclusion purification step. 0.5µl of load (L) and 3µl of the corresponding fractions were separated on 20% SDS-PAGE and stained with Coomassie Blue. The asterisk marks bands of MBP whereas the arrow marks bands corresponding to CVB4 $VP1_{20}$-$2A^{pro}$ (C110A).

4.1.4.5 Analysis of CVB4 $2A^{pro}$ (C110A) constructs with varying VP1 extensions by limited proteolysis

A critical factor for the crystallization of an enteroviral 2A protease with N-terminal VP1 extension clearly is the choice of an optimal length for the extension. Longer extensions might yield better affinity of the substrate to its binding site. However, extending the N-terminus too far will lead to an amino acid stretch that cannot be accommodated anymore and hence to a flexible N-terminus that might interfere with crystal packing. A method to detect exposed and elongated stretches is limited proteolysis. Processing of all four of our CVB4 constructs ($VP1_5$, $VP1_8$, $VP1_{12}$ and $VP1_{20}$ extension) with a range of standard proteases might reveal an exposed cleavage site for one or more of these proteases and thus indicate VP1 residues that cannot be accommodated in the substrate binding cleft any longer.

For this reason, 3µl (10mg/ml) of each construct was mixed with 3µl of a dilution of one of the six proteases (α-chymotrypsin, trypsin, elastase, papain, subtilisin and endoproteinase Glu-C) and incubated at 37°C for 60 min. The reaction was stopped by the addition of Laemmli sample buffer and heating at 95°C for 10 min before the samples were analyzed on 20% SDS-PAGE (Figure 4-59).

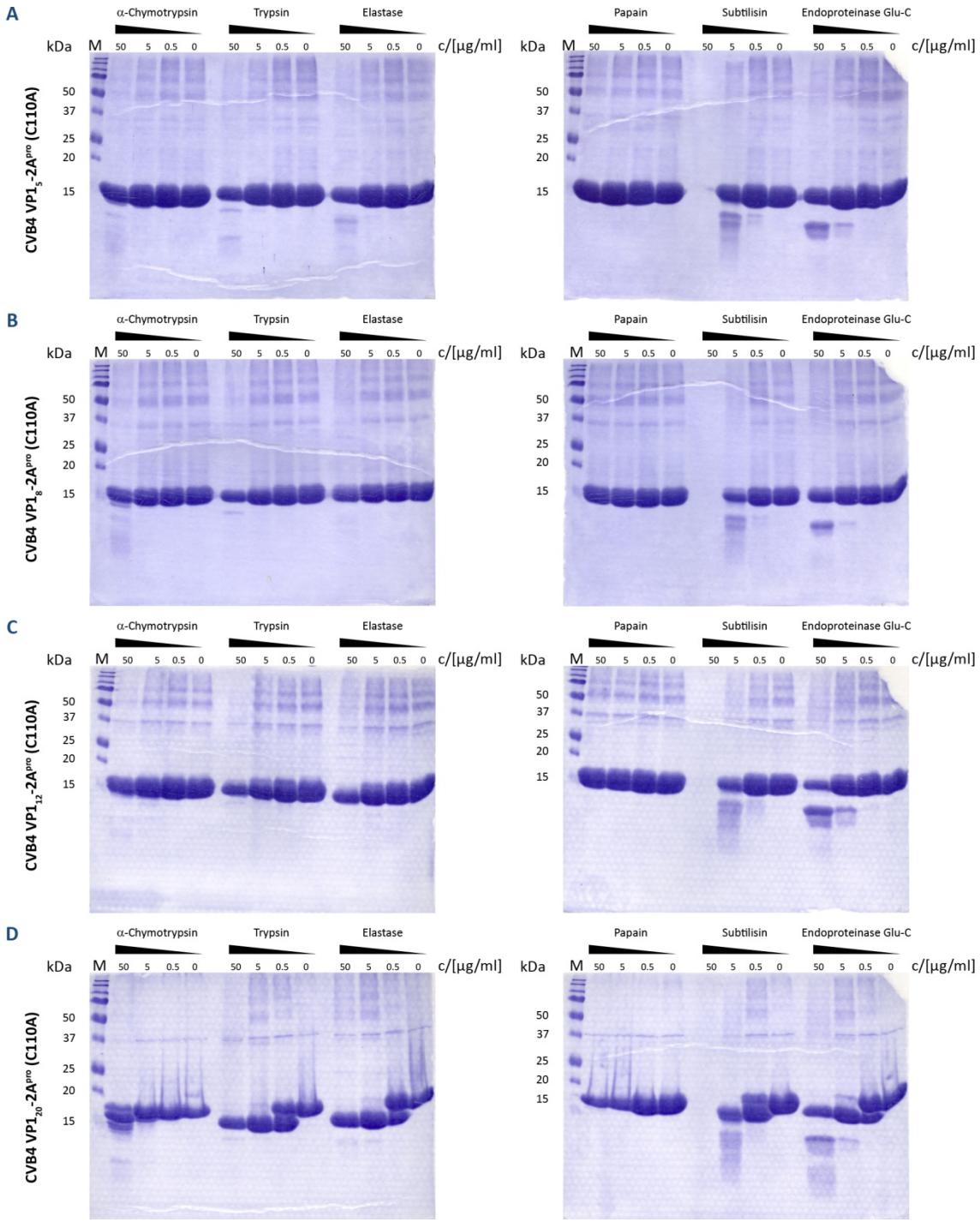


Figure 4-59 Limited proteolysis of CVB4 2A^{pro} (C110A) with four different VP1 extensions. 30µg of each protein in 3µl Buffer C (10mM Tris-HCl pH 8.1, 50mM NaCl, 5% v/v glycerol, 1mM TCEP) was digested with 3µl of α-chymotrypsin, trypsin, elastase, papain, subtilisin or endoproteinase Glu-C in different concentrations. The reactions were incubated at 37°C for 60 min before the reaction was stopped by the addition of Laemmli sample buffer and heat treatment at 95°C for 10 min. Samples were separated at 20% SDS-PAGE and stained with Coomassie Blue. **(A)** CVB4 VP1$_5$-2A^{pro} (C110A). **(B)** CVB4 VP1$_8$-2A^{pro} (C110A). **(C)** CVB4 VP1$_{12}$-2A^{pro} (C110A). **(D)** CVB4 VP1$_{20}$-2A^{pro} (C110A).

The gels for the $VP1_5$, $VP1_8$ and $VP1_{12}$ extended constructs showed very similar results. The recombinant proteins seemed to be quite resistant to processing by the tested proteases. Even at the highest concentrations of the proteases only a minor part of the input was cleaved. An exception was incubation with subtilisin. Subtilisin was capable of digesting the protein samples at concentrations as low as 500ng/ml as can be seen by the faint bands in the corresponding lanes. None of the input protein could be detected after the incubation with 50µg/ml of subtilisin, arguing for a complete degradation of our CVB4 constructs under these conditions.

In contrast, CVB4 $VP1_{20}$-$2A^{pro}$ (C110A) (Figure 4-59 D) was much more prone to cleavage by some of the proteases. Even the lowest concentrations of trypsin or elastase were capable to cleave around 50% of the input material. Endoproteinase Glu-C showed enhanced cleavage and subtilisin seemed to work even more efficient on CVB4 $VP1_{20}$-$2A^{pro}$ (C110A). Given the identical reaction conditions, these significant differences seemed to be a result of the longer VP1 extension. We therefore considered an extension of 20 VP1 amino acids as too long. On contrary, the $VP1_{12}$ construct seemed to behave identical to the $VP1_8$ and $VP1_5$ constructs. Thus, we decided to choose this construct for reductive lysine methylation in order to enhance the chances of crystal hits.

4.1.4.6 Reductive lysine methylation of CVB4 $VP1_{12}$-$2A^{pro}$ (C110A)

Recently, reductive methylation of surface lysine residues has developed into a routine technique for improving the likelihood of crystallization (Walter *et al.*, 2006) by forming new surface contacts that enhance the packing of a protein into a crystal lattice (Sledz *et al.*, 2010). CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) contains four lysine residues (K16, K68, K70 and K92). Analysis of the CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) NMR-structure (PDB 1Z8R) (Baxter *et al.*, 2006) showed that all four of them should be accessible to solvent. However, the ε-amino group of K16 is located on the bottom of a depression, making crystal contacts of this residue less likely. In contrast, the amino groups of K68, K70 and K92 are well exposed (Figure 4-60).

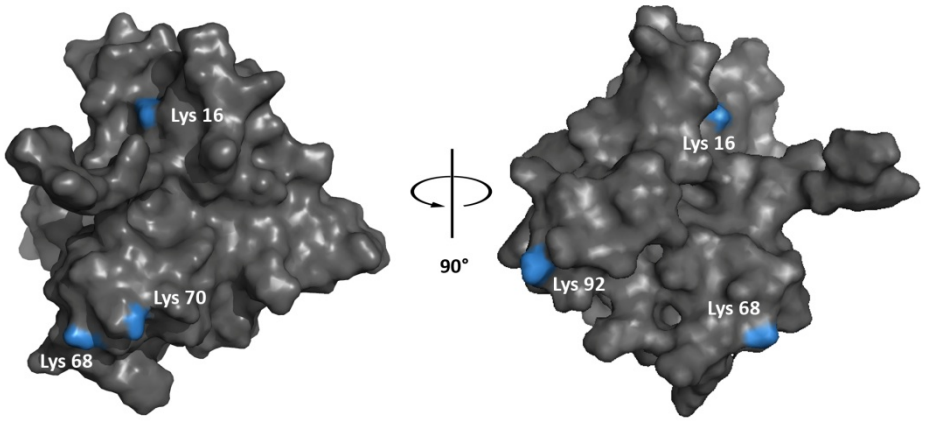


Figure 4-60 Surface representation of the CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) NMR structure (PDB 1Z8R). Model Nr. 1 (out of 17) is depicted as a Connolly surface representation. ε-amino groups of the four lysine residues are colored in blue and labeled. Rendered using PyMol (DeLano, 2002).

Protein was expressed and purified as described in chapter 4.1.4.3. Subsequently, the protein was re-buffered to 50mM HEPES pH 7.5, 250mM NaCl by using a HighPrep desalting 26/60 column (see chapter 3.6.7.2) and diluted to a concentration of 0.5mg/ml. Reductive lysine methylation was performed as described in material and methods (chapter 3.6.9). Subsequently, the sample was concentrated to a final volume of 1ml and re-purified on a HiLoad Superdex 75 prep grade 26/60 column that had been equilibrated with Buffer D/CVB4d (10mM Tris-HCl pH 8.1, 50mM NaCl). The protein eluted in a sharp and symmetric peak with a retention volume of 192ml (compared to 196.5ml for unmethylated CVB4 $VP1_{12}$-$2A^{pro}$ (C110A)) (Figure 4-61 A). Protein containing fractions were pooled and concentrated to a final concentration of 10mg/ml. A 2µl aliquot was loaded on 20% SDS-PAGE to estimate its purity (Figure 4-61 B). An aliquot of the sample was sent for mass spectrometry analysis to confirm the methylation status of the lysine residues. The results contained peptide matches with mono-, di- and trimethylated K16, K68 and K70. A peptide containing K92 could not be detected (data not shown).

A nanoliter crystallization robot (Phoenix) was used to set up crystallization screens in MRC-2 well plates using the commercial screens listed in APPENDIX XIV (see chapter 3.7.1). Drops were regularly monitored. However, crystal hits could not be detected.

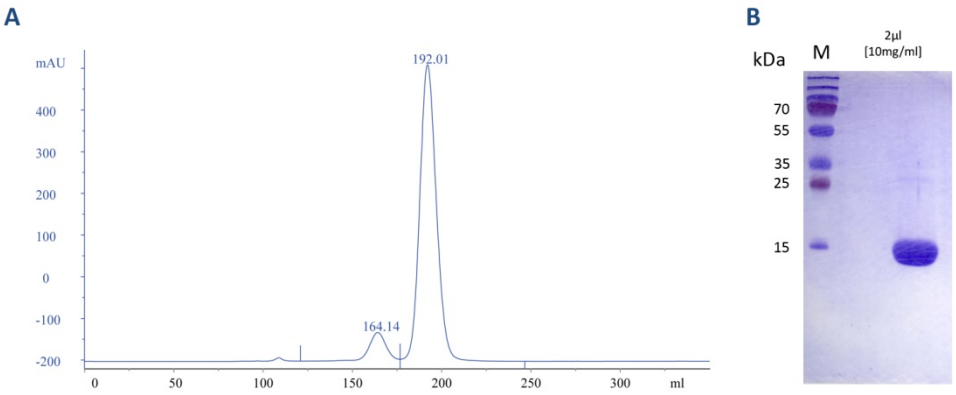


Figure 4-61 Final steps in methylated CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) purification. (A) Size exclusion profile of the final purification step using a HiLoad Superdex 75 prep grade 26/60 column. Methylated CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) elutes at 192ml. **(B)** Aliquot of final, methylated CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) sample in Buffer D/CVB4d (10mM Tris-HCl pH 8.1, 50mM NaCl) separated on 20% SDS-PAGE and stained with Coomassie Blue.

4.2 EXPRESSION STUDIES OF PV1 $2A^{PRO}$ WITHOUT N-TERMINAL EXTENSIONS

Our attempts to obtain a high resolution structure of PV1 $2A^{pro}$ in the state of self-processing failed. However, in the absence of any structural data of a PV $2A^{pro}$, also the structure of apo PV1-$2A^{pro}$ would be of great interest. We therefore tried to express PV1 $2A^{pro}$ without the N-terminal extension in a search for a more soluble variant. Table 4-4 lists the protein variants of the following section together with their most important parameters.

Virus	Plasmid	Protein variant alias	Protein variant expressed from the plasmid			$2A^{pro}$ variant after self-processing (if applicable)			Appendix
			Length [aa]	M_w [kDa]	pI (theoretical)	Length [aa]	M_w [kDa]	pI (theoretical)	
PV1	pET 3d PV1 $VP1_{28}$-$2A^{pro}$	PV1f	177	19.6	5.80	149	16.6	5.79	APPENDIX XI
	Mal pET PV1 $VP1_{8}$-$2A^{pro}$	PV1g	586	64.7	5.77	149	16.6	5.79	Compare APPENDIX III
	Mal pET PV1 $VP1_{8}$-$2A^{pro}$ short linker	PV1h	551	60.5	5.18	149	16.6	5.79	Compare APPENDIX IV
	pET 11d PV1 $2A^{pro}$ (C109A) aa 8-142	PV1i	135	14.9	6.57				APPENDIX XII
CVB4	pET 11d CVB4 $2A^{pro}$ (C110A) aa 7-143	CVB4f	137	15.0	5.53				APPENDIX XIII

Table 4-4 Overview of the protein variants covered in section 4.2. The names of the plasmids are given along their aliases used in the text in order to improve readability. The length, molecular weight and theoretical pI are given for all proteins expressed from the corresponding plasmids. If applicable, these specifications are also given for the $2A^{pro}$ variants after self-processing. Exact sequence information about all new variants is given in the appendix.

4.2.1.1 PV1f: pET 3d PV1 $VP1_{28}$-$2A^{pro}$

An *Xho*I site was introduced into pET 3d CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) (Baxter *et al.*, 2006) by replacing the *Bam*HI/*Blp*I fragment with an oligonucleotide cassette containing an *Xho*I site (TIM 1729/1730) (Table 3-1). A PCR fragment amplified from pCITE PV1 VP1-$2A^{pro}$ with the help of TIM 1718 (binding in the VP1 coding region 28 aa upstream of its C-terminus and carrying a *Bsp*HI site in its 5' non-binding part) and TIM 1728 (binding in the bla-coding region of pCITE PV1 VP1-$2A^{pro}$) was treated with *Bsp*HI and *Xho*I and ligated into a vector prepared with *Nco*I and *Xho*I from pET 3d CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A) carrying the *Xho*I site. A fusion of the last 28 amino acids of VP1 and active $2A^{pro}$ is expressed from this plasmid upon induction. As $2A^{pro}$ is in its active state, it will self-process the VP1-$2A^{pro}$ cleavage site, resulting in the production of active apo PV1 $2A^{pro}$ with an authentic N-terminus (APPENDIX XI).

In order to screen for the optimal expression strain, the plasmid was transformed into three different *E. coli* expression strains; BL21 (DE3), BL21 (DE3) pLysE, BL21 (DE3) Rosetta2 pLysS. Expression of the recombinant proteins followed the protocol described in chapter 3.6.1 using the following parameters. **Medium:** 1 liter LB-Amp; **[IPTG]:** 0.5mM; **Expression temperature:** 16°C; **Expression period:** 15 hours; **[$ZnCl_2$]:** 0µM. Cells were harvested by centrifugation and resuspended in 30ml of Buffer A/PV1f (Tris-HCl pH 7.5, 50mM NaCl). Aliquots of total, soluble and insoluble fractions were separated on 17.5% SDS-PAGE and stained with Coomassie Blue (data not shown). The gels did not show any overexpressed bands at the expected protein size. However, former experiments with HRV2 and CVB4 apo $2A^{pro}$ have shown that these proteases could not be detected in the crude cell extract by Coomassie staining. We therefore continued with fractionated ammonium sulfate precipitation. Proteins precipitating at 20% saturation of ammonium sulfate were removed by centrifugation and ammonium sulfate saturation in the supernatant was raised to 45% to precipitate further proteins possibly including PV1 $2A^{pro}$. However, no PV1 $2A^{pro}$ could be detected in the different steps of this process (see Figure 4-2, lanes 1, 4 and 7 as an example from an expression in BL21 (DE3) cells). A reason for PV1 being not expressed in *E. coli* cells could be its toxicity to the cells. However, even expression in BL21 (DE3) pLysE cells (a strain supposed to aid expression of toxic proteins) was not successful.

4.2.1.2 PV1g: Mal pET PV1 $VP1_8$-$2A^{pro}$ and PV1h: Mal pET PV1 $VP1_8$-$2A^{pro}$ (short linker)

In our attempts to express and purify inactive PV1 $2A^{pro}$ including N-terminal extensions fused to MBP as a solubility tag (see chapters 4.1.1.3 and 4.1.1.4), we included the construction of the identical plasmid lacking the inactivating C109A mutation. Primarily, this experiment was performed as a proof of principle that PV1 $2A^{pro}$ fused to MBP could fold correctly and self-process at its own N-terminus. Secondly, the strong over-expression of MBP fusion proteins could be a way to overcome the problem of PV1 $2A^{pro}$ not being expressed in former experiments (see chapter 4.2.1.1).

The strong over-expression of MBP indeed made expression of active PV1 $2A^{pro}$ possible. However, we were not able to detect PV1 $2A^{pro}$ in the soluble fraction (data not shown).

4.2.1.3 PV1i: pET 11d PV1 $2A^{pro}$ (C109A) aa 8-142 and CVB4f: pET 11d CVB4 $2A^{pro}$ (C110A) aa 7-143

As flexible stretches of proteins often hinder crystallization, we intended to trim the termini of $2A^{pro}$. The crystal structure of HRV2 $2A^{pro}$ (PDB 2HRV) (Petersen *et al.*, 1999) is lacking electron density for the last three amino acids (140-142) in both molecules of the asymmetric unit. As this absence of density is a sign of flexibility of the C-terminus, we decided to trim the C-terminus of CVB4 and PV1 to amino acids 143 and 142 respectively (equivalent positions to HRV2 Ala139). Baxter *et al.* reported reduction of signal intensity beyond the detection limit in the ^{1}H-^{15}N HSQC for all residues in the VP1 extension as well as residues 1-7 of the protease (Baxter *et al.*, 2006). We therefore decided to trim the CVB4 construct to amino acid 7 on the N-terminus. No such structural data is available for the PV1 construct. Hence, the N-terminus was trimmed to amino acid 8, the equivalent position to HRV2 Tyr6, which marks the start of the first β-strand in HRV2 $2A^{pro}$.

For the PV1 $2A^{pro}$ (C109A) aa 8-142 variant, primers TIM 1805 and TIM 1806 (flanking primers) were used together with TIM 1803 and TIM 1804 (mutagenesis primers) (Table 3-1) in a two step PCR mutagenesis (see chapter 3.3.7.2) reaction using MalpET PV1 $VP1_8$-$2A^{pro}$ (C109A) as a template. Mutagenesis destroys an *Nco*I site within the insert by introducing a silent mutation. Hence, the insert can be cloned with *Nco*I and *Bam*HI into the NcoI/*Bam*HI sites of the expression vector pET 11d (APPENDIX XII).

pET 11d CVB4 $2A^{pro}$ (C110A) aa 7-143 was constructed using primers TIM 1899 and TIM 1900 (Table 3-1) to amplify an appropriate fragment from pET MBP CVB4 $VP1_{20}$-$2A^{pro}$ (C110A). The PCR product was digested with *Bsp*HI and *Bam*HI and ligated into *Nco*I/*Bam*HI treated pET 11d vector (APPENDIX XIII).

Expression of the recombinant proteins followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strains:** *E. coli* BL21 (DE3) Rosetta2 pLysS and *E. coli* (DE3) pLysE; **Medium:** 2 liter LB-Amp; **[IPTG]:** 0.3mM; **Expression temperature:** 32°C; **Expression period:** 7 hours; **[$ZnCl_2$]:** 100µM. Cells were harvested by centrifugation and resuspended in 60ml Buffer F (50mM Tris-HCl pH 8.1 (CVB4) or pH 8.5 (PV1), 50mM NaCl, 5mM DTT, 5% v/v glycerol, 1mM EDTA). Cells were stored at -80°C and thawed at room temperature before cell disruption was finalized by sonication.

Inspection of total, soluble and insoluble aliquots on 20% SDS-PAGE stained with Coomassie Blue did not show over-expression of our constructs. However, our experience with the expression of active apo HRV2 and CVB4 $2A^{pro}$ showed a similar situation, and the presence of recombinant protein was only detected after the first purification steps for these constructs (see chapter 4.3.2). Hence, the supernatant was subjected to fractionated ammonium sulfate precipitation and anion exchange chromatography using a MonoQ 10/100 GL column. Unfortunately, no truncated $2A^{pro}$ could be detected in any of the fractions of these purification steps, suggesting that the constructs were indeed not expressed under the tested conditions (data not shown).

4.3 TOWARDS A HIGH RESOLUTION STRUCTURE OF $2A^{PRO}$ BOUND TO AN INHIBITOR

As our attempts to trap an enteroviral 2A protease during self-processing failed, we decided to turn towards a high resolution structure of $2A^{pro}$ covalently bound to an inhibitor. Such structural data could as well be essential for the development of future drugs.

The purification of active $2A^{pro}$ from two enteroviruses has been described, namely HRV2 $2A^{pro}$ and CVB4 $2A^{pro}$ of which HRV2 $2A^{pro}$ has been successfully crystallized and the structure determined to 1.95 Å (Liebig *et al.*, 1993; Petersen *et al.*, 1999). For this reason, we chose these two proteases as well as the two well characterized canonical inhibitors chymostatin and elastatinal for our experiments. Both of them have been reported to inhibit HRV2 $2A^{pro}$ in peptide

cleavage assays with IC_{50} values of 55µM for elastatinal and 25µM for chymostatin (Sommergruber *et al.*, 1992). Moreover, elastatinal was shown to act on the 2A proteinases of poliovirus and HRV14 (Molla *et al.*, 1993). The naturally occurring compound chymostatin was crystallized in complex with *Streptomyces griseus* protease A and the structure was solved to a resolution of 1.8 Å (Delbaere & Brayer, 1985).

Virus	Plasmid	**Protein variant alias**	**Length [aa]**	**M_w [kDa]**	**pI (theoretical)**
CVB4	pET 8c CVB4 $2A^{pro}$	CVB4g	150	16.5	4.90
HRV2	pET 8c HRV2 $2A^{pro}$	HRV2b	142	16.2	5.37

Table 4-5 Overview of the protein variants covered in section 4.3. The names of the plasmids are given along their aliases used in the text in order to improve readability. The length, molecular weight and theoretical pI are given for all proteins expressed from the corresponding plasmids.

4.3.1 CVB4g: pET 8c CVB4 $2A^{pro}$

4.3.1.1 Expression, purification and crystallization screens

The plasmid pET8c CVB4 $2A^{pro}$ has been described (Liebig *et al.*, 1993). Expression and purification of the protease followed a similar yet not identical protocol to the one described by Liebig *et al.* For expression we followed the protocol described in chapter 3.6.1 using the following parameters. **Expression strain:** *E. coli* BL21 (DE3) pLysE; **Medium:** 3 liter LB-Amp; **[IPTG]:** 0.3mM; **Expression temperature:** 32°C; **Expression period:** 5 hours; **[$ZnCl_2$]:** 100µM. Cells were harvested by centrifugation, resuspended in 50ml Buffer A/CVB4g (50mM Tris-HCl pH 8.1, 50mM NaCl, 1mM EDTA, 5mM DTT, 5% v/v glycerol) and broken by passing the suspension through a homogenizer for three times (see chapter 3.6.2.2). Aliquots of total, soluble and insoluble fractions were separated on 20% SDS-PAGE and stained with Coomassie Blue in order to monitor the expression of the recombinant protein (data not shown). Although no signs of over-expression were detected, the supernatant was passed on to fractionated ammonium sulfate precipitation, anion exchange chromatography and gel filtration as described by Liebig *et al.* The protease elutes with a retention volume of roughly 212ml from a HiLoad Superdex 75 prep grade 26/60 column (Figure 4-62 A). An aliquot of the pooled and mildly concentrated sample separated on 20% SDS-PAGE and stained with Coomassie Blue is shown in Figure 4-62 B.

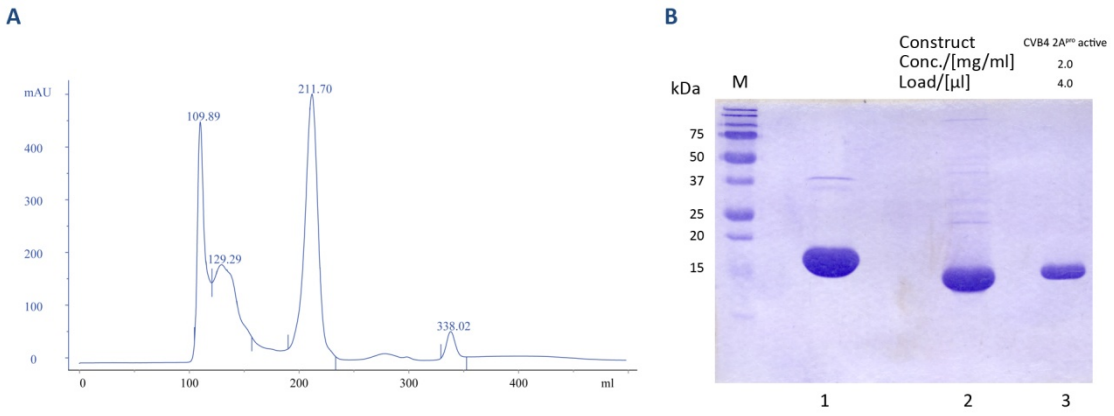


Figure 4-62 Final steps in the purification of active CVB4 2A^{pro}. (A) Size exclusion profile of CVB4 2A^{pro} eluting from a HiLoad Superdex 75 prep grade 26/60 column. The protein elutes in a sharp peak with a retention volume of roughly 212ml. **(B)** CVB4 2A^{pro} containing fractions were pooled, concentrated and separated on 20% SDS-PAGE in order to check the purity of the sample. Lanes 1+2 are protein samples from another experiment.

