
# Rational design of the cold active subtilisin-like serine protease VPR towards higher activity and thermostability

Kristinn Ragnar Óskarsson

**Faculty of Physical Sciences**
**University of Iceland**
**2015**

# Rational design of the cold active subtilisin-like serine protease VPR towards higher activity and thermostability

Kristinn Ragnar Óskarsson

90 ECTS thesis submitted in partial fulfillment of a
*Magister Scientiarum* degree in Biochemistry

Advisor(s)
Magnús Már Kristjánsson, Professor
Sigríður Helga Þorbjarnardóttir, Research specialist

Faculty Representative
Hörður Filippusson, Emeritus

Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, June 2015

Rational design of the cold active subtilisin-like serine protease VPR towards higher activity and thermostability
Rational design of VPR towards higher activity and thermostability
90 ECTS thesis submitted in partial fulfillment of a *Magister Scientiarum* degree in Biochemistry



Faculty of Physical Sciences
School of Engineering and Natural Sciences
University of Iceland
Hjarðarhagi 2-6
107, Reykjavik
Iceland

Telephone: 525 4700

Bibliographic information:
Kristinn Ragnar Óskarsson, 2015, *Rational design of the cold active subtilisin-like serine protease VPR towards higher activity and thermostability*, Master's thesis, Faculty of Physical Sciences, University of Iceland, pp. 99.

Printing: Háskólaprent ehf.
Reykjavik, Iceland, May 2015

# Abstract

This research project builds on research previously done on the subtilisin-like serine proteinase VPR, from a psychrotrophic *Vibrio* species and its structural homologue aqualysin I (AQUI) from the thermophile *Thermus aquaticus.* We set out to design a mutant of VPR using site directed mutagenesis that would be both more stable against heat denaturation and retain the high activity of the wild type enzyme. Starting with two different templates, one being a C-terminal truncated form of the enzyme ($VPR_{\Delta C}$), containing two proline mutations, N3P/I5P, close to the N-terminus of the protein, which had shown increased stability but loss of catalytic activity. The ΔC truncated form was produced by introducing a mutation as a stop codon at C277 to imitate the structure of AQUI in more detail. The other template contained the ΔC mutation and additional six mutations on a loop that may act as a hinge for movements that are postulated to be important for catalysis. The A116T/Q117R/A119H/S120R/G121R/S123A (6x) mutant had shown an increase in activity without losing stability to any degree. On top of these templates two mutations were added; N15D and Q142K. The N15D mutation had been shown to introduce a salt bridge yielding higher stability but with no detrimental effects on activity. The Q142K exchange on the other hand increased significantly the catalytic activity of the enzyme. Thus, we attempted to improve stability of the two mutants by introducing the N15D mutation, while the Q142K mutations was added with the purpose of increasing catalytic activity. The $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K mutant was a success, giving an 8°C rise in the $T_m$, a 10°C rise to the $T_{50\%}$ and the catalytic activity was slightly higher than that of the wild type enzyme. The $VPR_{\Delta C}$/6x/N15D/Q142K led to a 3°C rise in both $T_m$ and $T_{50\%}$. To examine the effects of the N3P/I5P mutation on the flexibility of the structure, fluorescence quenching with acrylamide was preformed comparing AQUI, $VPR_{\Delta C}$ and $VPR_{\Delta C}$/N3P/I5P. The results indicated that the environment of Trp6 in $VPR_{\Delta C}$/N3P/I5P is not as accessible as in $VPR_{\Delta C}$ probably due to tighter packing of the N-terminus.

# Útdráttur

Með því að byggja á fyrri rannsóknum sem gerðar hafa verið á samstofna ensímunum VPR úr kuldakærri *Vibrio* örveru og aqualysin I (AQUI) úr örverunni *Thermus aquaticus* var ákveðið að reyna að hanna ensímhvata sem hefði meiri stöðugleika gagnvart hitaafmyndun og sem jafnframt hefði meiri virkni en villigerðin af VPR. Byrjað var með tvö mismunandi stökkbrigði sem upphafsensím, annars vegar var um að ræða C-enda stytt afbrigði af VPR ($VPR_{\Delta C}$) sem innihélt tvær prólín innsetningar nálægt N-enda ensímsins, N3P/I5P, sem sýnt hafði verið að leiddi til mikillar aukningar í stöðugleika, en jafnframt umtalsverðu tapi í virkni. C-enda styttan hafði verið framleidd með því að setja inn stop tákna í set sem svarar til stöðu C277, með það að markmiði að líkja enn frekar eftir byggingu AQUI. Hitt upphafsensímið sem notað var í rannsókninni innihélt sexföldu breytinguna A116T/Q117R/A119H/S120R/G121R/S123A (6x) á lykkjusvæði sem tilgátur eru um að virki sem hjöruliður sem talinn er mikilvægur fyrir hreyfingar við hvarfstöð ensímsins. Þetta stökkbrigði hafði verið hannað með það að markmiði að líkja sem mest eftir mjög frábrugðinni samsetningu AQUI á þessu svæði. Ofan á þessi grunnensím var svo valið að bæta við tveimur breytingum; N15D og Q142K. N15D breytingin hafði áður sýnt fram á að saltbrú myndaðist sem jók á stöðugleika án þess þó að hafa neikvæð áhrif á virkni. Hins vegar hefur verið sýnt fram á að Q142K stökkbreytingin hefur í för með sér mikla virkniaukningu án þess þó að hafa áhrif á stöðugleika, hugsanlega með því að trufla nálæga saltbrú og liðka þannig fyrir hreyfingum á α-helix sem tengist inni hvarfstöðina. Hannaða stökkbrigðið $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K heppnaðist vel, þar sem umtalsverð hækkun á $T_m$ átti sér stað, eða um 8°C. Einnig var hækkun á $T_{50\%}$ sem mældist um 10°C og var hvötunargetan aðeins hærri en hjá villigerðinni. $VPR_{\Delta C}$/6x/N15D/Q142K leiddi til hækkunar á bæði $T_m$ og $T_{50\%}$ uppá 3°C, en virknin var ekki eins góð og búist var við. Til að skoða betur áhrif N5P/I5P á sveigjanleika byggingar VPR og þar helst í kringum Trp6 var notast við flúrljómunarbælingu með akrílamíði og voru bælingarferlar fyrir AQUI, $VPR_{\Delta C}$ og $VPR_{\Delta C}$/N3P/I5P bornir saman. Kom í ljós að hugsanlega er pökkun meiri í kringum N-enda $VPR_{\Delta C}$/N3P/I5P. Hugsanleg ástæða er sú aðgengi að Trp6 í VPR er minna en í prólín stökkbrigðinu sem bendir til þéttari pökkunar þessa hliðarhóps í stökkbrigðinu í samanburði við C-enda styttu ensímsins.

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgements

I like to thank Magnús M. Kristjánsson my professor for the guidance and good talks on various topics. I thank Sigríður Helga Þorbjarnardóttir for the help during all the work at Askja. I also would like to thank Bjarni Ásgeirsson, Hörður Filippusson, Lilja B. Jónsdóttir and Jens G. Hjörleifsson for interesting talks and help during the project. And last but not least I would like to thank the staff of the Science Institute.

# 1 Introduction

Stability of proteins plays a crucial role for their utilization as industrial, analytical or therapeutic tools [1]. Protein stability can refer to the stability of the native state against heat, pressure, pH, solvent, chemical modifications or proteolysis [2,3]. The stability of the native state can be viewed as twofold, thermodynamic, that is governed by the equilibrium constants and free-energy barriers between unfolded, folded and various intermediate states [4,5], or kinetic, i.e. how rapidly a protein unfolds under a set of conditions. Kinetic stability is particularly important for proteins that unfold irreversibly, as is often the case for proteins that work under harsh conditions [6,7].

The primary tool in the effort of increasing the stability of a protein for utilization, is mutagenesis with either directed evolution methods, based on generating diversity and selecting the more stable variants [8,9], or by using site-directed mutagenesis building on rational design based on attempts to interfere with or enhance certain physicochemical attributes of the three dimensional structure of the protein [10,11]. Over the years many different strategies have been developed to stabilize proteins, such as charge optimization and salt bridge incorporation [12,13], proline substitutions [14], α-helical stabilization [15], disulfide bond incorporation [16,17], cavity filling [18], replacing buried polar residues and/or replacing external hydrophobic residues [19,20], decreasing the entropy of the denatured state [21] and by replacing specific hydrogen bonds [22]. Apparently, however, every protein has its own way in increasing stability and the approaches mentioned above can have the very opposite effect in some systems. In the next section I will discuss temperature adaptation and the stability of the native fold focusing on different strategies Nature has come up with to make biologically active proteins at various temperatures.

## 1.1 Temperature adaptation

Life as we know it on Earth undoubtedly arose in an aqueous medium, as everywhere we find life, water is always present. As a result, the course of evolution has been shaped by the chemical and physical properties of water [23]. Under normal conditions the melting

point of water is 0°C and the boiling point 100°C, but can vary depending on factors such as pressure and salinity [23]. As a consequence of this, life has too adapt to the different attributes of water at different temperatures. So what are the limits of life here on Earth? To answer that question we can look at single cell organism and the extremes the cells themselves need to handle. Studies suggest that the lower temperature limit may be around -20°C as below that temperature no known organisms are known to reproduce [24]. For the upper temperature limit *Methanopyrus kandleri* currently holds the record and can proliferate at temperatures up to 122°C [25]. From temperature limits of growth single cell organisms can roughly be divided into four categories; psychrophilic ($0 < T_{opt} < 20$°C), mesophilic ($20 < T_{opt} < 50$°C), thermophilic ($50 < T_{opt} < 80$°C) and hyperthermophilic ($80 < T_{opt} < 120$°C) [26]. The fact that a cell´s life relies on their proteins to fold correctly and stay in an active native state, ultimately leads to the evolution highly diverse populations of protein species in different organisms in terms of thermostability [27]. So what governs the stability of proteins in general?

### 1.1.1 Stability of globular proteins

Proteins exist in many shapes and sizes, but they are generally composed of twenty different L-α-amino acids, covalently linked in a linear sequence and range in size from small single domain proteins to large assemblies consisting of multiple polypeptide chains [2]. Their functionalities differ greatly, they can be found as connective tissue, membrane receptors and transporters, to water soluble proteins acting as transporters, hormones or enzymes [2,23]. The structure of water soluble globular proteins is largely dependent on their primary structure i.e. properties of the peptide-bond and different side-chains which restrain movements around the α-carbon only allowing certain secondary structures (Fig. 1) [23].

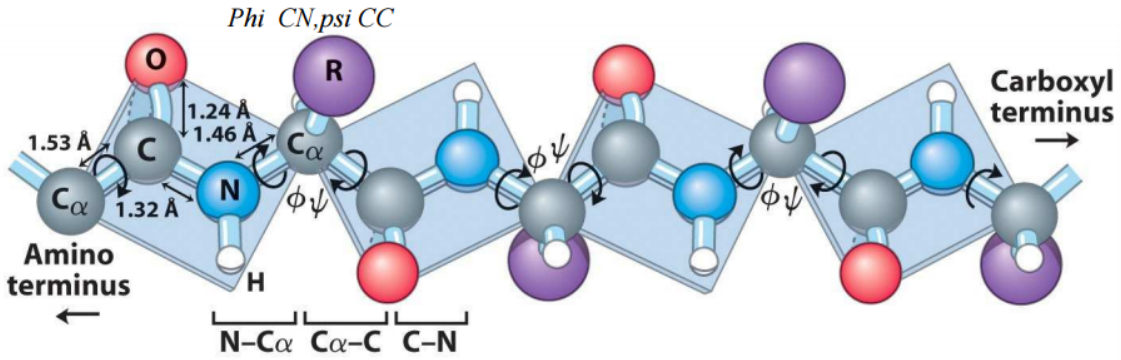


*Figure 1. A schematic representation of a peptide chain, the C-N bond is not free to rotate, while the N-Cα (φ) and Cα –C (ψ) can rotate to a certain degree varying with different R-groups [23].*

The most common secondary structures arising from these constraints are the α-helix and the β-strand and these structures are often connected by β-turns and loops. In globular proteins where relatively compact tertiary and quaternary structures are needed, turns and loops account for almost one third of their amino acid residues [23]. As a result the native globular state is highly ordered, but also needs to be dynamic. In terms of thermodynamics this structure derives its stability from two competing factors; enthalpy (ΔH) and entropy (ΔS) (Eq. I).

$$\Delta G = \Delta H - T\Delta S \qquad \text{(I)}$$

Contributing factors to the enthalpy are intramolecular non-covalent forces with the exception of energy from covalent disulfide bonds between two cysteine residues. Contributing factors to the non-covalent forces are; hydrogen-bonding, ion-ion interactions (such as salt bridges), van der Waals interactions and also factors like hydrophobic interactions between the aqueous phase and hydrophobic parts of the peptide chain and these forces are maximized in the native state. Enthalpy difference between the native and denatured state can differ by several hundred kcal/mol [2]. The other factor, entropy, derives from the second law of thermodynamics which states that energy is needed to create order and as mentioned above, native protein structures are highly ordered, so without the forces attributing to enthalpy, the denatured state is much more entropically favorable. The entropy difference between the native and the unfolded states can also differ by several hundred kcal/mol, thus resulting in a marginal stability in the native fold that is

only around 5-15 kcal/mol in terms of free energy for many proteins known from mesophilic organisms, which is of the same order of magnitude as hydrogen bonds (2-5 kcal/mol) [2].

### 1.1.2 Adaptation to high temperatures

Proteins from thermophilic and hyperthermophilic organisms (hyper- and thermophilic proteins) have to maintain their native and active form at these extreme temperatures, and do so by lowering the native conformational free energy (thermodynamic), by increasing the free energy barrier between the denatured and native states (kinetic) (Fig. 2) or a combination of both [28,29].

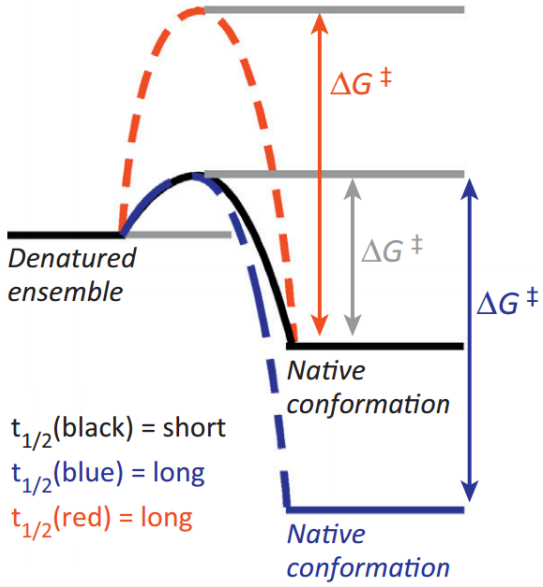


*Figure 2. A schematic representation of thermodynamic (blue) and kinetic (red) stabilization of the native conformation, compared with a marginally stable protein (black) [28].*

Thermodynamic stabilization of proteins is defined by the Gibbs free-energy change upon unfolding ($\Delta G_U$) (Fig. 3) quantified as (Eq. II),

$$\Delta G_U = -RTln(K_U) \qquad \text{(II)}$$

,where $K_U$ is the equilibrium constant between the native and denatured state. When viewed as a simple two-step transition the connection is defined by the rate constants of folding ($k_f$) and unfolding ($k_u$) involved in kinetic stabilization (Eq. III and IV) (where F is the native state and U is the denatured one) [29].

$$K_U = k_f / k_u \qquad \text{(III)}$$

$$F \overset{K_U}{\leftrightarrow} U \qquad \text{(IV)}$$

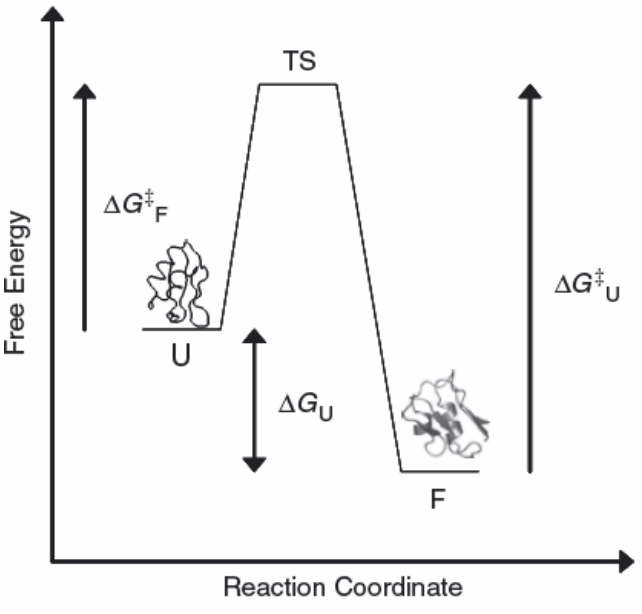


*Figure 3. A scheme linking the free energy of unfolding to the activation energy of folding ($\Delta G^{\ddagger}_F$) and activation energy of unfolding ($\Delta G^{\ddagger}_U$). Where: $\Delta G^{\ddagger}_U - \Delta G^{\ddagger}_F = \Delta G_U$ [29].*

Three models have been proposed for thermodynamic stabilization of thermophilic proteins (Fig. 4). The first model states that a thermostable protein would have a $\Delta G_U$ profile that is shifted upwards compared to a mesophilic homologue i.e. higher $\Delta G_U$ at every temperature and is therefore more stable. The second model proposes that the $\Delta G_U$ profile is displaced horizontally towards higher temperatures, leading to higher stability at higher temperatures, but lower stability at lower temperatures compared to a mesophilic counterpart. The third mechanism implicates a broadening of the $\Delta G_U$ profile, leading to more stability at lower and higher temperatures, but with the $\Delta G_U$ maxima at the same

temperature as the mesophilic proteins of the same family [29]. As a result the protein´s $\Delta G_U$ has a more shallow temperature dependence reflecting a lower specific heat capacity change ($\Delta C_p$) (Eq. V) [29,30].

$$\Delta G_U = \Delta H_U \left(1 - \frac{T}{T_m}\right) + \Delta C_p (T - T_m) - T\Delta C_p \ln\left(\frac{T}{T_m}\right) \qquad \text{(V)}$$

Using the Gibbs-Helmholtz equation (Eq. V) the temperature dependence of the conformational stability can be evaluated and described in terms of stability curves. In equation V $\Delta H_U$ is the enthalpy change at $T_m$ (melting point of the native conformation) and $\Delta C_p$ is the difference in the specific heat capacity at constant pressure between the native and the unfolded state of the protein [30].

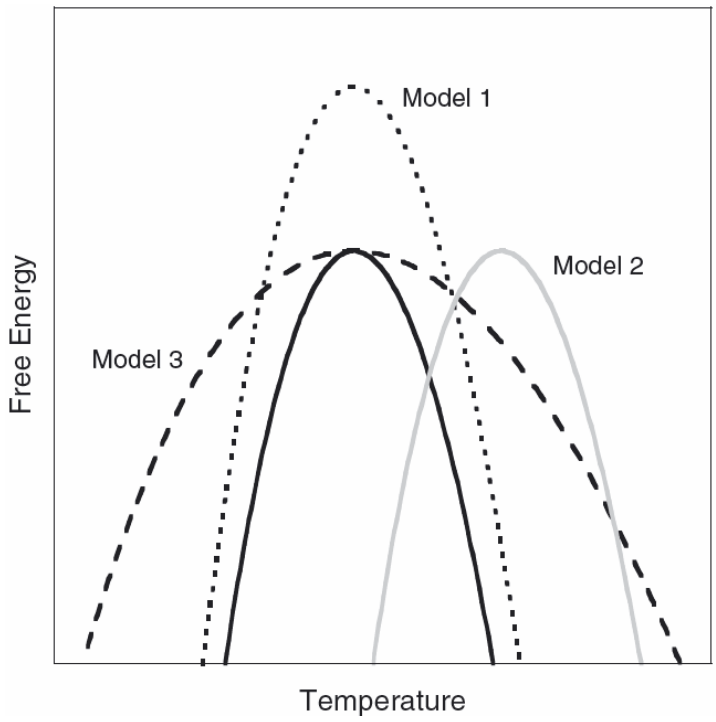


*Figure 4. Hypothetical stability curves for the three models proposed for thermodynamic stabilization, compared with a stability curve for a mesophilic protein (black solid line) [29].*

Kinetic stabilization has been shown to be an important aspect of protein stabilization, in models where the Anfinsen´s thermodynamic two state model does not apply (Eq. IV)

[31]. Two scenarios have been postulated for kinetic stability [7]. The first one can be described by Lumry-Eyring model (Eq. VI).[32].

$$F \overset{K_U}{\leftrightarrow} U \overset{k_2}{\rightarrow} D \qquad \text{(VI)}$$

It states that the folded (F) and a partially unfolded state (U) are in equilibrium that is governed by the equilibrium constant $K_U$ but the folded state is thermodynamically stable with respect to the partially unfolded state. The partially unfolded state can then unfold completely and irreversibly to a final unfolded state (D) determined by the rate constant $k_2$ [7]. In the other scenario the native state is not thermodynamically stable but is only stabilized by a high energy barrier between the two states and the life time is determined by the rate constant of unfolding ($k_u$) (Eq. VII) [7].

$$F \overset{k_u}{\rightarrow} U \qquad \text{(VII)}$$

To gain insight into which structural aspects contribute to this increased thermodynamic stability we need to look at comparative research of homologous proteins from psychrotrophic, mesophilic and thermophilic organisms. This approach has revealed some possible trends. In one such study [33] a computational analysis was carried out on structures from 25 protein families consisting of 64 mesophilic and 29 thermophilic proteins. While these proteins are structural homologues they often share only about 20-30% amino acid sequence identity with their counterparts [34]. Among the different characteristics observed in the sequence of thermophilic proteins compared to their mesophilic counterparts were that they contained fewer serine, methionine and asparagine residues, they had higher proline content at the expense of glycines in loops and a higher number of arginine residues at the cost of lysine residues. A net increase of all charged residues in hyperthermostable proteins was also observed [33]. For both hyperthermostable and thermostable proteins the strongest correlation seems to be with the number of ion pairs (salt bridges) and evidence for the relevance of these interactions for thermostabilizations are piling up [12,26,35,36]. Some research on salt bridges has indicated that they may have destabilizing, no, or little effects on protein stability [37-40]. These observations might be explained by pointing out that these experiments were mostly carried out at room temperatures, meaning that the association of the charges has a high desolvation penalty but at higher temperatures the desolvation penalty is much lower [41]

which may explain the higher abundance of salt bridges and even networks of salt bridges found in thermostable proteins [33]. Another correlation with thermostabilization was observed for the lower number of cavities in the protein structure, although the correlation was mostly found in hyperthermophilic proteins [33]. Secondary structure comparisons have shown a negative correlation of irregular/flexible areas (loops) with thermostabilization. Which lead to a higher β-sheet content in hyperthermophilic proteins and a higher α-helical content in thermophilic proteins, possibly affecting flexibility of the native structure [33]. A correlation has also been found with a higher number of polar residues exposed to the solvent, but the correlation was strongest with thermophilic proteins. Hydrophobic interactions probably contribute somewhat to the stability of thermophilic proteins as calculations suggest they should be maximized at 75°C. These calculations indicated that the free energy associated with these interactions should be entropic at room temperature but enthalpic at higher temperatures, with a maximum at 75°C [33,41,42]. These calculations come from a model that has many assumptions so the value for the enthalpic maxima is debatable and might even differ between proteins, the magnitude of these effects are marginal in comparison to electrostatic interactions that stabilize the native state [33,41,43]. Also according to the model of Elcock [41], breaking a salt bridge has a significant energy barrier that increases with temperature. Similar barriers are not observed for hydrophobic interactions indicating the importance of electrostatic interactions in thermophilic proteins [33]. Interestingly there did not seem to be a correlation between thermostability and the number of hydrogen bonds and the role of hydrogen bonds in thermostabilization of thermophilic proteins has been controversial [33,44,45]. In some cases an increase in the total number of hydrogen bonds in thermostable proteins has been reported, however [45]. The number of unsatisfied hydrogen bonding acceptors and donors seems to be slightly lower in thermophiles than in their mesophilic counterparts [33]. All in all, all these factors lead to a higher number of intramolecular interactions that lead to a higher degree of rigidity of the native state compared to mesophilic counterparts if measured at the same temperatures, but it has been proposed that their flexibility is almost identical at their respective optimum temperatures [46], as a certain degree of flexibility is often required to carry out their biological function.

Now having covered some aspects of thermodynamic stabilization, the next question is how do protein acquire kinetic stability? Probably the best known proteins that are kinetically stabilized are some proteases such as the α-lytic protease, were the native state has a $t_{1/2\ (unfolding)}$ of 1.2 years and $t_{1/2\ (folding)}$ 1800-2000 years corresponding to a 26-30 kcal/mol free energy barrier with an unfolded state that is more stable than the native state (Fig. 5) [6,47] and subtilisin [48].

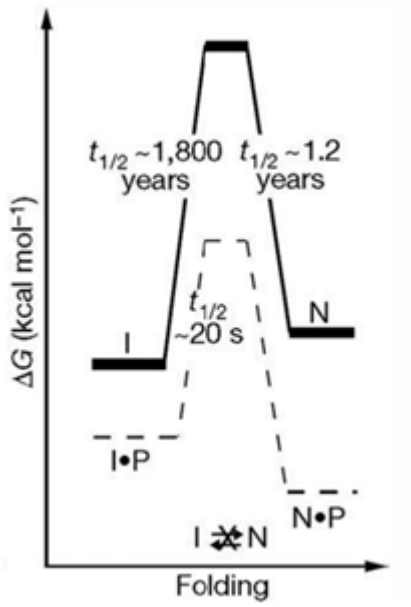


*Figure 5. The α-lytic protease gains its kinetic stabilization through an intramolecular chaperone mediated folding process, then the pro-domain is cleaved off it traps the native state in a kinetic trap [7].*

The way by which these proteins gain this stability may either be by post translational modification, or as is the case for some highly kinetically stabilized proteases, by cleaving of a pro-sequence (i.e. intramolecular chaperone) which assists in the folding of the protein, but which then cleaved off. The conformation of the protease is trapped in a local energy minimum that has high enough barriers for the protein to stay folded long enough to complete its task [7,48]. In relation to evolution, kinetically stable proteins might be the answer for proteins that have to manage at extremely harsh conditions such as extracellular proteases, or as some have speculated that the rate constants of unfolding might have a role as a biological clock in systems where a certain response is needed for a certain amount of time [7].

### 1.1.3 Adaptation to low temperatures

The conditions that psychrophilic proteins have to adjust to give rise to a whole new set of problems, such as overcoming the reduction in chemical reaction rates, reduction in membrane fluidity, increase in viscosity of biological fluids and other effects connected to lower temperatures [49]. For proteins that are adapted to high temperatures the main evolutionary pressure is on maintaining the native fold. While for cold adapted proteins they have to maintain biological function at these low temperatures (down to -20°C in some cases), therefore evolutionary pressure seems to be on maintaining molecular flexibility of the native structure [30,50]. As the largest part of the biomass on Earth is generated at cold temperatures (under 5°C) mostly from microorganisms in the oceans, it is safe to say that in terms of evolution adapting life to low temperature has been a success [30]. But how do they cope?

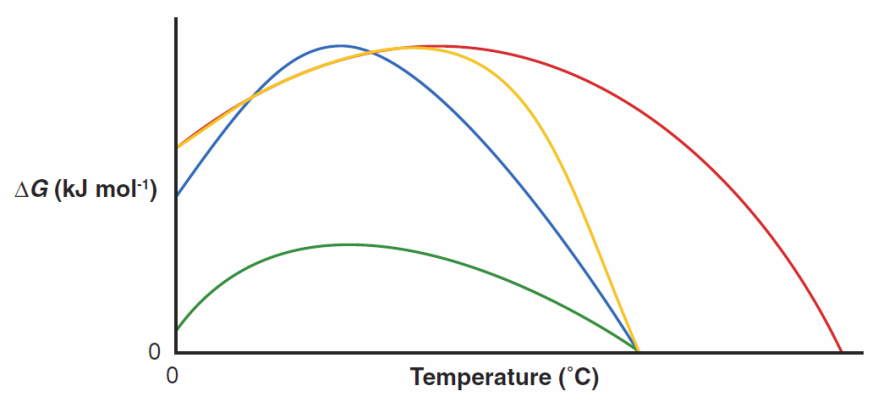


*Figure 6. Hypothetical stability curves for three different strategies to adapt to lower temperatures (green, Blue and yellow) compared to the stability curve of a thermostable homologue [30].*

The postulated thermodynamic ways to cope at these temperatures might be to shift the total stability curve to lower temperatures (Fig. 6; blue curve), or to lower the $T_m$ while maintaining the same $T_{max}$ and $\Delta G$ (Fig. 6; yellow curve) and the final model demonstrates overall lowering of the curve leading to lower $T_m$, $T_{max}$ and $\Delta G$ (Fig. 6; green curve) [30]. At subzero temperatures and temperatures under the $T_{max}$, cold-denaturation is a possible factor affecting the stability of these proteins [30,49], cold denaturation has not been extensively studied, but values for cold denaturation such as -5°C for the α-amylase form *Pseudoalteromonas haloplanktis* have been reported [30,51]. Reported data for the high

levels of ubiquitin-conjugated proteins in Antarctic fish, have also provided some evidence for the prevalence of cold denaturation of proteins *in vivo* [49], implying that cold-denaturation is a factor that needs to be considered in the evolution of cold adapted proteins.

When comparing the structures and amino acid composition of cold-adapted enzymes to their counterparts from warmer environments, several trends have been observed. Surface loops in cold-adapted enzymes contain fewer prolines that do have restrained conformations, imposing rigidity on the structure. They are also often longer and sometimes contain more charged residues that may induce flexibility through repulsive forces. This also causes that a larger proportion of the native structure is in the form of loops or is disordered [30,52-54]. An overall increase in surface charge, particularly negative charge, has also been reported for cold adapted proteins [30,55-57]. The increase in surface charge may reflect structural adaptation to the change of the dielectric constant of water which occurs with temperature, e.g. increases from 55.5 Debye at 100°C to 88 at 0°C [30,58]. Furthermore, at temperatures as low as 0°C the high viscosity and surface tension make the energetic cost of disrupting H-bond networks of water very costly [58]. Such energetic cost may therefore be offset by having charged or polar amino acids interacting with the water phase [55,59], resulting in better solvation of the protein and maintaining flexibility [30,58]. This might also reflect on why some cold-adapted proteins have lower Arg/Lys ratio than their homologs, although the opposite has been observed in some cases [30]. Also observed in some cold-adapted enzymes are rather large hydrophobic surface areas that destabilize the structure by lowering the entropy of the water molecules and decreasing hydrophobic packing, probably leading to higher flexibility [30]. The cores of cold adapted proteins also often contain fewer or smaller hydrophobic residues, affecting the hydrophobic interaction as well as van der Waals interactions that are highly distance sensitive [30]. Some psychrophilic enzymes lack aromatic-aromatic and aromatic-amino interactions compared to their thermophilic counterparts [30,55]. Others lack disulfide bridges [60], although that is not always the case [57]. Higher methionine content has been observed for several cold-adapted enzymes. These residues may confer flexibility upon the structure due to their high degree of freedom and lack of dipole interactions [61]. Lower binding constants for metal ligand ions such as $Ca^{2+}$ that often stabilize the structure have been observed [30]. Salt bridges in

psychrophilic proteins are less common, but even though the desolvation penalty is higher at low temperatures they are not unheard of [30,58,62]. Thus in general the cold-adapted proteins seem to be destabilized to gain higher flexibility. This may be reflected in a folding funnel that is both shallower and has more possible conformations for the native state than for thermophilic proteins (Fig. 7) [51].