The sample was further concentrated to 13mg/ml and dialyzed against Buffer B/CVB4g (10mM Tris-HCl pH 8.1, 10mM BME). A nanoliter crystallization robot was used to set up crystallization screens in MRC2-well plates using the commercial screens listed in APPENDIX XIV (see chapter 3.7.1). Unfortunately, no crystals could be detected within the period of observation. Consequently, crystal soaking experiments with chymostatin and elastatinal could not be performed.

For co-crystallization trials, we incubated the sample with a final concentration of 2.4µM of inhibitor (i.e. a 3-fold molar excess of either chymostatin or elastatinal to protein, from 10mM stocks in DMSO) at 4°C overnight. Subsequently the sample was dialyzed against Buffer B/CVB4g (10mM Tris-HCl pH 8.1, 10mM BME) in order to remove excess DMSO. Aliquots of these samples as well as untreated sample were separated on 20% SDS-PAGE in order to observe a possible band shift caused by irreversible binding of the inhibitors (Figure 4-63). Although no clear changes in migration in the SDS-PAGE could be observed crystal screens were set up in an identical way as described before using the commercial screens listed in APPENDIX XIV. Still, no crystal hits could be observed.

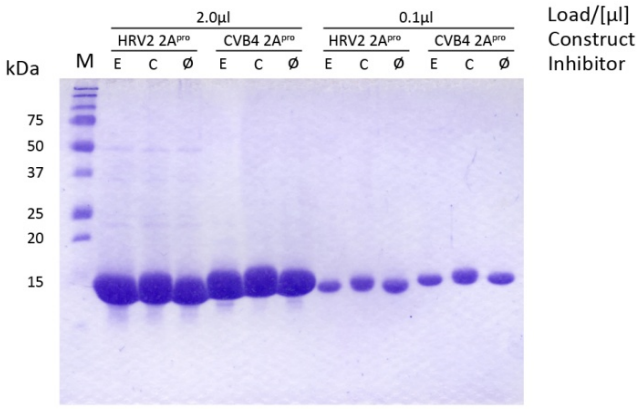


Figure 4-63 Comparing migration of active 2A^{pro} from CVB4 and HRV2 that were incubated with either chymostatin or elastatinal to the apo form. 2A^{pro} samples from either HRV2 (14mg/ml) or CVB4 (13mg/ml) were incubated with a 3-fold molar excess of inhibitors for 15 hours at 4°C. DMSO was removed by dialysis. 0.1µl or 2.0µl aliquots were separated on 20% SDS-PAGE and stained with Coomassie Blue to observe possible band shifts caused by the covalent binding of the inhibitors. (E) 2A^{pro} incubated with elastatinal, (C) 2A^{pro} incubated with chymostatin, (∅) 2A^{pro} without inhibitor.

4.3.2 HRV2b: pET 8c HRV2 2Apro

4.3.2.1 Expression and purification

The plasmid pET 8c HRV2 2A^{pro} has been described (Liebig *et al.*, 1993). Expression and purification of the protease followed the protocol described in chapter 4.3.1.1 for CVB4 2A^{pro}. Active HRV2 2A^{pro} elutes from the HiLoad Superdex 75 prep grade 26/60 column with a retention volume of 175ml (Figure 4-64 A). The substantial difference in retention volumes of CVB4 and HRV2 2A proteases was already pointed out by Liebig *et al.* in 1993. HRV2 2A^{pro} containing fractions were pooled and concentrated. A 4µl aliquot of a sample with 5mg/ml was analyzed on 20% SDS-PAGE (Figure 4-64 B).

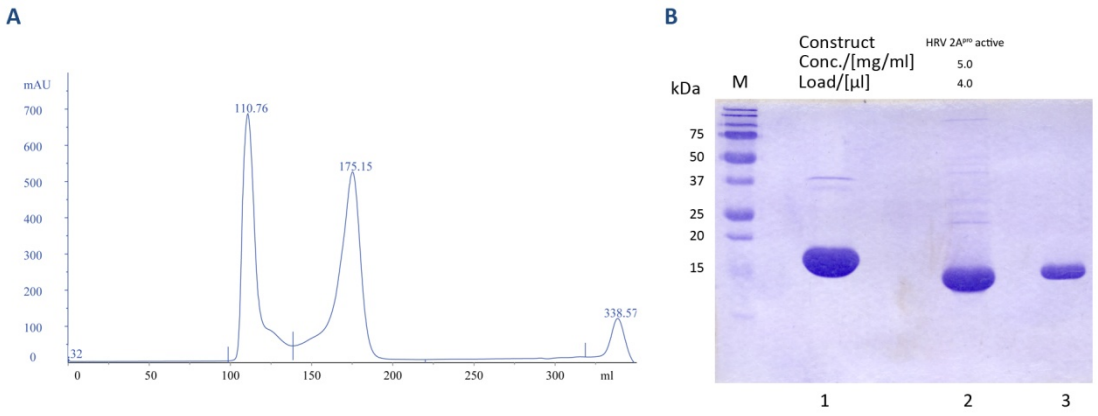


Figure 4-64 Final steps in the purification of active HRV2 2A^{pro}. (A) Size exclusion profile of HRV2 2A^{pro} eluting from a HiLoad Superdex 75 prep grade 26/60 column. The protein elutes in a slightly asymmetric peak with a retention volume of roughly 175ml. **(B)** HRV2 2A^{pro} containing fractions were pooled, concentrated and separated on 20% SDS-PAGE in order to check the purity of the sample. Lanes 1+3 are samples from a different experiment.

4.3.2.2 Co-crystallization trials of HRV2 2A^{pro} with chymostatin and elastatinal

Before starting co-crystallization experiments, we aimed to investigate inhibition of HRV2 2A^{pro} by chymostatin and elastatinal in an *in vitro* cleavage assay probing eIF4GI integrity. Protease was pre-incubated with different concentrations of either chymostatin or elastatinal for 15 minutes at 30°C. Subsequently, rabbit reticulocyte lysate (RRL) was added. The reactions were incubated at 30°C for another 60 minutes before Laemmli sample buffer was added followed by heating to 95°C for 10 minutes. Samples were separated on 6% SDS-PAGE, blotted onto PVDF membrane and investigated for eIF4GI cleavage (Figure 4-65) (see chapter 3.5.2.4). The blots revealed that the two inhibitors indeed could impair eIF4GI cleavage down to a concentration of 1µM, corresponding to an 8-fold molar excess of inhibitor to protease.

Having proven the inhibitory effect of chymostatin and elastatinal on the 2A^{pro} sample, we aimed to co-crystallize the protease with inhibitor. HRV2 2A^{pro} at a concentration of 14mg/ml was incubated with a 3-fold molar excess of either chymostatin or elastatinal (final concentration 2.6µM, from 10mM stocks of inhibitors in DMSO). The mixture was incubated at 4°C overnight. Samples of apo HRV2 2A^{pro} and 2A^{pro} incubated with inhibitors were analyzed on 20% SDS-PAGE in order to observe possible band shifts due to covalent binding of the inhibitors to the protease (Figure 4-63). As with active CVB4 2A^{pro}, no substantial band shift could be detected. However, as our *in vitro* assays showed an inhibitory effect, we set up crystallization experiments.

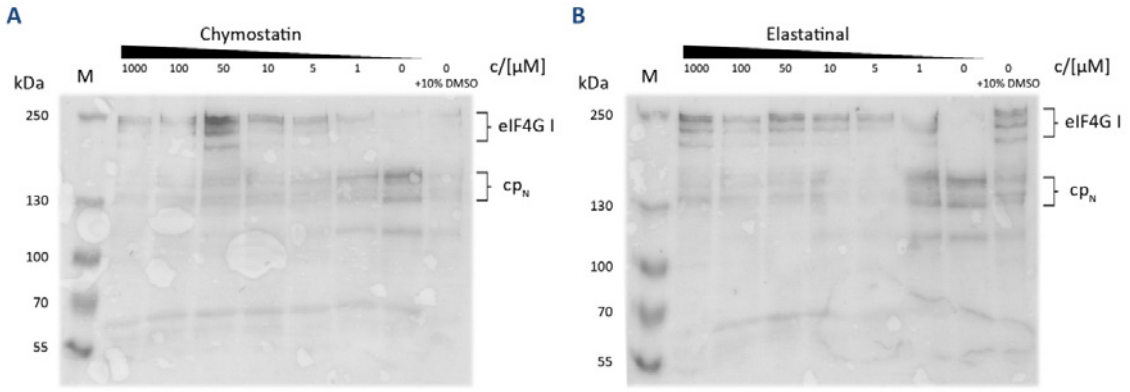


Figure 4-65 *In vitro* eIF4GI cleavage assay using recombinant HRV2 2A^{pro} with different concentrations of either chymostatin or elastatinal. eIF4GI present in rabbit reticulocyte lysate was used as a substrate to observe cleavage by HRV2 2A^{pro} under different concentrations of inhibitor. 20ng of protease and different amounts of inhibitors were pre-incubated and mixed with 4µl RRL's. The reactions were incubated at 30°C for 60 minutes and stopped by addition of Laemmli sample buffer and heat treatment at 95°C. A reaction without inhibitor but including 10% DMSO served as a control. Samples were separated on 6% SDS-PAGE and blotted onto PVDF membrane before eIF4GI cleavage was observed with an α-eIF4GI antiserum. Bands originating from eIF4GI or the N-terminal cleavage products are labeled.

The sample was dialyzed against a solution of 10mM BME (compare crystallization conditions of apo HRV2 2A^{pro} (Petersen *et al.*, 1999)). In a first crystallization experiment, the known crystallization conditions for apo HRV2 2A^{pro} were used. A hanging drop setup in 48-well plates was used to equilibrate drops containing of either 1µl protein and 1µl reservoir or 2µl protein and 1µl reservoir against 200µl of reservoir. The reservoir contained the crystallization buffer described in Petersen *et al.* (50mM Na^+/K^+ phosphate pH 8.0, 7.5% w/v PEG 8000, 300mM KCNS, 5% v/v tertiary amyl alcohol, 10% v/v glycerol, 10mM BME). Drops were incubated at 4°C or 22°C and monitored regularly. Unfortunately, no crystals could be observed. In parallel, a nanoliter crystallization robot (Phoenix) was used to setup crystallization trials in MRC2 well plates using the commercial screens listed in APPENDIX XIV (see chapter 3.7.1). However, no crystal hits could be observed.

Again, we became skeptical about the actual binding of the inhibitors to our protease. We therefore wanted to compare the eIF4GI cleavage activity of apo HRV2 2A^{pro} with the activity of a sample used for chymostatin co-crystallization. A number of protease dilutions was mixed with RRL's and incubated for 60 minutes at 30°C. After stopping the reaction with Laemmli sample buffer and heat treatment, the samples were separated on 6% SDS-PAGE and blotted on PVDF membrane in order to monitor the status of eIF4GI (Figure 4-66). When comparing the results, no substantial inhibition of eIF4GI cleavage by chymostatin could be detected. Hence, there seems to

be a problem with inhibitor binding in big batches compared to the cleavage assays performed before.

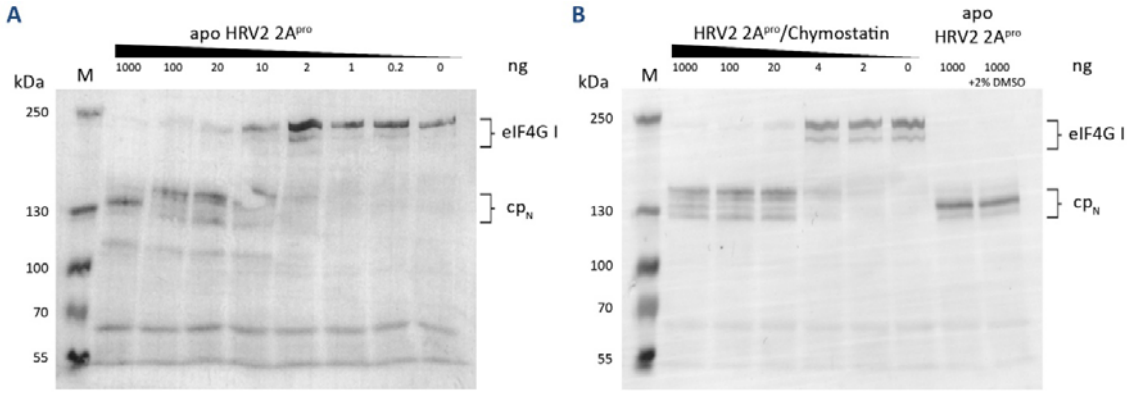


Figure 4-66 Comparing cleavage activity of apo HRV2 2A^{pro} with an HRV2 2A^{pro} sample incubated with chymostatin. 4µl of RRL were incubated with a serial dilution of the two protease samples at 30°C for 60 minutes. Subsequently the status of eIF4GI was observed by immunoblotting and taken as a measure for the activity of the protease. Bands arising from intact eIF4GI as well as the N-terminal cleavage product are labeled. **(A)** Status of eIF4GI after incubation with a serial dilution of apo HRV2 2A^{pro}. **(B)** Status of eIF4G after incubation with a serial dilution of HRV2 2A^{pro} that has been incubated with a 3-fold molar excess of chymostatin. Samples using 1000ng of apo HRV2 2A^{pro} with and without 2% DMSO served as positive controls.

4.3.2.3 Analysis of the apo HRV2 2A^{pro} structure (PDB 2HRV) for accessibility of the active site

Co-crystallization of CVB4 and HRV2 2A^{pro} with either chymostatin or elastatinal was not successful. Hence, we wanted to make use of the successful solution of the apo HRV2 2A^{pro} structure. We aimed to reproduce crystals from HRV2 2A^{pro} as published previously (Petersen *et al.*, 1999) and soak them with either chymostatin or elastatinal.

In order to successfully soak inhibitor into an apo HRV2 2A^{pro} crystal, the crystal packing must allow diffusion of the inhibitor into the crystal by sufficiently big channels. Moreover, the active site must not be blocked by crystal contacts in order to allow binding of the inhibitor. Analysis of the PDB entry 2HRV showed that both prerequisites are fulfilled. In a fully extended conformation, the two inhibitors exhibit a length of around 20 Å. The chemical structures of chymostatin and elastatinal are depicted in Figure 4-67 A and B. Figure 4-67 C shows the arrangement of 106 HRV2 2A^{pro} molecules within the crystal lattice. Big pores with diameters of at least 35Å are channeling the crystal, thereby allowing diffusion of the inhibitor into the crystal. Moreover, the active site residues His18, Asp35 and Cys106 (depicted in red) line the inner surface of the channels. A more detailed look at the crystal contacts shows that the active site is clearly accessible within the crystal. Figure 4-67 D shows two molecules of the asymmetric unit

(cartoon representation in black) with the active site residues colored in red. The surrounding surface represents residues of the surrounding molecules in the crystals (cutoff 15Å). The active site of the protease is not covered by neighboring molecules and thus should be able to bind inhibitor.

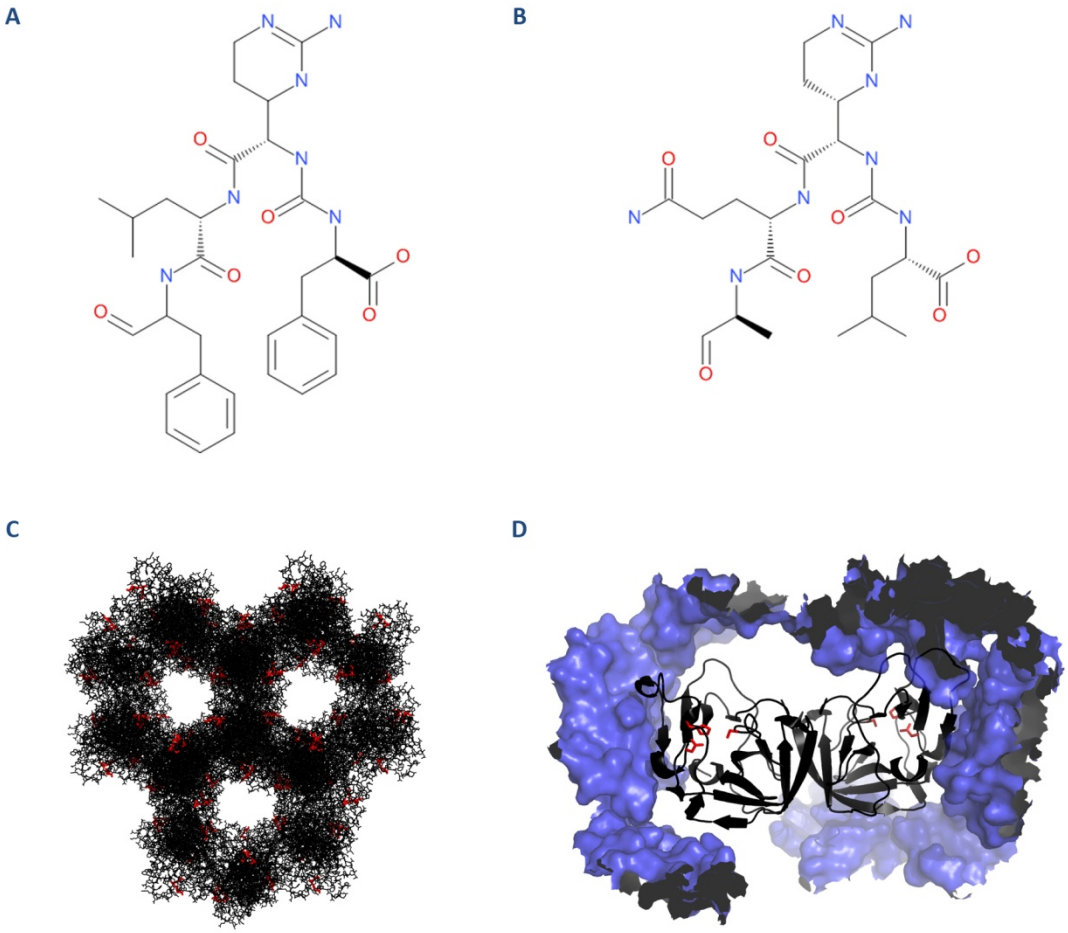


Figure 4-67 Analysis of the apo HRV2 2A^{pro} structure for inhibitor accessibility. The crystal packing of HRV2 2A^{pro} should allow access of the inhibitors upon crystal soaking. **(A)** Channels, at least 35Å in diameter, are passing through the crystal with the active site residues (red) lining the inner surface of the channels. **(B)** The two molecules within the asymmetric unit (black cartoon, active site residues in red) and all residues of neighboring molecules within 15Å (blue surface) are depicted. **(C)** Chemical structure of chymostatin. **(D)** Chemical structure of elastatinal. Representations rendered with PyMol (DeLano, 2002) and SymmyxDraw.

4.3.2.4 Crystallization of HRV2 2A^{pro} and soaking with chymostatin or elastatinal

As co-crystallization was unsuccessful, we aimed to reproduce apo HRV2 2A^{pro} crystals as described by Petersen *et al.* and soak them with either chymostatin or elastatinal. A fresh batch of HRV2 2A^{pro} (9mg/ml) in 10mM BME was used to set a number of hanging drops (see Table 4-6 for composition) that were equilibrated over 500µl of reservoir solution in 24-well plates. Plates were incubated at either 4°C or 22°C and monitored regularly. Light precipitate was detected within the drops after 2-3 days. The first crystals could be observed within 14 days after setup at 4°C that grew to full size in further 14 days (Figure 4-68). No crystals could be detected at 22°C.

	Reservoir	Protein solution
Row A	1.0µl	2.0µl
Row B	1.2µl	2.4µl
Row C	1.4µl	1.8µl
Row D	1.6µl	3.2µl

Table 4-6 Composition of drops for the crystallization of HRV2 2A^{pro} in a 24-well hanging drop array. 7 drops per well were equilibrated over 500µl of reservoir (50mM Na^+/K^+ phosphate pH 8.0, 7.5% w/v PEG 8000, 300mM KCNS, 5% v/v tertiary amyl alcohol, 10% v/v glycerol, 10mM BME) at 4°C or 22°C.

HRV2 2A^{pro} crystals grew as long rods with a typical length of about 0.5mm and hexagonal profile, being in good agreement with the results described by Petersen *et al.* Using the crystallization buffer including 21% v/v glycerol as cryo-solution, crystals diffracted to 2.5 Å using an in house X-Ray source (see chapter 3.8.1) (data not shown).

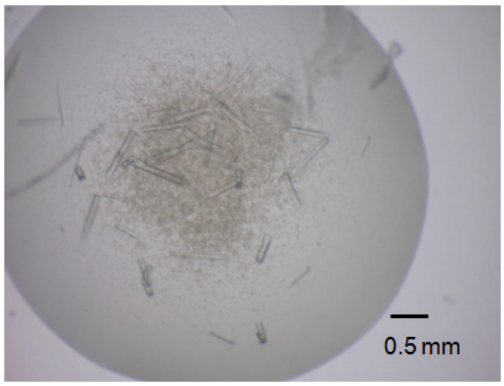


Figure 4-68 Crystals of apo HRV2 2A^{pro}. Crystals were grown using the hanging drop vapor diffusion method at 4°C in 24-well plates. 7 drops per well were equilibrated against 500µl or reservoir solution (50mM Na^+/K^+ phosphate pH 8.0, 7.5% w/v PEG 8000, 300mM KCNS, 5% v/v tertiary amyl alcohol, 10% v/v glycerol, 10mM BME). First crystals were detected 2 weeks after setup and were further grown till the end of week 4 after setup.

Given the promising quality of the crystals, we aimed to soak them with either chymostatin or elastatinal. Crystals were fished from the drops and washed in reservoir solution

before they were dropped into a dilution of the inhibitors (from 10mM DMSO stocks) for an appropriate time. Subsequently, crystals were dropped into cryo-solution and frozen in liquid nitrogen. Table 4-8 lists all combinations of inhibitor concentrations and soaking times tested as well as the maximum diffraction of the crystals (if still available) after the soaking experiment. Generally, crystals seemed to suffer from soaking as many of them dissolved or cracked during the procedure. Crystals staying intact nearly always exhibited slightly rounded edges, when compared to untreated crystals. Moreover, many of the soaked crystals did not diffract at all, whereas all still diffracting crystals showed worse diffraction than the non treated crystals.

Crystals with the identifiers 1-13 were sent to BESSY II in Berlin and complete datasets were collected from crystals 3, 4 and 9 at beamline BESSY_14.1 (λ=0.918Å) on a MarMosaic 225 detector. The datasets consisted of 200 images with 1° rotation for crystal 3 and 4 as well as 160 images with 1° rotation for crystal 9. Data was indexed and integrated using iMosflm (Leslie & Powell, 2007) and cut at 3.0Å (crystal 3), 3.3Å (crystal 4) and 3.2Å (crystal 9). The spacegroup was detected as $P3_221$ with unit cell dimensions as given in Table 4-7 and two molecules in the asymmetric unit (asu). Data was scaled with SCALA (CCP4, 1994) allowing the program to choose an R_{free} dataset of 5% of all reflections. Statistics of the datasets obtained are listed in Table 4-9.

	a	b	c	α	β	γ
Crystal 3	72.94	72.94	162.1	90°	90°	120°
Crystal 4	73.76	73.76	161.7	90°	90°	120°
Crystal 9	72.46	72.46	162.4	90°	90°	120°

Table 4-7 **Unit cell dimensions within the crystals.**

	Nr. of X-tals	[Inhibitor]	Soaking time	Nr. of X-tals with remaining diffraction	Max. resolution	X-tal identifier	Comments
Chymostatin	3	10.0 mM	-	0	-		X-tals dissolved immediately
	5	5.0 mM	15 h	0	-		4 X-tals dissolved
	5	3.3 mM	15 h	0	-		All X-tals dissolved
	3	6.6 mM	1 min	0	-		
	3	5.0 mM	1 min	0	-		
	3	3.3 mM	1 min	1	4.0 Å	**1**	
	4	100 µM	3 h	3	4.6 Å	**2**	
					3.6 Å	**3**	
					3.5 Å	**4**	
	3	100 µM	2 min	0	-		
	3	100 µM	20 h	2	4.4 Å	**5**	
					10.0 Å		
Elastatinal	3	10.0 mM	-	0	-		X-tals dissolved immediately
	5	5.0 mM	15 h	0	-		All X-tals dissolved
	5	3.3 mM	15 h	1	3.0 Å	**6**	2 X-tals dissolved
	3	6.6 mM	1 min	0	-		
	3	5.0 mM	1 min	0	-		
	3	3.3 mM	1 min	0	-		
	4	100 µM	3 h	4	4.6 Å	**7**	
					2.7 Å	**8**	
					4.0 Å	**9**	
					3.4 Å	**10**	
	2	100 µM	2 min	2	4.0 Å	**11**	
					3.5 Å	**12**	
	3	100 µM	20 h	1	4.3 Å	**13**	

Table 4-8 Overview of HRV2 2A^{pro} crystal soaking experiments. Crystals were fished from drops and washed in reservoir solution before they were dropped into a solution of inhibitor (Dilutions of a 10mM stock in DMSO with reservoir solution) for the indicated time. Crystals were then fished from this drop, shortly dropped into reservoir solution with 21% glycerol (cryo solution) and frozen in liquid nitrogen.

Initial phases were solved by molecular replacement using the program Phaser (McCoy *et al.*, 2007). The apo structure of HRV2 $2A^{pro}$ (PDB 2HRV) was used as a search ensemble and single solutions with log-likelihood gains (LLG) of 1567 (crystal 3), 1254 (crystal 4) and 1354 (crystal 9) were obtained. The initial electron density maps were scanned for positive difference peaks. The two most intense peaks were identified as the missing Zn^{2+}-ions which were subsequently added to the protein coordinates. The models including the Zn^{2+}-ions were rigid body refined followed by refinement including simulated annealing in PHENIX. The R_{work} and R_{free} values of the resulting models are given in Table 4-10.

	Crystal 3			Crystal 4			Crystal 9		
	Overall	Inner Shell	Outer Shell	Overall	Inner Shell	Outer Shell	Overall	Inner Shell	Outer Shell
Low resolution limit	41.07	41.07	3.16	41.20	41.20	3.48	40.98	40.98	3.37
High resolution limit	3.00	9.49	3.00	3.30	10.44	3.30	3.20	10.12	3.20
R_{merge}	0.114	0.043	0.286	0.142	0.039	0.337	0.123	0.035	0.289
R_{merge} in top int. bin	0.056	-	-	0.046	-	-	0.046	-	-
R_{meas} (within I+/I-)	0.119	0.045	0.298	0.148	0.042	0.352	0.130	0.037	0.305
R_{meas} (all I+ & I-)	0.119	0.045	0.298	0.148	0.042	0.352	0.130	0.037	0.305
R_{pim} (within I+/I-)	0.034	0.014	0.085	0.043	0.014	0.101	0.042	0.013	0.098
R_{pim} (all I+ & I-)	0.034	0.014	0.085	0.148	0.042	0.352	0.042	0.013	0.098
Frac. partial bias	-0.011	-0.010	-0.046	-0.041	-0.045	-0.069	-0.013	-0.001	-0.044
Tot. nr. of obs.	126515	3880	18569	95261	2858	14006	80335	2508	11488
Tot. nr. unique	10561	393	1515	8172	303	1170	8695	330	1208
Mean((I)/sd(I))	18.3	28.9	8.8	16.2	31.7	7.8	15.7	26.2	7.8
Completeness	100.0	99.0	100.0	100.0	98.7	100.0	99.9	98.8	99.9
Multiplicity	12.0	9.9	12.3	11.7	9.4	12.0	9.3	7.6	9.5

Table 4-9 Statistics of the X-ray datasets collected at BESSY _14.1 and processed using iMosflm and SCALA.

	R_{work}	R_{free}
HRV2 $2A^{pro}$, crystal 3 (chymostatin)	0.2009	0.2645
HRV2 $2A^{pro}$, crystal 4 (chymostatin)	0.1861	0.2424
HRV2 $2A^{pro}$, crystal 9 (elastatinal)	0.1849	0.2493

Table 4-10 R_{work} and R_{free} values of HRV2 $2A^{pro}$ models. Values as assessed by PHENIX after refining the models including simulated annealing.

Inspection of the electron density maps revealed extra density at the active site cysteine (Cys106) and along the active site cleft (Figure 4-69) possibly caused by the presence of the inhibitors. However, placement of the inhibitors into the density was not possible due to insufficient quality of the maps. For this reason we aimed to place a phenylalanyl or alanyl molecule (the P1 residues of chymostatin and elastatinal respectively) within the electron density blobs at P1, as density was of best quality at this position. Nevertheless, refinement of such models resulted in uninterpretable maps and a setback of R-factors.

Next, we aimed to calculate omit maps, excluding residues in proximity to the expected inhibitor binding site, in order to reduce phase bias. Four omit regions were defined as given in Table 4-11 and depicted in Figure 4-70. The AutoBuild GUI implemented in PHENIX was used to calculate omit maps including simulated annealing refinement for either single regions or combinations of regions. However, resulting maps did not show improved quality of electron density and were no better interpretable than previous maps (data not shown). For these reasons, a model of HRV2 2A^{pro} bound to chymostatin or elastatinal could not be obtained.

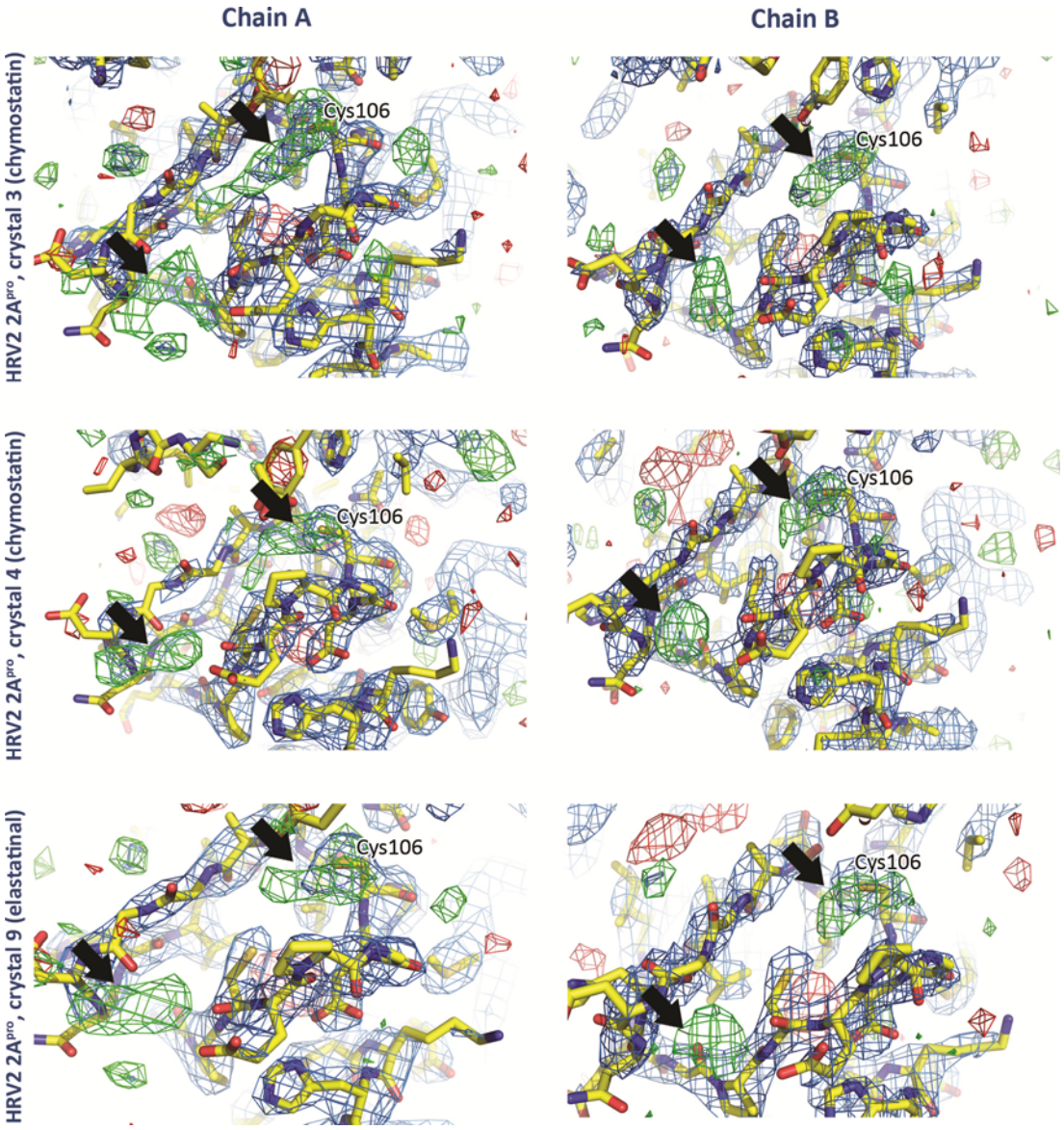


Figure 4-69 Electron densities and protein models at the active sites of HRV2 2A^{pro} soaked with either chymostatin or elastatinal. $2F_0$-F_c maps are depicted in blue at 2σ. Difference maps are shown at 3σ in green and red. The protein model is shown as sticks with carbon in yellow, oxygen in red, nitrogen in blue and sulfur in light brown. Arrows indicate extra electron density possibly arising from the presence of the inhibitors.

	Residue range	Location
Region 1	15-18	End of strand cI and helix I
Region 2	84-86	Loop connecting strands bII and cII (di-tyrosine flap)
Region 3	101-107	Loop connecting strands cII and dII
Region 4	119-125	End of strand eII and subsequent loop

Table 4-11 Definition of excluded regions for the calculation of omit maps. Regions in close proximity to the expected binding site of the inhibitors were chosen for the calculation of omit maps.

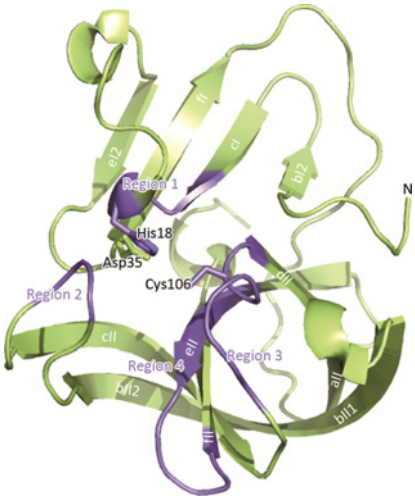


Figure 4-70 Illustration of excluded regions for the calculation of omit maps. The apo structure of HRV2 $2A^{pro}$ (PDB 2HRV) is shown as cartoon representation. Regions 1 to 4 as defined in Table 4-11 are shown in blue. Illustration rendered with PyMol (DeLano, 2002).

5 DISCUSSION

5.1 PV 2A^{PRO} IS INSOLUBLE WHEN EXPRESSED IN *E. COLI*

The purification of PV1 2A^{pro} has remained an unmet challenge for picornavirologists. The poor solubility of the protease has hampered the production of batches in the mg scale in the absence of detergent or solubility tags. Toyoda *et al.* managed to purify 50µg of protease from 5 liters of PV1 infected HeLa cell suspension using detergent (Toyoda *et al.*, 1986). In the 1990s, expression of recombinant PV 2A^{pro} in *E. coli* was used to obtain enough protease to perform cleavage assays on eIF4G and the TATA-binding protein. In these experiments 2A^{pro} was either fused to MBP as a solubility tag (Novoa *et al.*, 1997) or purified under denaturing conditions followed by refolding (Yalamanchili *et al.*, 1997). However, the expression and purification of PV 2A^{pro} in sufficient quality and quantity for structural analysis has been unsuccessful.

PV1 and CVB4 2A^{pro} share a sequence identity of 57% (73% similarity) (Figure 5-1). Furthermore, protein solubility predictions did only slightly favor CVB4 2A^{pro} over PV1 2A^{pro}. An algorithm according to Wilkinson *et al.* predicts a chance of solubility of 58.2% for PV1 and 68.3% for CVB4 2A^{pro} when expressed in *E. coli* (Wilkinson & Harrison, 1991). Moreover, the PROSO II prediction server ranks both proteases as soluble with scores of 0.706 for PV1 2A^{pro} and 0.758 for CVB4 2A^{pro} (Smialowski *et al.*). In spite of these similarities, CVB4 2A^{pro} behaved superbly in terms of solubility (protein concentrations of 1mM could be achieved (see chapter 4.1.4)) whereas the solubility of PV1 2A^{pro} was poor. What could be the reasons for the poor solubility of PV1 2A^{pro}?

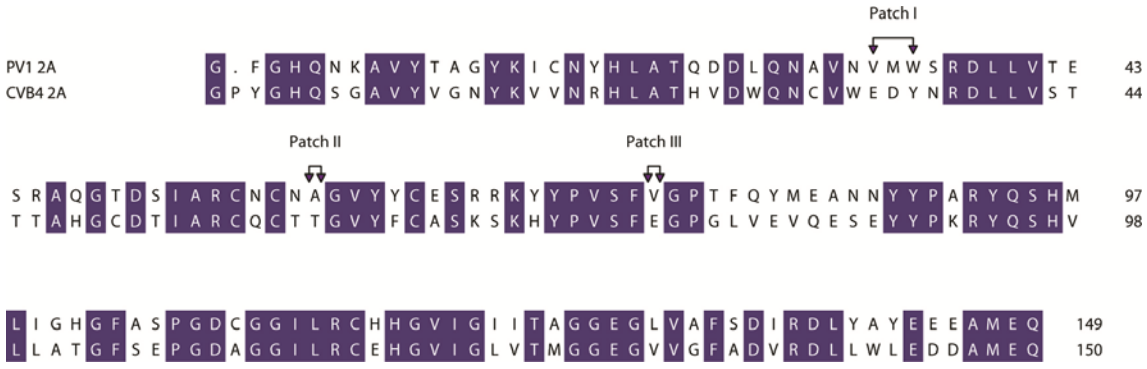


Figure 5-1 Sequence alignment of PV1 and CVB4 2A^{pro}. The amino acid sequences of PV1 and CVB4 2A^{pro} were aligned using ClustalW (Larkin *et al.*, 2007). Positions with identical amino acids are highlighted by reverse video in blue. The three patches that possibly cause the differences in solubility of CVB4 and PV1 2A^{pro} are indicated. Rendered with ALINE (Bond & Schuttelkopf, 2009).

The main determinant for the solubility of globular proteins in aqueous solution is the distribution of charged and hydrophilic residues on its surface. Dependent on the pH of the solution, a homogenous distribution of charges and hydrophilic groups enables interaction with water molecules. The theoretical pI values of CVB4 and PV1 $2A^{pro}$ are 4.90 and 5.79, respectively. The pH values of the buffers chosen in the experiments presented here (pH 8.0-8.5) should allow a negative net charge of the proteins. We therefore wanted to examine the possible existence of hydrophobic patches on the protein surface. In the absence of structural data for PV1 $2A^{pro}$, this was achieved by using a homology model. To this end, the polypeptide chain of PV1 $2A^{pro}$ was modeled on the lowest energy state of the CVB4 $2A^{pro}$ NMR-structure (PDB 1Z8R) using the automated Swiss-Model web-platform (Arnold *et al.*, 2006; Kiefer *et al.*, 2009; Peitsch *et al.*, 1995). Electrostatics of the surfaces were generated using the APBS plug-in in the molecular graphics program PyMol (Baker *et al.*, 2001; DeLano, 2002) (Figure 5-2 A). Furthermore, the hydrophobicity scale of Kyte & Doolittle was used to calculate mean hydrophobicity profiles of the two proteases (Figure 5-2 B) (Kyte & Doolittle, 1982).

Differences in electrostatic properties are easily detected at the face lying opposite to the active site. Inspection of the surfaces allows the definition of three patches with profound electrostatic variance. Patch I spans the residues ^{34}EDY^{36} (according to CVB4 nomenclature). The two negatively charged residues Glu and Asp as well as the Tyr side chain with its hydrophilic hydroxyl group are replaced by the hydrophobic amino acids ^{33}VMW^{35} in PV1 $2A^{pro}$ (see Figure 5-1 and Figure 5-2 A). Patch II is influenced by the presence of an Ala residue at position 59 in PV1 $2A^{pro}$, with a Thr residue being found at the corresponding position 60 in CVB4 $2A^{pro}$. Finally, patch III substantially varies by the presence of a glutamate residue at position 77 in CVB4 $2A^{pro}$ which is replaced by the hydrophobic amino acid valine in PV1 $2A^{pro}$. Patches I and III can also easily be identified on the mean hydrophobicity profiles (Figure 5-2 B). These two regions are the only stretches in which PV1 $2A^{pro}$ exhibits considerably higher mean hydrophobicity scores. Next, we investigated whether the three patches would be conserved amongst 2A proteases of the three PV serotypes (data not shown). Indeed, patch I is nearly completely conserved and is composed of hydrophobic amino acids in all three serotypes. Ala59 of patch II is exchanged to Thr in PV2 $2A^{pro}$ as well as Val76 of patch III. Taken together, patches I-III form a hydrophobic region on the back of the PV 2A proteases and might explain why PV $2A^{pro}$ is so poorly soluble. It remains unclear whether these differences in the electrostatic surface are connected to a special biological function. Mutational analysis of the involved amino acids will be needed to clarify their role in solubility.

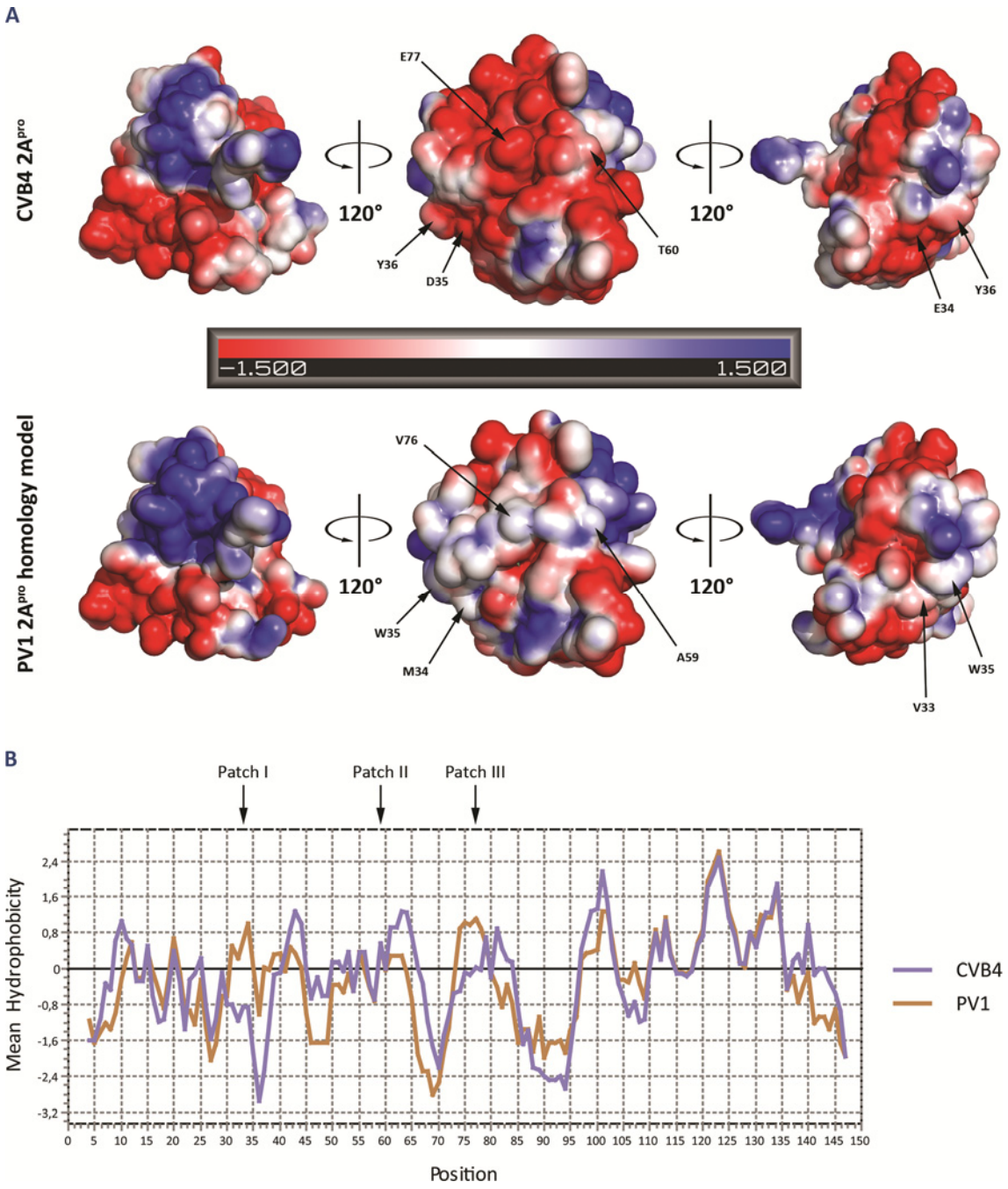


Figure 5-2 Comparing the 2A proteases of CVB4 and PV1 for factors relevant for protein solubility. **(A)** A homology model of PV1 2A^{pro} compiled using the lowest energy state of the CVB4 NMR structure (PDB 1Z8R) was used to compare the electrostatic surfaces of CVB4 and PV1 2A^{pro} calculated with the program APBS (Baker *et al.*, 2001). Residues with profound differences in electrostatics are labeled. **(B)** Comparison of the mean hydrophobicity profiles of CVB4 and PV1 2A^{pro}. Profiles were calculated using the Kyte & Doolittle scale with a scanning window size of 7 amino acids (Kyte & Doolittle, 1982).

5.2 SNAP-SHOOTING 2A^{PRO} DURING SELF-PROCESSING – A FAILED ATTEMPT

To our knowledge, no structural data on the intramolecular self-processing of a protease on its own N-terminus in order to free itself from a polyprotein is available. In contrast, a high resolution X-ray structure of the VP4 protein from *Tellina Birnavirus* with its own C-terminus bound in the active site has been published (Chung & Paetzel, 2011). Furthermore, the X-ray structure of the self-processing *Porcine Reproductive and Respiratory Syndrome Virus* protease Nsp1α was solved with its C-terminal extension bound to the substrate binding cleft (Sun *et al.*, 2009). These structures demonstrate the general possibility of snap-shooting a viral protease in the state of freeing itself from the viral polyprotein in *cis*.

Despite this proof of principle, crystallization of such a conformation remains a challenge. Not only one has to overcome the general bottleneck of crystallization, but also one has to face the problem of stabilizing the conformation of interest. This thesis describes crystallization trials on a number of constructs designed to overcome this problem. However, despite the large amount of crystal screens performed, all of our attempts to obtain diffraction-quality crystals failed. What might have been the reasons for this failure?

A number of algorithms that predict the likelihood of crystallization for a given amino acid sequence are available. We used the web-servers of XtalPred (Slabinski *et al.*, 2007a; Slabinski *et al.*, 2007b) and ParCrys (Overton *et al.*, 2008) to predict the crystallization propensities of our constructs (see Table 5-1). However, predictions by these two programs were highly diverse. Whereas ParCrys had a clear tendency to predict low crystallization likelihood (with exception of just three constructs), XtalPred ranked all constructs in high propensity groups. Furthermore, no tendency for a group of constructs could be observed within the results of one prediction method.

One main determinant of protein crystallization is the protein itself (Longenecker *et al.*, 2001). The distribution of charged, hydrophilic and hydrophobic residues on the surface determine whether suitable crystal contacts can be formed that favor the formation of well ordered crystals. Protein constructs can be altered in several ways in order to maximize the likelihood of crystallization.

Already in 1972, it was suggested that the problem of proteins recalcitrant to crystallization could be evaded by the use of homologous proteins (Campbell *et al.*, 1972). The idea has been proven successful in innumerable cases since. For this reason, we used 2A^{pro}

constructs from different enteroviruses. First experiments in the lab with HRV14 $2A^{pro}$ resulted in protein that seemed to be unfolded (Sousa, 2008). For this reason, we chose HRV2, PV1 and CVB4 $2A^{pro}$ for our experiments. However, none of the proteases yielded diffraction-quality crystals. Future experiments could make use of the availability of 2A sequences from a vast number of enteroviruses. A large scale screen on many of these proteases should allow the detection of constructs with satisfactory properties in solubility and crystallization propensity.

		ParCrys prediction	XtalPred prediction
PV1	$2A^{pro}$ (C109A) aa 8-142	High-Scoring	Optimal
	$2A^{pro}$ active	Recalcitrant	Suboptimal
	$VP1_8$-$2A^{pro}$ (C109A)	Recalcitrant	Suboptimal
	His_6-$VP1_8$-$2A^{pro}$ (C109A)	Recalcitrant	Optimal
	MBP-$VP1_8$-$2A^{pro}$ (C109A)	High-Scoring	Suboptimal
	MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker	Recalcitrant	Suboptimal
	MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker-His_6	Recalcitrant	Optimal
HRV2	$2A^{pro}$ active	Amenable	Suboptimal
	$VP1_9$-$2A^{pro}$ (C106A)	Recalcitrant	Optimal
CVB4	$2A^{pro}$ (C110A) aa 7-143	Recalcitrant	Optimal
	$2A^{pro}$ active	High-Scoring	Optimal
	$VP1_5$-$2A^{pro}$ (C110A)	Amenable	Optimal
	$VP1_8$-$2A^{pro}$ (C110A)	Amenable	Optimal
	$VP1_{12}$-$2A^{pro}$ (C110A)	Recalcitrant	Optimal
	$VP1_{20}$-$2A^{pro}$ (C110A)	Recalcitrant	Optimal

Table 5-1 Predictions on crystallization propensities according to the algorithms of ParCrys and XtalPred. ParCrys ranking from low to high crystallization propensity: recalcitrant, amendable and high-scoring. XtalPred ranking from low to high crystallization propensity: very difficult, difficult, average, suboptimal and optimal.

A different approach to manipulate the surface properties of a protein is reductive lysine methylation. A study by Walter *et al.* using a set of 10 proteins that could not be crystallized before showed that four of them yielded diffraction-quality crystals upon reductive lysine methylation, from which three led to the solution of a high resolution structure (Walter *et al.*, 2006). Unfortunately, the application of this method on one of our protein constructs did not enhance crystal formation in our hands (see chapter 4.1.4.6).

Flexible stretches of a protein are believed to disturb or even hinder the integration of a molecule into a crystal lattice. The N-terminal substrate extensions on our constructs were clearly problematic in this way. An NMR study by Baxter *et al.* aiming to investigate self-processing of CVB4 failed to record signals for the entire $VP1_8$ extension on the N-terminus as well as the first residues of the protease (Baxter *et al.*, 2006) (see chapter 2.4.2.2) due to internal flexibility of this part of the protein. Nevertheless, we rationalized that crystal contacts could possibly favor the occupancy of the substrate binding cleft with substrate thereby also reducing the degree of N-terminal motion in our constructs. However, our failure in obtaining crystal hits places strong doubts on our hypothesis. Even the usage of limited proteolysis in order to define an optimal length of the N-terminal extensions did not lead to a crystallizeable construct (see chapter 4.1.4.5).

We therefore aimed at obtaining crystals from a construct that was truncated on the N- as well as the C-terminus in order reduce internal flexibility (see chapter 4.2.1.3). A number of examples document the success of such an approach (De Kerpel *et al.*, 2006; Yeh *et al.*, 1996). However, the lack of over-expression in *E. coli* cells hampered the purification and crystallization of these proteins. Another procedure to obtain protein samples that are amendable to crystallization is the expression of single domains of a protein. However, $2A^{pro}$ is a small protein of no more than 150 amino acids composed of only two domains. The catalytic triad is split on these two domains. Expression of just one of these domains is therefore of no biological interest.

Finally, the use of fusion proteins has recently strongly been established. Whilst proteins of interest are mainly cleaved from their tags after usage as solubility enhancer or for purification, a few examples of MBP fusion protein crystals have been published so far (Center *et al.*, 1998; Liu *et al.*, 2001). The fusion of the C-terminal helix of MBP with an N-terminal helix on the protein of interest has been reported to favor crystallization of such a construct due to reduced flexibility of the linker. Indeed, we were able to obtain crystal hits with our PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) construct (see chapter 4.1.1.3). However, the crystals were unstable and could not be reproduced. The long and flexible linker between MBP and PV1 $2A^{pro}$ could be a major problem in crystallization of the fusion protein. As cleavage from the fusion protein was not possible due to the poor solubility of PV1 $2A^{pro}$, we aimed to shorten the linker to reduce the flexibility (see chapter 4.1.1.4 and 4.1.1.5). Unfortunately, problems in the purification of these constructs precluded intensive crystal screening.

5.3 CHYMOSTATIN AND ELASTATINAL: SUITABLE INHIBITORS FOR STRUCTURAL STUDIES OF 2A^{PRO} SUBSTRATE BINDING?

The two canonical inhibitors chymostatin and elastatinal were described to be efficient against chymostatin and elastase, respectively (Umezawa *et al.*, 1970; Umezawa *et al.*, 1973). Inhibition of HRV2 2A^{pro} by these two molecules was demonstrated using peptide cleavage assays and IC_{50} values of 25µM (chymostatin) and 55µM (elastatinal) have been calculated (Sommergruber *et al.*, 1992). No data is available on the affinity of these two inhibitors towards the enteroviral 2A^{pro} active site. However, the general affinity of HRV2 2A^{pro} towards its substrates does not seem to be very high. The K_m of HRV2 2A^{pro} towards a peptide spanning the P8-P8' region was measured to 540µM (Sommergruber *et al.*, 1992). Given the smaller size (see Figure 4-67 C) and the non-optimized side chains of the inhibitors, the affinity of chymostatin and elastatinal may not exceed the value for the P8-P8' peptide. However, the two inhibitors covalently bind to the active site of the protease, therefore possibly counteracting the relatively low affinity.

The high resolution structures of a number of protease-inhibitor complexes with chymostatin have been solved. Most interestingly, the structure of *Streptomyces griseus* protease A bound to chymostatin has been refined to 1.8Å (Delbaere & Brayer, 1985). The model of this simple chymotrypsin like protease revealed a tetrahedral intermediate upon nucleophilic attack of the active site serine. Furthermore, the conformation of the inhibitor and the accommodation of its side chains correlates well with the expected mode of substrate binding for HRV2 2A^{pro}. By superimposing this structure with the structure of apo HRV2 2A^{pro} (PDB 2HRV, molecule B) using DaliLite (Hasegawa & Holm, 2009), a crude model of how chymostatin could be accommodated within the HRV2 2A^{pro} substrate binding cleft was obtained (Figure 5-3). All of the chymostatin side-chains are directed towards their respective pockets with the side chain of P3 facing towards the solvent. However, whereas a Phe residue was shown to be relatively well accepted at P1, the Leu and Phe side chains at P2 and P4 strongly reduce relative cleavage efficiency or even abrogate the cleavage of peptides (Sommergruber *et al.*, 1994a; Sommergruber *et al.*, 1992). In contrast, the equivalent backbone of elastatinal is equipped with the side-chains of Leu (P1), Gln (P2) and Ala (P4). The even worse acceptance of these residues in peptide cleavage assays may be reflected in the somewhat higher IC_{50} value of elastatinal compared to chymostatin.