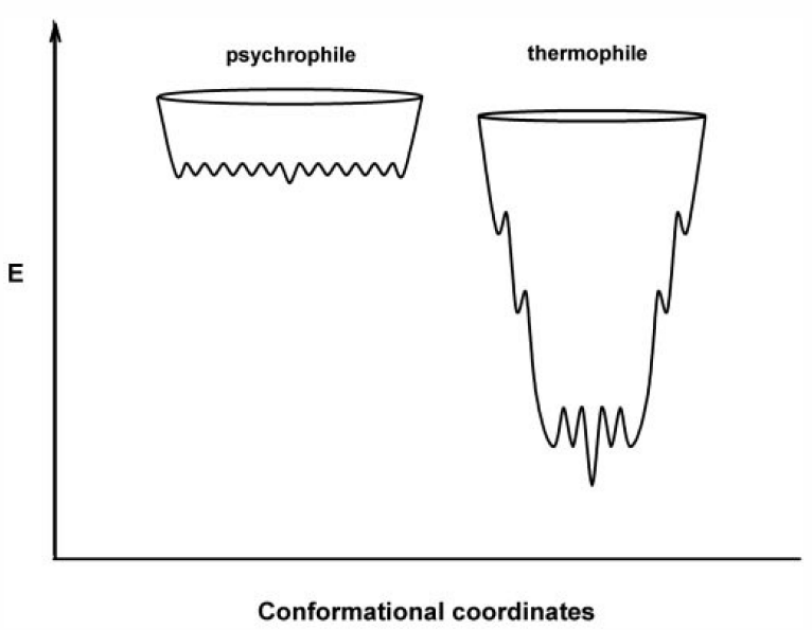


*Figure 7. Hypothetical models for the folding tunnels of psychrophilic and thermophilic proteins. Here E (the free energy) is shown as a function of conformational coordinates, where the width of the funnel represents conformational diversity [51].*

For enzymes found in psychrophilic organisms it is particularly important to counter the slower reaction rates as well as maintaining their native state and probably the most selective pressure on enzymes is exerted toward maintaining high reaction rates, even at low temperatures (Fig. 8) [50]. Enzymatic reactions that follow simple Michaelis-Menten mechanisms have their rates given as the first order rate constant $k_{cat}$, a parameter for the conversion of the enzyme-substrate complex to enzyme and product. This reaction is temperature dependent as described by the Arrhenius equation (modified according to the collision theory of reaction kinetics) (Eq. VIII) [30,63]:

$$k_{cat} = Z_p e^{-E_a/RT} \qquad \text{(VIII)}$$

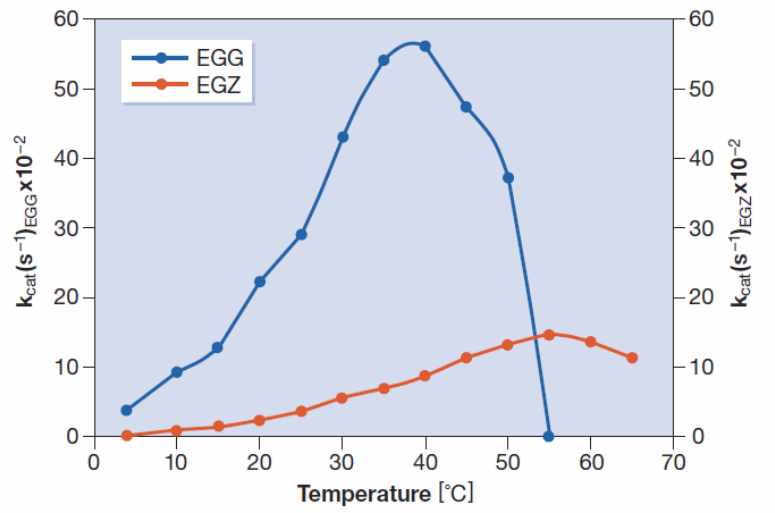


*Figure 8. Comparison of $k_{cat}$ over a certain temperature range for the cold-adapted cellulase (EGG) and its mesophilic homologue (EGZ) [50].*

where T is the absolute temperature in Kelvin degrees, R is the universal gas constant, $E_a$ is the activation energy, Z is the collision frequency and p is a steric factor. According to this at very low temperatures, there is insufficient kinetic energy in the system to overcome the reaction barrier [30]. To consider in what way enzymes can overcome that problem we need to look at a modified version of equation VIII (Eq. IX).

$$k_{cat} = \left(\frac{k_B T}{h}\right) e^{-\left(\left(\frac{\Delta H^{\#}}{RT}\right)+\left(\frac{\Delta S^{\#}}{R}\right)\right)} \quad \text{(IX)}$$

Here $k_B$ is the Boltzmann constant, h is the Planck constant, $\Delta H^{\#}$ is the activation enthalpy of the reaction and $\Delta S^{\#}$ is the activation entropy. From this equation it is apparent that $k_{cat}$ can either be increased by increasing $\Delta S^{\#}$ or by decreasing $\Delta H^{\#}$. In cold-adapted enzymes the mechanism chosen seems to be to decrease $\Delta H^{\#}$. This decrease is accomplished by increasing the flexibility of the active site and therefore reducing the number of enthalpy related interactions that need to be broken in the reaction [30]. But as every known mesophilic enzyme has a higher activation entropy [63-66] than their cold-adapted counterpart, this leads to the assumption that there is a trade off in a way that when $\Delta H^{\#}$ is decreased, $\Delta S^{\#}$ is also decreased, resulting in up to a tenfold increase in $k_{cat}$ compared with a thermophilic homologue [30]. If this trade off would not occur, a clear decrease in $\Delta H^{\#}$ by 20 kJ/mol would increase $k_{cat}$ by 50,000-fold [30,63]. Another important factor for

enzyme catalysis is the parameter $K_m$ which describes the stability of the enzyme-substrate (ES) complex reflecting the affinity of the enzyme to the substrate (Eq. X).

$$\Delta G_{ES} = -RT \ln(\frac{1}{K_m}) \quad \text{(X)}$$

As a result enzymes with low $K_m$ values have a more negative $\Delta G_{ES}$ [67] i.e. the energy valley that has to be passed before reaching transition state is deeper. As mentioned before cold-adapted enzymes often achieve higher $k_{cat}$ by decreasing $\Delta H^{\#}$ (i.e. lowering $\Delta G^{\#}$) which can be achieved by stabilizing the activated substrate transition state or by destabilizing the ES-complex [30,42]. As a result of this several cold-adapted enzymes have both higher $k_{cat}$ and $K_m$ than their thermophilic homologues, but there are however well known cold-adapted enzymes that have lower $K_m$ values [30].

### 1.1.4 Comments on activity and stability relationships

As discussed above many different ways exist for proteins to gain stability or to decrease their stability to gain higher activity as is the case for many enzymes. Also there seems to be a clear trade off in terms of activity and stability (Fig. 9) [34]. But can we use the lessons from nature to engineer enzymes that are both more active and more stable? The short answer to this question is yes. In the past years different groups have managed to develop mutants that tend to be more active and stable using either rational design or directed evolution [1,12,34,68]. There are still however many unanswered questions, for instance; how far can we manipulate a protein before it becomes completely inactive and there is a need to define better the different ways in which the different families of proteins can achieve the structural stability and activity needed to carry out their biological functions.

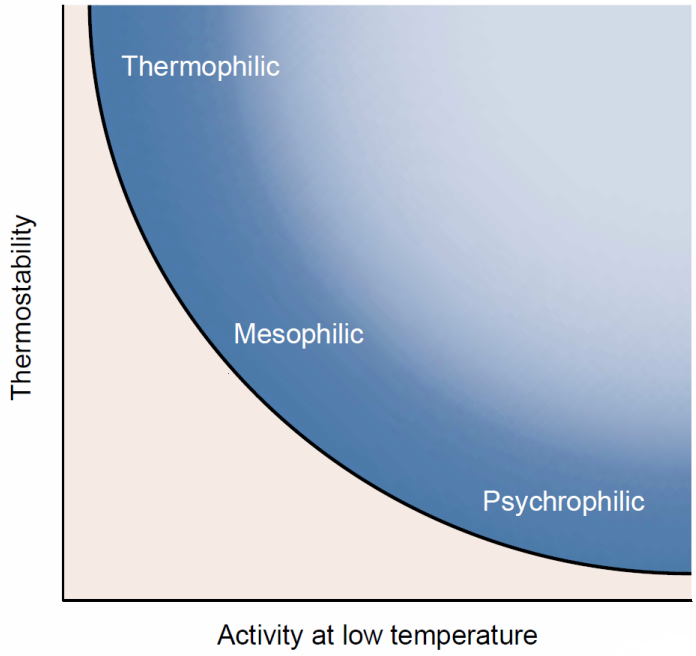


*Figure 9. When looking at homologous proteins in Nature that are adapted to different temperatures they show a trade-off between activity and stability. Naturally occurring proteins lie in the dark shaded blue area. The pink area is un-accessible as it defines the minimum stability needed for biological function. The light blue area represents enzymes that are both thermostable and highly active at low temperatures but they are usually not found in Nature. However with rational design and directed evolution this light blue area seems to be accessible but has no biological relevance [34].*

## 1.2 Proteases

Proteases (peptidases) are enzymes that facilitate the cleavage of peptide bonds of proteins into peptides and/or amino acids via hydrolysis and include both endo- and exopeptidases. According to the MEROPS peptidase database (merops.sanger.ac.uk) proteases can be divided into aspartic, cysteine, glutamic, metallo, asparagine, mixed, serine, threonine and unknown proteases. Each type can be divided into clans that contain all families, a family is a set of homologous proteolytic enzymes that have the same evolutionary origin. Typical genomes of mammalian cells have more than 2% of their genes encoding for proteolytic enzymes or their inhibitors [69]. In the next sections the focus will be on serine proteases.

### 1.2.1 Serine proteases

Over one third of known proteases are serine proteases and as a result they are the most abundant and functionally diverse of all the proteases and are found in all kingdoms of life as well as in some viral genomes [70]. Serine proteases are divided into 13 clans and 40 families [70], the name of the this group stems from the serine residue in the active site that acts as a nucleophile which attacks the carbonyl moiety of the peptide chain substrate forming the acyl-enzyme intermediate [70-72].

*Table 1. Classification of serine proteases into clans with number of known families in each clan, their representative member, catalytic residues and their primary specificity [70].*

| CLAN | NR. OF FAMILIES | REPRESENTATIVE MEMBER | CATALYTIC RESIDUES | PRIMARY SPECIFICITY |
|---|---|---|---|---|
| **PA** | 12 | Trypsin | His-Asp-Ser | A, E, F, G, K, Q, R, W, Y |
| **SB** | 2 | Subtilisin | Asp-His-Ser | F, W, Y |
| **SC** | 2 | Prolyl oligopeptidase | Ser-Asp-His | G, P |
| **SE** | 6 | D-A, D-A carboxypeptidase | Ser-Lys | D-A |
| **SF** | 3 | LexA peptidase | Ser-Lys/His | A |
| **SH** | 2 | Cytomegalovirus assembling | His-Ser-His | A |
| **SJ** | 1 | Lon peptidase | Ser-Lys | K, L, M, R, S |
| **SK** | 2 | Clp peptidase | Ser-His-Asp | A |
| **SP** | 3 | Nucleoporin | His-Ser | F |
| **SQ** | 1 | Aminopeptidase DmpA | Ser | A, G, K, R |
| **SR** | 1 | Lactoferrin | Lys-Ser | K, R |
| **SS** | 1 | L,D-carboxypeptidase | Ser-Glu-His | K |
| **ST** | 5 | Rhomboid | His-Ser | D, E |

Serine proteases were among the first enzymes that where studied extensively [73,74]. Sparking the interest in these proteases was their intensive involvement in physiological processes. The reaction mechanism of serine proteases is well studied and serine proteases perform their task very efficiently, with rates of catalyzed hydrolysis of a peptide bond almost $10^{10}$ –fold greater than that of the uncatalyzed reaction [72]. Three main obstacles stand in the way of the reaction, one is that amide bonds are extremely stable due to the electron donation from the amide nitrogen to the carbonyl (compared to alkyl esters that are 3000 times more reactive, and p-nitrophenyl esters that are 300000x more reactive). Proteases solve that via interactions of the carbonyl oxygen with a general acid that distorts

the resonance stabilization [72]. A second obstacle is that water is a poor nucleophile, which is solved via activation of water with a general base [72]. A third obstacle is the fact that that amines are poor leaving groups and this is solved in the mechanism of the serine proteases by protonation of the amine prior to the release [72].

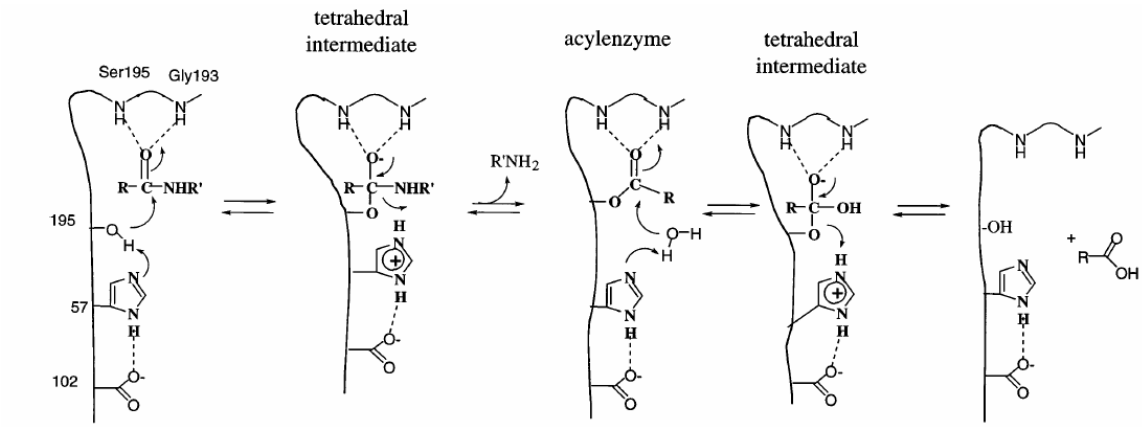


*Figure 10. The generally accepted reaction mechanism for the catalysis of serine-proteases (numbering of residues refers to the numbering in chymotrypsin) [72].*

In the mechanism of serine proteases the serine residue of the catalytic triad attacks the carbonyl of the peptide substrate. This is achieved by the activation of the serine by the side chain of a histidine that acts as a general base and pulls the proton from the serine residue making it a better nucleophile. The His-$H^+$ is stabilized by hydrogen bonds to the aspartate of the catalytic triad. The tetrahedral intermediate formed in reaction has an oxyanion which is stabilized by the NH-groups of the main-chain in the so called oxyanion hole. The next step is the collapse of the tetrahedral intermediate with the release of the peptide leaving group, assisted by the His-$H^+$ acting as a general acid. This then yields the acyl-enzyme. Deacylation is achieved by a water molecule that is assisted by His of the active site and attacks the acyl-enzyme forming a second tetrahedral intermediate. This intermediate is cleaved releasing a carboxylic acid product and freeing the active site serine [72].

As shown in table 1, serine proteases differ greatly in their specificity. This specificity is gained through different binding pockets around the active site that recognize different side

chains of the substrate [72]. PN and SN are on the acyl-enzyme side and PN´ and SN´ are on the leaving group side.

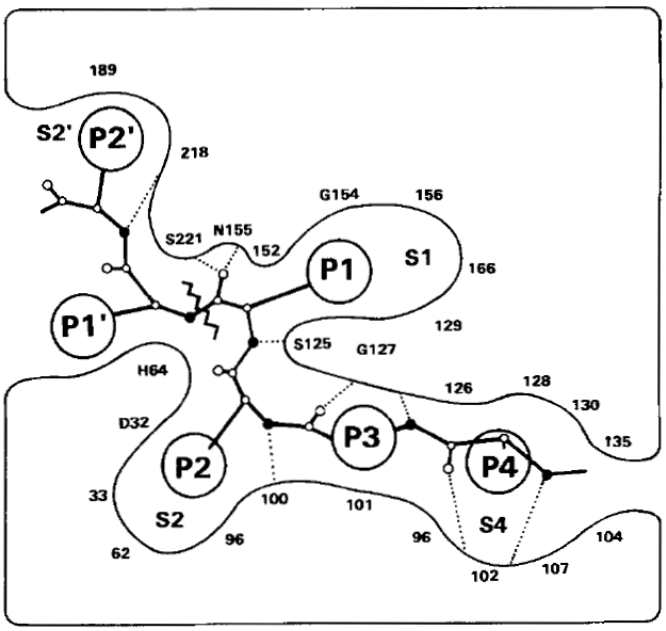


*Figure 11. A schematic representation of the active site of a serine protease describing the interactions between certain side chains in the substrate (PN and PN´) and their corresponding binding areas (SN and SN´). Residue numbering corresponds to that of substilisin BPN´. The cleavage point is represented by jagged lines. The catalytic triad is D32, H64 and S221, and the oxyanion hole residue N155 are indicated [75].*

Even though serine proteases are known in all kingdoms, their distribution is highly diverse. In eukaryotes the PA clan is highly represented and they mediate blood coagulation, immune response, digestion, fibrinolysis, development, fertilization and apoptosis [70,71]. Their specificities vary, greatly, trypsin-like proteases cleave at the C-terminal side of the positively charged residues Arg and/or Lys, chymotrypsin-like cleave at the C-terminal of the aromatic (large hydrophobic) Phe, Trp and Tyr and elastase-like proteases cleave at the C-terminal of small hydrophobic residues such as Ala, Gly and Val [71,72].

The PA clan contains the His, Asp, Ser (in that order) catalytic triad that is oriented in a specific manner, another clan the SB clan has the Asp, His, Ser (in that order) but remarkably the orientation of the catalytic triad site is the same, this is not due to some distant relation but rather an example of a convergent evolution as the structure itself bears no other similarities [71]. The SB clan has only two families. The S8 family contains the

subtilisins and the S53 family contains the sedolisins, these families are vital for protein processing in all metazoa, single cell organisms and some viruses [75,76]. The function of these peptidases tends to be nutrition oriented and they prefer hydrolyzing at the C-end of large hydrophobic residues (aromatic) [71]. Most of these enzymes are secreted out of the cell, with the exception of the tripeptidyl-peptidase that mediates the intracellular protein turnover [71,77]. These proteases have relatively few representatives in the animal kingdom but are prevalent in plants, bacterial and archaeal genomes [71]. The human genome contains at least ten proteases from the SB clan, nine subtilisins and one seldolisin (tripeptidyl-peptidase I, TPP-I) which are known to have roles in protein secretion. Furhtermore a pro-protein convertase PCKS9 (subtilisin-like kexin type 9) has been demonstrated to participate in regulation of low-density lipoproteins (LDL) receptors in the liver and in turn effect LDL levels in plasma [71,78].

### 1.2.2 Subtilisin-like serine proteases

The S8 family of the subtilisin-like serine proteases is the second largest family of serine proteases both in terms of known sequences and characterized enzymes according to the MEROPS database and are divided into two subfamilies S8A (type example subtilisin) and S8B (type example kexin). The family is distinguished by a high homology of their native structure that sets them apart from chymotrypsin-like serine proteases, which have a β/β structure for chymotrypsin, whereas the the subtilisin-like serine proteases are characterized by an α/β-structure [75]. The subtilisin fold is composed mainly of parallel β-sheets that are arranged in βαβ motifs, characterized by a three-layer αβα-sandwich [71,79]. This “sandwich” has seven β-sheets aligned in the order 2314567, with a rare left-hand crossover connecting sheets 2 and 3. The β-sheets are then surrounded by α-helices that are connected to the sheets with loops that lie on the surface of the protein, the active site is then located in a cleft that is positioned between two adjacent helices on the protein surface [79]. Almost all of the prokaryotic subtilisins are produced as preproproteins, consisting of a signal peptide needed for secretion and an intramolecular chaperone required for correct folding that is then auto-cleaved, leaving the processed native structure in a kinetic trap, therefore kinetically stabilizing them [79]. The only known exception to this is a subtilisin from *Aeromonas sorbia*, instead of an intramolecular chaperone a protein coded just downstream of that protease seems to be essential for folding [80]. Another aspect of subtilases is their calcium dependency, for example subtilisins BPN´ and E have

two binding sites [81,82]. Tk-subtililsin from the hyperthermophilic archaeon *Thermococcus kodakaraenis* has seven calcium binding sites, six of which are unique to that enzyme [83,84]. Subtilisin Tk-SP from *Thermococcus kodakaraenis* contains no calcium binding sites within its subtilase domain but has two of them in a β-jelly roll domain [85] and subtilase 3 (SBT3) from tomato that does not bind calcium ions [86]. All in all sixteen different calcium binding sites have been identified, but the true subtilisins usually contain two of them [79]. The calcium ligands are important for both stability and activity and possibly assist in folding [79]

Even though these proteases have such a high structural homology their sequence homology can be used to divide them into six groups of "sequence families" named after a well-studied enzyme of each group [75]. The subtilisin family mainly includes enzymes from *Bacilli*, and can be sub-grouped into true subtilisins, high-alkaline proteases and intracellular proteases. The thermitase family contains proteases form thermophiles and halophiles. The proteinase K family, is a large family of secreted endopeptidases found in fungi, yeast and gram-negative bacteria, some containing minor insertions or deletions and a varying number of $Ca^{2+}$ binding loops. The lantibiotic peptidase family is a small family but highly specialized, they cleave leader peptides from lantibiotics (a group of antimicrobial peptides) and are found in gram-positive bacteria. The kexin family is a large group of proprotein convertases and are widely distributed in the kingdoms of Life and also found in viruses. The sixth family is the pyrolysin family, a heterogeneous group of subtilases with low sequence homology but are characterized by large insertions and/or long C-terminal extensions [75].

### 1.2.3 The proteinase K-like serine proteases VPR and AQUI

The proteinase K family as mentioned before are mostly secreted endopeptidases found in fungi, yeast and gram-negative bacteria. The members from this group share a relatively high sequence identity, or over 37% [75]. As of now there have been four different calcium binding sites identified from members of this group [87]. The type example representative enzyme of this family is proteinase K (PRK), secreted by the mesophilic fungus *Engyodontium album* (formerly *Tritirachium album*). The mature form of the protease contains 279 amino acids (28.9 kDa), contains two disulfide bridges and two of the four calcium binding sites, the strong Ca1 and the weak Ca2 (according to numbering in VPR)

(Fig. 12) [79,88-90]. PRK is a highly active protease and has a pH optimum at pH 8.0 and is stable in a pH range of 4.0 – 12.5 [91] and has a $T_{50\%}$ value of 69°C [92].

Two other subtilisin-like serine proteases that are classified as proteinase K-like are the cold-adapted VPR and the thermophilic aqualysin I (AQUI). These proteases have been extensively studied with regard to thermostabilization of proteins, as they provide an excellent comparison model sharing around 60% sequence identity [12,14,35,57,79,92-95]. VPR is secreted by a psychrotropic *Vibrio* PA-44 species, though it retains considerable stability in the presence of calcium having a $T_{50\%}$ of 54°C [92]. AQUI is a highly thermostable alkaline subtilisin-like serine protease secreted by the gram-negative extreme thermophile *Thermus aquaticus* YT-1 and has an optimum pH and temperature in the presence of calcium of pH 10.0 and 80°C, but is active in the pH range of 6-11 [96] and has a $T_{50\%}$ of 95°C [92].

*Table 2. Amino acid composition of the mature proteases VPR and AQUI [94].*

| | **A** | **R** | **N** | **D** | **C** | **Q** | **E** | **G** | **H** | **I** | **L** | **K** | **M** | **F** | **P** | **S** | **T** | **W** | **Y** | **V** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **VPR** | 25 | 9 | 23 | 21 | 6 | 10 | 5 | 40 | 4 | 9 | 18 | 5 | 3 | 7 | 12 | 38 | 20 | 4 | 8 | 24 |
| **AQUI** | 40 | 15 | 19 | 13 | 4 | 5 | 3 | 37 | 5 | 9 | 19 | 2 | 2 | 3 | 11 | 29 | 25 | 3 | 12 | 25 |

Both proteases are secreted with a rather large pro-domain. As AQUI has to cross both the cytoplasmic and outer membrane of the Gram negative bacterium it is produced with an N-pre 14-residue signalling peptide, an N-prodomain, which acts as an intramolecular chaperone, consisting of 113 residues and a C-prosequence promoting extracellular secretion (Fig. 12). Via autolysis the mature protease consists of 281 residues (28.5 kDa) [96,97]. VPR is produced as a 530 residue peptide chain, thereof the C-prodomain is around 100 residues and the N-prodomain is 139 residues, thus yielding a 291 residue mature protease (29.7 kDa). This cleavage of VPR leads to a 2 residue shorter N-end and 15 residue longer C-terminus [94,95].

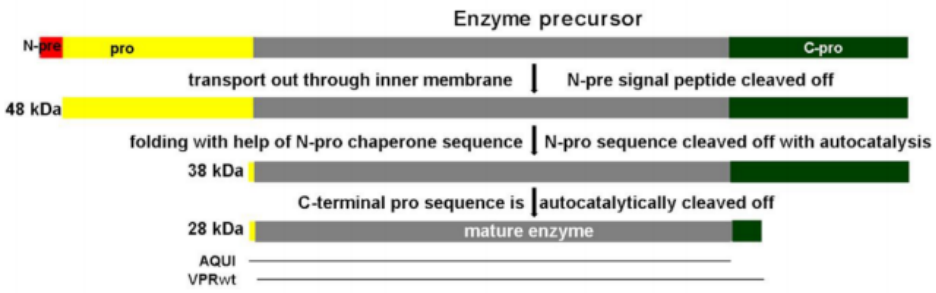


*Figure 12. Processing of the precursor proteins of AQUI and VPR. Signal peptide shown in red, N-pro sequence in yellow and the C-pro sequence in green [79].*

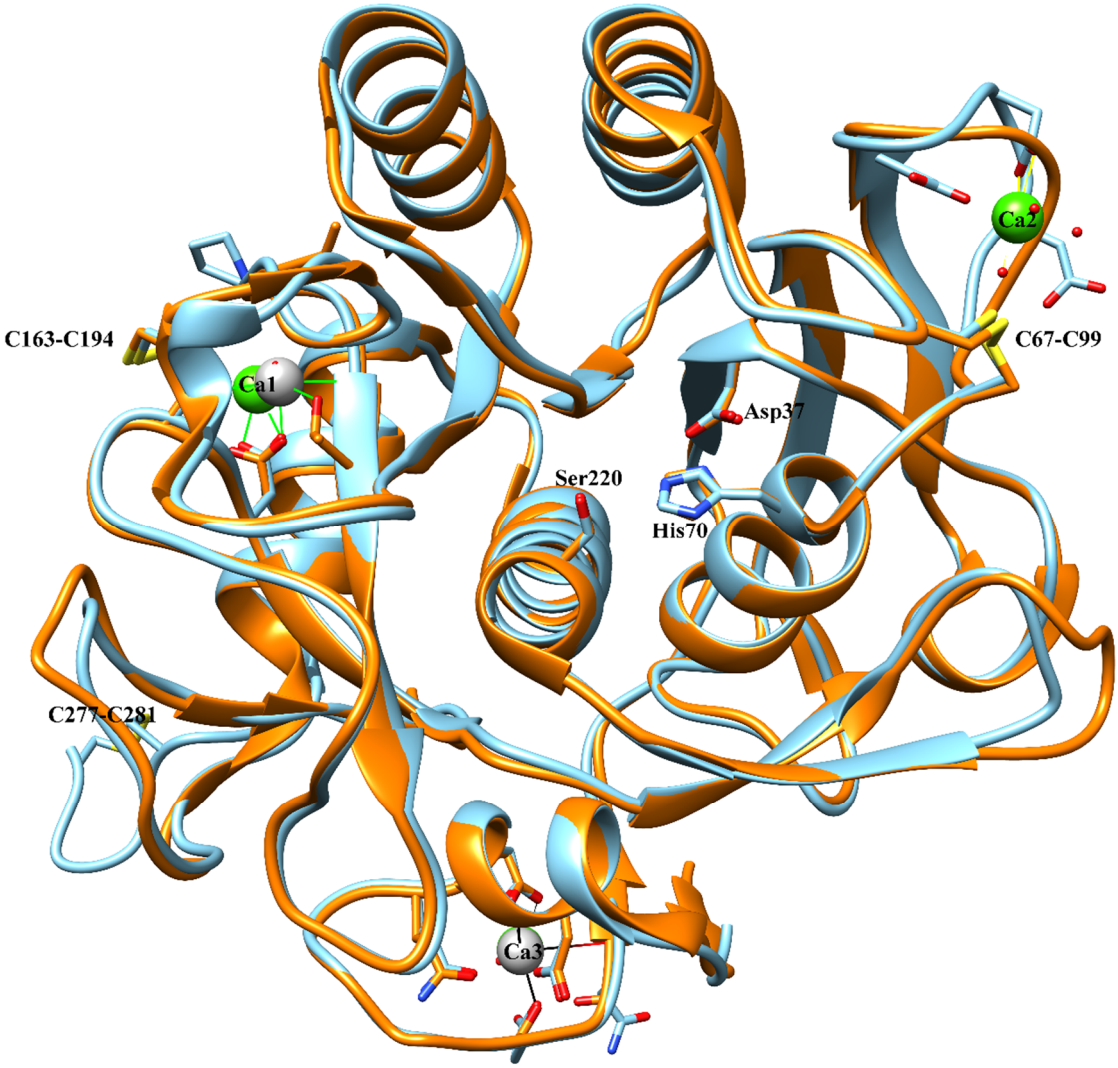


*Figure 13. Three-dimensional structure of the cold-adapted VPR, colored as light blue with green calcium ions (PDB code: 1SH7) and the thermostable AQUI, colored as light orange with grey calcium ions (PDB code: 4DZT). Numbering of residues follows the structure of VPR. The catalytic triad shown as sticks Asp37-His70-Ser220 and the disulfide bridges C64-C99, C163-C194 and C277-C281.Calsium numbering according to VPR. Atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

As mentioned before calcium binding is important for the stability of these proteases but interestingly there is not a correlation between the number of calcium sites and thermostability. Case in point there is that VPR contains three different calcium binding sites, whereas AQUI has only two (corresponding to Ca1 and Ca3) (Fig. 13), and proteinase K also has just two such sites (Ca1 and Ca2). However, in the absence of calcium the drop in $T_{50\%}$ has been measured and is around 26°C for AQUI, 12°C for PRK and 27°C for VPR (Fig. 14) [79,92].

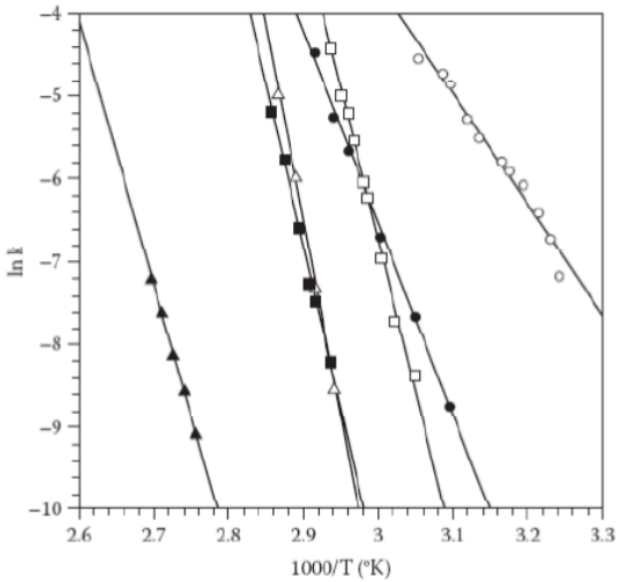


*Figure 14. Calcium dependency in regard of thermal inactivation of VPR (circle), PRK (square) and AQUI (triangle). Here blackened out symbols represents the presence of calcium and open symbols the absence of calcium [92].*

Other structural factors may contribute to the stability of the native form such as disulfide bridges. VPR and AQUI share two of them, but VPR has a third that is unique to VPR and is located on the C-terminus of the enzyme, an extension that is nonexistent in the mature enzyme structure of AQUI [57]. One of the common disulfide bridges is located near the Ca1 binding site, bridging the loop of the calcium site as well as the region harboring the residues of the S1 binding pocket (VPR: C163-C194) [57]. Another disulfide bridge lies near the Ca2 binding site in VPR (VPR: C67-C99), this bridge seems to be connected to the formation of the S4 binding pocket [57]. These bridges are somewhat conserved as they are both found in the proteinase K-like enzyme found in *Serratia* species (SPRK)

[87]. Proteinase K has two disulfide bridges but none of those correspond to those present in VPR, however [98]. Complicating matters even more, with regard to the contribution of disulfide bridges to the thermostability of these enzymes, is the fact that the thermophilic subtilase, thermitase (THM) has none [57].

Another aspect of possible stabilizing factors in the structure of these proteases are salt-bridges [99]. Even though not all of them are crucial to the stability one in particular has been shown to have drastic effects. The salt bridge Asp17-Arg259 in AQUI corresponds to an area in VPR that has none (Asn15-Lys257) and in a study where the bridge was removed in AQUI by mutagenesis it resulted in a lowering of 8-9°C in $T_{50\%}$ [35]. While in VPR when Asn15 was replaced by Asp15 the $T_{50\%}$ value improved by 3°C [12], indicating that this salt-bridge is very important for themostabilization of their native structure. Also observed in the structures and amino acid composition of VPR and AQUI is that AQUI has a higher aliphatic index and also that certain Ala residues in AQUI have been exchanged for Ser in VPR, in fact Ser-to-Ala exchanges was the most frequently observed amino acid exchange between the two enzymes [94]. As mentioned before that is in agreement to known ways of gaining higher stability through better packing or gaining higher activity through a more flexible native state.

The relationship between catalytic activity and stability/flexibility is an extremely complex one and poorly understood, with regard to time scales of certain fluctuations within the structure that are important for these factors. As of today cold-adapted proteins can be envisioned as having higher global flexibility compared to their thermophilic counterparts. Still both structures need high local flexibilities of certain areas that are important for their function, as well as there are areas that need to have as low flexibility as possible for the native state to stay folded [93]. VPR and AQUI are a good example for this theory (Fig. 15).

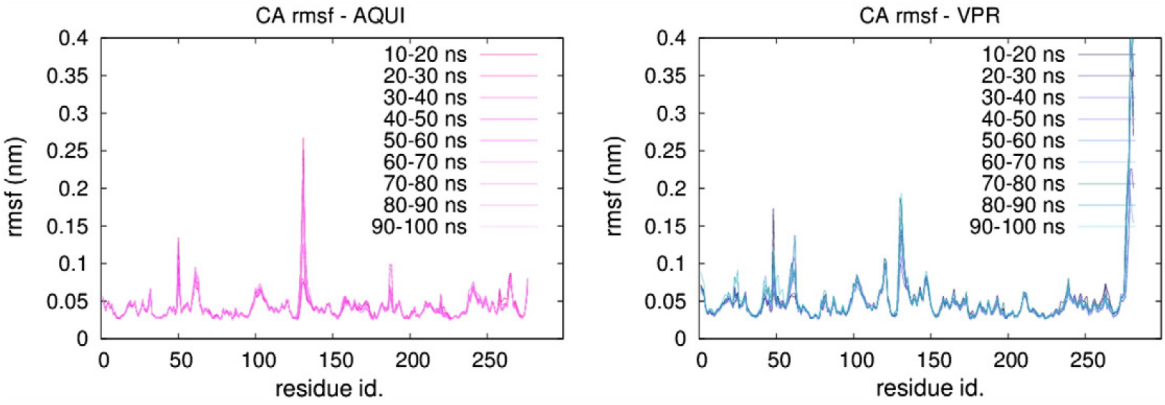


*Figure 15. Flexibility profiles of AQUI (left) and VPR (right) from molecular dynamics simulations on different time scales. The X-axis marks the residue number in the polypeptide chain and the Y-axis represents the rmsf (root mean square fluctuation) value of each residue as determined from molecular dynamics simulations [93].*

## 1.3 VPR mutations

Several mutations have been carried out on VPR in order to decipher which structural factors contribute to the stability of this protein or contribute to the temperature adaptations of these enzymes. Properties investigated involve the removal of the C-terminal of VPR to imitate the structure of AQUI in more detail ($VPR_{\Delta C}$) [12]. Ser/Ala mutations [14]. Increasing the rigidity of loops with insertions of Pro residues [14,95,100]. Insertions of multiple positively charged residues imitating a very special amino acid composition of a loop area in AQUI into VPR, which might be important for movements of the active site [101,102]. Reduction of aliphatic surface residues [102]. Mutations interfering with calcium binding sites [101,103]. Mutation affecting the hydrophobic packing of the native state [102]. Mutations have also been carried out aimed at incorporating salt-bridges into the structure of VPR [12,102,103]. Table. 3 lists some of the mutations done and the position of these mutations are shown in Fig. 16. In the next section notable mutations done on VPR will be discussed.