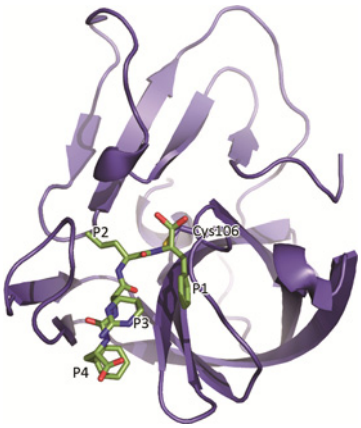


Figure 5-3 Crude model of a chymostatin- HRV2 2A^{pro} complex. The high resolution structure of *Streptomyces griseus* protease A complexed with chymostatin (PDB 1SGC) was superimposed on molecule B of the apo HRV2 2A^{pro} crystal structure (PDB 2HRV) using DaliLite (Hasegawa & Holm, 2009). The coordinates of the chymostatin molecule are shown as green (carbon), red (oxygen) and blue (nitrogen) sticks together with a cartoon representation of HRV2 2A^{pro}.

As described in chapter 4.3.2.4, we aimed to solve a high resolution structure of HRV2 2A^{pro} in complex with either chymostatin or elastatinal. Apo HRV2 2A^{pro} crystals were soaked in solutions of an inhibitor and full X-ray datasets were collected. Initial phases were obtained using molecular replacement. Indeed, additional density could be observed within the active site (see Figure 4-69). The quality of this density was however too poor to place the inhibitors. What might be the reasons for this failure?

Against our postulation the additional density could have emerged without the presence of inhibitor, which was generally strongest directly at the active site cysteine. Hence, the carry-over of DMSO by addition of inhibitors could have caused oxidation of the active site cysteine (Lipton & Bodwell, 1973). However, HRV2 2A^{pro} was crystallized in the presence of 10mM BME. Whether such a reaction would be possible in the presence of the reducing agent remains unclear. Furthermore, inspection of other cysteine residues within 2A^{pro} did not suggest oxidation of these side chains. Moreover, oxidation of Cys106 does not explain the presence of other additional electron density blobs within the substrate binding area. Another reason for the poor quality of the additional electron density could be low occupancy of the substrate binding site with inhibitor.

Finally, the probably low affinities of the inhibitors and their suboptimal side chains may give the most likely explanation. The inhibitors may intrude the substrate binding cleft followed by nucleophilic attack of their carbonyl at P1 by Cys106. The resulting covalent bond may allow the tracing of electron density at this site. However, the remainder of the molecule might not be stabilized in the active site cleft due to its low affinity and/or steric clashes of its side chains within the protease pockets. Hence, the molecule is fixed at P1 by a covalent bond but might wobble within or on top of the active site cleft leading to a failure in obtaining electron density for parts of the molecule. This hypothesis is further supported by the fact that electron density of the di-tyrosine flap (Glu79-Lys88) in 2A^{pro} was of poor quality. This loop connecting the bII2 and cII

strand of the C-terminal β-barrel contains Tyr85, which has been proposed to rotate away from its original position along its C_α-C_β bond in order to generate space for the accommodation of a P2 residue (Petersen *et al.*, 1999). The poor electron density quality of this loop might reflect flexibility of this region induced by steric clashes with the bulky side chains of leucine and glutamine at the P2 positions of chymostatin and elastatinal respectively. The bII2 to cII loop also exhibits the highest B-factors within the entire 2A protease.

5.4 A SUBSTRATE BINDING MODEL FOR HRV2 2A^{PRO}

As we could not reach our main goal to produce a high resolution structure of 2A^{pro} bound to substrate, we aimed to create a model for substrate binding of HRV2 2A^{pro}. Such a model could give new insights and ideas on how 2A^{pro} recognizes its substrates. To this end, we modified a strategy of model building that has been described (Petersen *et al.*, 1999) in order to produce a substrate binding model spanning the P5-P2' region. We used the DaliLite Server (Holm & Sander, 1993) to superimpose the structure of *Streptomyces griseus* protease B (SGPB) bound to the third domain of the turkey Kazal-type ovomucoid inhibitor (OMTKY3) (PDB 3SGB) with molecule B of the apo HRV2 2A^{pro} structure (PDB 2HRV). The two protease structures (especially their C-terminal β-barrels) superimpose reasonably well with 130 equivalent residues and an root mean square (RMS) of C_α atoms of 3.1Å (Figure 5-4 A). The most profound differences within the substrate binding area are found at the loop connecting strands bII2 and cII. Moreover, the complete strand aI and large parts of strand bI2 found in SGPB are missing in HRV2 2A^{pro}. However, the arrangement of the catalytic triads seems equivalent. For this reason, we used the coordinates of residues 14-20 of OMTKY3 together with the HRV2 2A^{pro} coordinates as our starting model. The side chains of the peptide were adapted to fit the P5-P2' substrate in HRV2 2A^{pro} self-processing (IleIleThrThrAla*GlyPro) (Figure 5-4 B).

The conformation and location of the peptide within the substrate binding cleft seems plausible. The carbonyl of the cleavable P1-P1' peptide bond is positioned above the active site cysteine. The peptide adopts a more or less extended conformation. Ile at P4 is buried within a deep pocket whereas Thr at P3 faces the solvent. Thr at P2 is buried underneath the di-tyrosine flap and Ala at P1 is positioned between the eII strand and the loop containing the active site Cys106. However, running the model on the MolProbity server (Chen *et al.*) in order to check geometry and contacts revealed several clashes (>0.4Å overlap) between the side chains of

substrate and the protease with a clash-score of 31.56 (Figure 5-5 A). The isoleucine residues at P5 and P4 clashed with Ile80, Gln81, Ala122 and the backbone of Gly123. The P2 Thr was found to overlap with Tyr85, Tyr86 as well as Thr121. Furthermore, the P1 Ala is positioned too close to Cys106 and the P' residues are in contact to His18, Leu19 and Gly109.

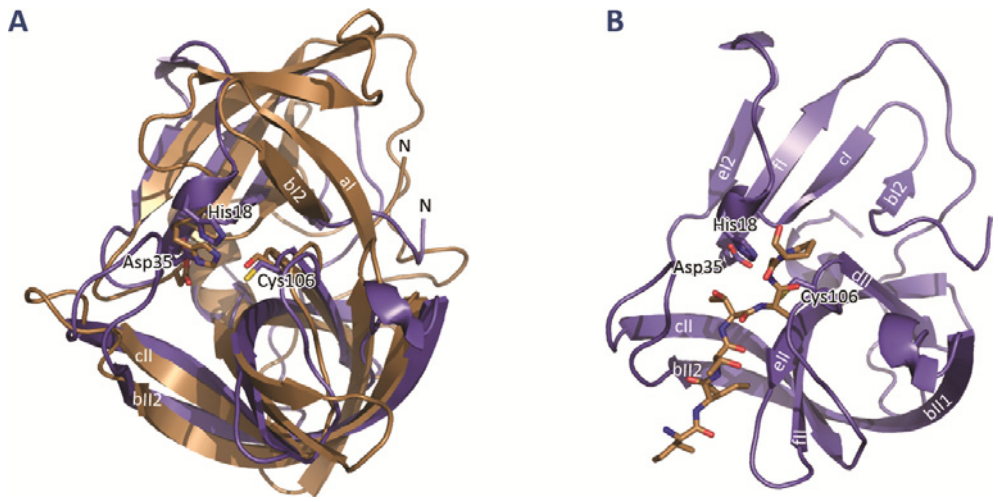


Figure 5-4 Construction of a substrate binding model for HRV2 2A^{pro}. (A) Chain B of the apo HRV2 2A^{pro} crystal structure (PDB 2HRV) was superimposed on SGPB bound to OMTKY3 (PDB 3SGB) using the DaliLite server. The proteases are shown as cartoon representations in blue (HRV2 2A^{pro}) and brown (SGPB) with the active site residues shown as sticks (labeling according to HRV2 2A^{pro}) **(B)** The structure of HRV2 2A^{pro} is shown together with residues 14-20 of the OMTKY3 (side chains have been changed to IleIleThrThrAla*GlyPro (the asterisk indicates the cleavable bond)) in brown sticks. Strands are labeled according to their homology to SGPB. Rendered with PyMol (DeLano, 2002).

Hence, we decided to refine the position and geometry of the substrate by minimizing the energy of the model. To this end, we used the YASARA energy minimization server (Krieger *et al.*, 2009) to obtain optimized coordinates. The server uses the self-parameterizing and knowledge-based YASARA force field that is derived from the AMBER force field family. The resulting coordinates were again analyzed with MolProbity. Indeed, steric clashes were significantly reduced within this model resulting in a clash-score of 0.44 (compared to 31.56 for the non-minimized model) and the elimination of all severe clashes with overlaps >0.4Å (Figure 5-5 B). Energy minimization triggered deeper penetration of the peptide into the binding cleft as demonstrated by an Rmsd of 1.22Å calculated on the C_α atoms of the peptides (Figure 5-5 C). The substrate binding cleft itself is being adapted by a 101° rotation of Tyr85 about its C_α-C_β bond (Figure 5-5 D). Our results herewith strengthen the hypothesis of Petersen *et al.* (Petersen *et al.*, 1999).

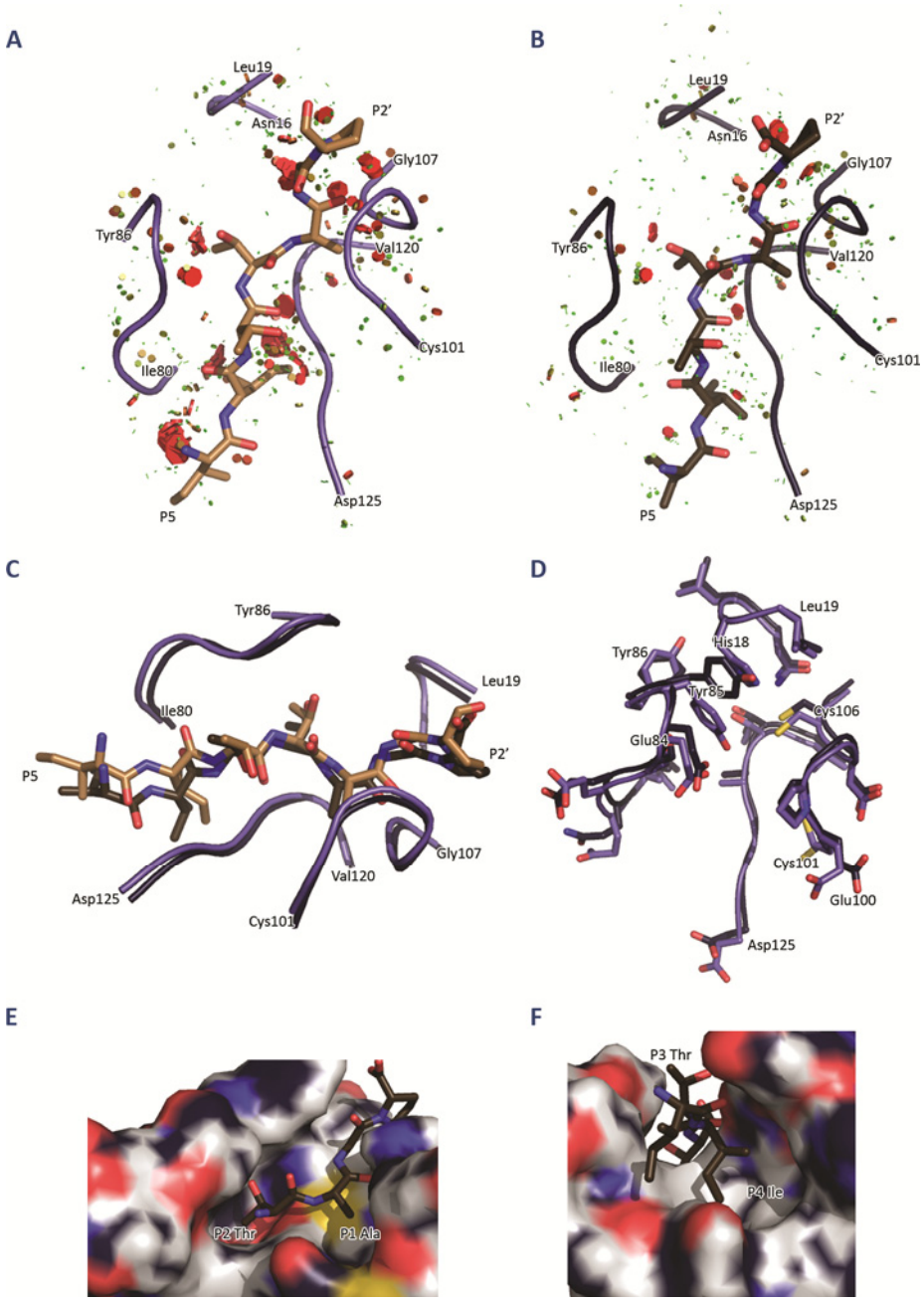


Figure 5-5 Structural features of an HRV2 2A^{pro} substrate binding model minimized with the YASARA force field. (A) Analysis of steric clashes between the peptide and the substrate binding cleft in the crude HRV2 2A^{pro} substrate binding model. The magnitude of the clashes is illustrated by the size of the red discs. Side chains of the substrate binding cleft are not drawn for simplicity. **(B)** Analysis of steric clashes between peptide and substrate binding cleft after energy minimization using the YASARA force field (Krieger *et al.*, 2009). **(C)** Side view on the overall position of the peptide before (protease in light blue, peptide in light brown) and after (protease in dark blue, peptide in dark brown) energy minimization. **(D)** Analysis of residues within the substrate binding cleft. The protease binding cleft is depicted as cartoon with side chains as sticks in light blue (before energy minimization) and dark blue (after energy minimization). The substrate is not drawn for simplicity. **(E+F)** Accommodation of the P4-P1 side chains within their substrate binding pockets. Substrate is shown as sticks whereas the substrate binding cleft is illustrated via its Connolly surface.

How are the single residues of the substrate accommodated within the binding cleft of the protease? Ala at P1 is buried within a pocket formed by the backbone of Glu102, Pro 103, Gly104, Ala122 and Gly123 as well as the side chain of Cys 101. The pocket leaves plenty of space for the accommodation of larger residues (Figure 5-5 E). The rotation of Tyr85 allows Thr at P2 to enter its pocket without major steric hindrance and is now close to Ser83 to which it possibly forms a hydrogen bond. Ile at P4 is deeply buried within a pocket built of Tyr78, Ile86, Leu94, Ile96 and Ala122 (Figure 5-5 F). The model correlates well with the model of Petersen *et al.* and nicely illustrates some features of the HRV2 $2A^{pro}$ substrate specificity. However, we aimed to test whether the model was suitable to tackle some unanswered questions of $2A^{pro}$ specificity.

Recently, we observed the inability of HRV2 $2A^{pro}$ to self-process on the sequence found on the HRV14 VP1-$2A^{pro}$ junction (AspIleLysSerTyr*GlyLeu) (Neubauer *et al.*, 2013). All of the amino acids differing from the original IleIleThrThrAla*GlyPro self-processing site were found to be well accepted by HRV2 $2A^{pro}$ when introduced one by one into the VP1-$2A^{pro}$ junction of HRV2. A combination of P2 Ser, P1 Tyr and P2' Leu was necessary to obtain complete inhibition of HRV2 $2A^{pro}$ self-processing. To our surprise, HRV2 $2A^{pro}$ was able to process the site at wt levels when the P1 Tyr was mutated back to the wt Ala. We therefore, mutated the peptide in our substrate binding model to IleIleThrSerTyr*GlyLeu and again minimized the model via the YASARA energy minimization server (Krieger *et al.*, 2009) (Figure 5-6).

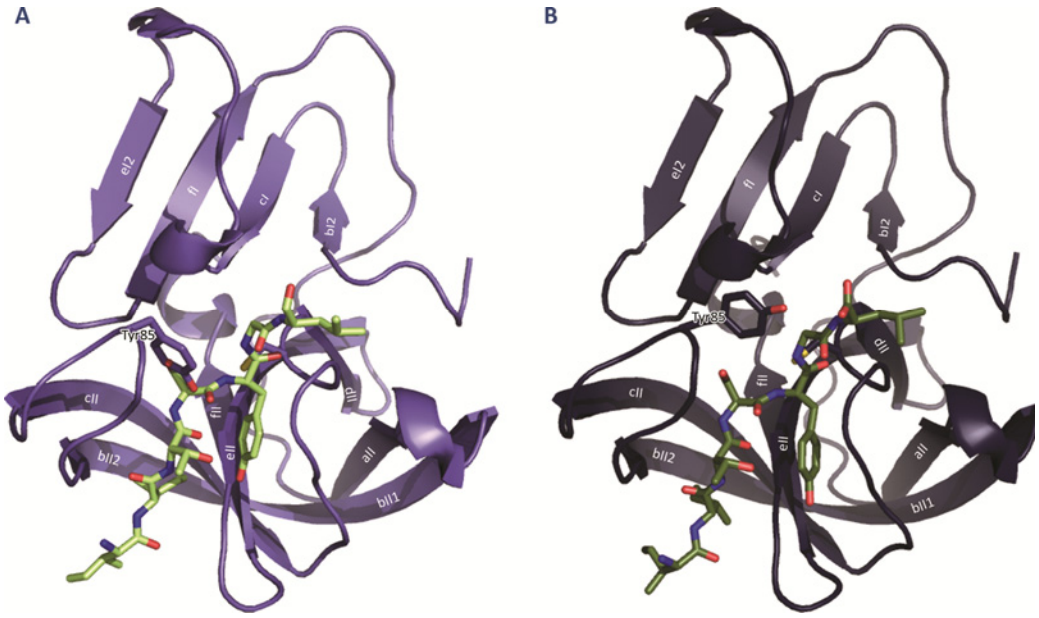


Figure 5-6 Model of HRV2 $2A^{pro}$ bound to an alternative peptide before and after energy minimization. (A) Substrate binding was modeled as described above. The side chains of the peptide were adjusted to fit the sequence IleIleThrSerTyr*GlyLeu. Protease is depicted as cartoon in light blue whereas the substrate is shown as light green sticks. **(B)** The model was energy minimized using the YASARA server (Krieger *et al.*, 2009). Protease and substrate are depicted in dark blue and dark green respectively. Representations rendered with PyMol (DeLano, 2002).

Again the geometry and clash-score of the model was drastically improved by the energy minimization and the peptide penetrated deeper into the substrate binding cleft. Moreover, Tyr85 is again rotated away from its stacked position with His18 in order to allow the P2 Ser side chain to enter its pocket. P2 Ser is found in a similar position and orientation than the wt Thr. In addition, the P1 Tyr is slightly rotated about its C_α-C_β bond perfectly orienting it within the S1 pocket. Finally, the position of Leu at P2' has to be judged with skepticism as the conformation in the model probably would not allow the N-terminus to span the region between the first sheet bI2 and the active site during self-processing. Hence, the model explains very well how the single amino acid changes in the P2, P1 and P2' positions of the HRV2 VP1-2A^{pro} junction can be accepted. However, we are not able to clearly explain why the change of all three residues leads to inhibition of self-processing. We speculate that Ser at P2, due to its smaller size compared to Thr, might be less powerful in forcing Tyr85 to rotate away from its original position. Additionally, the simultaneous presence of the bulky side chains at the P1 and P2' positions might interfere with proper substrate binding. Moreover, the model cannot clarify why HRV2 2A^{pro} can accept Arg at P1 during self-processing while HRV14 2A^{pro} cannot (Sousa *et al.*, 2006). We used our model to assign a color code reflecting the conservation between HRV2 and HRV14 2A^{pro} in order to detect differences between the two proteases (Figure 5-7). Despite the low conservation between HRV2 and HRV14 2A^{pro} on the amino acid sequence level (39% identity), the P1 pocket is well conserved. Indeed, the only variable amino acids are HRV2 Cys101 (Ala104 in HRV14) and the loop connecting the strands eII and fII. Position 101 has been investigated by Sousa *et al.* (Sousa *et al.*, 2006). However, HRV14 2A^{pro} A104C or A104S mutants were able to only partially restore activity on substrates with Arg at P1. Moreover, the side chains of the eII-fII loop (^{125}DNH^{127}) are facing the solvent and are oriented away from the P1 pocket.

For all these reasons, binding of substrates by 2A^{pro} might not be determined by the recognition of specific sequences or chemical groups. We speculate that specificity is rather dependent on the spatial properties of the substrate as defined by the overall combination of amino acids at the P4-P2' positions.

Why is Gly at P1' an absolute prerequisite for processing in *cis* whereas substrates with Ala at P1' can be accepted in *trans* (Neubauer *et al.*, 2013; Park *et al.*, 2010)? Petersen *et al.* suggested that any side chain at P1' would clash with the side chain of Leu19 (Petersen *et al.*, 1999). However, we have recently demonstrated that a L19A mutant of HRV2 2A^{pro} still self-processes very poorly on a substrate with Ala at P1' (Neubauer *et al.*, 2013). We therefore speculated that Leu19 is at least not the only residue determining the specificity at P1'.

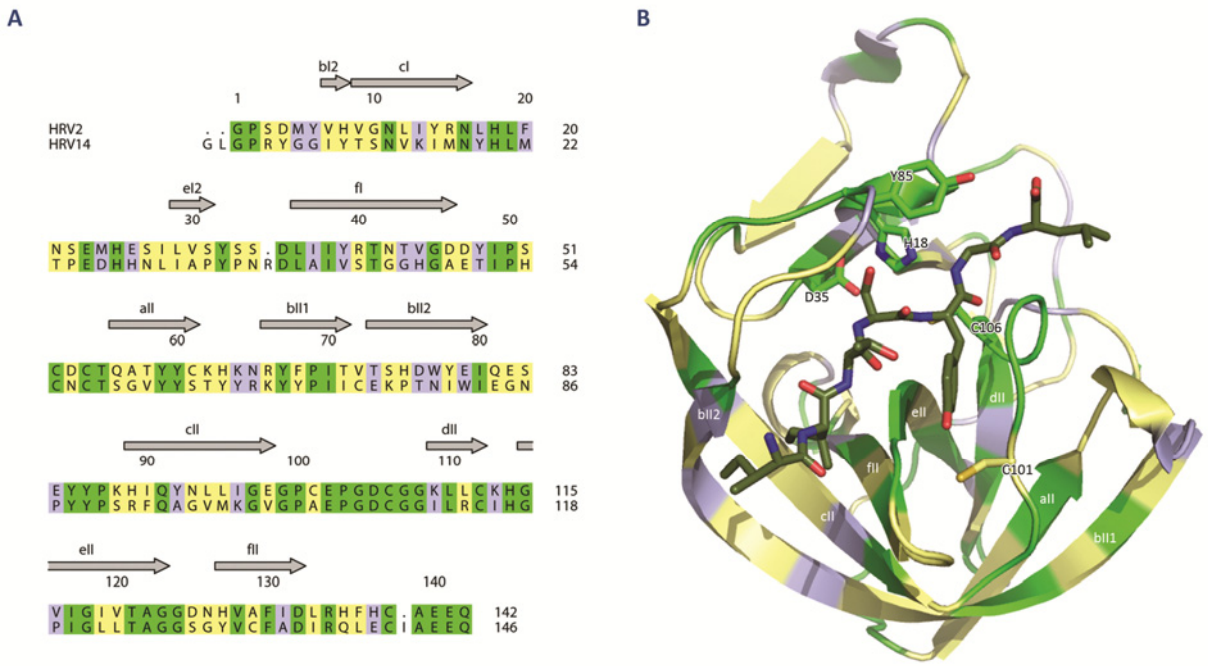


Figure 5-7 Sequence conservation between HRV2 and HRV14 2A^{pro}. (A) The amino acid sequences of HRV2 and HRV14 2A^{pro} have been aligned using ClustalW (Larkin *et al.*, 2007). A color code representing conservation was applied (green: conserved, yellow: similar, blue: unrelated). **(B)** The color code was assigned onto the HRV2 2A^{pro} substrate binding model. Rendered with ALINE (Bond & Schuttelkopf, 2009) and PyMOL (DeLano, 2002).

Hence, we adapted our substrate binding model and replaced Gly at P1' with Ala (Figure 5-8). At least in the conformation of our model the methyl group is found close to the side chains of the rotated Tyr85 and His18 rather than the side chain of Leu19. Clashes with His18 could result in a subtle reorientation of His18. Hence, the conformation of the catalytic triad would be disrupted leading to an inactivation of the protease. However, mutation of His18 to a less bulkier amino acid in order to provide more space for the methyl group would result in the complete inactivation of the protease. Hence, this idea cannot be followed up by using cleavage assays of mutant protease. Moreover, these thoughts are in contradiction with the finding that HRV2 2A^{pro} can cleave substrates with Ala at P1' in *trans*.

Tyr85 is found in even closer vicinity to P1' Ala than His18. Inspection of the model provides no obstacles against further rotation of Tyr85 about its C_α-C_β bond. Hence, clashes of the P1' Ala side chain with Tyr85 could be avoided allowing HRV2 2A^{pro} to cleave substrates with Ala at P1' in *trans*. Nevertheless, we also aimed to clarify why Ala cannot be accepted at P1' during self-processing. The answer to this question might be found in the mode of substrate binding. During self-processing the substrate is covalently bound to the N-terminus of the protease. Hence, substrate binding has to follow a step by step mechanism in which the amino acids are

more or less accomodated one by one in an ordered manner (…-P4'-P3'-P2'-P1'-P1-P2-P3-P4-…). Consequently, Ala at P1' would bind into its pocket before Tyr85 is forced to rotate by the presence of the P2 Thr thereby blocking the trajectory for the rotation of Tyr85. Substrates cleaved in *trans* are *per se* not limited to such a mode of binding. As a result, Tyr85 can first be pushed out by the P2 residue followed by possible binding of the P1' residue and further rotation of Tyr85. Tyr85 is strictly conserved within 2A proteases of enteroviruses. Mutation of Tyr85 to a smaller residue such as Thr was shown to severely impair self-processing (Sommergruber *et al.*, 1997). Hence, self-processing assays using mutant HRV2 $2A^{pro}$ may not be valuable to clarify this issue.

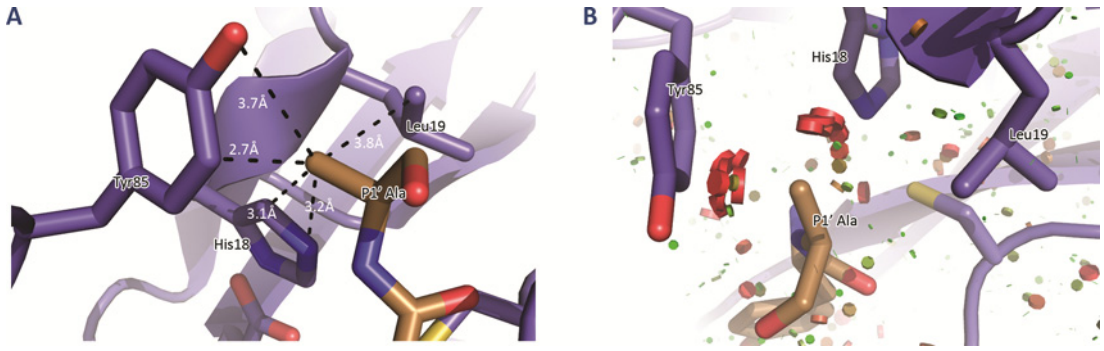


Figure 5-8 Analyzing the HRV2 $2A^{pro}$ substrate binding model with alanine at P1'. HRV2 $2A^{pro}$ is depicted as cartoon in blue with side chains of interest shown as sticks. Substrate is rendered as brown stick model. **(A)** Distances between C_β (P1' Ala) and C_ε/O (Tyr85), C_δ/N_ε (His18), C_δ (Leu19) are given. **(B)** The magnitude of steric clashes between the methyl group of alanine at P1' and the side chains of Tyr85, His18 and Leu19 was analyzed and is depicted by the size of red discs. Rendered with PyMol (DeLano, 2002).