*Table 3. A list of mutations done on VPR by Kristjánsson M. M. and coworkers.*

| | $K_m$ [mM] | $k_{cat}$ [$s^{-1}$] | $k_{cat}/K_m$ [$s^{-1}mM^{-1}$] | $T_{50\%}$ [°C] | $T_m$ [°C] |
|---|---|---|---|---|---|
| **$VPR_{wt}$** | 0.166 | 78.4 | 472 | 56.0 | 63.6 |
| **$VPR_{\Delta C}$** | 0.184 | 68.2 | 371 | 56.4 | 65.2 |
| ***Ser/Ala exchange*** | | | | | |
| **S51A** | 0.203 | 50.3 | 249 | 57.0 | 63.9 |
| **S64A** | 0.223 | 50.7 | 228 | 56.5 | 65.1 |
| **S110A** | 0.205 | 16.2 | 79 | 57.1 | 65.2 |
| **S51A/S64A** | 0.199 | 58.8 | 295 | 54.5 | 63.1 |
| **S51A/S64A/S110A** | 0.197 | 46.1 | 234 | 55.5 | 63.3 |
| ***Proline in loops*** | | | | | |
| **N238P** | 0.191 | 21.4 | 112 | 58.0 | 64.2 |
| **T265P** | 0.191 | 10.9 | 57 | 57.6 | 65.8 |
| **N238P/T265P** | 0.228 | 25.5 | 116 | 57.0 | 65.2 |
| **I5P** | 0.261 | 58.4 | 224 | 58.7 | 67.6 |
| **N3P/I5P** | 0.224 | 8.8 | 40 | 61.9 | 69.3 |
| ***Salt-bridges*** | | | | | |
| **$S172E_{\Delta C}$** | 0.190 | 103.7 | 535 | 56.4 | 64.4 |
| **$Q142K_{\Delta C}$** | 0.157 | 147.0 | 945 | 56.3 | 64.6 |
| **$Q142K/S172E_{\Delta C}$** | 0.170 | 54.7 | 320 | 55.6 | 61.2 |
| **N15D** | 0.173 | 137.3 | 803 | 58.8 | 66.3 |
| **$N15D_{\Delta C}$** | 0.179 | 64.9 | 365 | 59.6 | 68.2 |
| **N15D/K257R** | 0.159 | 105.8 | 677 | 56.2 | 65.1 |
| ***Flexibility of a hinge area*** | | | | | |
| **A119H/S120R/G121R (3X)** | 0.154 | 22.9 | 149 | 53.5 | 61.5 |
| **(3X) + S123A (4X)** | 0.175 | 47.6 | 263 | 54.2 | 63.1 |
| **(3X)+A116T/Q117R (5X)** | 0.151 | 61.1 | 405 | 54.5 | 64.3 |
| **(5X) + S123A (6X)** | 0.148 | 98.9 | 728 | 55.1 | 63.7 |
| ***Reduced aliphatic surface*** | | | | | |
| **F29R** | 0.176 | 29.0 | 175 | 53.8 | 65.6 |
| **L243A** | 0.162 | 15.1 | 94 | 53.8 | 63.5 |
| ***Hydrophobic packing*** | | | | | |
| **A33V** | 0.199 | 14.7 | 72.3 | 55.9 | 64.2 |

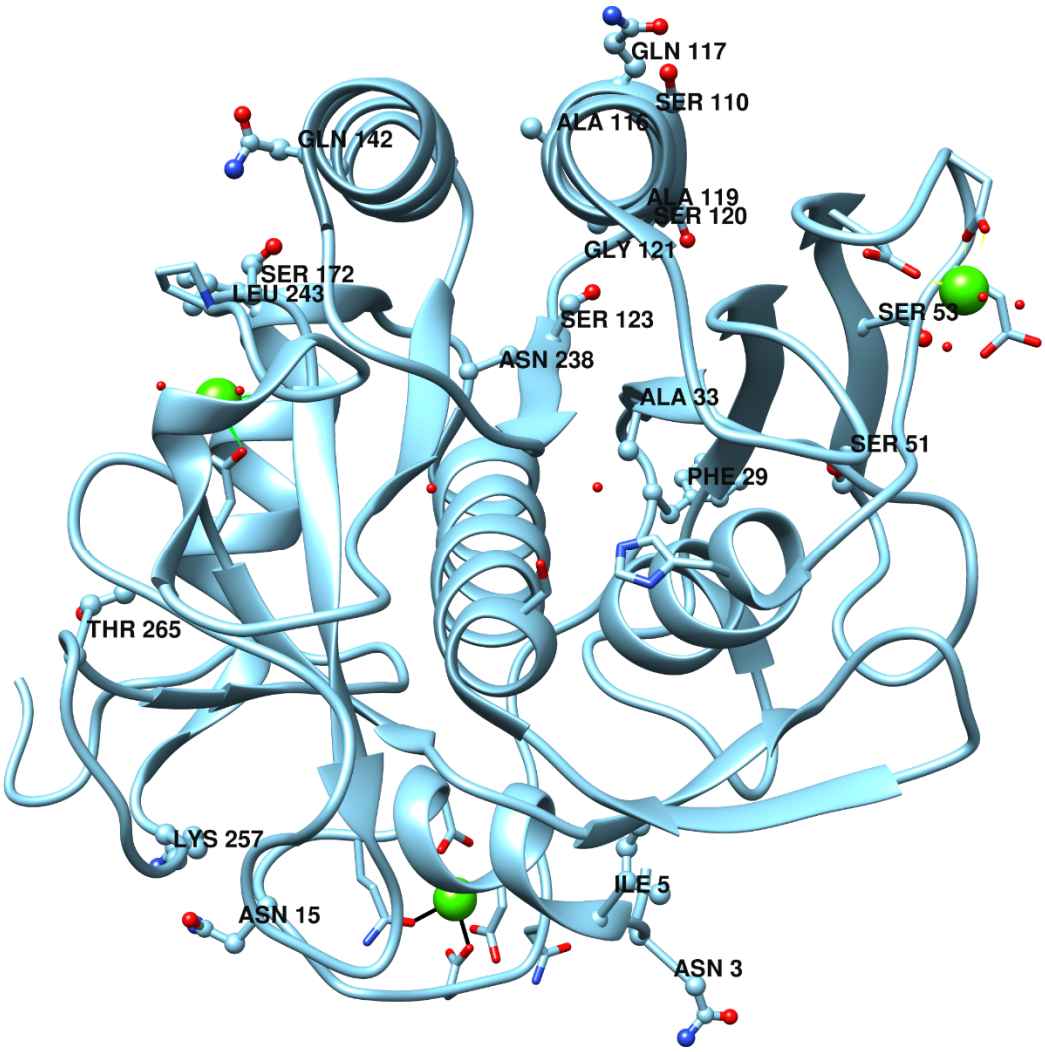


*Figure 16. The structure of VPR (PDB code: 1SH7) the native residues that have been mutated are shown as ball and sticks (see table 3 for mutations).Calcium ions are colored green and the atomic coloring code of the balls and sticks is as follows: red is O, blue is N and yellow is S.*

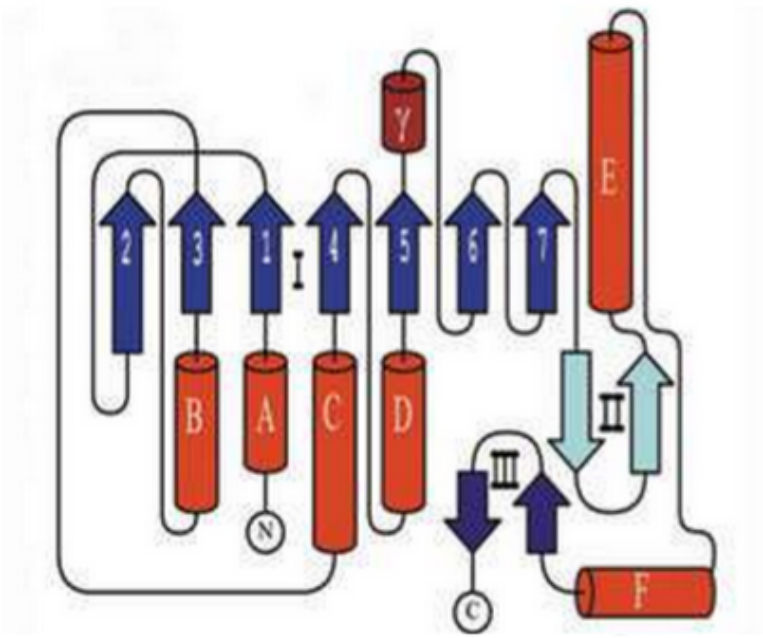


*Figure 17. A topology diagram of the structure of VPR. The red cylinders represent the α-helixes and are in alphabetical order in regard of their position in the peptide chain (with the exception of a three residue helical part in the Ca1 binding loop, named γ). The blue arrows represent the β-sheets and are numbered. The dark lines represent the loops in the structure [57].*

### 1.3.1 Mutations of VPR involved in the project

The $VPR_{\Delta C}$ mutant (C277_) (Table. 3) was made both to imitate the structure of AQUI in more detail, as this C-terminal tail found in VPR is about 15 residues longer than the C-terminus of AQUI (Fig. 18). But another aspect of the C-terminus in VPR is that it contains a disulfide bridge, has a very negative character (Glu278 and Asp280) and it has a very high beta-factor according to the crystal structure, indicative of high flexibility [12] as well as MS simulations indicating the same (Fig. 15) [93].

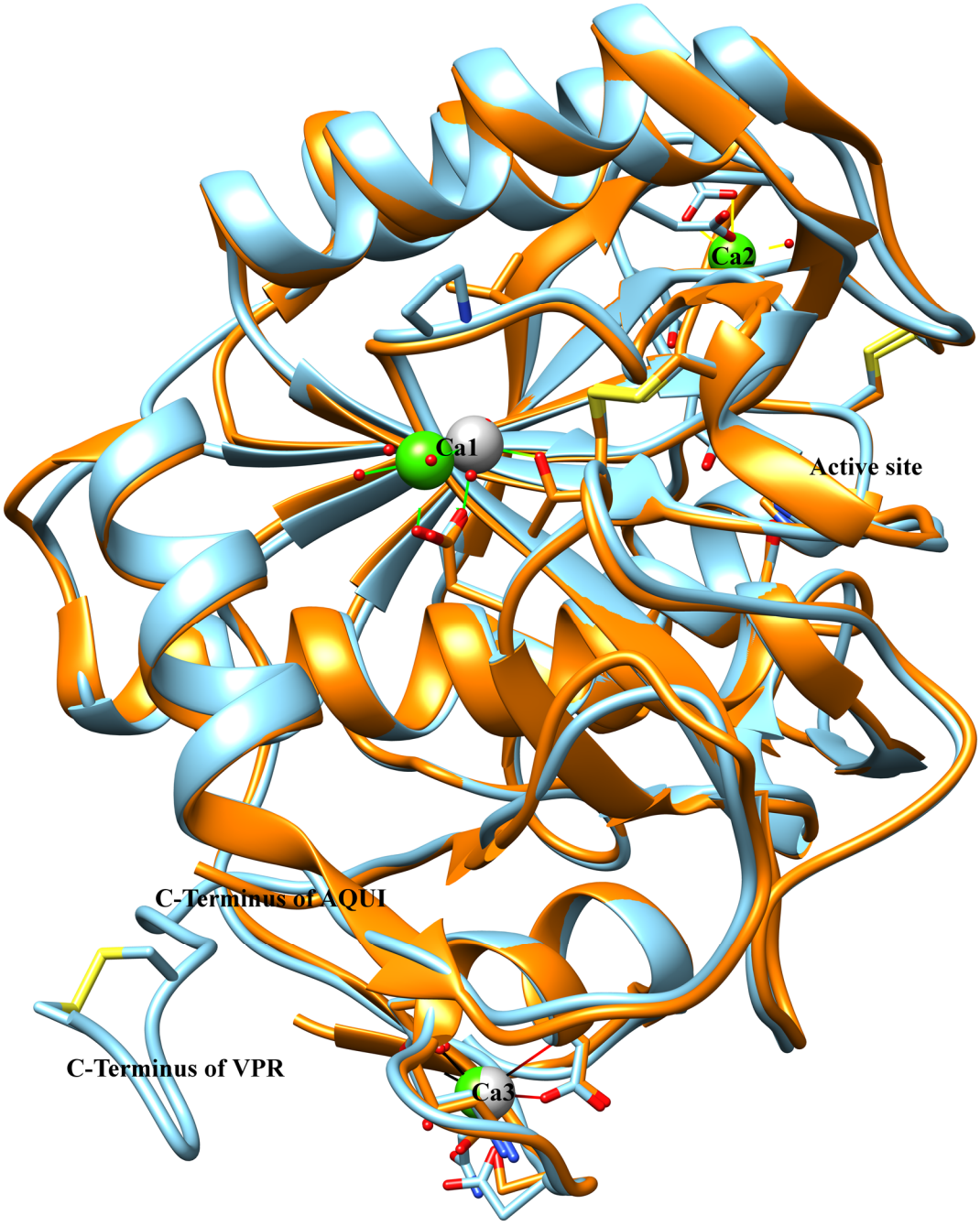


*Figure 18. Side view (90° horizontal turn comparing to Fig. 13) of VPR light blue with green calcium ions (PDB code: 1SH7) and AQUI light orange with grey calcium ions (PDB code: 4DZT). This picture shows clearly the difference between the C-ends of VPR and AQUI. Calcium numbering according to VPR and the atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

Removing the extension of the VPR C-terminus did not affect the properties drastically, but this mutant was a little more thermostable (0.6°C in $T_{50\%}$ and almost 2°C in $T_m$). A loss

of activity was also observed (lowering of $k_{cat}$ and increasing $K_m$) leading to a 1/5 loss in catalytic efficiency [12]. When introducing a negatively charged residue at position 15 (N15D) that is located on the loop of the Ca3 binding site, cooperative effects were observed. The $VPR_{\Delta C}$/N15D mutant was more stable but less active than the VPR/N15D mutant (Table. 3). The mutation N15D was meant to introduce a salt bridge between D15 and K257 and strong evidence is for the existence of this salt bridge [12]. When this mutation was carried out the wild type VPR it resulted also in a significant increase in the $k_{cat}$ of the enzyme, an effect not observed for the truncated form ($VPR_{\Delta C}$/N15D). Thus with the C-terminal tail in the proximity it may be interacting via long range ion-ion interactions or by inducing globular movements of the structure leading to increased activity and also destabilizing the new ion pair formed in VPR/N15D, but not in the $VPR_{\Delta C}$/N15D mutant, where these interactions are non-existent. This possibly explains the difference between these mutants and the effects of the C-terminus (Fig. 17 and 19) [12].

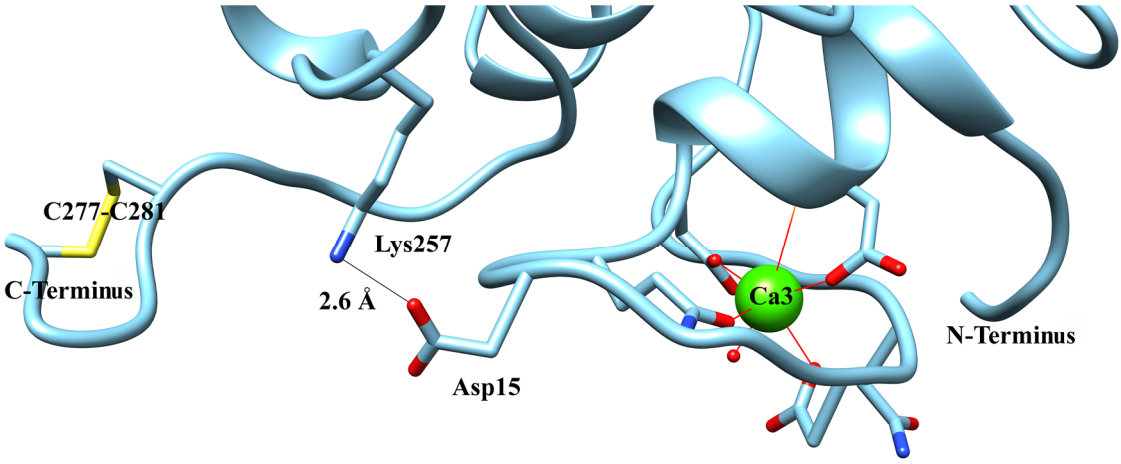


*Figure 19. View of the orientation of the N15D mutant of VPR (PDB code: 1SH7). The distance between the charged groups of Lys257 and Asp15 is 2.6Å. Calcium numbering according to VPR and are colored green. The atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

In an effort to explore other salt bridges that might have stabilizing effects on the structure another set of mutations have been performed. Imitating the structure of AQUI the mutations Q142K located on helix D (Fig. 17 and 20) and S172E on the Ca1 loop, were carried out as corresponding residues in AQUI are predicted to form a salt bridge [99].

Both of the single mutations significantly increased the activity of the enzyme, but Q142K mutant showed remarkable results, almost doubling the catalytic efficiency of the mutant (923 $s^{-1}M^{-1}$) over the wild-type enzyme (472 $s^{-1}M^{-1}$), while not affecting the stability of the protein to any extent. When combined negative cooperative effects were observed yielding a mutant which was less stable and also had lower activity [101,103]. This is hard to explain, but if there has been some interruption of the calcium binding site that might yield a less stable protein. If the putative salt-bridge between K142 and E172 reduces movements of the α-helix containing the Q142K mutation, movements in the active site important for catalysis might be affected, possibly explaining lower activity.

Q142K is an interesting mutation having these significant effects on the activity and almost none on stability. A turn away from K142 on the helix is an aspartic acid residue (D138) which forms a salt-bridge to R169, which is located on the Ca1 binding loop and according to MD simulations this salt bridge has a 100% persistence in the homologous enzyme AQUI. In AQUI, K142 also slightly interacts with D138 and has a prevalence of 5.5%. [99]. The introduction of the positively charged amino-group of K142 only one turn apart on the helix from D138 (Fig. 20), may lead to interference by the side group of K142 with the already existing salt-bridge between D138 and R269, possibly causing movements of the helix which may be transmitted to the active site regions of the enzymes, thus potentially increasing the activity of the mutant.

If the amino acid composition of the α-helix (helix-D (Fig. 17)) containing the mutation is compared to the corresponding helices in the structure of AQUI, PRK and SPRK it is revealed that a lysine in position 142 (numbering according to VPR) is not found in the other structures except in AQUI, where it is involved in a salt-bridge. The same observation can be made if VPR/Q142K is BLASTed against online databases, as a positively charged groups in position 142 are only observed if there is a negatively charged group at position 172 as is found as a part of a salt bridge in AQUI (Appendix 2). One may point out that there is an arginine in position 143 in PRK, but the orientation of that residue is towards the C α-helix and seem to be involved in a salt-bridge there with Asp112 of PRK. BLAST search for this mutation revealed no protein sequences containing a positive charge in position 142, making these beneficial effects of the mutation even more surprising.

*Table 4. Amino acid composition of the α-helix D of the homologous enzymes VPR, AQUI, PRK and SPRK. Revealing that K142 is not conserved at all, except it is found in AQUI but there it is part of a salt-bridge with E172.*

| | **Position** | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Enzyme** | 134 | 135 | 136 | 137 | 138 | 139 | 140 | 141 | 142 | 143 | 144 | 145 | 146 | 147 |
| **VPR** | S | T | A | L | D | S | A | V | Q | G | A | I | Q | S |
| **Q142K** | S | T | A | L | D | S | A | V | K | G | A | I | Q | S |
| **AQUI** | S | T | A | L | D | N | A | V | K | N | S | I | A | A |
| **PRK** | S | S | S | V | N | S | A | A | A | R | L | Q | S | S |
| **SPRK** | S | Q | A | T | D | D | A | V | N | A | A | V | A | A |

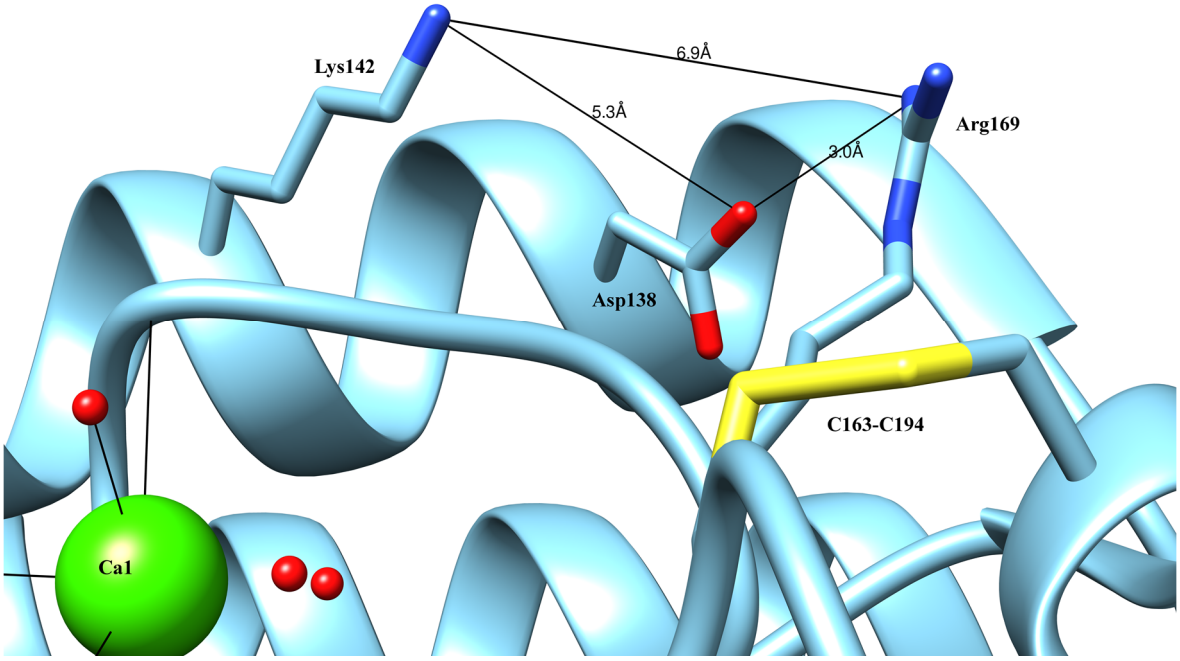


*Figure 20. Zoomed in view of helix D in the structure of VPR, showing the location of Lys142 mutation and its orientation according to calculations by the program UCSF Chimera. Also shown are residues Asp138 and Arg169 which are believed to form a salt-bridge. Calcium numbering is according to VPR and the atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

Proline insertions into loops seem to play an important role when it comes to flexibility of the structure, as all of the tested mutations have drastic effects on the catalytic efficiency, yielding much lower values as compared to the wild-type enzyme, varying from 224 $s^{-1}M^{-1}$ to 40 $s^{-1}M^{-1}$ (Table. 3). All of these mutations seem to contribute to some extent to the stability of the enzyme. The mutations that affected stability the most are those on the N-

end of the protein as two proline insertions in a N3P/I5P mutant raised the $T_m$ value to 69.3°C and $T_{50\%}$ to 61.9°C, which is a 4.1°C increase in $T_m$ and 5.5°C increase in $T_{50\%}$ compared to $VPR_{\Delta C,}$. Very detrimental effects were observed on the activity of the protein, however, as the $k_{cat}$ value dropped to 8.8 $s^{-1}$, which is almost nine fold lower than the value for the wild type enzyme [95]. The N-terminal region of VPR seems to be highly flexible compared to the rest of the structure as the β-value for the end is almost two fold higher than for the rest of the structure [57]. With this double proline insertion the autocatalytic cleavage site of the N-terminus was shifted by two residues (Fig. 12), resembling the 2 residue longer N-terminus in AQUI, PRK and SPRK (Fig. 21 and Table. 5) [95,100].

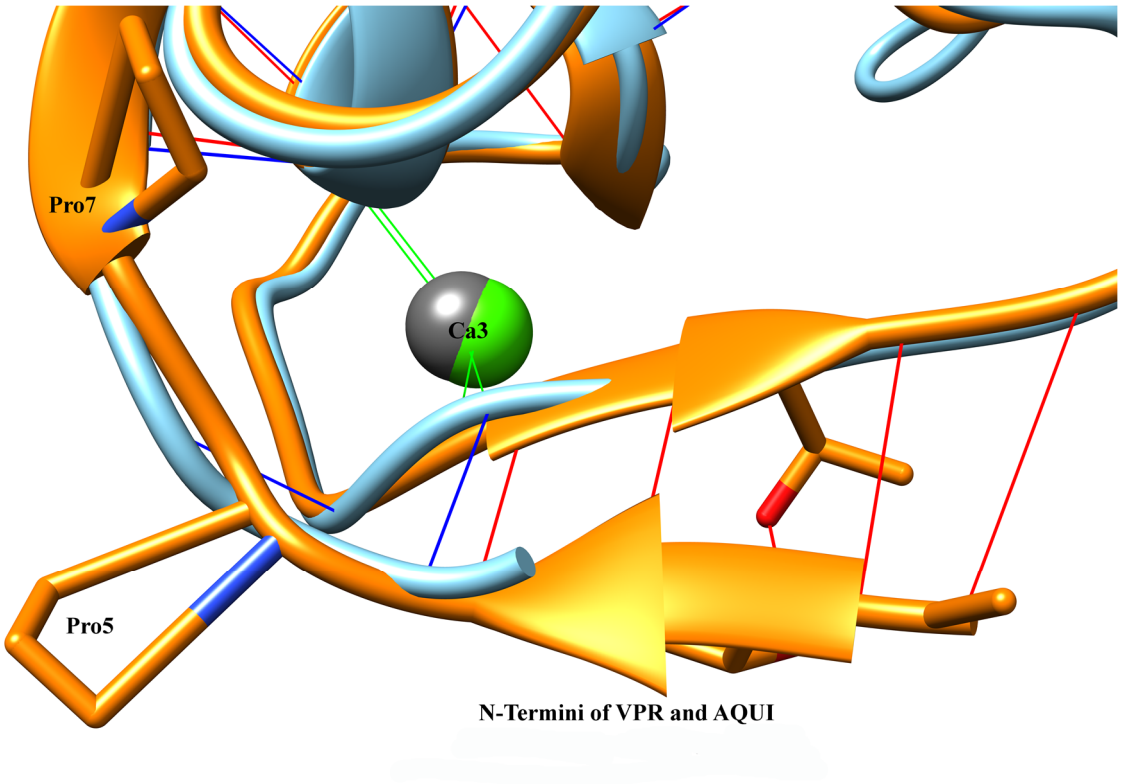


*Figure 21. Comparison of the N-termini of VPR (light blue with green calcium ions (PDB code: 1SH7) and AQUI (light orange with grey calcium ions) (PDB code: 4DZT). The N-end extension in the structure of AQUI leads to a β-sheet formation resulting in more hydrogen-bonds (red lines) than in VPR (blue lines). P5 and P7 are corresponding to positions 3 and 5 in VPR. Calcium numbering according to VPR and the atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

The fact that this two residue extension occurs in the N3P/I5P mutant of VPR is a good indicator of possible β-sheet formation at the N-end allowing for more hydrogen-bonds to be formed and therefore locking this flexible area of VPR, reducing flexibility and activity and increasing stability [95].

*Table 5. Residues at the N-terminus of VPR and its mutant N3P/I5P compared to homologous mesophilic and thermophilic enzymes. Numbering is according to VPR [95].*

| | **Position** | | | | | | |
|---|---|---|---|---|---|---|---|
| **Enzyme** | -2 | -1 | 1 | 2 | 3 | 4 | 5 |
| **VPR** | - | - | Q | S | N | A | I |
| **N3P/I5P** | A | V | Q | S | P | A | P |
| **AQUI** | A | T | Q | S | P | A | P |
| **PRK** | A | A | Q | T | N | A | P |
| **SPRK** | A | D | - | S | P | A | P |

The A116T/Q117R/A119H/S120R/G121R/S123A (6x) mutant is the product of series of mutations that were designed to imitate a very special amino acid composition of a loop region connecting α-helix C and β-sheet 4 (corresponding to the nomenclature in VPR) (Fig. 17 and 22). According to previous normal mode calculations on VPR that area in is one of the more flexible areas in the structure of VPR [102].

In the structure of AQUI this region is highly positively charged and Arg rich, but as mentioned before higher Arg content has been linked to thermostability (Table. 6). The first mutant produced in this series was A119H/S120R/G121R (3x). That mutant resulted in a decreased stability and had much lower activity [102]. Then more mutations where made, 4x, 5x and 6x, with more additions both activity and stability got higher comparing to the 3x mutant (Table. 3). Effects on activity where both higher $k_{cat}$ and lower $K_m$ resulting in almost twice as high catalytic efficiency, whereas the stability was almost on pair with the wild-type VPR.

*Table 6. Amino acid composition of loop area connecting helix C and sheet 4 in the structures of VPR, 6x mutant, AQUI, PRK and SPRK. Residues in the loop itself are in italics and underlined.*

| | **Position** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Enzyme** | 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | | 123 | 124 |
| **VPR** | V | A | Q | N | A | *S* | *G* | *P* | - | S | V |
| **6x** | V | T | R | N | H | *R* | *R* | *P* | - | A | V |
| **AQUI** | V | T | R | N | H | *R* | *R* | *P* | - | A | V |
| **PRK** | V | K | N | N | R | *N* | *C* | *P* | *K* | G | V |
| **SPRK** | V | K | N | N | A | *S* | *G* | *P* | - | A | V |

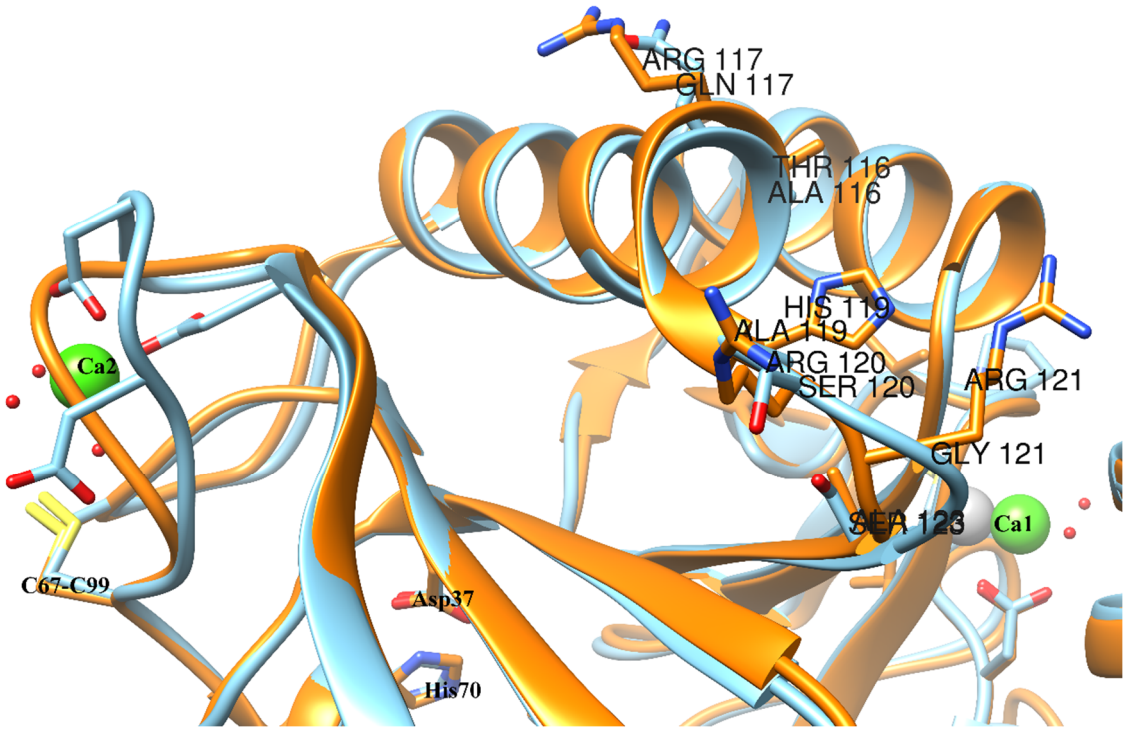


*Figure 22. Almost 180° horizontal turn of Fig. 13. Amino acids involved in the 6x mutation are marked in along with Asp37 and His70 of the active site. VPR is light blue with green calcium ions (PDB code: 1SH7) and AQUI is light orange with grey calcium ions (PDB code: 4DZT). Calcium numbering according to VPR as well as residue numbering. The atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

## 1.4 The aim of the project

This project is a part of a larger investigation of temperature adaptation of proteins conducted under Professor Magnús Már Kristjánsson. The project was meant to utilize work previously done on the subtilisin-like serine proteases to rationally design a multiple

mutant VPR that was both more stable and more active. For that purpose two different starting mutants were chosen as templates for the research the $VPR_{\Delta C}$/N3P/I5P and $VPR_{\Delta C}$/A116T/Q117R/A119H/S120R/G121R/S123A (6x). The rationalization for these particular mutants is that $VPR_{\Delta C}$/N3P/I5P had shown a big increase in stability but lacking activity and VPR/6x has been shown to have increased catalytic efficiency with only a little lower $T_{50\%}$ and $T_m$ values. The ΔC incorporation in all mutations was made to simplify the comparison model of VPR and AQUI and preventing complications of data analysis as shown in the case for some mutations such as N15D. On top of these templates two mutations were chosen for their effects on VPR. N15D was chosen for its stability effects on $VPR_{\Delta C}$ and Q142K was chosen for its remarkable effects on activity of $VPR_{\Delta C}$. These additions where added stepwise by site directed mutagenesis and all products were measured with regard to stability and activity. The mutants measured in this project were firstly $VPR_{\Delta C}$/N3P/I5P/N15D, $VPR_{\Delta C}$/N3P/I5P/Q142K and $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K, and secondly $VPR_{\Delta C}$/6x, $VPR_{\Delta C}$/6x/N15D, $VPR_{\Delta C}$/6x/Q142K and $VPR_{\Delta C}$/6x/N15D/Q142K.

Another part of this project was to investigate the flexibility change of N-terminus of $VPR_{\Delta C}$/N3P/I5P and compare it with $VPR_{\Delta C}$, AQUI and AQUI/Y191W. This was done via fluorescence quenching with acrylamide of the native forms of the enzymes at 25°C and 45°C and utilizing a tryptophan residue located at position 6 in VPR (corresponding to 8 in AQUI and $VPR_{\Delta C}$ /N3P/I5P) to gain insight into changes and/or lowered flexibility of the N-terminal region of $VPR_{\Delta C}$/N3P/I5P.

# 2 Materials and methods

## 2.1 Bacterial strains and plasmids

The bacterial strain used for production and expression of VPR and its mutants was the *E.coli* strain TOP10 from Invitrogen, which has the genotype:

*$F^-$ mcrA Δ(mmr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recAlaraD139 Δ(ara-leu)7697 gal/U gal/K rpsL ($Str^R$) endA1 nupG.*

The gene encoding for VPR originates from *Vibrio* PA44 [14,92]. The gene had earlier been cloned into a pBAD plasmid with a TOPO TA Cloning® kit from Invitrogen, the plasmid also contains an antibiotic resistance gene for ampicillin ($Amp^R$) [94]. The TOP10 strain which is ampicillin sensitive was also used for transformation, and grown on $L_{amp}$ agar plates to select for cells containing the plasmid. In this system L-(+)-arabinose from Sigma-Aldrich was used to induce expression.