5.4.1 Implications for the development of anti $2A^{pro}$ compounds

The aim of this thesis was to produce high resolution structures of enteroviral 2A proteases bound to substrates in *cis* or *trans*. Such structures should help to drive knowledge-based inhibitor design. However, the desired information could not be obtained and hypotheses on how HRV2 $2A^{pro}$ recognizes its substrates were obtained using substrate binding models. We speculated that $2A^{pro}$ substrate specificity is determined by combinatorial spatial requirements of sites on the primed as well as the unprimed side of the substrate rather than on the specific recognition of a sequence. A similar situation was pointed out for HIV-1 protease (Prabu-Jeyabalan *et al.*, 2002). The ten natural substrates of HIV-1 protease share little sequence homology which precluded exact understanding of how HIV-1 protease recognizes them. Prabu-

Jeyabalan and colleagues solved the X-ray structures of HIV-1 protease bound to decameric peptides corresponding to six different natural substrates. Analysis of these structures suggested that the substrates share a common shape filling a consensus volume. Direct H-bonds between substrate side chains and protease could not be observed. Hence, the protease recognizes the shape rather than a specific sequence.

Despite the fact that HIV-1 protease specificity is not determined by the sequence of the substrate, a number of inhibitors with excellent efficiency has been developed. Moreover, these findings led to the development of the substrate-envelope theory stating that inhibitors designed under substrate-envelope constrains are more robust against drug resistance (Nalam *et al.*). Hence, the finding that HRV2 $2A^{pro}$ specificity might also be determined by the spatial properties of the substrate is not precluding the development of anti $2A^{pro}$ compounds. However, these ideas also pinpoint the importance of high resolution structures of 2A proteases in complex with substrates as they are a prerequisite for the definition of a consensus volume.

6 REFERENCES

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.

Almstead, L. L. & Sarnow, P. (2007). Inhibition of U snRNP assembly by a virus-encoded proteinase. *Genes Dev* **21**, 1086-1097.

Alvarez, E., Castello, A., Carrasco, L. & Izquierdo, J. M. (2011). Alternative splicing, a new target to block cellular gene expression by poliovirus 2A protease. *Biochem Biophys Res Commun* **414**, 142-147.

Alvey, J. C., Wyckoff, E. E., Yu, S. F., Lloyd, R. & Ehrenfeld, E. (1991). cis- and trans-cleavage activities of poliovirus 2A protease expressed in Escherichia coli. *J Virol* **65**, 6077-6083.

Amako, K. & Dales, S. (1967). Cytopathology of Mengovirus infection. II. Proliferation of membranous cisternae. *Virology* **32**, 201-215.

Ambros, V. & Baltimore, D. (1978). Protein is linked to the 5' end of poliovirus RNA by a phosphodiester linkage to tyrosine. *J Biol Chem* **253**, 5263-5266.

Ambros, V. & Baltimore, D. (1980). Purification and properties of a HeLa cell enzyme able to remove the 5'-terminal protein from poliovirus RNA. *J Biol Chem* **255**, 6739-6744.

Andino, R., Rieckhof, G. E., Achacoso, P. L. & Baltimore, D. (1993). Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. *EMBO J* **12**, 3587-3598.

Andino, R., Rieckhof, G. E. & Baltimore, D. (1990). A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. *Cell* **63**, 369-380.

Antona, D., Leveque, N., Chomel, J. J., Dubrou, S., Levy-Bruhl, D. & Lina, B. (2007). Surveillance of enteroviruses in France, 2000-2004. *Eur J Clin Microbiol Infect Dis* **26**, 403-412.

Arden, K. E., McErlean, P., Nissen, M. D., Sloots, T. P. & Mackay, I. M. (2006). Frequent detection of human rhinoviruses, paramyxoviruses, coronaviruses, and bocavirus during acute respiratory tract infections. *J Med Virol* **78**, 1232-1240.

Arnesjo, B., Eden, T., Ihse, I., Nordenfelt, E. & Ursing, B. (1976). Enterovirus infections in acute pancreatitis - a possible etiological connection. *Scand J Gastroenterol* **11**, 645-649.

Arnold, E., Luo, M., Vriend, G., Rossmann, M. G., Palmenberg, A. C., Parks, G. D., Nicklin, M. J. & Wimmer, E. (1987). Implications of the picornavirus capsid structure for polyprotein processing. *Proc Natl Acad Sci U S A* **84**, 21-25.

Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201.

Arruda, E., Pitkaranta, A., Witek, T. J., Jr., Doyle, C. A. & Hayden, F. G. (1997). Frequency and natural history of rhinovirus infections in adults during autumn. *J Clin Microbiol* **35**, 2864-2868.

Aylward, R. B., Sutter, R. W. & Heymann, D. L. (2005). Policy. OPV cessation--the final step to a "polio-free" world. *Science* **310**, 625-626.

Baboonian, C., Davies, M. J., Booth, J. C. & McKenna, W. J. (1997). Coxsackie B viruses and human heart disease. *Curr Top Microbiol Immunol* **223**, 31-52.

Badorff, C., Berkely, N., Mehrotra, S., Talhouk, J. W., Rhoads, R. E. & Knowlton, K. U. (2000). Enteroviral protease 2A directly cleaves dystrophin and is inhibited by a dystrophin-based substrate analogue. *J Biol Chem* **275**, 11191-11197.

Badorff, C., Lee, G. H., Lamphear, B. J., Martone, M. E., Campbell, K. P., Rhoads, R. E. & Knowlton, K. U. (1999). Enteroviral protease 2A cleaves dystrophin: evidence of cytoskeletal disruption in an acquired cardiomyopathy. *Nat Med* **5**, 320-326.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* **98**, 10037-10041.

Barton, D. J. & Flanegan, J. B. (1993). Coupled translation and replication of poliovirus RNA in vitro: synthesis of functional 3D polymerase and infectious virus. *J Virol* **67**, 822-831.

Barton, D. J., Morasco, B. J. & Flanegan, J. B. (1999). Translating ribosomes inhibit poliovirus negative-strand RNA synthesis. *J Virol* **73**, 10104-10112.

Baxter, N. J., Roetzer, A., Liebig, H. D., Sedelnikova, S. E., Hounslow, A. M., Skern, T. & Waltho, J. P. (2006). Structure and dynamics of coxsackievirus B4 2A proteinase, an enyzme involved in the etiology of heart disease. *J Virol* **80**, 1451-1462.

Bazan, J. F. & Fletterick, R. J. (1988). Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional implications. *Proc Natl Acad Sci U S A* **85**, 7872-7876.

Belnap, D. M., McDermott, B. M., Jr., Filman, D. J., Cheng, N., Trus, B. L., Zuccola, H. J., Racaniello, V. R., Hogle, J. M. & Steven, A. C. (2000). Three-dimensional structure of poliovirus receptor bound to poliovirus. *Proc Natl Acad Sci U S A* **97**, 73-78.

Belov, G. A., Lidsky, P. V., Mikitas, O. V., Egger, D., Lukyanov, K. A., Bienz, K. & Agol, V. I. (2004). Bidirectional increase in permeability of nuclear envelope upon poliovirus infection and accompanying alterations of nuclear pores. *J Virol* **78**, 10166-10177.

Bergelson, J. M., Cunningham, J. A., Droguett, G., Kurt-Jones, E. A., Krithivas, A., Hong, J. S., Horwitz, M. S., Crowell, R. L. & Finberg, R. W. (1997). Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science* **275**, 1320-1323.

Bergmann, E. M., Mosimann, S. C., Chernaia, M. M., Malcolm, B. A. & James, M. N. (1997). The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. *J Virol* **71**, 2436-2448.

Bienz, K., Egger, D., Troxler, M. & Pasamontes, L. (1990). Structural organization of poliovirus RNA replication is mediated by viral proteins of the P2 genomic region. *J Virol* **64**, 1156-1163.

Bjellqvist, B., Basse, B., Olsen, E. & Celis, J. E. (1994). Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* **15**, 529-539.

Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J. C., Frutiger, S. & Hochstrasser, D. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* **14**, 1023-1031.

Blomqvist, S., Savolainen, C., Raman, L., Roivainen, M. & Hovi, T. (2002). Human rhinovirus 87 and enterovirus 68 represent a unique serotype with rhinovirus and enterovirus features. *J Clin Microbiol* **40**, 4218-4223.

Bodian, D. (1949). Histopathologic basis of clinical findings in poliomyelitis. *Am J Med* **6**, 563-578.

Bodian, D. (1956). Poliovirus in chimpanzee tissues after virus feeding. *Am J Hyg* **64**, 181-197.

Bodian, D. & Howe, A. (1940). An experimental study of the role of neurons in the dissemination of polyomyelitis virus in the nervous system. *Brain* **63**, 135-162.

Bond, C. S. & Schuttelkopf, A. W. (2009). ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* **65**, 510-512.

Borman, A. M., Kirchweger, R., Ziegler, E., Rhoads, R. E., Skern, T. & Kean, K. M. (1997). eIF4G and its proteolytic cleavage products: effect on initiation of protein synthesis from capped, uncapped, and IRES-containing mRNAs. *Rna* **3**, 186-196.

Bottino, C., Castriconi, R., Pende, D., Rivera, P., Nanni, M., Carnemolla, B., Cantoni, C., Grassi, J., Marcenaro, S., Reymond, N., Vitale, M., Moretta, L., Lopez, M. & Moretta, A. (2003). Identification of PVR (CD155) and Nectin-2 (CD112) as cell surface ligands for the human DNAM-1 (CD226) activating molecule. *J Exp Med* **198**, 557-567.

Burnet, F. M. & MacNamara, J. (1931). Immunological differences between strains of poliomyelitic viruses. *Br J Exp Pathol* **12**, 57-61.

Calandria, C., Irurzun, A., Barco, A. & Carrasco, L. (2004). Individual expression of poliovirus 2Apro and 3Cpro induces activation of caspase-3 and PARP cleavage in HeLa cells. *Virus Res* **104**, 39-49.

Caliguiri, L. A. & Tamm, I. (1969). Membranous structures associated with translation and transcription of poliovirus RNA. *Science* **166**, 885-886.

Campbell, J. W., Duee, E., Hodgson, G., Mercer, W. D., Stammers, D. K., Wendell, P. L., Muirhead, H. & Watson, H. C. (1972). X-ray diffraction studies on enzymes in the glycolytic pathway. *Cold Spring Harb Symp Quant Biol* **36**, 165-170.

Carrillo-Tripp, M., Shepherd, C. M., Borelli, I. A., Venkataraman, S., Lander, G., Natarajan, P., Johnson, J. E., Brooks, C. L., 3rd & Reddy, V. S. (2009). VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Res* **37**, D436-442.

Carson, S. D., Chapman, N. N. & Tracy, S. M. (1997). Purification of the putative coxsackievirus B receptor from HeLa cells. *Biochem Biophys Res Commun* **233**, 325-328.

Castello, A., Izquierdo, J. M., Welnowska, E. & Carrasco, L. (2009). RNA nuclear export is blocked by poliovirus 2A protease and is concomitant with nucleoporin cleavage. *J Cell Sci* **122**, 3799-3809.

CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763.

Cello, J., Paul, A. V. & Wimmer, E. (2002). Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* **297**, 1016-1018.

Center, R. J., Kobe, B., Wilson, K. A., Teh, T., Howlett, G. J., Kemp, B. E. & Poumbourios, P. (1998). Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. *Protein Sci* **7**, 1612-1619.

Centers for disease control and prevention (2000). Enterovirus surveillance--United States, 1997-1999. *MMWR Morb Mortal Wkly Rep* **49**, 913-916.

Centers for disease control and prevention (2006). Enterovirus surveillance--United States, 2002-2004. *MMWR Morb Mortal Wkly Rep* **55**, 153-156.

Centers for disease control and prevention (2008). Increased detections and severe neonatal disease associated with coxsackievirus B1 infection--United States, 2007. *MMWR Morb Mortal Wkly Rep* **57**, 553-556.

Centers for Disease Control and Prevention (CDC) (1999). Progress toward the global interruption of wild poliovirus type 2 transmission, 1999. *MMWR Morb Mortal Wkly Rep* **48**, 736-738, 747.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21.

Chen, Z. G., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kamer, G., Shanks, M., Lomonossoff, G. & Johnson, J. E. (1989). Protein-RNA interactions in an icosahedral virus at 3.0 A resolution. *Science* **245**, 154-159.

Chung, I. Y. & Paetzel, M. (2011). Crystal structure of a viral protease intramolecular acyl-enzyme complex: insights into cis-cleavage at the VP4/VP3 junction of Tellina birnavirus. *J Biol Chem* **286**, 12475-12482.

Clark, M. E., Hammerle, T., Wimmer, E. & Dasgupta, A. (1991). Poliovirus proteinase 3C converts an active form of transcription factor IIIC to an inactive form: a mechanism for inhibition of host cell polymerase III transcription by poliovirus. *EMBO J* **10**, 2941-2947.

Clark, M. E., Lieberman, P. M., Berk, A. J. & Dasgupta, A. (1993). Direct cleavage of human TATA-binding protein by poliovirus protease 3C in vivo and in vitro. *Mol Cell Biol* **13**, 1232-1237.

Cohen, C. J., Shieh, J. T., Pickles, R. J., Okegawa, T., Hsieh, J. T. & Bergelson, J. M. (2001). The coxsackievirus and adenovirus receptor is a transmembrane component of the tight junction. *Proc Natl Acad Sci U S A* **98**, 15191-15196.

Collis, P. S., O'Donnell, B. J., Barton, D. J., Rogers, J. A. & Flanegan, J. B. (1992). Replication of poliovirus RNA and subgenomic RNA transcripts in transfected cells. *J Virol* **66**, 6480-6488.

Cordey, S., Gerlach, D., Junier, T., Zdobnov, E. M., Kaiser, L. & Tapparel, C. (2008). The cis-acting replication elements define human enterovirus and rhinovirus species. *RNA* **14**, 1568-1578.

Couzin, J. (2006). Infectious diseases. Report concludes polio drugs are needed--after disease is eradicated. *Science* **311**, 1539.

Coyne, C. B., Kim, K. S. & Bergelson, J. M. (2007a). Poliovirus entry into human brain microvascular cells requires receptor-induced activation of SHP-2. *EMBO J* **26**, 4016-4028.

Coyne, C. B., Shen, L., Turner, J. R. & Bergelson, J. M. (2007b). Coxsackievirus entry across epithelial tight junctions requires occludin and the small GTPases Rab34 and Rab5. *Cell Host Microbe* **2**, 181-192.

Crowder, S. & Kirkegaard, K. (2005). Trans-dominant inhibition of RNA viral replication can slow growth of drug-resistant viruses. *Nat Genet* **37**, 701-709.

Cruz, R. R. (1984). Cuba: mass polio vaccination program, 1962-1982. *Rev Infect Dis* **6**, S408-S412.

Curry, S., Fry, E., Blakemore, W., Abu-Ghazaleh, R., Jackson, T., King, A., Lea, S., Newman, J. & Stuart, D. (1997). Dissecting the roles of VP0 cleavage and RNA packaging in picornavirus capsid stabilization: the structure of empty capsids of foot-and-mouth disease virus. *J Virol* **71**, 9743-9752.

Dales, S., Eggers, H. J., Tamm, I. & Palade, G. E. (1965). Electron Microscopic Study of the Formation of Poliovirus. *Virology* **26**, 379-389.

Dales, S. & Franklin, R. M. (1962). A comparison of the changes in fine structure of L cells during single cycles of viral multiplication, following their infection with the viruses of Mengo and encephalomyocarditis. *J Cell Biol* **14**, 281-302.

Dalldorf, G. & Sickles, G. M. (1948). An Unidentified, Filtrable Agent Isolated From the Feces of Children With Paralysis. *Science* **108**, 61-62.

Dalldorf, G., Sickles, G. M. & et al. (1949). A virus recovered from the feces of poliomyelitis patients pathogenic for suckling mice. *J Exp Med* **89**, 567-582.

Danthi, P., Tosteson, M., Li, Q. H. & Chow, M. (2003). Genome delivery and ion channel properties are altered in VP4 mutants of poliovirus. *J Virol* **77**, 5266-5274.

Davies, M. V., Pelletier, J., Meerovitch, K., Sonenberg, N. & Kaufman, R. J. (1991). The effect of poliovirus proteinase 2Apro expression on cellular metabolism. Inhibition of DNA replication, RNA polymerase II transcription, and translation. *J Biol Chem* **266**, 14714-14720.

de Breyne, S., Yu, Y., Unbehaun, A., Pestova, T. V. & Hellen, C. U. (2009). Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proc Natl Acad Sci U S A* **106**, 9197-9202.

De Kerpel, M., Van Molle, I., Brys, L., Wyns, L., De Greve, H. & Bouckaert, J. (2006). N-terminal truncation enables crystallization of the receptor-binding domain of the FedF bacterial adhesin. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 1278-1282.

De Sena, J. & Mandel, B. (1977). Studies on the in vitro uncoating of poliovirus. II. Characteristics of the membrane-modified particle. *Virology* **78**, 554-566.

DeLano, D. L. (2002). The PyMOL molecular graphics system. Palo Alto, CA, USA: DeLano Scientific.

Delbaere, L. T. & Brayer, G. D. (1985). The 1.8 A structure of the complex between chymostatin and Streptomyces griseus protease A. A model for serine protease catalytic tetrahedral intermediates. *J Mol Biol* **183**, 89-103.

Deszcz, L., Cencic, R., Sousa, C., Kuechler, E. & Skern, T. (2006). An antiviral peptide inhibitor that is active against picornavirus 2A proteinases but not cellular caspases. *J Virol* **80**, 9619-9627.

Deszcz, L., Seipelt, J., Vassilieva, E., Roetzer, A. & Kuechler, E. (2004). Antiviral activity of caspase inhibitors: effect on picornaviral 2A proteinase. *FEBS Lett* **560**, 51-55.

DeTulleo, L. & Kirchhausen, T. (1998). The clathrin endocytic pathway in viral infection. *EMBO J* **17**, 4585-4593.

Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.

Enders, J. F., Weller, T. H. & Robbins, F. C. (1949). Cultivation of the Lansing Strain of Poliomyelitis Virus in Cultures of Various Human Embryonic Tissues. *Science* **109**, 85-87.

Esfandiarei, M. & McManus, B. M. (2008). Molecular biology and pathogenesis of viral myocarditis. *Annu Rev Pathol* **3**, 127-155.

Etchison, D., Milburn, S. C., Edery, I., Sonenberg, N. & Hershey, J. W. (1982). Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. *J Biol Chem* **257**, 14806-14810.

Falah, N., Montserret, R., Lelogeais, V., Schuffenecker, I., Lina, B., Cortay, J. C. & Violot, S. (2012a). Blocking human enterovirus 71 replication by targeting viral 2A protease. *J Antimicrob Chemother*.

Falah, N., Violot, S., Decimo, D., Berri, F., Foucault-Grunenwald, M. L., Ohlmann, T., Schuffenecker, I., Morfin, F., Lina, B., Riteau, B. & Cortay, J. C. (2012b). Ex vivo and in vivo inhibition of human rhinovirus replication by a new pseudosubstrate of viral 2A protease. *J Virol* **86**, 691-704.

Fendrick, A. M., Monto, A. S., Nightengale, B. & Sarnes, M. (2003). The economic burden of non-influenza-related viral respiratory tract infection in the United States. *Arch Intern Med* **163**, 487-494.

Fenwick, M. L. & Cooper, P. D. (1962). Early interactions between poliovirus and ERK cells: some observations on the nature and significance of the rejected particles. *Virology* **18**, 212-223.

Fields, B. N., Knipe, D. M. & Howley, P. M. (2007). *Fields' virology*. Philadelphia, Pa. ; London: Wolters Kluwer/Lippincott Williams & Wilkins.

Finch, J. T. & Klug, A. (1959). Structure of poliomyelitis virus. *Nature* **183**, 1709-1714.

Flanegan, J. B., Petterson, R. F., Ambros, V., Hewlett, N. J. & Baltimore, D. (1977). Covalent linkage of a protein to a defined nucleotide sequence at the 5'-terminus of virion and replicative intermediate RNAs of poliovirus. *Proc Natl Acad Sci U S A* **74**, 961-965.

Foeger, N., Schmid, E. M. & Skern, T. (2003). Human rhinovirus 2 2Apro recognition of eukaryotic initiation factor 4GI. Involvement of an exosite. *J Biol Chem* **278**, 33200-33207.

Francis, T., Jr. (1955). Evaluation of the 1954 poliomyelitis vaccine field trial; further studies of results determining the effectiveness of poliomyelitis vaccine (Salk) in preventing paralytic poliomyelitis. *J Am Med Assoc* **158**, 1266-1270.

Fricks, C. E. & Hogle, J. M. (1990). Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. *J Virol* **64**, 1934-1945.

Fujito, T., Ikeda, W., Kakunaga, S., Minami, Y., Kajita, M., Sakamoto, Y., Monden, M. & Takai, Y. (2005). Inhibition of cell movement and proliferation by cell-cell contact-induced interaction of Necl-5 with nectin-3. *J Cell Biol* **171**, 165-173.

Gamarnik, A. V. & Andino, R. (1997). Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. *RNA* **3**, 882-892.

Gamarnik, A. V. & Andino, R. (1998). Switch from translation to RNA replication in a positive-stranded RNA virus. *Genes Dev* **12**, 2293-2304.

Gerber, K., Wimmer, E. & Paul, A. V. (2001). Biochemical and genetic studies of the initiation of human rhinovirus 2 RNA replication: identification of a cis-replicating element in the coding sequence of 2A(pro). *J Virol* **75**, 10979-10990.

Gern, J. E., Galagan, D. M., Jarjour, N. N., Dick, E. C. & Busse, W. W. (1997). Detection of rhinovirus RNA in lower airway cells during experimentally induced infection. *Am J Respir Crit Care Med* **155**, 1159-1161.

Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **182**, 319-326.

Glaser, W. & Skern, T. (2000). Extremely efficient cleavage of eIF4G by picornaviral proteinases L and 2A in vitro. *FEBS Lett* **480**, 151-155.

Glaser, W., Triendl, A. & Skern, T. (2003). The processing of eIF4GI by human rhinovirus type 2 2A(pro): relationship to self-cleavage and role of zinc. *J Virol* **77**, 5021-5025.

Goldstaub, D., Gradi, A., Bercovitch, Z., Grosmann, Z., Nophar, Y., Luria, S., Sonenberg, N. & Kahana, C. (2000). Poliovirus 2A protease induces apoptotic cell death. *Mol Cell Biol* **20**, 1271-1277.

Goodfellow, I., Chaudhry, Y., Richardson, A., Meredith, J., Almond, J. W., Barclay, W. & Evans, D. J. (2000). Identification of a cis-acting replication element within the poliovirus coding region. *J Virol* **74**, 4590-4600.

Gorbalenya, A. E., Donchenko, A. P., Blinov, V. M. & Koonin, E. V. (1989). Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. A distinct protein superfamily with a common structural fold. *FEBS Lett* **243**, 103-114.

Gorbalenya, A. E., Svitkin Yu, V. & Agol, V. I. (1981). Proteolytic activity of the nonstructural polypeptide p22 of encephalomyocarditis virus. *Biochem Biophys Res Commun* **98**, 952-960.

Gorbalenya, A. E., Svitkin, Y. V., Kazachkov, Y. A. & Agol, V. I. (1979). Encephalomyocarditis virus-specific polypeptide p22 is involved in the processing of the viral precursor polypeptides. *FEBS Lett* **108**, 1-5.

Gradi, A., Imataka, H., Svitkin, Y. V., Rom, E., Raught, B., Morino, S. & Sonenberg, N. (1998a). A novel functional human eukaryotic translation initiation factor 4G. *Mol Cell Biol* **18**, 334-342.

Gradi, A., Svitkin, Y. V., Imataka, H. & Sonenberg, N. (1998b). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc Natl Acad Sci U S A* **95**, 11089-11094.

Gradi, A., Svitkin, Y. V., Sommergruber, W., Imataka, H., Morino, S., Skern, T. & Sonenberg, N. (2003). Human rhinovirus 2A proteinase cleavage sites in eukaryotic initiation factors (eIF) 4GI and eIF4GII are different. *J Virol* **77**, 5026-5029.

Graham, K. L., Gustin, K. E., Rivera, C., Kuyumcu-Martinez, N. M., Choe, S. S., Lloyd, R. E., Sarnow, P. & Utz, P. J. (2004). Proteolytic cleavage of the catalytic subunit of DNA-dependent protein kinase during poliovirus infection. *J Virol* **78**, 6313-6321.

Greenberg, S. B., Allen, M., Wilson, J. & Atmar, R. L. (2000). Respiratory viral infections in adults with and without chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* **162**, 167-173.

Greve, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarck, M. E. & McClelland, A. (1989). The major human rhinovirus receptor is ICAM-1. *Cell* **56**, 839-847.

Grunert, H. P., Wolf, K. U., Langner, K. D., Sawitzky, D., Habermehl, K. O. & Zeichhardt, H. (1997). Internalization of human rhinovirus 14 into HeLa and ICAM-1-transfected BHK cells. *Med Microbiol Immunol* **186**, 1-9.

Gustin, K. E. & Sarnow, P. (2001). Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition. *Embo J* **20**, 240-249.

Hafenstein, S., Bowman, V. D., Chipman, P. R., Bator Kelly, C. M., Lin, F., Medof, M. E. & Rossmann, M. G. (2007). Interaction of decay-accelerating factor with coxsackievirus B3. *J Virol* **81**, 12927-12935.

Haghighat, A., Svitkin, Y., Novoa, I., Kuechler, E., Skern, T. & Sonenberg, N. (1996). The eIF4G-eIF4E complex is the target for direct cleavage by the rhinovirus 2A proteinase. *J Virol* **70**, 8444-8450.

Hamparian, V. V., Colonno, R. J., Cooney, M. K., Dick, E. C., Gwaltney, J. M., Jr., Hughes, J. H., Jordan, W. S., Jr., Kapikian, A. Z., Mogabgab, W. J., Monto, A. & et al. (1987). A collaborative report: rhinoviruses--extension of the numbering system from 89 to 100. *Virology* **159**, 191-192.

Hanecak, R., Semler, B. L., Anderson, C. W. & Wimmer, E. (1982). Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. *Proc Natl Acad Sci U S A* **79**, 3973-3977.

Harber, J. J., Bradley, J., Anderson, C. W. & Wimmer, E. (1991). Catalysis of poliovirus VP0 maturation cleavage is not mediated by serine 10 of VP2. *J Virol* **65**, 326-334.

Hasegawa, H. & Holm, L. (2009). Advances and pitfalls of protein structural alignment. *Curr Opin Struct Biol* **19**, 341-348.

He, Y., Bowman, V. D., Mueller, S., Bator, C. M., Bella, J., Peng, X., Baker, T. S., Wimmer, E., Kuhn, R. J. & Rossmann, M. G. (2000). Interaction of the poliovirus receptor with poliovirus. *Proc Natl Acad Sci U S A* **97**, 79-84.

He, Y., Chipman, P. R., Howitt, J., Bator, C. M., Whitt, M. A., Baker, T. S., Kuhn, R. J., Anderson, C. W., Freimuth, P. & Rossmann, M. G. (2001). Interaction of coxsackievirus B3 with the full length coxsackievirus-adenovirus receptor. *Nat Struct Biol* **8**, 874-878.

Hellen, C. U., Lee, C. K. & Wimmer, E. (1992). Determinants of substrate recognition by poliovirus 2A proteinase. *J Virol* **66**, 3330-3338.

Herold, J. & Andino, R. (2001). Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol Cell* **7**, 581-591.

Hewat, E. A., Neumann, E., Conway, J. F., Moser, R., Ronacher, B., Marlovits, T. C. & Blaas, D. (2000). The cellular receptor to human rhinovirus 2 binds around the 5-fold axis and not in the canyon: a structural view. *Embo J* **19**, 6317-6325.

Hindersson, M., Elshebani, A., Orn, A., Tuvemo, T. & Frisk, G. (2005). Simultaneous type 1 diabetes onset in mother and son coincident with an enteroviral infection. *J Clin Virol* **33**, 158-167.

Hindiyeh, M., Li, Q. H., Basavappa, R., Hogle, J. M. & Chow, M. (1999). Poliovirus mutants at histidine 195 of VP2 do not cleave VP0 into VP2 and VP4. *J Virol* **73**, 9072-9079.

Hofer, F., Gruenberger, M., Kowalski, H., Machat, H., Huettinger, M., Kuechler, E. & Blaas, D. (1994). Members of the low density lipoprotein receptor family mediate cell entry of a minor-group common cold virus. *Proc Natl Acad Sci U S A* **91**, 1839-1842.

Hogle, J. M., Chow, M. & Filman, D. J. (1985). Three-dimensional structure of poliovirus at 2.9 A resolution. *Science* **229**, 1358-1365.

Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233**, 123-138.

Honda, T., Saitoh, H., Masuko, M., Katagiri-Abe, T., Tominaga, K., Kozakai, I., Kobayashi, K., Kumanishi, T., Watanabe, Y. G., Odani, S. & Kuwano, R. (2000). The coxsackievirus-adenovirus receptor protein as a cell adhesion molecule in the developing mouse brain. *Brain Res Mol Brain Res* **77**, 19-28.

Hoover-Litty, H. & Greve, J. M. (1993). Formation of rhinovirus-soluble ICAM-1 complexes and conformational changes in the virion. *J Virol* **67**, 390-397.

Horsnell, C., Gama, R. E., Hughes, P. J. & Stanway, G. (1995). Molecular relationships between 21 human rhinovirus serotypes. *J Gen Virol* **76 (Pt 10)**, 2549-2555.

ICTV (2011). http://www.ictvonline.org/virusTaxonomy.asp?version=2011: International committee on taxonomy of viruses.

Ida-Hosonuma, M., Iwasaki, T., Yoshikawa, T., Nagata, N., Sato, Y., Sata, T., Yoneyama, M., Fujita, T., Taya, C., Yonekawa, H. & Koike, S. (2005). The alpha/beta interferon response controls tissue tropism and pathogenicity of poliovirus. *J Virol* **79**, 4460-4469.

Imrie, C. W., Ferguson, J. C. & Sommerville, R. G. (1977). Coxsackie and mumpsvirus infection in a prospective study of acute pancreatitis. *Gut* **18**, 53-56.

Jacobson, M. F. & Baltimore, D. (1968). Morphogenesis of poliovirus. I. Association of the viral RNA with coat protein. *J Mol Biol* **33**, 369-378.

Jang, S. K. & Wimmer, E. (1990). Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. *Genes Dev* **4**, 1560-1572.

Jenkins, O., Booth, J. D., Minor, P. D. & Almond, J. W. (1987). The complete nucleotide sequence of coxsackievirus B4 and its comparison to other members of the Picornaviridae. *J Gen Virol* **68 (Pt 7)**, 1835-1848.

Joachims, M., Van Breugel, P. C. & Lloyd, R. E. (1999). Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol* **73**, 718-727.

Johnston, S. L., Pattemore, P. K., Sanderson, G., Smith, S., Lampe, F., Josephs, L., Symington, P., O'Toole, S., Myint, S. H., Tyrrell, D. A. & et al. (1995). Community study of role of viral infections in exacerbations of asthma in 9-11 year old children. *BMJ* **310**, 1225-1229.

Jore, J., De Geus, B., Jackson, R. J., Pouwels, P. H. & Enger-Valk, B. E. (1988). Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. *J Gen Virol* **69 (Pt 7)**, 1627-1636.

Jurgens, C. K., Barton, D. J., Sharma, N., Morasco, B. J., Ogram, S. A. & Flanegan, J. B. (2006). 2Apro is a multifunctional protein that regulates the stability, translation and replication of poliovirus RNA. *Virology* **345**, 346-357.

Kaiser, L., Aubert, J. D., Pache, J. C., Deffernez, C., Rochat, T., Garbino, J., Wunderli, W., Meylan, P., Yerly, S., Perrin, L., Letovanec, I., Nicod, L., Tapparel, C. & Soccal, P. M. (2006). Chronic rhinoviral infection in lung transplant recipients. *Am J Respir Crit Care Med* **174**, 1392-1399.

Kallman, F., Williams, R. C., Dulbecco, R. & Vogt, M. (1958). Fine structure of changes produced in cultured cells sampled at specified intervals during a single growth cycle of polio virus. *J Biophys Biochem Cytol* **4**, 301-308.

Kempf, B. J. & Barton, D. J. (2008). Poliovirus 2A(Pro) increases viral mRNA and polysome stability coordinately in time with cleavage of eIF4G. *J Virol* **82**, 5847-5859.

Kerekatte, V., Keiper, B. D., Badorff, C., Cai, A., Knowlton, K. U. & Rhoads, R. E. (1999). Cleavage of Poly(A)-binding protein by coxsackievirus 2A protease in vitro and in vivo: another mechanism for host protein synthesis shutoff? *J Virol* **73**, 709-717.

Khan, A. G., Pichler, J., Rosemann, A. & Blaas, D. (2007). Human rhinovirus type 54 infection via heparan sulfate is less efficient and strictly dependent on low endosomal pH. *J Virol* **81**, 4625-4632.

Khan, A. G., Pickl-Herk, A., Gajdzik, L., Marlovits, T. C., Fuchs, R. & Blaas, D. Human rhinovirus 14 enters rhabdomyosarcoma cells expressing icam-1 by a clathrin-, caveolin-, and flotillin-independent pathway. *J Virol* **84**, 3984-3992.

Khetsuriani, N., Lu, X., Teague, W. G., Kazerouni, N., Anderson, L. J. & Erdman, D. D. (2008). Novel human rhinoviruses and exacerbation of asthma in children. *Emerg Infect Dis* **14**, 1793-1796.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L. & Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**, D387-392.

Kling, C., Wernstedt, W. & Pettersson, A. (1912a). Recherches sur le mode de propagation de la paralysie infantile épidémique (maladie de Heine-Medin). Deuxiéme mémoire. *Z Immunitätsforsch* **12**, 657-670.

Kling, C., Wernstedt, W. & Pettersson, A. (1912b). Recherches sur le mode de propagation de la paralysie infantile épidémique (maladie de Heine-Medin). Premier mémoire. *Z Immunitätsforsch* **12**, 316-323.

Kling, S., Donninger, H., Williams, Z., Vermeulen, J., Weinberg, E., Latiff, K., Ghildyal, R. & Bardin, P. (2005). Persistence of rhinovirus RNA after asthma exacerbation in children. *Clin Exp Allergy* **35**, 672-678.

Knowles, N. J., Hovi, T., Hyypiä, T., King, A. M. Q., Lindberg, A. M., Pallansch, M. A., Palmenberg, A. C., Simmonds, P., Skern, T., Stanway, G., Yamashita, T. & Zell, R. (2012). Picornaviridae. In *Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses*, pp. 855-880. Edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz. San Diego: Elsevier.

Knowlton, K. U. (2008). CVB infection and mechanisms of viral cardiomyopathy. *Curr Top Microbiol Immunol* **323**, 315-335.

Koike, S., Horie, H., Ise, I., Okitsu, A., Yoshida, M., Iizuka, N., Takeuchi, K., Takegami, T. & Nomoto, A. (1990). The poliovirus receptor protein is produced both as membrane-bound and secreted forms. *EMBO J* **9**, 3217-3224.

Kolatkar, P. R., Bella, J., Olson, N. H., Bator, C. M., Baker, T. S. & Rossmann, M. G. (1999). Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor. *Embo J* **18**, 6249-6259.

Krieger, E., Joo, K., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. & Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **77 Suppl 9**, 114-122.

Kuo, R. L., Kung, S. H., Hsu, Y. Y. & Liu, W. T. (2002). Infection with enterovirus 71 or expression of its 2A protease induces apoptotic cell death. *J Gen Virol* **83**, 1367-1376.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132.

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

Lal, S. M., Fowler, D., Losasso, C. J. & Berg, G. G. (1988). Coxsackie virus-induced acute pancreatitis in a long-term dialysis patient. *Am J Kidney Dis* **11**, 434-436.

Lamphear, B. J., Kirchweger, R., Skern, T. & Rhoads, R. E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J Biol Chem* **270**, 21975-21983.

Lamphear, B. J., Yan, R., Yang, F., Waters, D., Liebig, H. D., Klump, H., Kuechler, E., Skern, T. & Rhoads, R. E. (1993). Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of the 2A proteases from human Coxsackievirus and rhinovirus. *J Biol Chem* **268**, 19200-19203.

Lamson, D., Renwick, N., Kapoor, V., Liu, Z., Palacios, G., Ju, J., Dean, A., St George, K., Briese, T. & Lipkin, W. I. (2006). MassTag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004-2005. *J Infect Dis* **194**, 1398-1402.

Landsteiner, K. & Popper, E. (1909). Übertragung der Poliomyelitis acuta auf Affen. *Zeitsch Immunol* **2**, 377-390.

Lange, R., Peng, X., Wimmer, E., Lipp, M. & Bernhardt, G. (2001). The poliovirus receptor CD155 mediates cell-to-matrix contacts by specifically binding to vitronectin. *Virology* **285**, 218-227.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Lau, S. K., Yip, C. C., Tsoi, H. W., Lee, R. A., So, L. Y., Lau, Y. L., Chan, K. H., Woo, P. C. & Yuen, K. Y. (2007). Clinical features and complete genome characterization of a distinct human rhinovirus (HRV) genetic cluster, probably representing a previously undetected HRV species, HRV-C, associated with acute respiratory illness in children. *J Clin Microbiol* **45**, 3655-3664.

Lawson, M. A. & Semler, B. L. (1992). Alternate poliovirus nonstructural protein processing cascades generated by primary sites of 3C proteinase cleavage. *Virology* **191**, 309-320.

Lee, C. K. & Wimmer, E. (1988). Proteolytic processing of poliovirus polyprotein: elimination of 2Apro-mediated, alternative cleavage of polypeptide 3CD by in vitro mutagenesis. *Virology* **166**, 405-414.

Lee, P. W. & Colter, J. S. (1979). Further characterization of Mengo subviral particles: a new hypothesis for picornavirus assembly. *Virology* **97**, 266-274.

Lee, W. M., Kiesner, C., Pappas, T., Lee, I., Grindle, K., Jartti, T., Jakiela, B., Lemanske, R. F., Jr., Shult, P. A. & Gern, J. E. (2007). A diverse group of previously unrecognized human rhinoviruses are common causes of respiratory illnesses in infants. *PLoS One* **2**, e966.

Lee, Y. F., Nomoto, A., Detjen, B. M. & Wimmer, E. (1977). A protein covalently linked to poliovirus genome RNA. *Proc Natl Acad Sci U S A* **74**, 59-63.

Leslie, A. G. W. & Powell, H. R. (2007). Processing Diffraction Data with Mosflm. In *Evolving methods for macromolecular crystallography : the structural path to the understanding of the mechanisms of action of CBRN agents*

NATO science series Series II, Mathematics, physics, and chemistry v 245, pp. viii, 191 p. Edited by R. J. Read & J. Sussman. Dordrecht: Springer Verlag.

Li, X., Lu, H. H., Mueller, S. & Wimmer, E. (2001). The C-terminal residues of poliovirus proteinase 2A(pro) are critical for viral RNA replication but not for cis- or trans-proteolytic cleavage. *J Gen Virol* **82**, 397-408.

Liebig, H. D., Ziegler, E., Yan, R., Hartmuth, K., Klump, H., Kowalski, H., Blaas, D., Sommergruber, W., Frasel, L., Lamphear, B. & et al. (1993). Purification of two picornaviral 2A proteinases: interaction with eIF-4 gamma and influence on in vitro translation. *Biochemistry* **32**, 7581-7588.

Lipton, S. H. & Bodwell, C. E. (1973). Oxidation of amino acids by dimethyl sulfoxide. *J Agric Food Chem* **21**, 235-237.

Liu, Y., Manna, A., Li, R., Martin, W. E., Murphy, R. C., Cheung, A. L. & Zhang, G. (2001). Crystal structure of the SarR protein from Staphylococcus aureus. *Proc Natl Acad Sci U S A* **98**, 6877-6882.

Loeffler, F. & Frosch, P. (1897). Summarischer Bericht über die Ergebnisse der Untersuchungen der Kommission zur Erforschung der Maul- und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin. *Centralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten, Abt I* **22**, 257-259.

Loeffler, F. & Frosch, P. (1898). Berichte der Kommission zur Erforschung der MAul- und Klauenseuche bei dem Institut für Infektionskrankheiten in Berlin. *Centralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten, Abt I* **23**, 371-391.

Longenecker, K. L., Garrard, S. M., Sheffield, P. J. & Derewenda, Z. S. (2001). Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. *Acta Crystallogr D Biol Crystallogr* **57**, 679-688.

Longson, M., Cole, F. M. & Davies, D. (1969). Isolation of a Coxsackie virus group B, type 5, from the heart of a fatal case of myocarditis in an adult. *J Clin Pathol* **22**, 654-658.

Lublin, D. M. & Atkinson, J. P. (1989). Decay-accelerating factor: biochemistry, molecular biology, and function. *Annu Rev Immunol* **7**, 35-58.

Luz, N. & Beck, E. (1991). Interaction of a cellular 57-kilodalton protein with the internal translation initiation site of foot-and-mouth disease virus. *J Virol* **65**, 6486-6494.

Mader, S., Lee, H., Pause, A. & Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol Cell Biol* **15**, 4990-4997.

Makela, M. J., Puhakka, T., Ruuskanen, O., Leinonen, M., Saikku, P., Kimpimaki, M., Blomqvist, S., Hyypia, T. & Arstila, P. (1998). Viruses and bacteria in the etiology of the common cold. *J Clin Microbiol* **36**, 539-542.

Marongiu, M. E., Pani, A., Corrias, M. V., Sau, M. & La Colla, P. (1981). Poliovirus morphogenesis. I. Identification of 80S dissociable particles and evidence for the artifactual production of procapsids. *J Virol* **39**, 341-347.

Martínez-Salas, E. & Ryan, M. D. (2010). Translation and Protein Processing. In *The picornaviruses*, pp. 141-161. Edited by E. Ehrenfeld, E. Domingo & R. P. Roos: ASM Press.

Martino, T. A., Petric, M., Weingartl, H., Bergelson, J. M., Opavsky, M. A., Richardson, C. D., Modlin, J. F., Finberg, R. W., Kain, K. C., Willis, N., Gauntt, C. J. & Liu, P. P. (2000). The coxsackie-adenovirus receptor (CAR) is used by reference strains and clinical isolates representing all six serotypes of coxsackievirus group B and by swine vesicular disease virus. *Virology* **271**, 99-108.

Marvil, P., Knowles, N. J., Mockett, A. P., Britton, P., Brown, T. D. & Cavanagh, D. (1999). Avian encephalomyelitis virus is a picornavirus and is most closely related to hepatitis A virus. *J Gen Virol* **80 (Pt 3)**, 653-662.

Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., Villafranca, J. E., Janson, C. A., McElroy, H. E., Gribskov, C. L. & et al. (1994). Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell* **77**, 761-771.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.

McKnight, K. L. & Lemon, S. M. (1996). Capsid coding sequence is required for efficient replication of human rhinovirus 14 RNA. *J Virol* **70**, 1941-1952.

McKnight, K. L. & Lemon, S. M. (1998). The rhinovirus type 14 genome contains an internally located RNA structure that is required for viral replication. *RNA* **4**, 1569-1584.

Mendelsohn, C. L., Wimmer, E. & Racaniello, V. R. (1989). Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. *Cell* **56**, 855-865.

Merl, S., Michaelis, C., Jaschke, B., Vorpahl, M., Seidl, S. & Wessely, R. (2005). Targeting 2A protease by RNA interference attenuates coxsackieviral cytopathogenicity and promotes survival in highly susceptible mice. *Circulation* **111**, 1583-1592.

Miller, E. K., Edwards, K. M., Weinberg, G. A., Iwane, M. K., Griffin, M. R., Hall, C. B., Zhu, Y., Szilagyi, P. G., Morin, L. L., Heil, L. H., Lu, X. & Williams, J. V. (2009). A novel group of rhinoviruses is associated with asthma hospitalizations. *J Allergy Clin Immunol* **123**, 98-104 e101.

Milstone, A. M., Petrella, J., Sanchez, M. D., Mahmud, M., Whitbeck, J. C. & Bergelson, J. M. (2005). Interaction with coxsackievirus and adenovirus receptor, but not with decay-accelerating factor (DAF), induces A-particle formation in a DAF-binding coxsackievirus B3 isolate. *J Virol* **79**, 655-660.

Mischak, H., Neubauer, C., Berger, B., Kuechler, E. & Blaas, D. (1988). Detection of the human rhinovirus minor group receptor on renaturing Western blots. *J Gen Virol* **69 (Pt 10)**, 2653-2656.

Molla, A., Hellen, C. U. & Wimmer, E. (1993). Inhibition of proteolytic activity of poliovirus and rhinovirus 2A proteinases by elastase-specific inhibitors. *J Virol* **67**, 4688-4695.

Molla, A., Paul, A. V. & Wimmer, E. (1991). Cell-free, de novo synthesis of poliovirus. *Science* **254**, 1647-1651.

Morrison, J. M. & Racaniello, V. R. (2009). Proteinase 2Apro is essential for enterovirus replication in type I interferon-treated cells. *J Virol* **83**, 4412-4422.

Mosimann, S. C., Cherney, M. M., Sia, S., Plotch, S. & James, M. N. (1997). Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J Mol Biol* **273**, 1032-1047.

Mosser, A. G., Vrtis, R., Burchell, L., Lee, W. M., Dick, C. R., Weisshaar, E., Bock, D., Swenson, C. A., Cornwell, R. D., Meyer, K. C., Jarjour, N. N., Busse, W. W. & Gern, J. E. (2005). Quantitative and qualitative analysis of rhinovirus infection in bronchial tissues. *Am J Respir Crit Care Med* **171**, 645-651.

Muckelbauer, J. K., Kremer, M., Minor, I., Diana, G., Dutko, F. J., Groarke, J., Pevear, D. C. & Rossmann, M. G. (1995). The structure of coxsackievirus B3 at 3.5 A resolution. *Structure* **3**, 653-667.

Nairn, C., Galbraith, D. N., Taylor, K. W. & Clements, G. B. (1999). Enterovirus variants in the serum of children at the onset of Type 1 diabetes mellitus. *Diabet Med* **16**, 509-513.

Nalam, M. N., Ali, A., Altman, M. D., Reddy, G. S., Chellappan, S., Kairys, V., Ozen, A., Cao, H., Gilson, M. K., Tidor, B., Rana, T. M. & Schiffer, C. A. Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. *J Virol* **84**, 5368-5378.

Nathanson, N. & Bodian, D. (1961). Experimental poliomyelitis following intramuscular virus injection. I. The effect of neural block on a neurotropic and a pantropic strain. *Bull Johns Hopkins Hosp* **108**, 308-319.

Neubauer, C., Frasel, L., Kuechler, E. & Blaas, D. (1987). Mechanism of entry of human rhinovirus 2 into HeLa cells. *Virology* **158**, 255-258.

Neubauer, D. (2008). Investigating substrate specificities of picornaviral proteinases by mutational analysis. In *Department of medical biochemistry*, p. 109. Vienna: University of Vienna.

Neubauer, D., Aumayr, M., Gösler, I. & Skern, T. (2013). Specificity of human rhinovirus 2Apro is determined by combined spatial properties of four cleavage site residues. *submitted to J Gen Virol*.

Neutra, M. R., Pringault, E. & Kraehenbuhl, J. P. (1996). Antigen sampling across epithelial barriers and induction of mucosal immune responses. *Annu Rev Immunol* **14**, 275-300.

Nicholson, K. G., Kent, J. & Ireland, D. C. (1993). Respiratory viruses and exacerbations of asthma in adults. *BMJ* **307**, 982-986.

Nobis, P., Zibirre, R., Meyer, G., Kuhne, J., Warnecke, G. & Koch, G. (1985). Production of a monoclonal antibody against an epitope on HeLa cells that is the functional poliovirus binding site. *J Gen Virol* **66 (Pt 12)**, 2563-2569.

Nomoto, A., Lee, Y. F. & Wimmer, E. (1976). The 5' end of poliovirus mRNA is not capped with m7G(5')ppp(5')Np. *Proc Natl Acad Sci U S A* **73**, 375-380.

Novak, J. E. & Kirkegaard, K. (1991). Improved method for detecting poliovirus negative strands used to demonstrate specificity of positive-strand encapsidation and the ratio of positive to negative strands in infected cells. *J Virol* **65**, 3384-3387.

Novak, J. E. & Kirkegaard, K. (1994). Coupling between genome translation and replication in an RNA virus. *Genes Dev* **8**, 1726-1737.

Novoa, I., Martinez-Abarca, F., Fortes, P., Ortin, J. & Carrasco, L. (1997). Cleavage of p220 by purified poliovirus 2A(pro) in cell-free systems: effects on translation of capped and uncapped mRNAs. *Biochemistry* **36**, 7802-7809.

Nugent, C. I. & Kirkegaard, K. (1995). RNA binding properties of poliovirus subviral particles. *J Virol* **69**, 13-22.

Oberste, M. S., Maher, K., Schnurr, D., Flemister, M. R., Lovchik, J. C., Peters, H., Sessions, W., Kirk, C., Chatterjee, N., Fuller, S., Hanauer, J. M. & Pallansch, M. A. (2004). Enterovirus 68 is associated with respiratory illness and shares biological features with both the enteroviruses and the rhinoviruses. *J Gen Virol* **85**, 2577-2584.

Ohka, S., Matsuda, N., Tohyama, K., Oda, T., Morikawa, M., Kuge, S. & Nomoto, A. (2004). Receptor (CD155)-dependent endocytosis of poliovirus and retrograde axonal transport of the endosome. *J Virol* **78**, 7186-7198.

Ohka, S., Sakai, M., Bohnert, S., Igarashi, H., Deinhardt, K., Schiavo, G. & Nomoto, A. (2009). Receptor-dependent and -independent axonal retrograde transport of poliovirus in motor neurons. *J Virol* **83**, 4995-5004.

Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S. & Rossmann, M. G. (1993). Structure of a human rhinovirus complexed with its receptor molecule. *Proc Natl Acad Sci U S A* **90**, 507-511.

Onodera, S. & Phillips, B. A. (1987). A novel method for obtaining poliovirus 14 S pentamers from procapsids and their self-assembly into virus-like shells. *Virology* **159**, 278-287.

Overton, I. M., Padovani, G., Girolami, M. A. & Barton, G. J. (2008). ParCrys: a Parzen window density estimation approach to protein crystallization propensity prediction. *Bioinformatics* **24**, 901-907.

Owen, R. L. & Jones, A. L. (1974). Epithelial cell specialization within human Peyer's patches: an ultrastructural study of intestinal lymphoid follicles. *Gastroenterology* **66**, 189-203.

Palmenberg, A. C. (1982). In vitro synthesis and assembly of picornaviral capsid intermediate structures. *J Virol* **44**, 900-906.

Palmenberg, A. C., Neubauer, D. & Skern, T. (2010). Genome organization and encoded proteins. In *The picornaviruses*, pp. 3-18. Edited by E. Ehrenfeld, E. Domingo & R. P. Roos: ASM Press.

Palmenberg, A. C., Pallansch, M. A. & Rueckert, R. R. (1979). Protease required for processing picornaviral coat protein resides in the viral replicase gene. *J Virol* **32**, 770-778.

Palmenberg, A. C., Spiro, D., Kuzmickas, R., Wang, S., Djikeng, A., Rathe, J. A., Fraser-Liggett, C. M. & Liggett, S. B. (2009). Sequencing and Analyses of All Known Human Rhinovirus Genomes Reveals Structure and Evolution. *Science*.

Papadopoulos, N. G., Bates, P. J., Bardin, P. G., Papi, A., Leir, S. H., Fraenkel, D. J., Meyer, J., Lackie, P. M., Sanderson, G., Holgate, S. T. & Johnston, S. L. (2000). Rhinoviruses infect the lower airways. *J Infect Dis* **181**, 1875-1884.

Papadopoulos, N. G., Sanderson, G., Hunter, J. & Johnston, S. L. (1999). Rhinoviruses replicate effectively at lower airway temperatures. *J Med Virol* **58**, 100-104.

Parenti, D. M., Steinberg, W. & Kang, P. (1996). Infectious causes of acute pancreatitis. *Pancreas* **13**, 356-371.

Park, N., Katikaneni, P., Skern, T. & Gustin, K. E. (2008). Differential targeting of nuclear pore complex proteins in poliovirus-infected cells. *J Virol* **82**, 1647-1655.

Park, N., Skern, T. & Gustin, K. E. (2010). Specific cleavage of the nuclear pore complex protein Nup62 by a viral protease. *J Biol Chem* **285**, 28796-28805.

Parsley, T. B., Towner, J. S., Blyn, L. B., Ehrenfeld, E. & Semler, B. L. (1997). Poly (rC) binding protein 2 forms a ternary complex with the 5'-terminal sequences of poliovirus RNA and the viral 3CD proteinase. *RNA* **3**, 1124-1134.

Patel, K. P., Coyne, C. B. & Bergelson, J. M. (2009). Dynamin- and lipid raft-dependent entry of decay-accelerating factor (DAF)-binding and non-DAF-binding coxsackieviruses into nonpolarized cells. *J Virol* **83**, 11064-11077.

Paul, A. V., Peters, J., Mugavero, J., Yin, J., van Boom, J. H. & Wimmer, E. (2003a). Biochemical and genetic studies of the VPg uridylylation reaction catalyzed by the RNA polymerase of poliovirus. *J Virol* **77**, 891-904.

Paul, A. V., Rieder, E., Kim, D. W., van Boom, J. H. & Wimmer, E. (2000). Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. *J Virol* **74**, 10359-10370.

Paul, A. V., van Boom, J. H., Filippov, D. & Wimmer, E. (1998). Protein-primed RNA synthesis by purified poliovirus RNA polymerase. *Nature* **393**, 280-284.

Paul, A. V., Yin, J., Mugavero, J., Rieder, E., Liu, Y. & Wimmer, E. (2003b). A "slide-back" mechanism for the initiation of protein-primed RNA synthesis by the RNA polymerase of poliovirus. *J Biol Chem* **278**, 43951-43960.

Peitsch, M. C., Wells, T. N., Stampf, D. R. & Sussman, J. L. (1995). The Swiss-3DImage collection and PDB-Browser on the World-Wide Web. *Trends Biochem Sci* **20**, 82-84.