## 2.2 Site-directed mutagenesis (PCR)

Site-directed mutagenesis was performed on the VPR gene following instructions form the Quick Change® Site-Directed Mutagenesis Kit form Stratagene. For each mutation a specific primer was used (See. 2.2.1) and amplified in a typical PCR using a *Pfu* polymerase from Thermo Scientific. The reaction was carried out in an automatic block heater (Veriti® Thermal Cycler form Life Technologies) with a given number of cycles and temperatures (Table. 7). The preparation of the reaction mixture was a follows:

1.5µl DNA template containing the appropriate VPR template (ca. 50-100 ng of DNA).

1 µl of forward and reverse (fw/rv) primers (10 pmol/µl), from Eurofins MWG Operon.

1 µl dNTP (10 mM, final conc. 0.2 mM), from Fermentas.

1 µl *Pfu* DNA polymerase (2.5 U/µl), From Thermo Scientific.

Then the solution was diluted to 50 µl with dd$H_2O$.

*Table 7. Conditions in the PCR.*

| | Temperature (°C) | Time (min) | Number of cycles |
|---|---|---|---|
| **Initilization** | 95 | 1 | |
| **Denaturation** | 95 | 1 | |
| **Annealing** | X* | 1 | **30x cycles** |
| **Elongation** | 72 | 11 | |
| **Extension** | 72 | 11 | **1x cycle** |
| **Cooling** | 10 | ∞ | **End** |

*X stands for the annealing temperature.

The original methylated DNA of the PCR mixture was then digested with *DpnI* endonulease from Thermo Scientific at 37°C, overnight.

### 2.2.1 Primers

The mutagenic PCR primers were designed with the web-based program PrimerX (http://www.bioinformatics.org/primerx/), that calculates the melting point ($T_m$) and GC content. The primers that were used in this project were synthesized by Eurofins MGW Operon.

*Table 8. Forward (fw) and reverse (rv) primers used for site directed mutagenesis for the VPR mutants. The table lists the base composition of the primers as well as their GC content (GC%), melting point ($T_m$), annealing temperatures (An). The codon containing the mutation is underlined and the mutating bases are in bold.*

| Primer | Base sequence | GC % | $T_m$ (°C) | An (°C) |
|---|---|---|---|---|
| N15D fw | 5´-GAATAGATCAGAGA**G**ACCTTCCTTTGGATC-3´ | 43.3 | 73.6 | 60.4* |
| N15D rv | 5´-GATCCAAAGGAAG**C**TCTCTGATCTATTC-3´ | 43.3 | 73.6 | 60.4* |
| Q142K fw | 5´-CATTAGATAGCGCGGTG**A**AAGGCGCGATTCAATCTG-3´ | 50.0 | 80.2 | 66.7* |
| Q142K rv | 5´-CAGATTGAATCGCGCCTTT**T**CACCGCGCTATCTAATG-3´ | 50.0 | 80.2 | 66.7* |
| C277_ fw | 5´ GCAGACAGTGGTT**AA**GAGCCGGATTGCGG 3´ | 58.6 | 75.4 | 65.9* |
| C277_ rv | 5´ CCGCAATCCGGCTC**TT**AACCACTGTCTGC 3´ | 58.6 | 75.4 | 65.9* |

*Difficulties caused experimentation regarding annealing temperatures and the values seen in this table are the ones that provided the best results.

## 2.3 Medium

Media used were both LB medium (Luria-Bertani) and 2xYT medium (2xYeast extract and Tryptone (Table. 9). LB broth was used for transformation, starter cultures and agar plates. 2xYT broth was used for expression cultures. Media were sterilized immediately after preparation. When needed ampicillin was added to the media to a final conc. of 0.1 mg/mL prior to use to hinder the growth of non-transformed cells and unwanted bacteria. For the preparation of L agar plates, 7 g of agar (Fluka-Biochemika) were added to 1 L of LB medium and sterilized. For $L_{amp}$ agar plates, ampicillin was added to a molten agar solution (added when the temperature had cooled down to approximately 60°C) to a final conc. of 0.1 mg/mL.

*Table 9. Contents of LB and 2xYT media. Recipe for 1 L.*

| Medium | Tryptone (Bacto ™) | Yeast Extract (Bacto ™) | NaCl (Sigma) |
|---|---|---|---|
| **LB** | 10 g | 5 g | 10 g |
| **2xYT** | 16 g | 10 g | 5 g |

## 2.4 Transformation

*E. coli* Top10 cells had previously been made chemically competent [104] by other members of the research group. For transformation with the PCR product, 20-25 µl of the PCR solution was added to 200 µl of competent cells that had been thawed on ice. The solution was then mixed thoroughly and kept on ice for 10 min and after that heat shocked for 1 min at 42°C. After that treatment 1 mL of LB (without ampicillin) was added to the solution, mixed thoroughly and incubated at 37°C, 300 rpm for 1 hour. After incubation the samples where centrifuged at 13,000 rpm for 1 min, supernatant thrown away and the precipitate suspended in 200 µl of $dH_2O$, spread on $L_{amp}$ agar plates and incubated at 37°C overnight. For transformation with purified plasmids 2-5 µl of plasmid solution (40-80 ng/µl) were used instead and after incubation samples were not centrifuged, but 50 - 100 µl spread on the agar plate. Along with this negative and positive controls were also prepared, the negative control had only Top10 cells while the positive control was transformed with a purified pBAD plasmid.

## 2.5 Plasmid purification

After the transformed cells had been incubated at 37°C overnight a single colony was transferred into a tube containing 3-4 mL of $LB_{amp}$ broth and incubated overnight at 37°C (no longer than 16 hours). The purification itself was carried out using GeneJET™ Plasmid Miniprep Kit from Thermo Scientific, following their instruction.

Concentration of plasmid and their purity was measured with a NanoDrop 1000 Spectrophotometer by measuring absorbance at 230 nm, 260 nm and 280 nm. The concentration of the samples was evaluated from the absorbance at 260 nm according to the Beer-Lambert law (Eq. XI) (Were A is the absorbance, ε is the molar attenuation coefficient, c is the concentration and b is the path length).

$$A = \varepsilon * c * b \qquad \text{(XI)}$$

Estimation of purity of the sample was determined by the $A_{260}/A_{280}$ ratio and the $A_{260}/A_{230}$ ratio. $A_{260}/A_{280}$ ratio of 1.7 – 2.0 means a good quality sample, free of RNA and other contaminants absorbing at 280 nm. The $A_{260}/A_{230}$ ratio is indicative of the presence of organic compounds or chaotropic salts, a ratio of 1.5 or greater is therefore preferred. For size determination and further proof of a pure sample, agarose electrophoresis was conducted. Samples were prepared by mixing 8 µl of plasmid solution and 2 µl of 10x loading buffer from Invitrogen™ (65% (w/v) sucrose, 10 mM Tris-HCl (pH 7.5), 10mM EDTA and 0,3% Bromophenol Blue). Samples were loaded on 1% (w/v) agarose gel in 1xTAE buffer system (40mM Tris, 20mM acetic acid, pH 8.0 containing 1 mM EDTA). Running with the samples was also a positive control containing pBAD purified plasmid (containing the VPR gene). The electrophoresis was carried out at 90V for a minimum of 15 min. The resulting gel was then viewed using an UV light transilluminator. The plasmids were stored at -20°C between use.

## 2.6 Sequence analysis

Mutations done on VPR were all confirmed by Sanger DNA analysis done by Beckman Coulter Genomics (http://www.beckmangenomics.com), United Kingdom. Primers for the sequence analysis were universal pBAD primers provided by Beckman Coulter Genomics.

Both forward and reverse primers were used depending on the location of the mutation within the gene.

## 2.7 Cultivation and expression of VPR

For cultivation of expression batches of VPR 2xYT broth was used. A single colony from a petri dish that had been confirmed to contain the gene for VPR and the right mutations was placed in 20 mL of 2xYT containing 0.1 mg/mL of ampicillin ($2xYT_{amp}$). This starter was then cultivated overnight at 37°C, 180 rpm in a New Brunswick™ scientific Innova® 44 incubator shaker. The sample was then diluted to 1 L of $2xYT_{amp}$ and cultivated at 37°C and 180 rpm until $A_{600}$ was between 0.8-1.8. At that time the cultivation broth was made 10 mM $CaCl_2$ (Sigma) and 0.02% L-(+)-Arabinose to induce expression of the protein. The culture was then incubated at 18°C and 180 rpm for 20 - 24 hours. The 1 L cultures were then centrifuged in a Beckman Coulter Avanti® J-26XP centrifuge at 4000 *g* and 10°C for 15 minutes (JLA-8.1000 rotor). Following that, the supernatant was removed and the pellets resuspended and washed in 500 mL of $dH_2O$ and centrifuged at 6500 *g* at 10°C for 15 minutes. Pellets were then transferred into 50 mL falcon tubes and dissolved in 50 mL $dH_2O$ and centrifuged at 5500xg and 10°C for 25 minutes (JS-5.3 rotor). Supernatants were discarded and pellets dried as much as possible and then transferred to -25°C freezer prior to usage.

## 2.8 Purification

### 2.8.1 The purification of VPR

The purification of VPR was carried out mostly as previously described [92]. Cell pellets containing the VPR mutants were dissolved in 50 mL of a buffer containing 25 mM Tris-Cl (Sigma), 10 mM $CaCl_2$ (Sigma) and calibrated to pH 8.0 at 25°C with 6 M HCl (Buffer A, see appendix 1). The sample was then made 1mg/mL in lysozyme (Sigma) and 1 µg/mL DNAase (Sigma) and the mixture shaken gently for at least 2 hours prior to freezing with liquid nitrogen and thereafter allowed to thaw at 4°C overnight while being shaken.

The mixture was then frozen again and thawed at room temperature while being shaken, and this freeze/thaw cycle was repeated two times. The sample was then transferred to a 40°C water bath for 1 hour to obtain the fully active and mature protease.

The sample was then centrifuged at 20,000 *g* and 4°C for 45 minutes (JLA-16.250 rotor). The supernatant was collected and made to 80% saturation in ammonium sulfate (Sigma) ($(NH_4)_2SO_4$) and centrifuged again under the same conditions. The supernatant was discarded and the precipitate dissolved in 50 mL of Buffer A and kept at 4°C overnight.

The sample was then loaded onto a Z-D-Phe-TETA (N-carbobenzoxy-D-phenylalanyl-triethylenetetramine-Sepharose) column equilibrated with Buffer A. After the sample had been loaded onto the column it was eluted with Buffer A containing 0.5 M NaCl (Sigma) to elute proteins that had bound to the column in a non-specific way. The absorbance at 280 nm was observed and when no observable change in $A_{280}$ had been observed for some time (generally around $A_{280} < 0.05$) the column was washed with approximately two column volume equivalents. For the elution of the protein, 2 M GdmCl (Sigma) in Buffer A was used and collected in 2.5 mL aliquots that were diluted with 2 mL of 3 M ammonium sulfate.

The collected protease sample from the Z-D-Phe-TETA column was then loaded onto a phenyl Sepharose column previously equilibrated with Buffer A, containing 1 M ammonium sulfate. The column was then eluted with the same buffer as it was equilibrated and then the ammonium conc. lowered to 0.1 M eluting a sharp peak of contaminants and then the protease was eluted with 50% ethylene glycol in Buffer A. Prior to storage the samples were diluted to 25% ethylene glycol with buffer A and stored in a freezer at -25°C.

### 2.8.2 Purification of AQUI and AQUI/Y191W

Purification of AQUI and AQUI/Y191W had previously been carried out in the lab as described in [92,94,95].

### 2.8.3 Zaman-Verwilghen protein quantitation

To determine the protein concentrations in samples in the protein purification process, a Coomassie Brilliant Blue G250 protein staining assay was performed (Zaman-Verwilghen variation) [105]. The samples were prepared by combining 2.75 mL of the Coomassie Blue solution with 0.25 mL of a protein solution (or $dH_2O$ for the blank) and mixed thoroughly and incubated at room temperature for 15 minutes prior to measurements at 620 nm.

### 2.8.4 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Gel electrophoresis was conducted on pre-cast NuPAGE® Novex® 4-12% gradient Bis-Tris 15 well gel. The samples were prepared by mixing 72 μL of protein sample, 8 μL PMSF (25 mM) (Sigma) and 20 μL of a solution containing 0.5 M Tris/HCl, 10% w/v SDS, 50% w/v sucrose, 0.25 M dithiothreitol and 0.02% w/v bromophenol blue, pH 6.8. These samples were then boiled for 5 minutes prior to loading to the gel. 10 μL of each sample was loaded to the wells and 8 μL of the ladder (Spectra™ Multicolor Broad Range Protein Ladder). Following electrophoresis, the gels were stained with Coomassie Brillian Blue R-250 (0.25% (w/v) in 45% methanol, 7% acetic acid) overnight, and destained (45% (v/v) methanol and 7% (v/v) acetic acid).

## 2.9 Enzymatic activity assays

Activity of VPR and its mutants was assayed by using the synthetic substrate, succinyl-AlaAlaProPhe-p-nitroanilide (sAAPF-pNA) (Bachem). The substrate was kept dissolved in DMSO at concentration of 25 mM at 4°C prior to use. Assays were carried out at room temperature (25°C for Michaelis-Menten measurements) using 0.5 mM substrate that had been diluted in 100 mM Tris-Cl, 10 mM $CaCl_2$ and pH 8.6 at 25°C. The increase in absorbance at 410 nm, over a 30 second time interval and using the molar attenuation coefficient (ε) of 8480 $M^{-1}cm^{-1}$ to determine the reaction rate (V) (Eq. XII).

$$\frac{\Delta A/sec}{8480\ M^{-1}cm^{-1}} * \frac{1000\ mM}{M} = \frac{mM}{sec} = V \qquad \text{(XII)}$$

## 2.10 Michaelis-Menten kinetics

Prior to measurements, proteins samples were dialyzed against a 100 mM Tris-Cl, 10 mM $CaCl_2$, pH 8.6 (at 25°C), overnight at 4°C. Seven different substrate solutions were prepared by diluting a stock substrate solution to a final concentration of 0.075 mM, 0.10 mM, 0.15 mM, 0.25 mM, 0.50 mM, 0.75 mM and 1.00 mM into the same type of buffer as the protein was dialyzed in. The concentration of the protein solution was determined by measuring the absorbance at 280 nm and calculated according to equation XI. The molar attenuation coefficients for VPR, VPR mutants, AQUI and AQUI/Y191W were calculated using the web-based program ProtParam (http://web.expasy.org/protparam) (34,170 $M^{-}$

$^{1}cm^{-1}$ for VPR and all its mutants used in this project) [106]. Protein conc. was determined by at least three independent measurements of the protein stock on a Cary 50 Bio UV-Visible spectrometer (Varian).

The protein stock solutions were then diluted to have activity of ~1 U/mL (950 µL substrate against 50 µL of protein solution) at 0.5 mM substrate for each set. One set of measurements was triplicate measurements at each substrate concentration and usually nine sets for each mutant spread over three days with each stock prepared the day before. Measurements were carried out at 25°C using a Heλios α UV-Visible spectrometer from Thermo Electron Corporation equipped with a ThermoSpectronic Single Cell Peltier system.

For the determination of the constants $k_{cat}$ and $K_m$ the data points for each set were plotted using the program KaleidaGraph and Michaelis-Menten non-linear regression performed according to equation XIII.

$$V = \frac{V_{max} * [S]}{K_m * [S]} \qquad \text{(XIII)}$$

Where [S] stands for substrate and $V_{max}$ for the maximum reaction rate and $k_{cat}$ was calculated according to:

$$k_{cat} = \frac{V_{max}}{[E]_o} \qquad \text{(XIV)}$$

Where $[E]_0$ stands for the enzyme concentration in the reaction mixture. For each dataset, the catalytic efficiency was calculated with the ratio of $k_{cat}/K_m$. These constants were then expressed as the mean value of all the sets measured and standard deviation of the mean calculated to estimate the precision of the measurements.

## 2.11 Melting point determination ($T_m$)

To determine the melting point ($T_m$, the temperature where half of all protein molecules in the solution have been denatured), of the VPR mutants, circular dichroism (CD) measurements were performed [94]. Preparation of samples was as follows: Samples were inhibited to a final PMSF concentration of 1 mM and samples were concentrated with Amicon® Ultra -0.5 mL Centrifugal Filters with 10 kDa molecular weight cutoff

membrane up to a minimum concentration corresponding to $A_{280}$ of 1.2. Samples were then dialyzed against a buffer containing 25 mM glycine (Sigma), 100 mM NaCl, 15 mM $CaCl_2$, pH 8.6 (at 25°C), overnight at 4°C. Measurements were carried out on a JASCO-810 Circular Dichroism spectropolarimeter equipped with PTC-423S Peltier type single cell holder and temperature control system, using a 0.5 cm cuvette. The change in circular dichroism (mdeg) was monitored at 222 nm while raising the temperature from 20°C to 90°C with a heat gradient of 1°C/min.

To interpret the data, we assume a two-state unfolding process so the solution contains folded (F) and unfolded (U) states. This leads to the relationship:

$$f_F + f_U = 1 \qquad \text{(XV)}$$

Where $f_F$ and $f_U$ are the fractions of folded and unfolded states in the solution. When absorbance (mdeg) is plotted against temperature gives a sigmoidal curve (Eq. XVI). To normalize that cure, linear regression was made on the folded-state part and the unfolded-state part and extrapolated over the data set.

$$y = y_f f_F + y_u f_U \qquad \text{(XVI)}$$

Here $y_f$ and $y_u$ stand for the values obtained by the extrapolation. The curve was then normalized according to:

$$f_U = {(y_f - y)}\Big/{(y_F - y_U)} \qquad \text{(XVII)}$$

The normalized curve was then fitted by a sigmoidal curve with KaleidaGraph and the melting point ($T_m$) determined ($T_m$ = 0.5 $f_U$). Each mutant was measured at least in triplicate and expressed as the mean value with the standard deviation of the mean to determine the precision of the experiment.

## 2.12 Rate of thermal inactivation ($T_{50\%}$)

$T_{50\%}$ was determined by monitoring the thermal inactivation at selected temperatures (±4-5°C around $T_{50\%}$). Prior to measurements samples were dialyzed overnight at 4°C against a buffer containing 25 mM Tris-Cl, 15 mM $CaCl_2$, 100 mM NaCl, 1 mM EDTA (Sigma) and pH calibrated to 8.95.

Samples with activity adjusted to ~1 U/mL were incubated in a water bath at a constant temperature and activity measured at regular intervals over time (5-6 measurements for each temperature in every set) at room temperature using a Cary50 Bio UV-Visible spectrometer (Varian). To calculate the rate of thermal inactivation at each temperature the data was fitted by:

$$V = V_0 * e^{-kt} \qquad \text{(XVIII)}$$

Where V is the activity (enzyme velocity) at a given time, $V_0$ is the initial activity, k is the first order rate constant ($s^{-1}$) and t is the time. The natural logarithms of the rate constants at each temperature were then plotted versus 1000/K (K stands for temperature in Kelvin) yielding linear data and the data were then fitted with a modified version of the Arrhenius equation:

$$\ln(k) = 1000 * -\left({}^{E_a}/_{(R * T)}\right) * \ln(A) \qquad \text{(XIX)}$$

Where A is the pre-exponential factor, $E_a$ is the activation energy, R is the universal gas constant (8.314 J/mol*K) and T is the temperature in Kelvin.

$T_{50\%}$ is defined as the temperature where 50% of the initial activity has been lost over 30 minutes. To calculate the rate needed for 50% activity lost over 30 minutes ($k_{50\%}$) yields:

$$k_{50\%}(s^{-1}) = \frac{\ln(100) - \ln(50)}{30min * 60 {}^{s}/_{min}} \qquad \text{(XX)}$$

Therefore $T_{50\%}$ is:

$$T_{50\%} = \ln(k_{50\%}) - \left({}^{\ln(A)}/_{(E_a * R^{-1})}\right) \qquad \text{(XXI)}$$

Each set of measurements contained 5-8 data points for an Arrhenius plot. The final values contained at least three independent sets, expressed as the mean value with the standard deviation of the mean to determine the precision of the experiment.

## 2.13 Fluorescence experiments

Prior to fluorescence measurements protein samples were inhibited with PMSF to a final concentration of 1 mM and dialyzed overnight at 4°C against a buffer containing 50 mM

Tris-Cl, 10 mM $CaCl_2$ and pH calibrated to 8.0 at either 25°C or 45°C depending on the experiment at hand.

Fluorescence spectra were measured on Horiba FluoroMax-4 spectrofluormeter equipped with a Thermo Scientific Haake A25 water bath circulator. All samples were exited at 285 nm with a slit width of 3 nm and the emission spectra were measured between 300 nm and 400 nm, collecting data points with 0.5 nm intervals. As measurements on this spectrofluormeter should not exceed two million counts per second (CPS) the exit slit had to be varied (in the range of 4 nm to 8 nm). To correct for the varying exit slit widths, relative CPS was calculated according to:

$$Relative\ CPS = \frac{Measured\ CPS}{(ex.\ slit\ (nm))^2} \qquad \text{(XXII)}$$

Blanks were measured for all exit slits used and subtracted from the corresponding measurements of protein samples.

### 2.13.1 Emission spectra

Emission spectra of $VPR_{\Delta C}$, $VPR_{\Delta C}$ /N3P/I5P, AQUI at varying concentrations and exit slits were measured for both native and denatured states. For denaturing protein samples, $VPR_{\Delta C}$ and $VPR_{\Delta C}$/N3P/I5P, were incubated at 70°C for 30 minutes and AQUI and AQUI/Y191W at 95°C for 30 minutes. All samples were measured at 25°C and relative CPS/mol plotted to compare the emission spectra of the enzymes. Each set of measurements was done in quadruplicate, three sets were done for $VPR_{\Delta C}$ and $VPR_{\Delta C}$/N3P/I5P but one set for AQUI and AQUI/Y191W. The curves were then fitted with a 3$^{rd}$ degree polynomial to calculate the $\lambda_{max}$ of the enzymes.

### 2.13.2 Fluorescence quenching

Prior to dialysis, samples of $VPR_{\Delta C}$, $VPR_{\Delta C}$ /N3P/I5P, AQUI and AQUI/Y191W were either concentrated or diluted to a $A_{280}$ value of approximately 0.1 (to calculate the concentration of AQUI and AQUI/T191W the molar attenuation coefficients used were 34,630 $M^{-1}s^{-1}$ and 38,640 $M^{-1}s^{-1}$ respectively). Acrylamide (Sigma) was used as a quencher in these experiments, and was dissolved to a concentration of 2.5 M in the same buffer as the samples were dialyzed against. For each experiment 500 µL of protein were used and by adding 4 µL aliquots (11 in total for each sample) of the acrylamide stock added and the

emission spectra measured after each addition after thorough mixing and at least 45 second waiting time between measurements.

To translate the data into an estimate of flexibility, The Stern-Volmer equation was used (Eq. XXIII) [107].

$$\frac{F^0}{F} = 1 + k_q \tau_0 [Q] = 1 + K_{SV}[Q] \qquad \text{(XXIII)}$$

Where $F^0$ is the fluorescence of an unquenched sample, F is the fluorescence after each addition and [Q] is the quencher concentration. $K_{SV}$ is the Stern-Volmer quenching constant which is equal to the multiple of $k_q$ and $\tau_0$. $\tau_0$ is the unquenched lifetime and $k_q$ is the bimolecular quenching constant. A larger $K_{SV}$ represents higher accessibility of Trp residues within the structure, indicating higher flexibility of the structure in the vicinity of fluorophores in the protein. For calculations dilution of the protein sample after each addition of acrylamide was accounted for, as well as blanks with corresponding concentrations of acrylamide were measured. $K_{SV}$ determination for each enzyme was done in one to three sets, each set containing two to five measurements.

# 3 Results

## 3.1 Purification

Expression and purification of $VPR_{\Delta C}$ and all its mutants produced in this project were successful. Notably it was confirmed that for VPR it is crucial to add $CaCl_2$ to a final concentration of 10 mM to the expression cultures prior to inducing expression of the protein. For purification 70-90% yields should be expected, everything under 60% is unnecessary loss of protein and probably due to overloading the TETA column.

*Table 10. An example purification table for $VPR_{\Delta C}$.*

| Step | Volume [mL] | Conc. (C) [mg/mL] | Activity ([U/mL] | Total activity units (AU) [U] | Total protein ([mg] | Specific activity ([U/mg] | Yields (%) | Purification (x) |
|---|---|---|---|---|---|---|---|---|
| **Cell lysate** | 45 | 6.17 | 14.0 | 200 | 277.5 | 0.7 | 100 | 1 |
| **Lysate spin** | 40 | 2.49 | 15.6 | 200 | 99.5 | 2.0 | 100 | 3 |
| **$(NH_4)_2SO_2$ precipitate** | 53 | 2.04 | 20.2 | 233 | 108.0 | 2.2 | 117 | 3 |
| **TETA column** | 113 | 0.05 | 3.9 | 196 | 6.1 | 32.0 | 98 | 44 |
| **Phenyl sepharose column** | 36 | 0.22 | 7.9 | 177 | 0.8 | 220.6 | 89 | 306 |

## 3.2 The N3P/I5P additions

### 3.2.1 Addition of the N15D mutation to the $VPR_{\Delta C}$/N3P/I5P template

The first mutation on the $VPR_{\Delta C}$/N3P/I5P template was N15D and was meant to introduce a salt bridge into the structure. The resulting mutant had a $T_m$ of 70.4°C and $T_{50\%}$ of 65.5°C that is 5.2°C higher $T_m$ than for $VPR_{\Delta C}$ and 1.1°C higher than $VPR_{\Delta C}$ /N3P/I5P (Table 11). Even larger effects were observed when looking at the $T_{50\%}$ values, as the mutation resulted in a 9.1°C higher value than for $VPR_{\Delta C}$ and 3.6°C higher than for $VPR_{\Delta C}$/N3P/I5P. These results strongly support previous results indicating that this salt bridge is crucial for

thermal stability as it increases the stability of VPR and every mutant of VPR tested so far (Table 11).

*Table 11. Thermal stability of $VPR_{wt}$, VPR/N15D, $VPR_{\Delta C}$, $VPR_{\Delta C}$ /N15D, $VPR_{\Delta C}$ /N3P/I5P and $VPR_{\Delta C}$ /N3P/I5P/N15D (n = 3, for $T_{T50\%}$ and n = 4, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | $T_m$ (°C) | $T_{50\%}$ (°C) |
|---|---|---|
| **$VPR_{wt}$** | 63.6 ± 0.3 | 56.0 ± 0.2 |
| **VPR/N15D** | 66.3 ± 0.1 | 58.8 ± 0.5 |
| **$VPR_{\Delta C}$** | 65.2 ± 0.2 | 56.4 ± 0.1 |
| **$VPR_{\Delta C}$/N15D** | 68.2 ± 0.2 | 59.6 ± 0.2 |
| **$VPR_{\Delta C}$/N3P/I5P** | 69.3 ± 0.2 | 61.9 ± 0.2 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D** | 70.4 ± 0.1 | 65.5 ± 0.1 |

A most interesting observation for this mutant is the fact that the activity was measured to be almost five times higher than that of $VPR_{\Delta C}$/N3P/I5P in terms of $k_{cat}$. This is particularly interesting as the N15D mutation had only been shown to increase the activity of the wild-type VPR and was related to the presence of the C-terminus, as when it was removed as in $VPR_{\Delta C}$, the N15D mutation did not seem to have any notable effects on activity. Effects on $K_m$ were also observed, resulting in higher affinity for the substrate.

*Table 12. Kinetic parameters of $VPR_{wt}$, VPR/N15D, $VPR_{\Delta C}$, $VPR_{\Delta C}$ /N15D, $VPR_{\Delta C}$ /N3P/I5P and $VPR_{\Delta C}$ /N3P/I5P/N15D (n = 8). Expressed as mean values ± standard deviation of the mean.*

| | $k_{cat}$ ($s^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($s^{-1}mM^{-1}$) |
|---|---|---|---|
| **$VPR_{wt}$** | 74.6 ± 5.0 | 0.166 ± 0.017 | 449 ± 30 |
| **VPR/N15D** | 137 ± 17 | 0.173 ± 0.015 | 803 ± 135 |
| **$VPR_{\Delta C}$** | 68.2 ± 9.9 | 0.184 ± 0.017 | 371 ± 26 |
| **$VPR_{\Delta C}$/N15D** | 64.9 ± 4.0 | 0.179 ± 0.009 | 365 ± 7 |
| **$VPR_{\Delta C}$/N3P/I5P** | 8.8 ± 1.3 | 0.224 ± 0.010 | 40 ± 7 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D** | 51.2 ± 4.3 | 0.207±0.019 | 245 ± 19 |

### 3.2.2 The Q142K mutation to the N3P/I5P template

The second mutation on the $VPR_{\Delta C}$ /N3P/I5P template was the mutation Q142K located on α-helix D. This mutation is theorized to interact with a salt bridge in the vicinity and inducing flexibility as the mutant has shown to increase the activity almost twofold compared to $VPR_{\Delta C}$ and also resulted in lowering of the $K_m$ value leading to threefold increase in catalytic efficiency (Table. 14). The effects of this mutation on thermal stability

were almost negligible, although a little lowering in $T_m$ of $VPR_{\Delta C}$/Q142K as compared to $VPR_{\Delta C}$ was observed (Table 13). In combination with $VPR_{\Delta C}$/N3P/I5P, Q142K does not seem to affect the thermal stability of the native enzyme to any extent (Table 13).

*Table 13. Thermal stability of $VPR_{\Delta C}$, $VPR_{\Delta C}$/Q142K, $VPR_{\Delta C}$/N3P/I5P and $VPR_{\Delta C}$/N3P/I5P/Q142K (n = 3, for $T_{T50\%}$ and n = 4, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | **$T_m$ (°C)** | **$T_{50\%}$ (°C)** |
|---|---|---|
| **$VPR_{\Delta C}$** | 65.2 ± 0.2 | 56.4 ± 0.1 |
| **$VPR_{\Delta C}$/Q142K** | 64.6 ± 0.2 | 56.3 ± 0.3 |
| **$VPR_{\Delta C}$/N3P/I5P** | 69.3 ± 0.2 | 61.9 ± 0.2 |
| **$VPR_{\Delta C}$/N3P/I5P/Q142K** | 68.9 ± 0.3 | 61.8 ± 0.1 |

The Q142K mutation in $VPR_{\Delta C}$ /N3P/I5P/Q142K had a remarkable effect on the $k_{cat}$, however, increasing it more than tenfold compared to $VPR_{\Delta C}$ /N3P/I5P (Table. 14), but with no observable effects on $K_m$. These results give further evidence for activity promoting effect of this mutation.

*Table 14. Kinetic parameters of $VPR_{\Delta C}$, $VPR_{\Delta C}$/Q142K, $VPR_{\Delta C}$/N3P/I5P and $VPR_{\Delta C}$/N3P/I5P/Q142K (n = 4). Expressed as mean values ± standard deviation of the mean.*

| | **$k_{cat}$ ($s^{-1}$)** | **$K_m$ (mM)** | **$k_{cat}/K_m$ ($s^{-1}mM^{-1}$)** |
|---|---|---|---|
| **$VPR_{\Delta C}$** | 68.2 ± 9.9 | 0.184 ± 0.017 | 371 ± 26 |
| **$VPR_{\Delta C}$/Q142K** | 147.0 ± 7.7 | 0.157 ± 0.014 | 945 ± 51 |
| **$VPR_{\Delta C}$/N3P/I5P** | 8.8 ± 1.3 | 0.224 ± 0.010 | 40 ± 7 |
| **$VPR_{\Delta C}$/N3P/I5P/Q142K** | 94.1 ± 6.6 | 0.223 ± 0.014 | 422 ± 24 |

### 3.2.3 Combination of N15D and Q142K into $VPR_{\Delta C}$/N3P/I5P

The third and last mutation done on the $VPR_{\Delta C}$/N3P/I5P template was Q142K accompanied by N15D. The thermal stability of this mutant was almost on par with $VPR_{\Delta C}$/N3P/I5P/N15D, or somewhat higher (Fig. 23 and 24 and Table. 15). This result further establishes the importance of the salt bridge between the side chains of D15 and K257 formed with the N15D mutation for the stability of the native structure of VPR/N15D, as well as the thermophilic AQUI.

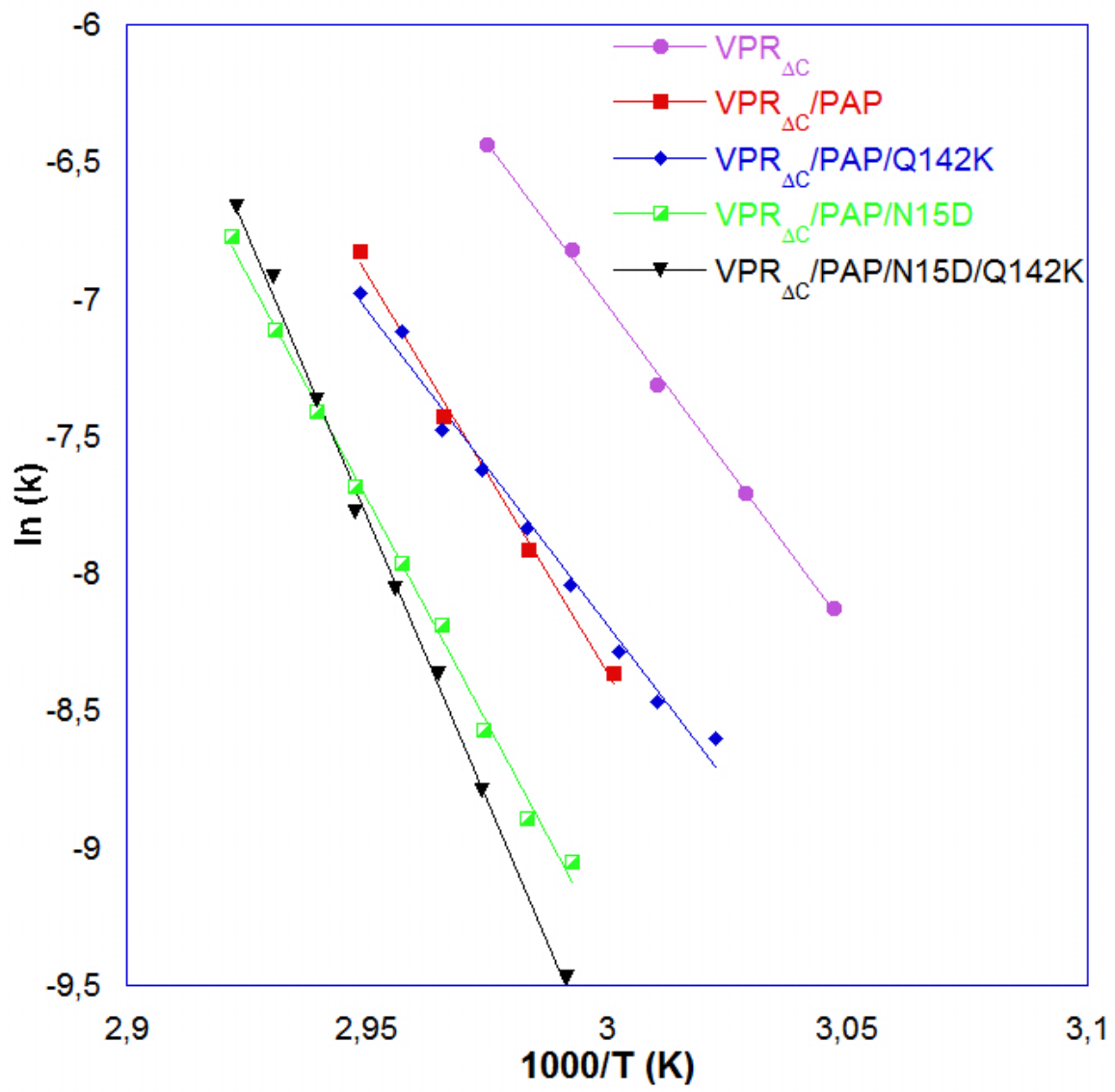


*Figure 23. Arrhenius plots the thermal inactivation of VPR$_{\Delta C}$ (purple circles), VPR$_{\Delta C}$/N3P/I5P (red squares), VPR$_{\Delta C}$/N3P/I5P/N15D (green half solid squares), VPR$_{\Delta C}$/N3P/I5P/Q142K (blue diamonds) and VPR$_{\Delta C}$/N3P/I5P/N15D/Q142K (black triangles). The rate of thermal inactivation determined by withdrawing aliquots at selected time intervals and measuring the remaining activity against the substrate sAAPF-pNA. The X-axis shows 1000/T (in Kelvin) and the Y-axis shows the natural logarithmic value of the rate constant k (s-1). PAP is a short for the N3P/I5P mutation.*

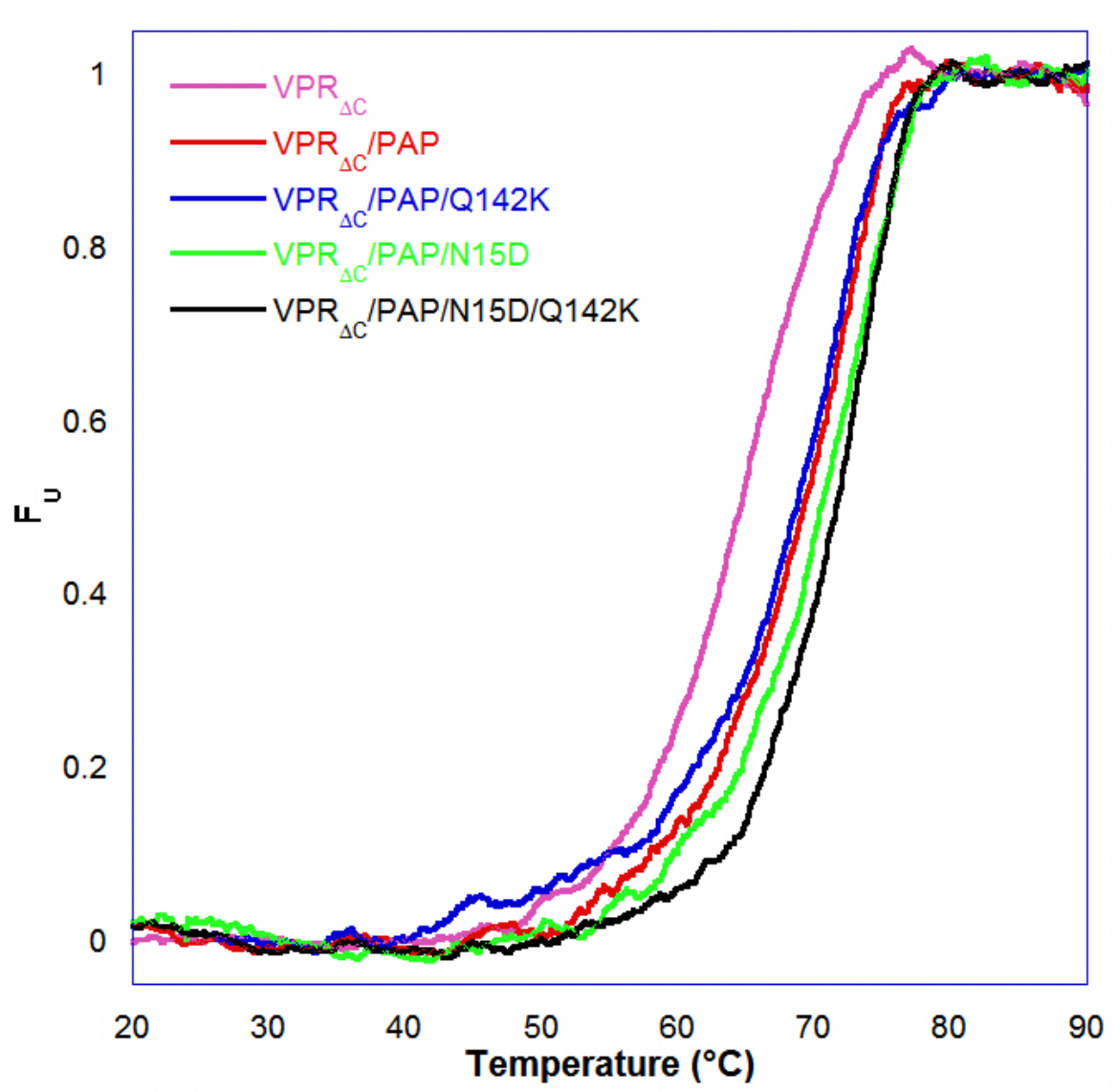


*Figure 24. Normalized melting curves of VPR$_{\Delta C}$ (purple), VPR$_{\Delta C}$/N3P/I5P (red), VPR$_{\Delta C}$/N3P/I5P/N15D (green), VPR$_{\Delta C}$/N3P/I5P/Q142K (blue) and VPR$_{\Delta C}$/N3P/I5P/N15D/Q142K (black, as measured by changes in circular dichroism at 222 nm (CD)). The X-axis shows the temperature in Celsius degrees and the Y-axis shows the fraction of unfolded protein at each temperature (on the scale 0 to 1). PAP is a short for the N3P/I5P mutation.*

The double proline substitution on the N-terminal region of VPR$_{\Delta C}$ leads to a significant stabilization, which also is further stabilized by the addition of N15D (Fig. 23 and Fig. 24). The Q142K does not seem to affect the thermal stability of the protein, although some indications cannot be fully ignored (see conclusions). Even though the stability of VPR$_{\Delta C}$/N3P/I5P/N15D/Q142K is just a little bit higher than VPR$_{\Delta C}$/N3P/I5P/N15D there are interesting effects observed in the kinetic parameters of that mutant (Table. 16). With

the N15D mutation some lowering in $K_m$ was observed, but none in the case of the mutant containing Q142K only. In the double mutant a lowering to 0.188 mM was observed as compared to that of 0.223 mM in the mutant containing only the Q142K added mutation. There is also an apparent decrease in $k_{cat}$, whereas the increase in $k_{cat}$ compared with $VPR_{\Delta C}$/N3P/I5P/N15D was slightly less than twofold.

*Table 15. Thermal stability of $VPR_{\Delta C}$/N3P/I5P, $VPR_{\Delta C}$/N3P/I5P/N15D, $VPR_{\Delta C}$/N3P/I5P/Q142K and $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K (n = 3, for $T_{T50\%}$ and n = 3, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | **$T_m$ (°C)** | **$T_{50\%}$ (°C)** |
|---|---|---|
| **$VPR_{\Delta C}$/N3P/I5P** | 69.3 ± 0.2 | 61.9 ± 0.2 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D** | 70.4 ± 0.1 | 65.5 ± 0.1 |
| **$VPR_{\Delta C}$/N3P/I5P/Q142K** | 68.9 ± 0.3 | 61.8 ± 0.1 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D/Q142K** | 71.1 ± 0.2 | 65.6 ± 0.2 |

*Table 16. Kinetic parameters of $VPR_{\Delta C}$/N3P/I5P, $VPR_{\Delta C}$/N3P/I5P/N15D, $VPR_{\Delta C}$/N3P/I5P/Q142K and $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K (n =6). Expressed as mean values ± standard deviation of the mean.*

| | **$k_{cat}$ ($s^{-1}$)** | **$K_m$ (mM)** | **$k_{cat}/K_m$ ($s^{-1}mM^{-1}$)** |
|---|---|---|---|
| **$VPR_{\Delta C}$/N3P/I5P** | 8.8 ± 1.3 | 0.224 ± 0.010 | 40 ± 7 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D** | 51.2 ± 4.3 | 0.207±0.019 | 245 ± 19 |
| **$VPR_{\Delta C}$/N3P/I5P/Q142K** | 94.1 ± 6.6 | 0.223 ± 0.014 | 422 ± 24 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D/Q142K** | 86.3 ± 6.4 | 0.188 ± 0.015 | 461 ± 20 |

## 3.3 The A116T/Q117R/A119H/G121R/S123A (6x) additions

The 6x template was made by introducing the C277_ (ΔC) mutation to VPR/6x. VPR/6x is the final product of a series of mutations done on a proposed hinge area, proposed to be important for movements related to activity of the enzyme. Mutants containing three, four and five residue exchanges had all negative effects on the enzyme, both with respect to catalytic efficiency and stability (Table. 3). The 6x mutant was the first in that series that had shown considerable increase in activity without affecting the stability drastically, although a little lowering was observed. The catalytic efficiency of the mutant was almost

twofold to that of the wild-type enzyme, which was caused by both an increase in $k_{cat}$, but also as a result of lowering in $K_m$.

Interesting results were obtained however after measuring $VPR_{\Delta C}$/6x. The $T_{50\%}$ value seems to be compatible with what was measured for $VPR_{\Delta C}$. Interestingly though, the melting point ($T_m$) of the mutant has decreased and is closer to that of the wild type enzyme rather than that of the $T_m$ of $VPR_{\Delta C}$. This is hard to explain and more research is needed to answer this question.

*Table 17. Thermal stability of VPR, VPR/6x, $VPR_{\Delta C}$ and $VPR_{\Delta C}$/6x (n = 3, for $T_{T50\%}$ and n = 4, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | **$T_m$ (°C)** | **$T_{50\%}$ (°C)** |
|---|---|---|
| **$VPR_{wt}$** | 63.6 ± 0.3 | 56.0 ± 0.2 |
| **VPR/6x** | 63.7 ± 0.1 | 55.1 ± 0.6 |
| **$VPR_{\Delta C}$** | 65.2 ± 0.2 | 56.4 ± 0.1 |
| **$VPR_{\Delta C}$/6x** | 62.8 ± 0.3 | 56.2 ± 0.2 |

The reason for using VPR/6x as a starting point was the interesting change in kinetic constants towards higher catalytic efficiency, but by removing the C-terminal arm drastic changes were observed mostly with regard to $k_{cat}$, with a twofold lowering compared to VPR/6x. It can be pointed out however that the lowering in $K_m$ that is observed is similar to the effects observed between VPR and VPR/6x.

*Table 18. Kinetic parameters of VPR, VPR/6x, $VPR_{\Delta C}$ and $VPR_{\Delta C}$/6x (n = 8). Expressed as mean values ± standard deviation of the mean.*

| | **$k_{cat}$ ($s^{-1}$)** | **$K_m$ (mM)** | **$k_{cat}/K_m$ ($s^{-1}mM^{-1}$)** |
|---|---|---|---|
| **$VPR_{wt}$** | 74.6 ± 5.0 | 0.166 ± 0.017 | 449 ± 30 |
| **VPR/6x** | 98.9 ± 13.4 | 0.148 ± 0.010 | 728 ± 64 |
| **$VPR_{\Delta C}$** | 68.2 ± 9.9 | 0.184 ± 0.017 | 371 ± 26 |
| **$VPR_{\Delta C}$/6x** | 50.3 ± 1.3 | 0.171 ± 0.009 | 293 ± 10 |

Despite these results it was decided to continue with this mutant and gain more insight into the effects of N15D and Q142K mutations on various templates.

### 3.3.1 The N15D mutation in the 6x template

The first mutation on $VPR_{\Delta C}$/6x was N15D, in order to introduce a salt bridge to the structure of the enzyme. Results from stability measurements strongly suggested that a salt bridge is introduced into the protein as in the case of previous mutants. The difference in values for $T_{50\%}$ and $T_m$ between $VPR_{\Delta C}$ and $VPR_{\Delta C}$/N15D and between $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/N15D are very similar or around 3°C in $T_m$ and around 2-3°C in $T_{50\%}$ (Table 19). These results seem to indicate that the insertion of this mutation leads to the forming of the salt bridge just as readily as in $VPR_{\Delta C}$ and that it has a similar effect on the stability of the enzyme.

*Table 19. Thermal stability of $VPR_{\Delta C}$, $VPR_{\Delta C}$/N15D, $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/N15D (n = 3, for $T_{T50\%}$ and n = 3, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | $T_m$ (°C) | $T_{50\%}$ (°C) |
|---|---|---|
| **$VPR_{\Delta C}$** | 65.2 ± 0.2 | 56.4 ± 0.1 |
| **$VPR_{\Delta C}$/N15D** | 68.2 ± 0.2 | 59.6 ± 0.2 |
| **$VPR_{\Delta C}$/6x** | 62.8 ± 0.3 | 56.2 ± 0.2 |
| **$VPR_{\Delta C}$/6x/N15D** | 65.2 ± 0.4 | 58.2 ± 0.1 |

Activity measurements of the mutant showed an increase in the $k_{cat}$ value of around 50% than for the template (Table 20). These results raised a question, as N15D had only been shown to increase activity in the wild type VPR, where the C-terminus was still present, but in the present experiments the N15D exchange increased both activity in the 6x template and the N3P/I5P template, indicating a more complex set of interactions caused by this mutation in terms of activity. Changes to $K_m$ were insignificant but suggested a slight lowering as earlier observed.

*Table 20. Kinetic parameters $VPR_{\Delta C}$, $VPR_{\Delta C}$/N15D, $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/N15D (n = 9). Expressed as mean values ± standard deviation of the mean.*

| | $k_{cat}$ ($s^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($s^{-1}mM^{-1}$) |
|---|---|---|---|
| **$VPR_{\Delta C}$** | 68.2 ± 9.9 | 0.184 ± 0.017 | 371 ± 26 |
| **$VPR_{\Delta C}$/N15D** | 64.9 ± 4.0 | 0.179 ± 0.009 | 365 ± 7 |
| **$VPR_{\Delta C}$/6x** | 50.3 ± 1.3 | 0.171 ± 0.009 | 293 ± 10 |
| **$VPR_{\Delta C}$/6x/N15D** | 79.6 ± 1.7 | 0.164 ± 0.005 | 485 ± 9 |

### 3.3.2 The Q142K mutation on the 6x template

The second mutation done on this template was as before Q142K. This mutation had previously not been reported to affect thermostability of VPR to any extent, but in this case of the $VPR_{\Delta C}$/6x/Q142K mutant it apparently had some impact on the stability with an increase of 1°C being observed for $T_{50\%}$ and 0.5°C for Tm, in case of the $VPR_{\Delta C}$/6x/Q142K mutant as compared to $VPR_{\Delta C}$/6x (Table 21).

*Table 21. Thermal stability of $VPR_{\Delta C}$, $VPR_{\Delta C}$/Q142K, $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/Q142K (n = 3, for $T_{T50\%}$ and n = 3, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | $T_m$ (°C) | $T_{50\%}$ (°C) |
|---|---|---|
| **$VPR_{\Delta C}$** | 65.2 ± 0.2 | 56.4 ± 0.1 |
| **$VPR_{\Delta C}$/Q142K** | 64.6 ± 0.2 | 56.3 ± 0.3 |
| **$VPR_{\Delta C}$/6x** | 62.8 ± 0.3 | 56.2 ± 0.2 |
| **$VPR_{\Delta C}$/6x/Q142K** | 63.9 ± 0.2 | 56.8 ± 0.3 |

Even though these changes are not relatively high, there is definitely a change as these mutants were measured side by side and the mutant containing Q142K was always measured to be more stable. A possible explanation of this stabilizing effect of Q142K in the 6x template might be related to the location of these mutations. The hinge area where the 6x mutations are located connects helix C to β-sheet 4 and β-sheet 4 is connected to α-helix D where Q142K is located. These helices also lie in close proximity of each other so presumably mutations effecting movements in that area may well affect each other, leading to complex dynamic changes around one of the most flexible areas in the enzyme (Fig. 15). As observed before in every mutant containing Q142K the activity increased greatly, with more than twofold increase in $k_{cat}$ observed as well as a little lowering in $K_m$, leading to more than 2.5 times increase in $k_{cat}$/Km for the mutant (Table 22).

*Table 22. Kinetic parameters $VPR_{\Delta C}$, $VPR_{\Delta C}$/Q142K, $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/Q142K (n = 9). Expressed as mean values ± standard deviation of the mean.*

| | $k_{cat}$ ($s^{-1}$) | $K_m$ (mM) | $k_{cat}/K_m$ ($s^{-1}mM^{-1}$) |
|---|---|---|---|
| **$VPR_{\Delta C}$** | 68.2 ± 9.9 | 0.184 ± 0.017 | 371 ± 26 |
| **$VPR_{\Delta C}$/Q142K** | 147.0 ± 7.7 | 0.157 ± 0.014 | 945 ± 51 |
| **$VPR_{\Delta C}$/6x** | 50.3 ± 1.3 | 0.171 ± 0.009 | 293 ± 10 |
| **$VPR_{\Delta C}$/6x/Q142K** | 123.3 ± 11.5 | 0.161 ± 0.006 | 767 ± 55 |

### 3.3.3 Combination of N15D and Q142K into $VPR_{\Delta C}$/6x

The final product of this part of the project was the combined mutant $VPR_{\Delta C}$/6x/N15D/Q142K. The resulting data obtained from stability measurements revealed the same trend as observed before, that is both mutations seem to increase the stability of the native structure as this mutant is ~1°C more stable than $VPR_{\Delta C}$/6x/N15D further establishing some role of Q142K in terms of stability of this mutant.

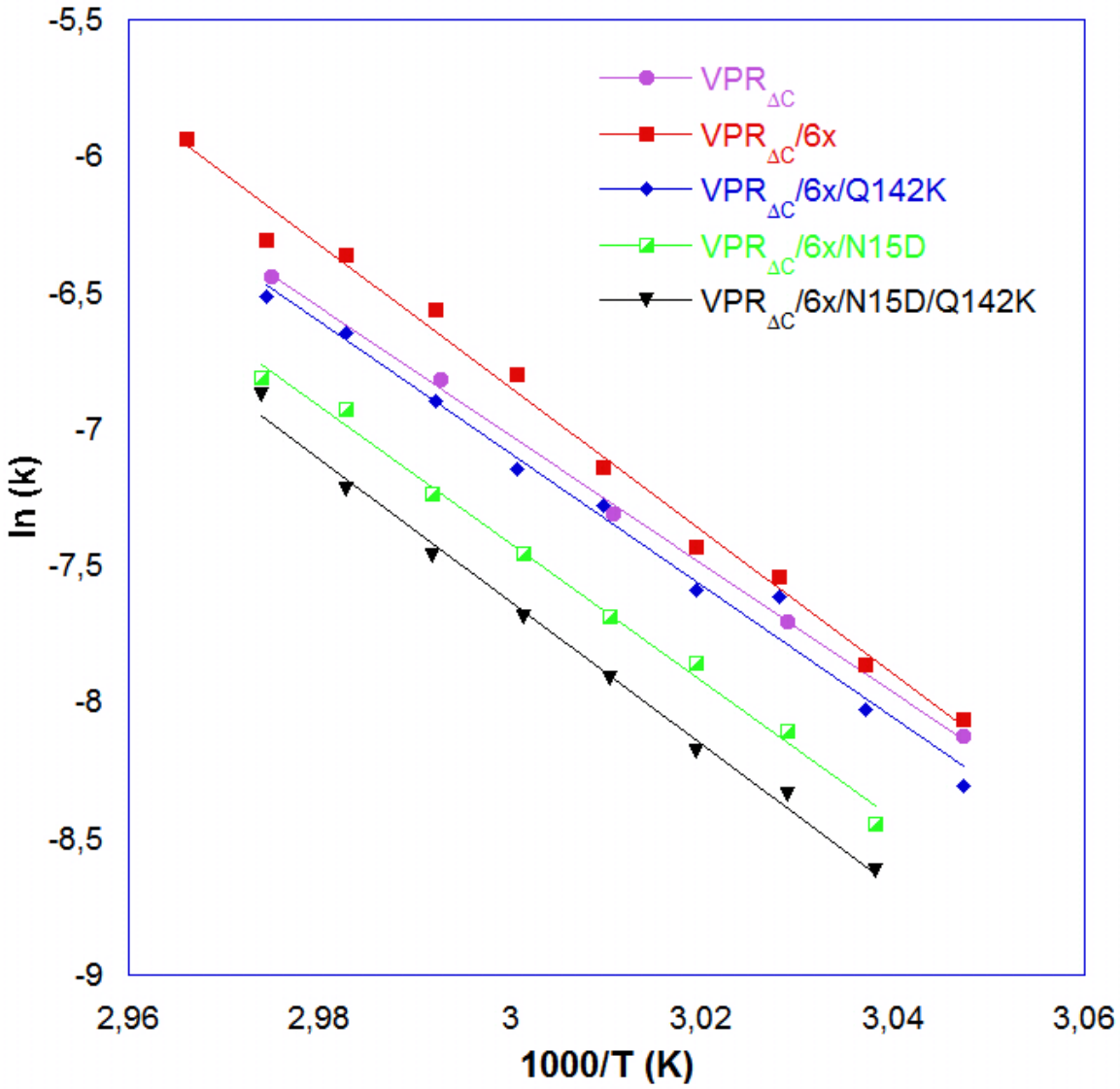


*Figure 25. Arrhenius plots the thermal inactivation of $VPR_{\Delta C}$ (purple circles), $VPR_{\Delta C}$/6x (red squares), $VPR_{\Delta C}$/6x/N15D (green half solid squares), $VPR_{\Delta C}$/6c/Q142K (blue diamonds) and $VPR_{\Delta C}$/6x/N15D/Q142K (black triangles). The rate of thermal inactivation determined by withdrawing aliquots at selected time intervals and measuring the remaining activity against the substrate sAAPF-pNA. The X-axis shows 1000/T (in Kelvin) and the Y-axis shows the natural logarithmic value of the rate constant k (s-1).6x is a short for the A116T/Q117R/A119H/S120R/G121R/S123A mutation.*

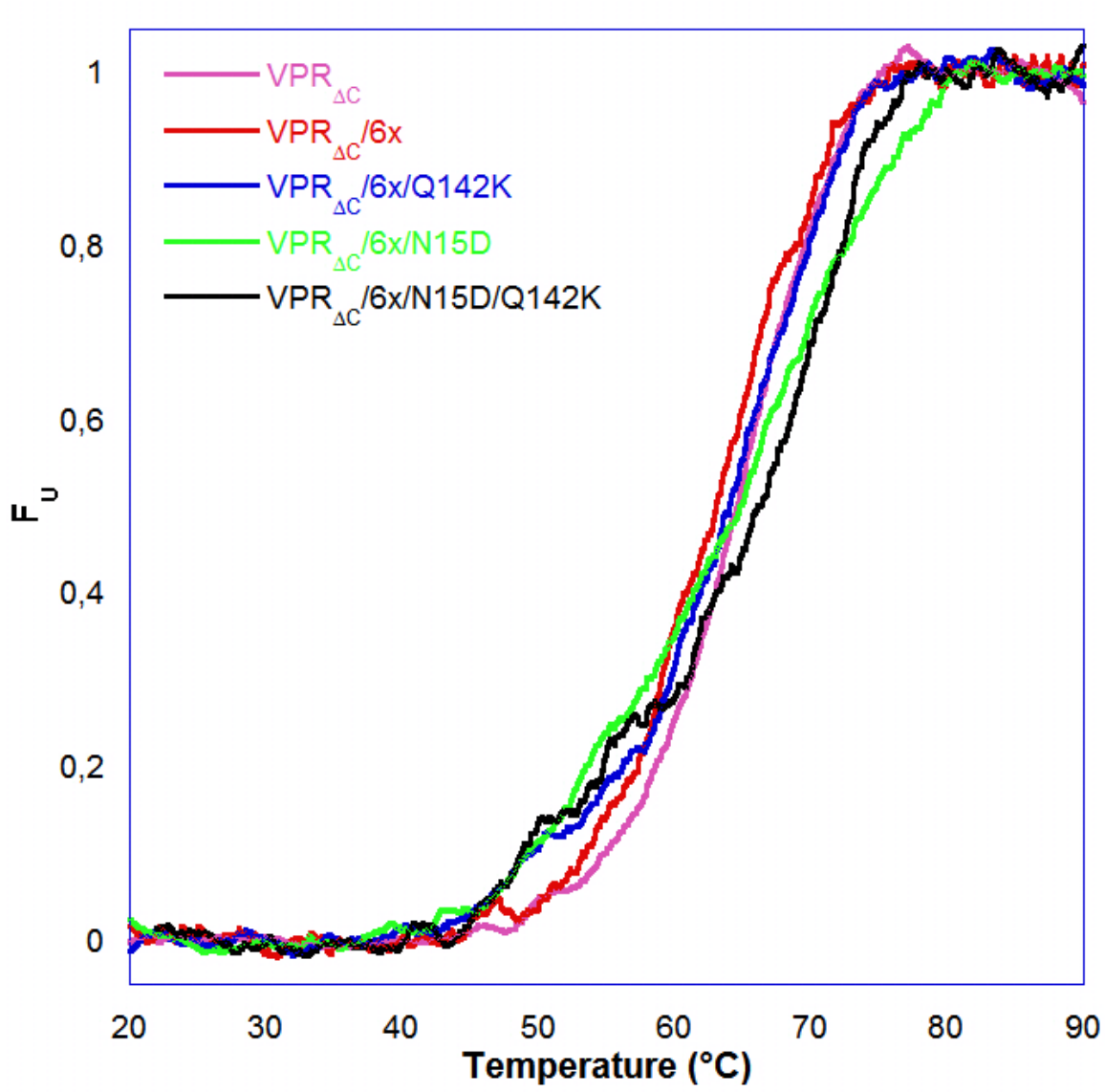


*Figure 26. Normalized melting curves of VPR$_{\Delta C}$ (purple), VPR$_{\Delta C}$/6x (red), VPR$_{\Delta C}$/6x/N15D (green), VPR$_{\Delta C}$/6x/Q142K (blue) and VPR$_{\Delta C}$/6x/N15D/Q142K (black as measured by changes in circular dichroism at 222 nm (CD)). The X-axis shows the temperature in Celsius degrees and the Y-axis shows the fraction of unfolded protein at each temperature (on the scale 0 to 1). 6x is a short for the A116T/Q117R/A119H/S120R/G121R/S123A mutation.*

The kinetic measurements did not however reveal expected trends, as the $k_{cat}$ value for this double mutant was similar to that of VPR$_{\Delta C}$/6x, but did not retain the high activity due to the Q142K mutation. In the light of these results these mutations seem to have non-additive effects when combined. What lies behind cannot be explained as of now and more research is needed. One cannot overlook certain trends however, i.e. that when N15D and

Q142K were combined on the $VPR_{\Delta C}$/N3P/I5P template the value for $k_{cat}$ is slightly lower, but the $K_m$ value has also lowered, which is similar to that when N15D and Q142K were combined on the $VPR_{\Delta C}$/6x template, although there the $k_{cat}$ value was lowered more drastically. This may indicate possible changes in the active site region when these mutations are combined also leading to lower values for $K_m$.

*Table 23. Thermal stability of $VPR_{\Delta C}$/6x, $VPR_{\Delta C}$/6x/N15D, $VPR_{\Delta C}$/6x/Q142K and $VPR_{\Delta C}$/6x/N15D/Q142K (n = 3, for $T_{T50\%}$ and n = 3, for $T_m$). Expressed as mean values ± standard deviation of the mean.*

| | **$T_m$ (°C)** | **$T_{50\%}$ (°C)** |
|---|---|---|
| **$VPR_{\Delta C}$/6x** | 62.8 ± 0.3 | 56.2 ± 0.2 |
| **$VPR_{\Delta C}$/6x/N15D** | 65.2 ± 0.4 | 58.2 ± 0.1 |
| **$VPR_{\Delta C}$/6x/Q142K** | 63.9 ± 0.2 | 56.8 ± 0.3 |
| **$VPR_{\Delta C}$/6x/N15D/Q142K** | 66.3 ± 0.2 | 59.3 ± 0.3 |

*Table 24. Kinetic parameters of $VPR_{\Delta C}$/6x, $VPR_{\Delta C}$/6x/N15D, $VPR_{\Delta C}$/6x/Q142K and $VPR_{\Delta C}$/6x/N15D/Q142K (n = 12). Expressed as mean values ± standard deviation of the mean.*

| | **$k_{cat}$ ($s^{-1}$)** | **$K_m$ (mM)** | **$k_{cat}/K_m$ ($s^{-1}mM^{-1}$)** |
|---|---|---|---|
| **$VPR_{\Delta C}$/6x** | 50.3 ± 1.3 | 0.171 ± 0.009 | 293 ± 10 |
| **$VPR_{\Delta C}$/6x/N15D** | 79.6 ± 1.7 | 0.164 ± 0.005 | 485 ± 9 |
| **$VPR_{\Delta C}$/6x/Q142K** | 123.3 ± 11.5 | 0.161 ± 0.006 | 767 ± 55 |
| **$VPR_{\Delta C}$/6x/N15D/Q142K** | 55.1 ± 2.4 | 0.151 ± 0.008 | 363 ± 13 |

# 3.4 Fluorescence

## 3.4.1 Fluorescence quenching of $VPR_{\Delta C}$, $VPR_{\Delta C}$/N3P/I5P and AQUI

It has been proposed that the molecular mechanisms of temperature adaptation involve adjustments of the molecular flexibility of proteins. The mutant $VPR_{\Delta C}$/N3P/I5P is of interest in this respect as its activity is much lower than that of the wild type enzyme, but it is also considerably more stable. To test the hypothesis that the changes in stability and activity of $VPR_{\Delta C}$/N3P/I5P compared to $VPR_{\Delta C}$ are due to increased rigidity of the native structure, fluorescence quenching measurements were carried out. For comparison the homologous enzyme AQUI was also used.

VPR contains four tryptophan residues: Trp6, Trp114, Trp191 and Trp208. AQUI on the other hand contains three Trp residues at three of the corresponding sites (Trp8, Trp114 and Trp208), but the residue at position 191 in AQUI is a tyrosine. For a better comparison a mutant of AQUI where Tyr191 was mutated to Trp (AQUI/Y191W) was also used in these experiments. At the excitation wavelength used in the study (285 nm) tryptophan absorbs most of the incoming light and as VPR only contains eight tyrosine residues (twelve for AQUI) Tyr emission should be miniscule compared to emission of the Trp residues in the spectra. Tyrosine is also often quenched in native proteins, which may be due to interactions with the peptide chain or energy transfer to tryptophan [107].

In quenching studies of these enzyme species a difference in the Stern-Volmer constants observed would therefore mostly reflect the accessibility to the Trp residues of the native structures, probably reflecting on the flexibility of the structure as well. Tryptophan residue at position 6 in VPR (Fig. 27) is most likely to be affected by the N3P/I5P mutation and any structural change in that area that might be observable via quenching. Trp6/Trp8 also show the most difference in solvent accessibility, Trp6 in VPR is highly solvent accessible whereas the corresponding tryptophan (Trp8) in AQUI is very poorly solvent accessible according to MD calculations [93]. Previous studies also showed that the quenching spectra of VPR and AQUI differs significantly over the temperature range 10-55°C thus indicating that AQUI has a more rigid structure over that temperature range [93].

Acrylamide is a well-known collisional quencher of indoles which is probably due to electron transfer from the indole to acrylamide, which does not occur in the ground state [107]. Quenching with acrylamide should therefore provide information on the accessibility changes of Trp residues in VPR.

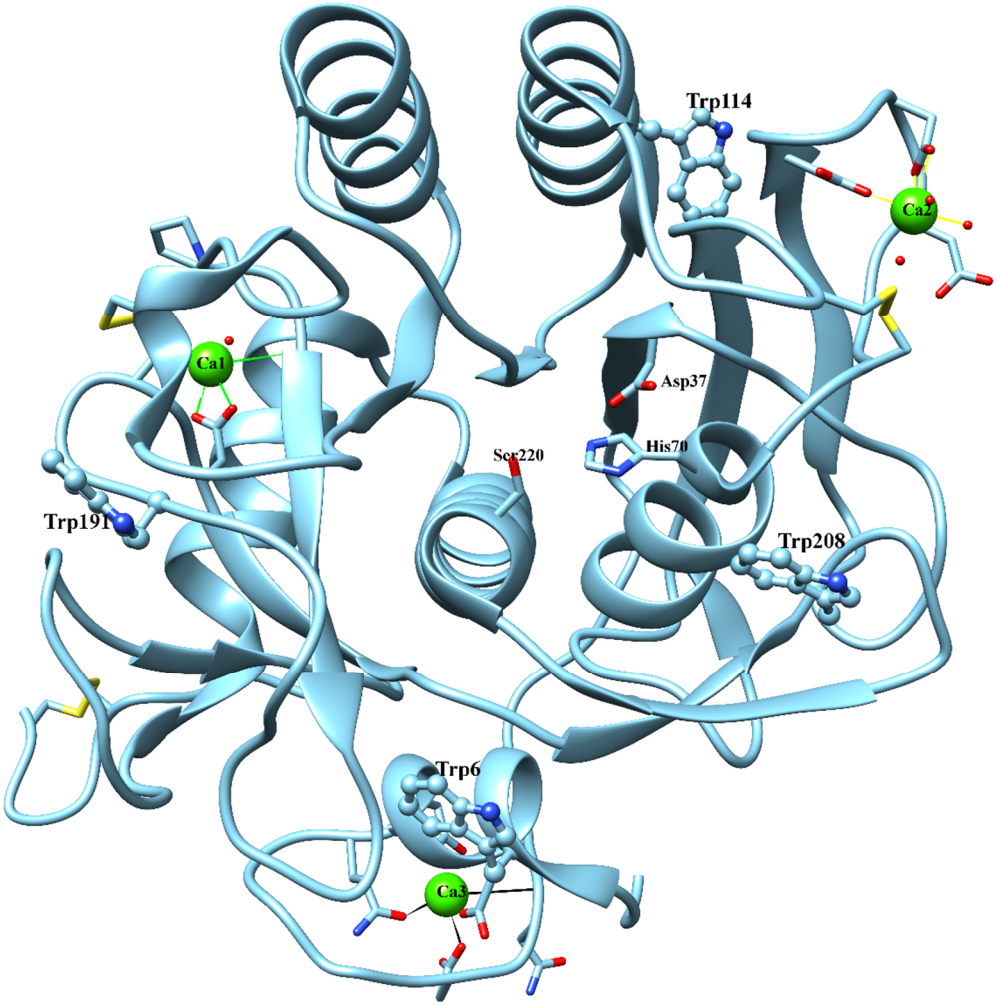


*Figure 27. The structure of VPR (PDB code: 1SH7) showing the location of the tryptophan residues within the structure of VPR expressed as balls and sticks (Trp 6, 114, 191 and 208), Residues of the active site are shown as sticks (Asp 37, His 70 and Ser220). Calcium ions are colored green and the atomic coloring code of the sticks and balls is as follows: red is O, blue is N and yellow is S.*

Results of the fluorescence studies indicated a dramatic change in quenching of $VPR_{\Delta C}$/N3P/I5P as compared to $VPR_{\Delta C}$, leading to a $K_{SV}$ constant that is much more similar to that of AQUI at both 25°C and 45°C (when observed at 345 nm).