Pelletier, J. & Sonenberg, N. (1988). Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* **334**, 320-325.

Pelletier, J. & Sonenberg, N. (1989). Internal binding of eucaryotic ribosomes on poliovirus RNA: translation in HeLa cell extracts. *J Virol* **63**, 441-444.

Perera, R., Daijogo, S., Walter, B. L., Nguyen, J. H. & Semler, B. L. (2007). Cellular protein modification by poliovirus: the two faces of poly(rC)-binding protein. *J Virol* **81**, 8919-8932.

Perez, L. & Carrasco, L. (1993). Entry of poliovirus into cells does not require a low-pH step. *J Virol* **67**, 4543-4548.

Pestova, T. V., Hellen, C. U. & Shatsky, I. N. (1996a). Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol Cell Biol* **16**, 6859-6869.

Pestova, T. V., Shatsky, I. N. & Hellen, C. U. (1996b). Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Mol Cell Biol* **16**, 6870-6878.

Petersen, J. F., Cherney, M. M., Liebig, H. D., Skern, T., Kuechler, E. & James, M. N. (1999). The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis. *Embo J* **18**, 5463-5475.

Piotrowska, Z., Vazquez, M., Shapiro, E. D., Weibel, C., Ferguson, D., Landry, M. L. & Kahn, J. S. (2009). Rhinoviruses are a major cause of wheezing and hospitalization in children less than 2 years of age. *Pediatr Infect Dis J* **28**, 25-29.

Prabu-Jeyabalan, M., Nalivaika, E. & Schiffer, C. A. (2002). Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure* **10**, 369-381.

Prchla, E., Kuechler, E., Blaas, D. & Fuchs, R. (1994). Uncoating of human rhinovirus serotype 2 from late endosomes. *J Virol* **68**, 3713-3723.

Prchla, E., Plank, C., Wagner, E., Blaas, D. & Fuchs, R. (1995). Virus-mediated release of endosomal content in vitro: different behavior of adenovirus and rhinovirus serotype 2. *J Cell Biol* **131**, 111-123.

Racaniello, V. R. & Baltimore, D. (1981). Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. *Proc Natl Acad Sci U S A* **78**, 4887-4891.

Redondo, N., Sanz, M. A., Welnowska, E. & Carrasco, L. (2011). Translation without eIF2 promoted by poliovirus 2A protease. *PLoS One* **6**, e25699.

Reich, E., Franklin, R. M., Shatkin, A. J. & Tatum, E. L. (1961). Effect of actinomycin D on cellular nucleic acid synthesis and virus production. *Science* **134**, 556-557.

Ren, R. & Racaniello, V. R. (1992). Poliovirus spreads from muscle to the central nervous system by neural pathways. *J Infect Dis* **166**, 747-752.

Ren, R. B., Costantini, F., Gorgacz, E. J., Lee, J. J. & Racaniello, V. R. (1990). Transgenic mice expressing a human poliovirus receptor: a new model for poliomyelitis. *Cell* **63**, 353-362.

Renwick, N., Schweiger, B., Kapoor, V., Liu, Z., Villari, J., Bullmann, R., Miething, R., Briese, T. & Lipkin, W. I. (2007). A recently identified rhinovirus genotype is associated with severe respiratory-tract infection in children in Germany. *J Infect Dis* **196**, 1754-1760.

Roberts, J. A., Kuiper, M. J., Thorley, B. R., Smooker, P. M. & Hung, A. (2012). Investigation of a predicted N-terminal amphipathic alpha-helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion. *J Mol Graph Model* **38**, 165-173.

Rohde, G., Wiethege, A., Borg, I., Kauth, M., Bauer, T. T., Gillissen, A., Bufe, A. & Schultze-Werninghaus, G. (2003). Respiratory viruses in exacerbations of chronic obstructive pulmonary disease requiring hospitalisation: a case-control study. *Thorax* **58**, 37-42.

Rombaut, B., Foriers, A. & Boeye, A. (1991). In vitro assembly of poliovirus 14 S subunits: identification of the assembly promoting activity of infected cell extracts. *Virology* **180**, 781-787.

Rombaut, B., Vrijsen, R. & Boeye, A. (1984). In vitro assembly of poliovirus empty capsids: antigenic consequences and immunological assay of the morphopoietic factor. *Virology* **135**, 546-550.

Rombaut, B., Vrijsen, R., Delgadillo, R., Vanden Berghe, D. & Boeye, A. (1985). Characterization and assembly of poliovirus-related 45 S particles. *Virology* **146**, 111-119.

Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G. & et al. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* **317**, 145-153.

Rothberg, P. G., Harris, T. J., Nomoto, A. & Wimmer, E. (1978). O4-(5'-uridylyl)tyrosine is the bond between the genome-linked protein and the RNA of poliovirus. *Proc Natl Acad Sci U S A* **75**, 4868-4872.

Rueckert, R. R. & Wimmer, E. (1984). Systematic nomenclature of picornavirus proteins. *J Virol* **50**, 957-959.

Sabin, A. B. (1957). Properties and behavior of orally administered attenuated poliovirus vaccine. *J Am Med Assoc* **164**, 1216-1223.

Sabin, A. B. & Ward, R. (1941). The Natural History of Human Poliomyelitis : I. Distribution of Virus in Nervous and Non-Nervous Tissues. *J Exp Med* **73**, 771-793.

Salk, J. E., Krech, U., Youngner, J. S., Bennett, B. L., Lewis, L. J. & Bazeley, P. L. (1954). Formaldehyde treatment and safety testing of experimental poliomyelitis vaccines. *Am J Public Health Nations Health* **44**, 563-570.

Sarkany, Z., Skern, T. & Polgar, L. (2000). Characterization of the active site thiol group of rhinovirus 2A proteinase. *FEBS Lett* **481**, 289-292.

Savolainen, C., Blomqvist, S., Mulders, M. N. & Hovi, T. (2002). Genetic clustering of all 102 human rhinovirus prototype strains: serotype 87 is close to human enterovirus 70. *J Gen Virol* **83**, 333-340.

Schechter, I. & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157-162.

Schneider, T. D. & Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**, 6097-6100.

Schober, D., Kronenberger, P., Prchla, E., Blaas, D. & Fuchs, R. (1998). Major and minor receptor group human rhinoviruses penetrate from endosomes by different mechanisms. *J Virol* **72**, 1354-1364.

Schroth, M. K., Grimm, E., Frindt, P., Galagan, D. M., Konno, S. I., Love, R. & Gern, J. E. (1999). Rhinovirus replication causes RANTES production in primary bronchial epithelial cells. *Am J Respir Cell Mol Biol* **20**, 1220-1228.

Seemungal, T., Harper-Owen, R., Bhowmik, A., Moric, I., Sanderson, G., Message, S., Maccallum, P., Meade, T. W., Jeffries, D. J., Johnston, S. L. & Wedzicha, J. A. (2001). Respiratory viruses, symptoms, and inflammatory markers in acute exacerbations and stable chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* **164**, 1618-1623.

Seipelt, J., Liebig, H. D., Sommergruber, W., Gerner, C. & Kuechler, E. (2000). 2A proteinase of human rhinovirus cleaves cytokeratin 8 in infected HeLa cells. *J Biol Chem* **275**, 20084-20089.

Seth, S., Maier, M. K., Qiu, Q., Ravens, I., Kremmer, E., Forster, R. & Bernhardt, G. (2007). The murine pan T cell marker CD96 is an adhesion receptor for CD155 and nectin-1. *Biochem Biophys Res Commun* **364**, 959-965.

Shatkin, A. J. (1962). Actinomycin inhibition of ribonucleic acid synthesis and poliovirus infection of HeLa cells. *Biochim Biophys Acta* **61**, 310-313.

Shen, Y., Igo, M., Yalamanchili, P., Berk, A. J. & Dasgupta, A. (1996). DNA binding domain and subunit interactions of transcription factor IIIC revealed by dissection with poliovirus 3C protease. *Mol Cell Biol* **16**, 4163-4171.

Simmonds, P., McIntyre, C., Savolainen-Kopra, C., Tapparel, C., Mackay, I. M. & Hovi, T. Proposals for the classification of human rhinovirus species C into genotypically assigned types. *J Gen Virol* **91**, 2409-2419.

Skern, T., Sommergruber, W., Auer, H., Volkmann, P., Zorn, M., Liebig, H. D., Fessl, F., Blaas, D. & Kuechler, E. (1991). Substrate requirements of a human rhinoviral 2A proteinase. *Virology* **181**, 46-54.

Skern, T., Sommergruber, W., Blaas, D., Gruendler, P., Fraundorfer, F., Pieler, C., Fogy, I. & Kuechler, E. (1985). Human rhinovirus 2: complete nucleotide sequence and proteolytic processing signals in the capsid protein region. *Nucleic Acids Res* **13**, 2111-2126.

Skinner, M. S., Halperen, S. & Harkin, J. C. (1968). Cytoplasmic membrane-bound vesicles in echovirus 12-infected cells. *Virology* **36**, 241-253.

Slabinski, L., Jaroszewski, L., Rodrigues, A. P., Rychlewski, L., Wilson, I. A., Lesley, S. A. & Godzik, A. (2007a). The challenge of protein structure determination--lessons from structural genomics. *Protein Sci* **16**, 2472-2482.

Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A. & Godzik, A. (2007b). XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* **23**, 3403-3405.

Sledz, P., Zheng, H., Murzyn, K., Chruszcz, M., Zimmerman, M. D., Chordia, M. D., Joachimiak, A. & Minor, W. (2010). New surface contacts formed upon reductive lysine methylation: improving the probability of protein crystallization. *Protein Sci* **19**, 1395-1404.

Sloan, K. E., Stewart, J. K., Treloar, A. F., Matthews, R. T. & Jay, D. G. (2005). CD155/PVR enhances glioma cell dispersal by regulating adhesion signaling and focal adhesion dynamics. *Cancer Res* **65**, 10930-10937.

Smialowski, P., Doose, G., Torkler, P., Kaufmann, S. & Frishman, D. PROSO II--a new method for protein solubility prediction. *FEBS J* **279**, 2192-2200.

Smith, C. P., Clements, G. B., Riding, M. H., Collins, P., Bottazzo, G. F. & Taylor, K. W. (1998). Simultaneous onset of type 1 diabetes mellitus in identical infant twins with enterovirus infection. *Diabet Med* **15**, 515-517.

Snyers, L., Zwickl, H. & Blaas, D. (2003). Human rhinovirus type 2 is internalized by clathrin-mediated endocytosis. *J Virol* **77**, 5360-5369.

Sommergruber, W., Ahorn, H., Klump, H., Seipelt, J., Zoephel, A., Fessl, F., Krystek, E., Blaas, D., Kuechler, E., Liebig, H. D. & et al. (1994a). 2A proteinases of coxsackie- and rhinovirus cleave peptides derived from eIF-4 gamma via a common recognition motif. *Virology* **198**, 741-745.

Sommergruber, W., Ahorn, H., Zophel, A., Maurer-Fogy, I., Fessl, F., Schnorrenberg, G., Liebig, H. D., Blaas, D., Kuechler, E. & Skern, T. (1992). Cleavage specificity on synthetic peptide substrates of human rhinovirus 2 proteinase 2A. *J Biol Chem* **267**, 22639-22644.

Sommergruber, W., Casari, G., Fessl, F., Seipelt, J. & Skern, T. (1994b). The 2A proteinase of human rhinovirus is a zinc containing enzyme. *Virology* **204**, 815-818.

Sommergruber, W., Seipelt, J., Fessl, F., Skern, T., Liebig, H. D. & Casari, G. (1997). Mutational analyses support a model for the HRV2 2A proteinase. *Virology* **234**, 203-214.

Sommergruber, W., Zorn, M., Blaas, D., Fessl, F., Volkmann, P., Maurer-Fogy, I., Pallai, P., Merluzzi, V., Matteo, M., Skern, T. & et al. (1989). Polypeptide 2A of human rhinovirus

type 2: identification as a protease and characterization by mutational analysis. *Virology* **169**, 68-77.

Sonenberg, N., Morgan, M. A., Merrick, W. C. & Shatkin, A. J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc Natl Acad Sci U S A* **75**, 4843-4847.

Sousa, C. (2008). Investigating the specificity of human rhinoviruses 2A proteinase. In *Department for medical biochemistry*, p. 147. Vienna: University of Vienna.

Sousa, C., Schmid, E. M. & Skern, T. (2006). Defining residues involved in human rhinovirus 2A proteinase substrate recognition. *FEBS Lett* **580**, 5713-5717.

Stanway, G., Hughes, P. J., Mountford, R. C., Minor, P. D. & Almond, J. W. (1984). The complete nucleotide sequence of a common cold virus: human rhinovirus 14. *Nucleic Acids Res* **12**, 7859-7875.

Staunton, D. E., Marlin, S. D., Stratowa, C., Dustin, M. L. & Springer, T. A. (1988). Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families. *Cell* **52**, 925-933.

Staunton, D. E., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D. & Springer, T. A. (1989). A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell* **56**, 849-853.

Stott, E. J. & Heath, G. F. (1970). Factors affecting the growth of Rhinovirus 2 in suspension cultures of L132 cells. *J Gen Virol* **6**, 15-24.

Su, R. T. & Taylor, M. W. (1976). Morphogenesis of picornaviruses: characterization and assembly of bovine enterovirus subviral particles. *J Gen Virol* **30**, 317-328.

Sun, Y., Xue, F., Guo, Y., Ma, M., Hao, N., Zhang, X. C., Lou, Z., Li, X. & Rao, Z. (2009). Crystal structure of porcine reproductive and respiratory syndrome virus leader protease Nsp1alpha. *J Virol* **83**, 10931-10940.

Svitkin, Y. V., Gorbalenya, A. E., Kazachkov, Y. A. & Agol, V. I. (1979). Encephalomyocarditis virus-specific polypeptide p22 possessing a proteolytic activity: preliminary mapping on the viral genome. *FEBS Lett* **108**, 6-9.

Svitkin, Y. V., Pause, A., Haghighat, A., Pyronnet, S., Witherell, G., Belsham, G. J. & Sonenberg, N. (2001). The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *RNA* **7**, 382-394.

The Committee on Typing of the National Foundation for Infantile Paralysis (1951). Immunologic classification of poliomyelitis viruses. *Am J Hyg* **54**, 191-274.

Todd, S., Towner, J. S. & Semler, B. L. (1997). Translation and replication properties of the human rhinovirus genome in vivo and in vitro. *Virology* **229**, 90-97.

Tomassini, J. E., Graham, D., DeWitt, C. M., Lineberger, D. W., Rodkey, J. A. & Colonno, R. J. (1989). cDNA cloning reveals that the major group rhinovirus receptor on HeLa cells is intercellular adhesion molecule 1. *Proc Natl Acad Sci U S A* **86**, 4907-4911.

Tomko, R. P., Xu, R. & Philipson, L. (1997). HCAR and MCAR: the human and mouse cellular receptors for subgroup C adenoviruses and group B coxsackieviruses. *Proc Natl Acad Sci U S A* **94**, 3352-3356.

Toyoda, H., Nicklin, M. J., Murray, M. G., Anderson, C. W., Dunn, J. J., Studier, F. W. & Wimmer, E. (1986). A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell* **45**, 761-770.

Tsang, S. K., McDermott, B. M., Racaniello, V. R. & Hogle, J. M. (2001). Kinetic analysis of the effect of poliovirus receptor on viral uncoating: the receptor as a catalyst. *J Virol* **75**, 4984-4989.

Tuthill, T. J., Bubeck, D., Rowlands, D. J. & Hogle, J. M. (2006). Characterization of early steps in the poliovirus infection process: receptor-decorated liposomes induce conversion of the virus to membrane-anchored entry-intermediate particles. *J Virol* **80**, 172-180.

Umezawa, H., Aoyagi, T., Morishima, H., Kunimoto, S. & Matsuzaki, M. (1970). Chymostatin, a new chymotrypsin inhibitor produced by actinomycetes. *J Antibiot (Tokyo)* **23**, 425-427.

Umezawa, H., Aoyagi, T., Okura, A., Morishima, H. & Takeuchi, T. (1973). Letter: Elastatinal, a new elastase inhibitor produced by actinomycetes. *J Antibiot (Tokyo)* **26**, 787-789.

Vagin, A. & Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22-25.

van Ooij, M. J., Vogt, D. A., Paul, A., Castro, C., Kuijpers, J., van Kuppeveld, F. J., Cameron, C. E., Wimmer, E., Andino, R. & Melchers, W. J. (2006). Structural and functional characterization of the coxsackievirus B3 CRE(2C): role of CRE(2C) in negative- and positive-strand RNA synthesis. *J Gen Virol* **87**, 103-113.

Ventoso, I., MacMillan, S. E., Hershey, J. W. & Carrasco, L. (1998). Poliovirus 2A proteinase cleaves directly the eIF-4G subunit of eIF-4F complex. *FEBS Lett* **435**, 79-83.

Verdaguer, N., Fita, I., Reithmayer, M., Moser, R. & Blaas, D. (2004). X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. *Nat Struct Mol Biol* **11**, 429-434.

Virgen-Slane, R., Rozovics, J. M., Fitzgerald, K. D., Ngo, T., Chou, W., van der Heden van Noort, G. J., Filippov, D. V., Gershon, P. D. & Semler, B. L. An RNA virus hijacks an incognito function of a DNA repair enzyme. *Proc Natl Acad Sci U S A* **109**, 14634-14639.

Vlasak, M., Goesler, I. & Blaas, D. (2005). Human rhinovirus type 89 variants use heparan sulfate proteoglycan for cell attachment. *J Virol* **79**, 5963-5970.

Voss, T., Meyer, R. & Sommergruber, W. (1995). Spectroscopic characterization of rhinoviral protease 2A: Zn is essential for the structural integrity. *Protein Sci* **4**, 2526-2531.

Walter, T. S., Meier, C., Assenberg, R., Au, K. F., Ren, J., Verma, A., Nettleship, J. E., Owens, R. J., Stuart, D. I. & Grimes, J. M. (2006). Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617-1622.

Wang, Q. M., Johnson, R. B., Jungheim, L. N., Cohen, J. D. & Villarreal, E. C. (1998a). Dual inhibition of human rhinovirus 2A and 3C proteases by homophthalimides. *Antimicrob Agents Chemother* **42**, 916-920.

Wang, Q. M., Johnson, R. B., Sommergruber, W. & Shepherd, T. A. (1998b). Development of in vitro peptide substrates for human rhinovirus-14 2A protease. *Arch Biochem Biophys* **356**, 12-18.

Wenner, H. A., Kamitsuka, P., Lenahan, M. & Archetti, I. (1960). The pathogenesis of poliomyelitis. Sites of multiplication of poliovirus in cynomolgus monkeys after alimentary infection. *Arch Gesamte Virusforsch* **9**, 537-558.

Westendorf, W. (1992). *Erwachen der Heilkunst. Die Medizin im alten Ägypten.*: Artemis & Winkler.

Wetter, L. R. & Deutsch, H. F. (1951). Immunological studies on egg white proteins. IV. Immunochemical and physical studies of lysozyme. *J Biol Chem* **192**, 237-242.

WHO (2012). Polioeradication.org.

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* **112**, 531-552.

Wilkinson, D. L. & Harrison, R. G. (1991). Predicting the solubility of recombinant proteins in Escherichia coli. *Biotechnology (N Y)* **9**, 443-448.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D %@ 0907-4449* **67**, 235-242.

Wisdom, A., Leitch, E. C., Gaunt, E., Harvala, H. & Simmonds, P. (2009). Screening respiratory samples for detection of human rhinoviruses (HRVs) and enteroviruses: comprehensive

VP4-VP2 typing reveals high incidence and genetic diversity of HRV species C. *J Clin Microbiol* **47**, 3958-3967.

Wutz, G., Auer, H., Nowotny, N., Grosse, B., Skern, T. & Kuechler, E. (1996). Equine rhinovirus serotypes 1 and 2: relationship to each other and to aphthoviruses and cardioviruses. *J Gen Virol* **77 (Pt 8)**, 1719-1730.

Xing, L., Tjarnlund, K., Lindqvist, B., Kaplan, G. G., Feigelstock, D., Cheng, R. H. & Casasnovas, J. M. (2000). Distinct cellular receptor interactions in poliovirus and rhinoviruses. *EMBO J* **19**, 1207-1216.

Yalamanchili, P., Banerjee, R. & Dasgupta, A. (1997). Poliovirus-encoded protease 2APro cleaves the TATA-binding protein but does not inhibit host cell RNA polymerase II transcription in vitro. *J Virol* **71**, 6881-6886.

Yalamanchili, P., Harris, K., Wimmer, E. & Dasgupta, A. (1996). Inhibition of basal transcription by poliovirus: a virus- encoded protease (3Cpro) inhibits formation of TBP-TATA box complex in vitro. *J Virol* **70**, 2922-2929.

Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. & Russell, D. W. (1984). The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* **39**, 27-38.

Yang, W. X., Terasaki, T., Shiroki, K., Ohka, S., Aoki, J., Tanabe, S., Nomura, T., Terada, E., Sugiyama, Y. & Nomoto, A. (1997). Efficient delivery of circulating poliovirus to the central nervous system independently of poliovirus receptor. *Virology* **229**, 421-428.

Yeh, J. I., Biemann, H. P., Prive, G. G., Pandit, J., Koshland, D. E., Jr. & Kim, S. H. (1996). High-resolution structures of the ligand binding domain of the wild-type bacterial aspartate receptor. *J Mol Biol* **262**, 186-201.

Ypma-Wong, M. F., Dewalt, P. G., Johnson, V. H., Lamb, J. G. & Semler, B. L. (1988). Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. *Virology* **166**, 265-270.

Yu, S. F., Benton, P., Bovee, M., Sessions, J. & Lloyd, R. E. (1995). Defective RNA replication by poliovirus mutants deficient in 2A protease cleavage activity. *J Virol* **69**, 247-252.

Zautner, A. E., Jahn, B., Hammerschmidt, E., Wutzler, P. & Schmidtke, M. (2006). N- and 6-O-sulfated heparan sulfates mediate internalization of coxsackievirus B3 variant PD into CHO-K1 cells. *J Virol* **80**, 6629-6636.

Zautner, A. E., Korner, U., Henke, A., Badorff, C. & Schmidtke, M. (2003). Heparan sulfates and coxsackievirus-adenovirus receptor: each one mediates coxsackievirus B3 PD infection. *J Virol* **77**, 10071-10077.

Zhang, P., Mueller, S., Morais, M. C., Bator, C. M., Bowman, V. D., Hafenstein, S., Wimmer, E. & Rossmann, M. G. (2008). Crystal structure of CD155 and electron microscopic studies of its complexes with polioviruses. *Proc Natl Acad Sci U S A* **105**, 18284-18289.

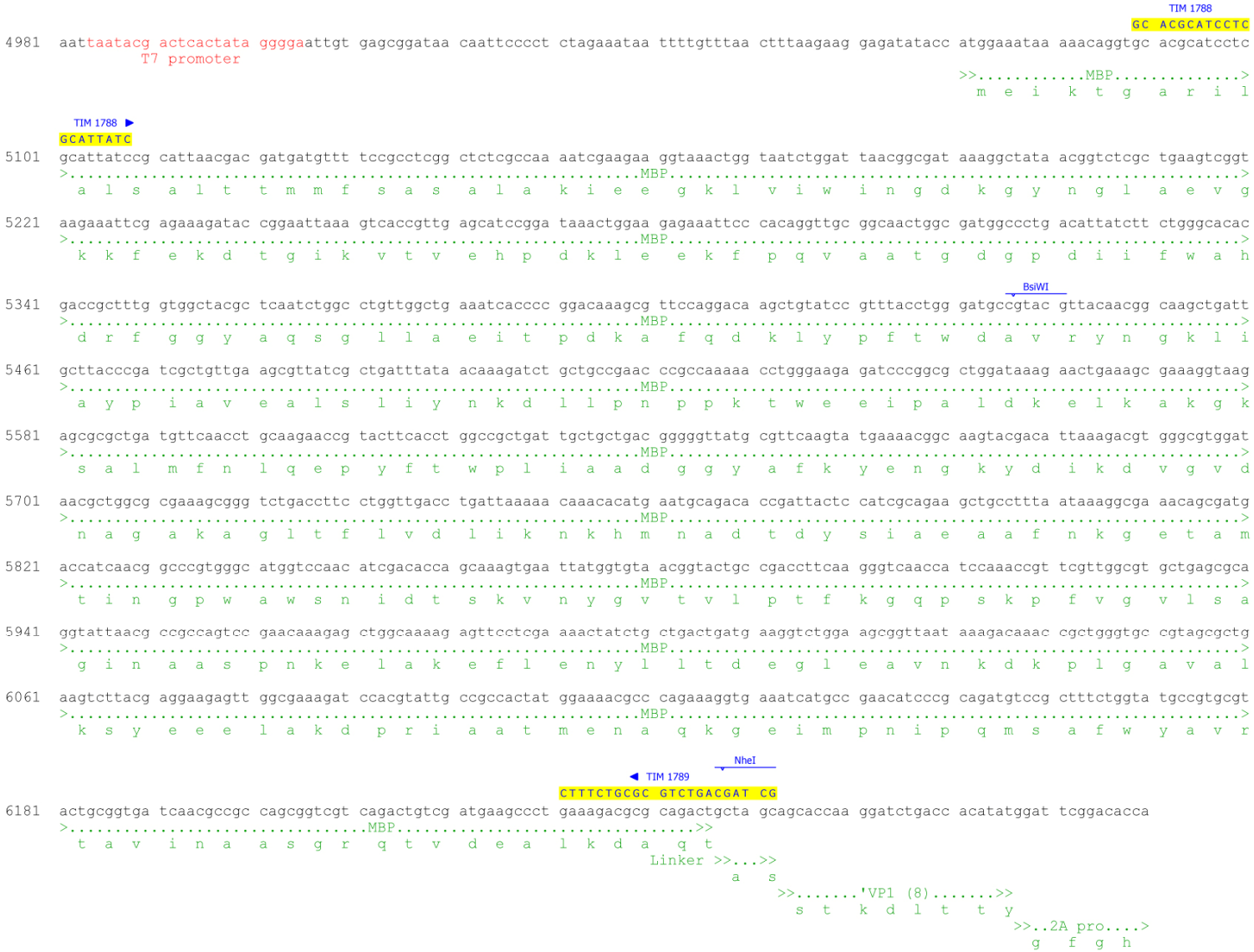
Zoll, J., Heus, H. A., van Kuppeveld, F. J. & Melchers, W. J. (2009). The structure-function relationship of the enterovirus 3'-UTR. *Virus Res* **139**, 209-216.