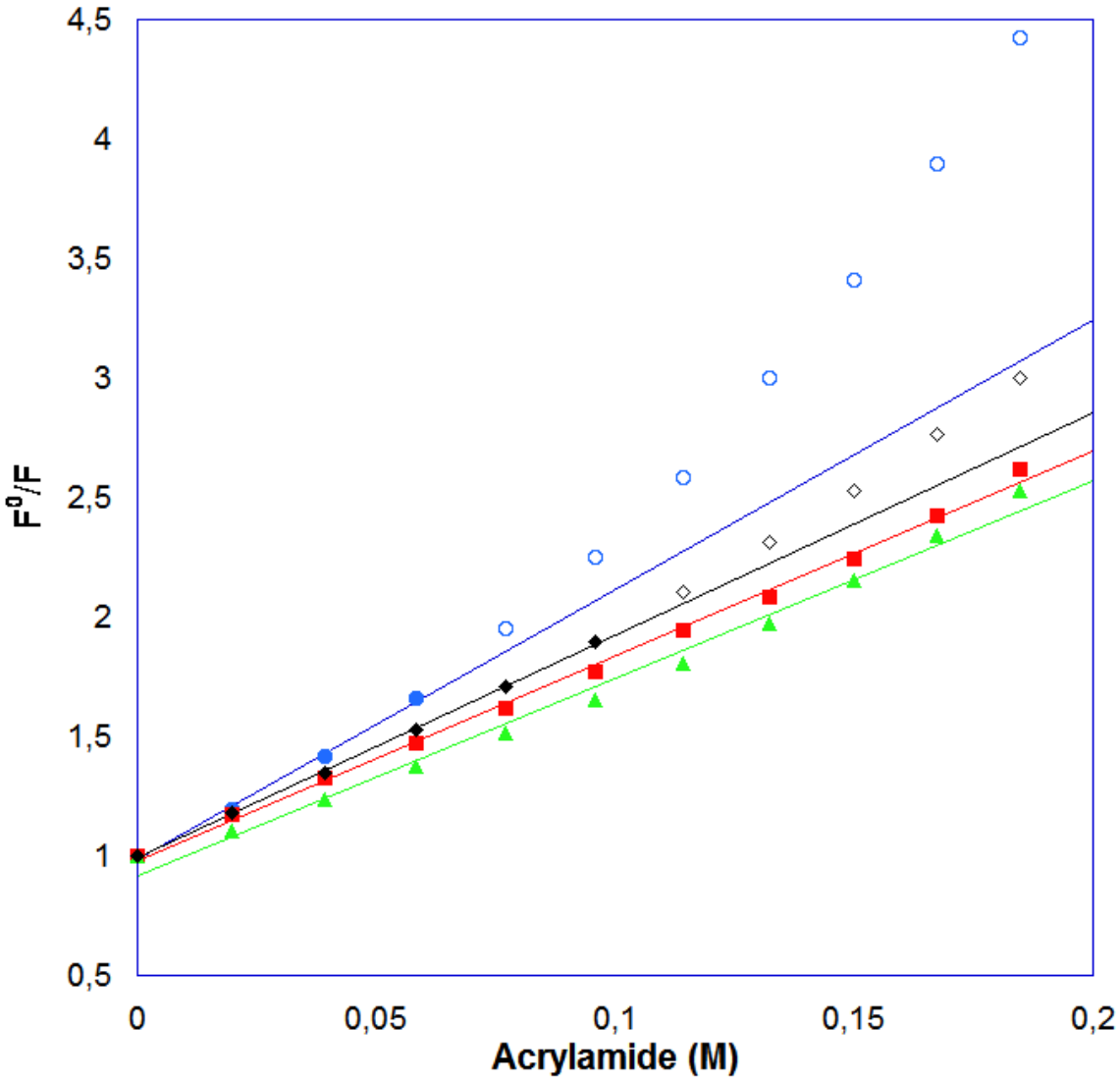


*Figure 28. Stern-Volmer plots of $VPR_{\Delta C}$ (blue circles), $VPR_{\Delta C}$/N3P/I5P (red squares), $AQUI_{wt}$ (green triangles) and AQUI/Y191W (black diamonds) at 345 nm and 25°C. Open symbols represent data points not included in $K_{sv}$ calculations.*

As can be seen in figure 28, accessibility of the fluorophores in $VPR_{\Delta C}$/N3P/I5P resembles much more that observed for AQUI when measured at 345 nm. Another notable fact is the upward curving character for $VPR_{\Delta C}$, which can only be explained by static quenching taking place. Static quenching arises from a complex formation between the fluorophore and the quencher, this complex being non-fluorescent [107]. This observation is not easy to interpret, but in the case of $VPR_{\Delta C}$/N3P/I5P and $VPR_{\Delta C}$ we can rule out that the effects observed may be attributable to quencher/protein ratio difference between the different samples, for that the difference is too great. One may speculate that this difference might

be due to much better accessibility of the quencher to Trp6 in $VPR_{\Delta C}$, so that acrylamide molecules can more easily get in such close proximity that a non-fluorescent complex is formed. In the case of AQUI/Y191W, little static quench seems to be taking place. Those results are however more likely to be due to quencher/protein ratio difference as there seems also to be a little upward character in the data points of AQUI.

The relationship between static quench and dynamic quench to the measured fluorescence is as stated in equation XXIII.

$$\frac{F^0}{F} = (1 + K_D[Q])(1 + K_S[Q]) \qquad \text{(XXIII)}$$

Here $K_D$ is the dynamic quenching constant ($K_D = K_{SV}$ in complete dynamic quenching as $K_S$ is 0 in such a case) and $K_S$ is the static quenching constant. The best way to distinguish between the two constants in data that shows upward character would be to take life time measurements while quenching, as dynamic quench affects the life time but static quench does not. But in this project a rough estimation was used, based on the fact that at low quencher concentrations the effects of $K_S$ approaches 0. Thus a linear fit through data points at low quencher concentrations was used to estimate the Stern-Volmer constants (Table 25).

*Table 25. Stern-Volmer constants at 345 nm for $VPR_{\Delta C}$ (n = 6 at 25°C and n = 5 at 45°C), $VPR_{\Delta C}$/N3P/I5P (n = 4 at 25°C and n = 4 at 45°C), $AQUI_{wt}$ (n = 3 at 25°C and n = 3 at 45°C) and AQUI/Y191W (n = 7 at 25°C and n = 3 at 45°C) Expressed as mean values ± standard deviation of the mean.*

| | **$K_{SV}$ ($M^{-1}$)** | |
|---|---|---|
| | 25°C | 45°C |
| **$VPR_{\Delta C}$** | 11.3 ± 0.6 | 17.6 ± 1.1 |
| **$VPR_{\Delta C}$/N3P/I5P** | 8.6 ± 0.2 | 10.4 ± 0.8 |
| **$AQUI_{wt}$** | 8.3 ± 1.0 | 10.1 ± 0.4 |
| **AQUI/Y191W** | 9.3 ± 0.1 | 10.5 ± 0.3 |

The difference in quenching at 45°C was even greater than for 25°C. This can be interpreted as at 45°C the structure of $VPR_{\Delta C}$ is much more sensitive to temperature than $VPR_{\Delta C}$/N3P/I5P and AQUI leading to the assumption that the proline mutations on the N-terminus make the structure more rigid or less accessible around the fluorophores.

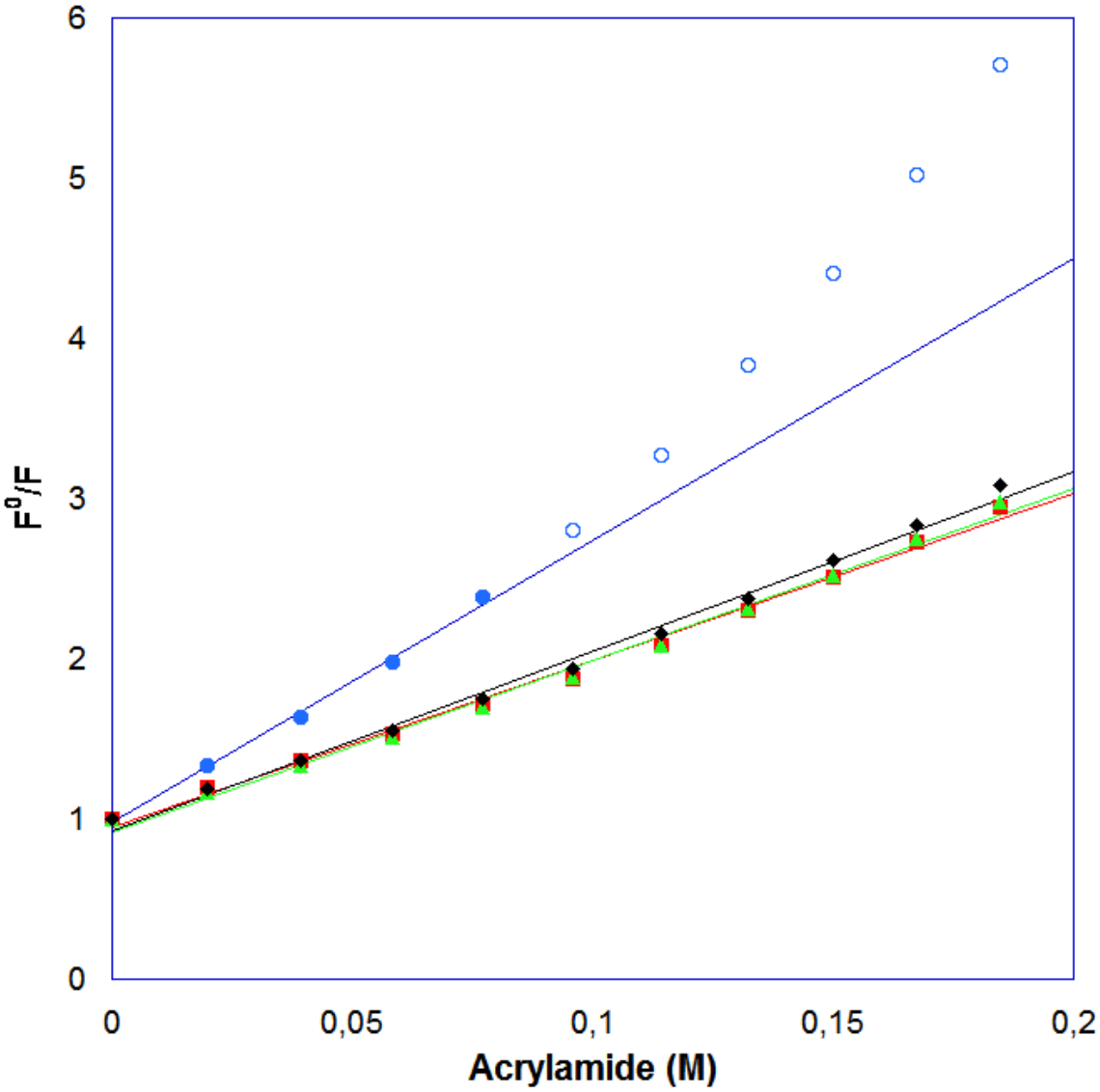


*Figure 29. Stern-Volmer plots of $VPR_{\Delta C}$ (blue circles), $VPR_{\Delta C}$/N3P/I5P (red squares), $AQUI_{wt}$ (green triangles) and AQUI/Y191W (black dimonds) at 345 nm and 45°C. Open symbols represent data points not included in $K_{sv}$ calculations.*

The data for 45°C establishes more evidence for static quench as at this temperature the static contribution to the overall quenching is much less than for 25°C. This is in good agreement with the theory of static quenching. According to that theory static quenching should decrease with temperature, as at higher temperatures the complexes causing the static quench are much less stable, whereas dynamic quenching increases with temperature as collisions between quencher and fluorophore are more frequent.

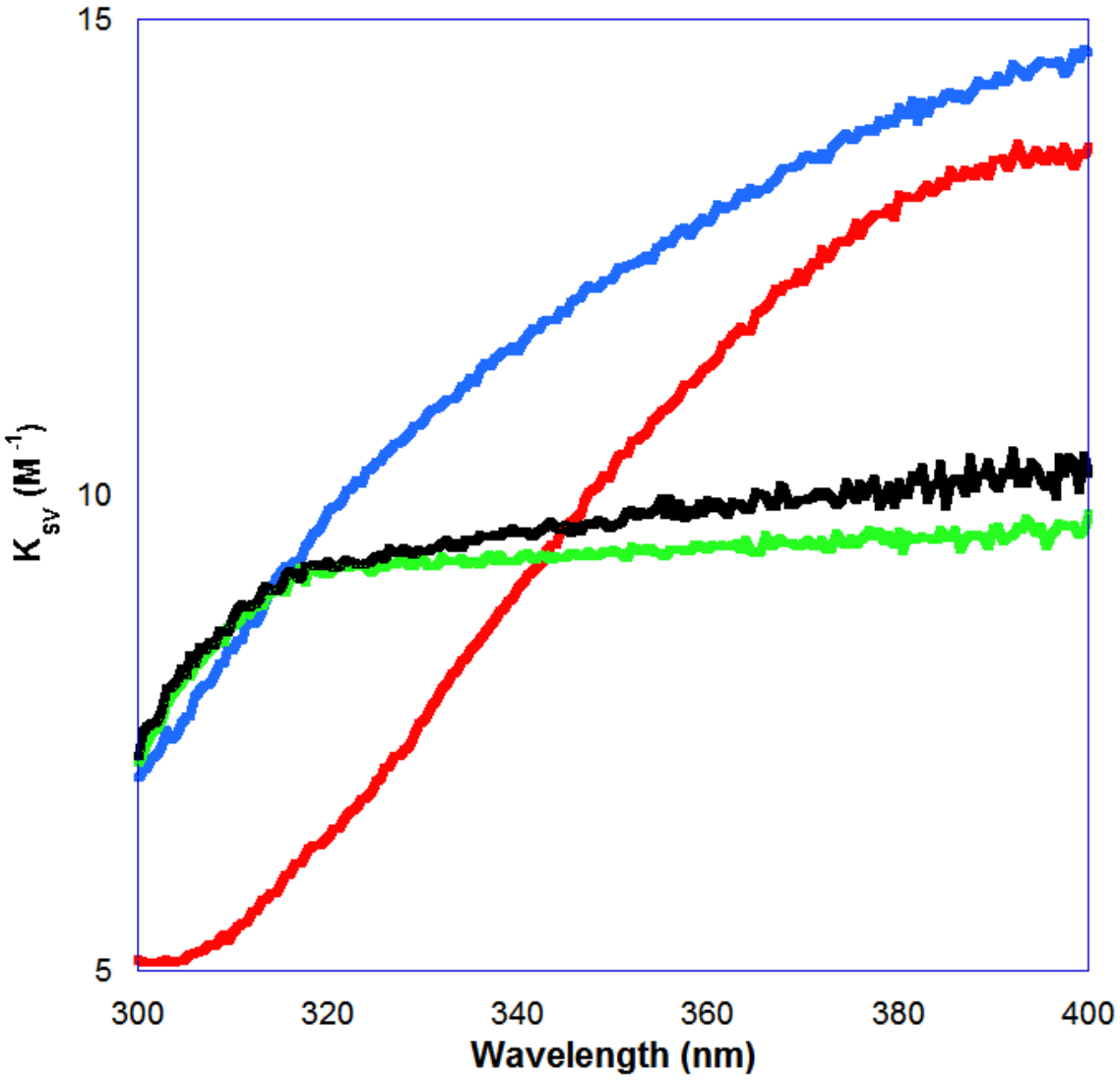


*Figure 30. Quenching-resolved emission spectra of $VPR_{\Delta C}$ (blue), $VPR_{\Delta C}$/N3P/I5P (red), $AQUI_{wt}$ (green) and AQUI/Y191W (Black) at 25°C. Represented as Stern-Volmer constants against wavelength.*

Emission spectra of proteins that contain more than one fluorophore often shift in the presence of quenchers, this is because the various Trp residues in the structure are differently exposed to the aqueous phase and are thus differently accessible to the quenchers. This leads to different Stern-Volmer constants at different wavelengths whereas a single fluorophore would have the same Stern-Volmer constant over the whole emission spectra. As seen in Figure 30 the change between $VPR_{\Delta C}$/N3P/I5P and $VPR_{\Delta C}$ is observed over all wavelengths and the curvature is also quite different as quenching at lower wavelengths has decreased more possibly indicating a burying of a fluorophore that leads

to lower quenching at lower wavelengths. For AQUI and AQUI/Y191W the difference in $K_{SV}$ over the emission spectra is much more subtle so tryptophan residue 191 in AQUI/Y191W does not seem to affect the emission much. Also notable in the figure 30 is that the resolved spectrum of AQUI is much simpler. A possible reason is that the emitting fluorophores in AQUI all share similar environments i.e. they all have similar Stern-Volmer constants in contrast to VPR where emitting fluorophores have vastly varying Stern-Volmer constants, probably due to more solvent accessibility to some of the tryptophan residues.

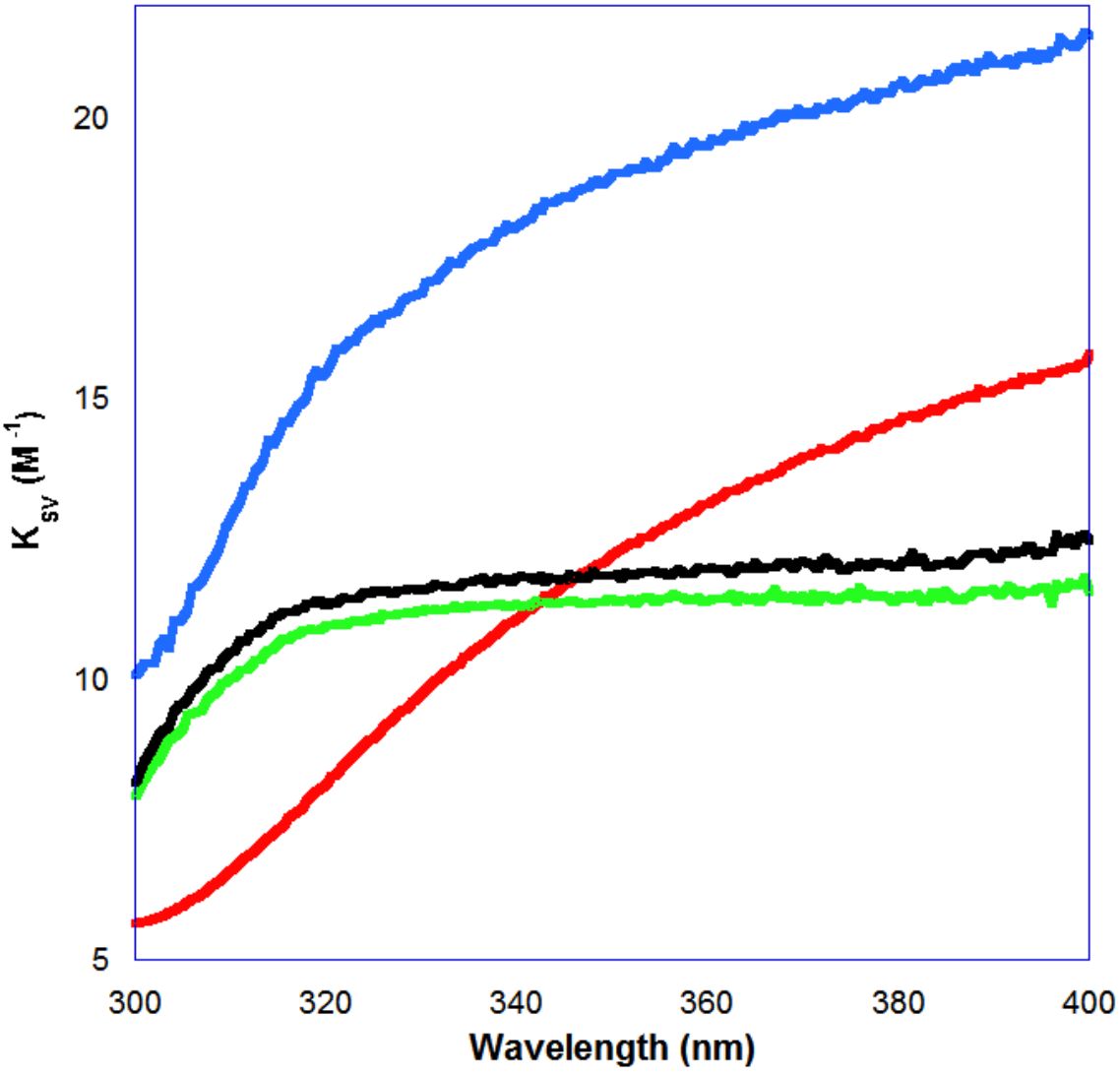


*Figure 31. Quenching-resolved emission spectra of $VPR_{\Delta C}$ (blue), $VPR_{\Delta C}$/N3P/I5P (red), $AQUI_{wt}$ (green) and AQUI/Y191W (Black) at 45°C. Represented as Stern-Volmer constants against wavelength.*

From figure 31 it is apparent that $VPR_{\Delta C}$ is much more sensitive to quenching than the other enzymes, which strengthens the argument that the proline mutations affect movements in $VPR_{\Delta C}$ making the structure more rigid and less accessible to quenchers. $VPR_{\Delta C}$/N3P/I5P clearly has a fluorophore that is more readily quenched at higher wavelengths meaning that the structure contains a fluorophore(s) that is more accessible than that in AQUI.

### 3.4.2 Fluorescence emission spectra of $VPR_{\Delta C}$, $VPR_{\Delta C}$/N3P/I5P and AQUI

Fluorescence spectra of proteins give information about the environment that the fluorophores are in within the protein structure. This has been demonstrated for example for free tryptophan in polar and apolar solvents. Apolar environment causes a blue shift in the emission spectra whereas polar solvents such as water red shifts them [107].

*Table 25. $\lambda_{max}$ values for the emission spectra of $VPR_{\Delta C}$, $VPR_{\Delta C}$/N3P/I5P, $AQUI_{wt}$ and AQUI/Y191W. Calculated from the emission spectra of the enzymes.*

| | **Native $\lambda_{max}$ (nm)** | **Unfolded $\lambda_{max}$ (nm)** |
|---|---|---|
| **$VPR_{\Delta C}$** | 340 | 356 |
| **$VPR_{\Delta C}$/N3P/I5P** | 338 | 357 |
| **$AQUI_{wt}$** | 330 | 341 |
| **AQUI/Y191W** | 332 | 355 |

The values in table 25 AQUI has much lower value for $\lambda_{max}$ suggesting that the environment of the fluorophores is much more apolar than in VPR possibly due to better water accessibility to Trp residues in VPR, or that neighboring residues are more polar than in AQUI. The difference in $\lambda_{max}$ for $VPR_{\Delta C}$ and $VPR_{\Delta C}$/N3P/I5P is barely measureable, but still is indicative of burying of a fluorophore in $VPR_{\Delta C}$/N3P/I5P.

A most interesting aspect of the unquenched fluorescence spectra of VPR and AQUI is the fact that VPR is much less fluorescent than AQUI. In fact the difference is so large that it cannot be explained by the differences in the polarity of the environment, nor the number of Tyr residues. Thus measurements were also taken of the denatured proteins to see if the native state of VPR has its Trp residues highly quenched by neighboring residues. These measurements revealed little difference between the unfolded states of AQUI/Y191W and VPR confirming intrinsic quenching in VPR (Fig. 32, 33, 34).

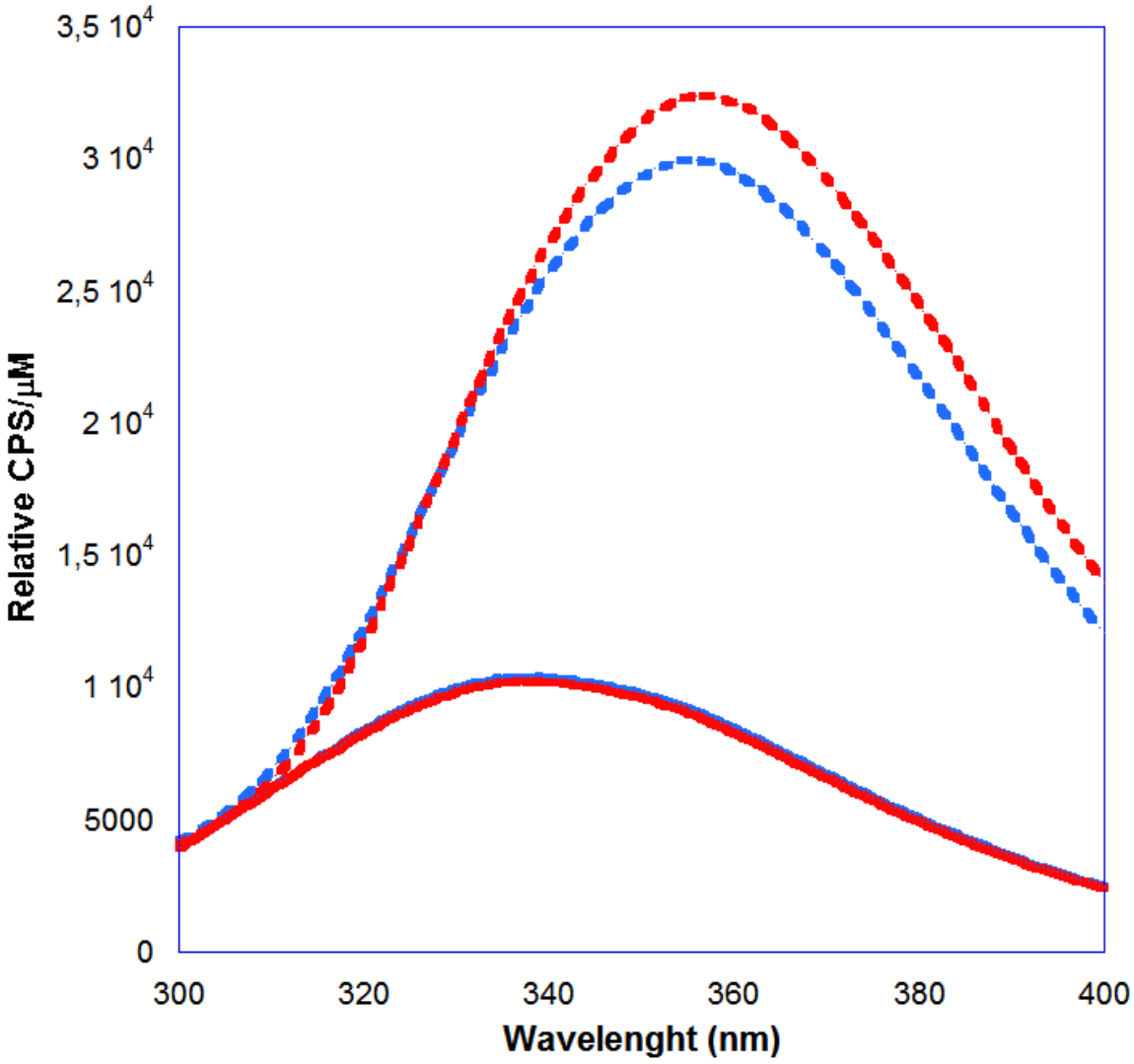


*Figure 32. Fluorescence emission spectra of VPR$_{\Delta C}$ (blue sold line for the native state and dotted line for the unfolded state) and VPR$_{\Delta C}$/N3P/I5P (red sold line for the native state and dotted line for the unfolded state).*

No statistical difference was observed between native spectra for VPR$_{\Delta C}$ and VPR$_{\Delta C}$/N3P/I5P and the small difference seen between the spectra for the unfolded states these enzymes is probably just due to error in concentration measurements (Fig.32).

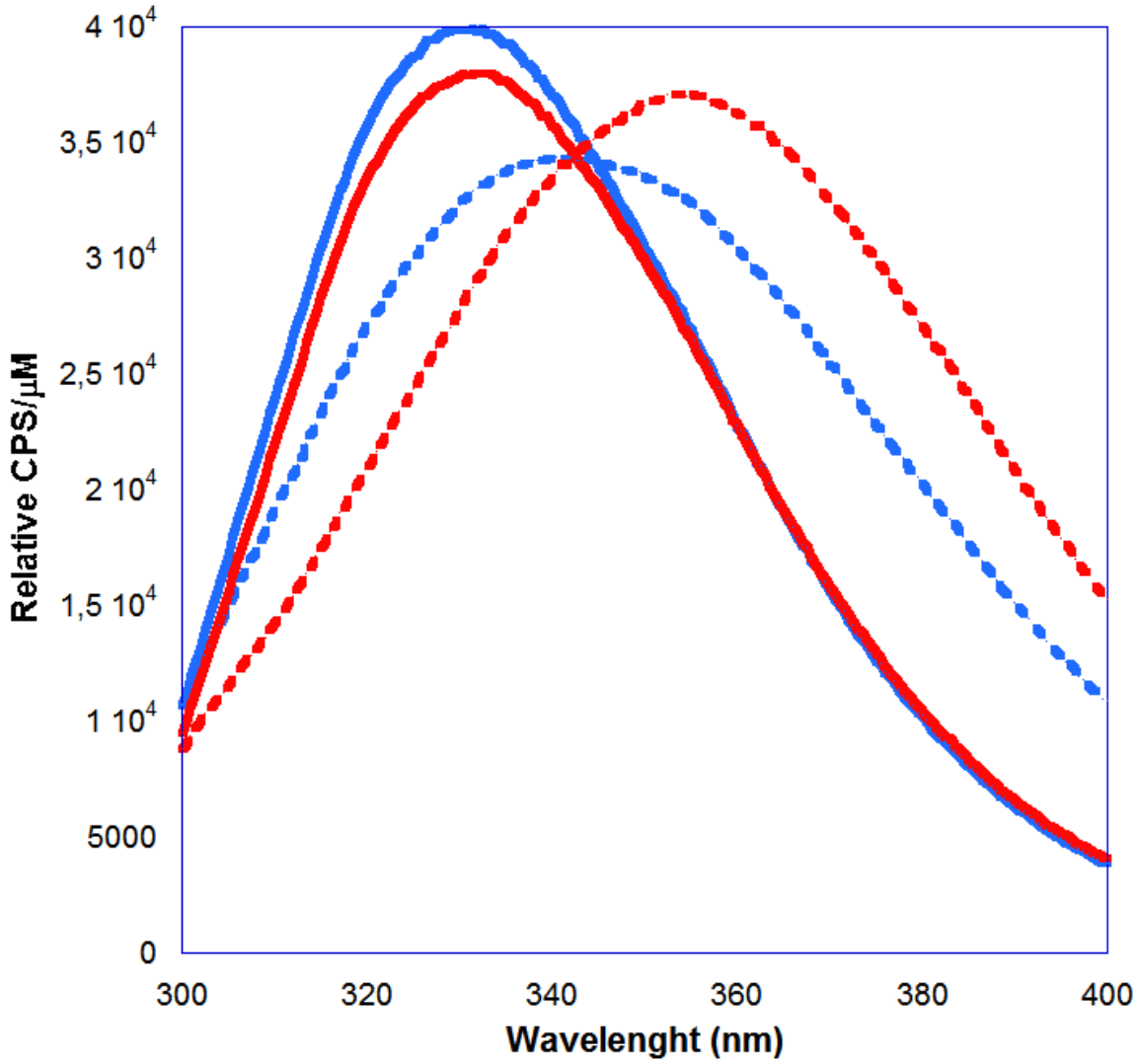


*Figure 33. Fluorescence emission spectra of AQUI$_{wt}$ (blue sold line for the native state and dotted line for the unfolded state) and AQUI/Y191W (red sold line for the native state and dotted line for the unfolded state).*

Differences in the emission spectra for the native states of AQUI$_{wt}$ and AQUI/Y191W are seemingly very small if any indicating that Trp 191 in the AQUI mutant is intrinsically quenched to a high degree. The spectrum for the unfolded state of AQUI/Y191W shows a rather extensive red shift and in fact becomes compatible with the spectrum for VPR, as might be expected (Fig. 32).

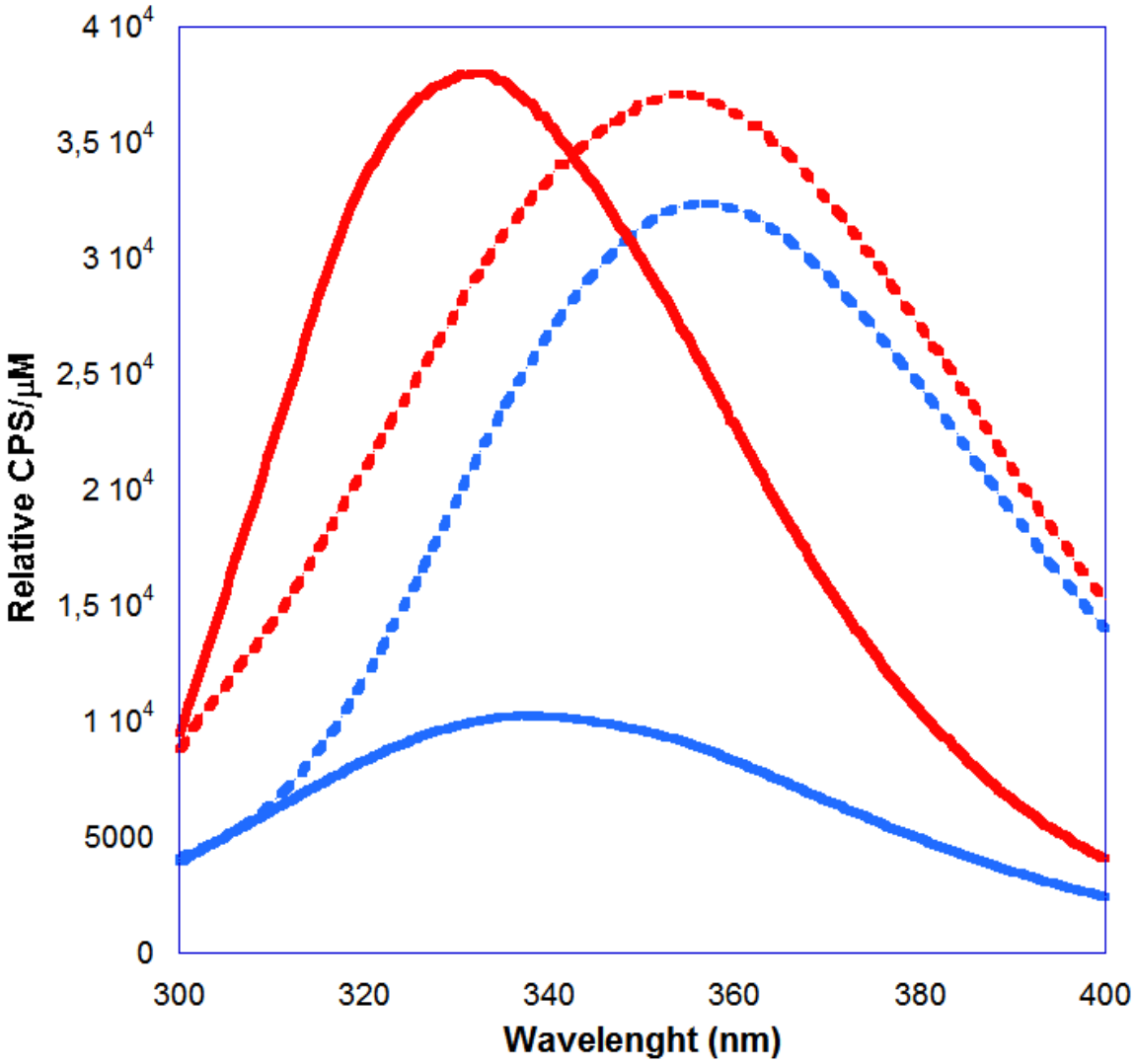


*Figure 34. Fluorescence emission spectra of VPR$_{\Delta C}$/N3P/I5P (blue sold line for the native state and dotted line for the unfolded state) and AQUI/Y191W (red sold line for the native state and dotted line for the unfolded state).*

It is quite clear from figure 34 that the emission spectra for AQUI/Y191W and VPR$_{\Delta C}$/N3P/I5P are very different for their native states, both in terms of $\lambda_{max}$ and intensity. The difference between the spectra for their unfolded states is much more subtle, e.g. having similar $\lambda_{max}$ values. The reason for the intensity difference might be due to extra tyrosine residues present in AQUI, which is supported by the fact that the difference is greater at lower wavelengths as tyrosine emits at lower wavelengths than tryptophan.

The fact that fluorescent emission of VPR is much higher for the unfolded state, which is rather rare for proteins but not unheard of, can only mean that fluorophores within VPR are

to a large extent quenched in the native state by neighboring residues. Besides being subjected to quenching by acrylamide and iodide, tryptophan fluorescence is also quenched by nearby disulfide groups as well as electron-deficient groups like $–NH_3^+$, also $–CO_2H$ and protonated histidine residues, are known quenchers [107].

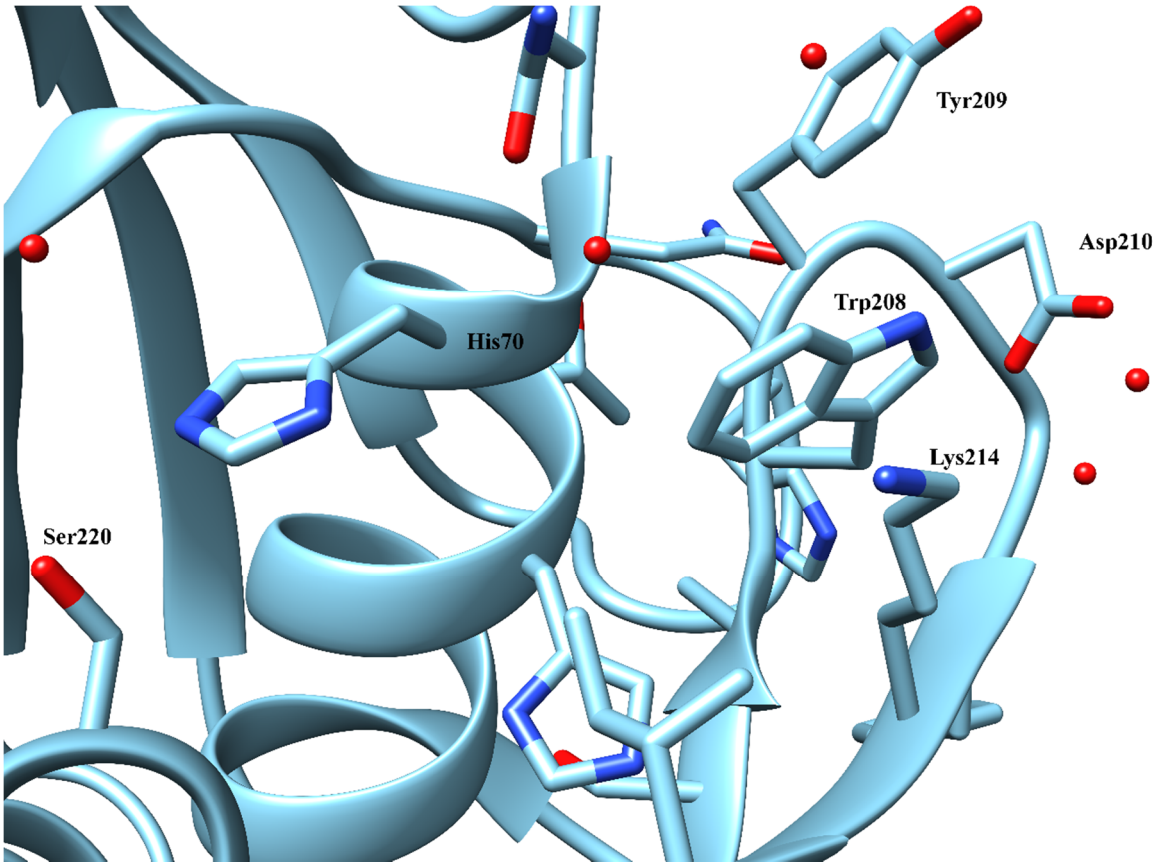


*Figure 35. Neighboring residues of tryptophan 208 in the crystal structure of VPR (PDB code: 1SH7) residue numbering according to VPR. The atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

If the crystal structure of VPR is examined closely it can be observed that most of the Trp residues in the structure are surrounded by groups that are known intrinsic quenchers. In close vicinity to Trp208 (Fig. 35) are residues like Lys214 and Asp210 (4 Å), both which could readily quench fluorescence of that Trp residue. In AQUI neither of these residues are present but Gln 216 (according to numbering in AQUI) is found in aprox 4.5 Å distance from the corresponding Trp residue in AQUI. This suggest that Trp208 is probably much more quenched in VPR than AQUI.

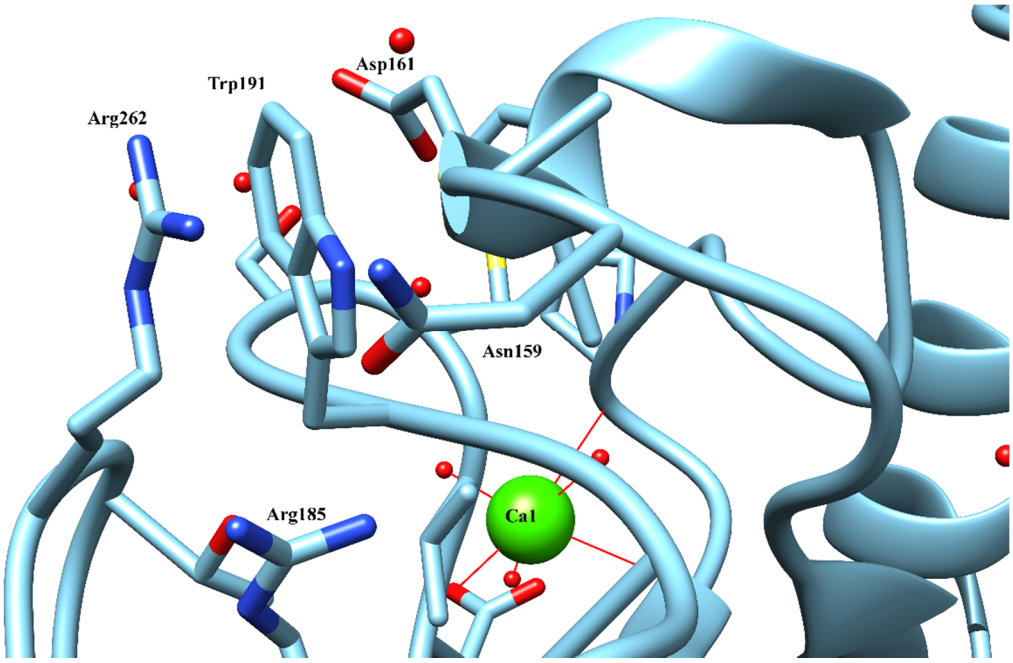


*Figure 36. Neighboring residues of tryptophan 191 in the crystal structure of VPR (PDB code: 1SH7) Calcium numbering according to VPR as well as residue numbering. The atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

Tryptophan 191 is most likely highly quenched in the structure of VPR, being surrounded by residues that are candidates for quenching. Arg262 being 3.5 Å away from Trp191, but Asp161 located 4 Å away and Asn159 within 4.5 – 5 Å distance. In AQUI/Y191W Asp161, Asn159 and Arg185 are present and as evident from the emission spectra of AQUI and AQUI/Y191W, Trp191 seems to almost fully quenched (Figure 33).

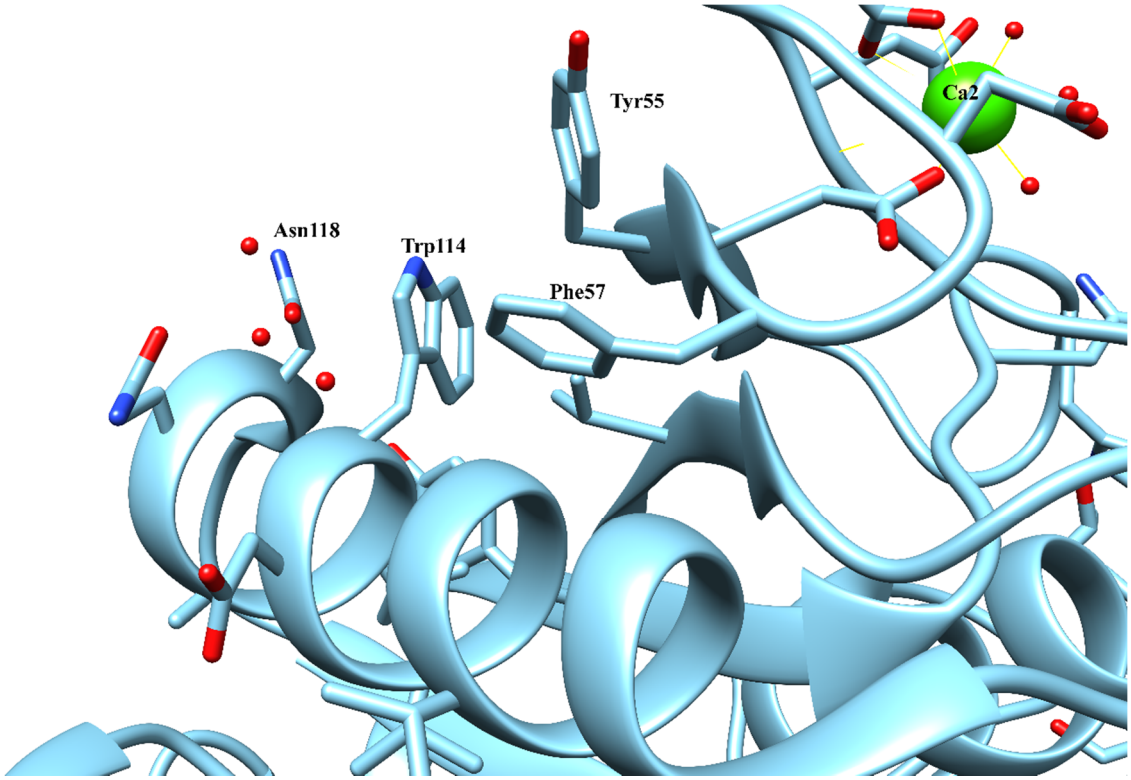


*Figure 37. Neighboring residues of tryptophan 114 in the crystal structure of VPR (PDB code: 1SH7) Calcium numbering according to VPR as well as residue numbering. The atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

The fluorescence of Trp114 in VPR might be quenched by the side group of Asn118 which is located at a distance of 4 Å away. But as is apparent from the quenching-resolved emission spectra of VPR (Figs. 30 and 31) there are more than one emitting fluorophore and compared to Trp208 and Trp191, Trp114 is much more likely to contribute to the fluorescence of VPR. The neighboring residues of Trp114 in AQUI are quite similar except that AQUI contains Arg in position 117, but it is over 5 Å away.

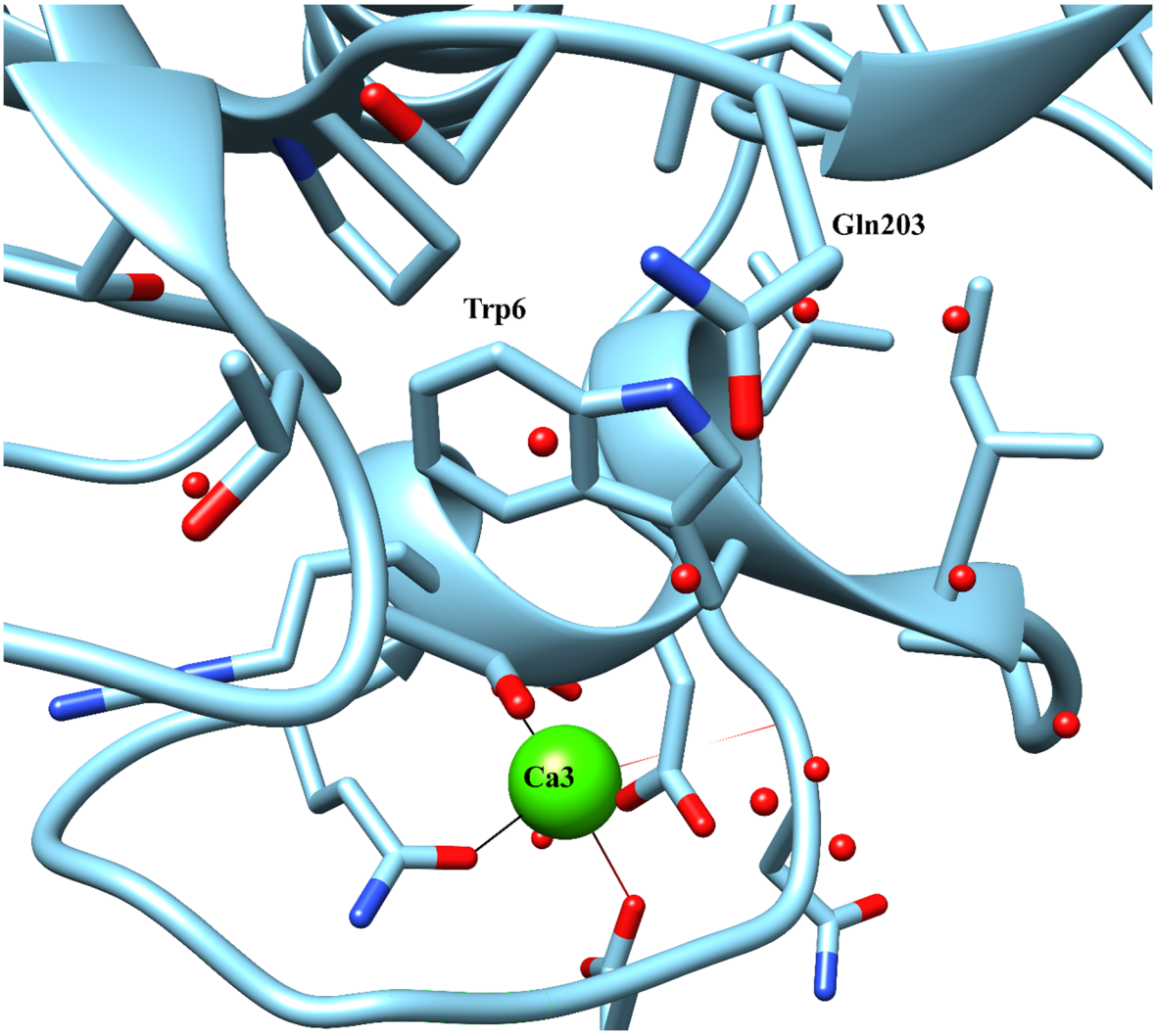


*Figure 38. Neighboring residues of tryptophan 6 in the crystal structure of VPR (PDB code: 1SH7) Calcium numbering according to VPR as well as residue numbering. The atomic coloring code of the sticks is as follows: red is O, blue is N and yellow is S.*

Trp6 in VPR is probably the best candidate for an emitting fluorophore in the native state of VPR. The only residue in the vicinity that might cause some partial quenching is Gln203 but the distance between side groups of Trp6 and Gln203 is around 6 Å in the crystal structure so if there is any quenching is occurring between these groups it is probably minimal, compared to that of Trp208 and Trp191. In AQUI the corresponding tryptophan residue is Trp8. In the vicinity there are no good candidates to quench the fluorescence of that residue so Trp8 is probably highly fluorescent. So the crystal structure of VPR gives strong evidence to intrinsic quenching effects that can explain this difference between AQUI and VPR. These results suggest then, that the quenching data is mostly reflecting on the flexibility of the N-termini in the enzymes.

# 4 Conclusions

## 4.1 Mutational studies

In the case of $VPR_{\Delta C}$/N3P/I5P the project can be deemed as a success as the mutant $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K is both more stable and active than the wild type VPR. Q142K has extensive effects on that template and increases the activity almost tenfold, probably due to interactions to a nearby salt bridge on helix D (D138-R169). The mutation most likely induces flexibility at this site which may be transmitted at least to areas around the active site as this mutation affects the stability to a very low extent. Introducing a salt bridge via the mutation N15D appears to be successful in every case as stability always increased with this mutation. However N15D does also affect the activity of $VPR_{\Delta C}$/N3P/I5P increasing it almost fivefold. This is interesting as N15D mutation on $VPR_{\Delta C}$ does not affect the activity at all, but had, however, been shown to affect activity in wild type VPR, which was attributed to interactions to the C-terminus [12]. These results suggest therefore that there are some other interactions at work, although these interactions remain unknown. Further supporting this are the effects observed when the mutations N15D and Q142K are combined, they have cumulative effects on $K_m$ i.e. lowering it as well as probably lowering the $k_{cat}$ value compared with $VPR_{\Delta C}$/N3P/I5P/Q142K. Although the difference in $k_{cat}$ is barely observable, results from mutational experiments with $VPR_{\Delta C}$/6x as a template suggest the same. Also accompanied with this slight change in activity are effects that may contribute to stability. Although the effects are almost negligible, similar effects are observed when these mutations are incorporated into $VPR_{\Delta C}$/6x, but how these mutations are interacting with the structure of VPR is completely unknown.

Stability and catalytic properties of $VPR_{\Delta C}$/6x is different from VPR/6x, being less active and less thermostable. Why this is so, is unclear and clearly more experiments are required to determine what reasons lie behind that. But $VPR_{\Delta C}$/6x shows lowering of $K_m$ compared with $VPR_{\Delta C}$ which is the same as has been observed between VPR and VPR/6x. As in $VPR_{\Delta C}$/N3P/I5P/N15D, $VPR_{\Delta C}$/6x/N15D seems to increase $k_{cat}$ giving another example where N15D increases activity without the C-terminus being present. Comparing thermal

stability of $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/N15D it is safe to say that the salt bridge is formed in this case as well. $VPR_{\Delta C}$/6x/Q142K increases the catalytic efficiency of the mutant compared with $VPR_{\Delta C}$/6x and this is in good agreement with results that have been obtained for $VPR_{\Delta C}$ and $VPR_{\Delta C}$/Q142K. However Q142K mutation also increases the stability of the $VPR_{\Delta C}$/6x template. There is no clear reason for this increase in stability of $VPR_{\Delta C}$/6x/Q142K as compared to $VPR_{\Delta C}$/6x, but the hinge area where the 6x mutations are located connects helix C to β-sheet 4 and β-sheet 4 is connected to α-helix D where Q142K is located. These helices also lie in close proximity to each other so presumably mutations affecting movements in that area might well influence each other, leading to complex dynamic changes in the enzyme. The stability of $VPR_{\Delta C}$/6x/N15D/Q142K also shows this trend, whereas that mutant is more stable than $VPR_{\Delta C}$/6x/N15D. But the difference in thermal stability of $VPR_{\Delta C}$/6x/N15D and $VPR_{\Delta C}$/6x/N15D/Q142K seems to be a little greater than between $VPR_{\Delta C}$/6x and $VPR_{\Delta C}$/6x/N15D, this is interestingly similar to measured values of $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K. Another aspect of $VPR_{\Delta C}$/6x/N15D/Q124K are effects observed on $k_{cat}$ and $K_m$ i.e. lowering of $K_m$ and $k_{cat}$ which is similar to what was observed for $VPR_{\Delta C}$/N3P/I5P/N15D/Q142K but with regards to $k_{cat}$ the effects are much more drastic and lowering the $k_{cat}$ value to that similar of $VPR_{\Delta C}$/6x.

*Table 26. An overview of kinetic constants and results from thermal stability measurements of mutants measured and produced in this study and the mutants the project is based on.*

| | $T_m$ (°C) | $T_{50\%}$ (°C) | $k_{cat}$ ($s^{-1}$) | $K_M$ (mM) | $k_{cat}/K_M$ ($s^{-1}mM^{-1}$) |
|---|---|---|---|---|---|
| **$VPR_{wt}$** | 63.6 ± 0.3 | 56.0 ± 0.2 | 74.6 ± 5.0 | 0.166 ± 0.017 | 449 ± 30 |
| **$VPR_{\Delta C}$** | 65.2 ± 0.2 | 56.4 ± 0.1 | 68.2 ± 9.9 | 0.184 ± 0.017 | 371 ± 26 |
| **VPR/6x** | 63.7 ± 0.1 | 55.1 ± 0.6 | 98.9 ± 13.4 | 0.148 ± 0.010 | 728 ± 64 |
| **$VPR_{\Delta C}$/Q142K** | 64.6 ± 0.2 | 56.3 ± 0.3 | 147.0 ± 7.7 | 0.157 ± 0.014 | 945 ± 51 |
| **$VPR_{\Delta C}$/N15D** | 68.2 ± 0.2 | 59.6 ± 0.2 | 64.9 ± 4.0 | 0.179 ± 0.009 | 365 ± 7 |
| **$VPR_{\Delta C}$/N3P/I5P** | 69.3 ± 0.2 | 61.9 ±0.2 | 8.8 ± 1.3 | 0.224 ± 0.010 | 40 ±7 |
| **$VPR_{\Delta C}$/N3P/I5P/Q142K*** | 68.9 ± 0.3 | 61.8 ± 0.2 | 94.1 ± 6.6 | 0.223±0.014 | 422 ± 24 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D*** | 70.4 ± 0.1 | 65.5 ± 0.1 | 51.2 ± 4.3 | 0.207±0.019 | 245 ± 19 |
| **$VPR_{\Delta C}$/N3P/I5P/N15D/Q142K*** | 71.1 ± 0.2 | 65.6 ± 0.2 | 86.3 ± 6.4 | 0.188 ± 0.015 | 461 ± 20 |
| **$VPR_{\Delta C}$/6x*** | 62.8 ± 0.3 | 56.2 ± 0.2 | 50.3 ± 1.3 | 0.171 ± 0.009 | 293 ± 10 |
| **$VPR_{\Delta C}$/6x/Q142K*** | 63.9 ± 0.2 | 56.8 ± 0.3 | 123.3 ± 11.5 | 0.161 ± 0.006 | 767 ± 55 |
| **$VPR_{\Delta C}$/6x/N15D*** | 65.2 ± 0.4 | 58.2 ± 0.1 | 79.6 ± 1.7 | 0.164 ± 0.005 | 485 ± 9 |
| **$VPR_{\Delta C}$/6x/N15D/Q142K*** | 66.3 ± 0.2 | 59.3 ± 0.3 | 55.1 ± 2.4 | 0.151 ± 0.008 | 363 ± 13 |

*Mutants measured and produced in this project.

A lot of speculations can be made from these results but any definite answers are far away as effects on flexibility of the enzyme are not understood nor is the role of global and local flexibility fully understood. But MD simulations are ongoing for the $VPR_{\Delta C}$/Q142K mutant that hopefully will shed some light on the increased activity of mutants containing that mutation.

Also changes in $K_m$ cannot be explained fully as it is not known if the mutations that are causing changes in $K_m$ are affecting enthalpy or entropy of the substrate binding. This calls for the design of ITC (isothermal titration calorimeter) experiment where VPR and interesting mutants are measured against a substrate analog inhibitor, this would provide extra information on changes in the binding site and possibly determine if the change is in enthalpy or entropy.

## 4.2 Fluorescence studies

The results from fluorescence studies strongly indicate extensive changes in flexibility of $VPR_{\Delta C}$/N3P/I5P, probably mostly around the N-terminus, but the data could also be indicating that the accessibility of all emitting fluorophores within the structure are not as readily quenched as in $VPR_{\Delta C}$, as without mutagenesis studies it is hard to say which fluorophores are emitting. The fluorescence quenching properties of $VPR_{\Delta C}$/N3P/I5P are also much less affected by higher temperatures than $VPR_{\Delta C}$ and behaves more similarly to AQUI in this respect, indicating that its structure is much more rigid than for $VPR_{\Delta C}$.

The fact that the fluorescence of VPR is highly intrinsically quenched raises some questions, such as exactly what residues are quenched, even though the crystal structure can give good estimations of this, mutational studies are needed to confirm those speculations. Also as amino-groups are in the vicinity of so many of the Trp residues and probably quenching the fluorescence, one wonders if amino-aromatic/amino-cation interactions play an important role in stabilization of VPR. This work also provides a good basis for further work for the examination of molecular movements in VPR and the interpretation of data acquired from further experimentation.

# 5 References


1. Lamazares, E., Clemente, I., Bueno, M., Velazquez-Campoy, A. & Sancho, J. (2015) Rational stabilization of complex proteins: a divide and combine approach, *Sci Rep.* **5**, 9129.

2. Branden, C. & Tooze, J. (1992) *Introduction to protein structure. Second Edition*, Garland Publishing, Inc., New York.

3. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W. & Katayama, D. S. (2010) Stability of protein pharmaceuticals: an update, *Pharm Res.* **27**, 544-75.

4. Dill, K. A. & MacCallum, J. L. (2012) The protein-folding problem, 50 years on, *Science.* **338**, 1042-6.

5. Sancho, J. (2013) The stability of 2-state, 3-state and more-state proteins from simple spectroscopic techniques... plus the structure of the equilibrium intermediates at the same time, *Arch Biochem Biophys.* **531**, 4-13.

6. Sohl, J. L., Jaswal, S. S. & Agard, D. A. (1998) Unfolded conformations of alpha-lytic protease are more stable than its native state, *Nature.* **395**, 817-9.

7. Sanchez-Ruiz, J. M. (2010) Protein kinetic stability, *Biophys Chem.* **148**, 1-15.

8. Dalby, P. A. (2011) Strategy and success for the directed evolution of enzymes, *Curr Opin Struct Biol.* **21**, 473-80.

9. Traxlmayr, M. W. & Obinger, C. (2012) Directed evolution of proteins for increased stability and expression using yeast display, *Arch Biochem Biophys.* **526**, 174-80.

10. Borgo, B. & Havranek, J. J. (2012) Automated selection of stabilizing mutations in designed and natural proteins, *Proc Natl Acad Sci U S A.* **109**, 1494-9.

11. Wijma, H. J., Floor, R. J., Jekel, P. A., Baker, D., Marrink, S. J. & Janssen, D. B. (2014) Computationally designed libraries for rapid enzyme stabilization, *Protein Eng Des Sel.* **27**, 49-58.

12. Sigurdardottir, A. G., Arnorsdottir, J., Thorbjarnardottir, S. H., Eggertsson, G., Suhre, K. & Kristjansson, M. M. (2009) Characteristics of mutants designed to incorporate a new ion pair into the structure of a cold adapted subtilisin-like serine proteinase, *Biochim Biophys Acta.* **1794**, 512-8.

13. Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M. & Makhatadze, G. I. (1999) Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface, *Biochemistry.* **38**, 16419-23.

14. Arnorsdottir, J., Helgadottir, S., Thorbjarnardottir, S. H., Eggertsson, G. & Kristjansson, M. M. (2007) Effect of selected Ser/Ala and Xaa/Pro mutations on the stability and catalytic properties of a cold adapted subtilisin-like serine proteinase, *Biochim Biophys Acta.* **1774**, 749-55.

15. Serrano, L., Sancho, J., Hirshberg, M. & Fersht, A. R. (1992) Alpha-helix stability in proteins. I. Empirical correlations concerning substitution of side-chains at the N and C-caps and the replacement of alanine by glycine or serine at solvent-exposed surfaces, *J Mol Biol.* **227**, 544-59.

16. Perry, L. J. & Wetzel, R. (1984) Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation, *Science.* **226**, 555-7.

17. Grana-Montes, R., de Groot, N. S., Castillo, V., Sancho, J., Velazquez-Campoy, A. & Ventura, S. (2012) Contribution of disulfide bonds to stability, folding, and amyloid fibril formation: the PI3-SH3 domain case, *Antioxid Redox Signal.* **16**, 1-15.

18. Ishikawa, K., Nakamura, H., Morikawa, K. & Kanaya, S. (1993) Stabilization of Escherichia coli ribonuclease HI by cavity-filling mutations within a hydrophobic core, *Biochemistry.* **32**, 6171-8.

19. Ayuso-Tejedor, S., Abian, O. & Sancho, J. (2011) Underexposed polar residues and protein stabilization, *Protein Eng Des Sel.* **24**, 171-7.

20. Wigley, D. B., Clarke, A. R., Dunn, C. R., Barstow, D. A., Atkinson, T., Chia, W. N., Muirhead, H. & Holbrook, J. J. (1987) The engineering of a more thermally stable lactate dehydrogenase by reduction of the area of a water-accessible hydrophobic surface, *Biochim Biophys Acta.* **916**, 145-8.

21. Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding, *Proc Natl Acad Sci U S A.* **84**, 6663-7.

22. Irun, M. P., Maldonado, S. & Sancho, J. (2001) Stabilization of apoflavodoxin by replacing hydrogen-bonded charged Asp or Glu residues by the neutral isosteric Asn or Gln, *Protein Eng.* **14**, 173-81.

23. Nelson, D. L., Nelson, D. L., Lehninger, A. L. & Cox, M. M. (2008) *Lehninger principles of biochemistry*, W.H. Freeman, New York.

24. Clarke, A., Morris, G. J., Fonseca, F., Murray, B. J., Acton, E. & Price, H. C. (2013) A Low Temperature Limit for Life on Earth, *PLoS ONE.* **8**, e66207.

25. Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., Hirayama, H., Nakagawa, S., Nunoura, T. & Horikoshi, K. (2008) Cell proliferation at 122°C and isotopically heavy CH(4) production by a hyperthermophilic methanogen under high-pressure cultivation, *Proc Natl Acad Sci U S A.* **105**, 10949-10954.

26. Karshikoff, A. & Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads, *Trends Biochem Sci.* **26**, 550-6.

27. Travaglini-Allocatelli, C., Ivarsson, Y., Jemth, P. & Gianni, S. (2009) Folding and stability of globular proteins and implications for function, *Curr Opin Struct Biol.* **19**, 3-7.

28. Braselmann, E., Chaney, J. L. & Clark, P. L. (2013) Folding the proteome, *Trends Biochem Sci.* **38**, 337-344.

29. Luke, K. A., Higgins, C. L. & Wittung-Stafshede, P. (2007) Thermodynamic stability and folding of proteins from hyperthermophilic organisms, *FEBS J.* **274**, 4023-4033.

30. Siddiqui, K. S. & Cavicchioli, R. (2006) Cold-adapted enzymes, *Annu Rev Biochem.* **75**, 403-33.

31. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains, *Science.* **181**, 223-30.

32. Lumry, R. & Eyring, H. (1954) Conformation Changes of Proteins, *J Phys Chem.* **58**, 110-120.

33. Szilágyi, A. & Závodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey, *Structure.* **8**, 493-504.

34. Arnold, F. H., Wintrode, P. L., Miyazaki, K. & Gershenson, A. (2001) How enzymes adapt: lessons from directed evolution, *Trends Biochem Sci.* **26**, 100-106.

35. Arnorsdottir, J., Magnusdottir, M., Fridjonsson, O. H. & Kristjansson, M. M. (2011) The effect of deleting a putative salt bridge on the properties of the thermostable subtilisin-like proteinase, aqualysin I, *Protein Pept Lett.* **18**, 545-51.

36. Tomschy, A., Bohm, G. & Jaenicke, R. (1994) The effect of ion pairs on the thermal stability of D-glyceraldehyde 3-phosphate dehydrogenase from the hyperthermophilic bacterium Thermotoga maritima, *Protein Eng.* **7**, 1471-8.

37. Sali, D., Bycroft, M. & Fersht, A. R. (1991) Surface electrostatic interactions contribute little of stability of barnase, *J Mol Biol.* **220**, 779-88.

38. Sun, D. P., Sauer, U., Nicholson, H. & Matthews, B. W. (1991) Contributions of engineered surface salt bridges to the stability of T4 lysozyme determined by directed mutagenesis, *Biochemistry.* **30**, 7142-53.

39. Hendsch, Z. S. & Tidor, B. (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis, *Protein Sci.* **3**, 211-26.

40. Waldburger, C. D., Schildbach, J. F. & Sauer, R. T. (1995) Are buried salt bridges important for protein stability and conformational specificity?, *Nat Struct Biol.* **2**, 122-8.

41. Elcock, A. H. (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins, *J Mol Biol.* **284**, 489-502.

42. Carlow, D. C., Short, S. A. & Wolfenden, R. (1996) Role of glutamate-104 in generating a transition state analogue inhibitor at the active site of cytidine deaminase, *Biochemistry.* **35**, 948-54.

43. Makhatadze, G. I. & Privalov, P. L. (1995) Energetics of protein structure, *Adv Protein Chem.* **47**, 307-425.

44. Pace, C. N., Shirley, B. A., McNutt, M. & Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins, *FASEB J.* **10**, 75-83.