7 APPENDIX

APPENDIX I **pET 3d PV1 $VP1_8$-$2A^{pro}$ (C109A)**

APPENDIX II **pET 3d PV1 His$_6$-VP1$_8$-2A^{pro} (C109A)**

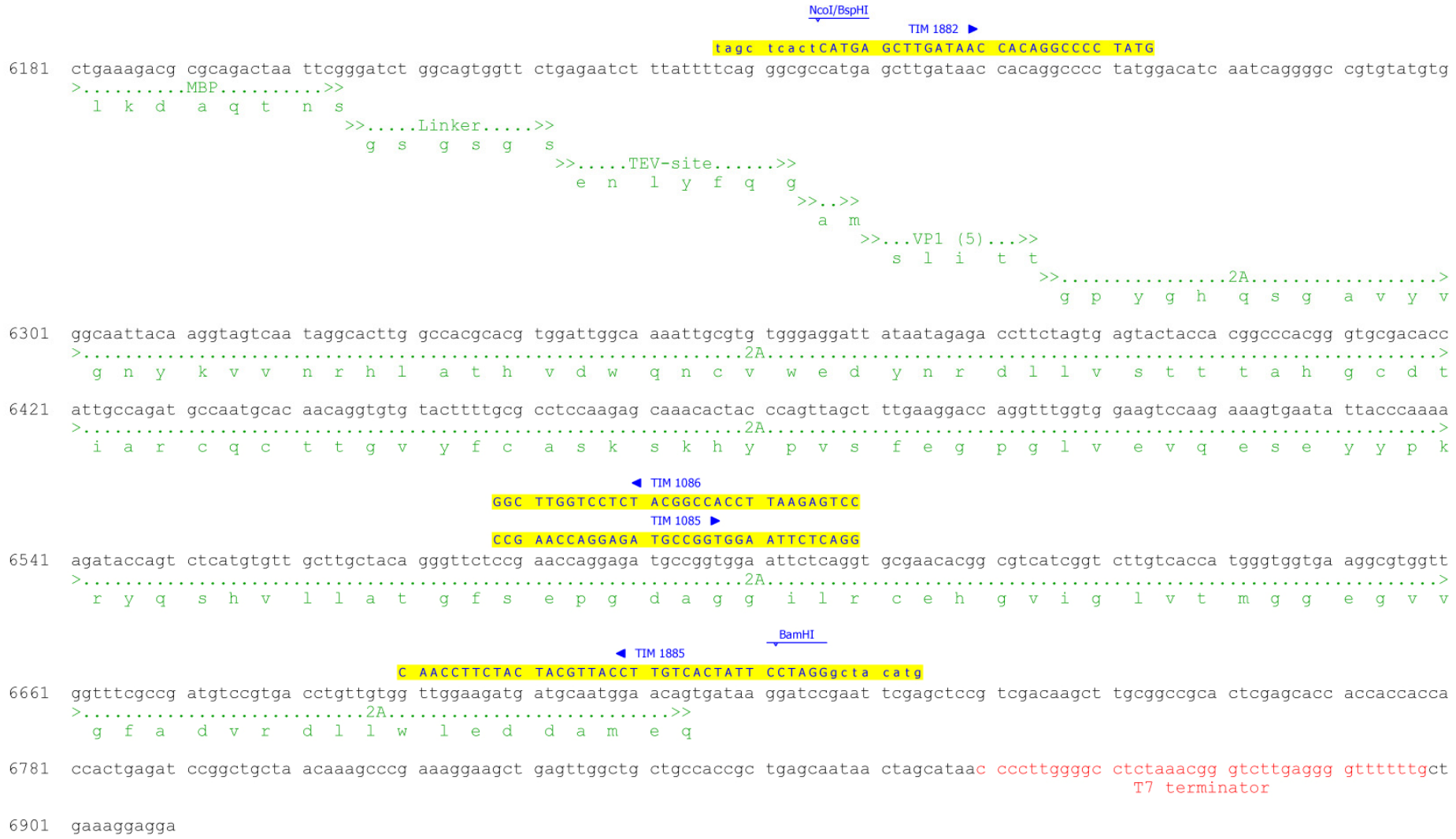
APPENDIX III **Mal pET PV1 VP1$_8$-2A^{pro} (C109A)**



APPENDIX IV **Mal pET PV1 VP1$_8$-2A^{pro} (C109A) short linker**

APPENDIX V **Mal pET PV1 $VP1_8$-$2A^{pro}$ (C109A) short linker His_6**

APPENDIX VI pET MBP HRV2 $VP1_9$-$2A^{pro}$ (C106A)

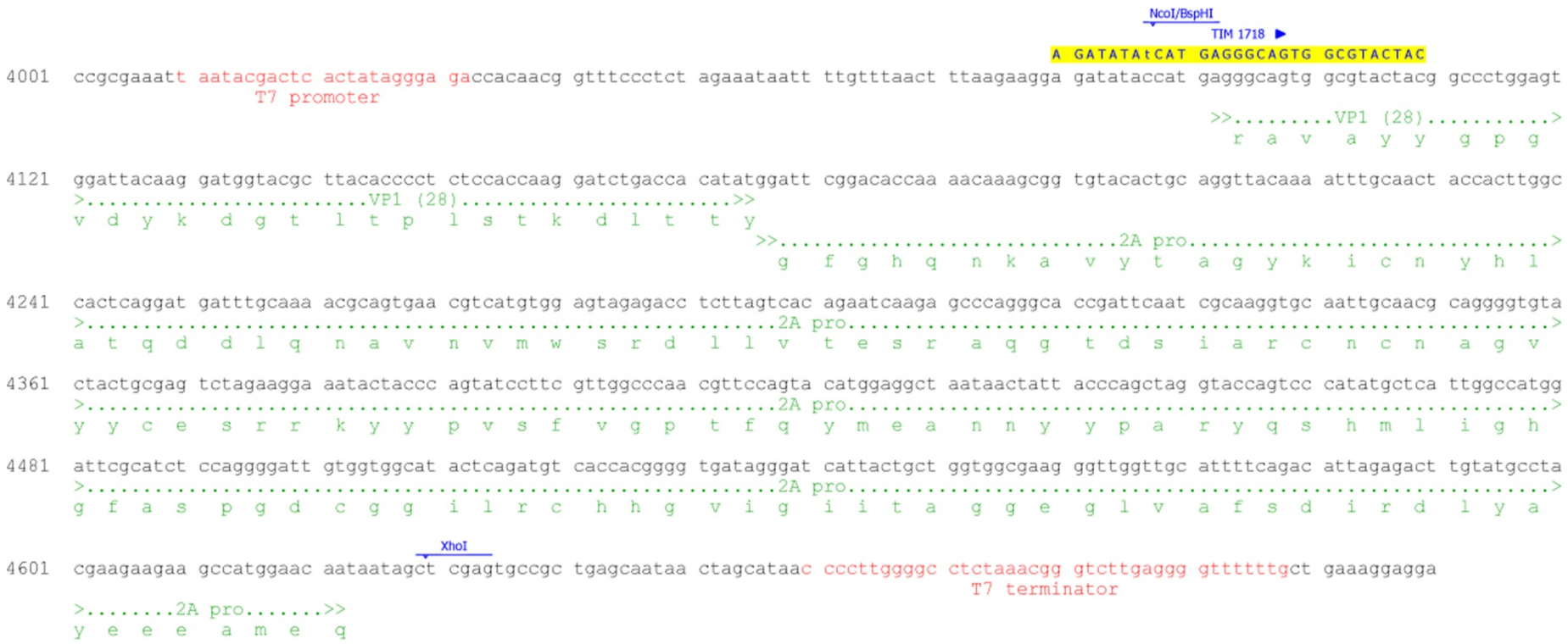


APPENDIX VII pET MBP CVB4 VP1$_5$-2A^{pro} (C110A)

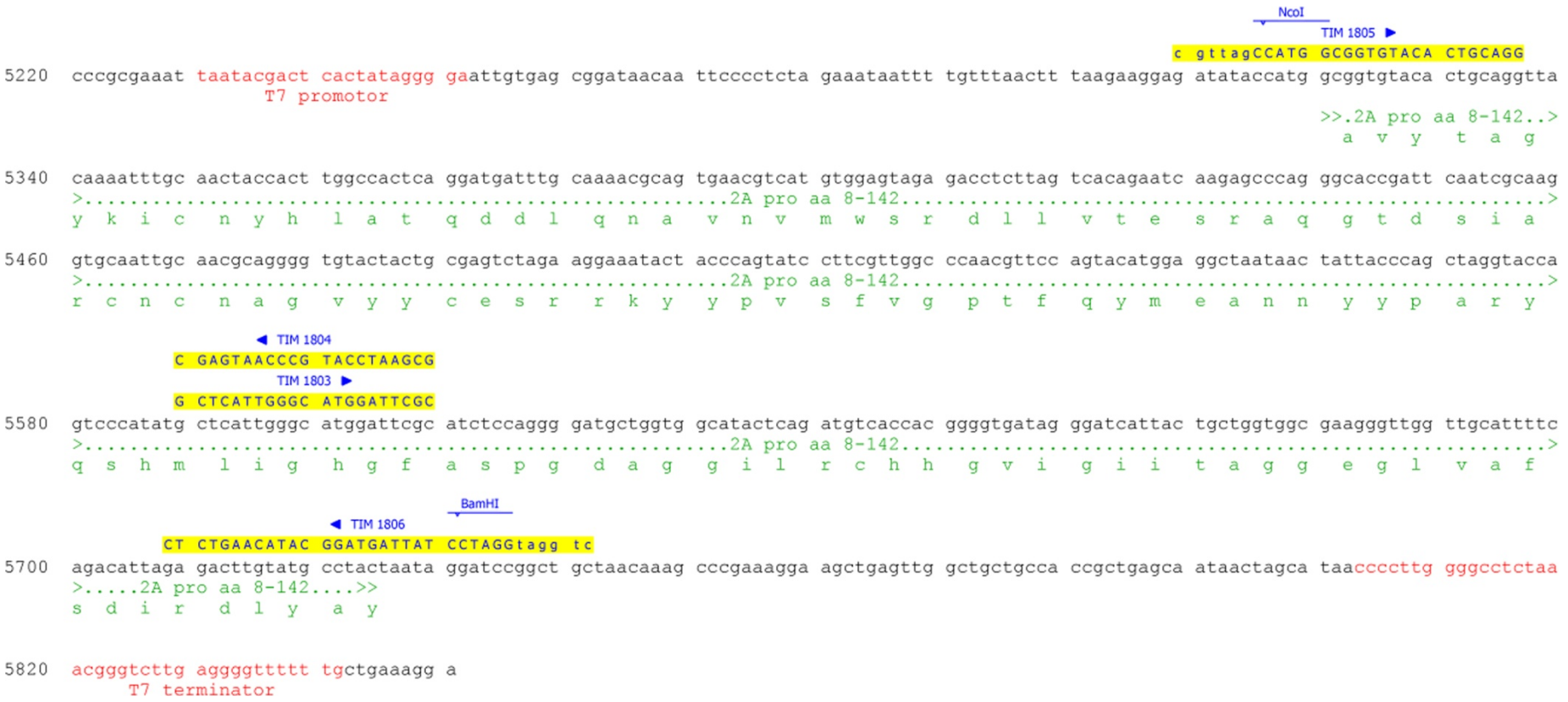
APPENDIX VIII Mal pET CVB4 VP1$_8$-2A^{pro} (C110A)

APPENDIX IX pET MBP CVB4 VP1$_{12}$-2A^{pro} (C110A)

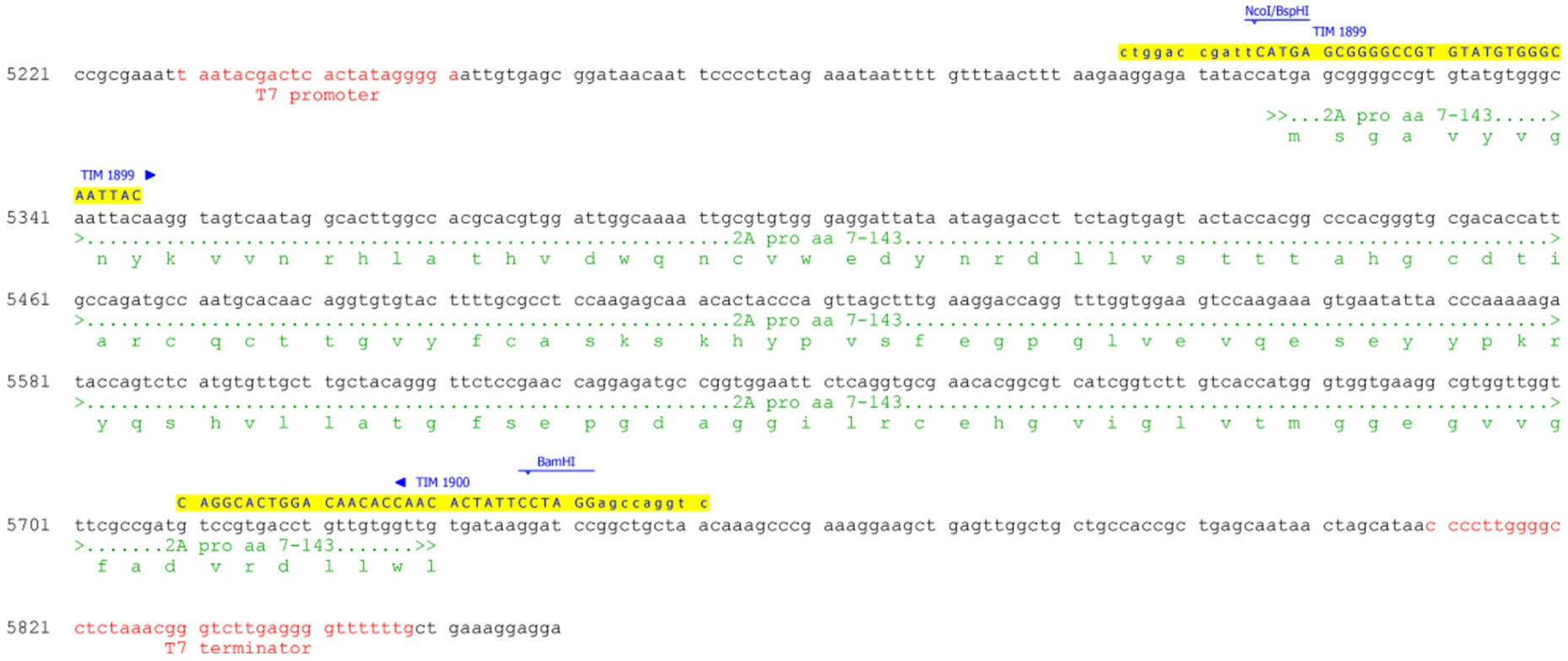
APPENDIX X **pET MBP CVB4 VP1$_{20}$-2A^{pro} (C110A)**



APPENDIX XI **pET 3d PV1 $VP1_{28}$-$2A^{pro}$**



APPENDIX XII pET 11d PV1 2A^{pro} (C109A) aa 8-142



APPENDIX XIII pET 11d CVB4 2A^{pro} (C110A) aa 7-143

		Hampton Research			Molecular Dimensions		QIAGEN	Emerald Biosystems	
Protein variant	**Chapter**	**Crystal Screen I+II**	**Index I + II**	**SaltRx I+II**	**Pact I+II**	**Morpheus**	**JCSG + Suite**	**Cryo I+II**	**Wizard I+II**
PV1 MBP-His_{10}-$VP1_8$-$2A^{pro}$ (C109A)	*4.1.1.3*	x	x		x		x	x	x
PV1 MBP-$VP1_8$-$2A^{pro}$ (C109A) short linker-His6	*4.1.1.5*	x	x	x	x	x	x	x	x
CVB4 His_6-$VP1_8$-$2A^{pro}$ (C110A)	*4.1.3.1*	x	x		x		x	x	x
CVB4 $VP1_5$-$2A^{pro}$ (C110A)	*4.1.4.1*	x	x		x		x	x	x
CVB4 $VP1_8$-$2A^{pro}$ (C110A)	*4.1.4.2*	x	x			x	x	x	x
CVB4 $VP1_{12}$-$2A^{pro}$ (C110A)	*4.1.4.3*	x	x		x	x	x	x	x
CVB4 $VP1_{20}$-$2A^{pro}$ (C110A)	*4.1.4.4*	x	x		x	x	x	x	x
CVB4 $VP1_{12}$-$2A^{pro}$ (C110A) methylated	*4.1.4.6*	x			x	x	x	x	x
CVB4 $2A^{pro}$ active	*4.3.1.1*	x	x		x				x

APPENDIX XIV **Overview of commercial crystal screens used with different protein variants. The crystallization setup is described in materials and methods.**

APPENDIX XV List of abbreviations

aa	Amino acid
Amp	Ampicillin
APS	Ammoniumpersulfate
BBB	Blood brain barrier
BC	Before christ
BCIP	5-Bromo-4-chloro-3-indolyl phosphate
BME	β-mercaptoethanol
bp	Base pair
CAR	Coxsackie and adenovirus receptor
CIP	Calf intestine phosphatase
CITE	Cap-independent translation enhancer
CNS	Central nervous system
COPD	Chronic obstructive pulmonary disease
CRE	*Cis* replicative element
CV	Coxsackievirus
DAF	Decay-accelerating factor
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNA-PK(CS)	Catalytic subunit of the DNA-dependent protein kinase
DTT	Dithiothreitol
eIF	Eeukaryotic initiation factor
EM	Electron microscope
EMCV	Encephalomyocarditis virus
EV	Enterovirus
FMDV	Foot and mouth disease virus
FPLC	Fast protein liquid chromatography
GFP	Green fluorescent protein
Gluc	Glucose
gpi	Glycosyl phosphatidylinositol
GST	Glutathione S-transferase
HAV	Hepatitis A virus
HeLa	Henrietta Lacks
HIV	Human immunodeficiency virus
HRV	Human Rhinovirus
IC_{50}	Half maximal inhibitory concentration

ICAM-1	Intracellular adhesion molecule 1
ICTV	International committee on taxonomy of viruses
IFN-α	Interferon alpha
Ig	Immunoglobulin
IPTG	Isopropyl-β-D-thiogalactopyranosid
IPV	Inactivated polio vaccine
IRES	Internal ribosomal entry site
ISG	Interferon stimulated genes
ITAF	IRES *trans*-acting factors
Kan	Kanamycin
kb	Kilo bases
kDa	Kilo dalton
LB	Luria Bertani
LDLR	Low density lipoprotein receptor
LFA-1	Function-associated antigen-1
Mac-1	Macrophage-1 antigen
MBP	Maltose binding protein
M-cells	Microfold cells
mRNA	Messenger RNA
NBT	Nitro blue tetrazolium chloride
NEM	N-Ethylmaleimide
NMR	Nuclear magnetic resonance
NPC	Nuclear pore complex
nt	Nucleotide
Nup	Nuclear pore protein
OD_{600}	Optical density (600nm)
ORF	Open reading frame
PABP	Poly(A) binding protein
PAGE	Poly Acrylamide gel electrophoresis
PARP	Poly (ADP-ribose) polymerase
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline / Tween
PCBP2	Protein poly(rC)-binding protein
PCR	Polymerase chain reaction
PNK	Poly-nucleotide kinase
Poly(A)	Poly adenosin
PTB	Polypyrimidine tract-binding protein

PV	Polio virus
PVDF	Polyvinylidenfluorid
PVR	Polio virus receptor
RdRP	RNA dependent RNA polymerase
Rmsd	Root mean square deviation
RNA	Ribonucleic acid
SCR	Short consensus repeat
SDS	Sodium dodecyl sulfate
SGPB	Streptomyces griseus protease B
siRNA	Small interfering RNA
snRNP	Small nuclear ribonucleoproteins
T1D	Insulin dependent (type 1) diabetes mellitus
TDP2	5'-tyrosyl-DNA phosphodiesterase-2
TEMED	N,N,N’,N’,-Tetramethylethylendiamin
TEV	Tobacco etch virus
TLCK	Tosyl lysyl chloromethyl ketone
tOPV	Trivalent oral polio vaccine
tRNA	Transport RNA
UTR	Untranslated region
UV	Ultra violet
VLDLR	Very low density lipoprotein receptor
VPg	Viral protein, genome linked
WHO	World health organization
WPV	Wild polio virus
wt	Wild type

APPENDIX XVI List of amino acids

Amino acid	One letter code	Three letter code
Glycine	G	Gly
Alanine	A	Ala
Valine	V	Val
Leucine	L	Leu
Isoleucine	I	Ile
Proline	P	Pro
Methionine	M	Met
Histidine	H	His
Cysteine	C	Cys
Phenylalanine	F	Phe
Tryptophane	W	Trp
Tyrosine	Y	Tyr
Serine	S	Ser
Threonine	T	Thr
Asparagine	N	Asn
Glutamine	Q	Gln
Aspertate	D	Asp
Glutamate	E	Glu
Lysine	K	Lys
Arginine	R	Arg

CURRICULUM VITAE

David Neubauer

Vogelsanggasse 43/32
1050 Vienna
t: +43 650 5314959
m: david.neubauer@aon.at

PERSONAL INFORMATION

DATE OF BIRTH	16.11.1981
PLACE OF BIRTH	Linz
CITIZENSHIP	Austrian
PERSONAL STATUS	Unmarried

EDUCATION

NOVEMBER 2008 – to date	**PhD student at the University of Vienna**, FWF funded collegiate in the DK plus "Structure and Interaction of Biological Macromolecules"; Laboratory of ao. Univ-Prof. Dr. Timothy Skern. Title of PhD thesis: *"Investigating the substrate specificity of enteroviral 2A proteases – an attempt to solve high resolution structures of enzyme-substrate complexes."* (To be submitted).
29. OCTOBER 2008	**Diploma exam passed with honours, Mag. rer. nat.**
MAY 2007- OCTOBER 2008	Diploma thesis in the laboratory of ao. Univ.-Prof. Dr. Timothy Skern (Medical University of Vienna, Department of Medical Biochemistry): *"Investigating substrate specificity of picornaviral proteinases by mutational analysis."*
2000-2008	**Student of chemistry at the University of Vienna**
1995-1999	High school (BORG Honauerstraße) in Linz; Matura with first-class honours

ADDITIONAL EDUCATION

FEBRUARY 2013	"Quality management in pharmaceutical industry, production of medical devices and laboraatories." WIFI, Vienna.

PUBLICATIONS

Neubauer, D., Aumayr, M., Gösler, I. & Skern, T. (2012) Specificity of human rhinovirus $2A^{pro}$ is determined by combined spatial properties of four cleavage site residues. (Accepted for publication in J. Gen. Virol.).

Palmenberg, A. C., Neubauer, D. & Skern, T. (2010). Genome organization and encoded proteins. In *The picornaviruses*, pp. 3-18. Edited by E. Ehrenfeld, E. Domingo & R. P. Roos: ASM Press.

Neubauer, D., Steinberger, J. & Skern, T. (2009). Picornaviruses. In *Viral proteases and antiviral protease inhibitor therapy*, 1 edn, pp. 101-130. Edited by a. N. M. H. Uwe Lendeckel: Springer.

Mayer, C., Neubauer, D., Nchinda, A. T., Cencic, R., Trompf, K. & Skern, T. (2008). Residue L143 of the foot-and-mouth disease virus leader proteinase is a determinant of cleavage specificity. *J Virol* **82**, 4656-4659.

PRESENTATIONS / WORKSHOPS

JULY 2012	Attendance and poster presentation at EUROPIC 2012, St. Raphael, France.
SEPTEMBER 2010	Attendance and poster presentation at EUROPIC 2010, St. Andrews, Scotland.
SEPTEMBER 2009	Attendance and poster presentation at the annual meeting of the ÖGMBT (Austrian society for molecular biology and biotechnology), Innsbruck, Austria.
SEPTEMBER 2008	Attendance and poster presentation at the annual meeting of the ÖGBM (Austrian society for biochemistry and molecular biology), Graz, Austria.
AUGUST 2008	Attendance and poster presentation at the VIZIER/SPINE2 workshop "Structural virology", Vienna, Austria.
AUGUST 2007	Attendance at the EMBO workshop "RNA viruses: replication, evolution and drug design", Vienna, Austria.

ADDITIONAL SKILLS

LANGUAGES	German (first language), English (fluent, spoken and written)
COMPUTER SKILLS	Microsoft Windows, Microsoft Office, Adobe Photoshop, Linux, PyMOL, Clone Manager, CCP4 (basic), phenix (basic), …

LEBENSLAUF

David Neubauer

Vogelsanggasse 43/32
1050 Vienna
t: +43 650 5314959
m: david.neubauer@aon.at

PERSÖNLICHE INFORMATIONEN

GEBURTSDATUM	16.11.1981
GEBURTSORT	Linz
STAATSBÜRGERSCHAFT	Österreich
FAMILIENSTAND	Ledig

AUSBILDUNG

NOVEMBER 2008 – bis dato	**Doktoratsstudium an der Universität Wien**, vom FWF finanzierter Student im DK plus "Structure and Interaction of Biological Macromolecules"; Labor von ao. Univ.-Prof. Dr. Timothy Skern. Titel der Doktorarbeit: *„Investigating the substrate specificity of enteroviral 2A proteases – an attempt to solve high resolution structures of enzyme-substrate complexes."* (vor Einreichung).
29. OKTOBER 2008	**Diplomprüfung mit Auszeichnung bestanden, Mag. rer. nat.**
MAI 2007- OKTOBER 2008	Diplomarbeit im Labor von ao. Univ.-Prof. Dr. Timothy Skern (Medizinische Universität Wien, Department für medizinische Biochemie): *„Investigating substrate specificity of picornaviral proteinases by mutational analysis."*
2000-2008	**Studium der Chemie an der Universität Wien**
1995-1999	BORG Honauerstraße, Linz; Matura mit ausgezeichnetem Erfolg

ZUSÄTZLICHE AUSBILDUNG

FEBRUAR 2013	"Qualitätsmanagement für Pharmaindustrie, Medizinproduktehersteller und Labore." WIFI Wien.

PUBLIKATIONEN

Neubauer, D., Aumayr, M., Gösler, I. & Skern, T. (2012) Specificity of human rhinovirus $2A^{pro}$ is determined by combined spatial properties of four cleavage site residues. (Angenommen bei J. Gen. Virol.).

Palmenberg, A. C., Neubauer, D. & Skern, T. (2010). Genome organization and encoded proteins. In *The picornaviruses*, pp. 3-18. Edited by E. Ehrenfeld, E. Domingo & R. P. Roos: ASM Press.

Neubauer, D., Steinberger, J. & Skern, T. (2009). Picornaviruses. In *Viral proteases and antiviral protease inhibitor therapy*, 1 edn, pp. 101-130. Edited by a. N. M. H. Uwe Lendeckel: Springer.

Mayer, C., Neubauer, D., Nchinda, A. T., Cencic, R., Trompf, K. & Skern, T. (2008). Residue L143 of the foot-and-mouth disease virus leader proteinase is a determinant of cleavage specificity. *J Virol* **82**, 4656-4659.

KONFERENZEN / WORKSHOPS

JULI 2012	Teilnahme (inkl. Poster Präsentation) an EUROPIC 2012, St. Raphael, Frankreich.
SEPTEMBER 2010	Teilnahme (inkl. Poster Präsentation) an EUROPIC 2010, St. Andrews, Schottland.
SEPTEMBER 2009	Teilnahme (inkl. Poster Präsentation) an der Jahrestagung der ÖGMBT (Österreichische Gesellschaft für molekulare Biowissenschaften und
SEPTEMBER 2008	Teilnahme (inkl. Poster Präsentation) an der Jahrestagung der ÖGBM (Österreichische Gesellschaft für Biochemie und Molekularbiologie), Graz, Österreich. Biotechnologie), Innsbruck, Österreich.
AUGUST 2008	Teilnahme (inkl. Poster Präsentation) am VIZIER/SPINE2 workshop “Structural virology”, Wien, Österreich.
AUGUST 2007	Teilnahme am EMBO workshop “RNA viruses: replication, evolution and drug design”, Wien, Österreich.

ZUSÄTZLICHE FÄHIGKEITEN

SPRACHEN	Deutsch (Muttersprache), Englisch (fließend in Wort und Schrift)
EDV-KENNTNISSE	Microsoft Windows, Microsoft Office, Adobe Photoshop, Linux, PyMOL, Clone Manager, CCP4 (basic), phenix (basic), …