45. Vogt, G. & Argos, P. (1997) Protein thermal stability: hydrogen bonds or internal packing?, *Fold Des.* **2**, S40-6.

46. Závodszky, P., Kardos, J., Svingor, Á. & Petsko, G. A. (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins, *Proc Natl Acad Sci U S A.* **95**, 7406-7411.

47. Sauter, N. K., Mau, T., Rader, S. D. & Agard, D. A. (1998) Structure of alpha-lytic protease complexed with its pro region, *Nat Struct Biol.* **5**, 945-50.

48. Baker, D. & Agard, D. A. (1994) Kinetics versus thermodynamics in protein folding, *Biochemistry.* **33**, 7505-9.

49. Todgham, A., Hoaglund, E. & Hofmann, G. (2007) Is cold the new hot? Elevated ubiquitin-conjugated protein levels in tissues of Antarctic fish as evidence for cold-denaturation of proteins in vivo, *J Comp Physiol B.* **177**, 857-866.

50. D'Amico, S., Collins, T., Marx, J.-C., Feller, G. & Gerday, C. (2006) Psychrophilic microorganisms: challenges for life, *EMBO Reports.* **7**, 385-389.

51. D'Amico, S., Marx, J.-C., Gerday, C. & Feller, G. (2003) Activity-Stability Relationships in Extremophilic Enzymes, *J Biol Chem.* **278**, 7891-7896.

52. Russell, R. J., Gerike, U., Danson, M. J., Hough, D. W. & Taylor, G. L. (1998) Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium, *Structure.* **6**, 351-61.

53. Wallon, G., Lovett, S. T., Magyar, C., Svingor, A., Szilagyi, A., Zavodszky, P., Ringe, D. & Petsko, G. A. (1997) Sequence and homology model of 3-isopropylmalate dehydrogenase from the psychrotrophic bacterium Vibrio sp. I5 suggest reasons for thermal instability, *Protein Eng.* **10**, 665-72.

54. Gerike, U., Danson, M. J., Russell, N. J. & Hough, D. W. (1997) Sequencing and expression of the gene encoding a cold-active citrate synthase from an Antarctic bacterium, strain DS2-3R, *Eur J Biochem.* **248**, 49-57.

55. Davail, S., Feller, G., Narinx, E. & Gerday, C. (1994) Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the antarctic psychrophile Bacillus TA41, *J Biol Chem.* **269**, 17448-53.

56. Narinx, E., Davail, S., Feller, G. & Gerday, C. (1992) Nucleotide and derived amino acid sequence of the subtilisin from the antarctic psychrotroph Bacillus TA39, *Biochim Biophys Acta.* **1131**, 111-3.

57. Arnorsdottir, J., Kristjansson, M. M. & Ficner, R. (2005) Crystal structure of a subtilisin-like serine proteinase from a psychrotrophic Vibrio species reveals structural aspects of cold adaptation, *FEBS J.* **272**, 832-45.

58. Kumar, S. & Nussinov, R. (2004) Different roles of electrostatics in heat and in cold: adaptation by citrate synthase, *ChemBioChem.* **5**, 280-90.

59. Fields, P. A. (2001) Review: Protein function at thermal extremes: balancing stability and flexibility, *Comp Biochem Physiol A Mol Integr Physiol.* **129**, 417-31.

60. Asgeirsson, B., Nielsen, B. N. & Hojrup, P. (2003) Amino acid sequence of the cold-active alkaline phosphatase from Atlantic cod (Gadus morhua), *Comp Biochem Physiol B Biochem Mol Biol.* **136**, 45-60.

61. Thomas, T. & Cavicchioli, R. (1998) Archaeal cold-adapted proteins: structural and evolutionary analysis of the elongation factor 2 proteins from psychrophilic, mesophilic and thermophilic methanogens, *FEBS Lett.* **439**, 281-6.

62. Kumar, S., Ma, B., Tsai, C. J. & Nussinov, R. (2000) Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers, *Proteins.* **38**, 368-83.

63. Lonhienne, T., Gerday, C. & Feller, G. (2000) Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility, *Biochim Biophys Acta.* **1543**, 1-10.

64. Feller, G. (2003) Molecular adaptations to cold in psychrophilic enzymes, *Cell Mol Life Sci.* **60**, 648-62.

65. Feller, G. & Gerday, C. (1997) Psychrophilic enzymes: molecular basis of cold adaptation, *Cell Mol Life Sci.* **53**, 830-41.

66. Gerday, C., Aittaleb, M., Arpigny, J. L., Baise, E., Chessa, J. P., Garsoux, G., Petrescu, I. & Feller, G. (1997) Psychrophilic enzymes: a thermodynamic challenge, *Biochim Biophys Acta.* **1342**, 119-31.

67. Tsuruta, H. & Aizono, Y. (2003) Catalytic Efficiency and Some Structural Properties of Cold-Active Protein-Tyrosine-Phosphatase, *J Biochem.* **133**, 225-230.

68. Boulain, J.-C., Dassa, J., Mesta, L., Savatier, A., Costa, N., Muller, B. H., L'hostis, G., Stura, E. A., Troesch, A. & Ducancel, F. (2013) Mutants with higher stability and specific activity from a single thermosensitive variant of T7 RNA polymerase, *Protein Eng Des Sel.* **26**, 725-734.

69. Puente, X. S., Sanchez, L. M., Gutierrez-Fernandez, A., Velasco, G. & Lopez-Otin, C. (2005) A genomic view of the complexity of mammalian proteolytic systems, *Biochem Soc Trans.* **33**, 331-4.

70. Di Cera, E. (2009) Serine Proteases, *IUBMB life.* **61**, 510-515.

71. Page, M. J. & Di Cera, E. (2008) Serine peptidases: classification, structure and function, *Cell Mol Life Sci.* **65**, 1220-36.

72. Hedstrom, L. (2002) Serine Protease Mechanism and Specificity, *Chemical Reviews.* **102**, 4501-4524.

73. Perona, J. J. & Craik, C. S. (1995) Structural basis of substrate specificity in the serine proteases, *Protein Sci.* **4**, 337-60.

74. Neurath, H. (1985) Proteolytic enzymes, past and present, *Fed Proc.* **44**, 2907-13.

75. Siezen, R. J. & Leunissen, J. A. (1997) Subtilases: the superfamily of subtilisin-like serine proteases, *Protein Sci.* **6**, 501-23.

76. Rockwell, N. C. & Thorner, J. W. (2004) The kindest cuts of all: crystal structures of Kex2 and furin reveal secrets of precursor processing, *Trends Biochem Sci.* **29**, 80-7.

77. Seifert, U., Maranon, C., Shmueli, A., Desoutter, J. F., Wesoloski, L., Janek, K., Henklein, P., Diescher, S., Andrieu, M., de la Salle, H., Weinschenk, T., Schild, H.,

Laderach, D., Galy, A., Haas, G., Kloetzel, P. M., Reiss, Y. & Hosmalin, A. (2003) An essential role for tripeptidyl peptidase in the generation of an MHC class I epitope, *Nat Immunol.* **4**, 375-9.

78. Abifadel, M., Varret, M., Rabes, J. P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D., Derre, A., Villeger, L., Farnier, M., Beucler, I., Bruckert, E., Chambaz, J., Chanu, B., Lecerf, J. M., Luc, G., Moulin, P., Weissenbach, J., Prat, A., Krempf, M., Junien, C., Seidah, N. G. & Boileau, C. (2003) Mutations in PCSK9 cause autosomal dominant hypercholesterolemia, *Nat Genet.* **34**, 154-6.

79. Kristjánsson, M. M. (2012) *Thermostable Subtilases (Subtilisin-Like Serine Proteases), In: Thermostable Proteins: Structural Stability and Design*, (Sen, S and Nilsson, L, eds.) CRC Press, Taylor and Francis, USA.

80. Kobayashi, H., Yoshida, T., Miyakawa, T., Tashiro, M., Okamoto, K., Yamanaka, H., Tanokura, M. & Tsuge, H. (2015) Structural Basis for Action of the External Chaperone for a Propeptide-deficient Serine Protease from Aeromonas sobria, *J Biol Chem.* **290**, 11130-43.

81. Jain, S. C., Shinde, U., Li, Y., Inouye, M. & Berman, H. M. (1998) The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 å resolution1, *J Mol Biol.* **284**, 137-144.

82. Takeuchi, Y., Satow, Y., Nakamura, K. T. & Mitsui, Y. (1991) Refined crystal structure of the complex of subtilisin BPN′ and Streptomyces subtilisin inhibitor at 1·8Åresolution, *J Mol Biol.* **221**, 309-325.

83. Tanaka, S., Matsumura, H., Koga, Y., Takano, K. & Kanaya, S. (2007) Four new crystal structures of Tk-subtilisin in unautoprocessed, autoprocessed and mature forms: insight into structural changes during maturation, *J Mol Biol.* **372**, 1055-69.

84. Tanaka, S.-i., Saito, K., Chon, H., Matsumura, H., Koga, Y., Takano, K. & Kanaya, S. (2007) Crystal Structure of Unautoprocessed Precursor of Subtilisin from a Hyperthermophilic Archaeon: Evidence for $Ca^{2+}$ induced folding, *J Biol Chem.* **282**, 8246-8255.

85. Foophow, T., Tanaka, S., Angkawidjaja, C., Koga, Y., Takano, K. & Kanaya, S. (2010) Crystal structure of a subtilisin homologue, Tk-SP, from Thermococcus kodakaraensis: requirement of a C-terminal beta-jelly roll domain for hyperstability, *J Mol Biol.* **400**, 865-77.

86. Ottmann, C., Rose, R., Huttenlocher, F., Cedzich, A., Hauske, P., Kaiser, M., Huber, R. & Schaller, A. (2009) Structural basis for $Ca^{2+}$-independence and activation by homodimerization of tomato subtilase 3, *Proc Natl Acad Sci U S A.* **106**, 17223-8.

87. Helland, R., Larsen, A. N., Smalas, A. O. & Willassen, N. P. (2006) The 1.8 A crystal structure of a proteinase K-like enzyme from a psychrotroph Serratia species, *FEBS J.* **273**, 61-71.

88. Singh, R. K., Gourinath, S., Sharma, S., Roy, I., Gupta, M. N., Betzel, C., Srinivasan, A. & Singh, T. P. (2001) Enhancement of enzyme activity through three-phase partitioning: crystal structure of a modified serine proteinase at 1.5 A resolution, *Protein Eng.* **14**, 307-13.

89. Jany, K. D. & Mayer, B. (1985) Proteinase K from Tritirachium album limber. I. Molecular mass and sequence around the active site serine residue, *Biol Chem Hoppe Seyler.* **366**, 485-92.

90. Muller, A. & Saenger, W. (1993) Studies on the inhibitory action of mercury upon proteinase K, *J Biol Chem.* **268**, 26150-4.

91. Ebeling, W., Hennrich, N., Klockow, M., Metz, H., Orth, H. D. & Lang, H. (1974) Proteinase K from Tritirachium album Limber, *Eur J Biochem.* **47**, 91-7.

92. Kristjansson, M. M., Magnusson, O. T., Gudmundsson, H. M., Alfredsson, G. A. & Matsuzawa, H. (1999) Properties of a subtilisin-like proteinase from a psychrotrophic Vibrio species comparison with proteinase K and aqualysin I, *Eur J Biochem.* **260**, 752-60.

93. Sigtryggsdóttir, Á. R., Papaleo, E., Thorbjarnardóttir, S. H. & Kristjánsson, M. M. (2014) Flexibility of cold- and heat-adapted subtilisin-like serine proteinases evaluated with fluorescence quenching and molecular dynamics, *Biochim Biophys Acta.* **1844**, 705-712.

94. Arnorsdottir, J., Smaradottir, R. B., Magnusson, O. T., Thorbjarnardottir, S. H., Eggertsson, G. & Kristjansson, M. M. (2002) Characterization of a cloned subtilisin-like serine proteinase from a psychrotrophic Vibrio species, *Eur J Biochem.* **269**, 5536-46.

95. Arnórsdóttir, J., Sigtryggsdóttir, Á. R., Thorbjarnardóttir, S. H. & Kristjánsson, M. M. (2009) Effect of proline substitutions on stability and kinetic properties of a Cold Adapted Subtilase, *J Biochem.* **145**, 325-329.

96. Matsuzawa, H., Tokugawa, K., Hamaoki, M., Mizoguchi, M., Taguchi, H., Terada, I., Kwon, S.-T. & Ohta, T. (1988) Purification and characterization of aqualysin I (a

thermophilic alkaline serine protease) produced by Thermus aquaticus YT-1, *Eur J Biochem.* **171**, 441-447.

97. Terada, I., Kwon, S. T., Miyata, Y., Matsuzawa, H. & Ohta, T. (1990) Unique precursor structure of an extracellular protease, aqualysin I, with NH2- and COOH-terminal pro-sequences and its processing in Escherichia coli, *J Biol Chem.* **265**, 6576-81.

98. Betzel, C., Pal, G. P. & Saenger, W. (1988) Three-dimensional structure of proteinase K at 0.15-nm resolution, *Eur J Biochem.* **178**, 155-171.

99. Jonsdottir, L. B., Ellertsson, B. O., Invernizzi, G., Magnusdottir, M., Thorbjarnardottir, S. H., Papaleo, E. & Kristjansson, M. M. (2014) The role of salt bridges on the temperature adaptation of aqualysin I, a thermostable subtilisin-like proteinase, *Biochim Biophys Acta.* **1844**, 2174-2181.

100. Sigtryggsdóttir, Á. R. (2010) *Hlutverk sameinda-sveigjanleika í hitastigsaðlögun subtilisín-líkra serín próteinasa*, Masters Theisis, Háskóli Íslands, Reykjavík.

101. Kristjánsson, M. M. Unpublished Work in

102. Sigurðardóttir, A. G. (2007) *Áhrif markvissra stökkbreytinga á hitastigsaðlögun VPR, subilisín-líks serín próteinasa úr kuldakærri Vibrio-tegund*, Masters Theisis Háskóli Íslands, Reykjavík.

103. Ellertsson, B. Ö. (2012) *Áhrif saltbrúa og yfirborðshleðslna á hitastigsaðlögun subtilísin-líkra serín próteinasa*, Masters Theisis, Háskóli Íslands, Reykjavík.

104. Inoue, H., Nojima, H. & Okayama, H. (1990) High efficiency transformation of Escherichia coli with plasmids, *Gene.* **96**, 23-28.

105. Zaman, Z. & Verwilghen, R. L. (1979) Quantitation of proteins solubilized in sodium dodecyl sulfate-mercaptoethanol-tris electrophoresis buffer, *Anal Biochem.* **100**, 64-69.

106. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci.* **4**, 2411-23.

107. Lakowicz, J. R. (2006) *Principles of Fluorescence Spectroscopy*, 3 edn, Springer Science, New York, USA.

# 6 Appendix 1

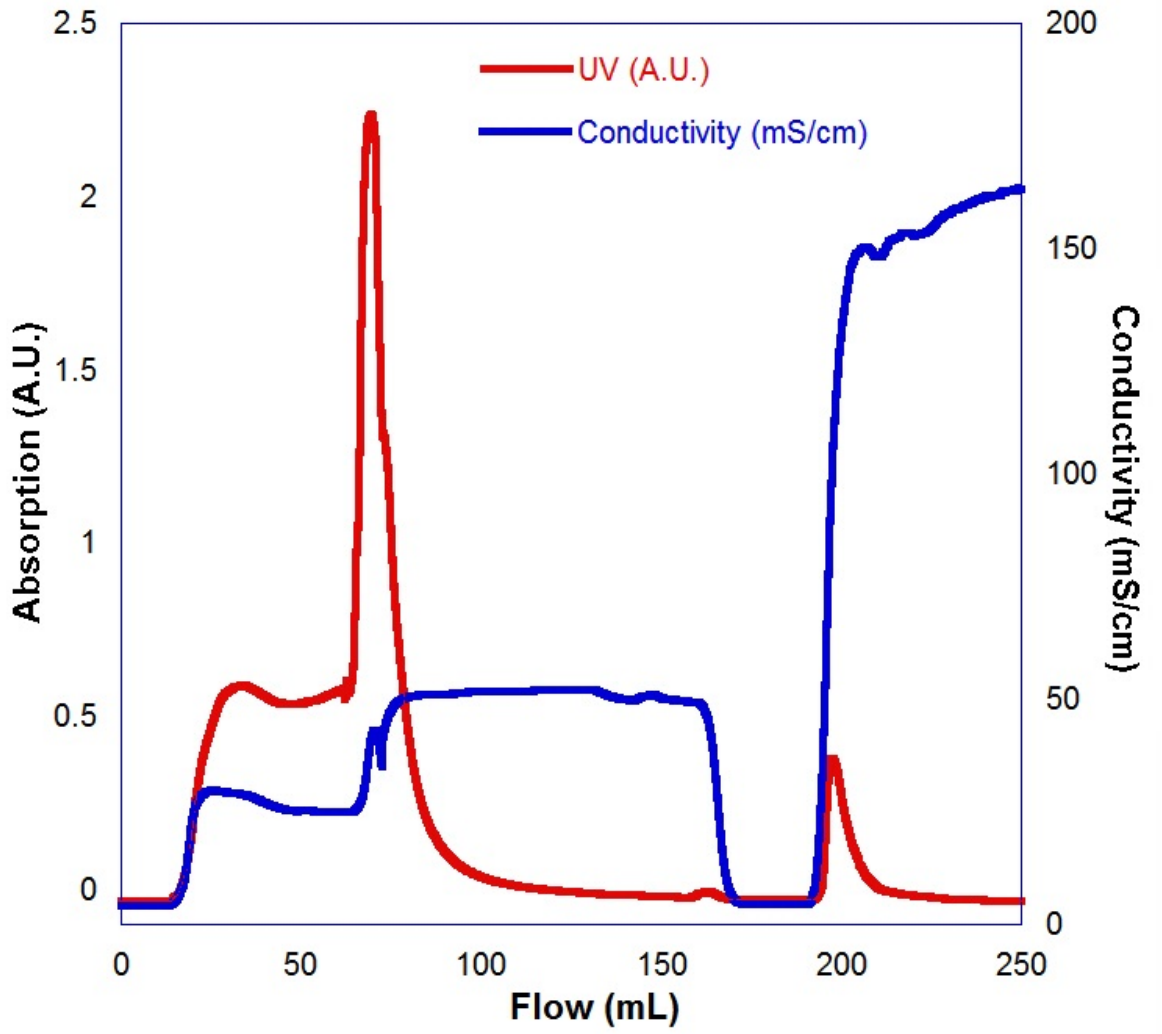


*Figure 39. Typical purification of VPR and its mutants on a Z-D-Phe-TETA column absorption shown in red (left Y-axis) and conductivity shown in blue (right Y-axis). See chapter 2.8.1 for details. Activity only in the last peak.*

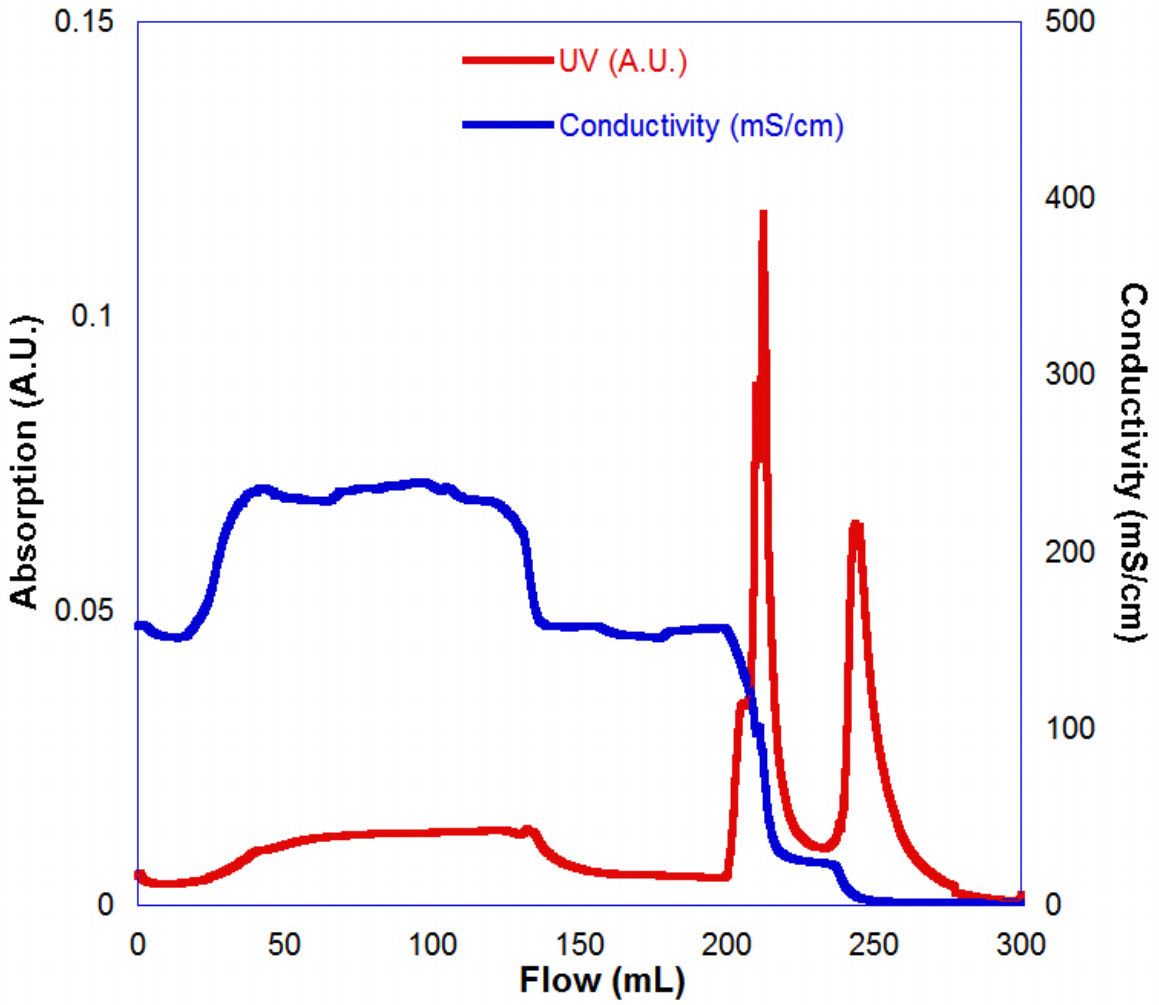


*Figure 40. Typical purification of VPR and its mutants on a phenyl Sepharose column absorption shown in red (left Y-axis) and conductivity shown in blue (right Y-axis). See chapter 2.8.1 for details. Considerable activity only in the last peak.*

*Table 27. Composition of buffers used in the project.*

| | |
|---|---|
| **Buffer A** | 25 mM Tris-Cl<br>10 mM $CaCl_2$<br>pH 8.0 at 25°C |
| **$T_{50\%}$ Buffer** | 25 mM Tris-Cl<br>100 mM NaCl<br>15 mM $CaCl_2$<br>1 mM EDTA<br>pH 8.95 at 25°C |
| **$T_m$ Buffer** | 25 mM glycine<br>100 mM NaCl<br>15 mM $CaCl_2$<br>pH 8.6 at 25°C |
| **Michaelis-Menten Buffer** | 100 mM Tris-Cl<br>10 mM $CaCl_2$<br>pH 8.6 at 25°C |
| **Fluorescence Buffer** | 50 mM Tris-Cl<br>10 mM $CaCl_2$<br>pH calibrated to 8.0 for the experimental temperature at hand |

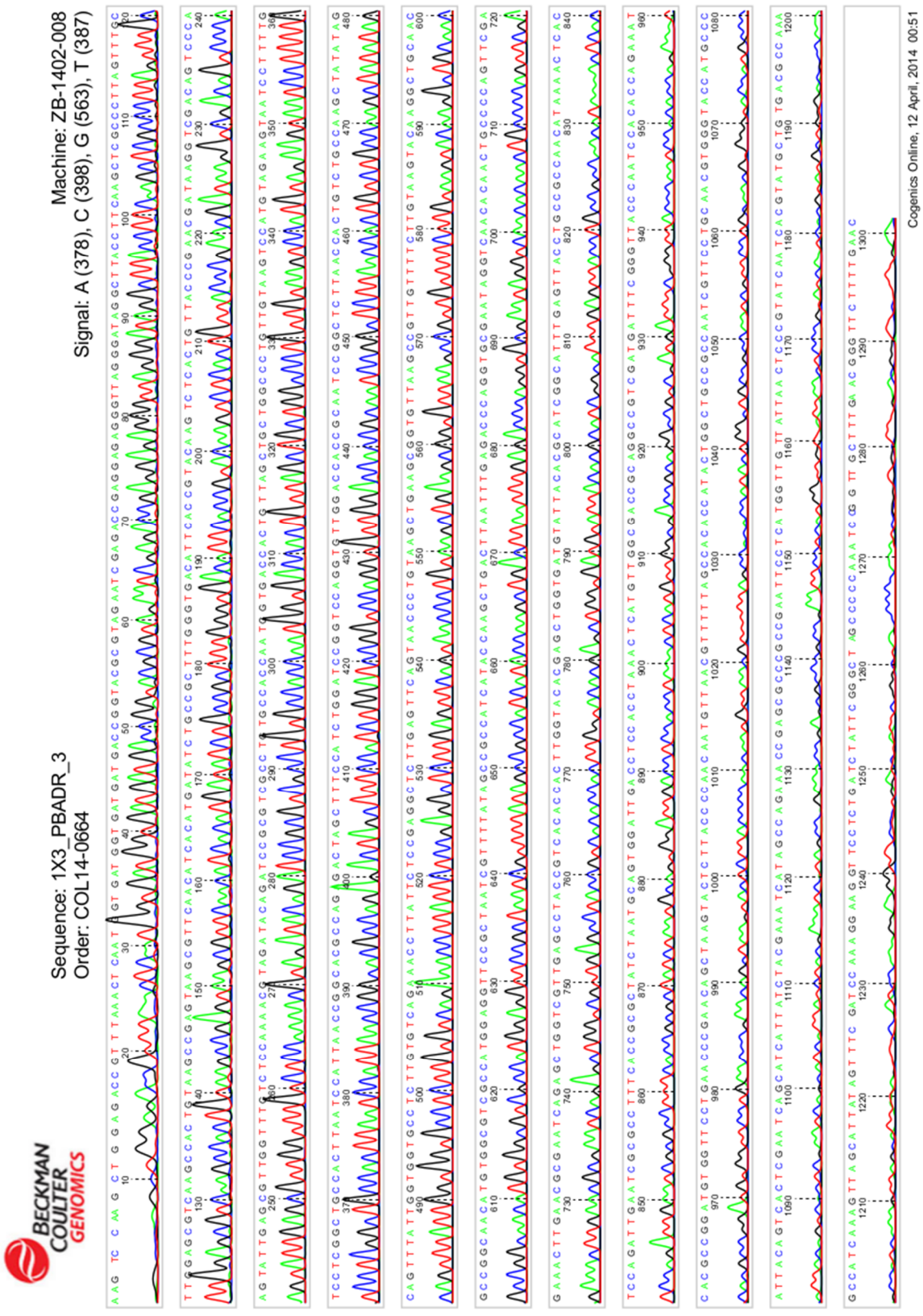


*Figure 41. Typical results from a reverse sequencing of a pBAD plasmid containing the VPR gene.*

*Figure 42. Typical results from a forward sequencing of a pBAD plasmid containing the VPR gene.*

# 7 Appendix 2


*Figure 43. Sequence alignment of the area around residue 142 (according to numbering in VPR, highlighted in grey). As seen in the alignment none of the most related proteases (see figure 45) have a negatively charged residue in the position corresponding to Q142 in VPR. Only proteases more related to AQUI have a negatively charged residue. An * (asterisk) indicates positions which have a single, fully conserved residue. A : (colon) indicates conservation between groups of strongly similar properties. A . (period) indicates conservation between groups of weakly similar properties. Organism of origin shown for all except VPR and AQUI.*

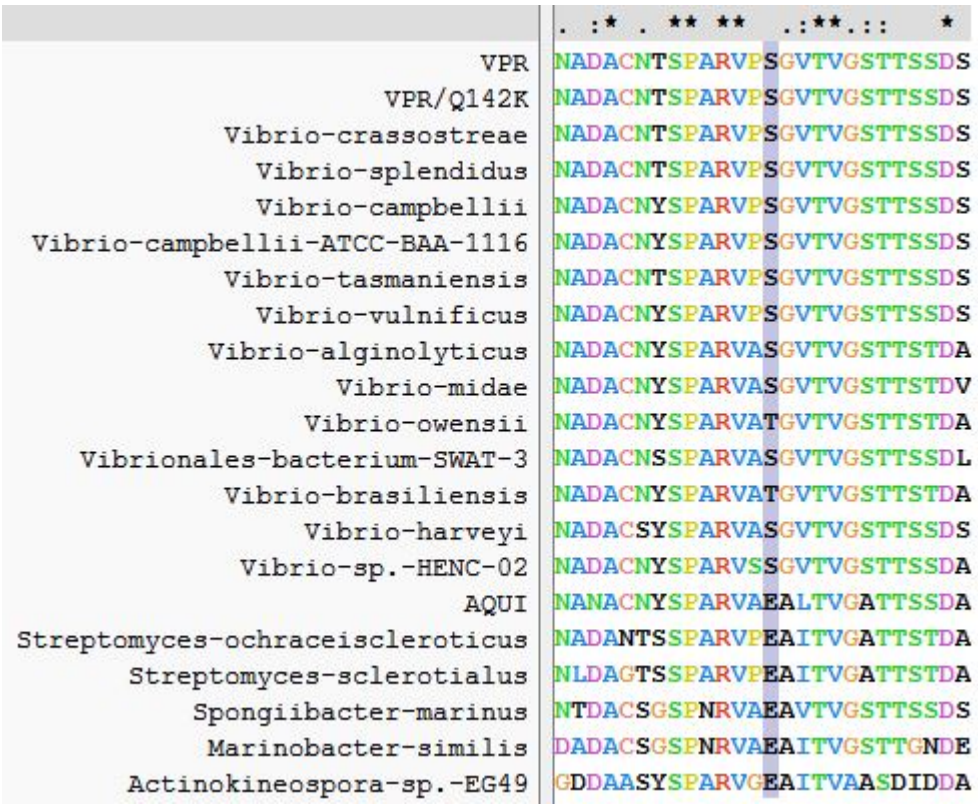


*Figure 44. Sequence alignment of the area around residue 172 (According to numbering in AQUI, highlighted in grey). As seen in the alignment all the proteases that have a negatively charged residue in position 142 (VPR numbering) have a positively charged residue in position 172, that is likely forming a salt bridge as in AQUI. An * (asterisk) indicates positions which have a single, fully conserved residue. A : (colon) indicates conservation between groups of strongly similar properties. A . (period) indicates conservation between groups of weakly similar properties. Organism of origin shown for all except VPR and AQUI.*

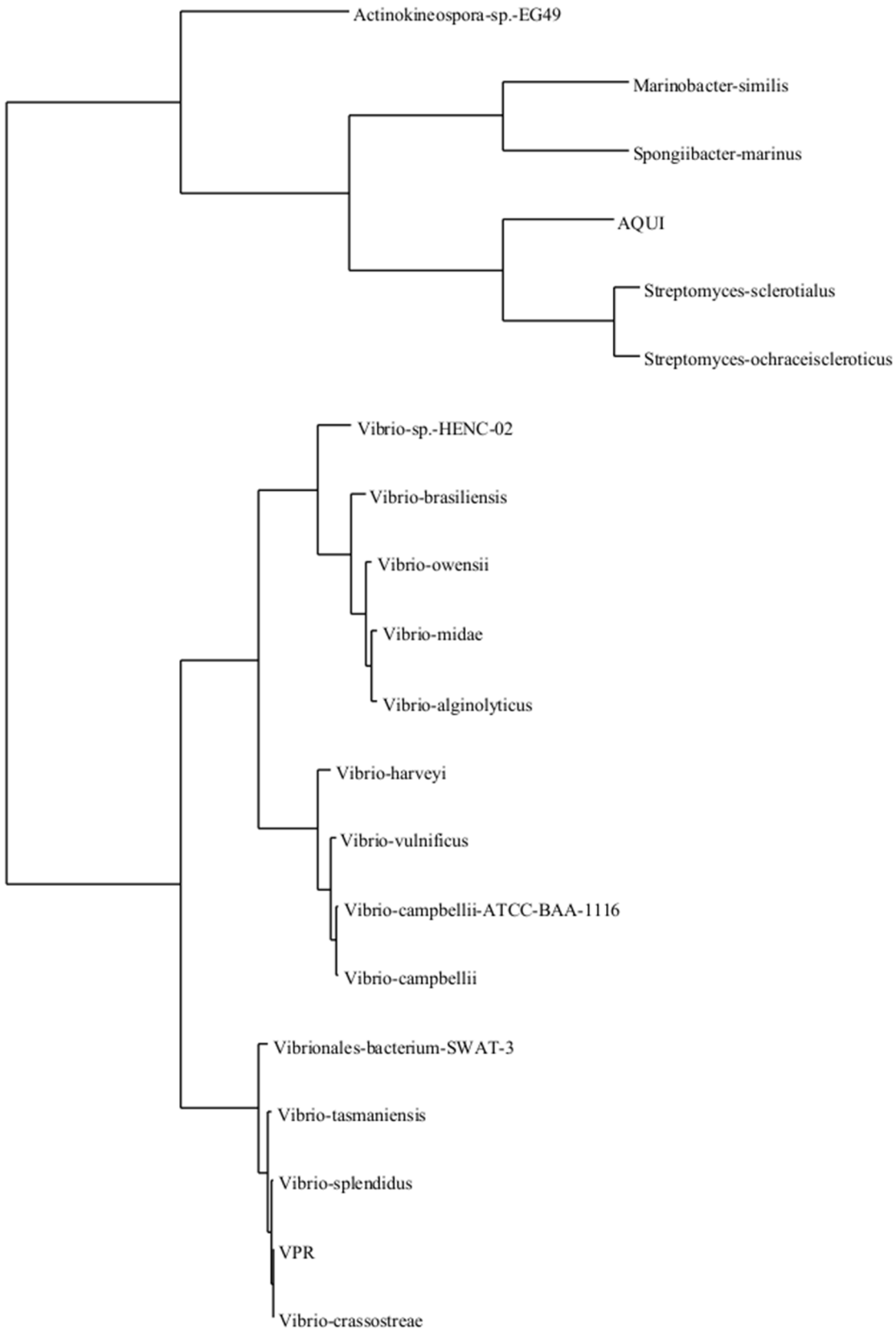


*Figure 45. Phylogenic tree for subtilisin-like serine proteases from the S8 family build on their amino acid identity. The tree was made using ClustalX2, using UPGMA clustering algorithm. Organism of origin shown for all except VPR and AQUI.*